労災疾病臨床研究事業費補助金

職場における腰痛の効果的な治療法等に関する研究

平成26年度～28年度　総合研究報告書

主任研究者　松平　浩

平成29年3月

労災疾病臨床研究事業費補助金

職場における腰痛の効果的な治療法等に関する研究

平成 26 年度～28 年度　総合研究報告書

主任研究者　松平浩

平成 29 年 3 月

目　　次

Ⅰ.総合研究報告

Ⅱ. 総合研究分担報告

Ⅰ. 総合研究報告

労災疾病臨床研究事業費補助金

平成 26 年度～28 年度総合研究分担報告書

職場における腰痛の効果的な治療法等に関する研究

研究代表者　松平浩　東京大学医学部附属病院 22 世紀医療センター
運動器疼痛メディカルリサーチ&マネジメント講座

研究要旨

本研究では、世界的にみてもいまだ克服されていない腰痛対策をテーマに、特に介護看護従事労働者をターゲットとして、疫学的手法を用いたリスク因子の同定、発症予防に役立つ体操や福祉機器および両立支援手法の開発ついての取り組みを、3 年計画で包括的に推進した。本研究の検討結果は以下の通りである。

1)作業支障腰痛が遷延するリスクとして、職場を主とする心理社会的要因が大きく関与することが明らかになった。

2)開発した福祉機器（新たなコンセプトの体幹装具）は、脊柱起立筋の活動を減少させ、体幹深部筋（腹横筋）の筋厚を増大させたことから、腰痛予防に有益な可能性が高いことが示唆された。さらに装着（歩行時）時の高次運動野の活動への影響を近赤外光脳機能計測装置により検討した結果、装具の着用で垂直姿勢の維持や運動制御に関わる高次運動野の活動が修飾されることが明らかになった。

3)姿勢と腰部負担との関連を、三次元動作分析装置を用いて明らかにした。適度な骨盤前傾と体幹伸展の姿勢では腰部負担が小さくなることがわかったため、この姿勢をとることをフィードバックする新たな福祉機器「不良姿勢チェッカー」を作製し、その有用性を確認した。

4) 腰痛スクリーニングに活用可能な腰椎 MRI 画像診断アルゴリズムを構築した。

5)簡易で即実践できる体操に加え、産業理学療法士からの科学的根拠に基づいた腰痛教育の有益性を、全国の 12 労災病院をクラスターとして大規模介入比較試験を行い検証した。ベースライン調査では 4,767 名にアンケートを配布、アンケート回収数は 3,439 名、解析には 3,381 名分のアンケートを利用した。6 か月後の追跡調査時の、各群の回収数は 2,406 名であり、追跡率は 70.0%あった。、腰痛の自覚症状改善の割合は,コントロール群に比して、腰椎伸展体操の普及・実践群、+産業理学療法士による腰痛教育・相談いずれの介入によっても上昇していた。多変量を調整した Logistic 回帰分析の結果、両介入群とも有意に腰痛を改善（コントロール群の約 2 倍）することが分かった。

6)産業理学療法士によるメール指導は、相談者の腰痛予防のための行動変容を促すのに効果的であり、両立支援手法の一手段となる可能性があると思われた。

最終的な本研究グループの活動成果として効率的かつ包括的な作業関連性腰痛の予防対策の提言を作成した。同提言は、社会・医療経済面、更には労災補償面にも大きく貢献するものと考えている。

＜研究分担者＞

東京大学医学部附属病院22世紀医療センター
運動器疼痛メディカルリサーチ&マネジメント講座　特任准教授　岡敬之　（H28年度に助教より職位変更）

金城大学医療健康学部　教授　小山善子

新潟医療福祉大学医療技術学部義肢装具自立支援学科　准教授　勝平純司
（国際医療福祉大学専任講師よりH28年度より所属変更）

関東労災病院　整形外科医師　唐司寿一

東京大学医学部附属病院　特任助教　四津 有人

横浜労災病院整形外科部長　三好光太

関西福祉科学大学理学療法学部　教授　野村卓生

関西労災病院 治療就労両立支援センター　主任理学療法士　高野 賢一郎

大阪労災病院治療就労両立支援センター　主任理学療法士　浅田　史成

東京大学大学院医学系研究科　成人看護学　講師　高井ゆかり（転任に伴い H28 年度より分担研究者を降任）

＜研究協力者＞

東京大学医学系研究科精神保健学/精神看護学分野　准教授　島津明人

労働安全衛生総合研究所上席研究員　高橋正也

労働安全衛生総合研究所　岩切一幸

川村義肢株式会社製造本部　安井匡

東京大学医学部附属病院22世紀医療センター運動器疼痛メディカルリサーチ&マネジメント講座
登録研究員　川又　華代

東京大学医学部附属病院22世紀医療センター運動器疼痛メディカルリサーチ&マネジメント講座
登録研究員　吉本隆彦

長崎労災病院副院長　小西宏昭

千葉労災病院副院長　山縣正庸
関東労災病院副院長　岡﨑裕司

中国労災病院治療就労両立支援センター　産業理学療法士　仁田靖彦

東北労災病院治療就労両立支援センター　産業理学療法士　佐藤友則

関西福祉科学大学理学療法学　明崎禎輝

研究分担者の研究期間を通じてのテーマを以下の表に示す。

研究者名	H26	H27	H28	研究期間を通じてのテーマ
松平浩	○	○	○	研究統括・提言作成
岡敬之	○	○	○	腰痛に関わる実態およびリスクの同定
小山 善子	○	○	○	
勝平 純司	○	○	○	予防に有用な福祉機器等の開発
唐司 寿一		○	○	
四津 有人			○	
三好 光太	○	○		介護看護従事者への予防介入
野村 卓生	○	○	○	個人と職場の双方に有益な腰痛治療と職業生活との両立支援手法の開発
高野 賢一郎		○	○	
浅田 史成		○	○	
高井ゆかり	○	○		

A. 研究目的

厚生労働省が公表する「国民生活基礎調査の概要」において，国民の代表的愁訴（有訴者率）が，腰痛（男 1 位，女 2 位），肩こり（男 2 位，女 1 位）であることはよく知られているが、同じく厚生労働省が公表する業務上疾病発生状況等調査によると、休業 4 日以上の業務上疾病の発生件数のうち腰痛は、長年に渡り全職業性疾病の約 6 割を占め第 1 位である。平成 23 年の腰痛全届け出のうち社会福祉施設が 19%を占め、10 年で 2.7 倍という最も顕著な増加となった背景を踏まえ、19 年ぶりに改訂された「職場における腰痛予防対策指針」（平成 25 年、厚生労働省）では、重症心身障

害児施設等に限定されていた適用を、福祉・医療等における介護・看護作業全般に拡大し、内容を充実させるに至った。つまり、介護・看護従事者への腰痛対策は、産業衛生領域の喫緊の課題といえる。また世界疾病負担研究にて 289 の疾患や傷病のうち、腰痛が Years Lived with Disability (YLDs)、つまり健康でない状態で生活する年数を指標とする統計のトップにランクされるなど、社会的損失や健康面への影響の大きい腰痛への対策はグローバルにも重要な課題として位置づけられている。

研究代表者は、昨年度まで行われた「労災疾病等 13 分野研究」の本分野において世界標準のエビデンスを踏まえつつ独創的な研究を展開し、近年、国際的にも評価される業績を公表してきた（13 分野研究の総括事後評価点数：5 点満点で 4.9)。

本研究では、世界的にみてもいまだ克服されていない腰痛対策に関し、臨床専門の医師のみならず産業医学・産業保健、看護、人間工学、福祉工学、統計学といった様々な分野のスペシャリストを分担研究者、研究協力者として多数招聘し、これまでの主任研究者の実績と研究基盤をさらに発展させる。特に介護看護従事労働者をターゲットとして、疫学的手法を用いたリスク因子の同定、発症予防を目的とした介入法の構築、福祉用具の開発や利用および職業と治療の両立支援法の作成を、PDCA サイクルも有効に活用しつつ包括的に推進することとした。

具体的には、３年計画で以下のサブテーマに基づき遂行する予定とした。①腰痛に関わる実態およびリスクの同定、②予防に有用な福祉機器等の開発、③介護看護従事者への予防介入法とマネジメントシステムの構築、④個人と職場の双方に有益な腰痛治療と職業生活との両立支援手法の開発、以上を踏まえた⑤労働安全衛生マネジメントシステム構築を視野に入れた提言作成。

3 年間推進した研究に関し、上①～⑤のサブテーマ毎に報告する。年度をまたぎ継続した課題もあるため、本報告では特に年度毎の記載を行うのではなく、H26-28 年度通じての成果を記載する。なお研究代表者である松平と分担研究者の岡は全ての分担研究に参画し、研究デザイン・統計解析を行っている。

B. 研究方法

① 腰痛に関わる実態およびリスクの同定

1. 石川県内の医療・介護施設 125 か所に調査用紙を送り、本調査研究の趣意に賛同を得た医療・介護施設 95 か所及びそこで働く無作為に選定された 1 施設 20 名を対象とし、腰痛状況、職業性ストレス簡易調査票を含む多面的な自記式質問紙による実態調査を行った。
2. 社会福祉懇談会の会員である介護施設（50 施設、4,105 名）に対して、自記式質問紙による調査を行い、作業支障腰痛と作業負荷や勤務体制等の関連性等を検討した。
3. 前向きに追跡したコホート研究である JOB (Japan epidemiological research of Occupation-related Back pain) study の 1 年間の追跡データを用い、ベースライン時、仕事に支障のある腰痛を直近１ヶ月に経験した労働者が、１年間のフォローアップ期間に同等の腰痛を３ヶ月以上経験していた場合を遷延化と定義し、遷延化する因子を検討した。
4. 長野県の介護福祉施設(2 施設を積極介入群、1 施設をコントロール群）として、研究代表者が開発した‘これだけ体操’の効果を検証した（**予防介入法に導入可能か一般的な施設における実態の検証を行った。**）
5. 単一医療介護施設の職員を対象に、自記式調査票を用いた横断研究において、腰痛の現状およびその関連因子について心理社会的要因を中心に探索した。対象は、某医療法人社団の職員 280 名とし、無記名の自記式質問票を郵送にて回収した。
6. 心身の健康に関するインターネット調査にて労働力人口として現役世代である 20-64 歳の

慢性腰痛のを持つ 20-64 歳の日本人約 3,000 人を対象に健康関連 QOL と身体化，抑うつ症状の関連を評価した.解析では抑うつ,年齢,性別,BMI,喫煙,婚姻,学歴,定期的な運動,雇用状況，通院していた疾患の個数（0－25）を調整した.

②予防に有用な福祉機器等の開発

1. 従来の体幹装具は装着することで体幹運動を制限することで腰部負担の軽減を目指している。しかしながら、先行研究では体幹装具装着による腰部負担の軽減効果を示すことができていない。また、体幹装具は長期間装着すると体幹周囲筋、特に側腹筋の弱化につながると指摘されている。そこで継手の抗力により胸部を前方から押す力を与えることで腹筋の活動を促通し、背筋の活動を低減する新たなコンセプトの体幹装具 Trunk solution（以下 TS）を開発し、2014 年の Good Design 賞を受賞した。本研究では、健常若年成人を対象とし、TS の歩行における脊柱起立筋の活動に対する効果と外した後の持ち越し効果および TS 装着時の側腹筋の筋厚を、従来の体幹装具の中で最も汎用されているダーメンコルセットと比較検討を行った。対象は健常成人男性 24 名で、TS 装着群(TS 群)とダーメンコルセット装着群(ダーメン群)の 2 群に、無作為に選別した。
2. 腰部負担の大きい重量物挙上における持ち上げ姿勢と，負担は小さくとも蓄積されると腰痛の発症につながると考え，立位姿勢も対象として良姿勢が腰部負担に与える影響を三次元動作分析装置を用いて明らかにした.適度な骨盤前傾と体幹伸展の姿勢では腰部負担が小さくなることがわかったため，この姿勢をとることをフィードバックする「不良姿勢チェッカー」を作製した.
3. 腰椎の画像所見と腰痛とが必ずしも一致しない症例が、臨床的には散見される。なかでも腰椎 MRI は空間分解能も高く、優れたモダリティであるが労働者における疫学的な検討は不十分である。、関東労災病院に勤務する職員にて画像データベースを構築し腰椎 MRI 所見と過去の高度な腰痛の既往との関連を分析した。
4. 開発した福祉機器（新たなコンセプトの体幹装具）装着（歩行時）時の高次運動野の活動への影響を、8 名の被験者にて近赤外光脳機能計測装置により検討した。

③介護看護従事者への予防介入法とマネジメントシステムの構築

主任研究者は、勤務中多忙な介護看護従事者が簡易で即実践できる腰痛予防体操（腰を反らす“これだけ体操”）を、ポピュレーションアプローチとして実践することにより職場の腰痛状況を改善できる可能性を先行研究で示している。本研究では、産業衛生領域の喫緊の課題である腰痛予防対策を効率的に行うマネジメントシステムを構築する基盤として、簡易で即実践できる体操に加え、産業理学療法士からの科学的根拠に基づいた腰痛教育の有益性を、大規模介入比較試験により検証した。全国の 12 労災病院をクラスターとして、A:対照（無介入）、B：腰椎伸展体操の普及・実践、C:B+産業理学療法士による腰痛教育・相談の実践の 3 群を実施する臨床研究を実施した。

- 施設をクラスターとした無作為比較試験
- 対照（無介入）、腰椎伸展体操の普及・実践、B の介入+産業理学療法士による腰痛教育・相談の実践の 3 群
- 北海道中央（予定看護師数:155）、東北（421）、関東（610）、横浜（585）、新潟（261）、浜松（236）、旭（189）、大阪（662）、関西（619）、中国（363）、愛媛（180）、長崎（300）、総計 4,581 名をリクルート。以上 12 労災病院（施設）のをクラスターとし、病床・看護師数、看護師の男女数・平均年齢を割付調整因子と

し、コンピューターの乱数表を用い、3群（4施設ごと）に無作為割付する非盲検試験

- エンドポイント：腰痛の有無および仕事への支障度を勘案した腰痛 grade（重症度）の改善
- 介入期間：1年
- 選択基準:選定された労災病院に勤務する成人（20歳以上）看護師、本研究の趣旨に賛同し同意を得た者。
- 除外基準：妊婦，あるいは妊娠の疑いがある場合、腰椎伸展により症状が誘発される腰部脊柱管狭窄症と診断されたことがある者、研究の同意を撤回した者。

④個人と職場の双方に有益な腰痛治療と職業生活との両立支援手法の開発

業務上疾病の約6割を占める腰痛には、人間工学的要因のみならず心理社会的要因も関与することが科学的根拠のある事項として認識され、さらには正しい情報の提供や周囲の励ます態度などは腰痛を軽快させることが明らかになりつつある。

一方、腰痛予防に関しても、特定健診・保健指導で用いられるメール指導による腰痛予防効果の有効性が期待されたため、両立支援手法の一手段として産業理学療法士主導で取り組んできた。その結果、メール指導前後において労働者が各自の職務をどれほど上手にできているかを表す指標である Work Ability Index（WAI）の有意な向上、腰痛に関わる就労状況を含めた予後規定因子としてグローバルに最も重要視されている恐怖回避の思考・行動を表す Fear-Avoidance Beliefs Questionnaire（FABQ）の改善傾向を認め、産業理学療法士による腰痛予防を主軸とする両立支援を目的としたメール指導の一定の効果を確認してきた。

本研究では多数の労働者を対象としてメール指導ができるように独自のシステムを開発した．また，世界における労働者を対象とした理学療法（産業理学療法）の介入やその教育課程に関して情報収集を行った．さらに，腰痛予防教育教材を開発した．専用のシステムは，産業理学療法指導システム「Consulting system for physical therapy in occupational health: Compo」と命名した．

開発したシステムを用いて，保健衛生業に従事する労働者を対象として，メール指導による腰痛予防効果を検証した．また，世界の産業理学療法の情報収集を継続した．さらに，開発した腰痛予防教育教材の普及を行った．

Compo を用いた介入研究は，臨床試験登録システム UMIN-CTR（UMIN000018450）に登録した．30 歳から 65 歳までの保健衛生業に従事する者を対象として，Compo を用いて指導を行う群（介入群）と介入を行わない群（対照群）の 2 群に振り分け，研究を実施した．研究手順として，国内の 11 施設の協力を得た．介入・観察期間は 6 カ月である．まず，指導者から対象者へメールを送信し，以降，1 カ月に 1 回，指導者から対象者へメールを送信することを原則とした（計 7 回）．

⑤労働安全衛生マネジメントシステム構築を視野に入れた提言作成

上述した全ての研究より得られた知見から研究代表者が、多くの施設で導入可能な簡易な「マネジメントシステムを視野に入れた提言」を作成した。

（倫理面への配慮）

独立行政法人労働者健康福祉機および関連各労災病院、国際医療福祉大学、関西福祉科学大学、東京大学医学部附属病院の倫理審査の承認を得ている．被験者に対してはデータを ID 化して管理するなど個人情報には十分配慮している。

C. 研究結果

① 腰痛に関わる実態およびリスクの同定

1. アンケートの回答者（施設回収率 76.0%）は 1704 名で、男性 420 名、女性 1281 名（性別未回答 3 人）であった。雇用状況は、職種では介護福祉士が 48.0%、ホームヘルパー14.0%、看護師・保健師は 12.6%、理学療法士・作業療法士が 1.2%、ケアワーカー6.6%である。看護・介護業務の経験年数は 10 年以上～20 年未満が 32.3%と一番多く、ついで 5 年以上～10 年未満 26.8%、2 年以上～5 年未満 17.7%、20 年以上が 11.3%、2 年未満では 10.6%で、うち 1 年以上～2 年未満 6.0%、1 年未満 4.6%であった。腰痛状況に関してだが、過去 1 か月で腰痛がないと回答した人は 441 人（25.9%）のみで、1133 人（66.5%）の 7 割弱の人が腰痛を認めていた（以下、腰痛あり群)。腰痛有訴者をさらに分析すると、腰痛のため仕事に支障をきたしたとの回答者は 258 名（全回答者の 15.1%）にみられ、40 名（全回答者の 2.3%）が腰痛のため欠勤（休職）していた。さらに腰痛ありの人で腰痛を患わってから 3 か月以上経過している、つまり慢性化している人が 74.4%と腰痛有訴者の 2/3 を占めた。また、腰痛のため、連続して 4 日以上休んだ経験のある人は、11.2%存在した。腰痛に関連する要因を、職業性ストレス簡易調査票の項目を主に説明変数として多変量調整ロジスティック回帰分析で検討したところ、身体的負担度、ストレス反応の活気、疲労感、抑うつ感、身体的愁訴が負の方向に有意な因子であった。3,155 名より回答が得られた（回答率 77%)。
2. 作業に支障をきたす腰痛の訴えは、自覚的な労働負荷が強く夜勤時間が長くなるにつれて増加した。加えて、介護職 1 名あたりの利用者数が多く夜勤時間が長いと、作業支障腰痛が増加した。
3. ベースライン時に回答した 5,310 名中,3,811 名が 1 年フォローアップに回答していた（追跡率：71.8%)。171 名にベースラインでの直近 1 カ月に仕事に支障のある腰痛があり、そのうち 46%が、看護・介護あるいは 20kg 以上の重量物を取り扱う作業に従事していた。171 名中 29 名（17.0%）がフォローアップ期間中、遷延化していた。ロジスティック回帰分析の結果、働きがい、不安感、および仕事や生活の満足度が、他要因を調整しても遷延化の有意な危険因子として挙げられた。
4. 両群併せて 167 名（積極介入群：89 名、対照群：78 名）の介護士をエントリーした。167 名の平均年齢は 37.5 才で、女性が 65 名（81.4%）であった。過去 1 年間に腰痛がなかった割合は 29.9%にとどまり、全体の 10.2%が過去 1 年を総合的に考えて腰痛にため仕事に支障をきたしていた。自覚的改善度、対策の実行度とも対照群と比較し、積極介入群のほうが統計的に有意に優れていた。
5. アンケートの有効回答者数 203 名（72.5%）であった。平均年齢は 39.8 歳（SD 12.2)、性別は 70%が女性であった。対象者のうち、仕事に支障をきたした腰痛経験者は 36 名（17.7%）であった。上記腰痛経験者の群とそれ以外の群で、各調査項目について群間比較を行ったところ、FABQ（恐怖回避思考、p = 0.037)、SSS-8（身体症状、p = 0.0003)、職場での対人関係ストレス（p = 0.022）が統計学的な有意差を認めた。年齢、性別、BMI、職業を調整したロジスティック解析においても、上記 3 要因が有意な因子として抽出された。
6. 参加者は平均 44.5±11.2 歳で，48%が女性であった.PHQ-2 は 1576 人(51%) が 0,632 人(20%)が 1, 892 人 (29%)が 2 であった. EQ-5D は平均 0.78±0.18 であり， PHQ-2 の点数が高くなるほど低かった． SSS-8 の平均スコアは 9.67±6.68 で，PHQ-2 の点数が高くなるほど，SSS-8 が非常に高い（≥16）の割合

が高かった．SSS-8 と PHQ-2 の交互作用は統計的に優位でなかったので，最終モデルには含めなかった．最終的な多変量モデルでは，年齢，性別，BMI，喫煙，婚姻，学歴（短大以上か），定期的な運動の有無，雇用状況（正規雇用かそれ以外か），通院していた疾患の個数（0－25）を調整した．PHQ-2 の点数は EQ-5D のスコアと有意に関連していた．SSS-8 のどのカテゴリに属するかは，PHQ-2 の点数や他の共変数を調整しても，EQ-5D のスコアと有意に関連していた．すなわち，身体化傾向の高いグループほど，EQ-5D のスコアが低かった．

②予防に有用な福祉機器等の開発

1. 体幹装具（TS)を装着することにより、歩行時の脊柱起立筋の活動が減少するとともに外した直後においてもその効果が維持された。さらには、体幹深部筋である腹横筋の筋厚が増加した。これらの効果はダーメンコルセットを装着しても認められなかった。
2. 体幹部と腰部との関係性に，体幹部と頭部との関係を付加考慮することで，頸部に起因するストレートネック等の関連する症状も把握できることがわかった．これらの機能を組み込んだ第 2 次「治療モデル不良姿勢チェッカー」を試作評価した結果，不良姿勢改善の再現性において良好な結果を得ることができた．第 2 次「治療モデル不良姿勢チェッカー」を介護現場での腰痛の前兆となる腰痛リスクの可視化を検討したが，介護現場においては重量物の運搬等，姿勢のみでは判断できない介護者への負担という新しいリスクを考慮する必要があることがわかった． また運搬等の動作分析から負担部位を特定し，それらの部位を主動している脊柱起立筋の筋活動を介護業務中に監視する機能を有する，「予防モデル不良姿勢チェッカー」を作製した．「予防モデル不良姿勢チェッカー」では脊柱起立筋を指標として作業現場で姿勢のみならず腰部負担を計測し，かつ負担量が大きい作業員の位置を特定する仕組みまで構築した．
3. 腰椎 MRI 所見と過去の高度な腰痛の既往との関連を年齢・性を調整して解析した結果、Pfirrmann 分類≧3、椎間板膨隆あり、High intensity zone(HIZ)あり、が過去の高度な腰痛の既往と関連していた。 これらの知見は、腰痛スクリーニングに活用可能であり、予防アルゴリズムの一角を担うものと考えている。
4. 開発した福祉機器（新たなコンセプトの体幹装具）の着用で、補足運動野は歩行安定期により早期に有意に低下した。また左右の運動前野は、歩行安定期により晩期に低下した。以上により装具の着用で垂直姿勢の維持や運動制御に関わる高次運動野の活動が修飾されることが明らかになった。

③介護看護従事者への予防介入とマネジメントシステムの構築

6 か月後の追跡調査時の､各群の回収数は A 群 949 名､B 群 706 名､C 群 751 名､計 2,406 名であり、追跡率はそれぞれ 71.9%、70.6%、67.0%で、全体では 70.0%あった。腰痛と関連情報を把握するためのアンケート調査を行った結果、腰痛の自覚症状改善の割合は,A 群で 13.3%､B 群で 23.5%､C 群で 22.6%と介入群で上昇していた。また腰痛予防対策の実行度はコントロール群で低くなっていた。腰痛の改善を目的変数として、背景を調整しても介入治療効果が認められるかに関して多変量解析（Logistic 回帰分析）を用いて検討した。雇用の安定等に関する法律（高年齢者雇用安定法）をもとに、45 歳以上を「中高年齢者」と、また BMI 25 以上を肥満と定義した。この結果、B の介入（腰椎伸展体操の普及・実践)、C の介入（B の介入+産業理学療法士による腰痛教育・相談の実践）とも有意に腰痛を改善（コントロール群の

約2倍）することが分かった。またFABQが15点未満であることは腰痛改善の因子であることが明らかになった。

④個人と職場の双方に有益な腰痛治療と職業生活との両立支援手法の開発

産業理学療法士によるメール指導は、相談者の腰痛予防のための行動変容を促すのに有用と思われているものの介入群および対照群において，全ての一般特性の項目に有意な差は認めなかった．各群において，介入/観察前後の全ての項目に有意差は認めなかった．また，2群間の変化量についても全ての項目に有意な差を認めなかった．

腰痛予防を目的とした教育教材を開発し，これらをインターネットやSNSを利用して普及啓発した．今後は，これらの効果的な使い方（労働者への教育方法や労働環境への導入方法など）を検討し，人的手段あるいはインターネットによる教育教材を使用する側（指導者側）の教育も継続しなければならないと考えている．

⑤労働安全衛生マネジメントシステム構築を視野に入れた提言作成

作成した成果物を本稿章末に提示する。

D. 考察

介護の現場では、腰痛有訴者が多く、身体的負荷のみならず、ストレス反応としての活気、疲労感、抑うつ感、身体愁訴が関連することが浮き彫りになった。

また、作業に支障をきたす腰痛の訴えは、自覚的な労働負荷が強く夜勤時間がが長くなるにつれて増加することに加え、介護職1名あたりの利用者数が多くかつ夜勤時間が長い場合も作業支障腰痛が増加することが示唆されたことから、介護労働に伴う身体的・精神的負荷を軽減することが夜勤の短縮とともに重要になると考えられた。そのためには，福祉機器の活用や，施設内で安全衛生のレベルを向上させるマネジメントシステムの確立などが求められる。

作業に支障をきたす腰痛が遷延化することに影響する要因は、仕事や生活での満足度が低い、働きがいが乏しい、不安感が強いといった、職場を主とする心理社会的要因であることが明らかになったが、本知見は、欧米のエビデンスと矛盾しない。メカニズムとしては、心理社会的要因がストレッサーとなり，中脳辺縁系ドパミン・オピオイドシステムの機能異常に続発する下行性疼痛調節系や自律神経系のアンバランスに伴う痛覚過敏や局所の血流低下・筋攣縮などが考えられる。その結果として、複数の身体愁訴（いわゆる身体化、腰痛はその一症状）が出現したり遷延化する場合があるのだろう。

我が国の産業衛生分野において、人間工学的なアプローチによる腰痛の予防や対策が主流であり、重要なアプローチであることは疑いない。しかしながら、厚生労働省業務上疾病発生状況等調査にて、腰痛における休業4日以上の業務上疾病の発生件数をはじめとする統計データが長年にわたり好転していない現状を踏まえると、作業支障腰痛の遷延化による職場における労働力の損失を予防・緩和するための今後の対策として、心理社会的要因へのアプローチも人間工学的アプローチと並行して考慮する必要がある。本研究では疫学的な分析を積み重ね、リスク要因に関するエビデンスを構築し提言を作成するまでにいたった。この成果物は今後の腰痛予防のマイルストーンになるものと自負している。

予防に有効な福祉機器の開発に関してであるが、新たなコンセプトで開発した体幹装具（TS）を装着することにより，歩行時の脊柱起立筋の活動が減少するとともに外した直後においてもその効果が維持され、かつ、体幹深部筋である腹横筋の筋厚が増加するという従来のダーメンコルセット装着では得られない効果が認められた。脊柱起立筋の過活動は腰痛につながり，体幹深部筋のエクササイズは腰痛予防に効果があると報告

されていることから、体幹装具（TS)を装着することによる脊柱起立筋の活動を低減しながら、体幹深部筋の収縮を促通する効果は腰痛予防の福祉機器となりうると考えられた。

また TS 着用で、補足運動野は歩行安定期により早期に有意に低下し、左右の運動前野は、歩行安定期により晩期に低下したことから装具の着用で垂直姿勢の維持や運動制御に関わる高次運動野の活動が修飾されることが明らかになった。

本研究ではさらに今適度な骨盤前傾と体幹伸展の姿勢では腰部負担が小さくなることがわかったため，この姿勢をとることをフィードバックする「不良姿勢チェッカー」を作製した。当初目的としていたエビデンスに基づく体幹装具に代わる新たな姿勢を修正する機器の開発だけでなく，これを発展させた機器の開発まで 3 か年内に達成することができ，本プロジェクトは順調に推移したものと考えている。

また本研究では、国内外に類を見ない腰痛予防運動の大規模な介入研究を施行した。これは 12 労災病院の協力なくしては得られなかった成果であり、研究代表者が提唱する腰痛予防法のエビデンスを確立するための代表的な研究となることが予想される。

さらには、腰痛に関わる両立支援を推進するうえで、運動器およリハビリテーション医学の領域に加え、産業保険分野に精通している産業理学療法士が実施するメール指導は、腰痛予防のための行動変容を促すことから、その質の高いシステムを構築された。

以上の成果物をもとに、これらが普及することによって腰痛対策を新たなステージに進めることができるのではないかと考えている。今後、広報も含め積極的な展開を予定している。

E. 結論

最終的な本研究グループの活動成果として効率的かつ包括的な作業関連性腰痛の予防対策の提言を作成した。同提言は、社会・医療経済面、更には労災補償面にも大きく貢献するものと考えている。

F. 健康危険情報

該当なし

G. 研究発表

論文発表

1. Fukushima M, Oka H, Hara N, Oshima Y, Chikuda H, Tanaka S, Takeshita K, Matsudaira K. Prognostic factors associated with the surgical indication for lumbar spinal stenosis patients less responsive to conservative treatments. J Orthop Sci (in press)
2. Oka H, Matsudaira K, Kikuchi N, Haga Y, Sawada T, Katsuhira J, Yoshimoto T, Kawamata K, Tonosu J, Sumitani M, Kasahara S, Tanaka S: Estimated risk for chronic pain determined using the generic STarT Back 5-item screening tool. J Pain Res (in press)
3. Yamada K, Matsuadira K, Tanaka E, Oka H, Katsuhira J, Iso H. Sex-specific impact of early-life adversity on chronic pain: A large population-based study in Japan. J Pain Res (in press)
4. Matsudaira K, Oka H, Kawaguchi M, Murakami M, Fukudo S, Hashizume M, Löwe, B. Development of a Japanese Version of the Somatic Symptom Scale-8: Psychometric Validity and Internal Consistency. Gen Hosp Psychiatry 45:7–11, 2017
5. Wakaizumi K, Yamada K, Oka H, Kosugi S, Morisaki H, Shibata M, Matsudaira K. Fear-avoidance beliefs are independently associated with the prevalence of chronic pain in Japanese workers. J Anesth. 2017, Jan 3. [Epub ahead of print]
6. Tonosu J, Oka H, Matsudaira K, Higashikawa A, Okazaki H, Tanaka S. The relationship between

the findings on magnetic resonance imaging and previous history of low back pain. J Pain Res 10: 47-52, 2017

7. Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S. The Japanese version of the STarT Back Tool predicts 6-month clinical outcomes of low back pain. J Orthop Sci. 2016 Dec 23. [Epub ahead of print]
8. Katsuhira J, Matsudaira K, Oka H, Iijima S, Itou A, Yasui T, Yozu A. Efficacy of a trunk orthosis with joints providing resistive force on low back load during level walking in elderly persons. Clin Interv Aging 11: 1589-1597, 2016
9. Tonosu J, Inanami H, Oka H, Katsuhira J, Takano Y, Koga H, Yuzawa Y, Shiboi R, Oshima Y, Baba S, Tanaka S, Matsudaira K. Diagnosing Discogenic Low Back Pain Associated with Degenerative Disc Disease Using a Medical Interview. PLoS One 11: e0166031, 2016
10. Tsuji T, Matsudaira K, Sato H, Vietri J. The impact of depression among chronic low back pain patients in Japan. BMC Musculoskelet Disord 17: 447, 2016
11. Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Vargas-Prada S, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SS, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Sarquis LM, Marziale MH, Harari F, Freire R, Harari N, Monroy MV, Quintana LA, Rojas M, Harris EC, Serra C, Martinez JM, Delclos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Freimann T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VC, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A, Salazar Vega EJ.: Epidemiological differences between localised and non-localised low back pain. Spine, 2016 Nov4.[[Epub ahead of print]
12. Isomura T, Sumitani M, Matsudaira K, Kawaguchi M, Inoue R, Hozumi J, Tanaka T, Oshima H, Mori K, Taketomi S, Inui H, Tahara K, Yamagami R, Hayakawa K. Development of the Japanese Version of the Leeds Assessment of the Neuropathic Symptoms and Signs Pain Scale (LANSS-J): Diagnostic Utility in a Clinical Setting. Pain Practice, 2016 Oct 22. [Epub ahead of print]
13. Taniguchi Y, Takahashi M, Matsudaira K, Oka H, Momose T. Potential use of 18F-FDG-PET/CT to visualize muscle pain in patients with adult spinal deformity: A case report. Skeletal Radiol 45:1577-81,2016
14. Hara N, Matsudaira K, Masuda K, Tohnosu J, Takeshita K, Kobayashi A, Murakami M, Kawamura N, Yamakawa K, Terayama S, Ogihara S, Shiono H, Morii J, Hayakawa K, Kato S, Nakamura K, Oka H, Sawada T, Inuzuka K, Kikuchi N. Psychometric Assessment of the Japanese Version of the Zurich Claudication Questionnaire (ZCQ): Reliability and Validity. PLoS One 11:e0160183, 2016
15. Iwahashi H, Yoshimura N, Hashizume H, Yamada H, Oka H, Matsudaira K, Shinto K, Ishimoto Y, Nagata K, Teraguchi M, Kagotani R, Muraki S, Akune T, Tanaka S, Kawaguchi H, Nakamura K, Minamide A, Nakagawa Y, Yoshida M. The Association between the Cross-Sectional Area of the Dural Sac and Low Back Pain in a Large Population: The Wakayama Spine Study. PLoS One 11: e0160002, 2016
16. Yamada K, Matsudaira K, Imano H, Kitamura A, Iso H. Influence of work-related psychosocial factors on the prevalence of chronic pain and quality of life in chronic pain patients. BMJ Open 6: e010365, 2016

17. Hayashi S, Katsuhira J, Matsudaira K, Maruyama H: Effect of pelvic forward tilt on low back compressive and shear forces during a manual lifting task. J Phys Ther Sci 28: 802-6, 2016
18. Vargas-Prada S, Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SS, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Sarquis LM, Marziale MH, Harari F, Freire R, Harari N, Monroy MV, Quintana LA, Rojas M, Harris EC, Serra C, Martinez JM, Delclos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Freimann T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VC, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A, Vega EJ. Descriptive Epidemiology of Somatising Tendency: Findings from the CUPID Study. PLoS One 11 :e0153748, 2016
19. Tonosu J, Matsudaira K, Oka H, Okazaki H, Oshio T, Hanaoka I, Muraoka Y, Midorikawa M, Wakabayashi K, Tanaka S. A population approach to analyze the effectiveness of a back extension exercise "One Stretch" in patients with low back pain: A replication study. J Orthop Sci 21:414-8, 2016
20. Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S. Psychometric Properties of the Japanese Version of the STarT Back Tool in Patients with Low Back Pain. Plos One 11:e0152019, 2016
21. Nomura T, Asada F, Takano K, Matsudaira K. The current state along with outstanding issues related to email-based guidance by physical therapists aiming to prevent low back pain among workers. JJOMT 64: 113-8, 2016
22. Matsudaira K, Hara N, Oka H, Kunogi J, Yamazaki T, Takeshita K, Seichi S, Tanaka S. Predictive factors for subjective improvement in lumbar spinal stenosis patients with nonsurgical treatment: a 3-year prospective cohort study. Plos One 11:e0148584, 2016
23. Shimazu A, Matsudaira K, De Jonge J, Tosaka N, Watanabe K, Takahashi M: Psychological Detachment from Work during Nonwork Time: Linear or Curvilinear Relations with Mental Health and Work Engagement? Ind Health 54 :282-92, 2016
24. Sawada T, Matsudaira K, Muto Y, Koga T, Takahashi M: Potential risk factors for onset of severe neck and shoulder discomfort (Katakori) in Urban Japanese workers. Ind Health 54: 230-6, 2016
25. Sarquis LM, Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Vargas-Prada S, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SS, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Marziale MH, Harari F, Freire R, Harari N, Monroy MV, Quintana LA, Rojas M, Harris EC, Serra C, Martinez JM, Delclos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Freimann T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VC, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A, Salazar Vega EJ.: Classification of neck/shoulder pain in epidemiological research a comparison of personal and occupational characteristics, disability and prognosis among 12,195 workers from 18 countries. Pain 157: 1028-36, 2016
26. Matsudaira K, Hiroe M, Kikkawa M, Suzuki M, Isomura T, Oka H, Hiroe K, Hiroe K. Can

standing back extension exercise improve or prevent low back pain in Japanese care workers? J Man Manip Ther 23: 205-9, 2015

27. Katsuhira J, Matsudaira K, Yasui T, Iijima S, Ito A: Efficacy of a trunk orthosis with joints providing resistive force on low back load in elderly persons during static standing. Clin Interv Aging 10: 1413-20, 2015
28. Kikuchi N, Matsudaira K, Sawada T, Oka H: Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK) in patients with whiplash neck injury pain and/or low back pain. J Orthop Sci 20: 985-92, 2015
29. Oka H, Matsudaira K, Fujii T, Okazaki H, Shinkai Y, Tsuji Y, Tanaka S, Kato R: Risk factors for prolonged treatment of whiplash-associated disorders. Plos One 10: e0132191,2015
30. Takahashi M, Matsudaira K, Shimazu A: Disabling low back pain associated with night shift duration: sleep problems as a potentiator. Am J Ind Med 58: 1300-10, 2015.
31. Matsudaira K, Kawaguchi M, Isomura T, Inuzuka K, Koga T, Miyoshi K, Konishi H: Assessment of psychosocial risk factors for the development of non-specific chronic disabling low back pain in Japanese workers- Findings from the Japan epidemiological research of Occupation-related Back pain (JOB) study. Ind Health 53: 368-77, 2015
32. Ohya J, Miyoshi K, Oka H, Matsudaira K, Fukushima M, Nagata K: Optimal measurement for "posterolateral protrusion" of the vertebral artery at the craniovertebral junction using computed tomography angiography. J Craniovertebr Junction Spine 5: 151-6, 2014
33. Hasegawa T, Katsuhira J, Matsudaira K, Iwakiri K, Maruyama H. Biomechanical Analysis of Low Back Load when Sneezing. Gait Posture 40: 670-675, 2014
34. Matsudaira K, Konishi H, Miyoshi K, Isomura T, Inuzuka K: Potential risk factors of persistent low back pain developing from mild low back pain in urban Japanese workers. PLos One 9: e93924, 2014
35. Yamada K, Matsudaira K, Takeshita K, Oka H, Hara N, Takagi Y: Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey. Mod Rheumatol 24: 343-348, 2014
36. Omata Y, Hagiwara F, Nishino J, Matsudaira K, Kadono Y, Juji T, Mori T, Nakayama H, Nagase Y, Hirose J, Yasui T, Matsumoto T, Matsui T, Tohma S, Tanaka S: Vertebral fractures affect functional status in postmenopausal rheumatoid arthritis patients J Bone Miner Medab 32: 725-731, 2014
37. Matsudaira K, Kikuchi N, Murakami A, Isomura T: Psychometric properties of the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ). J Orthop Sci 19: 26-32, 2014

（発表誌名巻号・頁・発行年等も記入）

2. 学会発表

1. Matsudaira K, Suzuki M, Sawada T, Sato E, Isomura T. Usefulness of “One Stretch”, a simple, daily, standing back extension exercise, for the prevention of onset or aggravation of low back pain in care workers. ISSLS Annual Meeting in Seoul, Korea, 2014.06.3-7
2. Matsudaira K, Sawada T, Kikuchi N,Sato E, Suzuki M. Workaholism as a risk factor for depression and disabling back pain among Japanese workers. ISSLS Annual Meeting in Seoul, Korea, 2014.06.3-7

H. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

Ⅱ.総合研究分担報告

労災疾病臨床研究事業費補助金

総合研究分担報告書

サブテーマ①腰痛に関わる実態およびリスクの同定

「医療介護職場における腰痛の実態の検討」

研究分担者　岡敬之　東京大学医学部附属病院 22 世紀医療センター
運動器疼痛メディカルリサーチ&マネジメント講座

研究要旨

研究分担者の専門の一分野は統計解析であり、本研究班においても **H26-28 年度**を通じて、各サブテーマで統計解析を担当した。自身のサブテーマは腰痛に関わる実態およびリスクの同定であり、**H26 年度**には、研究分担者の三上らと看護師を含む労働者を対象に，前向きに追跡したコホート研究である JOB (Japan epidemiological research of Occupation-related Back pain) study の 1 年間の追跡データを解析に用いて、仕事に支障をきたす腰痛の遷延化に関わる危険因子を探索的に検討した。この結果，働きがい，不安感，および仕事や生活の満足度が，遷延化の危険因子であることを明らかにした．**H27 年度**には、研究分担者の唐司とともに長野県の介護福祉施設（2 施設を積極介入群、1 施設をコントロール群）として、研究代表者が開発した‘これだけ体操’の効果を検証した。自覚的改善度、対策の実行度とも対照群と比較し、積極介入群のほうが統計的に有意に優れてており‘これだけ体操’の有用性が示唆された

H28 年度には単一医療介護施設の職員を対象に、自記式調査票を用いた横断研究において、腰痛の現状およびその関連因子について心理社会的要因を中心に探索した。この結果、FABQ（恐怖回避思考が強い）、SSS-8（身体症状が強い）、職場での対人関係でのストレスが有意な因子として抽出された。職場での腰痛対策には、上記の心理社会面に配慮した介入が必要であることが示唆された。

A. 研究目的

腰痛は世界共通の健康問題であり、特に勤労者においては腰痛が労働生産性の低下の主要因とされている。また、世界疾病負担研究において、腰痛は、障害生存年数（Years Lived with Disability)、つまり健康でない状態で生活する年数を指標とする統計で、289 の疾患や傷病の中でトップに位置しており、社会的損失や健康面への影響が大きい腰痛への対策は世界的に重要な課題といえる。

厚生労働省の発表した業務上疾病発生状況等調査によると、腰痛により 4 日以上を休職した業務上疾病の発生件数は、全職業性疾病の約 6 割を占め第 1 位となっている。業種別にみると、運輸交通業、保健衛生業、製造業、商業・金融・広告業、貨物取扱業などが、業務上疾病による腰痛発生が多く、特に近年では、保健衛生業の腰痛が 10 年で 2.7 倍という最も顕著な増加となっており、医療介護現場での腰痛対策は喫緊の課題といえる。

上記を踏まえ、平成 6 年に厚生労働省から発表されていた「腰痛予防対策指針」（2013 年）が 19 年ぶりに改訂され、新指針では適用対象を拡大し、福祉・医療分野における介護・看護作業も対象となった。また、新指針での変更点で注目すべきは、腰痛の新規発症要因として「動作要因」「環境要

因」「個人的要因」に加えて、「心理・社会的要因」が追加されたことである。これまでの腰痛発症に関する研究は、身体的負荷など人間工学的側面に重点を置いて検討されていたが、近年では精神的ストレスや職場での対人関係など心理社会的要因も腰痛の発症に影響することが明らかとなってきている。

保健衛生業領域における適切な腰痛対策の基礎資料とするため、1) **H26 年度** 腰痛の遷延化に関わる危険因子の探索的検討、2) **H27 年度** 介護福祉施設における研究代表者が開発した‘これだけ体操’の効果の検証、3) **H28 年度** 医療介護現場における腰痛の実態調査および腰痛に関連する心理社会的要因の網羅的探索をテーマに本研究を遂行した。

B. 研究方法

H26 年度 腰痛の遷延化に関わる危険因子の探索的検討

本邦勤労者の「仕事に支障をきたす非特異的腰痛」の危険因子について心理社会的要因を含む多因子の中から探索することを主目的とし，JOB（Japan epidemiological research of Occupation-related Back pain）study を実施した。今回は，看護師を含む JOB study のデータベースから，当該分野の今後の両立支援に役立てる情報として，LBD 慢性（遷延）化に関わる危険因子を探索的に検討した。

JOB study では，腰痛およびそれに関連しうる多要因（多くの個人的要因，人間工学的要因，心理社会的要因等）を網羅した自己記入式調査票を作成し，郵送法により実施することとした。腰痛の範囲は，肋骨縁より下部で下殿溝より上部として明確に図示し，これに下肢痛を伴う場合も含むと定義した。

また，仕事への支障度によって腰痛の程度を，Von Korff らの grading を参照に，4 段階で評価した。このうち，grade 2 と 3 を仕事に支障をきたす支障度の強い腰痛（LBD）とした。以下，具体的な grade の定義を示す。

Grade 0:腰痛を伴うことはなかった

Grade 1:腰痛を伴うことはあったが，仕事に支障をきたすことはなかった

Grade 2:腰痛のため仕事に支障をきたしたこともあったが，休職はしなかった

Grade 3:腰痛のため休職をした

心理社会的要因に関する質問票としては，厚生労働省の職業性ストレス簡易調査票を用いた。

主に首都圏の 4 労働者健康福祉労災病院（看護師）を含む 16 事業所の勤労者 6,140 名（18 歳以上）に前向き調査を依頼，ベースラインの調査では 5310 人から回答を得て（回収率 86%）翌 1 年の腰痛状況（前述の grade）等の追跡調査を行った。そのうちベースライン（調査時直近の 1 ヶ月）に LBD があった者を抽出し，翌 1 年間に同等の LBD を 3 ヶ月以上経験したこと，つまり LBD が遷延化したこと（従属変数）の危険因子をベースライン時に収集した変数を用い探索した。なお，追跡した 1 年間に腰痛以外の理由で業務内容に変更のあった人，交通外傷で腰痛になった人，骨折等特異的な疾患で腰痛を伴った人は解析するにあたり除外した。

独立変数は，ベースライン時の個人的要因として年齢（40 代未満/40 代/50 代以上），性別，肥満（Body Mass Index：BMI 25kg/㎡以上を肥満ありと定義），喫煙習慣（Brinkmann 係数 400 以上をヘビースモーカーと定義），学歴（最終学歴を中学・高校卒等と大学・短大・専門学校等で区分），平均睡眠時間（5 時間未満を短睡眠時間と定義），運動習慣（過去 1 年間の定期的な運動習慣の有無），腰痛体操習慣の有無，前屈の柔軟性（指先が足首に届かない場合を柔軟性が乏しいと定義），通院が必要な併存症の有無，生活習慣病の有無，腰痛既往の有無，作業の経験年数（5 年未満か以上か），人間工学的（身体への負荷）要因として，前屈動作，捻り動作，持ち上げ動作，物を押す動作（それぞれ 1 日作業時間おいて半分以上行っているか否か），重量物取り扱い作業（取り扱いなし（主

にデスクワーク）/取り扱い 20kg 未満/20kg 以上あるいは介護作業）とした。

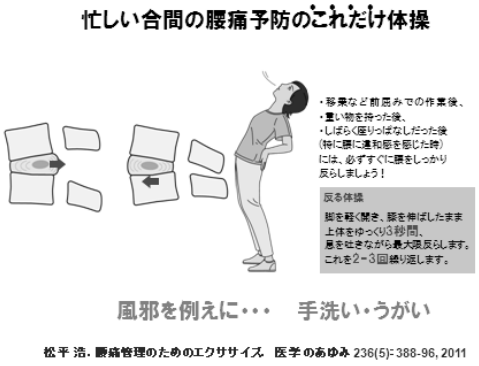
心理社会的要因としては，前述した職業性ストレス簡易調査票から算出した 19 因子（心理的な仕事の負担，職場での対人関係でのストレス，仕事のコントロール度, 働きがい, 疲労感, 不安感, 抑うつ感，上司及び同僚からのサポート等）の 5 段階評価の結果を 2 群に分けた評価（例：「仕事の量的負担」は，1-3 段階（少ない/やや少ない/普通）と 4-5 段階（やや多い/多い）の 2 群）に加え，単調な反復作業と感じているか否か，勤務形態（日中の勤務か夜勤のある不規則な交代制か），仕事時間（週 60 時間未満か以上か），雇用形態（正社員か否か），職場での腰痛対策の有無，小児期（14 歳以下）の心的外傷歴があり現在もそれが影響しているか否か，生活や仕事に支障をきたす腰痛の家族歴の有無とした。

統計解析についてはロジスティック回帰分析によりオッズ比を求め，危険因子の評価指標とした. 単変量解析により粗オッズ比と 95%信頼区間を求めた後，統計的に有意な関連を認めた要因から多重共線性と解釈可能性を考慮して選択したものを独立変数として多変量解析を行ない，要因調整オッズ比（性・年齢でも調整）とその 95%信頼区間を算出した。統計的検定は両側で行い，有意水準は 5%とした。統計パッケージは STATA 9.0 (StataCorp, LP, College Station, TX)を用いた。

H27 年度 介護福祉施設における研究代表者が開発した‘これだけ体操’の効果の検証

対象であるが、腰痛対策のスペシャリストである分担研究者が、伸展体操プログラムを推進するにあたり妥当である「前かがみ作業」に従事することが多い社会福祉法人の介護士とターゲットとした。具体的には長野県の 3 つの介護福祉施設（依田窪、みまき、ベルポート）のうち 2 施設を積極介入群、1 施設をコントロール群とした。両群間で施設の入居者数、入居者の平均介護度、認知症者の割合、障害の程度の割合、職員の数、性差、年齢に差はなかった。積極介入群は、体操を主とするマニュアルを配布するだけでなく、マニュアルの内容に関する 30 分の講義を受け、さらに参加型形式で業務中の伸展体操を積極的に行う群とした。各施設の担当者と本研究の分担研究者の協議により、予防体操（具体的には立位で腰を反らす‘これだけ体操’）を勤務中に習慣化する仕組みを構築していただいた。対照群は、マニュアルの配布のみとした。尚、本介入研究は、対象とした地域にちなんで「しなのプロジェクト」と名づけた。

以下に、**‘これだけ体操’**の内容を提示する。



シンプルな本体操のコンセプトは、動作・姿勢に依存する、言い換えればメカニカルな要素が明確な場合にターゲットを絞った診断・治療体系の代表である McKenzie 法（mechanical diagnosis and therapy）の derangement syndrome（椎間板内における髄核の変位に依存するとした腰痛モデル）に基づいている。Derangement syndrome はメカニカルストレスに伴う最もポピュラーな腰痛のサブグループであると考えられ、その中でも最も多いのが、前屈姿勢・作業により誘発される後方 derangement（髄核の後方への移動・陥凹）と想定されるパターンであり、伸展運動により改善しやすい。Zou J らは，立位での動的な MRI による検討により、変性が乏しい椎間板では McKenzie の理論モデル（後方 derangement syndrome のメカニズム）が妥当で

あることを報告している。

アウトカム評価は、介入直前（ベースライン）と介入開始から1年後に自記式調査票を用い行った。主要評価項目は、介入1年後の**自覚的改善度**（腰痛の状態は1年前と比較し、改善／不変／悪化）及び対策の**自覚的実行度**（実行／未実行）とした。副次的評価として、ベースライン時と介入1年後の**腰痛による通院状況の推移**及び**Oswestry Disability Index** (ODI：最小値0、最大値100)を評価した。先行研究によると機能障害を伴う腰痛はODI値が12以上であることから、ODI≧12である割合も比較した。

H28年度 医療介護現場における腰痛の実態調査および腰痛に関連する心理社会的要因の網羅的探索

研究デザインは横断研究とした。対象は、某医療法人社団の職員280名とし、無記名の自記式質問票を郵送にて回収した。

調査項目は、基本情報（年齢・性別・BMI・職種)、腰痛の有無、仕事のストレス要因（職業性ストレス簡易調査票より抜粋)、心の健康状態（K6)、恐怖回避思考（Fear-Avoidance Beliefs Questionnaire:FABQ)、身体症状(Somatic Symptom Scale-8：SSS-8)、仕事依存度である。腰痛におけるdisabilityの程度は4段階とした(Grade 1、腰痛はなかった；Grade 2、腰痛はあったが仕事に支障はなかった；Grade 3、腰痛のため仕事に支障をきたしたこともあったが休職はしなかった；Grade 4、腰痛のため休職したことがある)。本研究では、disabilityの強い（仕事への支障度が高い）腰痛に注目するため、仕事に支障をきたす群（Grade 3、4）とそうでない群（Grade 1、2）の2群に分けて解析を行った。仕事に支障をきたす腰痛の有無により、各評価項目を群間比較（t検定、Fisher正確検定）し、その後、腰痛の有無を目的変数、各評価項目を説明変数として、傾向スコアで調整するロジスティック回帰モデルで解析を行った。

C. 研究結果

H26年度 腰痛の遷延化に関わる危険因子の探索的検討

ベースライン時に回答した参加者 5,310 名中，3,811 名が 1 年フォローアップに回答した（追跡率：71.8%)。平均年齢 42.9 歳（標準偏差 10.1)，男性が 80.6%を占めた。追跡できなかった集団と，年齢，性別，BMI は，近似しており統計学的有意差はなかった。

171 名（平均年齢 41.5 歳，標準偏差 10.2，男性 71.4%，平均 BMI 23.0 kg/m2）がベースラインでの直近 1 カ月に LBD をもっており，そのうち 79 名（46.2%）が，看護・介護あるいは 20kg 以上の重量物を取り扱う作業に従事していた。

171 名中 29 名（17.0%）がフォローアップ期間中，３ヶ月以上の LBD（仕事に支障のある腰痛）を経験していた。

ロジスティック回帰分析の結果，働きがい（オッズ比 3.62,95%信頼区間 1.17 - 11.2，p ＝0.025)，不安感(オッズ比 2.89,95%信頼区間 0.97 - 8.57，p ＝0.056，および仕事や生活の満足度（オッズ比 4.14，95%信頼区間 1.18-14.58，p ＝0.027）が，他要因を調整しても LBD 遷延化の危険因子として挙げられた。

H27年度 介護福祉施設における研究代表者が開発した‘これだけ体操’の効果の検証

両群併せて167名（積極介入群：89名、対照群：78名）の介護士をエントリーした。167名の平均年齢は37.5才で、女性が65名（81.4%）であった。過去1年間に腰痛がなかった割合は29.9%にとどまり、全体の10.2%が過去1年を総合的に考えて腰痛にため仕事に支障をきたしていた。

尚、介入終了後のアウトカム評価は、1年後に行った自記式調査票への記入を完了した例を対象に分析を行った。

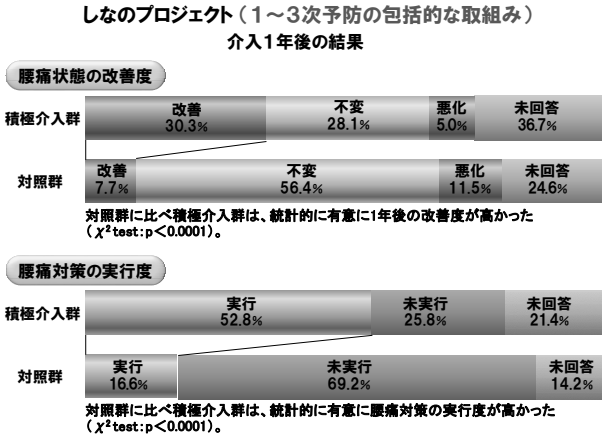
以下に、両群における分析対象者のベースライ

ン時での背景情報を示す。年齢、性別、BMI、喫煙習慣に加え、腰痛に対する通院状況、ベースライン時での直近１カ月の腰痛状態（程度)、ODIに関し、両群の間に統計的有意差はなかった。

Table 1.

	積極介入群 (n = 89)	対照群 (n = 78)	P
年齢	37.5 ± 12.4	37.6 ± 11.6	0.97
女性（%）	74 (83.1)	62 (79.5)	0.56
BMI	22.5 ± 3.8	22.3 ± 3.5	0.74
喫煙習慣（%）	39 (43.8)	37 (47.4)	0.28
腰痛による通院あり（%）	9 (10.1)	16 (20.5)	0.06
直近１ヵ月の腰痛状況			
● 腰痛なし	26 (29.2)	24 (30.8)	
● 腰痛はあるが仕事に支障はない	54 (60.7)	40 (51.3)	0.37
● 腰痛があり仕事に支障がある	4 (4.5)	11 (14.1)	
● 腰痛があり休職を要する	1 (1.1)	1 (1.3)	
ODI	9.8 ± 1.0	11.5 ± 1.1	0.28
ODI≧12	30 (33.7)	32 (41.0)	0.33

次に、主要評価項目の群間比較結果を示す。自覚的改善度、対策の実行度とも対照群と比較し、積極介入群のほうが統計的に有意に優れていた。



また、副次的評価項目とした腰痛による通院状況および ODI を以下に示す。積極介入群では腰痛による医療機関への通院機会は有意に減少していた。また、介入後には ODI≧12 である割合が、積極介入群でコントロール群よりも優位に少なかった。

Table 2.

	積極介入群 (n = 89)	対照群 (n = 78)	P
腰痛による通院あり			
介入前（%）	9 (10.1)	16 (20.5)	0.06
介入後（%）	3 (3.4)	12 (15.4)	*0.007
ODI≧12			
介入前（%）	30 (33.7)	32 (41.0)	0.33
介入後（%）	26 (29.2)	37 (47.4)	*0.04

次に、腰痛体操の実行度による腰痛の自覚的改善度を示す。積極介入群、対照群ともに実行度が高いほど腰痛の自覚的改善が得られた。

Table 3. 腰痛体操実行度による自覚的腰痛の変化

実行度		n	改善	不変	悪化
実行	積極介入群（%）	38	22 (57.9)	12 (31.6)	4 (10.5)
	対照群（%）	10	3 (30.0)	7 (70.0)	0 (0)
未実行	積極介入群（%）	17	4 (23.5)	12 (70.6)	1 (7.5)
	対照群（%）	48	3 (6.3)	37 (77.1)	8 (16.7)

H28 年度 医療介護現場における腰痛の実態調査および腰痛に関連する心理社会的要因の網羅的探索

アンケートの有効回答者数 203 名（72.5%）であった。平均年齢は 39.8 歳（SD 12.2)、性別は 70%が女性であった。63.1%が看護および介護業務職であった（表 1)。

表 1　対象者の属性（n = 203）

		n (%)
年齢(歳), 平均(SD)		39.8 (12.2)
性別	男性	61 (30.0)
	女性	142 (70.0)
BMI, 平均(SD)		22.6 (4.1)
職業	看護・介護関係	128 (63.1)
	それ以外	75 (36.9)
恐怖回避思考(FABQ)	Low	172 (85.6)
	High	29 (14.4)
仕事満足度	Not satisfied	51 (26.2)
	Satisfied	144 (73.8)
仕事負担量	Not stressed	126 (62.7)
	Stressed	75 (37.3)
職場での対人関係ストレス	Not stressed	163 (81.5)
	Stressed	37 (18.5)
仕事のコントロール度	Cotrolled	147 (72.8)
	Not controlled	55 (27.2)
上司からのサポート	Supported	114 (57.6)
	Not supported	84 (42.4)
同僚からのサポート	Supported	151 (75.9)
	Not supported	48 (24.1)
家族、友人からのサポート	Supported	58 (29.1)
	Not supported	141 (70.9)
心の健康状態(K6)	Low	103 (50.7)
	Middle	54 (26.6)
	High	46 (22.7)
身体症状(SSS-8)	other	128 (64.0)
	Very high	72 (36.0)
仕事依存度	Low	63 (31.2)
	Middle	73 (36.1)
	High	66 (32.7)

対象者のうち、仕事に支障をきたした腰痛経験者は 36 名（17.7%）であった。上記腰痛経験者の群とそれ以外の群で、各調査項目について群間比較を行ったところ、FABQ（恐怖回避思考、p = 0.037)、SSS-8（身体症状、p = 0.0003)、職場での対人関係ストレス（p = 0.022）が統計学的な有意差を認めた。年齢、性別、BMI、職業を調整したロジスティック解析においても、上記 3 要因が有意な因子として抽出された（表 2)。

表 2　多重ロジスティック解析により抽出された要因

要因	Adjusted OR	95% CI	*p*-value
恐怖回避思考	2.619	1.003-6.538	0.049
対人関係ストレス	2.619	1.067-6.224	0.036
身体症状	4.034	1.819-9.337	< 0.001

D. 考察

H26 年度 腰痛の遷延化に関わる危険因子の探索的検討

LBD が慢性(遷延)化に影響する要因は，仕事や生活での満足度が低い，働きがいが乏しい，不安感が強いといった、職場を主とする心理社会的要因であることが，今回の検討から示唆された。本知見は，欧米のエビデンスと矛盾しない。

メカニズムとしては，心理社会的要因がストレッサーとなり，中脳辺縁系ドパミン・オピオイドシステムの機能異常に続発する下行性疼痛調節系や自律神経系のアンバランスに伴う中枢性感作（痛覚過敏）や局所の血流低下・筋攣縮などが考えられる。

我が国の産業衛生分野において，人間工学的なアプローチによる腰痛の予防や対策が主流であった。しかしながら，厚生労働省業務上疾病発生状況等調査にて、腰痛における休業 4 日以上の業務上疾病の発生件数は，長年にわたり全職業性疾病の約 6 割を占め第 1 位であり，近年の世界疾病負担研究においても 289 の疾患や傷病のうち，腰痛が Years Lived with Disability (YLDs)のトップにランクているなど，我が国においてもグローバルにも労働現場での LBD は減少していない。言い換えれば，従来の対策が奏功しているとは言い難い。

今後は，LBD 遷延化による職場における労働力の損失を予防・緩和するための今後の対策として，心理社会的要因へのアプローチも重要視する必要があると思われる。

結果の一般化には限界があるが，職業そのものではなく，作業形態を調整したことにより，この結果が介護・看護作業はもちろんのこと，他の業種についても参考にしてよい知見であると思われる。

H27 年度 介護福祉施設における研究代表者が開発した‘これだけ体操’の効果の検証

本研究での主要な介入は、McKenzie 法

（mechanical diagnosis and therapy）の後方 derangement の予防対策である立位で腰を反らす‘これだけ体操’を、勤務中に習慣化する仕組みを構築したことである。対策実行者が、対照群に対して積極介入群では優位に多かったことから仕組みの構築は概ね成功したと思われ、そのことが腰痛状況の改善につながったと解釈できる。腰痛状況の改善には、積極介入群では腰痛による通院状況も有意に改善したことも含まれ、本介入は医療経済的にも有益な効果をもたらしうることが示唆された。また、各群ともに本体操の実行度が高いほど腰痛の自覚的改善が得られたことから、本体操自体の有効性を示すことができた。

H28 年度 医療介護現場における腰痛の実態調査および腰痛に関連する心理社会的要因の網羅的探索

産業衛生領域の腰痛対策を効率的に行うための基礎資料として、単一医療介護施設職員に対して、腰痛の実態およびその関連要因について心理社会的要因を中心にアンケート調査を行った。対象者の 17.7%が仕事に支障をきたす腰痛を経験していた。腰痛の関連要因として、恐怖回避思考、身体症状、職場での対人関係ストレスが抽出された。

恐怖回避思考とは、痛みに対する強い不安感や恐怖感から活動を過剰に制限（回避）してしまう思考のことをいう。腰痛の慢性化の予後規定因子である心理社会的要因（yellow flag sign）の中でも、この恐怖回避思考は機能障害や就業状況の予後に強く影響し、最も重要視すべきものとされている。勤労者の腰痛を慢性化させないために、上記概念を考慮した早期のスクリーニングが必要と考えられる。

今回の研究では、身体症状が高いオッズ比を示した。身体症状は、一般に精神の症状が身体の不調・不具合として身体化したものであり、頭痛、眩暈、胃腸の不調、疲労・活力低下といった愁訴として報告されている。これらは心理的ストレスが脳機能に影響を与えることにより生じる症状であり、腰痛にも心理的ストレスによる脳機能の不具合（dysfunction）を介し、筋緊張などの局所症状が強まる可能性があると思われる。身体症状は筋骨格系疼痛などの健康状態と関連があるとされており、また腰痛慢性化の一因であるとの報告が増えてきている。心理・社会的要因の強い腰痛では、さまざまな身体症状をあわせもつ場合が想定されるため、診療では注意深く問診することが必要であると考えている。

本結果では、職場での対人関係ストレスも仕事に支障をきたす腰痛との関連が示唆された。我が国の 5310 名の勤労者を対象としたコホート研究において、仕事関連ストレスが腰痛の発症および慢性化に影響することが示されており、職場での腰痛対策にはストレス要因を包含する必要がある。厚生労働省が発表している「腰痛予防対策指針（2013)」の中にも、「職場の対人ストレスに代表される心理社会的要因」との記載があり、対人関係ストレスが腰痛における重要な因子であることがわかる。

E. 結論

以上により腰痛に関わる実態およびリスクの同定ための次の知見を得た。

1) 仕事に支障をきたす非特異的腰痛が遷延することの危険因子は，仕事や生活での満足度が低い，働きがいが乏しい，不安感が強いといった、職場を主とする心理社会的要因であることが示唆された。

2) 腰痛を抱える対象者に対し、運動器のエキスパートでない現場の指導員が簡便に迷い無く指導できる指針予防的・治療的体操メニューを考案し、その実行のための集団アプローチを行い、その有用性を検証した。

3) 医療介護職場における腰痛の関連因子は、恐怖回避思考・身体症状・対人関係でのストレスであ

った。職場での腰痛対策には、上記の心理社会面に配慮した介入が必要であることが示唆された。

F. 健康危険情報

該当なし

G. 研究発表

1. 論文発表

1. Muraki S, Akune T, En-yo Y, Yoshida M, Tanaka S, Kawaguchi H, Nakamura K, Oka H, Yoshimura N. Association of dietary intake with joint space narrowing and osteophytosis at the knee in Japanese men and women: The ROAD Study. ***Mod Rheumatol*** 24, 236-242, 2014
2. Muraki S, Akune T, Nagata K, Ishimoto Y, Yoshida M, Tokimura F, Tanaka S, Oka H, Kawaguchi H, Nakamura K, Yoshimura N. Association of knee osteoarthritis with onset and resolution of pain and physical functional disability: The ROAD study. ***Mod Rheumatol*** 24:966-973, 2014
3. Hashizume H, Yoshimura N, Nagata K, Miyazaki N, Ishimoto Y, Nishiyama R, Oka H, Yamada H, Yoshida M: Development and evaluation of a video exercise program for locomotive syndrome in the elderly. ***Mod Rheumatol*** 24, 250-257, 2014
4. Yoshimura N, Nagata K, Muraki S, Oka H, Yoshida M, Enyo Y, Kagotani R, Hashizume H, Yamada H, Ishimoto Y, Teraguchi M, Tanaka S, Kawaguchi H, Toyama Y, Nakamura K, Akune T: Prevalence and progression of the radiographic ossification of posterior longitudinal ligament and its associated factors in the Japanese populations: A 3-year follow-up of the ROAD study. ***Osteoporos Int*** 25, 1089-1098, 2014
5. Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Tokimura F, Yoshida H, Suzuki T, Nakamura K, Yoshimura N: Incidence of certified needed care in the long-term care insurance system and its risk factors in the elderly of Japanese population-based cohorts: the ROAD study. ***Geriatrics & Gerontology International*** 14, 695-701, 2014
6. Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Tokimura F, Yoshida H, Suzuki T, Nakamura K, Yoshimura N. Association of physical activities of daily living with the incidence of certified need of care in the long-term care insurance system of Japan: the ROAD study. ***J Orthop Sci*** 19:489-496, 2014
7. Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Yoshimura N. Exercise habits during middle age are associated with lower prevalence of sarcopenia: the ROAD study. ***Osteoporos Int*** 25:1081-1088, 2014
8. Yamada K, Matsudaira K, Takeshita K, Oka H, Hara N, Takagi Y: Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey. ***Mod Rheumatol*** 24, 343-348, 2014
9. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, Oka H, Ishimoto Y, Nagata K, Kagotani R, Takiguchi N, Akune T, Kawaguchi H, Nakamura K, Yoshida M. Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study. ***Osteoarthritis Cartilage*** 22:104-110, 2014
10. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, Minamide A, Nakagawa H, Ishimoto Y, Nagata K, Kagotani R, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshida M. The association of combination of disc degeneration, endplate signal change, and Schmorl node with low back pain in a large population study: the Wakayama Spine Study. ***Spine J*** pii: S1529-9430(14)01758-6. doi: 10.1016/j.spinee.2014.11.012, 2014
11. Nagata K, Yoshimura N, Hashizume H, Muraki S, Ishimoto Y, Yamada H, Takiguchi N, Nakagawa Y, Minamide A, Oka H, Kawaguchi H, Nakamura K, Akune T, Yoshida M. The prevalence of cervical myelopathy among subjects with narrow cervical spinal canal in a population-based magnetic resonance imaging study: the Wakayama Spine Study. ***Spine J*** 14:2811-2817, 2014
12. Yoshimura N, Akune T, Fujiwara S, Shimizu Y, Yoshida H, Omori G, Sudo A, Nishiwaki Y, Yoshida M, Shimokata H, Suzuki T, Muraki S, Oka H, Nakamura K. Prevalence of knee pain, lumbar pain and its coexistence in Japanese men and women: The Longitudinal Cohorts of Motor System Organ (LOCOMO) study. ***J Bone Miner Metab*** 32:524-532, 2014
13. Ohya J, Miyoshi K, Oka H, Matsudaira K, Fukushima M, Nagata K. Optimal measurement for "posterolateral protrusion" of the vertebral artery at the craniovertebral junction using computed tomography angiography. ***J Craniovertebr Junction Spine*** 5:151-156, 2014
14. Ito H, Takatori Y, Moro T, Oshima H, Oka H, Tanaka S. Total Hip Arthroplasty After Rotational Acetabular Osteotomy. ***J Arthroplasty*** 2014 Oct 8. pii: S0883-5403(14)00746-3. doi: 10.1016/j.arth.2014.10.002, 2014
15. Minamide A, Yoshida M, Yamada H, Hashizume H, Nakagawa Y, Nishi H, Iwasaki H, Tsutsui S,

Okada O, Okada S, Oka H. Efficacy of Posterior Segmental Decompression Surgery for Pincer Mechanism in Cervical Spondylotic Myelopathy: A Retrospective Case-controlled Study Using Propensity Score Matching. ***Spine (Phila Pa 1976)*** 40:1807-1815, 2015.

16. Kikuchi N, Matsudaira K, Sawada T, Oka H. Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury pain and/or low back pain. ***J Orthop Sci*** 20:985-992, 2015.
17. Oka H, Matsudaira K, Fujii T, Okazaki H, Shinkai Y, Tsuji Y, Tanaka S, Kato R. Risk Factors for Prolonged Treatment of Whiplash-Associated Disorders. ***PLoS One*** 10:e0132191, 2015.
18. Muraki S, Akune T, Teraguchi M, Kagotani R, Asai Y, Yoshida M, Tokimura F, Tanaka S, Oka H, Kawaguchi H, Nakamura K, Yoshimura N. Quadriceps muscle strength, radiographic knee osteoarthritis and knee pain: the ROAD study. ***BMC Musculoskelet Disord*** 16:305, 2015.
19. Yamada H, Oka H, Iwasaki H, Endo T, Kioka M, Ishimoto Y, Nagata K, Takiguchi N, Hashizume H, Minamide A, Nakagawa Y, Kawai M, Tsutsui S, Yoshida M. Development of a support tool for the clinical diagnosis of symptomatic lumbar intra- and/or extra-foraminal stenosis. ***J Orthop Sci*** 20:811-817, 2015.
20. Minamide A, Yoshida M, Yamada H, Hashizume H, Nakagawa Y, Nishi H, Iwasaki H, Tsutsui S, Okada O, Okada S, Oka H. Efficacy of Posterior Segmental Decompression Surgery for Pincer Mechanism in Cervical Spondylotic Myelopathy: A Retrospective Case-controlled Study Using Propensity Score Matching. ***Spine (Phila Pa 1976)*** 40:1807-1815, 2015.
21. Yoshimura N, Muraki S, Oka H, Tanaka S, Ogata T, Kawaguchi H, Akune T, Nakamura K. Association between new indices in the locomotive syndrome risk test and decline in mobility: third survey of the ROAD study. ***J Orthop Sci*** 20:896-905, 2015.
22. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T. Factors affecting changes in the serum levels of 25-hydroxyvitamin D: a 3-year follow-up of the ROAD study. ***Osteoporos Int*** 26:2597-2605, 2015.
23. Mure K, Yoshimura N, Hashimoto M, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Takeshita T. Urinary 8-iso-prostaglandin F2α as a marker of metabolic risks in the general Japanese population: The ROAD study. ***Obesity (Silver Spring)*** 23:1517-1524, 2015.
24. Nakamura M, Hashizume H, Oka H, Okada M, Takakura R, Hisari A, Yoshida M, Utsunomiya H. Physical Performance Measures Associated With Locomotive Syndrome in Middle-Aged and Older Japanese Women. ***J Geriatr Phys Ther*** 38:202-207, 2015.
25. Yamada H, Terada M, Iwasaki H, Endo T, Okada M, Nakao S, Hashizume H, Minamide A, Nakagawa Y, Nishi H, Tsutsui S, Oka H, Yoshida M. Improved accuracy of diagnosis of lumbar intra and/or extra-foraminal stenosis by use of three-dimensional MR imaging: comparison with conventional MR imaging. ***J Orthop Sci*** 20:287-294, 2015.
26. Ito H, Takatori Y, Moro T, Oshima H, Oka H, Tanaka S. Total hip arthroplasty after rotational acetabular osteotomy. ***J Arthroplasty*** 30:403-406, 2015.
27. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, Minamide A, Nakagawa H, Ishimoto Y, Nagata K, Kagotani R, Tanaka S, Kawaguchi H, Nakamura K, Akune

T, Yoshida M. The association of combination of disc degeneration, end plate signal change, and Schmorl node with low back pain in a large population study: the Wakayama Spine Study. ***Spine J*** 15:622-628, 2015

28. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T. Mutual associations among musculoskeletal diseases and metabolic syndrome components: A 3-year follow-up of the ROAD study. ***Mod Rheumatol*** 25:438-448, 2015.
29. Kato S, Oshima Y, Oka H, Chikuda H, Takeshita Y, Miyoshi K, Kawamura N, Masuda K, Kunogi J, Okazaki R, Azuma S, Hara N, Tanaka S, Takeshita K. Comparison of the Japanese Orthopaedic Association (JOA) score and modified JOA (mJOA) score for the assessment of cervical myelopathy: a multicenter observational study. ***PLoS One*** 10:e0123022, 2015.
30. Muraki S, Akune T, En-Yo Y, Yoshida M, Suzuki T, Yoshida H, Ishibashi H, Tokimura F, Yamamoto S, Tanaka S, Nakamura K, Kawaguchi H, Oka H, Yoshimura N. Joint space narrowing, body mass index, and knee pain: the ROAD study (OAC1839R1). ***Osteoarthritis Cartilage*** 23:874-881,2015.
31. Ogihara S, Yamazaki T, Maruyama T, Oka H, Miyoshi K, Azuma S, Yamada T, Murakami M, Kawamura N, Hara N, Terayama S, Morii J, Kato S, Tanaka S. Prospective multicenter surveillance and risk factor analysis of deep surgical site infection after posterior thoracic and/or lumbar spinal surgery in adults. ***J Orthop Sci*** 20:71-77,2015.
32. Ohya J, Oshima Y, Takeshita K, Oka H, Chikuda H, Taniguchi Y, Matsubayashi Y, Tanaka S. Patient satisfaction with double-door laminoplasty for cervical compression myelopathy. ***J Orthop Sci*** 20:64-70,2015
33. Muraki S, Akune T, Nagata K, Ishimoto Y, Yoshida M, Tokimura F, Tanaka S, Kawaguchi H, Nakamura K, Oka H, Yoshimura N. Does osteophytosis at the knee predict health-related quality of life decline? A 3-year follow-up of the ROAD study. ***Clin Rheumatol*** 34:1589-1597, 2015.
34. Yoshimura N, Akune T, Fujiwara S, Shimizu Y, Yoshida H, Nishiwaki Y, Sudo A, Omori G, Yoshida M, Shimokata H, Suzuki T, Muraki S, Oka H, Nakamura K. Incidence of disability and its associated factors in Japanese men and women: the Longitudinal Cohorts of Motor System Organ (LOCOMO) study. ***J Bone Miner Metab*** 33:186-1891, 2015.
35. Katsuhira J, Matsudaira K, Oka H, Iijima S, Ito A, Yasui T, Yozu A. Efficacy of a trunk orthosis with joints providing resistive force on low back load during level walking in elderly persons. ***Clin Interv Aging***11:1589-159, 2016 .
36. Tonosu J, Inanami H, Oka H, Katsuhira J, Takano Y, Koga H, Yuzawa Y, Shiboi R, Oshima Y, Baba S, Tanaka S, Matsudaira K. Diagnosing Discogenic Low Back Pain Associated with Degenerative Disc Disease Using a Medical Interview. ***PLoS One***11:e0166031, 2016.
37. Nakamura M, Kobashi Y, Hashizume H, Oka H, Kono R, Nomura S, Maeno A, Yoshida M, Utsunomiya H. Locomotive syndrome is associated with body composition and cardiometabolic disorders in elderly Japanese women. ***BMC Geriat*** 16:166, 2016.
38. Teraguchi M, Samartzis D, Hashizume H, Yamada H, Muraki S, Oka H, Cheung JP, Kagotani R, Iwahashi H, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Cheung KM, Yoshimura N, Yoshida M. Classification of High Intensity Zones of the Lumbar Spine and Their Association with Other Spinal MRI Phenotypes: The Wakayama Spine Study. ***PLoS One***

11:e0160111, 2016.

39. Ono K, Ohashi S, Oka H, Kadono Y, Yasui T, Omata Y, Nishino J, Tanaka S, Tohma S. The impact of joint disease on the Modified Health Assessment Questionnaire scores in rheumatoid arthritis patients: A cross-sectional study using the National Database of Rheumatic Diseases by iR-net in Japan. ***Mod Rheumatol*** 26:529-33, 2016
40. Oshima Y, Takeshita K, Taniguchi Y, Matsubayashi Y, Doi T, Ohya J, Soma K, Kato S, Oka H, Chikuda H, Tanaka S. Effect of Preoperative Sagittal Balance on Cervical Laminoplasty Outcomes. ***Spine (Phila Pa 1976)*** 41:E1265-E1270, 2016.
41. Ohya J, Oshima Y, Oka H, Saiki F, Taniguchi Y, Matsubayashi Y, Tanaka S, Chikuda H, Takeshita K. Patient Satisfaction with Posterior Decompression Surgery for Cervical Ossification of the Posterior Longitudinal Ligament: Prognostic Radiographic Factors and Patient-Reported Outcomes for the Effectiveness of Surgical Treatment. ***World Neurosurg*** 96:272-279, 2016.
42. Tonosu J, Matsudaira K, Oka H, Okazaki H, Oshio T, Hanaoka I, Muraoka Y, Midorikawa M, Wakabayashi K, Tanaka S. A population approach to analyze the effectiveness of a back extension exercise "One Stretch" in patients with low back pain: A replication study. ***J Orthop Sci*** 21: 414-418,2016.
43. Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S. Psychometric Properties of the Japanese Version of the STarT Back Tool in Patients with Low Back Pain. ***PLoS One*** 11:e0152019, 2016.
44. Matsudaira K, Hiroe M, Kikkawa M, Sawada T, Suzuki M, Isomura T, Oka H, Hiroe K, Hiroe K. Can standing back extension exercise improve or prevent low back pain in Japanese care workers? J ***Man Manip The***r 23:205-209, 2015.
45. Kodama R, Muraki S, Oka H, Iidaka T, Teraguchi M, Kagotani R, Asai Y, Yoshida M, Morizaki Y, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshimura N. Prevalence of hand osteoarthritis and its relationship to hand pain and grip strength in Japan: The third survey of the ROAD study. ***Mod Rheumatol.*** 26:767-773, 2016
46. Matsudaira K, Hara N, Oka H, Kunogi J, Yamazaki T, Takeshita K, Atsushi S, Tanaka S. Predictive Factors for Subjective Improvement in Lumbar Spinal Stenosis Patients with Nonsurgical Treatment: A 3-Year Prospective Cohort Study. ***PLoS One*** 11:e0148584,2016.
47. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, Minamide A, Ishimoto Y, Nagata K, Kagotani R, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshida M. Metabolic Syndrome Components Are Associated with Intervertebral Disc Degeneration: The Wakayama Spine Study. ***PLoS One*** 11:e0147565, 2016.
48. Ohashi S, Ohnishi I, Oka H, Matsumoto T, Bessho M, Nakamura K, Tanaka S. The effect of cartilage degeneration on ultrasound speed in human articular cartilage. ***Mod Rheumatol*** 26:426-434, 2016.
49. Iidaka T, Muraki S, Akune T, Oka H, Kodama R, Tanaka S, Kawaguchi H, Nakamura K, Yoshimura N. Prevalence of radiographic hip osteoarthritis and its association with hip pain in Japanese men and women: the ROAD study. ***Osteoarthritis Cartilage*** 24:117-123, 2016.
50. Kikuchi N, Matsudaira K, Sawada T, Oka H. Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury

pain and/or low back pain. ***J Orthop Sci*** 20:985-992, 2016

51. Iwahashi H, Yoshimura N, Hashizume H, Yamada H, Oka H, Matsudaira K, Shinto K, Ishimoto Y, Nagata K, Teraguchi M, Kagotani R, Muraki S, Akune T, Tanaka S, Kawaguchi H, Nakamura K, Minamide A, Nakagawa Y, Yoshida M. The Association between the Cross-Sectional Area of the Dural Sac and Low Back Pain in a Large Population: The Wakayama Spine Study. ***PLoS One*** 11:e0160002, 2016.
52. Hara N, Matsudaira K, Masuda K, Tohnosu J, Takeshita K, Kobayashi A, Murakami M, Kawamura N, Yamakawa K, Terayama S, Ogihara S, Shiono H, Morii J, Hayakawa K, Kato S, Nakamura K, Oka H, Sawada T, Inuzuka K, Kikuchi N. Psychometric Assessment of the Japanese Version of the Zurich Claudication Questionnaire (ZCQ): Reliability and Validity. ***PLoS One***11:e0160183, 2016.
53. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, Takiguchi N, Minamide A, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshida M. Association of Lumbar Spondylolisthesis with Low Back Pain and Symptomatic Lumbar Spinal Stenosis in a Population-based Cohort: The Wakayama Spine Study. ***Spine (Phila Pa 1976)*** , in press.
54. Yoshimura N, Muraki S, Oka H, Iidaka T, Kodama R, Kawaguchi H, Nakamura K, Tanaka S, Akune T. Is osteoporosis a predictor for future sarcopenia or vice versa? Four-year observations between the second and third ROAD study surveys.***Osteoporos Int***, in press.

Yasui T, Oka H, Omata Y, Kadono Y, Tanaka S. Relationship between roentgenographic joint destruction in the hands and functional disorders among patients with rheumatoid arthritis. ***Mod Rheumatol***, in press.

2. 学会発表

第 9 回日本運動器疼痛学会、東京、2016

H. 知的財産権の出願・登録状況（予定を含む）

該当なし

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし

労災疾病臨床研究事業費補助金
総合研究分担報告書

サブテーマ①腰痛に関わる実態およびリスクの同定
「慢性腰痛患者のQOL心理的要因の検討」

研究分担者　小山善子　金城大学医療健康学部

研究要旨

H26-27年度に心理・社会的要因を考慮して介護施設労働者における腰痛の関連要因を多施設横断研究で介護施設に働く労働者を対象とし，自記式の質問表を使用して検討した．多変量ロジスティック回帰分析にて，ストレス要因による身体愁訴，STarT Back スクリーニングツールがハイリスク（心理的要因が強い）,TSK-J（恐怖回避思考が強い）が有意な結果で，心理・社会的要因の関与が示唆された．

しばしば抑うつに合併する身体愁訴＝身体化（somatization）は，文献的にも腰痛に影響していることが示唆される腰痛関連因子である．H28年度には心身の健康に関するインターネット調査にて労働力人口として現役世代である20-64歳の慢性腰痛のを持つ20-64歳の日本人約3,000人を対象に健康関連QOLと身体化，抑うつ症状の関連を評価した．解析では抑うつ，年齢，性別，BMI，喫煙，婚姻，学歴，定期的な運動，雇用状況，通院していた疾患の個数（0－25）を調整した．

身体化は抑うつを調整しても健康関連QOLと有意に関係していた．すなわち，慢性腰痛のある人で，身体化傾向の強い人ほど，健康関連QOLが低く慢性腰痛リスクとして身体化を視野に入れる必要がある．

A.研究目的

腰痛は病気の名前ではなく，症状の総称である．具体的には助骨縁より下部で，下殿溝より上記に限定的に起こる疼痛や不快感と定義される．その分類は，医師の診察や画像検査で痛みの原因が特定できる「特異的腰痛」と，さまざまな検査をうけても原因を特定しきれない「非特異的腰痛」に大別される．その割合は特異的腰痛については腰痛を理由に医療機関を受診する 10%程度であり，およそ 85%が非特異的腰痛に該当する．

近年，この非特異的腰痛の実態を分析するため，各国の研究者によりさまざまな角度から研究が行われている．その結果，非特異的腰痛のリスク要因は「重い荷物を持つ」,「姿勢の悪さ」など，腰自体への負担に関わる問題に加え，さまざまな心理・社会的な要因（心理的ストレッサー）が重要であることが明らかになってきた．特に難治化する要因の多くが心理・社会的要因である報告もある．

その中でも抑うつは腰痛発症の危険因子であるとともに，腰痛の慢性化の予測因子でもある．慢性腰痛の患者の中でも，抑うつのある患者ではない患者よりも生活の質が低いという報告がある．また直接的な医療費も，抑うつのある腰痛患者の方が無い患者よりも高い．そのため，医療現場で腰痛患者の抑うつ症状を評価することは，予後の判定や治療の選択に重要であると考えられる．

そこで我々は，多面的な心理・社会的要因を考

慮して介護施設労働者における腰痛の関連要因を検討する研究を実施した．腰痛は一般的な症状であるため本研究では，特に重症度の高い腰痛に注目して，H26-27 年度には介護施設に働く労働者の自記式質問表にてて介護施設労働者における腰痛の関連要因を検討した．

H28 年度には、前年度までの調査や文献的考察にて腰痛関連因子として推察された身体愁訴すなわち，身体化（somatization）に関する検討をインターネット調査にて行った．身体化とは，心理的ストレスに反応して，身体症状の不安を訴え，治療を求める傾向であるとされ，しばしば抑うつに合併する．過去の我々の研究で，身体化傾向は軽度の腰痛のあった人達の中で，持続する腰痛を発症することの予測因子であった．他の研究では，身体化傾向が腰痛患者の治療アウトカムと関連していた．腰痛と抑うつについての過去の研究は散見されるものの，慢性腰痛と身体化の関連についてはあまりよく分かっていない．

頻度の高い身体症状を評価するための，自己申告式質問票は複数存在するが，最近のシステマティックレビューは Patient Health Questionnaire-15 (PHQ-15) と 12 項目の Symptom Checklist-90 somatization scale が大規模研究では最も適切であると報告している．この 2 つの質問票は，重要な身体症状を含んでいながら比較的短く，かつ心理学的特性が確立されている．Somatic Symptom Scale-8 (SSS-8) は PHQ-15 の 8 項目から作成された．最終年度における目的は，労働力人口として現役世代である 20-64 歳の慢性腰痛のある日本人において，SSS-8 をもちいて評価した身体化傾向が，抑うつを考慮しても，健康関連 QOL と関連しているかを検討し、腰痛リスクとして重要視すべきかを検討することである．

B. 研究方法

H26-27 年度

研究デザインは多施設共同横断研究とした．対象は介護施設に働く労働者とし，施設の選定は石川県内の介護施設 125 箇所に調査用紙を郵送して本研究の趣旨に賛同をえた 95 箇所とした．1 施設 20 人を対象とし，無記名の自記式質問表を郵送にて回収した．

調査項目は，個人的要因（性別，年齢，学歴，婚姻），生活習慣（喫煙，運動習慣，睡眠時間，睡眠の質），労働要因（雇用形態，経験年数，職種，労働時間，夜勤の回数），心理・社会的要因（職業性ストレス調査票，日本語版 STarT Back スクリーニングツール，TSK-J）とした．

年代は，“10 代～35 歳未満”，“35 歳以上～50 歳未満 “，“50 歳以上”と区分した．学歴は，“中卒・高卒”，“専門学校・高専・短大卒・大学卒・大学院卒”と区分した．婚姻は，“未婚”，“既婚”と区分した．

喫煙は，“現在喫煙”，“現在非喫煙”と区分した．運動習慣は，30 分以上の運動を過去 1 ヵ月に平均週 2 回以上の実施とした．睡眠時間は過去 1 ヵ月の平均で，“5 時間未満”，“5 時間以上 6 時間未満”，“6 時間以上”と区分した．睡眠の質は次の 3 つの質問：①“入眠が 30 分以上”，②“中途覚醒が週 3 回以上”，③“早期覚醒が週 3 回以上”で調査し，該当数を“2 つ以上”，“1 つ”，“0 個”と区分した．

雇用形態は，“正社員”，“正社員以外”と区分した．経験年数は，“1 年未満”，“1 年以上 5 年未満”，“5 年以上”と区分した．職種は，“介護福祉士・ホームヘルパー”，“介護福祉士・ホームヘルパー以外”と区分した．労働時間は過去 1 ヵ月の 1 週間あたりの残業時間を含む労働時間が，“40 時間以内”，“41 時間以上 60 時間以内”，“61 時間以上”と区分した．夜勤の回数は過去 1 ヵ月の 1

週間あたりの回数が，“2 回以上”，“1 回”，“0 回”と区分した．

心理・社会的要因であるストレスの測定は，職業性ストレス調査票を用いた．57 の質問で構成され，①ストレスの原因と考えられる因子，②ストレスによっておこる心身の反応，③ストレス反応に影響を与える他の因子の 3 つの領域に大別され，さらに細かくは 18 の尺度に分類される．得点化は標準化得点法（下図）に沿って，各尺度について状態が一番悪い分類（灰色マーカー部）を“High”，それ以外を“Low”と区分した．

尺度	計算 No.は質問項目得点	男性					女性				
		低い／少い	やや低い／少い	普通	やや高い／多い	高い／多い	低い／少い	やや低い／少い	普通	やや高い／多い	高い／多い
		質問項目合計得点 下段は分布 (n=15, 933)					質問項目合計得点 下段は分布 (n=8, 447)				
【ストレスの原因と考えられる因子】											
心理的な仕事の負担（量）	15-(No.1+No.2+No.3)	3-5	6-7	8-9	10-11	12	3-4	5-6	7-9	10-11	12
		7.2%	18.9%	40.8%	22.7%	10.4%	6.6%	20.4%	51.7%	15.6%	5.8%
心理的な仕事の負担（質）	15-(No.4+No.5+No.6)	3-5	6-7	8-9	10-11	12	3-4	5-6	7-8	9-10	11-12
		4.5%	20.6%	43.4%	25.7%	5.7%	4.9%	17.5%	38.2%	29.1%	10.3%
自覚的な身体的負担度	5-No.7		1	2	3	4		1	2	3	4
			33.8%	39.3%	18.7%	8.2%		37.0%	33.7%	19.7%	9.6%
職場の対人関係でのストレス	10-(No.12+No.13)+No.14	3	4-5	6-7	8-9	10-12	3	4-5	6-7	8-9	10-12
		5.7%	24.8%	47.5%	17.6%	4.5%	7.3%	26.8%	41.0%	18.4%	6.4%
職場環境によるストレス	5-No.15		1	2	3	4	1		2	3	4
			25.1%	38.0%	23.1%	13.8%	17.7%		31.7%	28.8%	21.7%
仕事のコントロール度	15-(No.8+No.9+No.10)	3-4	5-6	7-8	9-10	11-12	3	4-5	6-8	9-10	11-12
		5.4%	16.6%	37.1%	32.4%	8.5%	5.5%	16.0%	48.8%	23.3%	6.3%
技能の活用度	No.11	1	2	3	4		1	2	3	4	
		4.5%	18.2%	49.4%	27.9%		9.1%	26.7%	45.6%	18.6%	
仕事の適性度	5-No.16	1	2	3		4	1	2	3		4
		6.4%	23.3%	54.9%		15.4%	9.3%	25.9%	49.7%		15.1%
働きがい	5-No.17	1	2	3		4	1	2	3		4
		7.3%	24.2%	51.4%		17.0%	13.1%	29.3%	44.5%		13.1%
【ストレスによっておこる心身の反応】											
活気	No.1+No.2+No.3	3	4-5	6-7	8-9	10-12	3	4-5	6-7	8-9	10-12
		10.9%	14.3%	41.6%	24.5%	8.7%	13.4%	19.2%	37.3%	21.3%	8.8%
イライラ感	No.4+No.5+No.6	3	4-5	6-7	8-9	10-12	3	4-5	6-8	9-10	11-12
		10.3%	20.9%	38.2%	22.7%	7.8%	7.6%	18.2%	45.1%	20.3%	8.8%
疲労感	No.7+No.8+No.9	3	4	5-7	8-10	11-12	3	4-5	6-8	9-11	12
		9.7%	12.2%	47.4%	23.3%	7.4%	6.2%	23.2%	40.1%	23.1%	7.4%
不安感	No.10+No.11+No.12	3	4	5-7	8-9	10-12	3	4	5-7	8-10	11-12
		8.3%	14.9%	51.9%	17.8%	7.1%	12.3%	15.6%	44.7%	21.6%	5.8%
抑うつ感	No.13~No.18の合計	6	7-8	9-12	13-16	17-24	6	7-8	9-12	13-17	18-24
		15.1	21.6%	40.6%	16.2%	6.5%	12.40%	18.9%	39.3%	22.3%	7.2%
身体愁訴	No.19~No.29の合計	11	12-15	16-21	22-26	27-44	11-13	14-17	18-23	24-29	30-44
		5.3%	31.0%	40.5%	15.9%	7.4%	8.3%	23.6%	38.6%	21.7%	7.8%
【ストレス反応に影響を与える他の因子】											
上司からのサポート	15-(No.1+No.4+No.7)	3-4	5-6	7-8	9-10	11-12	3	4-5	6-7	8-10	11-12
		6.9%	27.0%	32.8%	24.7%	8.7%	7.5%	22.0%	38.9%	26.7%	4.9%
同僚からのサポート	15-(No.2+No.5+No.8)	3-5	6-7	8-9	10-11	12	3-5	6-7	8-9	10-11	12
		6.1%	32.4%	39.9%	16.3%	5.3%	8.1%	31.3%	35.3%	17.9%	7.4%
家族・友人からのサポート	15-(No.3+No.6+No.9)	3-6	7-8	9	10-11	12	3-6	7-8	9	10-11	12
		6.9%	13.9%	20.3%	28.4%	30.6%	4.4%	10.6%	16.0%	28.6%	40.4%
仕事や生活の満足度	10-(No.1+No.2)	2-3	4	5-6	7	8	2-3	4	5-6	7	8
		5.0%	12.3%	57.2%	17.4%	8.1%	6.4%	15.4%	57.8%	15.4%	5.0%

さらに、近年、世界的に腰痛の難治化する予後を予想するツールとして注目されている STarT(Subgroup for Targeted Treatment) Back スクリーニングツール日本語版も測定した．5 つの質問から構成され，“そうだ”もしくは“とても/極めて”を 1 点として，本ツールを開発した英国キール大学の推奨に沿って（下頁図が領域得点と呼ばれる心理的要因の設問項目)、領域得点 4 点以上を“High risk”,4 点未満を“Low risk”と区分した．

	そうではない 0	そうだ 1
私のような体の状態の人は、体を動かし活動的であることは決して安全とはいえない	□	□
最近2週間は、心配事が心に浮かぶことが多かった	□	□
私の腰痛はひどく、決して良くならないと思う	□	□
以前は楽しめたことが、最近2週間は楽しめない	□	□
全般的に考えて、ここ2週の間に腰痛をどの程度煩わしく感じましたか？	全然 □ 0 / 少し □ 0 / 中等度 □ 0 / とても □ 1 / 極めて □ 1	

さらに，腰痛の最も重要な予後規定因子ともされる心理的要因である恐怖回避思考について，TSK-J（下図）にて測定した．11 の質問から構成され，該当するを 1 点として，加算した総得点を 3 分位して，“High”，“Medium”，“Low”とした．

1. 運動すると体を痛めてしまうかもしれないと不安になる
2. 痛みが増すので何もしたくない
3. 私の体には何か非常に悪いところがあると感じている
4. 他の人は私の体の状態のことなど真剣に考えてくれない
5. アクシデント(痛みが起こったきっかけ)のせいで，私は一生痛みが起こりうる体になった
6. 痛みを感じるのは，私の体を痛めたことが原因である
7. 不必要な動作を行わないよう，とにかく気をつけることが，私の痛みを悪化させないためにできる最も確実なことである
8. この強い痛みは私の体に何か非常に悪いことが起こっているからにちがいない
9. 体を痛めないために，痛みを感じたら私は運動をやめる
10. 私はとても体を痛めやすいので，全てのことを普通の人と同じようにできるわけではない
11. 痛みがある時は，誰であっても運動することを強要されるべきではない

腰痛の程度は 4 段階（grade 0：腰痛無し，grade 1：支障のない腰痛，grade 2：支障はあるが欠勤しなかった腰痛，grade 3：腰痛のため欠勤したことがある）で評価し，grade2 以上かつ 3 ヵ月以上継続した腰痛を重症度の高い腰痛と定義した．

統計解析は記述統計と，重症度の高い腰痛との関連要因を検討するためロジスティック回帰を用いた．ロジスティック回帰の結果はオッズ比と 95%信頼区間で示した．まず各要因について重症度の高い腰痛ありをイベントとしたロジスティック回帰にて粗オッズ比を算出した．粗オッズ比の p 値が 0.05 以下の要因について，変数同士の関連性を検討した．独立と思われる要因を用いて，多変量ロジスティック回帰のステップワイズで p 値 0.05 を基準として要因を検討した．全ての解析は両側で有意水準を 0.05 とした．解析ソフトは SAS9.3 を使用した．

H28 年度

最終年度の研究の対象は慢性腰痛のある 20-64 歳の日本人 3100 人である．2015 年にインターネットにて実施された「心身の健康に関する調査」のデータを用いた．参加者はインターネット調査会社（United Inc.）により募集された. 年齢が 20-64 歳である約 125 万人の中から, 27 万人を無作為に抽出し，e-mail でオンライン調査への協力を求めた.

調査では，過去 4 週間に生理，妊娠，または風邪に関係しない腰痛があったかをたずねた．そのために日常生活や社会活動に支障を来した腰痛が 3 か月以上続いたものを，慢性腰痛ありと定義した.

身体化は日本語版 SSS-8 を用いて評価した. SSS-8 は自己記入式の質問票で，身体症状の有無と重症度を評価するものである．SSS-8 は DSM-5 のフィールドトライアルで，身体表現性障害（somatic symptom disorder） の診断を容易にするために用いられ，その後 SSS-8 は世界各地で使用されている．ドイツ語版の SSS-8 はドイツの一般国民において，信頼性と妥当性が良好であったと報告されている．我々は英語版の SSS-8 を日本語に訳し，言語的，心理測定的妥当性を確認した. SSS-8 の合計点は，Gierk らの文献と同様に，無‐僅か (0-3)，低 (4-7)，中 (8-11)，高 (12-15) ，非常に高 (16-32)に分けた.

抑うつ症状は Patient Health Questionnaire-9（PHQ-9）から抽出した 2 問からなる PHQ-2 を用いて評価した．これは，過去 2 週間に抑うつや失感情症を体験したかを問うものである．PHQ-2 の原著では各質問は 0－3 のスケールで評価されるが，我々は各質問を有り／無しの 2 段階で評価する，国立精神・神経医療研究センターによるものを用いた．よって PHQ-2 の合計点は 0，1，または 2 である.

健康関連 QOL は，全般的な健康状態を測定する，EuroQol 5 Dimension (EQ-5D) を用いて評価した．これは移動，セルフケア，日常活動，痛み／不快, 不安／抑うつを問う, 5 つの質問からなる．回答は全般的な健康状態を表す，-0.11 から 1.00 までの 1 つのスコアに変換される．1 は完全に健康な状態で，0 は死である．日本語版 EQ-5D は EuroQol グループに承認されており，広く研究に用いられている.

調査では，年齢，性別，身長，体重，婚姻，学歴，雇用状況，喫煙の有無についてたずねた．身長, 体重より BMI を計算した. また過去 1 年間に, 30 分以上の運動をした頻度（週 2 回以上，週 1 回程度，月 1-2 回程度，していなかった）をたずねた．週 2 回以上と答えたものを，定期的な運動ありと定義した．また，27 の疾患について，通院しているかをたずねた（心臓の病気，高血圧，高脂血症，肺の病気，糖尿病，胃腸の病気，腎臓の病気，肝臓の病気，貧血などの血液の病気，甲状腺の病気，ガン，うつなどのメンタルの病気，婦人科系の病気，泌尿器科系の病気，皮膚の病気，睡眠時無呼吸症候群，耳鼻科の病気，眼科の病気，虫歯や口腔内の病気, 変形性関節症, 腰痛, 頭痛, 関節リウマチ，線維筋痛症，骨粗しょう症，肥満症，その他）．腰痛とうつなどのメンタルヘルスの病気以外の 25 の疾患のうちで，通院ありと答えた疾患の個数を求めた.

SSS-8（5 カテゴリ）と EQ-5D スコアの関係は，線形回帰モデルで検討した．抑うつ（PHQ-2）も同じモデルに入れて解析した．SSS-8 と PHQ-2 の

交互作用は統計的に優位でなかったので，最終モデルには含めなかった．最終的な多変量モデルでは，年齢，性別，BMI，喫煙，婚姻，学歴（短大以上か），定期的な運動の有無，雇用状況（正規雇用かそれ以外か），通院していた疾患の個数（0－25）を調整した．これらの変数は，この研究のデータで統計学的に有意であるかではなく，過去の文献をもとにあらかじめ決定した．これらの変数の VIF から，あきらかな多重共線性の問題はみとめられなかった．解析はすべて SAS9.4 を用いて行い，両側検定で $p<0.05$ を統計学的に有意とみなした．

C. 研究結果

H26-27 年度

調査票は 95 施設，1,704 名より回答を得て，全てを解析対象とした．平均年齢は 40.2 歳（SD 11.7），性別は女性が 75.3%であった．

腰痛の程度は grade 0 が 28.0 %，grade 1 が 55.6%，grade 2 が 13.9 %，grade 3 が 2.5%であった．grade2 以上で 3 ヵ月以上継続した重症度の高い腰痛は 205 名（12.0%）に認められた．

重症度の高い腰痛に関する粗オッズ比を求めた．p 値が 0.05 以下の関連が疑われる要因は，年代，睡眠時間，睡眠の質，職場満足度，心理的な仕事の負担（量），心理的な仕事の負担（質），自覚的な身体的負担度，職場での対人関係のストレス，職場環境によるストレス，仕事の適性度，働きがい，上司からのサポート，活気，イライラ感，疲労感，不安感，抑うつ，身体愁訴，STarT Back スクリーニングツール，TSK-J であった．

変数間の相関係数を算出するとともに，臨床的な見地から，年代，睡眠時間，睡眠の質，職場満足度，心理的な仕事の負担（量），職場での対人関係のストレス，上司からのサポート，身体愁訴，STarT Back スクリーニングツール，TSK-J を独立な変数とした．

多変量ロジスティック回帰の結果，身体愁訴，STarT Back スクリーニングツール，TSK-J が選択され，本研究における重症度の高い腰痛の関連要因とした（下図）．

要因		Odds ratio (95%CI)
身体愁訴	Low	1.00
	High	2.00 (1.43- 2.77)
STarT	Low risk	1.00
	High risk	4.06 (2.67- 6.18)
TSK-J	Low	1.00
	Medium	2.69 (1.37- 5.30)
	High	3.18 (1.65- 6.16)

H28 年度

参加者の特徴を表 1 に示す．参加者は平均 44.5±11.2 歳で，48%が女性であった．PHQ-2 は 1576 人(51%) が 0, 632 人(20%)が 1, 892 人 (29%) が 2 であった．EQ-5D は平均 0.78±0.18 であり，PHQ-2 の点数が高くなるほど低かった． SSS-8 の平均スコアは 9.67±6.68 で，PHQ-2 の点数が高くなるほど，SSS-8 が非常に高い（≥16）の割合が高かった．

多変量解析の結果を表 2 に示す．PHQ-2 の点数は EQ-5D のスコアと有意に関連していた．SSS-8 のどのカテゴリに属するかは，PHQ-2 の点数や他の共変数を調整しても，EQ-5D のスコアと有意に関連していた．すなわち，身体化傾向の高いグループほど，EQ-5D のスコアが低かった．

D. 考察

H26-27 年度

職業性ストレス調査票における身体愁訴が高いオッズ比を示した．身体愁訴の質問には“めまい”，“体のふしぶしが痛む”，“頭が重かったり頭痛がする”，“首筋や肩がこる”，“腰が痛い”，“目が疲れる”，“動悸や息切れがする”，“胃腸の具合

が悪い”,“食欲がない”,“下痢や便秘をする”,“よく眠れない”が該当する．これらはストレスに伴う自律神経失調様の機能的な症状であり，専門科で器質的な原因が明らかにされない臓器系の症状に加え，運動器系の症状も含まれている．これらは心理的ストレスが脳機能に影響を与えることによって起こってくる症状であり，腰痛にも心理的ストレスによる脳機能の不具合（dysfunction）を介し，筋緊張や局所の動脈でのスパズムが強まって起こるタイプがあるという認識を持っている．心理・社会的要因の強い腰痛では，さまざまな身体化徴候をあわせもつ場合が想定されるため，診療では注意深く問診することが必要であると考えている．

H28 年度

ポピュレーション研究のレビューにおいては，身体症状の総合スコアは，抑うつ，不安症，一般的な疾病を調整しても，医療機関の利用と関連しており，健康状態の予測因子であったと報告している．我々の研究では，日本人の慢性腰痛のある人に限っても，身体化傾向が，抑うつや並存疾患を考慮したうえで，健康関連 QOL と有意に関係していることが示された．先行研究から，身体化が腰痛のアウトカムに影響している可能性が示唆されている．松平らは，軽度の腰痛をもつ日本の都市部の勤労者において，職業性ストレス簡易調査表により評価した身体化傾向が，持続性腰痛の発症を予測したことを報告した．また，海外の研究では，カイロプラクティックで治療された腰痛患者で，身体化が痛みの強さ，身体機能，自覚的回復と関連していた．病院で治療された腰痛患者でも，手術あるいは保存的治療を受けたかによらず，ベースラインの身体化は，1 年後の SF-36 と相関し，痛みが 50%以上軽減したかどうかと関連していた．しかしこれらの先行研究では，抑うつは必ずしも調整されていない． 抑うつは腰痛の発症と慢性化の危険因子であり，身体化はしばしば抑うつに合併するので，身体化が抑うつと独立して，腰痛のアウトカムと関連しているかを明らかにすることは容易ではないかもしれない． この研究では，標本数が多く，重要な共変数も調整している．参加者は医療機関の患者ではないため，治療を求めたという特異性によるバイアスの可能性も低い．しかし，抑うつの評価は 2 つの質問によるもので，誤分類があり得る．また本研究では不安症は評価していない．そのため，抑うつや不安症による交絡が残存している可能性はある．また，この研究の参加者はインターネットで募集されたため，結果は日本人全体に一般化することは出来ないかもしれない．

E. 結論

慢性腰痛のある人で，身体化傾向の強い人ほど，健康関連 QOL が低く慢性腰痛リスクとして身体化を視野に入れる必要がある．

F．健康危険情報

特記すべき事項なし。

G．研究発表

現時点ではなし。

H．知的財産権の出願・登録状況

現時点ではなし。

表 1 慢性腰痛のある参加者の特徴

	全員 (n=3100)	**PHQ-2=0 (n=1576)**	**PHQ-2=1 (n=632)**	**PHQ-2=2 (n=892)**	***p 値**
平均年齢(SD)	44.5 (11.2)	45.8 (11.0)	44.5 (11.5)	42.1 (11.1)	<0.001
女性 (%)	1483 (47.8)	743 (47.1)	311 (49.2)	429 (48.1)	0.669
BMI (%)					0.018
<25	2333 (75.3)	1184 (75.1)	484 (76.6)	665 (74.6)	
25 - 29	589 (19.0)	320 (20.3)	103 (16.3)	166 (18.6)	
≥30	178 (5.7)	72 (4.6)	45 (7.1)	61 (6.8)	
現在の喫煙 (%)					<0.001
あり	1064 (34.3)	489 (31.0)	233 (36.9)	342 (38.3)	
なし	2036 (65.7)	1087 (69.0)	399 (63.1)	550 (61.7)	
現在の婚姻 (%)					<0.001
なし	1327 (42.8)	544 (34.5)	299 (47.3)	484 (54.3)	
あし	1773 (57.2)	1032 (65.5)	333 (52.7)	408 (45.7)	
学歴 (%)					0.043
短大未満	1650 (53.2)	804 (51.0)	350 (55.4)	496 (55.6)	
短大以上	1450 (46.8)	772 (49.0)	282 (44.6)	396 (44.4)	
定期的な運動 (%)					0.001
なし	2487 (80.2)	1222 (77.5)	524 (82.9)	741 (83.1)	
あり	613 (19.8)	354 (22.5)	108 (17.1)	151 (16.9)	
雇用形態 (%)					<0.001
正規職員	1271 (41.0)	700 (44.4)	221 (35)	350 (39.2)	
それ以外	1829 (59.0)	876 (55.6)	411 (65)	542 (60.8)	
通院疾患の平均個数 (SD)	1.2 (2.0)	1.0 (1.8)	1.3 (2.0)	1.4 (2.3)	<0.001
EQ-5D 平均 (SD)	0.78 (0.18)	0.84 (0.16)	0.75 (0.17)	0.70 (0.17)	<0.001
SSS-8 (%)					<0.001
無 - 僅か	590 (19.0)	445 (28.2)	77 (12.2)	68 (7.6)	
低	785 (25.3)	488 (31.0)	152 (24.1)	145 (16.3)	
中	616 (19.9)	297 (18.9)	156 (24.7)	163 (18.3)	
高	505 (16.3)	186 (11.8)	122 (19.3)	197 (22.1)	
非常に高	604 (19.5)	160 (10.2)	125 (19.8)	319 (35.8)	

表 2. EQ-5D と SSS-8 の関係（多変量解析）

	回帰係数	標準誤差	p 値	トレンド p
切片	0.815	0.014	<0.001	
SSS-8				<0.001
無 - 僅か	0.218	0.009	<0.001	
低	0.142	0.008	<0.001	
中	0.098	0.009	<0.001	
高	0.040	0.009	<0.001	
非常に高	レファレンス			
PHQ-2				
0	レファレンス			
1	-0.042	0.007	<0.001	
2	-0.066	0.007	<0.001	

労災疾病臨床研究事業費補助金

総合研究分担報告書

サブテーマ②予防に有用な福祉機器等の開発

エビデンスに基づいた腰痛治療・予防を目指した「不良姿勢チェッカー」の開発

研究分担者　勝平　純司　　新潟医療福祉大学医療技術学部義肢装具自立支援学科

研究要旨

腰ベルトやコルセットに代表される体幹装具使用による腰痛の治療と予防は産業衛生分野においても行われている．しかしながら，体幹装具使用による腰痛予防や治療のエビデンスは乏しく，最近では長期間装着すると姿勢安定に寄与する体幹深部筋の弱化を招いてしまうという報告もある．そこで我々は，腰部負担計測によるエビデンスに基づいた新しい姿勢を改善し，腰痛治療と予防を目指した「不良姿勢チェッカー」の開発に着手した．

初年度（H26）と２年目（H27）については既存の機器の評価，腰部負担の大きい重量物挙上における持ち上げ姿勢と，負担は小さくとも蓄積されると腰痛の発症につながると考え，立位姿勢も対象として良姿勢が腰部負担に与える影響を三次元動作分析装置を用いて明らかにした．**H28 年度**には，適度な骨盤前傾と体幹伸展の姿勢では腰部負担が小さくなることがわかったため，この姿勢をとることをフィードバックする「不良姿勢チェッカー」を作製した．また２次試作では骨盤と体幹だけでなく，頸部の姿勢の計測とフィードバックも可能とした．その後新たな機能として脊柱起立筋を指標として作業現場における姿勢のみならず腰部負担を計測し，かつ負担量が大きい作業員の位置を特定する仕組みまで構築することができた．

A．研究目的

産業衛生分野においても腰痛の予防や軽減を目的として腰ベルトやコルセットに代表される体幹装具が使用されている．コルセットや他の現存の体幹装具は腹部を圧迫する装具や３点で固定する装具に大別される．体幹装具は我が国でも腰痛の予防や治療を目的として，数多く使用されているが，装着による効果のエビデンスは十分でない．また，これらの装具を装着すると体幹の周囲筋が弱化するといわれてきたが，明確な証拠は得られていなかった．しかしながら，2014 年に Rostami らによって行われた超音波画像診断装置を用いた計測により，コルセットを４週間使用すると体幹深部筋の中でも脊柱の安定性に重要な役割を果たすといわれている腹横筋と多裂筋の厚さや断面積が減少することが明らかになった．予防や治療のため体幹装具を長期間使用すると外して生活することが難しくなるケースが多くみられるが，体幹深部筋の弱化がその原因となっていることが考えられる．

そこで我々は，コルセットや腰ベルトに代わる姿勢改善を行う新たな機器「不良姿勢チェッカー」の開発に着手した．本研究の目的は３か年の中で初年度については体幹深部筋を弱化させない機構を有する体幹装具の評価，２年目については腰部負担計測をとして「不良姿勢チェッカー」がターゲットとすべきパラメータや使用場面の抽出，３年目については２年間の成果から得たデータを基に「不良

姿勢チェッカー」を完成させることを目的に３か年の研究を実施した.

Ｂ．研究方法

抗力を具備した継手付き体幹装具の評価

勝平らは，体幹支持体と骨盤支持体を抗力を具備した継手により連結した体幹装具 Trunk Solution(以下 TS)を新たに開発し，2014 年度の Good Design 賞を受賞した．TS は装着した際に継手の抗力によって体幹を伸展方向に押す力を与えて姿勢を整えることで体幹深部筋の弱化を防ぐことを目的としており，従来の体幹装具とは異なる機構と効果を有している．初年度は TS の評価を行い，その効果を検証することで「不良姿勢チェッカー」に付加すべき機能の抽出を行うこととした.

1)TS の有無による歩行時の脊柱起立筋の計測

対象は健常成人男性 24 名(年齢 22.0±1.6 歳，身長 171.4±4.㎝，体重 62.6±7.3kg)とした．なお，対象は整形外科疾患や中枢神経疾患を有さない者とした．被検者は，TS 装着群(以下：TS 群)(年齢 22.4±1.9 歳，身長 171.5±6.0cm，体重 63.5±9.2kg)とダーメンコルセット装着群(以下：ダーメン群)(年齢 21.5±1.5 歳，身長 171.3±5.2cm，体重 61.8±6.7kg)の 2 群にランダムに選別した.

歩行解析機器として，Gait Judge System（パシフィックサプライ社製)を用いた.Gait Judge System とは，wifi コンバータと wifi を受信する専用アプリをダウンロードした iPad からなり,iPad にリアルタイムに動画や筋活動等が記録される．筋活動の計測には，サンプリング周波数 1000Hz の表面筋電計(パシフィックサプライ社製)を使用した.記録用電極はメッツ社製 Ambu Blue Sensor P を使用し，先行研究に従い，被験者の皮膚処理を施した後に双極性表面電極 2 個を電極中心間隔 3 ㎝で貼付した．筋活動の測定筋は，右側の脊柱起立筋とした.

計測課題は歩行とした．被験者は計測室内に設けられた約 10m の直線歩行路を自由速度にて歩行を行った．被験者をランダムに TS 群，ダーメン群に分け，歩行は同一被験者に対して①装具装着前，②装具装着時，③外した直後の計 3 条件で計測を行った．計測プロトコールとして，装具装着前に歩行計測後，装具を装着し 5 分間安静後に計測を実施，TS を外して 5 分間安静後に計測を実施した．歩行計測実施後，各筋の最大随意収縮(Maximal Voluntary Contraction：MVC)の筋活動を測定し，歩行時の筋活動を正規化するために用いた.

2) TS 装着時における側腹筋の計測

対象は体幹，下肢に神経学的・整形外科的な既往のない健常男性 27 名（平均年齢 22±2.3 歳，平均身長 170.0±0.5 ㎝，平均体重 61.5±6.5 ㎏ BMI21.0±1.8）とした．なお全対象者は事前に本研究の目的と方法を説明した.

側腹筋の測定には超音波画像診断装置 Sonosite（FUJIFILM 社製）を用いた．再現性を得るためにプローブは右側の前腋窩線上にて肋骨辺縁と腸骨稜の中央部にあてるようにして腹横筋を測定した．計測は，非装着時,TS 装着時，ダーメンコルセット装着時の計 3 回をランダムに測定した.

「不良姿勢チェッカー」の開発のために，上記の実験研究を通して腰部負担軽減に効果的な介入方法のエビデンスを得た.

持ち上げ動作における不良姿勢改善の効果

腰痛は労働時に生じる疾患で最も多いとされており，作業内容では重量物を持ち上げる作業が腰痛を発症する危険性が最も高いといわれている.「不良姿勢チェッカー」が効果を

発揮する場面として，持ち上げ動作に代表される重量物運搬を伴う作業が想定された．そこで我々は重量物挙上を行う際に不良姿勢を改善する指導を与えることが，どの程度腰部負担を軽減するのかについて明らかにした．

対象は健常成人男性 10 名（年齢 20.9±0.5 歳　身長 174.9±4.3 ㎝　体重 64.1±4.8 ㎏）とした．計測条件については以下の 4 条件を設定した．①squat 条件（股関節と膝関節を屈曲）②stoop 条件（股関節を屈曲，膝関節を伸展），　③squat 骨盤前傾条件（squat 条件よりも骨盤を前傾させるように指示），④stoop 骨盤前傾条件(stoop 条件よりも骨盤を前傾させるように指示)．

重量物は 11.3 ㎏に設定し,被験者の足先から足長の 1/2 の距離に設置するように統一した．

測定機器は三次元動作解析装置 VICON MX（VICON 社製），床反力計（AMTI 社製）4 枚，赤外線カメラ(周波数 100Hz)10 台を用いた．被験者には 45 個の赤外線反射マーカーを貼付し,動作中の椎間板圧縮力,椎間板剪断力，骨盤前傾角度，腰部関節中心と重量物の重心との距離を算出した．椎間板圧縮力と椎間板剪断力は体重で除して正規化した値で比較・検討を行った．

統計処理は挙上方法の違いと骨盤前傾指示の有無を要因とした繰り返しのある二元配置分散分析反復測定法を用いた．また，通常の条件と骨盤前傾を指示した条件での差を判定するために，対応のある t 検定を用いた．さらに，4 条件の中でどの姿勢が最も腰部負担が小さいのかを明らかにするため，椎間板圧縮力と剪断力については一元配置分散分析反復測定法と多重比較検定(Bonferroni 法)を併せて行った．有意水準は 5%とした．

立位姿勢における不良姿勢改善の効果

職場環境において，長時間の立位姿勢は腰痛を発症するリスクファクターになることや，不良姿勢が腰痛に関連することも指摘されている．いわゆる不良姿勢を想定すると，腰部関節中心と上半身重心位置との距離が離れることで，椎間板への力学的負担が増大することが予想できる．本研究では，立位における不良姿勢の改善が腰部負担に与える影響を明らかにした．

若年男性 20 名（23.9±3.3 歳，172.2±6.8 cm，62.9±8.9 ㎏）を対象者とした．計測を行うに先立ち，被験者には研究内容を十分に説明し，書面にて研究への同意を得た．

計測条件は被験者が習慣的にとっている安楽立位姿勢とした．被験者は片脚ずつ床反力計上に肩幅程度に足を広げて立ち 10 秒間 3 試行の立位姿勢を計測した.目線は 5m 先に設置した目線の高さの目印を見るように指示した.安楽立位姿勢と直立姿勢の比較を行った．本研究では直立姿勢に特化した姿勢介入を行い，大規模な介入研究も行われているアレクサンダーテクニック(以下 AT)を用いることとした.介入には AT 教師養成校の認定を受け，日本アレクサンダーテクニーク協会（Japan Alexander Technique Society）に所属する AT 教師（資格取得後 5 年）が実施した．姿勢介入は声による教示と徒手による姿勢の修正によって行われた．介入時間は 5 分以内とした．介入を行った後は，被験者には介入内容を意識して立位姿勢を保持するよう指示し，通常の立位姿勢の計測と同様に10秒間3試行の立位姿勢を計測した．計測順序は，先に介入無し条件の安楽立位姿勢を計測後に直立姿勢をとる介入有り条件の計測を行った．

計測には三次元動作分析装置(VICON MX, VICON,UK)，床反力計(AMTI,USA)を使用した．被験者の立位姿勢を三次元的にとらえるためにサンプリング周波数 100Hz の床反力計 2 枚赤外線カメラを 10 台用いた三次元動作解析

装置を使用した．三次元動作解析装置で得られたパラメータは，10 秒間の値を平均し，3 試行の値を平均した値を分析に用いた．スパイナルマウスから得られた値は 3 試行の平均値を用いた．関節モーメント，椎間板圧縮力は被験者の体重で除し正規化した．

安楽立位姿勢と直立姿勢における姿勢と腰部負担パラメータの比較を，対応のある t 検定を用いて比較した．有意水準は 5%とした．

前述したすべての実験研究は，国際医療福祉大学の倫理委員会の承認を得て実施した．また，対象者全員に対して，本研究の目的と内容を十分に説明し，書面による同意を得た後に計測を行った．

不良姿勢チェッカーの設計方針の決定

上記の実験研究では腰部の負担を軽減する方法として「良姿勢」に着目し，3 次元動作解析，筋電計，超音波画像診断装置を指標とした計測を行い「良姿勢」をとらせることの効果を客観的に明らかにした．

腰部への負担を増加させる不良姿勢を，3 次元動作解析装置がない臨床現場や作業現場で使用でき，尚且つそれらの状況をフィードバックできる機器を開発することを考えた．当初は TS 自体を改良することで，機器自体に腰部負担軽減効果を持たせることを考えていたが，TS には姿勢修正効果や腰部負担軽減効果はあるものの，試用による調査において作業場面で常用するのは難しいというコメントが得られたことから，より簡便で受け入れられやすいフィードバック機器を開発するという方針となった．また，腰痛対策として治療と予防という観点が必要になることから，「不良姿勢チェッカー」に治療モデルと予防モデルという軸を持たせることにした．

C．研究結果

抗力を具備した継手付き体幹装具の評価

1) 歩行時の脊柱起立筋の計測

脊柱起立筋に関しては，1 歩行周期において TS 装着前と比較して TS 装着時に有意に減少した．また，TS 装着前と比較して TS を外した直後でも有意に減少した．遊脚期において TS 装着前と比較して TS 装着時に有意に減少した．ダーメン群では，1 歩行周期・各歩行周期において有意差は認められなかった．

2) 側腹筋の計測

腹横筋では，装具なしに比べ TS 装着時に有意な筋厚の増加が認められた．さらにダーメンコルセット装着時より TS 装着時にも有意な筋厚の増加が認められた．しかし，装具なしとダーメンコルセット装着時の間には有意差は認められなかった．内腹斜筋では，安静時の装具なし，ダーメンコルセット装着時，TS 装着時には有意差は認められなかった．外腹斜筋では，装具なしに比べ TS 装着時に有意な筋厚の増加が認められた．しかし，装具なしとダーメンコルセット装着時の間，ダーメンコルセット装着時と TS 装着時の間には有意差は認められなかった．

持ち上げ動作における不良姿勢改善の効果

二元配置分散分析反復測定法の結果，挙上方法の違いと骨盤前傾指示の有無による交互作用がみられた．対応のある t 検定の結果，椎間板圧縮力は squat 条件よりも squat 骨盤前傾条件で有意に小さくなった．一方，stoop 条件では骨盤前傾指示の有無で有意な差はみられなかった．また，一元配置分散分析と多重比較検定の結果，squat 条件は他の 3 つの条件よりも有意に椎間板圧縮力が大きく，squat 条件以外の 3 つの条件の間に有意差はみられなかった．

骨盤の前傾角度を示す．二元配置分散分析反復測定法の結果，挙上方法の違いと骨盤前傾指示の有無による交互作用がみられた．また，対応のある t 検定の結果，squat 条件よ

りも squat 骨盤前傾条件において骨盤の前傾角度は有意に大きく，stoop 条件と stoop 骨盤前傾条件では有意差はみられなかった.

立位姿勢における不良姿勢改善の効果

椎間板圧縮力は直立姿勢において有意に低い値を示した. 腰部モーメントに関しては, 3 軸まわりの腰部モーメントの内，腰部屈伸モーメントは低下傾向であったが（p=0.0889)，有意差は認められなかった. 腰部側屈モーメントが有意に低い値を示した.

関節角度に関しては，頭部屈曲角度，体幹伸展角度,骨盤前傾角度に有意な差を認めた. 脊柱彎曲角度に関しては，胸椎彎曲角度，腰椎彎曲角度に有意差を認めた. 頭部角度は屈曲方向に，体幹角度は伸展方向に増加した. 胸椎後彎角度は減少し，腰椎前彎角度は増加していた. 骨盤前傾角度は増加していた.

不良姿勢チェッカーの設計

上記の実験結果をまとめて不良姿勢チェッカーの設計を行った.

１）治療モデル不良姿勢チェッカー

体幹上部の肩部と腰部に電気的に角度を計測する「治療モデル不良姿勢チェッカー」を試作した.「治療モデル不良姿勢チェッカー」は，スマートフォンと連動しており電子的に送信される角度情報からスマートフォン内の疑似モデルが対象者と同様の姿勢を表現することで，自身の姿勢の状況の理解が可能となる仕組みである.

２）予防モデル不良姿勢チェッカー

運搬等の動作分析から負担部位を特定し，それらの部位を主動している脊柱起立筋の筋活動を介護や労働業務中に監視する機能を有する,「予防モデル不良姿勢チェッカー」を作製した. 生体信号計測ユニットは，左右の脊柱起立筋の活動を 1000Hz でモニタリングし，筋活動データに対して RMS(二乗平均平方根)処理を行い定量化する. １sec 当たりの定量化された筋活動をリスク指標として用い，位置情報と同時に記録，表示することで位置と腰痛リスクを時系列で把握するシステムを開発した.

D．考察

実験研究を通して得られたエビデンスを基に２種類の「不良姿勢チェッカー」を開発した.

初年度に行った TS の評価を通して,良姿勢を学習させることが出来れば，脊柱起立筋の活動を抑えつつ，体幹の深部筋の弱化を防ぐことができると考えられた. また TS を外した後も脊柱起立筋の活動が有意に低下していたことから，良姿勢の学習が進めば，必ずしも姿勢を修正する効果を持つ機器を直接身に着ける必要がないと考えた.

2 年目に実施した持ち上げと立位姿勢における姿勢と腰部負担の関係を調べた研究では,骨盤前傾，体幹伸展の良姿勢をとらせると筋活動のみならず，椎間板を押しつぶす力である圧縮力が減少することが明らかになった.

上記の実験結果から，良い姿勢を使用者にフィードバックすることができれば，脊柱起立筋の筋活動を低減させつつ，椎間板圧縮力を軽減することができること，持ち上げなどの重労働にかかわるものだけでなく，オフィスワーカーのように小さな負担が蓄積されて発症する腰痛に関しても効果があると考えられた.

２年間の研究を通して明らかにした腰部負担増大に関わる姿勢因子の中で特に，体幹部と腰部の位置関係が腰痛に関連することがわかった. そこでこれらの因子をフィードバックに用いるため，体幹部と腰部の関係性を視覚化できる「治療モデル不良姿勢チェッカー」を作製した. さらに体幹部と腰部との関係性に，体幹部と頭部との関係を付加考慮するこ

とで，頸部に起因するストレートネック等の関連する症状も把握できることがわかった．これらの機能を組み込んだ第 2 次「治療モデル不良姿勢チェッカー」を試作評価した結果，不良姿勢改善の再現性において良好な結果を得ることができた．第 2 次「治療モデル不良姿勢チェッカー」を介護現場での腰痛の前兆となる腰痛リスクの可視化を検討したが，介護現場においては重量物の運搬等，姿勢のみでは判断できない介護者への負担という新しいリスクを考慮する必要があることがわかった．

運搬等の動作分析から負担部位を特定し，それらの部位を主動している脊柱起立筋の筋活動を介護業務中に監視する機能を有する，「予防モデル不良姿勢チェッカー」を作製した．「予防モデル不良姿勢チェッカー」では脊柱起立筋を指標として作業現場で姿勢のみならず腰部負担を計測し，かつ負担量が大きい作業員の位置を特定する仕組みまで構築した．

当初目的としていたエビデンスに基づく体幹装具に代わる新たな姿勢を修正する機器の開発だけでなく，これを発展させた機器の開発まで 3 か年内に達成することができ，本プロジェクトは順調に推移した．本研究の限界としては，作製した 2 つの「不良姿勢チェッカー」をトライアルレベルではなく，ある程度の規模と期間をもって評価を行うことが出来なかった点にある．これについては今後の課題とする．

E．結論

体幹装具に代わる姿勢を修正する機器の開発を目指し，良姿勢に関する客観的な分析によるエビデンスを理解したうえで，日常でも良姿勢を再現することができる「治療モデル不良姿勢チェッカー」を設計した．治療の場面で必要な，不良姿勢の把握から良姿勢の再現までの治療モデルに必要な一連の機器の開発を行うことができた．また腰痛リスクを可視化する脊柱起立筋と位置情報を指標とした「予防モデル不良姿勢チェッカー」も設計した．今後はこれら二つの不良姿勢チェッカーの腰痛発生抑制効果について評価を進めていく予定である．

F．健康危険情報

該当なし．

G．研究発表

論文発表

1. 勝平純司，福祉支援工学の基礎（体幹装具 Trunk Solution）．Biophilia 3: 26-30, 2014
2. Hasegawa T, Katsuhira J, Matsudaira K, Iwakiri K, Maruyama H. Biomechanical Analysis of Low Back Load when Sneezing. Gait Posture 40: 670-675,2014
3. 伊藤晃洋, 勝平純司, 飯島進乃, 地域在住高齢者の静止立位における足圧中心位置に影響を与える変数の検討-装具を用いた体幹への介入効果を含めて -, 臨床バイオメカニクス　35：357-361，2014
4. Hayashi　S, Katsuhira J, Matsudaira K, Maruyama H. Effect of pelvic forward tilt on low back compressive and shear forces during a manual lifting task. Physical Therapy Science 28:802-806, 2016
5. 勝平純司，動作分析の活用-住環境整備，移乗介助への応用-. 臨床歩行分析研究会誌 2： 17-22，2015
6. Katsuhira J, Matsudaira K, Yasui T, Iijima S, Ito A. Efficacy of a trunk orthosis with joints providing resistive force on low-back load in elderly persons

during static standing. Clinical intervention in aging (10):1413-1420, 2015

7. 小川幸宏，勝平純司，金子純一朗，前田和也，石坂正大，抗力を具備した継手付き体幹装具の装着が腹横筋に及ぼす影響，日本義肢装具学会誌１：41-44，2016
8. 屋嘉比 章紘，勝平 純司，新井 健介，曽部 健太，西川 智洋，藤原 光伸，遠藤 裕伽，齊藤 愛弥，下重 絢香，米川 茉希, トレッドミル歩行時の酸素摂取量に対する継手付き体幹装具とダイエットベルトの効果，理学療法科学 31(3)，455-459，2016
9. 伊藤 将円，勝平 純司，野村 高弘 高齢者における歩行補助具使用時の歩行分析：前額面における関節モーメントの比較, 臨床バイオメカニクス 37，359-363，2016
10. 伊藤 晃洋，勝平 純司，飯島 進乃，地域在住高齢者の静止立位における前額面上の足圧中心位置に影響を与える変数の検討：装具を用いた体幹への介入効果を含めて，臨床バイオメカニクス 37，373-377，2016
11. Oka H, Matsudaira K, Kikuchi N, Haga Y, Sawada T, Katsuhira J, Yoshimoto T, Kawamata K, Tonosu J, Sumitani M, Kasahara S, Tanaka S: Estimated risk for chronic pain determined using the generic STarT Back 5-item screening tool. J Pain Res (in press)
12. Yamada K, Matsuadira K, Tanaka E, Oka H, Katsuhira J, Iso H. Sex-specific impact of early-life adversity on chronic pain: A large population-based study in Japan. J Pain Res (in press)
13. Katsuhira J, Matsudaira K, Oka H, Iijima S, Itou A, Yasui T, Yozu A. Efficacy of a trunk orthosis with joints providing resistive force on low back load during level walking in elderly persons. Clin Interv Aging 11: 1589-1597, 2016
14. Tonosu J, Inanami H, Oka H, Katsuhira J, Takano Y, Koga H, Yuzawa Y, Shiboi R, Oshima Y, Baba S, Tanaka S, Matsudaira K. Diagnosing Discogenic Low Back Pain Associated with Degenerative Disc Disease Using a Medical Interview. PLoS One 11: e0166031, 2016

H. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし

労災疾病臨床研究事業費補助金
総合研究分担報告書

サブテーマ②予防に有用な福祉機器等の開発

予防アルゴリズム構築に向けた画像と腰痛の関連因子探索

研究分担者　唐司寿一　関東労災病院　整形外科

研究要旨
腰椎の画像所見と腰痛とが必ずしも一致しない症例が、臨床的には散見される。なかでも腰椎 MRI は空間分解能も高く、優れたモダリティであるが労働者における疫学的な検討は不十分である。上述したテーマに取り組むべく研究分担者は、「予防アルゴリズム構築に向けた画像と腰痛の関連因子探索」を行うため、H27 年度より本研究班に参画した。初年度（H27 年）には、関東労災病院に勤務する職員にて画像データベースを構築した。H28 年度には、このデータベースを用いて腰椎 MRI 所見と過去の高度な腰痛の既往との関連を分析した。年齢・性を調整して解析した結果、Pfirrmann 分類≧3、椎間板膨隆あり、High intensity zone(HIZ)あり、が過去の高度な腰痛の既往と関連していた。 これらの知見は、腰痛スクリーニングに活用可能であり、予防アルゴリズムの一角を担うものと考えている

A. 研究目的

国民生活基礎調査では、腰痛は有訴率、通院率とも常に上位にある。我々が行った調査では、一生のうちに腰痛に罹患する割合は 83%、直近 4 週間での腰痛の罹患率は 36%である[1]。

腰椎の Magnetic Resonance Imaging(MRI)は腰痛の病態を評価するのに役立つが、椎間板の変性所見と撮影時に存在する腰痛との関連についてはまだ議論の一致がなく、椎間板変性所見が現在の腰痛と関連するという報告[2]と関連しないという報告[3]がある。

慢性腰痛はさまざまな経過をたどることが知られており、持続的な腰痛を呈する例の他に、寛解と再燃を繰り返す間欠的な腰痛を呈する例も存在する[4]。我々は、もし高度な腰痛が再燃することを推測させる MRI 所見を知ることができれば、そのような患者に選択的に腰痛予防の指導介入が可能になることを期待した。そこで、「MRI で椎間板変性所見があり、かつ撮影の時点で腰痛がないならば、椎間板変性の所見は過去の腰痛の既往を示し、高度な腰痛が再燃する可能性を示唆する」という仮説を立てた。本研究の目的は、撮影時に腰痛のない症例の腰椎 MRI 所見と過去の腰痛の既往との関連を調べることである。

B．研究方法

研究分担者は、「予防アルゴリズム構築に向けた画像と腰痛の関連因子探索」を行うため、H27 年度より本研究班に参画した。初年度（H27 年）には、関東労災病院に勤務する職員にて画像データベースを構築した。H28 年度には関東労災病院に勤務する職員で、MRI 撮影時に「腰痛がない」と申告された 91 例にて解析を行った。「現在腰痛がない」ことは「１ヵ月以内に肋骨下縁から殿裂までの間の痛みがないもの」と定義した[5]。「過去に腰痛があった」ことは、ある程度高度な腰痛の既往があったことに限定するために、「医療機関へ通院するほどの腰痛があったもの」と定義した。自記式質問票を用いて年齢、性別、身長、体重を調査した。

MRI 所見の読影は、T12/L1 から L5/S1 の 6 椎間についてそれぞれ椎間板変性、椎間板膨隆、High intensity zone(HIZ)、すべりの有無を評価した。

各所見について、少なくとも 1 椎間でみられるものを所見ありとした。

椎間板変性は Pfirrmann 分類（5 段階：1-5）[6]で 3、4、5 であるものとした。椎間板膨隆は 3mm 未満の椎間板腔の膨隆で矢状面像にて前後ともに同様に膨隆しているものとした[7]。HIZ は椎間板後方部分に、高信号を示す白い点状の所見があるものとした[8]。すべりは 5mm 以上すべっているものとした。

検者内信頼性を評価するために、無作為に選択された 20 例の MRI を 1 ヵ月以上の間隔を空けて 2 回読影した。検者間信頼性を評価するために、同様に無作為に選択された 20 例の MRI を 2 名の脊椎脊髄病指導医が読影した。検者内・検者間信頼性は κ 値を用いて評価した。

MRI 撮影時に腰痛のない参加者 91 例を「過去に腰痛があった群」と「過去にも腰痛がなかった群」に分けて、MRI 所見との関連を評価した。さらに、単変量解析および年齢・性を調整した解析を行い、各 MRI 所見のオッズ比を算出した。

（倫理面への配慮）

関東労災病院医学研究倫理審査の承認を得て推進した。被験者に対してはデータを ID 化して管理するなど個人情報には十分配慮すること、同意後もいつでも同意撤回が可能であること等を説明後、書面での同意を取得した。

C．研究結果

91 名の参加者のうち 27 名には過去の腰痛の既往があり、64 名には過去の腰痛の既往がなかった。参加者全体の年齢は 34.9±10.6 才、女性 47 名・男性 44 名、BMI は 21.8±3.0 kg/m² であった。過去の腰痛がある群の平均年齢は 38.3 才、ない群は 33.5 才と有意差がみられた。性別、BMI は両群間で有意差がなかった(Table 1)。

Table 1 患者背景

Table 1. 患者背景

	全体 (n = 91)	過去の腰痛あり (n = 27)	過去の腰痛なし (n = 64)	P
年齢	34.9±10.6	38.3±10.7	33.5±10.4	0.0486*
女性 (%)	47 (51.6)	12 (44.4)	35 (54.7)	0.3718
BMI (kg/m²)	21.8±3.0	21.8±0.6	21.7±0.4	0.9639

検者内信頼性と検者間信頼性は、Pfirrmann 分類、椎間板膨隆はいずれも”moderate”、HIZ はいずれも”substantial”であり、一致度が高いことが示された[9]。すべりについては、2 名の読影者のうち 1 名の読影所見ですべりがあるとされた例がゼロであったため、κ 値の計算が不能だった(Table 2)。

Table 2　MRI 読影所見の検者内信頼性・検者間信頼性

Table 2.　MRI 読影の検者内信頼性・検者間信頼性

MRI 所見		読影数	Kappa 値	95%信頼区間
Pfirrmann 分類				
	検者内信頼性	20 vs 20	0.66	0.55-0.77
	検者間信頼性	20 vs 20	0.64	0.52-0.76
椎間板膨隆				
	検者内信頼性	20 vs 20	0.60	0.39-0.81
	検者間信頼性	20 vs 20	0.67	0.48-0.87
High intensity zone (HIZ)				
	検者内信頼性	20 vs 20	0.85	0.64-1.06
	検者間信頼性	20 vs 20	0.93	0.79-1.07
すべり				
	検者内信頼性	20 vs 20	NA	NA
	検者間信頼性	20 vs 20	NA	NA

Fisher 正確検定の結果、過去の腰痛がある群は、ない群と比べて、有意に Pfirrmann 分類≧3 (p=0.0026)、椎間板膨隆(p=0.0019)がみられた。HIZ とすべりには有意差がみられなかった(Table 3)。

Table 3 MRI 所見

Table 3. MRI 所見

	全体 (n = 91)	過去の腰痛あり (n = 27)	過去の腰痛なし (n = 64)	P
Pfirrmann 分類≧3	69 (75.8)	26 (96.3)	43 (67.2)	0.0026*
椎間板膨隆（＋）	48 (52.3)	21 (77.8)	27 (42.2)	0.0019*
High intensity zone (HIZ)（＋）	19 (20.9)	9 (33.3)	10 (15.6)	0.0883
すべり（＋）	4 (4.4)	3 (11.1)	1 (1.6)	0.0766

各椎間について着目すると、Pfirrmann 分類≧3 は T12/L1、L3/4、L4/5、L5/S1、椎間板膨隆は L2/3、L3/4、L4/5、L5/S1 で有意差がみられた（Table 4)。HIZ はほとんどすべて L4/5 または L5/S1 でみられた。すべりは L4/5 と L5/S1 のみでみられた。

Table 4 各椎間の Pfirrmann 分類と椎間板膨隆（MRI 所見）

Table 4. 各椎間の Pfirrmann 分類と椎間板膨隆（MRI 所見）

MRI 所見	高位	全体 (n = 91)	過去の腰痛あり (n = 27)	過去の腰痛なし (n = 64)	*P*
Pfirrmann 分類≧3	T12/L1	18 (19.8)	9 (33.3)	9 (14.1)	0.0350*
	L1/2	22 (24.2)	9 (33.3)	13 (20.3)	0.1851
	L2/3	30 (33.0)	10 (37.0)	20 (31.3)	0.5917
	L3/4	44 (48.4)	18 (66.7)	26 (40.6)	0.0232*
	L4/5	56 (61.5)	24 (88.9)	32 (50.0)	0.0005*
	L5/S1	56 (61.5)	23 (85.2)	33 (51.6)	0.0026*
椎間板膨隆（＋）	T12/L1	2 (2.2)	1 (3.7)	1 (1.6)	0.5245
	L1/2	1 (1.1)	1 (3.7)	0 (0.0)	0.1216
	L2/3	2 (2.2)	2 (7.4)	0 (0.0)	0.0277*
	L3/4	5 (5.5)	4 (14.8)	1 (1.6)	0.0113*
	L4/5	35 (38.5)	17 (63.0)	18 (28.1)	0.0018*
	L5/S1	35 (38.5)	16 (59.3)	19 (29.7)	0.0081*

単変量解析の結果、各オッズ比は Pfirrmann 分類≧3　12.7、椎間板膨隆 4.8、HIZ2.7、すべり 7.9 であり、Pfirrmann 分類≧3 (p=0.0009)と椎間板膨隆(p=0.0015)で有意差がみられた。年齢・性調整オッズ比を計算すると、各オッズ比は Pfirrmann 分類≧3　10.5、椎間板膨隆 4.2、HIZ3.1、すべり 6.6 であり、Pfirrmann 分類≧3 (p=0.0065)、椎間板膨隆(p=0.0047)、HIZ(p=0.0405)で有意差がみられた(Table 5)。

Table 5 単変量解析と年齢・性調節解析

Table 5. 単変量解析と年齢・性調節解析

	単変量解析			年齢・性調節解析		
	オッズ比	95%信頼区間	*P* 値	オッズ比	95%信頼区間	*P* 値
Pfirrmann 分類≧3	12.7	2.43-234.18	0.0009*	10.5	1.78-202.09	0.0065*
椎間板膨隆（＋）	4.8	1.79-14.55	0.0015*	4.2	1.54-13.15	0.0047*
High intensity zone (HIZ)（＋）	2.7	0.94-7.78	0.0652	3.1	1.05-9.42	0.0405*
すべり（＋）	7.9	0.96-163.50	0.0551	6.6	0.74-141.71	0.0923

D. 考察

両群間の背景には年齢以外に有意差がなかった。また、検者内信頼性と検者間信頼性は各所見について概ね良好とみなすことができた。

Pfirrmann 分類≧3 は特にオッズ比が 10 以上であり、過去の腰痛の既往と強く関連していた。椎間板変性は L5/S1 と L4/5 で生じやすいという過去の報告[10]と同様、本研究でも特に下位腰椎で Pfirrmann 分類≧3 の所見がみられた。下位腰椎は上位腰椎と比較して可動域が小さいため[11]、椎間板に対する負荷が増大し変性を惹起するものと考えられた。椎間板膨隆も過去の腰痛の既往と関連していた。p 値は下位腰痛ほど低値になるものの、L2/3 以下のすべての椎間板レベルで椎間板膨隆は過去の腰痛の既往との関連がみられた。HIZ は Fisher 正確検定と単変量解析では過去の腰痛の既往との関連はないという結果だったが、年齢・性を調整して解析すると有意に関連があるという結果になった。すべりは過去の腰痛の既往と関連がないという結果だったが、すべりのある例が少なかったことが結果に影響した可能性はある。

E．結論

腰椎 MRI における Pfirrmann 分類≧3、椎間板変性、HIZ は過去の腰痛の既往と関連があった。すべりは関連がなかった。関連が示された所見は、高度な腰痛が再発する可能性を予測する所見の一部と考えられた。

F．健康危険情報

特記すべき事項なし。

G．研究発表

Tonosu J, Oka H, Matsudaira K, Higashikawa A, Okazaki H, Tanaka S. The relationship between findings on magnetic resonance imaging and previous history of low back pain. Journal of Pain Resarch. 2017;10: 47-52.

H．知的財産権の出願・登録状況

現時点ではなし。

I．参考文献

1. Fujii T, Matsudaira K. Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. Eur Spine J. 2013;22:432-438.
2. Cheung KM, Karppinen J, Chan D, Ho DW, Song YQ, Sham P, et al. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine. 2009;34: 934-940.
3. Endean A, Palmer KT, Coggon D. Potential of magnetic resonance imaging findings to refine case definition for mechanical low back pain in epidemiological studies: a systematic review. Spine. 2011;36: 160-169.
4. Tamcan O, Mannion AF, Eisenring C, et al. The course of chronic and recurrent low back pain in the general population. Pain 2010;150:451-7.
5. Dionne CE, Dunn KM, Croft PR, et al. A consensus approach toward the

standardization of back pain definitions for use in prevalence studies. Spine 2008;33:95-103.
6. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine. 2001;26:1873-1878.
7. Fardon DF, Milette PC; Combined Task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. Nomenclature and classification of lumbar disc pathology. Recommendations of the Combined task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. Spine. 2001;26:E93-E113.
8. Aprill C, Bogduk N. High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. Br J Radiol 1992;65:361-9.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
10. Cheung KM, Karppinen J, Chan D, et al. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine 2009;34:934-40.
11. Lee SH, Daffner SD, Wang JC. Does lumbar disk degeneration increase segmental mobility in vivo? Segmental motion analysis of the whole lumbar spine using kinetic MRI. J Spinal Disord Tech 2014;27:111-6.

労労災疾病臨床研究事業費補助金
総合研究分担報告書

サブテーマ②予防に有用な福祉機器等の開発
抗力を具備した継手付き体幹装具による，歩行時の脳機能へ効果

研究分担者　四津有人　東京大学医学部附属病院

研究要旨
腰痛は労災疾病対策の重要課題であり，その背後には不良姿勢がある事が多い．近年，不良姿勢を改善するための，継手に抗力を具備する体幹装具が開発された．本装具の装着により，静止立位時や歩行時の姿勢が改善し，歩行パフォーマンスも向上することが確認されているが，神経系への影響は未だ調べられていなかった．研究分担者は脳機能計測の専門かであり H28 年度より研究班に参画した。本研究では健常者を対象に，抗力付き体幹装具装着による，歩行時の高次運動野の活動への影響を検討した．結果，歩行安定期に補足運動野の活動は装着下でより早期に低下し，運動前野の活動は装着下でより晩期に低下した．抗力付き体幹装具の装着により，垂直姿勢の維持や運動制御に関わる高次運動野の活動が修飾されたと考える．

A．研究目的

厚生労働省が公表する業務上疾病発生状況等調査によると，休業 4 日以上の業務上疾病の発生件数のうち腰痛は，長年に渡り全職業性疾病の約 6 割を占め第 1 位であった．また，世界疾病負担研究でも腰痛が Years Lived with Disability のトップにランクされている．このように，腰痛は日本国内のみならず世界的にも頻度が高く，これによる社会的損失は大きい．腰痛は労災疾病対策の重要課題である．

腰痛の背後には，不良姿勢があることが多い．不良姿勢を矯正する旧来の体幹装具は，腹部を圧迫して固定する方式か，胸部・腰背部・恥骨部の 3 点で固定する方式であり，これらの方式を用いた体幹装具を長期間使用すると体幹筋群が弱化すると報告されている．近年，継手の抗力によって体幹を伸展方向に回転させる力と骨盤を前傾方向に回転させる力を与えて姿勢を矯正し，腹筋群の活動を促す新たな抗力を具備した継手付き体幹装具（以下抗力付き体幹装具）が開発された[1]．本装具の装着により，静止立位時や歩行時の姿勢が改善し，歩行パフォーマンスも向上することが確認されている[2-4]．

ヒトの姿勢・動作は神経系によって制御されており，大脳皮質では一次運動野のみならず，補足運動野や運動前野などの高次運動野も関与している[5,6]．抗力付き体幹装具の装着で姿勢・動作が変化した際には，これらの皮質の活動も変化していることが予想されるが，これを示した研究は未だ無い．そこで本研究では，抗力付き体幹装具装着による，歩行時の高次運動野の活動への影響を調べた．

B．研究方法

対象を健常者 10 人とした．各人に，抗力付き体幹装具非装着と装着の 2 条件でトレッドミル歩行を行ってもらい，その際の脳活動を調べた．

脳活動の計測には，近赤外光脳機能計測装置（OEG-17APD, Spectratech 社）を用いた．プローベは運動野・補足運動野・運動前野を計測できるように３×８で配置した．各条件で，安静立位時をベースラインとし歩行時の活動を計測した．

解析は，ノイズ等で適切な信号を得られなかった 2 名を除き，8 名で行った.

（倫理面への配慮）

実施に際し，新潟医療福祉大学倫理委員会での承認を得た．調査への参加は完全な任意であり，インフォームドコンセントの上行われた.

C．研究結果

補足運動野は，歩行安定期に，抗力付き体幹装具非装着下に比べ装着下ではより早期に有意に低下した.

左右の運動前野は，歩行安定期に，抗力付き体幹装具非装着下に比べ装着下ではより晩期に低下した.

D．考察

補足運動野は垂直姿勢の維持に関わるとされ，歩行安定期に抗力付き体幹装具装着下でより早期に有意に低下したのは，垂直姿勢が抗力付き体幹装具によって補完されたためと考える．運動前野の背側部は視覚情報による運動の空間的制御に関わるとされ，抗力付き体幹装具装着下の歩行安定期の低下がより晩期に起こったのは，姿勢の変化による視覚座標系の変更による可能性がある.

垂直姿勢を保つことができない不良姿勢は筋骨格系だけでなく，高次運動野にも影響を及ぼしている可能性がある．今後は，実際に不良姿勢を呈する人や腰痛患者を対象に計測を行っていく予定である.

E．結論

抗力付き体幹装具の装着により，垂直姿勢の維持や運動制御に関わる高次運動野の活動が修飾されたと考える.

F．健康危険情報

該当なし.

G．研究発表

1. 論文発表　なし

2. 学会発表　なし

（発表誌名巻号・頁・発行年等も記入）

H．知的財産権の出願・登録状況（予定を含む.）

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし

I．参考文献

[1] 飯島進乃他：“抗力を具備した骨盤前傾を促す継手付き体幹装具が高齢者の歩行に与える影響”，理学療法学，Vol. 41，No. 6，pp.355-363，2014.

[2] J. Katsuhira, N. Miura, T. Yasui, T. Mitomi, and S. Yamamoto, "Efficacy of a newly designed trunk orthosis with joints providing resistive force in adults with post-stroke hemiparesis," Prosthet Orthot Int, vol. 40, no. 1, pp. 129-136, 2016.

[3] J. Katsuhira, K. Matsudaira, H. Oka et al., "Efficacy of a trunk orthosis with joints providing resistive force on low back load during level walking in elderly persons," Clin Interv Aging, vol. 11, pp. 1589-1597, 2016.

[4] J. Katsuhira, K. Matsudaira, T. Yasui, S. Iijima, and A. Ito, "Efficacy of a trunk orthosis with joints providing resistive force on low-back load in elderly persons during static standing," Clin Interv Aging, vol. 10, pp. 1413-1420, 2015.

[5] M. Mihara, I. Miyai, N. Hattori et al., "Cortical control of postural balance in patients with hemiplegic stroke," Neuroreport, vol. 23, no. 5, pp. 314-319, 2012.

[6] I. Miyai, H. C. Tanabe, I. Sase et al., "Cortical mapping of gait in humans: a near-infrared spectroscopic topography study," Neuroimage, vol. 14, no. 5, pp. 1186-1192, 2001.

労災疾病臨床研究事業費補助金

総合研究分担報告書

サブテーマ③：介護看護従事者への予防介入とマネジメントシステムの構築

「労災病院に勤務する看護師に対する腰痛予防の大規模介入研究」

研究分担者 三好光太 横浜労災病院 整形外科

研究要旨

厚生労働省調査にて、業務上疾病の発生件数は、腰痛が全職業性疾病の約 6 割を占め第 1 位であること、平成 23 年の腰痛全届け出のうち社会福祉施設で腰痛が顕著な増加を辿っていることなどから介護・看護従事者への腰痛対策は、産業衛生領域の喫緊の課題といえる。

本研究では、産業衛生領域の喫緊の課題である腰痛対策を効率的に行うために、簡易で即実践できる体操に加え、産業理学療法士からの科学的根拠に基づいた教育の有益性に大規模介入比較試験を施行した。

H26 年度には参加施設の職員数からサンプルサイズ設計を行い、**H27 年度**には全国 12 労災病院をクラスターとして、A:対照（無介入)、B：腰椎伸展体操の普及・実践、C:B+産業理学療法士による腰痛教育・相談の実践の 3 群の無作為比較試験を開始、介入前のベースライン調査を行った（3,381 名分の有効回答)。**H28 年度**に実施した 6 か月後の追跡調査時の、各群の回収数は A 群 949 名、B 群 706 名、C 群 751 名、計 2,406 名であり、追跡率はそれぞれ 71.9%、70.6%、67.0%で、全体では 70.0%あった。

腰痛と関連情報を把握するためのアンケート調査を行った結果、腰痛の自覚症状改善の割合は,A 群で 13.3%、B 群で 23.5%、C 群で 22.6%と介入群で上昇していた。また腰痛予防対策の実行度はコントロール群で低くなっていた。

A. 研究目的

厚生労働省業務上疾病発生状況等調査にて、腰痛における休業 4 日以上の業務上疾病の発生件数は、全職業性疾病の約 6 割を占め第 1 位となっている。平成 23 年の腰痛全届け出のうち社会福祉施設が 19%を占め、10 年で 2.7 倍という最も顕著な増加となった背景を踏まえ、19 年ぶりに改訂された「職場における腰痛予防対策指針」(平成 25 年、厚生労働省)では、重症心身障害児施設等に限定されていた適用を、福祉・医療等における介護・看護作業全般に拡大し、内容を充実させるに至った。つまり、介護・看護従事者への腰痛対策は、産業衛生領域の喫緊の課題といえる。また世界疾病負担研究にて 289 の疾患や傷病のうち、腰痛が Years Lived with Disability (YLDs)のトップにランクされるなど、社会的損失や健康面への影響の大きい腰痛への対策は global にも重要な課題として位置づけられている。

また疾患の対策としては、高リスク群のみに限定して対策を行うハイリスク・アプローチは、高リスクと考えられなかった大多数集団が潜在的なリスクを抱えたていた場合、効果的な手法とは言えない。このため対象を一部に限定せずに集団全体へアプローチをし、全体としてリスクを下げ集

団としての健康状態を向上させるポピュレーションアプローチが注目を集めている。

本研究では、産業衛生領域の喫緊の課題である腰痛対策を効率的に行うために、簡易で即実践できる体操に加え、産業理学療法士からの科学的根拠に基づいた教育の有益性をポピュレーションアプローチに基づいた大規模介入比較試験で検討した。

B. 研究方法

H26 年度には参加施設の職員数からサンプルサイズ設計を行い、H27 年度には全国 12 労災病院をクラスターとして、A:対照（無介入)、B：腰椎伸展体操の普及・実践、C:B+産業理学療法士による腰痛教育・相談の実践の 3 群を実施するため、統計学的な見地を踏まえデザインを行い、介入を実施、H28 年度には追跡調査が終了した。

（論理面への配慮）

本研究は、研究対象者の組み入れ前であるが、同意取得やデータは匿名化の方法は確立しており、研究遂行にあたり倫理面での問題はないとの承認を、全国労災病院倫理委員会より得ている。

C. 研究結果

以下の研究プロトコールの通りに、ベースライン調査を実施した。

①施設をクラスターとした無作為比較試験

選択基準:選定された労災病院に勤務する成人（20 歳以上）看護師、本研究の趣旨に賛同し同意を得た者

除外基準:妊婦, あるいは妊娠の疑いがある場合、腰椎伸展により症状が誘発される腰部脊柱管狭窄症と診断されたことがある者、研究の同意を撤回した者

②対照（無介入)、腰椎伸展体操の普及・実践、Bの介入+産業理学療法士による腰痛教育・相談の実践の 3 群

③北海道中央（看護師数：156)、東北（407)、関東（562)、横浜（667)、新潟（274)、浜松（256)、旭（182)、大阪（720)、関西（674)、中国（391)、愛媛（193)、長崎（285)、総計 4,767 名。以上 12 労災病院（施設）のをクラスターとし、病床・看護師数、看護師の男女数・平均年齢を割付調整因子とし、コンピューターの乱数表を用い、3 群（4 施設ごと）に無作為割付する非盲検試験を行った。

④A 群は北海道中央、横浜、大阪、浜松の 1,799 名、B 群は関東、旭、中国、長崎の 1,420 名、C 群は東北、新潟、関西、愛媛の 1,548 名、全体で 4,767 名にアンケートを配布した。全体でのアンケート回収数は 3,439 名分で、回収率は 72.1%だった。各群の回収数は A 群 1,319 名、B 群 1,000 名、C 群 1,120 名であり、回収率はそれぞれ 73.3%、70.4%、72.4%であった。

回収したアンケートのうち 58 名に不備があったためベースライン解析には 3,381 名分のアンケートを利用した(A 群 1,292 名、B 群 987 名、C 群 1,102 名)。

ベースライン調査での各群の背景情報は以下のとおりである：

	A 群	B 群	C 群
年齢	35.5 (35.0-36.1)	35.1 (34.5-35.8)	35.5 (34.9-36.1)
性 男性（%）	6.7	5.3	4.2
BMI	21.2 (21.0-21.3)	21.5 (21.3-21.6)	21.1 (20.9-21.3)
StarTBack high risk(%)	2.2	2.8	2.2
FABQ 15 点以上（%）	27.7	30.2	29.6

EQ5D	0.88 (0.87-0.89)	0.87 (0.86-0.88)	0.88 (0.87-0.89)

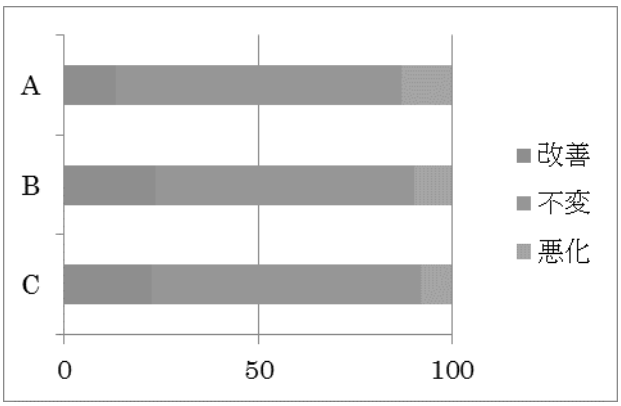
上表内の（　）には 95%信頼区間を示した。
各群の背景情報の分布は上表に示すとおりであり、全ての群で似通った傾向であった。

⑤6 か月後の追跡調査時の、各群の回収数は A 群 949 名、B 群 706 名、C 群 751 名、計 2,406 名であり、追跡率はそれぞれ 71.9%、70.6%、67.0%で、全体では 70.0%あった。以下に追跡可能だった症例のベースライン時における各群での背景情報を記載する。

	A 群	B 群	C 群
年齢	36.8 (36.1-37.5)	36.1 (35.3-36.9)	35.1 (34.7-36.3)
性 男性（%）	7.1	6.3	6.2
BMI	21.3 (21.1-21.5)	21.6 (21.3-21.8)	21.2 (20.9-21.4)
StarTBack high risk(%)	2.0	2.1	1.9
FABQ 15 点以上-BL （%）	26.8	28.9	29.1
EQ5D-BL	0.88 (0.87-0.89)	0.87 (0.86-0.88)	0.88 (0.87-0.89)
EQ5D-6M	0.88 (0.87-0.89)	0.87 (0.86-0.88)	0.89 (0.89-0.90)

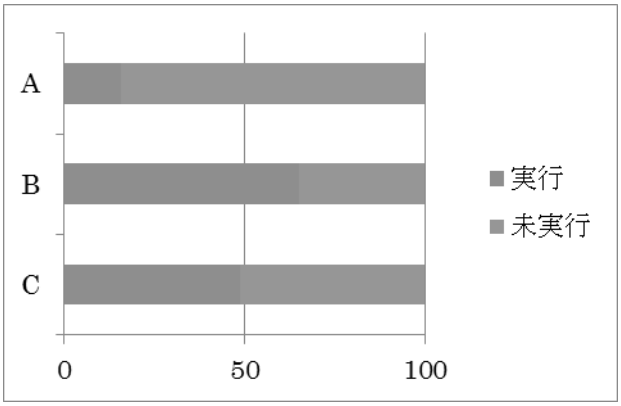
上表内の（　）には 95%信頼区間を示した。
各群の背景情報の分布は上表に示すとおりであり、全ての群で似通った傾向であった。前述した全例でのベースライン調査での各群の背景情報追跡可能例とでは、背景情報の傾向は異ならなかった。

⑥本研究の主要評価項目は腰痛の自覚的改善度である。



各群の改善、不変、悪化の割合(%)を上図に示す。A,B,C 群での改善の割合は,13.3%、23.5%、22.6%であった。悪化の割合は 13.0%、9.6%、8.1%と介入の度合いが高いほど減少していた（Cochran-Armitage の傾向検定:P< 0.0001）。

腰痛予防対策の実行度(%)を下図に示す



A,B,C 群での実行度の割合は 15.6%、64.9%、48.8%であり A 群（コントロール群）での実行度が低くなっていた（カイ 2 乗検定：残差分析 p<0.05）。

腰痛の改善を目的変数として、背景を調整しても介入治療効果が認められるかに関して多変量解析（Logistic 回帰分析）を用いて検討した。雇用の安定等に関する法律（高年齢者雇用安定法）をも

とに、45 歳以上を「中高年齢者」と、また BMI 25 以上を肥満と定義した。

	Odds 比	95%信頼区間	p 値
性（男性）	1.0	(0.7-1.7)	0.89
中高年齢者	0.9	(0.7-1.1)	0.36
肥満	0.8	(0.6-1.1)	0.10
StarTBack high risk	0.8	(0.4-1.6)	0.70
FABQ 15 点未満	1.4	(1.1-1.7)	0.01
A 群 vs. B 群	2.1	(1.6-2.7)	<0.0001
A 群 vs. C 群	2.0	(1.5-2.6)	<0.0001

多重共線性を検討するために、各説明変数の分散拡大係数:variance inflation factor (VIF) を算出した。この結果、性（男性）・中高年齢者・肥満・StarTBack high risk・FABQ15 点未満・治療 B 群・治療 C 群で、それぞれ 1.0、1.0、1.0、2.1、1.1、1.3、1.3 でいずれも 10 を超えておらず、説明変数間に多重共線性は生じていないものと判定した。

多変量を調整した Logistic 回帰分析の結果、B の介入（腰椎伸展体操の普及・実践）、C の介入（B の介入+産業理学療法士による腰痛教育・相談の実践）とも有意に腰痛を改善（コントロール群の約 2 倍）することが分かった。また FABQ が 15 点未満であることは腰痛改善の因子であることが明らかになった。

D. 考察

産業衛生領域の喫緊の課題である腰痛対策を効率的に行うために、簡易で即実践できる体操に加え、産業理学療法士からの科学的根拠に基づいた教育の有益性を検証するために大規模介入比較試験を施行した。研究３年目となる本年度は、統計学的な検討に基づいた割付を行い、6 か月の期間をおき、腰痛と関連情報を把握するためのアンケート調査を行った。この結果、腰痛の自覚症状改善の割合は,コントロール群で 13.3%、腰椎伸展体操の普及・実践群で 23.5%、+産業理学療法士による腰痛教育・相談の実践 22.6%と、いずれの介入によっても上昇していた。

また腰痛予防対策の実行度はコントロール群で低くなっていた。多変量を調整した Logistic 回帰分析の結果、両介入群とも有意に腰痛を改善（コントロール群の約 2 倍）することが分かった。また FABQ が 15 点未満であることは腰痛改善の因子であることが明らかになった。

E. 結論

ポピュレーションアプローチに基づいた介入研究を行い、介入群で腰痛の自覚的改善度、腰痛予防対策の実行度が高くなっていることが明らかになった。

F. 健康危険情報

該当なし

G. 研究発表

1. 論文発表

現時点でなし

2. 学会発表

現時点でなし

H. 知的財産権の出願・登録状況(予定を含む)

特許取得　実用新案登録

現時点でなし

労災疾病臨床研究事業費補助金

総合研究分担報告書

サブテーマ④個人と職場の双方に有益な腰痛治療と職業生活との両立支援手法の開発

「腰痛予防への効率的かつ効果的な理学療法介入に関する研究」

研究分担者

野村卓生　　関西福祉科学大学 保健医療学部 リハビリテーション学科

浅田史成　　大阪労災病院 治療就労両立支援センター

高野賢一郎　　関西労災病院 治療就労両立支援センター

研究要旨

平成 26 年度においては，我々が過去に行った課題名「いつでもフィジカルコンサルティング」(Physical Consultant 研究）のデータベースを利用し，指導者（理学療法士）へのアンケート結果等から，インターネットや携帯端末機器を利用した効果的なメール指導を行う上での基礎資料を得ることを目的に検討を行った.

平成 27 年度においては，メール指導を効率的かつ効果的に行うためのシステム（Consulting system for physical therapy in occupational health, Compo）の開発，腰痛予防教育教材（スクリーンセイバー，動画）の作製を行った．また，労働者への理学療法（産業理学療法）に関する先進国の一つであるオーストラリアにおいて視察を行った.

平成 28 年度においては，まず，Compo を用いて，介護職員を対象としたメール指導の効果検証を無作為化比較デザインで行った．ついで，諸外国の情報をふまえ，日本の理学療法士が腰痛予防へ関わっていく上での現状と課題を分析した．最後に，腰痛予防の重要性を広く認知してもらうために，これまでに作製した腰痛予防の教育教材をインターネットや SNS を利用して普及啓発を行い，さらにこれらをどのように活用していけばよいか検討した.

A. 研究目的

平成 26 年度には我々が過去に行った Physical Consultant 研究（PCo 研究）のデータベースを用いて，事務系職員を対象とした理学療法士による腰痛予防を目的としたメール指導の現状と問題点を分析した．メール指導に一定の効果を認めるが，多数の労働者を対象にする場合には，簡便に使用可能で，かつ多数のデータを管理するためのデータベース・システムの構築が必要と考えられた.

平成 27 年度にはメール指導を効率的かつ効果的に行うためのシステムとして，産業理学療法システム「Consulting system for physical therapy in occupational health: Compo」の開発を行った．また，腰痛予防に対する理学療法の情報を国内外から収集，腰痛予防の重要性や具体的な予防方法を普及啓発させることを目的として腰痛予防のための教育教材の作製

を行った.

平成 28 年度においては，まず，Compo を用いたメール指導の効果検証について介護職員を対象とし，無作為化比較デザインで行うことを目的とした．ついで，諸外国の情報をふまえ，日本の理学療法士が腰痛予防へ関わる上での現状と課題を分析することを目的とした．最後に，腰痛予防の重要性を広く認知してもらうために，これまでに作製した腰痛予防の教育教材をインターネットや SNS を利用して普及啓発を行い，さらにこれらをどのように活用していけばよいか検討することを目的とした.

B. 研究方法

1) 平成 26 年度

PCo 研究のデータベースを利用して，メール指導のあり方やその効果を検討した．具体的には，PCo 研究究終了後に行った指導者へのアンケート調査結果を分析した．指導者は臨床経験 3 年以上で現職を有する理学療法士である.

アンケート項目は，①相談者の個人情報の必要性，②指導者が相談者に返信するまでの期間，③相談者からの返信に対して指導者が回答文章の作成に要する時間，④相談者が指導者に返信したメールの回数，⑤指導者が考える相談者からの相談に対して返信内容を考える理想的な時間，⑥腰痛以外の相談内容，⑦指導者が考える 1 ヵ月に対応可能な相談者の人数，⑧指導者が考える PCo 研究での問題点とした.

2) 平成 27 年度

多数の労働者を対象としてメール指導ができるように独自のシステムを開発した．また，世界における労働者を対象とした理学療法（産業理学療法）の介入やその教育課程に関して情報収集を行った．さらに，腰痛予防教育教材を開発した.

専用のシステムは，産業理学療法指導システム「Consulting system for physical therapy in occupational health: Compo」と命名した．産業理学療法先進国の視察については，世界理学療法連盟において，Network for Occupational Health and Ergonomics の代表を務める理学療法士の Dr. Rose Boucaut 氏（南オーストラリア大学）の協力を得て，オーストラリアにおける産業理学療法を視察することとした．腰痛予防教育教材は，スクリーンセイバーと動画（映像）を作製することとした.

3) 平成 28 年度

開発したシステムを用いて，保健衛生業に従事する労働者を対象として，メール指導による腰痛予防効果を検証した．また，世界の産業理学療法の情報収集を継続した．さらに，開発した腰痛予防教育教材の普及を行った.

Compo を用いた介入研究は，臨床試験登録システム UMIN-CTR(UMIN000018450)に登録した．30 歳から 65 歳までの保健衛生業に従事する者を対象として，Compo を用いて指導を行う群（介入群）と介入を行わない群（対照群）の 2 群に振り分け，研究を実施した．研究手順として，国内の 11 施設の協力を得た．介入・観察期間は 6 カ月である．まず，指導者から対象者へメールを送信し，以降，1 カ月に 1 回，指導者から対象者へメールを送信することを原則とした（計 7 回）．指導者は対象者からの相談に対して個別に対応することとした．尚，対照群についても観察期間終了後，介入群と同様に指導を行うこととした.

対象者の一般特性として，性別，年齢，身長，体重，喫煙の有無，介護業務の経験年数，管理職の有無，夜勤の有無，1 週間当たりの労働時間などのデータ収集を行った．また，介入あるいは観察前後で，ここ 4 週間の腰痛の程度を 0（まったく痛みのない状態）-10（想像しうる最悪の痛み）点法の Visual Analog Scale(VAS)，

ここ 30 日の仕事の出来を 0（最低）-10（最高）点法，抑うつの状態を K6 質問票日本版，心理社会的要因を Subgrouping for Targeted Treatment (STarT) Back スクリーニングツール日本語版，腰痛に対する恐怖回避思考を日本版 Fear-Avoidance Beliefs Quetionnaire（FABQ-J），健康関連の生活の質（health-related quality of life, HRQoL）を日本語版 EuroQoL で評価し効用値を算出した．さらに，介入群では，介入終了後に，理学療法士によるメール相談によって腰痛予防の効果があると感じたか，理学療法士によるメール相談の満足度についてなどの聴取を行った．

C. 研究結果

1) 平成 26 年度

指導者へのアンケート結果について，①相談者の個人情報の必要性については，全ての指導者が相談者の個人情報が必要と回答した．②指導者が相談者に返信するまでの期間については，相談者から返信があった際に，指導者が返信までにかかる期間の平均は，2.0±0.4 日であった．③相談者からの返信に対して指導者が回答文章の作成に要する時間については，指導者が回答の文章を作成するのに要した時間の平均は，22.0±9.7 分であった．④相談者が指導者に返信したメールの回数については，相談者が指導者へ返信したメール回数の平均は 4.7±4.2 回であった．⑤指導者が考える相談者からの相談に対して返信内容を考える理想的な時間について，指導者が考える理想的な時間の平均は，20.3±13.7 分であった．⑥腰痛以外の相談内容については，肩凝り・肩の痛み：2 件，冷え性：1 件，朝起きた際のこわばり：1 件，家族の介護：1 件，肥満：1 件，運動不足：1 件，下肢の痛み：1 件あった．⑦指導者が考える 1 ヵ月に対応可能な相談者の人数については，現所属を有する指導者が 1 ヵ月に対応可能な相談者の人数の平均は，5.3±7.4 名であった．⑧指導者が考える PCo 研究での問題点と今後の改善策について，複数に認めた意見として，事前に対象者の情報が欲しい：4 件，メール自体を確認してもらったかどうかが不明である：2 件，メール指導のガイドラインや具体例が必要である：2 件，返信がない場合の指導者の対処方法を決めて欲しい：2 件があった．その他の意見として，相手方が必要としている情報を提供できているのか不明：1 件，指導に対する効果の判定が難しい：1 件，指導者によって差がでた場合，相談者の不満に繋がる可能性がある：1 件，返信のない相談者に対する対応に迷う，相手の質問に対してどこまで踏み込むべきか不明である：1 件，匿名であるので相談者と信頼関係を築くのが難しい：1 件，定量的な相談者と指導者が共通認識できるデータやツールが欲しい：1 件があった．

2) 平成 27 年度

開発した Compo の機能概要は以下の通りである．パソコンでも，携帯電話でも使用可能である．特定の URL を入力し，個別の ID とパスワードでログインする．相談者は担当の指導者へテキストで相談を送信することができ，画像などの種々のファイルも添付可能である．指導者は相談者からの相談内容に応じて返信を行う．相談と指導のやり取りは，システムを通すし，仮名設定を前提とするので，相互の個人情報が開示・他者から確認されることはない．指導者からはアンケートなども一斉送信で容易に実施可能となっており，その結果も CSV でダウンロード可能で，研究事務局で一括管理できる．対象者から相談のあった場合に指導者へ通知されるアラート機能（登録したメールアドレスへシステム上に相談者から連絡のあった場合，リアルタイムに通知される．同様に指導者からシステム上で返信した場合，相談者へリアルタ

イムに登録したメールアドレスへ通知される）を装備している．

産業理学療法先進諸国の視察では，南オーストラリア大学（アデレード）ならびにシドニー大学（シドニー）において，産業理学療法に関する教育，研究，理学療法の実際や理学療法士の役割に関する視察を行った．南オーストラリア大学理学療法学科では4年次に講義・現場での実習を含め，多くの時間をかけて「産業保健と安全管理(occupational health and safety)」に関する理学療法，理学療法士の役割が教授される．南オーストラリア大学理学療法学科の 1 学年の定員は 100 名であり，1 学年全体で受講する講義のほか，Tutorial などは少人数制できめ細かく行われる（Occupational Health and Safety であれば，当該科目をもつ教員においては，学生は異なるが同じ内容を6回開講する）．例えば，ワイナリーに勤務する勤労者の健康管理と安全対策をテーマにする学生では，そのワイナリーへ実際に何度も出向き，詳細に仕事・作業の内容を調査して問題を抽出する．一連の調査内容を数十枚にわたるレポートにまとめ，学内で他の学生や教員と議論の上，エビデンスをふまえて対策方法を現場の安全管理者や従業員へ提案し議論を行っていた．

腰痛予防教育教材について，スクリーンセイバーはスライド枚数・全5枚で作製した．看護師を中心とした保健衛生業に従事する者を対象にして，松平の開発した「これだけ体操」の実施を促す構成とした．動画（映像）についても同様に保健衛生業に従事する者を対象にして，腰痛の発生を誘引しないことを目的としたトランスファーの技術指導，これだけ体操や体幹の屈曲・伸展・回旋運動などの腰痛予防体操の実施を促す3部構成で作製した．

3) 平成 28 年度

介入群および対照群において，全ての一般特性の項目に有意な差は認めなかった．各群において，介入/観察前後の全ての項目に有意な差は認めなかった．また，2 群間の変化量についても全ての項目に有意な差を認めなかった．介入群における介入終了後の対象者の感想については，理学療法士によるメール相談によって腰痛予防の効果があると感じたかについての質問において，「あまり効果がなかった」，「まったく効果がなかった」と回答した6名およびメール相談の満足度について，「不満がある」と回答した2名の自由記載の意見に関して代表的な内容を以下に示す：腰痛自体は軽い方だが，一度仕事中にぎっくり腰になった際，具体的な対処方法などの指導はなく腰痛防止には役立っていなかった；現在のところ痛みがない為，相談らしい相談を行うことができなかった；メールのやり取りがスムーズでなく，確認の頻度が減っていった．

世界の産業理学療法の情報収集について，理学療法士が対象者へ関わるためには日本では医師の処方が必要である．一方，英国では1978年から医師の処方箋がなくとも理学療法士が必要と判断して行う治療も国民保健サービスでカバーされ，開業の有無や届出等も関係するが対象者による理学療法士へのダイレクトアクセスが可能である国がある．理学療法士の教育制度について，米国では大学卒業後に約3年をかけての大学院教育，豪国では4年制の大学教育で行われるなど一定に統制された高等教育で理学療法士養成を行っている国がある一方，日本では3年制および4年制の大学・専門学校混合教育で理学療法士の養成が行われている．日本では，産業保健分野の理学療法に関しては，ほぼ全ての養成校で教育されていないのが現状である．理学療法士が行うことのできる業務範囲については開業権が認められている国や消炎鎮痛薬の処方が認められる国があるなど，各国で異なるのが実情である．

腰痛予防教材の普及に関して，スクリーンセイバーについては，（一社）産業理学療法研究会の会員に無料で提供することとし，自らの職場や研究フィールドでの活用を促した．さらに，スクリーンセイバーについては，問い合わせのあった場合には無償提供しており，今後は研究会のホームページを通して会員以外にも無料で提供する予定である．動画については，3 部構成の一部である「様々な状況を想定し腰痛発生の予防を目指したトランスファー技術」の一部について，YouTube で公開した．今後，更なる動画の普及を計画している．一方，腰痛予防を目的とした教育教材を開発し，無料で，かつ全国的に利用できるようにインターネットや SNS を活用して普及啓発しているが，指導者となる会員などからは，教育教材を用いて，どのように労働者へ教育すれば効果的なのか，教育教材の活用方法に関する研修会が必要との意見が多数認められた．

D. 考察

PCo 研究の方法論上，個人情報の保護の観点から，研究終了後にも指導者は相談者を特定できない，相談者は指導者を特定できないという点に配慮して，双方にマスキングを行った．これについては，全ての指導者から個人情報の必要であるとの回答を得たが，今後の展開も鑑み（例えばスマートフォンのアプリ化をした場合など），不特定多数からの相談でも対応できるような指導方法のマニュアル化が必要と考えられた．もちろん，例えば企業との契約となった際などは，企業の同意や個人の同意を得たうえで，相談者の個人情報を指導者に提供することも可能であるが，逆に相談者への指導者の個人情報の提供に関しても検討することが必要と考えられた．

指導者が考える相談者からの相談に対して返信内容を考える理想的な時間の平均は 20.3±13.7 分，実際に返信までに要した時間は 22.0±9.7 分であり，理想的な時間通りに返信できていたことが伺えた．指導者が返信までに要した期間の平均は 2.0±0.4 日であり，円滑に返信ができていたと想定される．1 ヵ月に何名の相談者を担当できるかについては平均で 5.3±7.4 名と幅が大きいが，現職を有する今回の理学療法士で 100 名程度の相談者へのメール指導は同期間に可能と考えられた．PCo 研究では相談者から腰痛以外の相談もあり，その概要は運動器の症状から，生活習慣の問題や家族の介護の問題まで幅広かった．これら医療・介護の問題に対応するには，リハビリテーションの領域に精通している理学療法士等でなければ難しいと思われ，産業保健分野においても理学療法士のニーズは高いと思われた．

PCo 研究の成果を基盤として，より効果的・効率的なメール指導が行うことを目的として開発した Compo を用い，腰痛の発生が多い業種である介護職員を対象にして，メール指導の効果を検証した．中間解析の結果，メール指導の有効性は有意差として認められなかった．現在の分析は中間解析の結果であり，対象数が増えれば結果が異なる可能性があるが，現状の分析においても効果を認めない，あるいは効果が表れにくい対象がいることは確実と考えられる．どのような業種，どのような身体的・心理社会的要因を有する対象，またその他要因を有する対象者にメール指導の効果があるのかを検証することは，メール指導の対象となるターゲットを明確にするためにも重要であり，今後の研究課題である．

理学療法士の社会的地位が高い欧米諸国において，4 年間の教育で理学療法士免許が授与される国がある中で，日本も修業年限だけは 3～4 年間であるので，何をもって教育内容が充実，レベルの違いがあるかについては諸外国と日本との単純比較は難しい．歴史的背景，法制

度をふまえて，日本の理学療法士の卒前教育において産業理学療法が教授されてこなかったことについては不適切とは言及できないが，日本の理学療法士が産業保健分野で活躍するためには，産業理学療法に関する教育が必要不可欠である．コアカリキュラムが設定される卒前教育において，新たな産業理学療法のカリキュラムを設定して，それに多くの時間を費やすのは現状では難しいため，卒後教育において産業理学療法に関する教育内容をいかに充実させていくかが課題と考える．

エビデンスの構築にあたっては，労働者を雇用する側および保険組合へ如何に理学療法の必要性や有効性，理学療法士が関わることでのメリットを示していくかが重要であると考える．例えば，雇用する側と保険組合に対しては，某企業において理学療法士が関わることにより，労働力の損失を防止，勤労者の生産性の向上に寄与することができるかを示せれば雇用する側にとって有用であり，加えて医療費の削減効果を示すことができれば保険組合としても有用である．これらの観点をふまえて，理学療法のエビデンスを構築していくことができれば，産業保健分野での理学療法士の活躍の場は広がると考えられる．

我々は，腰痛予防を目的とした教育教材を開発し，これらをインターネットや SNS を利用して普及啓発した．今後は，これらの効果的な使い方（労働者への教育方法や労働環境への導入方法など）を検討し，人的手段あるいはインターネットによる教育教材を使用する側（指導者側）の教育も継続しなければならないと考えている．

E. 結論

1. 理学療法士による腰痛予防を目的としたメール指導は，対象者の満足度が比較的高い．
2. 理学療法士が腰痛予防を目的としたメール指導を行えば，どのような業種，どのような対象にでも効果を認めるとは研究結果からは言及できない．
3. 今後，メール指導に効果のある業種やターゲット層を明確にする必要があり，効果が検証される必要がある．
4. 法制度上，理学療法士養成のカリキュラムにおいて，日本と諸外国では違いがあり，腰痛予防を目的とする介入を行うには卒前教育だけでは不十分である．
5. 日本の理学療法士が腰痛予防に関わるためには卒後教育の充実化が必要であり，エビデンスの構築と共に社会に発信することが重要である．
6. 腰痛予防を目的した教育教材を普及啓発させるにはインターネットや SNS を利用することが効果的である．
7. 一方で，開発した教育教材を労働者に適応，現場へ効果的に導入するには，指導者側への教育も必要と考えられた．

F. 研究発表

論文発表

1. 高野賢一郎．産業理学療法の展開．総合リハビリテーション 43：527-534，2015
2. 野村卓生．産業保健と理学療法．理学療法ジャーナル 50：83-85，2016
3. 高野賢一郎，浅田史成，野村卓生，明崎禎輝，松平浩，山縣英久．いつでもどこでもフィジカルコンサルティング．日職災医誌，64：101-106，2016
4. Nomura T, Asada F, Takano K, Matsudaira K. The current state along with outstanding issues related to email-based guidance by physical therapists aiming to prevent low back

pain among workers. JJOMT 64: 113-118, 2016
5. Asada F, Nomura T, Kubota M, Ohashi M, Ito K. Evaluation of a physical activity promotion program using the “Exercise Guide 2006” aimed at preventing lifestyle-related diseases among working people. JJOMT 64: 162-172, 2016
6. 野村卓生. 一次予防領域における健康管理への理学療法士の貢献. 日衛誌 71: 107-110, 2016
7. 浅田史成, 高野賢一郎. 勤労者の運動器障害に対する理学療法について. 日衛誌 71: 111-118, 2016

G. 知的財産権の出願・登録状況

1. 特許取得
   なし
2. 実用新案登録
   なし
3. その他
   なし

謝辞

本研究を行うにあたって協力いただいた（一社）産業理学療法研究会の会員, 関係者各位に深謝する. また, Compo を用いた研究において, 対象者の募集やシステム管理に貢献した森本信三氏（白浜はまゆう病院・理学療法士), 明崎禎輝氏（四国がんセンター・理学療法士）に感謝する.

労災疾病臨床研究事業費補助金

総合研究分担報告書

サブテーマ④個人と職場の双方に有益な腰痛治療と職業生活との両立支援手法の開発

「介護職の勤務体制と健康状況との関連性の検討」

研究分担者　高井ゆかり　東京大学大学院医学系研究科健康科学・看護学専攻成人看護/緩和ケア看護学分野　（研究協力者　高橋正也 独立行政法人 労働安全衛生総合研究所）

研究要旨

介護職で健康で働けることは提供するサービスの向上につながると考えられる。本調査では，H26-27 年度に高齢者介護施設で働く介護職を対象に，労働負荷を考慮しながら，勤務体制と健康状況（作業支障腰痛および不眠と精神的不調）との関連を検討した。調査を進めるにあたり，勤務時間が長くなるにつれて，また労働負荷が高くなるにつれて，健康への望ましくない影響が現れるという仮説を設定し，介護施設（155 施設）に対して，本調査への参加を打診した。回答のあった 86 施設のうち，50 施設が参加意思を示した。参加意思を示した施設には所属する介護職の人数分の調査票を送付し，3,155 名より回答が得られた（回答率 77%）。分析の結果，労働負荷を主観評価に基づいた場合，作業支障腰痛の訴えは夜勤が長くなるにつれて増加することが分かった。それに対して，不眠と精神的不調の訴えは夜勤の長さにかかわらず，主観的労働負荷が高いと増加した。一方，労働負荷を介護職 1 名あたりの利用者数に基づいて推定した場合，夜勤が長い（16 時間以上の）時にのみ，有意な関連が認められた。労働負荷の 2 つの指標は意味するところが異なるかもしれないが，介護職の健康確保という点では，介護労働に伴う精神的・肉体的負荷を軽減することが夜勤の短縮とともに重要になると考えられた。そのためには，福祉機器の活用や，施設内で安全衛生のレベルを向上させるマネジメントシステムの確立などが求められる。

A. 研究目的

松平らの大規模全国調査によると，わが国の腰痛の生涯有病率は，8 割を超え，4 人に 1 人は、仕事等の社会活動を休んだ経験があり，腰痛は，誰でも経験しうる国民的愁訴であるといえる。加えて，厚生労働省業務上疾病発生状況等調査によると，腰痛における休業 4 日以上の業務上疾病の発生件数は，全職業性疾病の約 6 割を占め第 1 位となっている。平成 23 年の腰痛全届け出のうち社会福祉施設が 19%を占め，10 年で 2.7 倍という最も顕著な増加となった背景を踏まえ，19 年ぶりに改訂された「職場における腰痛予防対策指針」(平成 25 年、厚生労働省)では，重症心身障害児施設等に限定されていた適用を，福祉・医療等における介護・看護作業全般に拡大し，内容を充実させるに至った。つまり，社会福祉法人等の介護の現場で活躍する介護福祉士，ケアワーカーといった介護職への腰痛対策は，産業衛生領域の喫緊の課題といえる。さらには労働負担が少なくないと想定される介護職が健康で働ける環境を提供することはサービスの向上につながると考えられる。一方，近年，腰痛と心理社会的問題に伴う精神的不調は，密接な関連があることが明らかになり，先述「職場における腰痛予防対策指針」(平成 25 年、厚生労働省)でもこの問題が指摘された。以上の背景から本調査では，高齢者介護施設で働

く介護職を対象に，労働負荷を考慮しながら，勤務スケジュールと健康状態（作業支障腰痛および不眠と精神的不調）との関連を検討した。調査を進めるにあたり，勤務時間が長くなるにつれて，また労働負荷が高くなるにつれて，健康状態への望ましくない影響が現れるという仮説を設定した。

B. 研究方法

対象

社会福祉懇談会の会員である介護施設（155 施設）に対して，本調査への参加を打診した。回答のあった 86 施設のうち，50 施設が参加意思を示した（参加拒否 5 施設，高齢者介護を行わない施設 31 施設）。参加意思を示した施設には所属する介護職の人数分の調査票を送付し（合計 4,105 名分），配布回収を求めた。

調査項目

匿名の自記式調査票では次の項目を測定した。

勤務の状況：働く施設（特別養護老人ホーム［特養］，老人保健施設［老健］，介護療養型施設，グループホーム，その他），この 1 ヶ月における勤務の状況（早出，日勤，遅出／準夜勤，夜勤の頻度，始業・終了時刻，担当介護職の人数，利用者の数），主観的労働負荷（夜勤，夜勤以外），夜勤中仮眠（頻度，長さ），雇用形態，職種，週労働時間，仕事のストレス要因（仕事の量的負荷，裁量権，上司・同僚からの社会的支援），現在の勤務体制に対する主観的適応度等。

健康の状況：過去 1 ヶ月の腰痛（仕事に支障をきたしたり，欠勤をしたりした腰痛を作業支障腰痛ありと定義。腰痛の範囲は，肋骨縁より下部で下殿溝より上部として明確に図示し，これに下肢痛を伴う場合も含むと定義），睡眠（入眠困難，中途覚醒，早朝覚醒のいずれかが週に 3 回以上ある場合を不眠ありと定義，睡眠時間，起床時疲労感，昼間の強い眠気），精神的不調（K6 で 5 点以上を不調ありと定義），生活習慣等。

統計解析

夜勤の長さを 3 群に分けた（短:9 時間まで，中：9.1-15.9 時間，長：16 時間以上）。労働負荷は主観的評価と介護職 1 名あたりの利用者数からそれぞれ評価した。前者は 4 項目の合計得点（4～36 点）の中央値（28）で低高二分した。後者はその中央値（20）で低高二分した。これらに基づいて，6 群（短低，中低，長低，短高，中高，長高）を 2 組作成した。夜勤以外については各勤務の長さの平均値を求め，9 時間未満か以上で二分した。主観的労働負荷は中央値（26）で二分した。各勤務における介護職 1 名あたりの利用者数の平均値を求め，その中央値（8.9）で低高二分した。これらに基づいて，4 群（短低，長低，短高，長高）を 2 組作成した。夜勤あるいは夜勤以外の長さと労働負荷が作業支障腰痛，不眠，精神的不調にどのように関連するかを調べるために，短低群を参照としたロジスティック回帰分析を行った。年齢，性別，週労働時間，特養か否か（全施設の分析時のみ）による影響は統計的に調整した。また，施設による違いを検討するために，上記の分析を特養，老健，グループホーム別に行った。

（論理面への配慮）

独立行政法人労働安全衛生総合研究所の倫理審査の承認を得て推進した。被験者に対してはデータを ID 化して管理するなど個人情報には十分配慮すること等を説明し，書面での同意を取得した。

C. 研究結果

3,155 名より回答が得られた（回答率 77%）。

全施設（2,218 名）における夜勤の長さ（短中長）と主観的労働負荷の結果であるが，作業支障腰痛の訴えは労働負荷が高いと，夜勤が長くなるにつれて増加した。不眠と精神的不調の訴えは夜勤の長さにかかわらず，労働負荷が高いと増加した。この傾向は，特養のみ（1,572 名），老健のみ（192 名）でも同様であった。グループホームのみ（239 名）では，作業支障腰痛，不眠，精神的不調の訴えはいずれも，夜勤の長さと労働負荷との関連は

認められなかった。

全施設（2,262 名）における夜勤の長さ（短中長）と介護職 1 名あたりの利用者数に関する分析結果に関してであるが，作業支障腰痛の訴えは介護職 1 名あたりの利用者数が高く，夜勤が長いと増加した。不眠と精神的不調の訴えについて，夜勤の長さおよび労働負荷との関連は認められなかった。特養のみ（1,607 名）でも作業支障腰痛の訴えは介護職 1 名あたりの利用者数が高く，夜勤が長いと増加し，不眠の訴えについては，夜勤の長さおよび労働負荷との関連は認められなかったが，精神的不調の訴えは労働負荷が高く，夜勤が長いと増加した。老健のみ（195 名）では，作業支障腰痛，不眠，精神的不調の訴えはいずれも，夜勤の長さと介護職 1 名あたりの利用者数との関連は認められなかった。なお，夜勤が中で労働負荷が高である群は 2 名であったため，分析から除外した。グループホームのみ（284 名）においては，介護職 1 名あたりの利用者数該当者が少なく各群 5 名未満であったため，分析から除外した。各指標について介護職 1 名あたりの利用者数が低となる群のみでみても，夜勤の長さとの関連は認められなかった。

全施設（2,832 名）での夜勤以外の長さ（短長）と主観的労働負荷に関する分析結果であるが，作業支障腰痛，不眠，精神的不調の訴えはいずれも，労働負荷にかかわらず，夜勤以外の長さが長いと増加した。この傾向は，特養のみ（1,949 名）でも同様であった。老健のみ（252 名）では，作業支障腰痛，不眠，精神的不調の訴えはいずれも，夜勤以外の長さおよび労働負荷との関連は認められなかった。グループホームのみ（288 名）では，作業支障腰痛と不眠の訴えは夜勤以外の長さが長いと増加するような傾向があったが，統計的には有意ではなかった。精神的不調の訴えは労働負荷が高く，夜勤以外の長さが長いと増加したが，結果は不安定であった。

全施設（2,734 名）における夜勤以外の長さ（短長）と介護職 1 名あたりの利用者数の分析に関してであるが，作業支障腰痛，不眠，精神的不調の訴えのいずれも，夜勤以外の長さおよび介護職 1 名あたりの利用者数との関連は認められなかった。特養のみ（1,895 名）でも同様に，3 指標いずれも，夜勤以外の長さおよび介護職 1 名あたりの利用者数との関連は認められなかった。老健のみ（241 名）でも，作業支障腰痛と不眠の訴えは，夜勤以外の長さおよび介護職 1 名あたりの利用者数との関連は認められなかった。精神的不調の訴えについては，参照群となる夜勤以外の長さが短で介護職 1 名あたりの利用者数が低である群が 5 名未満であったため，分析から除外した。グループホームのみ（284 名）でもまた，作業支障腰痛と不眠の訴えはいずれも，夜勤以外の長さおよび介護職 1 名あたりの利用者数との関連は認められなかった。精神的不調の訴えでは，労働負荷が高く，夜勤以外の長さが長いと増加したが，結果は不安定であった。なお，夜勤以外の長さ長で介護職 1 名あたりの利用者数が低である群が 5 名未満であったため，分析から除外した。

D. 考察

労働負荷を主観評価に基づいた場合，作業支障腰痛の訴えは夜勤が長くなるにつれて増加することが分かった。それに対して，不眠と精神的不調の訴えは夜勤の長さにかかわらず，主観的労働負荷が高いと増加した。一方，労働負荷を介護職 1 名あたりの利用者数に基づいて推定した場合，夜勤が長い（16 時間以上の）時にのみ，有意な関連が認められた。労働負荷の 2 つの指標は意味するところが異なるかもしれないが，介護職の健康確保という点では，介護労働に伴う精神的・肉体的負荷を軽減することが夜勤の短縮とともに重要になると考えられた。そのためには，福祉機器の活用や，施設内で安全衛生のレベルを向上させるマネジメントシステムの確立などが求められる。それらを通じて，介護職の健康のみならず，介護の質的改善や介護職の離職防止も期待できる。早出，日勤，遅出／準夜勤など夜勤以外の勤

務では，労働負荷よりむしろ，勤務時間の長さがカギになるように思われた。従って，夜勤以外であっても，勤務時間の不必要な延長には注意が必要と言える。

施設の型ごとに調べると，勤務時間および労働負荷と腰痛，不眠，精神的不調との関連は特養，老健，グループホームによって異なることが示唆された。本調査では特養で働く介護職が大半であり，それ以外の施設で働く介護職の人数は少なかった。この人数の問題は大きく関わっている可能性はある。また，対象となった施設の介護度という面では，特養 3.9±SD 0.3，老健 3.3±0.3，グループホーム 2.8±0.4 と各群の間で有意差もあった（$P<0.05$）。従って，各施設の介護業務の特徴に即して勤務条件を適正に調整することが肝要と考えられた。

E. 結論

高齢者介護施設で働く介護職を対象に，勤務体制と健康状況（作業支障腰痛および不眠と精神的不調）との関連を検討した結果，作業支障腰痛の訴えは夜勤が長くなるにつれて増加した。不眠と精神的不調の訴えは夜勤の長さにかかわらず，主観的労働負荷が高いと増加した。

F. 健康危険情報

該当なし

G. 研究発表

1.論文発表

- Takai, Y., Yamamoto-Mitani, N., Abe Y. & Suzuki, M. Literature review of pain management for people with chronic pain, Japan Journal of Nursing Science, Article first published online: 19 NOV 2014 | DOI: 10.1111/jjns.12065
- Takai, Y., Yamamoto-Mitani, N., Kawakami, S., Abe, Y., Kamiyama, M. & Saito, S. Differences between Nurses' and Care workers' Estimations of Pain Prevalence among Elderly Residents, Pain Management Nursing, 16 (1): 20-32, 2015.
- Takai, Y., Abe, A., Torimoto-Sasai, Y., Okamoto, Y., Kamiyama, M. & Yamamoto-Mitani, N. Compliance of urinary continence care for Japanese residents on short-stay respite services, International Journal of Urological Nursing, 8(2):90-96, 2014.
- Takai, Y., Yamamoto-Mitani, N., & Ko, A. Prevalence of and factors related to pain among elderly Japanese residents in long-term healthcare facilities, Geriatrics & Gerontology International, 14(2):481-9, 2014
- Takai, Y., Yamamoto-Mitani, N., Ko, A., & Heilemann, M.V. Differences in pain measures by Mini-Mental State Examination scores of residents in aged care facilities: Examining the usability of the Abbey Pain Scale-Japanese version, Pain Management Nursing, 15(1): 236-245, 2014.
- Takai, Y., Yamamoto-Mitani, N., Chiba, Y., & Kato, A. Feasibility and clinical utility of the Japanese Version of the Abbey Pain Scale in Japanese aged care, Pain Management Nursing, 15(2): 439-448, 2014.
- 鈴木みずえ，古田良江，高井ゆかり，佐藤文美，大城一，山本則子，金森雅夫，認知症高齢者における疼痛の有症率と疼痛が認知症の行動・心理症状(BPSD)に及ぼす影響，老年看護学，19(1):25-33，2014
- 古田良江，鈴木みずえ，高井ゆかり，在宅高齢者の痛みと健康関連 QOL の関連と痛み緩和対策の実態，日本早期認知症学会誌 7(2):26-35，2014
- 鈴木 みずえ，山本 則子，高井 ゆかり，古田 良江，鈴木 有希，金森 雅夫：認知症高齢者の痛みに関するアセスメントツールとケア介

入，日本早期認知症学会誌，7(1)：53-58，2014

- 古田良江，鈴木みずえ，高井ゆかり，在宅虚弱高齢者である二次予防事業参加者の疼痛有症率と疼痛の状況が健康関連 QOL に及ぼす影響，老年看護学，18(2)：48-57，2014.

２．学会発表

- Takai Y & Yamamoto-Mitani N:A feasibility study of developed chronic pain management standards for older residents，第 34 回日本看護科学学会学術集会，名古屋，2014 年 11 月
- 高井ゆかり，古田良江，阿部吉樹，鈴木みずえ，交流集会「高齢者の慢性痛へのセルフケアを支える看護」，第 34 回日本看護科学学会学術集会，名古屋，2014 年 11 月
- Takai Y，Yamamoto-Mitani N,：A qualitative study of expectations and attitudes toward doctors/therapists among elderly Japanese patients with chronic pain，14th World Congress on Pain (International Association for the Study of Pain: IASP)，Buenos Aires，Argentina，2014 年 10 月
- 高井ゆかり，山本則子，認知症高齢者を在宅で介護する家族員が介護サービスを決定・利用する際のパターン，日本家族看護学会第 21 回学術集会，倉敷市，2014 年 8 月
- 古田良江，鈴木みずえ，高井ゆかり，二次予防事業参加者の疼痛緩和対策の実態，日本老年看護学会第 19 回学術集会，名古屋，2014 年 6 月
- 高井ゆかり，山本則子，千葉いくみ，体の痛みと共に老いを生きる 慢性痛のある高齢者の経験の探索，日本老年看護学会第 19 回学術集会，名古屋，2014 年 6 月
- 山本則子，高井ゆかり，鈴木美穂，五十嵐歩，竹原君江，長期療養施設におけるケアの質向上の取り組みに関する文献レビュー，日本老年看護学会第 19 回学術集会，名古屋，2014 年 6 月

H. 知的財産権の出願・登録状況(予定を含む)

１．特許取得

なし

２．実用新案登録

なし

３．その他

Ⅲ. 研究成果の刊行に関する一覧

研究成果の刊行に関する一覧表
【H26.4.1～H29.3.31】

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Ohya J, Miyoshi K, Oka H, Matsudaira K, Fukushima M, Nagata K.	Optimal measurement for "posterolateral protrusion" of the vertebral artery at the craniovertebral junction using computed tomography angiography.	J Craniovertebr Junction Spine	5	151-156	2014
Hasegawa T, Katsuhira J, Matsudaira K, Iwakiri K, Maruyama H.	Biomechanical Analysis of Low Back Load when Sneezing.	Gait Posture	40	670-675	2014
Matsudaira K, Konishi H, Miyoshi K, Isomura T, Inuzuka K.	Potential risk factors of persistent low back pain developing from mild low back pain in urban Japanese workers.	PLoS One	9	e93924	2014
Yamada K, Matsudaira K, Takeshita K, Oka H, Hara N, Takagi Y.	Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey.	Mod Rheumatol	24	343-348	2014
Omata Y, Hagiwara F, Nishino J, Matsudaira K, Kadono Y, Juji T, Mori T, Nakayama H, Nagase Y, Hirose J, Yasui T, Matsumoto T, Matsui T, Tohma S, Tanaka S.	Vertebral fractures affect functional status in postmenopausal rheumatoid arthritis patients.	J Bone Miner Medab	32	725-731	2014
Matsudaira K, Kikuchi N, Murakami A, Isomura T.	Psychometric properties of the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ).	J Orthop Sci	19	26-32	2014
Yoshimura N, Akune T, Fujiwara S, Shimizu Y, Yoshida H, Omori G, Sudo A, Nishiwaki Y, Yoshida M, Shimokata H, Suzuki T, Muraki S, Oka H, Nakamura K.	Prevalence of knee pain, lumbar pain and its coexistence in Japanese men and women: The Longitudinal Cohorts of Motor System Organ (LOCOMO) study.	J Bone Miner Medab	32	524-532	2014
Nagata K, Yoshimura N, Hashizume H, Muraki S, Ishimoto Y, Yamada H, Takiguchi N, Nakagawa Y, Minamide A, Oka H, Kawaguchi H, Nakamura K, Akune T, Yoshida M.	The prevalence of cervical myelopathy among subjects with narrow cervical spinal canal in a population-based magnetic resonance imaging study: the Wakayama Spine Study.	Spine J	14	2811-2817	2014

Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, Oka H, Ishimoto Y, Nagata K, Kagotani R, Takiguchi N, Akune T, Kawaguchi H, Nakamura K, Yoshida M.	Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study.	Osteoarthritis Cartilage	22	104-110	2014
Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Yoshimura N.	Exercise habits during middle age are associated with lower prevalence of sarcopenia: the ROAD study.	Osteoporos Int	25	1081-1088	2014
Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Tokimura F, Yoshida H, Suzuki T, Nakamura K, Yoshimura N.	Association of physical activities of daily living with the incidence of certified need of care in the long-term care insurance system of Japan: the ROAD study.	J Orthop Sci	19	489-496	2014
Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Tokimura F, Yoshida H, Suzuki T, Nakamura K, Yoshimura N.	Incidence of certified needed care in the long-term care insurance system and its risk factors in the elderly of Japanese population-based cohorts: the ROAD study.	Geriatr Gerontol Int	14	695-701	2014
Yoshimura N, Nagata K, Muraki S, Oka H, Yoshida M, Enyo Y, Kagotani R, Hashizume H, Yamada H, Ishimoto Y, Teraguchi M, Tanaka S, Kawaguchi H, Toyama Y, Nakamura K, Akune T.	Prevalence and progression of the radiographic ossification of posterior longitudinal ligament and its associated factors in the Japanese populations: A 3-year follow-up of the ROAD study.	Osteoporos Int	25	1089-1098	2014
Hashizume H, Yoshimura N, Nagata K, Miyazaki N, Ishimoto Y, Nishiyama R, Oka H, Yamada H, Yoshida M.	Development and evaluation of a video exercise program for locomotive syndrome in the elderly.	Mod Rheumatol	24	250-257	2014
Muraki S, Akune T, Nagata K, Ishimoto Y, Yoshida M, Tokimura F, Tanaka S, Oka H, Kawaguchi H, Nakamura K, Yoshimura N.	Association of knee osteoarthritis with onset and resolution of pain and physical functional disability: The ROAD study.	Mod Rheumatol	24	966-973	2014
Muraki S, Akune T, En-yo Y, Yoshida M, Tanaka S, Kawaguchi H, Nakamura K, Oka H, Yoshimura N.	Association of dietary intake with joint space narrowing and osteophytosis at the knee in Japanese men and women: The ROAD Study.	Mod Rheumatol	24	236-242	2014
Matsudaira K, Hiroe M, Kikkawa M, Suzuki M, Isomura T, Oka H, Hiroe K, Hiroe K.	Can standing back extension exercise improve or prevent low back pain in Japanese care workers?	J Man Manip Ther	23	205-209	2015

Katsuhira J, Matsudaira K, Yasui T, Iijima S, Ito A.	Efficacy of a trunk orthosis with joints providing resistive force on low back load in elderly persons during static standing.	Clin Interv Aging	10	1413-1420	2015
Kikuchi N, Matsudaira K, Sawada T, Oka H.	Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK) in patients with whiplash neck injury pain and/or low back pain.	J Orthop Sci	20	985-992	2015
Oka H, Matsudaira K, Fujii T, Okazaki H, Shinkai Y, Tsuji Y, Tanaka S, Kato R.	Risk factors for prolonged treatment of whiplash-associated disorders.	PLoS One	10	e0132191	2015
Takahashi M, Matsudaira K, Shimazu A.	Disabling low back pain associated with night shift duration: sleep problems as a potentiator.	Am J Ind Med	58	1300-1310	2015
Matsudaira K, Kawaguchi M, Isomura T, Inuzuka K, Koga T, Miyoshi K, Konishi H.	Assessment of psychosocial risk factors for the development of non-specific chronic disabling low back pain in Japanese workers- Findings from the Japan epidemiological research of Occupation-related Back pain (JOB) study.	Ind Health	53	368-377	2015
Yoshimura N, Akune T, Fujiwara S, Shimizu Y, Yoshida H, Nishiwaki Y, Sudo A, Omori G, Yoshida M, Shimokata H, Suzuki T, Muraki S, Oka H, Nakamura K.	Incidence of disability and its associated factors in Japanese men and women: the Longitudinal Cohorts of Motor System Organ (LOCOMO) study.	J Bone Miner Medab	33	186-191	2015
Muraki S, Akune T, Nagata K, Ishimoto Y, Yoshida M, Tokimura F, Tanaka S, Kawaguchi H, Nakamura K, Oka H, Yoshimura N.	Does osteophytosis at the knee predict health-related quality of life decline? A 3-year follow-up of the ROAD study.	Clin Rheumatol	34	1589-1597	2015
Ohya J, Oshima Y, Takeshita K, Oka H, Chikuda H, Taniguchi Y, Matsubayashi Y, Tanaka S.	Patient satisfaction with double-door laminoplasty for cervical compression myelopathy.	J Orthop Sci	20	64-70	2015
Ogihara S, Yamazaki T, Maruyama T, Oka H, Miyoshi K, Azuma S, Yamada T, Murakami M, Kawamura N, Hara N, Terayama S, Morii J, Kato S, Tanaka S.	Prospective multicenter surveillance and risk factor analysis of deep surgical site infection after posterior thoracic and/or lumbar spinal surgery in adults.	J Orthop Sci	20	71-77	2015

Muraki S, Akune T, En-Yo Y, Yoshida M, Suzuki T, Yoshida H, Ishibashi H, Tokimura F, Yamamoto S, Tanaka S, Nakamura K, Kawaguchi H, Oka H, Yoshimura N.	Joint space narrowing, body mass index, and knee pain: the ROAD study (OAC1839R1).	Osteoarthritis Cartilage	23	874-881	2015
Kato S, Oshima Y, Oka H, Chikuda H, Takeshita Y, Miyoshi K, Kawamura N, Masuda K, Kunogi J, Okazaki R, Azuma S, Hara N, Tanaka S, Takeshita K.	Comparison of the Japanese Orthopaedic Association (JOA) score and modified JOA (mJOA) score for the assessment of cervical myelopathy: a multicenter observational study.	PLoS One	10	e0123022	2015
Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T.	Mutual associations among musculoskeletal diseases and metabolic syndrome components: A 3-year follow-up of the ROAD study.	Mod Rheumatol	25	438-448	2015
Yamada H, Terada M, Iwasaki H, Endo T, Okada M, Nakao S, Hashizume H, Minamide A, Nakagawa Y, Nishi H, Tsutsui S, Oka H, Yoshida M.	Improved accuracy of diagnosis of lumbar intra and/or extra-foraminal stenosis by use of three-dimensional MR imaging: comparison with conventional MR imaging.	J Orthop Sci	20	287-294	2015
Nakamura M, Hashizume H, Oka H, Okada M, Takakura R, Hisari A, Yoshida M, Utsunomiya H.	Physical Performance Measures Associated With Locomotive Syndrome in Middle-Aged and Older Japanese Women.	J Geriatr Phys Ther	38	202-207	2015
Mure K, Yoshimura N, Hashimoto M, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Takeshita T.	Urinary 8-iso-prostaglandin F2α as a marker of metabolic risks in the general Japanese population: The ROAD study	Obesity	23	1517-1524	2015
Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T.	Factors affecting changes in the serum levels of 25-hydroxyvitamin D: a 3-year follow-up of the ROAD study.	Osteoporos Int	26	2597-2605	2015
Yoshimura N, Muraki S, Oka H, Tanaka S, Ogata T, Kawaguchi H, Akune T, Nakamura K.	Association between new indices in the locomotive syndrome risk test and decline in mobility: third survey of the ROAD study.	J Orthop Sci	20	896-905	2015
Yamada H, Oka H, Iwasaki H, Endo T, Kioka M, Ishimoto Y, Nagata K, Takiguchi N, Hashizume H, Minamide A, Nakagawa Y, Kawai M, Tsutsui S, Yoshida M.	Development of a support tool for the clinical diagnosis of symptomatic lumbar intra- and/or extra-foraminal stenosis.	J Orthop Sci	20	811-817	2015

Muraki S, Akune T, Teraguchi M, Kagotani R, Asai Y, Yoshida M, Tokimura F, Tanaka S, Oka H, Kawaguchi H, Nakamura K, Yoshimura N.	Quadriceps muscle strength, radiographic knee osteoarthritis and knee pain: the ROAD study.	BMC Musculoskelet Disord	16	305	2015
Minamide A, Yoshida M, Yamada H, Hashizume H, Nakagawa Y, Nishi H, Iwasaki H, Tsutsui S, Okada O, Okada S, Oka H.	Efficacy of Posterior Segmental Decompression Surgery for Pincer Mechanism in Cervical Spondylotic Myelopathy: A Retrospective Case-controlled Study Using Propensity Score Matching.	Spine	40	1807-1815	2015
Ito H, Takatori Y, Moro T, Oshima H, Oka H, Tanaka S.	Total Hip Arthroplasty After Rotational Acetabular Osteotomy.	J Arthroplasty	30	403-406	2015
Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, Minamide A, Nakagawa H, Ishimoto Y, Nagata K, Kagotani R, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshida M.	The association of combination of disc degeneration, endplate signal change, and Schmorl node with low back pain in a large population study: the Wakayama Spine Study.	Spine J	15	622-628	2015
Katsuhira J, Matsudaira K, Oka H, Iijima S, Itou A, Yasui T, Yozu A.	Efficacy of a trunk orthosis with joints providing resistive force on low back load during level walking in elderly persons.	Clin Interv Aging	11	1589-1597	2016
Tonosu J, Inanami H, Oka H, Katsuhira J, Takano Y, Koga H, Yuzawa Y, Shiboi R, Oshima Y, Baba S, Tanaka S, Matsudaira K.	Diagnosing Discogenic Low Back Pain Associated with Degenerative Disc Disease Using a Medical Interview.	PLoS One	11	e0166031	2016
Tsuji T, Matsudaira K, Sato H, Vietri J.	The impact of depression among chronic low back pain patients in Japan.	BMC Musculoskelet Disord	17	447	2016
Taniguchi Y, Takahashi M, Matsudaira K, Oka H, Momose T.	Potential use of 18F-FDG-PET/CT to visualize muscle pain in patients with adult spinal deformity: A case report.	Skeletal Radiol	45	1577-1581	2016
Hara N, Matsudaira K, Masuda K, Tohnosu J, Takeshita K, Kobayashi A, Murakami M, Kawamura N, Yamakawa K, Terayama S, Ogihara S, Shiono H, Morii J, Hayakawa K, Kato S, Nakamura K, Oka H, Sawada T, Inuzuka K, Kikuchi N.	Psychometric Assessment of the Japanese Version of the Zurich Claudication Questionnaire (ZCQ): Reliability and Validity.	PLoS One	11	e0160183	2016

Iwahashi H, Yoshimura N, Hashizume H, Yamada H, Oka H, Matsudaira K, Shinto K, Ishimoto Y, Nagata K, Teraguchi M, Kagotani R, Muraki S, Akune T, Tanaka S, Kawaguchi H, Nakamura K, Minamide A, Nakagawa Y, Yoshida M.	The Association between the Cross-Sectional Area of the Dural Sac and Low Back Pain in a Large Population: The Wakayama Spine Study.	PLoS One	11	e0160002	2016
Yamada K, Matsudaira K, Imano H, Kitamura A, Iso H.	Influence of work-related psychosocial factors on the prevalence of chronic pain and quality of life in chronic pain patients.	BMJ Open	6	e010365	2016
Hayashi S, Katsuhira J, Matsudaira K, Maruyama H.	Effect of pelvic forward tilt on low back compressive and shear forces during a manual lifting task.	J Phys Ther Sci	28	802-806	2016
Vargas-Prada S, Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SS, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Sarquis LM, Marziale MH, Harari F, Freire R, Harari N, Monroy MV, Quintana LA, Rojas M, Harris EC, Serra C, Martinez JM, Delclos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Freimann T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VC, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A, Vega EJ.	Descriptive Epidemiology of Somatising Tendency: Findings from the CUPID Study.	PLoS One	11	e0153748	2016
Tonosu J, Matsudaira K, Oka H, Okazaki H, Oshio T, Hanaoka I, Muraoka Y, Midorikawa M, Wakabayashi K, Tanaka S.	A population approach to analyze the effectiveness of a back extension exercise "One Stretch" in patients with low back pain: A replication study.	J Orthop Sci	21	414-418	2016
Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S.	Psychometric Properties of the Japanese Version of the STarT Back Tool in Patients with Low Back Pain.	PLoS One	11	e0152019	2016

Nomura T, Asada F, Takano K, Matsudaira K.	The current state along with outstanding issues related to email-based guidance by physical therapists aiming to prevent low back pain among workers.	JJOMT	64	113-118	2016
Matsudaira K, Hara N, Oka H, Kunogi J, Yamazaki T, Takeshita K, Seichi S, Tanaka S.	Predictive factors for subjective improvement in lumbar spinal stenosis patients with nonsurgical treatment: a 3-year prospective cohort study.	PLoS One	11	e0148584	2016
Shimazu A, Matsudaira K, De Jonge J, Tosaka N, Watanabe K, Takahashi M.	Psychological Detachment from Work during Nonwork Time: Linear or Curvilinear Relations with Mental Health and Work Engagement?	Ind Health	54	282-292	2016
Sawada T, Matsudaira K, Muto Y, Koga T, Takahashi M.	Potential risk factors for onset of severe neck and shoulder discomfort (Katakori) in Urban Japanese workers.	Ind Health	54	230-236	2016
Sarquis LM, Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Vargas-Prada S, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SS, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Marziale MH, Harari F, Freire R, Harari N, Monroy MV, Quintana LA, Rojas M, Harris EC, Serra C, Martinez JM, Delclos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Freimann T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VC, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A, Salazar Vega EJ	Classification of neck/shoulder pain in epidemiological research a comparison of personal and occupational characteristics, disability and prognosis among 12,195 workers from 18 countries.	Pain	157	1028-1036	2016
Iidaka T, Muraki S, Akune T, Oka H, Kodama R, Tanaka S, Kawaguchi H, Nakamura K, Yoshimura N.	Prevalence of radiographic hip osteoarthritis and its association with hip pain in Japanese men and women: the ROAD study.	Osteoarthritis Cartilage	24	117-123	2016

Ohashi S, Ohnishi I, Oka H, Matsumoto T, Bessho M, Nakamura K, Tanaka S.	The effect of cartilage degeneration on ultrasound speed in human articular cartilage.	Mod Rheumatol	26	426-434	2016
Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, Minamide A, Ishimoto Y, Nagata K, Kagotani R, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshida M.	Metabolic Syndrome Components Are Associated with Intervertebral Disc Degeneration: The Wakayama Spine Study.	PLoS One	11	e0147565	2016
Kodama R, Muraki S, Oka H, Iidaka T, Teraguchi M, Kagotani R, Asai Y, Yoshida M, Morizaki Y, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshimura N.	Prevalence of hand osteoarthritis and its relationship to hand pain and grip strength in Japan: The third survey of the ROAD study.	Mod Rheumatol	26	767-773	2016
Ohya J, Oshima Y, Oka H, Saiki F, Taniguchi Y, Matsubayashi Y, Tanaka S, Chikuda H, Takeshita K.	Patient Satisfaction with Posterior Decompression Surgery for Cervical Ossification of the Posterior Longitudinal Ligament: Prognostic Radiographic Factors and Patient-Reported Outcomes for the Effectiveness of Surgical Treatment.	World Neurosurg	96	272-279	2016
Oshima Y, Takeshita K, Taniguchi Y, Matsubayashi Y, Doi T, Ohya J, Soma K, Kato S, Oka H, Chikuda H, Tanaka S.	Effect of Preoperative Sagittal Balance on Cervical Laminoplasty Outcomes.	Spine	41	E1265-E1270	2016
Ono K, Ohashi S, Oka H, Kadono Y, Yasui T, Omata Y, Nishino J, Tanaka S, Tohma S.	The impact of joint disease on the Modified Health Assessment Questionnaire scores in rheumatoid arthritis patients: A cross-sectional study using the National Database of Rheumatic Diseases by iR-net in Japan.	Mod Rheumatol	26	529-533	2016
Teraguchi M, Samartzis D, Hashizume H, Yamada H, Muraki S, Oka H, Cheung JP, Kagotani R, Iwahashi H, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Cheung KM, Yoshimura N, Yoshida M.	Classification of High Intensity Zones of the Lumbar Spine and Their Association with Other Spinal MRI Phenotypes: The Wakayama Spine Study.	PLoS One	11	e0160111	2016
Nakamura M, Kobashi Y, Hashizume H, Oka H, Kono R, Nomura S, Maeno A, Yoshida M.	Utsunomiya H. Locomotive syndrome is associated with body composition and cardiometabolic disorders in elderly Japanese women.	BMC Geriat	16	166	2016

Matsudaira K, Oka H, Kawaguchi M, Murakami M, Fukudo S, Hashizume M, Löwe, B.	Development of a Japanese Version of the Somatic Symptom Scale-8: Psychometric Validity and Internal Consistency.	Gen Hosp Psychiatry	45	7-11	2017
Tonosu J, Oka H, Matsudaira K, Higashikawa A, Okazaki H, Tanaka S.	The relationship between the findings on magnetic resonance imaging and previous history of low back pain.	J Pain Res	10	47-52	2017
Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S.	The Japanese version of the STarT Back Tool predicts 6-month clinical outcomes of low back pain.	J Orthop Sci			2016 Epub ahead of print
Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Vargas-Prada S, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SS, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Sarquis LM, Marziale MH, Harari F, Freire R, Harari N, Monroy MV, Quintana LA, Rojas M, Harris EC, Serra C, Martinez JM, Delclos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Freimann T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VC, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A, Salazar Vega EJ.	Epidemiological differences between localised and non-localised low back pain.	Spine			2016 Epub ahead of print
Isomura T, Sumitani M, Matsudaira K, Kawaguchi M, Inoue R, Hozumi J, Tanaka T, Oshima H, Mori K, Taketomi S, Inui H, Tahara K, Yamagami R, Hayakawa K.	Development of the Japanese Version of the Leeds Assessment of the Neuropathic Symptoms and Signs Pain Scale (LANSS-J): Diagnostic Utility in a Clinical Setting.	Pain Practice			2016 Epub ahead of print
Wakaizumi K, Yamada K, Oka H, Kosugi S, Morisaki H, Shibata M, Matsudaira K.	Fear-avoidance beliefs are independently associated with the prevalence of chronic pain in Japanese workers.	J Anesth			2017 Epub ahead of print

Fukushima M, Oka H, Hara N, Oshima Y, Chikuda H, Tanaka S, Takeshita K, Matsudaira K.	Prognostic factors associated with the surgical indication for lumbar spinal stenosis patients less responsive to conservative treatments.	J Orthop Sci			in press
Oka H, Matsudaira K, Kikuchi N, Haga Y, Sawada T, Katsuhira J, Yoshimoto T, Kawamata K, Tonosu J, Sumitani M, Kasahara S, Tanaka S.	Estimated risk for chronic pain determined using the generic STarT Back 5-item screening tool.	J Pain Res			in press
Yamada K, Matsuadira K, Tanaka E, Oka H, Katsuhira J, Iso H.	Sex-specific impact of early-life adversity on chronic pain: A large population-based study in Japan.	J Pain Res			in press
Yasui T, Oka H, Omata Y, Kadono Y, Tanaka S.	Relationship between roentgenographic joint destruction in the hands and functional disorders among patients with rheumatoid arthritis.	Mod Rheumatol			in press
Yoshimura N, Muraki S, Oka H, Iidaka T, Kodama R, Kawaguchi H, Nakamura K, Tanaka S, Akune T.	Is osteoporosis a predictor for future sarcopenia or vice versa? Four-year observations between the second and third ROAD study surveys.	Osteoporos Int			in press
Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, Takiguchi N, Minamide A, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T, Yoshida M.	Association of Lumbar Spondylolisthesis with Low Back Pain and Symptomatic Lumbar Spinal Stenosis in a Population-based Cohort: The Wakayama Spine Study.	Spine			in press

Ⅳ. 研究成果の刊行物・別刷

J C J S V

Editor-in-Chief : Atul Goel (INDIA)

Open Access HTML Format

For entire Editorial Board visit : http://www.jcvjs.com/editorialboard.asp

Journal of Craniovertebral Junction and Spine

Original Article

Optimal measurement for "posterolateral protrusion" of the vertebral artery at the craniovertebral junction using computed tomography angiography

Junichi Ohya[1], Kota Miyoshi[1], Hiroyuki Oka[2], KO Matsudaira[2], Masayoshi Fukushima[1], Kosei Nagata[1]

[1]Department of Orthopaedic surgery, Yokohama Rosai Hospital,Yokohama, Japan, [2]Department of Medical Reserch and Management for Musculoskeltal Pain 22nd Century Medical and Reserch Center, The University of Tokyo, Tokyo, Japan

Corresponding author: Dr. Junichi Ohya, Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo - 113-0033, Japan. E-mail: oyaj-ort@h.u-tokyo.ac.jp

Journal of Craniovertebral Junction and Spine 2014, 5:38

Abstract

Purpose: Among extraosseous abnormalities of the vertebral artery (VA) at the craniovertebral junction (CVJ), available evidence regarding "posterolateral protrusion," the VA running distant from the groove over the superior surface of the posterior arch of the atlas, is limited. The purpose of this study was to determine the optimal measurement to indicate posterolateral protrusion of the VA. **Materials and Methods:** Computed tomography angiography (CTA) images of 40 consecutive patients with cervical disease were reviewed. Ultimately, 66 arteries were included in this study. Five parameters predicted to indicate posterolateral protrusion of the VA were defined (A–E) and measured by two surgeons twice over a 2-week interval. Intraclass correlation coefficients (ICC) were used to examine intra-observer reproducibility and inter-observer reliability. Receiver operating characteristic (ROC) curve analysis was performed to determine the most optimal parameter to predict posterolateral protrusion of the VA. **Results:** Excellent inter-observer reliability and intra-observer reproducibility were obtained for all parameters (ICC = 0.87-0.99). Among them, parameter A, defined as the maximal length from the outer surface of the VA to the outer surface of the posterior arch of the atlas, was most accurately described posterolateral protrusion of the VA. The optimal cut-off value of parameter A obtained with ROC curves was 8.3 mm (sensitivity 97.5%, specificity 100%). **Conclusions:** The measurement in this study can quantitatively evaluate the posterolateral protrusion of the VA. Before posterior surgery at the CVJ, pre-operative CTA can help surgeons detect anomalous VA and reduce the risk of intra-operative VA injury.

Key words: Atlas, computed tomography angiography, posterolateral protrusion, vertebral artery

Access this article online	
Quick Response Code:	Website: www.jcvjs.com
	DOI: 10.4103/0974-8237.147077

INTRODUCTION

Among extraosseous abnormalities of the vertebral artery (VA) at the craniovertebral junction (CVJ), available evidence regarding "posterolateral protrusion," the VA running distant from the groove over the superior surface of the posterior arch of the atlas, is limited.[1] The purpose of this study was to determine the optimal measurement to indicate posterolateral protrusion of the VA. We modified previously reported parameters to have greater clinical

relevance and to quantitatively evaluate this condition. We also briefly present some cases of posterolateral protrusion of the VA.

MATERIALS AND METHODS

Data source

Computed tomography angiography (CTA) images of 40 consecutive patients with cervical disease that were collected in Yokohama Rosai Hospital between March 2006 and March 2013 were retrospectively reviewed. Four cases with injuries involving atlantoaxial lesions including fractures and traumatic atlantoaxial dislocation were excluded. Images from two patients were not adequate for assessment. Two VAs could not be measured due to their occlusion. Of the 80 arteries in 40 patients, 66 arteries were ultimately included in this study. The patients included 17 men and 17 women ranging from 15 to 79 years of age.

CTA imaging conditions

CTA was performed with a 64-slice computed tomography (CT) scanner (Aquillion; Toshiba Medical Systems, Tokyo, Japan). Imaging parameters were as follows: 0.5 mm slice thickness, 0.75 s/rotation, 120 kV, and 300 mA. Reconstruction was performed based on images with a slice thickness of 1.0 mm. Image scanning was acquired 15 s after intravenous injection of 53 ml non-ionic contrast medium at a rate of 4 ml/s. For measurement, reconstructed axial slices were created parallel to the line connecting the anterior and posterior arch of the atlas.

Definition of parameters

The five parameters predicted to reflect posterolateral protrusion of the VA were measured on reconstructed axial CTA images. These parameters were defined according to a modification of the parameters of Yamaguchi *et al.,*[1] as follows [Figure 1]: A) distance from the outer surface of the VA to the outer surface of the posterior arch of the atlas; B) distance from the midline of the atlas to the most protrusive part of the VA; C) distance from the midline to the intersection of the outer surface of the VA with the outer cortex of the posterior arch of the atlas; D) distance from the posterior tubercle of the posterior arch of the atlas to the intersection described for parameter C; E) distance from the posterior surface of the superior facet of the atlas to the posterior edge of the protrusion.

CTA measurements

Two observers measured the five parameters to evaluate inter-observer reliability. The second measurements were collected 2 weeks after the first, and the two sets were compared to evaluate intra-observer reproducibility. A spine surgeon diagnosed the posterolateral protrusion of the VA by reviewing CTA. Sensitivity and specificity were calculated for the accuracy of the five parameters. Receiver operator characteristic (ROC) analysis was performed for each group.

Statistical analysis

The intraclass correlation coefficient (ICC) was used to assess the inter-observer reliability and intra-observer reproducibility in evaluating the posterolateral protrusion. ICC less than 0.40 was defined as poor, 0.40-0.60 as fair, 0.60-0.74 as good, and 0.75-1.00 as excellent.[2] ROC curves and the corresponding area under the curve (AUC) were used to evaluate the performance of the prediction model using the CTA measurements. ROC curves plot the true-positive rate (sensitivity) vs. the false-positive rate (1-specificity) at a continuum of thresholds; participants were classified as having a posterolateral protrusion if their estimated probability of protrusion exceeded a particular threshold. Statistical Package for the Social Sciences (SPSS) version 20 (SPSS, Chicago, IL, USA) was used for all statistical analyses.

RESULTS

Measurements

Five parameters were measured in 66 arteries by two observers using CTA. The mean lengths of the five parameters are shown in Table 1.

Reliability and reproducibility of measurements

Intra-observer reproducibility between the two sets of measurements by the senior and junior observers and inter-observer reliability between measurements carried out by the two observers are shown in Table 2. Intra-observer reproducibility and inter-observer reliability for all parameters were excellent (ICC = 0.96-0.99 and ICC = 0.87-0.99, respectively).

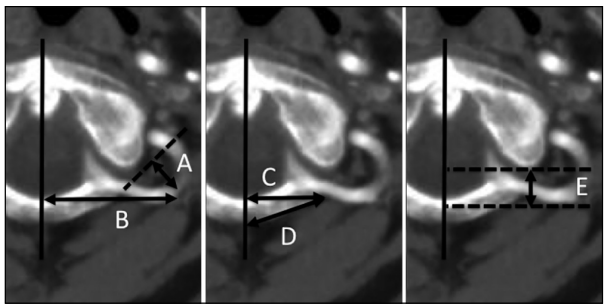


Figure 1: Reconstructed computed tomography angiography (CTA) at the atlas: Arrows in each panel represent the parameters. Five parameters (A, B, C, D, and E) are defined. See Methods for details on the definition of parameters

Table 1: Five parameters measured by the two observers in duplicate using computed tomography angiography (mm ± SD)

Parameters	Observer 1				Observer 2			
	Session 1		Session 2		Session 3		Session 4	
Parameter A	7.6±	3.3	7.5±	3.3	8.2±	3.2	8.3±	3.4
Parameter B	29.6±	3.2	29.5±	3.2	29.8±	3.2	29.5±	3.1
Parameter C	19.3±	3.0	19.4±	3.3	19.2±	3.0	19.0±	3.0
Parameter D	21.4±	3.8	21.5±	3.7	21.5±	3.5	21.2±	3.9
Parameter E	8.6±	2.2	8.6±	2.2	8.0±	2.2	8.1±	2.3

SD = standard deviation

Relationships between parameters

Strong correlations were found between parameters A and B, and between parameters C and D [r = 0.78 and 0.81, Table 3]. Moderate correlations were found between parameters B and C, parameters B and D, and parameters A and E [r = 0.42 to 0.49, Table 3]. The correlations between A and C, A and D, B and E, C and E, and D and E were weak [r = -0.30 to 0.29, Table 3].

ROC analysis

ROC curves illustrated the accuracy of the five parameters measured using CTA to predict posterolateral protrusion of the VA [Figure 2]. The AUC for parameter A was 0.998 (95% confidence interval [CI], 0.99-0.9998). AUC comparison revealed that parameter A significantly differed from the other parameters, with AUC values of 0.81 (95% CI, 0.69-0.89) for B, 0.63 (95%CI, 0.50-0.74) for C, 0.57 (95%CI, 0.44-0.69) for D, and 0.75 (95%CI, 0.62-0.84) for E. The cut-off value of parameter A obtained with ROC analysis was 8.3 (sensitivity 97.5%, specificity 100%). Cut-off values of the other parameters were 30.6 for B (sensitivity 77.5%, specificity 85.0%), 19.0 for C (sensitivity 60.0%, specificity 75.0%), 25.1 for D (sensitivity 17.5%, specificity 100%), and 8.2 for E (sensitivity 75.0%, specificity 72.5%).

CASE PRESENTATION

Case 1

An 81-year-old woman with incomplete spinal cord injury due to subaxial cervical spine fracture underwent fusion surgery. Pre-operative axial reconstruction CTA showed posterolateral protrusion of the bilateral VA [Figure 3a], which differed from the well-known imaging finding of bilateral VA contained in the groove of the posterior arch of the atlas [Figure 3b]. Parameter A measured using axial reconstruction CTA was 12.3 mm for the right VA and 11.9 mm for the left VA.

Table 2: Intra-observer reproducibility and inter-observer reliability: ICC values regarding posterolateral protrusion of the VA evaluated for each parameter (ICC, 95% CI)

ICC values	Intra-observer		Inter-observer	
Parameter A	0.99	(0.98-0.997)	0.99	(0.96-0.99)
Parameter B	0.98	(0.96-0.99)	0.89	(0.75-0.97)
Parameter C	0.96	(0.91-0.98)	0.87	(0.70-0.95)
Parameter D	0.96	(0.91-0.98)	0.87	(0.71-0.95)
Parameter E	0.97	(0.93-0.99)	0.87	(0.71-0.95)

ICC = intraclass correlation coefficient; CI = confidence interval

Case 2

A 65-year-old woman presented with progressive myelopathy secondary to atlantoaxial instability. Axial reconstruction CTA revealed bilateral VA running posterolaterally [Figure 4a]. The 3D reconstruction images from CTA also showed that the posterolateral portion of the bilateral VA deviated from the groove of the posterior arch of the atlas. Particularly, left dominant VA was recessed, caught between the posterior arch of the atlas and occipital bone [Figure 4b]. Pre-operative CTA revealed that the left VA, running slightly caudally, could suffer intra-operative injury during the procedure to insert a left C1 lateral mass screw. The results of measurement for parameter A were 11.0 mm for the right VA and 14.3 mm for the left VA.

Case 3

A 70-year-old man presented with myelopathy in association with atlantoaxial instability. Axial and 3D reconstruction CTA showed posterolateral protrusion of the right VA, whereas the left side was in the groove [Figure 5a and b]. Parameter A of the right VA was 9.4 mm, whereas on the left it was 2.7 mm. Atlantoaxial fusion surgery was performed successfully, with attention to VA injury on exposure of the posterior arch of the atlas, especially in the right lateral direction.

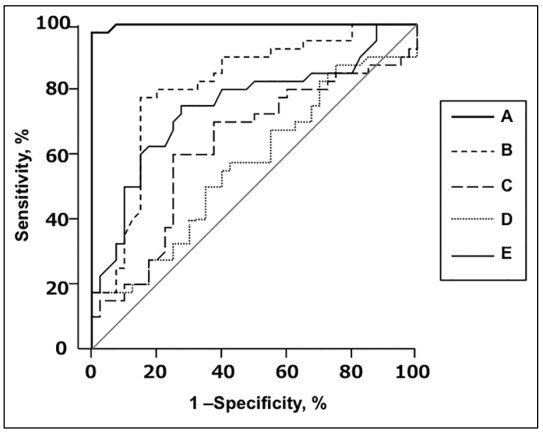


Figure 2: Receiver operator characteristic (ROC) curves demonstrated the relative ability of the five parameters to predict posterolateral protrusion of the vertebral artery (VA). Area under the curve (AUC) for parameter A was 0.998, whereas 0.81 for B, 0.63 for C, 0.57 for D, and 0.75 for E

Table 3: Matrix of correlations between each parameter: R values after the pearson test (r, 95% CI)

R values	Parameter B		Parameter C		Parameter D		Parameter E	
Parameter A	0.78	(0.67-0.85)	0.02	(–0.20-0.24)	0.29	(0.08-0.48)	0.48	(0.29-0.63)
Parameter B			0.42	(0.22-0.58)	0.49	(0.30-0.64)	0.18	(–0.04-0.38)
Parameter C					0.81	(0.71-0.87)	–0.30	(–0.49-0.08)
Parameter D							–0.24	(–0.44-0.02)

CI = confidence interval

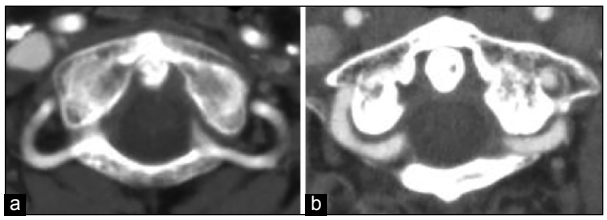


Figure 3: Axial computed tomography angiography (CTA) in Case 1 (a) showing posterolateral protrusion of bilateral vertebral artery (VA), which was differentiated compared to CTA in a different patient (b), whose bilateral VA was contained in the groove of the posterior arch of the atlas

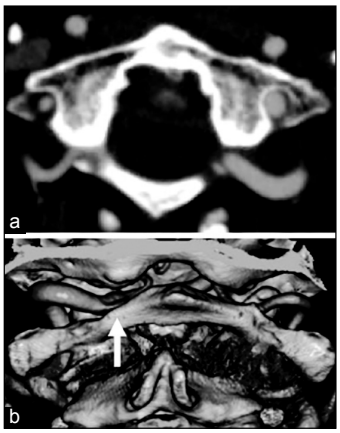


Figure 4: Axial computed tomography angiography (CTA) in Case 2 (a) showing posterolateral protrusion of bilateral vertebral artery (VA). 3D CTA images in Case 2 (b) showing left dominant VA recessed by being caught between the posterior arch of the atlas and occipital bone (white arrow)

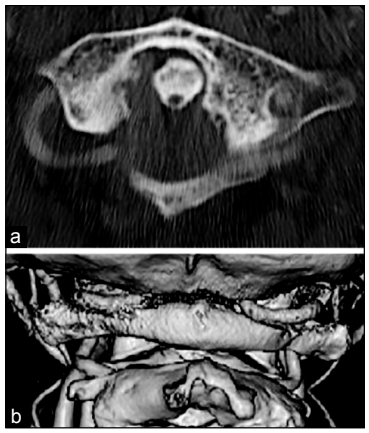


Figure 5: Axial and 3D computed tomography angiography (CTA) in Case 3 (a, b) showing posterolateral protrusion of the right vertebral artery (VA), whereas left VA was in the groove of the posterior arch of the atlas

DISCUSSION

This study is the first report regarding quantitative measurements to detect posterolateral protrusion of the VA using image examination. Although instrumentation surgery for cervical disorders has become widespread, these surgical procedures may carry the risk of VA injury. Intraoperative VA injury can cause severe complications such as cerebral infarction, massive bleeding, and even death.[3-7] Therefore, preoperative evaluations to determine the route of the VA at cervical vertebrae are of great significance to prevent VA injury and subsequent consequences. In this study, five parameters predicted to represent this condition were measured, and all parameters were assessed with excellent agreement with respect to inter-observer reliability and intra-observer reproducibility. ROC curve analysis revealed the predictive accuracy of the parameters, and the AUC for parameter A in particular was excellent and significantly greater than those of the other parameters. The findings of this study established a quantitative measurement strategy to identify posterolateral protrusion of the VA.

Extraosseous parts of the VA, as well as the intraosseous parts, are exposed to a high risk for VA injury. First, fusion surgery using screws at C1 or C2 requires wide exposure at the craniovertebral junction, which can cause VA injury to the extraosseous region during posterior exposure.[8] Also, a case of VA injury due to superior tap deviation during C1 lateral mass screw fixation has been reported.[9] Thus, surgeons should also exercise great caution against extraosseous VA injury when inserting such devices. Moreover, novel navigation system-based techniques for cervical instrument insertion cannot prevent injury to the extraosseous VA, although they can help surgeons insert instrumentation without injury to the intraosseous VA.[10,11] A recent large-scale study demonstrated the timing of VA injury during operative procedures and revealed that 20% of VA injuries occurred during surgical exposure.[12] In particular, posterior exposure was reported to be a relatively common situation of VA injury following posterior instrumentation of the upper cervical spine and anterior corpectomy. These findings suggest that surgeons should also pay attention to extraosseous abnormality of the VA, which carries the risk for VA injury during exposure, in addition to intraosseous abnormality such as high-riding VA, which should be a focus of attention during instrumentation. Because VA anomalies in the extraosseous region, including posterolateral protrusion, cannot be detected intra-operatively even with a navigation system, pre-operative evaluation of the VA course is important.

Although recent studies on extraosseous abnormality of the VA at the CVJ have been reported,[8,13,14] available evidence regarding posterolateral protrusion is limited. A few cadaver studies described lateral protrusion of the VA. Cassiola *et al.,* reported that the VA did not occupy the entire vertebral artery groove on the inferior surface of the superior articular facet and over the posterior arch of the atlas in their cadaver study.[15] Previous cadaveric and angiographic study demonstrated that the VA of elderly patients and the dominant VA had a tendency to be ectatic, bulging out from the C-1 groove, and therefore carried greater risk of injury during lateral exposure of the posterior skull base.[16] However, available studies on image findings quantitatively evaluating this phenomenon are limited. Yamaguchi *et al.,* termed a VA adopting a protrusive

course posterolaterally over the posterior arch of the atlas as posterolateral protrusion of the vertebral artery.[1] Although measurements for this condition in their study using CTA included five parameters, two of those parameters were similar, reflecting the distance from the midline to the intersection of the VA outer surface with the cortex of the posterior arch of the atlas. The only difference was whether outer or inner cortex of the posterior arch of the atlas was considered. However, the latter is not significant clinically, because exposure of the craniovertebral junction is performed through the outer side of the atlas. Therefore, we omitted this parameter and added two additional parameters, parameter D and E in this study, representing a view of the surgical field from posterior approach and the posterior component of protrusion. Finally, we identified five relatively more clinically important parameters.

The measurements of posterolateral protrusion were quantitatively evaluated in this study. The interobserver reliability and intraobserver reproducibility of quantitative posterolateral protrusion measurements in this study were excellent for all parameters, exceeding the ICC threshold of 0.75 indicating acceptability.[17] Previous studies have not evaluated the reliability and reproducibility of measurement parameters. Some parameters demonstrated strong associations upon analysis of correlation coefficients. Such strong correlations may result from measuring the same aspect of the VA protrusion. Parameter B measured the distance from midline to the posterolateral edge of the protrusion defined by parameter A. Additionally, parameters C and D measured the distance from a certain point to the intersection of the outer surface of the VA and outer cortex of the posterior arch of the atlas. These results suggested that measuring similar parameters that represent the same aspect appears unnecessary.

The accuracy of the five parameters to predict posterolateral protrusion of the VA on CTA was evaluated using ROC analysis. Parameters A, B, and E were above the threshold for acceptability (AUC >0.7),[18] whereas parameters C and D were not. Among the acceptable parameters, parameter A predicted posterolateral protrusion of the VA with the highest accuracy, and a cut-off value of 8.3 mm (sensitivity 97.5%, specificity 100%) was determined as the most useful to define this condition on CTA.

The measurement of posterolateral protrusion in this study using CTA is relatively easy because CTA has already become a routine tool before surgical treatment at the CVJ based on the superiority of CTA over MRI in terms of accurate depiction of the VA, surrounding osseous tissue, and their reciprocal anatomy, which aids in spatial analysis with unrestricted image reconstruction.[8,14] Surgeons should evaluate the directionality of the VA with preoperative imaging and be well versed in various VA anomalies at the CVJ.[8] When considering surgery at the CVJ in patients with Down syndrome, pre-operative CTA was reported to be of further importance for precisely identifying abnormal courses of the VA, which are more prevalent in patients with some congenital disease.[14] However, whether posterolateral protrusion of the VA is associated with a specific disease condition remains unknown.[1] We believe that the measurements in this study provide a basis for future research to examine the association between this phenomenon and disease conditions.

This study has several limitations. First, our sampling strategy may cause selection bias, because CTA was used to assess patients with cervical disease. However, such bias may be unavoidable considering that the inclusion of healthy participants in this study would have ethical problems due to the invasiveness of contrast radiography. Second, the two observers were able to identify the presence of this phenomenon in appearance while measuring the parameters using CTA, which can lead to diagnostic suspicion bias. Such bias could not be prevented in a quantitative evaluation using image examination such as the present study, because no blind method was available. Finally, observers measured the parameters on two-dimensional reconstructed axial slices. In some cases, the VA ran caudally in addition to posterolaterally distant from the groove of the atlas. For more accurate measurement of posterolateral protrusion of the VA, three-dimensional images may be effective.

In conclusion, the measurement in this study can evaluate one extraosseous abnormality of the VA, posterolateral protrusion, quantitatively. Before posterior surgery with instrumentation or wide exposure at the CVJ, preoperative CTA should be performed to detect this VA anomaly and reduce the risk of intraoperative VA injury.

REFERENCES

1. Yamaguchi S, Eguchi K, Kiura Y, Takeda M, Kurisu K. Posterolateral protrusion of the vertebral artery over the posterior arch of the atlas: Quantitative anatomical study using three-dimensional computed tomography angiography. J Neurosurg Spine 2008;9:167-74.
2. Fleiss JL. Reliability of measurement. In: Fleiss JL, editor. The Design and Analysis of Clinical Experiments. Toronto: Wiley; 1986. p. 1-32.
3. Burke JP, Gerstzen PC, Welch WC. Iatrogenic vertebral artery injury during anterior cervical spine surgery. Spine J 2005;5:508-14.
4. Neo M, Fujibayashi S, Miyata M, Takemoto M, Nakamura T. Vertebral artery injury during cervical spine surgery: A survey of more than 5600 operations. Spine (Phila Pa 1976) 2008;33:779-85.
5. Russo VM, Graziano F, Peris-Celda M, Russo A, Ulm AJ. The V (2) segment of the vertebral artery: Anatomical considerations and surgical implications. J Neurosurg Spine 2011;15:610-9.
6. Park HK, Jho HD. The management of vertebral artery injury in anterior cervical spine operation: A systematic review of published cases. Eur Spine J 2012;21:2475-85.
7. Wang M. Death due to extensive cervicomedullary infarction following iatrogenic vertebral artery occlusion. Forensic Sci Med Pathol 2012;8:334-7.
8. Yamazaki M, Okawa A, Furuya T, Sakuma T, Takahashi H, Kato K, *et al.* Anomalous vertebral arteries in the extra- and intraosseous regions of the craniovertebral junction visualized by 3-dimensional computed tomographic angiography: Analysis of 100 consecutive surgical cases and review of the literature. Spine (Phila Pa 1976) 2012;37:E1389-97.
9. Aota Y, Honda A, Uesugi M, Yamashita T, Baba N, Niwa T, *et al.* Vertebral artery injury in C-1 lateral mass screw fixation. Case illustration. J Neurosurg Spine 2006;5:554.
10. Ishikawa Y, Kanemura T, Yoshida G, Matsumoto A, Ito Z, Tauchi R, *et al.* Intraoperative, full-rotation, three-dimensional image (O-arm)-based

navigation system for cervical pedicle screw insertion. J Neurosurg Spine 2011;15:472-8.
11. Uehara M, Takahashi J, Hirabayashi H, Hashidate H, Ogihara N, Mukaiyama K, *et al*. Computer-assisted C1-C2 transarticular screw fixation "Magerl Technique" for Atlantoaxial instability. Asian Spine J 2012;6:168-77.
12. Lunardini DJ, Eskander MS, Even JL, Dunlap JT, Chen AF, Lee JY, *et al*. Vertebral artery injuries in cervical spine surgery. Spine J 2014;14:1520-5.
13. Hong JT, Lee SW, Son BC, Sung JH, Yang SH, Kim IS, *et al*. Analysis of anatomical variations of bone 1 and vascular structures around the posterior atlantal arch using three-dimensional computed tomography angiography. J Neurosurg Spine 2008;8:230-6.
14. Yamazaki M, Okawa A, Hashimoto M, Aiba A, Someya Y, Koda M. Abnormal course of the vertebral artery at the craniovertebral junction in patients with Down syndrome visualized by three-dimensional CT angiography. Neuroradiology 2008;50:485-90.
15. Cacciola F, Phalke U, Goel A. Vertebral artery in relationship to C1-C2 vertebrae: An anatomical study. Neurol India 2004;52:178-84.
16. Ulm AJ, Quiroqa M, Russo A, Russo VM, Graziano F, Velasquez A, *et al*. Normal anatomical variations of the V_3 segment of the vertebral artery: Surgical implications. J Neurosurg Spine 2010;13:451-60.
17. Portney LG, Watkins MP. Part II: Concepts of Measurement, 5. Reliability of Measurements. In Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. 3rd ed. Upper Saddle River. 2009.
18. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Hosmer DW, Lemeshow S, editors. Applied logistic regression, 2nd edn. New York: Wiley; 2000. p. 143–202

How to cite this article: Ohya J, Miyoshi K, Oka H, Matsudaira K, Fukushima M, Nagata K. Optimal measurement for "posterolateral protrusion" of the vertebral artery at the craniovertebral junction using computed tomography angiography. J Craniovert Jun Spine 2014;5:151-6.

Source of Support: Nil, **Conflict of Interest:** None declared.

Gait & Posture 40 (2014) 670–675

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Biomechanical analysis of low back load when sneezing

Tetsuya Hasegawa [a,*], Junji Katsuhira [b], Ko Matsudaira [c], Kazuyuki Iwakiri [d], Hitoshi Maruyama [a]

[a] Graduate School, International University of Health and Welfare, 2600-1 Kitakanemaru, Otawara, Tochigi 324-8501, Japan
[b] Department of Odawara Health Science, International University of Health and Welfare, 1-2-25 Shiroyama, Odawara, Kanagawa 250-8588, Japan
[c] Department of Medical research and Management for Musculoskeletal pain,22nd Century Medical and Research Center, The University of Tokyo,7-3-1,Hongo,Bunkyo-ku,Tokyo,113-8655, Japan
[d] National Institute of Occupational Safety and Health, Nagao 6-21-1, Tama-ku, Kawasaki, Kanagawa 214-8585, Japan

ARTICLE INFO

Article history:
Received 2 September 2013
Received in revised form 29 May 2014
Accepted 22 July 2014

Keywords:
Sneeze
Low back load
Low back pain
Intervertebral disk compressive force
Low back moment

ABSTRACT

Background: Although sneezing is known to induce low back pain, there is no objective data of the load generated when sneezing. Moreover, the approaches often recommended for reducing low back pain, such as leaning with both hands against a wall, are not supported by objective evidence.
Methods: Participants were 12 healthy young men (mean age 23.25 ± 1.54 years) with no history of spinal column pain or low back pain. Measurements were taken using a three-dimensional motion capture system and surface electromyograms in three experimental conditions: normal for sneezing, characterized by forward trunk inclination; stand, in which the body was deliberately maintained in an upright posture when sneezing; and table, in which the participants leaned with both hands on a table when sneezing. We analyzed and compared the intervertebral disk compressive force, low back moment, ground reaction force, trunk inclination angle, and co-contraction of the rectus abdominis and erector spinae muscles in the three conditions.
Findings: The intervertebral disk compressive force and ground reaction force were significantly lower in the stand and table conditions than in the normal condition. The co-contraction index value was significantly higher in the stand condition than in the normal and table conditions.
Interpretation: When sneezing, body posture in the stand or table condition can reduce load on the low back compared with body posture in the normal sneezing condition. Thus, placing both hands on a table or otherwise maintaining an upright body posture appears to be beneficial for reducing low back load when sneezing.

1. Introduction

Low back pain (LBP) is a common and major health problem, which can have sizeable socioeconomic impacts due to substantial direct and indirect social costs associated with LBP-related disability and loss of work [1,2]. In fact, most adults at some point in their lives experience some degree of LBP, of which approximately 85–90% of cases are classified as non-specific LBP [3,4]. In some instances, LBP is characterized as recurrent [5,6]. A recent report in Japan suggested that the lifetime prevalence of LBP was as high as 83% and the 4-week prevalence was 36%, making it the fifth-most common reason for medical consultation among outpatients [7].

Various factors can cause acute onset of non-specific LBP, including lifting and bending [8], and strategies for reducing low back load during such actions have been investigated from a biomechanical viewpoint using indicators for low back load such as the low back moment (LBM) and intervertebral disk compressive force (CF) in the lower back [9,10]. In clinical practice, sneezing is often reported to aggravate LBP. Indeed, Walker et al. reported sneezing to be an indicator of mechanical LBP [11], and Vroomen et al. [12] observed that 33% (40/122) of patients with LBP radiating in the leg but without radicular syndrome felt more pain on coughing, sneezing, or straining.

Sneezing occurs frequently as a respiratory reflex triggered to expel foreign bodies that mechanically irritate the nasal mucosa [13,14]. Characterized by explosive exhaling, sneezing is said to

* Corresponding author. Tel.: +81 0287 24 3000; fax: +81 0287 24 3100.
E-mail address: hasegawatetsuya1986@gmail.com (T. Hasegawa).

http://dx.doi.org/10.1016/j.gaitpost.2014.07.020
0966-6362/

cause strong concentric contraction of the rectus abdominis (RA) muscles and often sudden forward inclination of the trunk when in an upright posture. This forward inclination increases the lever arm from the center of rotation of the lower back to the center of mass in the upper body, thereby increasing the LBM. Moreover, since the forward trunk inclination angle (TA) is suddenly increased while sneezing, it is assumed that the acceleration applied to the center of gravity (COG) of the trunk also increases. This increase in acceleration entails a strong increase in the force that bends the trunk, so the erector spinae (ES) muscles must contract to maintain posture. Forward trunk inclination and ES contraction are reported to increase the CF [15], and therefore sneezing can be regarded as an action that increases low back load. However, no studies to date have reported objective measurement and biomechanical analysis of the low back load when sneezing.

Various types of media targeting people with LBP often recommend maintaining an upright posture or leaning with both hands on a table when sneezing to counter such pain [16]. These recommendations are made despite the lack of evidence for their efficacy. In this study, we conducted biomechanical tests to verify the hypothesis that maintaining an upright position or leaning with both hands on a table when sneezing reduces the low back load.

2. Methods

2.1. Subjects

Participants were 12 healthy young men (mean age, 23.25 SD 1.54 years; mean height, 170.30 SD 4.00 cm; mean weight, 60.90 SD 7.39 kg) with no history of LBP or spinal column pain. All provided written consent to participate after the study protocol was approved by institutional ethics committees.

2.2. Experimental conditions

Measurements were conducted under the following three conditions (Fig. 1): NORMAL condition for sneezing, characterized by forward trunk inclination; STAND condition, deliberately maintaining an upright posture of the trunk when sneezing; and TABLE condition, bending the trunk and leaning with both hands on a table when sneezing. Subjects stood on force plates and freely chose the distance between their feet and the position of their hands on the table. Subjects induced sneezing by irritating the nasal mucosa with a long, thin strip of tissue paper [17]. Measurements were taken 3 times under each experimental condition. In total, 9 trials with 1-min recovery intervals were conducted.

2.3. Experimental setup

Fig. 1 shows the measurement system used. Movement was recorded with a three-dimensional (3D) motion capture system (Vicon 612, Vicon, Oxford, UK) consisting of four force plates (AMTI, Watertown, MA) and 12 infrared cameras with a sampling rate of 120 Hz. Thirty-two infrared (IR)-reflective markers (diameter, 14 mm) were attached to each subject: top of the head, C7 spinous process, T10 spinous process, L5 spinous process, manubrium sterni, xiphoid process and bilaterally on the acromion process, lateral epicondyle, ulnar styloid process, anterior and posterior superior iliac spine, iliac crest, acetabulofemoral joint, medial knee joint, lateral knee joint, medial and lateral malleoli, and the first and fifth metacarpophalangeal joints. The obtained physical coordinates and ground reaction force (GRF) data were processed with a 6 Hz and 18 Hz second-order low-pass Butterworth filter (dual-pass for zero lag), respectively [18].

To measure muscle activity during movement, electromyograms were obtained (Biometrics, Newport, UK) at a sampling rate of 1000 Hz for the right RA (1 cm to the side of the umbilical region and 2 cm to the side of the medial line) [19] and the right ES (2 cm to the side between the L4 and 5 vertebrae) [20]. Electrodes were attached to only the right side because the left and right sides were expected to behave in a similar manner. Electromyography signals were prefiltered, producing a bandwidth of 20–460 Hz, and amplified with a differential amplifier (common-mode rejection ratio $>$ 96 dB at 60 Hz, input impedance $>$ 10 TΩ). Subjects wore a wristband connected to the grounding electrodes on the right hand. Subjects performed in the supine position against gravity with maximum resistance applied by the experimenter to obtain the maximum voluntary contraction of the RA (sit-up with straight leg while imposing resistance to the breast region) and in the prone position to obtain the maximum voluntary contraction of the ES (back extension with their hand resting on their head while imposing resistance to the scapular region) [21]. The subjects were required to produce maximal isometric extension efforts while resistance was provided by a single examiner with a physical therapy license.

Pressure sensors (DKH, Tokyo, Japan) were connected to the electromyographs and force plates to synchronize the

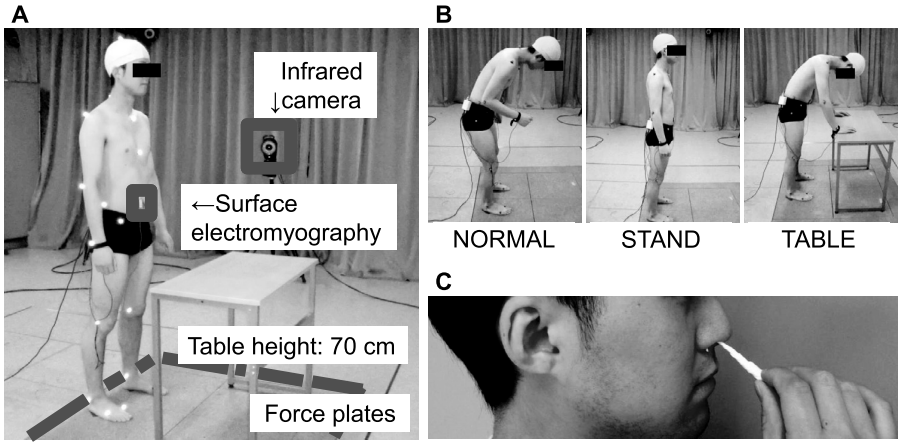


Fig. 1. (A) Experimental setup. (B) The three experimental sneezing conditions examined. In the NORMAL condition, subjects sneezed with no instructions. In the STAND condition, subjects were instructed to maintain an upright position as long as possible. In the TABLE condition, subjects were instructed to immediately place both hands on the table when they felt they would sneeze. (C) To promote sneezing, each volunteer irritated his nasal area using a roll made by twisting a sheet of tissue paper.

electromyograms and graphs obtained with the Vicon system. The observer input an analog electrical pulse as the synchronization marker to send to all systems to identify a common temporal reference point at the beginning of the measurement.

2.4. Data analysis

Data were analyzed using the 3D motion analysis software package Vicon Body Builder (Vicon). The method of Katsuhira et al. [9,10] was used to calculate the LBM. Briefly, the LBM was calculated using inverse dynamic analysis based on the Newton–Euler method from the GRF data obtained from the coordinates of the IR-reflective markers and force plates. In the analysis, segments were regarded as rigid and the joint moment was calculated using a link segment model in which segments were connected together at nodal points. To compute the joint moment, muscle coordinate data were added to the GRF data, in which the position of the center of mass, the weight portion, and the moment of inertia of each segment were used as parameters. The measurement data reported by Winter et al. [18], Okada et al. [23] and Jorgensen et al. [24] were used as the body parameters necessary for calculating the LBM. The method by Katsuhira et al. [25] was used to calculate the CF. Because LBP is often reported to occur between L4 and L5 [26], the L4–5 interspace was taken as the center of rotation for the LBM. The moment arm of the intervertebral disks and muscles was taken as the distance between the intervertebral disks and RA upon generation of the low back flexion moment and as the distance between the disks [23] and ES upon generation of the low back extension moment [24].

When calculating the CF in TABLE, the table was set to straddle the force plates, and the weight of the table was excluded from the calculations. In addition, the GRF readings obtained from the force plates on which the table was mounted were decomposed in accordance with the TA, and the reaction force obtained from the table was calculated by subtracting the result from the CF. Formula 1 refers to the CF [25]. Here, 20, 13, 8, and 23 are inverse numbers of the moment arms [23–25]. The low back joint compression force was obtained by multiplying the inverse number of the moment arms by the absolute value of the low back joint moments for each axis and adding the resolved gravitational force applied to the COG of the head, arm, and trunk (HAT) and the TA (θ):

$$\begin{aligned}&\text{Intervertebral disk CF}\\&+20|\text{Extension moment}|\text{or}13|\text{Flexion moment}|\\&+8|\text{Side flexion moment}|\\&+23|\text{Rotation moment}|\\&+\text{Gravitational force applied to COG of HAT}\cdot\cos\theta\\&-\text{Reaction force from table}\cdot\cos\theta\end{aligned} \quad (1)$$

The LBM and CF calculated with the above methods were taken as indicators of low back load. By taking the markers on both shoulders and manubrium sterni as indicators, the TA was measured as the change in angle when standing and the angle at peak CF when sneezing.

The co-contraction index (CCI) was calculated according to Falconer and Winter [27] to evaluate the co-contraction of the RA and ES when sneezing. Electromyographic data from these muscles were integrated over 1000 frames from a 1-s period (0.5 s before to 0.5 s after) of the peak CF recorded when sneezing. Using the obtained integral value, we calculated the portion corresponding to co-contraction of the muscles, which was taken as CCI. The computation of I_{ant}, which refers to the integral of the electromyogram of the antagonist muscle, shows the signal was stronger for the RA than for the ES between $t1$ and $t2$, and vice versa between $t2$ and $t3$, where t is timing. EMG_{AB} and EMG_{ES} indicate the activities of the RA and ES, respectively.

Consequently, the calculation was as follows:

$$I_{ant} = \int_{t1}^{t2} EMG_{AB}(t)dt + \int_{t2}^{t3} EMG_{ES}(t)dt \quad (2)$$

here I_{total} denotes the added integral values for these muscles, and EMG_{agon} and EMG_{ant} denote the electromyogram of the agonist and antagonist muscles, respectively. CCI was calculated from these values as follows:

$$I_{total} = \int_{t4}^{t3} [EMG_{agon} + EMG_{ant}](t)dt \quad (3)$$

$$CCI = \frac{2I_{ant}}{I_{total}} \times 100\% \quad (4)$$

Data for the CF, GRF, LBM, and TA were extracted at peak CF, and the CF, GRF, and LBM were normalized by body weight (mean of three measurements) to decrease individual differences.

2.5. Statistical analysis

Statistical analysis was performed using the mean values of the parameters for each participant and comparing the CF, LBM, GRF, TA, and CCI for the three experimental conditions. Verification was performed using repeated measures ANOVA, and variables showing a significant difference were subjected to multiple comparisons with Bonferroni correction. Significance was set at 5%. Intra-class correlation coefficients (ICC) of peak low back CF from the three trials were calculated for each condition. Statistical analysis was performed using SPSS 20 (SPSS Inc., Chicago, IL).

3. Results

3.1. Intervertebral disk compressive force and low back moment

The CF waveform in NORMAL shows two peaks, peak 1 and peak 2 (Fig. 2). The LBM waveform shows the flexion moment generated first, followed by the extension moment. Both the CF and LBM showed similar tendencies in all conditions.

Fig. 3 shows the mean CF for each condition. ICCs indicated moderate reliability in each condition. The force in STAND and TABLE was about half that in NORMAL. Table 1 shows the CF for peak 1, peak 2, and over a sneeze normalized by each subject's weight. Compared with NORMAL, these forces were significantly lower in STAND and TABLE ($p < 0.05$).

Table 1 shows the values for the LBM normalized by subject weight at peak CF. The force peaked when the low back extension moment was generated in NORMAL and TABLE and when the low back flexion moment was generated in STAND.

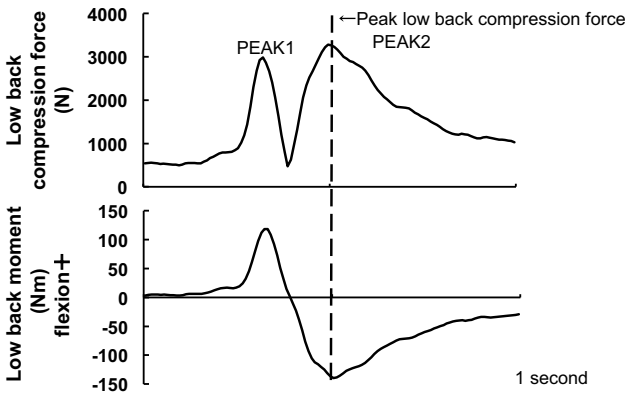


Fig. 2. Single data for low back compression force and joint moment when sneezing in the normal condition. Wave patterns 0.5 s before and after peak low back compression force are shown because this duration included the start and end of the sneeze in all subjects using a wave form of compressive force. The start and end of the sneeze was therefore defined as 0.5 s before and 0.5 s after peak intervertebral disk compressive force.

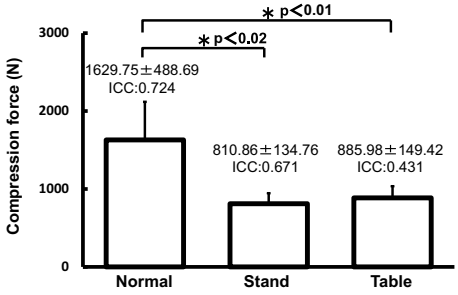


Fig. 3. Comparison of intervertebral disk compressive force of mean peak values and intra-class correlation coefficients (ICCs) in each experimental condition.

3.2. Vertical ground reaction force

Table 1 shows the GRF values normalized by subject weight at peak CF. The vertical component decreased significantly in STAND and TABLE compared with NORMAL ($p < 0.001$). No significant difference was seen between STAND and TABLE.

3.3. Change in trunk inclination angle

Table 1 shows the changes in the TA between standing posture and at peak CF. The positive direction is taken in the direction of flexion. The change was in the direction of flexion in NORMAL and TABLE, but in the direction of extension in STAND ($p < 0.001$). No significant difference was observed between NORMAL and TABLE.

3.4. Electromyograms and CCI

The electromyogram waveform for NORMAL indicates high activity for both the RA and ES when sneezing (Fig. 4). Furthermore, the CF and ES activity peaked at roughly the same time, and CCI was significantly higher in STAND than in NORMAL and TABLE ($p < 0.001$) (Table 1).

4. Discussion

4.1. Low back moment and intervertebral disk compressive force when sneezing

Two peaks were found in the plots of CF and LBM when sneezing, indicating that the RA is highly active during the characteristic forceful exhalation of sneezing. Such muscle activity induces flexion of the trunk, which activates the ES to maintain posture. Electromyograms also showed that since the activity of the RA peaked before that of the ES, the former is predominantly active during generation of the low back flexion moment, while the latter is more active during generation of the low back extension moment.

The mean CF when sneezing in NORMAL for a young man of approximately 60 kg is about 1600 N. This is roughly equivalent to holding a 20-kg load in a stationary upright position, which results in an estimated 3- or 4-fold increase in CF on the L4–5 intervertebral disks during static standing [15]. In other words, although sneezing is a momentary action, the load exerted on the intervertebral disks might aggravate or cause recurrent LBP.

Among the three experimental conditions, the CF and LBM were significantly lower in STAND and TABLE. The LBM was estimated from the GRF using the inverse dynamics method. This moment is influenced by the TA, and the GRF reflects the acceleration generated as a result of trunk movement. For this reason, the CF can probably be decreased by reducing these two parameters. There are a number of possible reasons for the significantly lower CF in STAND and TABLE. First, the change in TA was comparatively small in STAND, meaning that the moment arm of the center of mass of the upper body with respect to the intervertebral disks is small, so it can be considered to reduce the LBM. In the aforementioned study measuring the CF [15], the force increased with flexion of the trunk, a tendency similar to that observed in the present study. In addition, the vertical GRF was small compared with NORMAL. This might have resulted from deliberately maintaining an upright posture, where the acceleration of the trunk was suppressed by consciously stopping the trunk from moving.

Second, the CF peaked during generation of the flexion moment only in STAND. This force is considered to peak when the RA is active. Because this moment arm is about 1.5-fold longer than that for the ES [24], the tensile force exerted by the RA is smaller, which reduces the CF.

Third, in TABLE, the vertical GRF was reduced and the LBM was significantly reduced compared with NORMAL. No significant difference was seen in the magnitude of TA change. Furthermore, compared with NORMAL, the GRF acting on the feet as a result of leaning with both hands on the table was reduced, which suppressed movement of the trunk when sneezing.

Table 1
Comparison with the normal sneezing posture of mean peak values at peak intervertebral disk compressive force and standard deviations of each of the parameters measured in the standing upright posture and leaning with hands on a table posture. The waveform of intervertebral disk compressive force shows two peaks. PEAK1 and PEAK2 indicate the first and second peak of the compression force, respectively. Verification was performed with repeated measures ANOVA using the different sneeze conditions as factors.

	Normal	Stand	Table	*p*-value
Compression force (N/kg) (PEAK1)	16.37 SD 5.09	12.36 SD 1.99[a1]	8.88 SD 3.50[a2]	$p < 0.001$* a1: $p < 0.001$ a2: $p < 0.031$
Compression force (N/kg) (PEAK2)	26.09 SD 6.16	11.198 SD 2.65[a1]	9.37 SD 3.00[a2]	$p < 0.001$* a1: $p < 0.001$ a2: $p < 0.001$
Compression force (N/kg) (Over sneezing)	26.75 SD 6.44	13.24 SD 2.32[a1]	14.04 SD 1.50[a2]	$p < 0.001$* a1: $p < 0.001$ a2: $p < 0.001$
Moment (Nm/kg) (Extension+)	−0.90 SD 0.38	0.27 SD 0.21[b]	−0.45 SD 0.33	$p < 0.001$* b: $p < 0.001$
Ground reaction force (N/kg)	10.77 SD 0.55	9.67 SD 0.58[a1]	8.93 SD 0.86[a2]	$p < 0.001$* a1: $p < 0.001$ a2: $p < 0.001$
Co-contraction index (%)	31.99 SD 8.07	44.83 SD 8.15[b]	31.00 SD 7.71	$p < 0.001$* b: $p < 0.001$
Trunk angle (°) (Flexion+)	31.05 SD 12.24	−4.44 SD 6.25[a]	36.47 SD 6.29	$p < 0.001$* a: $p < 0.001$

* One-way analysis of variance: (a) significantly smaller than in the normal condition on multiple comparison ($p < 0.05$); (b) larger than the other two conditions on multiple comparison ($p < 0.05$).

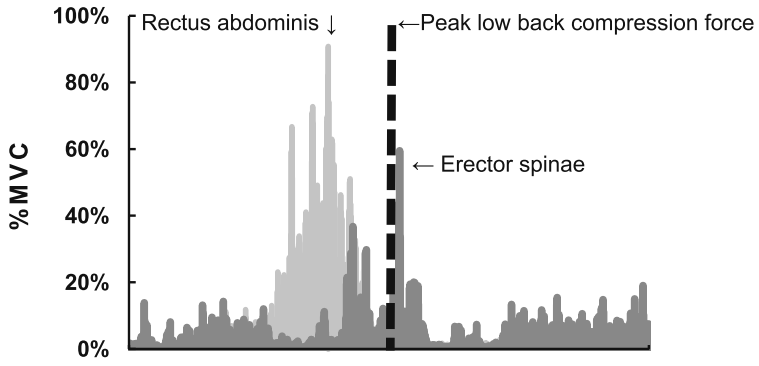


Fig. 4. Surface electromyography of the erector spinae and rectus abdominis muscles showing maximal voluntary contraction. During acquisition, we performed full-wave rectification using WAD analysis software (DKH, Tokyo, Japan) and a band-pass filter (20–420 Hz) to decrease noise according to the methods reported by Cholewicki et al. [22]. The obtained electromyograms were normalized using maximal voluntary contraction during isometric contraction. The wave pattern shows 0.5 s before and after the peak low back compression force.

4.2. Differences between conditions induced by muscle activity

Comparing the three conditions from the viewpoint of muscle activity, we found that CCI was significantly higher in STAND. In other words, there is greater co-contraction of the RA and ES in STAND. Co-contraction of antagonist muscles in the trunk increases the CF and stabilizes the upper trunk [28,29]. Arjmand et al. [29] suggested that co-contraction of antagonist muscles in the trunk is effective for stabilizing the upper trunk while lifting a heavy load. However, the low back load while sneezing was not as large as that when lifting a heavy load.

The present calculations transforming joint moment to muscle force did not separately clarify the magnitude of muscle force generated by the agonist and antagonist muscles. The CF might be greater in STAND than in NORMAL due to the CF generated by muscle co-contraction. Given that CCI was significantly higher in the STAND, greater CF seems to be generated by muscle co-contraction. Therefore, this force is more likely to be reduced in TABLE than in STAND. In other words, leaning with both hands on a table is more suitable for reducing the risk of low back load generated when sneezing than deliberately maintaining an upright posture.

This study has some limitations. First, the subjects were healthy young men, so it will be necessary to conduct a further study considering sex, age, and morphological differences and include subjects with LBP. Second, since the tensile forces exerted by the agonist and antagonist trunk muscles were not calculated separately, the CF generated by muscle co-contraction would not be entirely correct. Third, a previous study reported that high intra-abdominal pressure (IAP) might harm the lumbar tissues and cause LBP. Our biomechanical model accounted for the LBM including the effect of IAP but not the direct effect of IAP on low back load. Fourth, other factors, such as neck position and internal muscular or cardiovascular pressures, while sneezing should be examined.

Prevention measures for low back disability require continuous awareness of the fear avoidance (FA) model, because making the patient aware of posture while sneezing might cause an opposite effect to that desired [30]. Although we could not show the effect of FA in this study, future studies should consider the FA model when observing the effects of preventive measures for low back load while sneezing.

Conflicts of interest statement

None.

Acknowledgments

This study was supported by the Japan Labor Health and Welfare Organization's dissemination project Thirteen Fields of Occupational Injuries and Illness.

References

[1] Manchikanti L, Singh V, Datta S, Cohen SP, Hirsc JA. American society of interventional pain physicians. Comprehensive review of epidemiology, scope, and impact of spinal pain. Pain Physician 2009;12:E35–70.
[2] Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J 2008;8:8–20.
[3] Krismer M, van Tulder M. Low back pain (non-specific). Best Pract Res Clin Rheumatol 2007;21:77–91.
[4] Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? JAMA 1992;268:760–5.
[5] Carey TS, Garrett JM, Jackman A, Hadler N. Recurrence and care seeking after acute back pain: results of a long-term follow-up study. North Carolina Back Pain Project. Med Care 1999;37:157–64.
[6] Pengel L, Herbert R, Maher CG, Refshauge KM. Acute low back pain: a systematic review of its prognosis. BMJ 2003;327:323–7.
[7] Fujii T, Matsudaira K. Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. Eur Spine J 2013;22(2):432–8 [Epub 2012 August 7].
[8] Vingard E, Nachemson AL. Work-related influences on neck and low back pain. In: Nachemson AJ, Jonsson E, editors. Neck and back pain: the scientific evidence of causes, diagnosis and treatment. Philadelphia, PA: Lippincot Williams & Wilkins; 2000. p. 97–126.
[9] Katsuhira J, Sasaki H, Asahara S, Ikegami T, Ishihara H, Kikuchi T, et al. Comparison of low back joint movement using a dynamic 3D biomechanical model in different transferring tasks wearing low back belt. Gait Posture 2008;28:258–64.
[10] Katsuhira J, Tomita S, Haraguchi T, Harada S, Isikawa E, Kubo T, et al. Effect of use and type of assistive devices and posture while using them on the low back load in transferring Tasks. Jpn J Ergon 2010;46:157–65 [in Japanese].
[11] Walker BF, Williamson OD. Mechanical or inflammatory low back pain. What are the potential signs and symptoms? Man Ther 2009;14:314–20.
[12] Vroomen PCAJ, de Krom MCTFM, Wilmink JT, Kester ADM, Knottnerus JA. Diagnostic value of history and physical examination in patients suspected of lumbosacral nerve root compression. J Neurol Neurosurg Psychiatry 2002;72:630–4.
[13] Batsel HL, Lines AJ. Neural mechanism of sneeze. Am J Physiol 1975;229:770–6.
[14] Richardson PS, Peatfield AC. Reflexes concerned in the defence of the lungs. Bull Eur Physiopathol Respir 1981;17:979–1012.
[15] Wilke HJ, Neef P, Caimi M, Hoogland T, Claes LE. New in vivo measurements of pressures in the intervertebral disc in daily life. Spine 1999;24:755–62.
[16] Japan Labour Health and Welfare Organization website. Available from: http://www.research12.jp/22_kin/docs/manual.pdf [accessed 17.11.11] [in Japanese].
[17] Nishimura H, Sakata S, Kaga A. A new methodology for studying dynamics of aerosol particles in sneeze and cough using a digital high-vision, high-speed video system and vector analyses. PLOS ONE 2013;8(11):e80244.
[18] Winter DA. Biomechanics and motor control of human movement. 3rd ed. New York: John Wiley & Sons, Inc.; 2004.
[19] Ng JK, Kippers V, Richardson CA. Muscle fibre orientation of abdominal muscles and suggested surface EMG electrode positions. Electromyogr Clin Neurophysiol 1998;38:51–8.

[20] De Foa JL, Forrest W, Biedermann HJ. Muscle fibre direction of longissimus, iliocostalis and multifidus: landmark-derived reference line. J Anat 1989;163: 243–7.
[21] Hislop HJ, Montgomery J. Daniels and Worthingham's muscle testing: techniques of manual examination. 8th ed. Philadelphia: Saunders; 2007.
[22] Cholewicki J, McGill KC, Shah KR, Lee AS. The effects of a three-week use of lumbosacral orthoses on trunk muscle activity and on the muscular response to trunk perturbations. BMC Musculoskelet Disord 2010;11(1): 154.
[23] Okada H, Michiyoshi AE, Fujii N, Morioka Y. Body segment inertia properties of Japanese elderly. Biomechanisms 1996;13:125–39 [in Japanese].
[24] Jorgensen MJ, Marras WS, Granata KP, Wiand JW. MRI-derived moment-arms of the female and male spine loading muscles. Clin Biomech (Bristol Avon) 2001;16:182–93.
[25] Katsuhira J, Matsudaira K, Iwakiri K, Kimura V, Ohashi T, Ono R, et al. Effect of mental processing on low back load while lifting an object. Spine 2013;38:832–9.
[26] Chaffin DB, Anderson GBJ, Martin BJ. Occupational biomechanics. New York: John Wiley & Sons, Inc.; 1999.
[27] Falconer K, Winter DA. Quantitative assessment of co-contraction at the ankle joint in walking. Electromyogr Clin Neurophysiol 1985;25:135–49.
[28] Cholewicki J, Ivancic PC, Radebold A. Can increased intra-abdominal pressure in humans be decoupled from trunk muscle co-contraction during steady state isometric exertions? Eur J Appl Physiol 2002;87:127–33.
[29] Arjmand N, Shirazi-Adl A, Parnianpour M. Relative efficiency of abdominal muscles in spine stability. Comput Methods Biomech Biomed Eng 2008;11:291–9.
[30] Fujii T, Matsudaira K, Oka H. Factors associated with fear-avoidance beliefs about low back pain. J Orthop Sci 2013;18:909–15.

OPEN ACCESS Freely available online

PLOS ONE

Potential Risk Factors of Persistent Low Back Pain Developing from Mild Low Back Pain in Urban Japanese Workers

Ko Matsudaira[1*¤a], Hiroaki Konishi[2], Kota Miyoshi[3¤b], Tatsuya Isomura[4¤c], Kyoko Inuzuka[4]

1 Clinical Research Center for Occupational Musculoskeletal Disorders, Kanto Rosai Hospital, Kawasaki, Kanagawa, Japan, **2** Department of Orthopaedic Surgery, Nagasaki Rosai Hospital, Sasebo, Nagasaki, Japan, **3** Spine Center, Yokohama Rosai Hospital, Yokohama, Kanagawa, Japan, **4** Clinical Research Department, CLINICAL STUDY SUPPORT, Inc., Nagoya, Aichi, Japan

Abstract

Study Design: Two-year, prospective cohort data from the Japan epidemiological research of occupation-related back pain study in urban settings were used for this analysis.

Objective: To examine the association between aggravated low back pain and psychosocial factors among Japanese workers with mild low back pain.

Summary of Background Data: Although psychosocial factors are strongly indicated as yellow flags of low back pain (LBP) leading to disability, the association between aggravated LBP and psychosocial factors has not been well assessed in Japanese workers.

Methods: At baseline, 5,310 participants responded to a self-administered questionnaire including questions about individual characteristics, ergonomic work demands, and work-related psychosocial factors (response rate: 86.5%), with 3,811 respondents completing the 1-year follow-up questionnaire. The target outcome was aggravation of mild LBP into persistent LBP during the follow-up period. Incidence was calculated for the participants with mild LBP during the past year at baseline. Logistic regression was used to explore risk factors associated with persistent LBP.

Results: Of 1,675 participants who had mild LBP during the preceding year, 43 (2.6%) developed persistent LBP during the follow-up year. Multivariate analyses adjusted for individual factors and an ergonomic factor found statistically significant or almost significant associations of the following psychosocial factors with persistent LBP: interpersonal stress at work [adjusted odds ratio (OR): 1.96 and 95% confidence interval (95%CI): 1.00–3.82], job satisfaction (OR: 2.34, 95%CI: 1.21–4.54), depression (OR: 1.92, 95%CI: 1.00–3.69), somatic symptoms (OR: 2.78, 95%CI: 1.44–5.40), support from supervisors (OR: 2.01, 95%CI: 1.05–3.85), previous sick-leave due to LBP (OR: 1.94, 95%CI: 0.98–3.86) and family history of LBP with disability (OR: 1.98, 95%CI: 1.04–3.78).

Conclusions: Psychosocial factors are important risk factors for persistent LBP in urban Japanese workers. It may be necessary to take psychosocial factors into account, along with physical work demands, to reduce LBP related disability.

Citation: Matsudaira K, Konishi H, Miyoshi K, Isomura T, Inuzuka K (2014) Potential Risk Factors of Persistent Low Back Pain Developing from Mild Low Back Pain in Urban Japanese Workers. PLoS ONE 9(4): e93924. doi:10.1371/journal.pone.0093924

Editor: Laxmaiah Manchikanti, University of Louisville, United States of America

Received December 11, 2013; **Accepted** March 10, 2014; **Published** April 8, 2014

Funding: The study was a part of clinical research projects conducted by the Japan Labor Health and Welfare Organization. The research projects aimed to resolve occupational health issues and disseminate the research findings. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: TI is a founder of CLINICAL STUDY SUPPORT, Inc. KI is an employee of CLINICAL STUDY SUPPORT, Inc. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: kohart801@gmail.com

¤a Current address: Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan
¤b Current address: Department of Orthopaedic Surgery, Yokoyama Rosai Hospital, Yokohama, Kanagawa, Japan
¤c Current address: Clinical Research Department, CLINICAL STUDY SUPPORT, Inc., Nagoya, Aichi, Japan; Division of Clinical Research Consultation, Institute of Medical Science, Tokyo Medical University, Tokyo, Japan

Introduction

Low back pain (LBP) is a common musculoskeletal occupational health problem in industrialized countries and was found to be the leading specific cause of years lived with disability [1]. Japan is no exception, and LBP is one of the five most common health complaints of the Japanese general population [2]. Typically, 85–90% of the cases are classified as 'non-specific' [3,4], and the majority of LBP is mild, so they do not become severely disabled [5,6]. However, in terms of cost and work loss, the small proportion of people who become disabled due to LBP account for the largest occupational health care cost and the greatest number of work days lost around the world [7,8]. Therefore,

clarifying potential risk factors that could aggravate the LBP condition and lead to disability to work would be very important.

Many epidemiological studies of LBP have been conducted worldwide for decades. Psychosocial factors such as low job satisfaction, depression, or the tendency to somatize have been strongly indicated as 'yellow flags' for LBP leading to disability, as have ergonomic factors such as physical work demands [8–11], although the magnitude or intensity of each factor may vary across cultures or work environments [12]. Based on the above evidence, recently in Japan psychosocial factors began to be considered as a major risk for aggravating LBP. However, to our knowledge, the association between aggravation of Japanese workers' back pain and psychosocial factors has not been thoroughly assessed in prospective epidemiological research studies.

Previously, we reported potential risk factors for new onset of back pain disability in Japanese workers enrolled in a prospective cohort study in urban settings [13]. Data regarding various potential risk factors at baseline, as well as LBP-related outcomes, were collected prospectively. The cohort study focused mainly on LBP that caused work disability, a subject of critical importance to employers as well as workers, in terms of occupational health care.

The present study was designed to ascertain whether various psychosocial factors are associated with aggravating mild LBP into persistent LBP in workers with a 1-year history of mild LBP, using data from the previously reported cohort study; the findings of this further data analysis are reported here. This study was part of a series of clinical research projects conducted by the Japan Labor, Health and Welfare Organization related to 13 fields of occupational injuries and illnesses, including musculoskeletal disorders, mental health, and cancer. The research projects were conducted to help resolve occupational health issues and to disseminate the findings.

Materials and Methods

Data source

Data were extracted from a prospective cohort of the "The Japan epidemiological research of Occupation-related Back pain (JOB)" study. Participants were recruited from 16 workplaces in various occupational fields, located in or near Tokyo. The major occupational groups at these workplaces were office workers, nurses, sales/marketing personnel, and manufacturing engineers. Each participating organization was asked to distribute a self-administrated questionnaire to their workers, along with a cover letter from the study administration office. Respondents were asked to return their completed questionnaires by post, including their names and mailing addresses, which were used to send follow-up questionnaires directly from the study administration office. A total of 6,140 baseline questionnaires were distributed during September 2005 and February 2006, and 5,310 completed questionnaires were returned (response rate: 86.5%).

The baseline questionnaire included questions about the severity of the respondent's LBP and various individual and work-related factors. LBP severity was evaluated by the respondents themselves, who were asked to quantify the severity into one of four grades: grade 0, no LBP; grade 1, LBP not interfering with work; grade 2, LBP interfering with work; and grade 3, LBP interfering with work and leading to sick leave. The grades were determined with reference to Von Korff's grading method [14]. LBP was defined as pain localized between the costal margin and the inferior gluteal folds [3], and the area was depicted in the questionnaire. The baseline questionnaire included questions about the following: individual characteristics, including gender, age, obesity, smoking habits, history of LBP, and previous sick leave due to LBP; ergonomic work demands, such as frequency of bending, twisting or lifting at work; and psychosocial factors, such as depression, interpersonal stress at work, job control, job satisfaction, and somatization. A brief job stress questionnaire (BJSQ) was used to evaluate the major psychosocial factors [15,16]. The BJSQ is a self-administered scale having a total of 57 items, developed by a research working group organized by the Japan Labour, Health and Welfare Organization. Question items for the questionnaire were extracted from standard questionnaires commonly used for evaluating stress related factors, psychological stress response, depression, anxiety, and somatization [17–23]. The questionnaire was assessed using standardized scores, which were classified into 19 work-related stress factors: mental workload (quantitative aspect), mental workload (qualitative aspect), physical workload, interpersonal stress at work, environmental work stress, job control, utilization of skills and expertise, physical fitness, job satisfaction, vigor, irritability, fatigue, anxiety, depression, somatic symptoms, support from supervisors, support from co-workers, support from family or friends, and daily-life satisfaction. For each factor above, standardized scores were developed on a 5-point scale ranging from 1 (lowest) to 5 (highest) based on a sample of more than 10,000 Japanese workers. The questionnaire has demonstrated moderate reliability, high internal consistency, and its criterion validity has been assessed with respect to the Job Content Questionnaire (JCQ) and The National Institute for Occupational Safety and Health (NIOSH) [24].

The follow-up questionnaire was distributed 1 year after the baseline questionnaire was administered. Of the 5,310 participants who completed the baseline questionnaire, 3,811 successfully completed and returned the follow-up questionnaire, resulting in a follow-up rate of 71.8%. The follow-up questionnaire included questions relating to LBP, such as severity of LBP during the past year, length of sick-leave due to LBP, whether medical care was sought, pain duration, and onset pattern. LBP severity was assessed by the respondents themselves, using the same categories as those of the baseline questionnaire.

Ethical approval for the study was provided by the review board of the Japan Labour, Health and Welfare Organization. Informed consent was obtained in writing from all participants.

Data analysis

The outcome of interest was occurrence of persistent LBP during the 1-year follow-up period. In this study, persistent LBP was categorized as LBP interfering with work (grade 2 or grade 3), with disability lasting for longer than 3 months. Incidence was calculated for the participants who reported mild LBP (grade 1) during the past year at baseline. Participants were excluded from the analysis if they met any of the following criteria: a job change for reasons other than LBP; LBP due to a traffic accident; or LBP caused by a tumor, including metastasis, infection or fracture.

In addition to the compilation of simple, descriptive statistics, univariate and multivariate logistic regression analyses were used to explore risk factors associated with persistent LBP. Associations found by logistic regression analysis were summarized as odds ratios (ORs) with 95% confidence intervals (CIs). For the assessment of potential risk factors, crude ORs initially were estimated. Next, factors with P-values$<$0.1 were adjusted for individual factors, and also adjusted for individual factors and an ergonomic factor, in order to explore their potential risk factors. Factors with adjusted ORs that were statistically significant were considered to be potential risk factors. The following factors were used as adjusting factors because they are considered to be representative of individual and ergonomic factors: age, sex, obesity, smoking habits, education, and manual handling of

objects [25–27]. Additionally, the above psychosocial risk factors were grouped by their correlations to explore multicollinearity, and then a statistically significant factor that had the highest adjusted ORs were selected from each group and applied to multivariate regression analysis. Statistical significance was assumed at the 5% level if the 95% CI did not overlap 1. All statistical calculations were carried out using the STATA 9.0 software package.

Results

Baseline characteristics of study participants

Of the 3,811 participants who responded to the 1-year follow-up questionnaires, 1,675 (excluding 43 who did not answer the question on LBP severity on their follow-up questionnaire) reported mild LBP during the past year at baseline and met the selection criteria. The mean age was 43.1 years (SD 10.1 years) and 1,342 (78.6%) were male. The mean BMI was 23.1 kg/m^2 (SD 3.4 kg/m^2). Of these participants, 1,165 (68.2%) were categorized as non-manual laborers; 147 (8.6%) as manual handlers of $<$ 20-kg objects; 338 (19.8%) as manual handlers of $\geq$ 20-kg objects or as caregivers; and 58 (3.4%) were lacking job description data. In each category, the most common occupations were office work in the non-manual laborer category; manufacturing/engineering in the manual handler of $<$ 20-kg objects category; and nurse in the manual handler of $\geq$ 20-kg objects or caregiver category.

The baseline characteristics of the 3,811 participants who provided follow-up data appeared to be not much different from those who did not. The mean (SD) ages were 42.9 (10.1) years and 38.0 (10.2) years, respectively, and the majority were male in both groups (80.6% and 82.8%, respectively). Those who completed the study had a mean (SD) BMI of 23.1 (3.3) while the values for dropouts were 22.9 (4.1). In the follow-up group (vs. the drop-out group), 78.6% (vs. 75.5%) were categorized as manually handling $<$ 20-kg objects or not manually handling any objects in their work, 17.8% (vs. 18.9%) manually handled $\geq$ 20-kg objects or were working as caregivers, and data were lacking for 3.6% (vs. 5.6%). In both groups, the most common occupational fields in the categories of "manual handling of $<$ 20-kg objects or not manually handling any objects", and "manual handling of $\geq$ 20-kg objects or working as a caregiver" were office worker and nurse, respectively.

Incidence of persistent LBP

Of the 1,675 eligible participants, 43 (2.6%) reported persistent LBP within the 1-year follow-up period. Of the 43 participants reporting persistent LBP, 76.7% had pain that persisted for longer than 6 months.

Association between persistent LBP and potential risk factors

Crude ORs for persistent LBP, their 95% CIs, and P-values are shown in Table S1. The "somatic symptoms" risk factor was associated with an approximately 2.5-fold higher risk of suffering from persistent LBP. Associations of persistent LBP, with about a 2-fold risk increase, were also found with the following 5 psychosocial factors: interpersonal stress at work, job satisfaction, depression, support from supervisors, and daily-life satisfaction factors. An approximately 2-fold risk increase was found for the following 2 factors: previous sick-leave due to LBP and family history of LBP with work disability. Of the ergonomic factors, 7 (manual handling of objects at work, frequent bending, twisting, lifting, or pushing, hours of desk work, and physical workload) were associated with about a 3- to 4-fold higher risk of developing persistent LBP. These 15 factors were chosen for multivariate logistic regressions, and the results are shown in Table 1. Most of the ergonomic factors were significant with the ORs adjusted for individual factors. Five factors from the BJSQ (interpersonal stress at work, job satisfaction, depression, somatic symptoms, and support from supervisors), as well as previous sick-leave due to LBP and family history of LBP with disability, remained statistically significant or almost significant by adjusted ORs. The magnitudes of adjusted ORs of these factors did not markedly change from our crude OR analyses. Among the 5 factors from the BJSQ, interpersonal stress at work, job satisfaction, and support from supervisors tended to correlate to each other, and depression and somatic symptoms tended to correlate to each other (Spearman's rho, data not shown). Additional multivariate regression analysis included job satisfaction and somatic symptoms from the BJSQ psychosocial factors and family history of LBP with disability, chosen by the statistical significance of the adjusted OR. As shown in Table 2, all of the factors remained statistically significant or almost significant in the multivariate analysis.

Discussion

Potential risk factors for people with LBP that could aggravate the condition and cause too much disability to work were explored in a cohort of urban Japanese workers. The incidence of persistent LBP developing from mild LBP was 2.6%. ORs adjusted for individual factors and an ergonomic factor (manual handling of objects) showed that low job satisfaction, lack of support from supervisors, interpersonal stress at work, depression, somatic symptoms, and a family history of LBP with disability were significant risk factors, and previous sick leave a nearly significant risk factor, for development of persistent from mild LBP. Our results indicate that these psychosocial factors are important in urban Japanese workers who have made the transition from mild to persistent LBP.

In this study, the definition of persistent LBP was disability longer than 3 months, and the index for disability was LBP interfering with work, with or without sick leave. In Western countries, 'absence from work' is often used as an outcome measurement for disability. The number of participants who were absent due to LBP (grade 3) was relatively small. Our previous international epidemiological study showed that taking sick leave due to musculoskeletal disorders, mostly LBP, appears to be less common among Japanese workers than British workers [28]. The lower percentage of absence due to LBP in Japanese workers compared to workers in European countries may be due to a difference in concerns about being absent, such as worries that it might affect employment, salary increases, or evaluations of work performance. In fact, the proportion of Japanese workers with disability irrespective of taking sick leave (sick leave defined as any unplanned absence from work) was approximately the same as the proportion of UK workers with sickness-related absences. Additionally, in another international cross-sectional study, the prevalence of disabling LBP varied markedly across countries, and the Japanese workers showed the lower prevalence than in other countries [29]. Therefore, when assessing Japanese workers, it seems appropriate to define LBP disability as LBP interfering with work, with or without sick leave.

Among the five factors from the BJSQ (low job satisfaction, little support from supervisors, interpersonal stress at work, depression, and somatic symptoms), low job satisfaction, little support from supervisors, and interpersonal stress at work tend to relate to each other, and depression and somatic symptoms tend to relate to each

Table 1. Adjusted odds ratios of the baseline factors for persistent low back pain (LBP) with work disability; factors with crude odds ratio P values<0.1.

Factors		%	OR Adjusted for individual factors[a]		OR Adjusted for individual factors and an ergonomic factor[b]	
			OR	95%CI	OR	95%CI
Previous sick leave due to LBP	No previous sick leave	76.5	1.00		1.00	
	Previous sick leave	23.5	1.92	0.99–3.74	1.94	0.98–3.86
Manual handling of materials at work	Manual handling of < 20-kg objects including desk work	79.5	1.00			
	Manual handling of ≥ 20-kg objects or working as a caregiver	20.5	2.70	1.98–8.67	-	-
Bending[c]	Infrequent	88.7	1.00			
	Frequent	11.3	3.45	1.54–7.72	-	-
Twisting[c]	Infrequent	94.6	1.00			
	Frequent	5.4	4.35	1.80–10.52	-	-
Lifting[c]	Infrequent	89.6	1.00			
	Frequent	10.4	2.81	1.18–6.66	-	-
Pushing[c]	Infrequent	95.2	1.00			
	Frequent	4.8	3.48	1.24–9.76	-	-
Hours of desk work[d]	< 6 hours per day	53.9	1.00		1.00	
	≥ 6 hours per day	46.1	0.45	0.23–0.88	0.66	0.31–1.40
Physical workload[e]	No stress	61.9	1.00		1.00	
	Stress	38.1	2.22	1.16–4.23	1.53	0.70–3.33
Interpersonal stress at work[e]	No stress	78.8	1.00		1.00	
	Stress	21.2	2.04	1.06–3.93	1.96	1.00–3.82
Job satisfaction[e]	Satisfied	77.3	1.00		1.00	
	Not satisfied	22.7	2.48	1.31–4.70	2.34	1.21–4.54
Depression[e]	Not feeling depressed	64.6	1.00		1.00	
	Depressed	35.4	2.09	1.10–3.99	1.92	1.00–3.69
Somatic symptoms[e]	No somatic symptoms	63.4	1.00		1.00	
	Somatic symptoms	36.6	2.99	1.55–5.75	2.78	1.44–5.40
Support from supervisors[e]	Support	74.0	1.00		1.00	
	No support	26.0	1.97	1.04–3.73	2.01	1.05–3.85
Daily-life satisfaction[e]	Satisfied	68.7	1.00		1.00	
	Not satisfied	31.3	1.81	0.97–3.40	1.61	0.84–3.08
Family history of LBP with disability	No LBP with disability	74.6	1.00		1.00	
	LBP with disability	25.4	2.02	1.07–3.81	1.98	1.04–3.78

OR: odds ratio, CI: confidence interval, LBP: low back pain
[a]Adjusted for age, gender, obesity, smoking habits, and education.
[b]Adjusted for age, gender, obesity, smoking habits, education, and manual handling of materials at work.
[c]Bending, twisting, lifting, and pushing: ≥ half of the day was considered frequent.
[d]Hours of desk work: longer than 6 hours per day was considered to be static posture.
[e]Work-related stress factors assessed with the brief job stress questionnaire: not feeling stressed, feeling stressed: the 5 original responses were reclassified into "not feeling stressed", where low, slightly low and moderate were combined, and "feeling stressed", where slightly high and high were combined.
doi:10.1371/journal.pone.0093924.t001

other. The first three factors (e.g., low job satisfaction) could be considered stressful conditions that directly and negatively affect the individual, and the latter two factors (e.g., depression) as symptoms of both physical and mental stress. Generally, the symptoms of somatization are headaches, neck and shoulder discomfort, dizziness, palpitations or shortness of breath, diarrhea or constipation, and back pain, and these symptoms are triggered by emotional discomfort and psychosocial distress [30]. Individuals with somatization often complain of pain in various locations, functional disturbance of various organ systems, and are depressed or overwhelmed by these symptoms. Patients falling into such a situation are usually said to suffer from functional somatic syndrome (FSS) [31,32]. Our results could suggest that workers with mild LBP, under frazzled, depressed, or somatizing conditions, accompanied by emotional discomfort and psychosocial distress (e.g., low job satisfaction, little social support from

Table 2. Multivariate-adjusted odds ratios for the persistent low back pain (LBP).

Factors		Adjusted OR[a]	95%CI	P value
Job satisfaction	Satisfied	1.00		
	Not satisfied	2.03	1.01–4.07	0.046
Somatic symptoms	No somatic symptoms	1.00		
	Somatic symptoms	2.46	1.25–4.83	0.009
Family history of LBP with disability	No LBP with disability	1.00		
	LBP with disability	2.00	1.03–3.88	0.042

OR: odds ratio, CI: confidence interval, LBP: low back pain.
[a]Adjusted for individual factors (age, gender, obesity, smoking habits, and education) and an ergonomic factor (manual handling).
doi:10.1371/journal.pone.0093924.t002

supervisors, and interpersonal stress at work), did not manifest disabling back pain as a symptom of FSS at baseline, but the pain became disabling during the following year.

A family history of persistent LBP was also suggested as a psychosocial risk factor in this analysis. Second-hand experience of LBP among people with whom a worker is in very close contact (families, friends, or partners) may make it easier to imagine how mild LBP transforms to persistent LBP. Previous research has revealed that some people can share another person's physical pain experience, in both emotional and sensory components, by just observing the other person's pain [33,34]. Family members, therefore, may provide reinforcement for sick behavior [35], even though these family members do not have had any disorders, such as back pain [36–39].

Psychosocial intervention has been reported to improve overall well-being, as well as reducing distress and physical complaints, in patients with LBP in Western countries [40]. This intervention is based on the hypothesis that psychosocial factors are associated with the transition to persistent LBP, and should be examined in future research studies in Japan.

Limitations of the current study should be mentioned. One is the fact that the majority of the subjects were males, and that a broad range of Japanese occupations was not represented. The study cohort was not a representative sample of the entire Japanese workers in urban areas; therefore, the generalizability of the findings may be limited. Secondly, although cognitive and emotional aspects of back pain are known to influence disability aggravation, some important psychosocial factors, such as the attitudes of health care providers, and catastrophizing and fear-avoidance beliefs, were not included in this analysis. This was because appropriate questionnaires were not available in the Japanese language. Future studies should include additional self-reported outcome measures, such as results of the Fear-Avoidance Belief Questionnaire (FABQ) [41,42] or the Tampa Scale of Kinesiophobia (TSK) [43,44], to assess the impact of these factors in Japanese workers. The Japanese versions of these questionnaires are now being developed.

Psychosocial factors are one of the most important risk factors for making the transition to persistent LBP from mild LBP in urban Japanese workers. In the future, preventive strategies for reducing persistent LBP in the workplace should deal not only with physical work demands, which is already well-understood, but potentially should incorporate psychosocial management techniques as well.

Supporting Information

Table S1 Crude odds ratios of the baseline factors for persistent low back pain (LBP) with work disability. OR: odds ratio, CI: confidence interval, BMI: body mass index, LBP: low back pain. [a] Obesity: BMI of ≥ 25 is defined as obesity in Japan. [b] Smoking habits: Brinkmann index of ≥ 400 was defined as heavy smoker, calculated from the total number of cigarettes smoked per day multiplied by duration of smoking in years [45]. [c] Working hours: ≥ 60 hours per week was assumed to be uncontrolled overtime. [d] Bending, twisting, lifting, and pushing: ≥ half of the day was considered frequent. [e] Hours of desk work: longer than 6 hours per day was considered as static posture. [f] Work-related stress factors assessed with the brief job stress questionnaire: not feeling stressed, feeling stressed: the 5 original responses were reclassified into "not feeling stressed", where low, slightly low and moderate were combined, and "feeling stressed", where slightly high and high were combined. [g] Monotonous task: feelings of monotony or boredom at work.
(DOC)

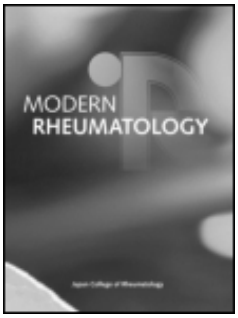
Author Contributions

Conceived and designed the experiments: K. Matsudaira HK K. Miyoshi. Performed the experiments: K. Matsudaira HK K. Miyoshi. Analyzed the data: K. Matsudaira TI KI. Wrote the paper: K. Matsudaira TI KI.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, et al. (2010) Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study. Lancet 380: 2163–2196.
2. Ministry of Health, Labour and Welfare (2010) Survey of Living Conditions. Available: http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/3-1.html. Accessed 2012 May 21.
3. Krismer M, van Tulder M (2007) Low back pain (non-specific). Best Practice & Research Clinical Rheumatology 21: 77–91.
4. Deyo RA, Rainville J, Kent DL (1992) What can the history and physical examination tell us about low back pain? JAMA 268: 760–765.
5. Schmidt CO, Raspe H, Pfingsten M, Hasenbring M, Basler HD, et al. (2007) Back pain in the German adult population: prevalence, severity, and sociodemographic correlates in a multiregional survey. Spine (Phila Pa 1976) 32: 2005–2011.
6. Walker BF, Muller R, Grant WD (2004) Low back pain in Australian adults: prevalence and associated disability. J Manipulative Physiol Ther 27:238–244.
7. Snook SH (2004) Work-related low back pain: secondary intervention. J Electromyogr Kinesiol 14: 153–160.
8. Maetzel A, Li L (2002) The economic burden of low back pain: a review of studies published between 1996 and 2001. Best Pract Res Clin Rheumatol 16: 23–30.
9. Waddell G, Burton AK (2001) Occupational health guidelines for the management of low back pain at work: evidence review. Occup Med 51: 124–135.

10. Papageorgiou AC, Croft PR, Thomas E, Ferry S, Jayson MI, et al. (1996) Influence of previous pain experience on the episode incident of low back pain: results the South Manchester Back Pain Study. Pain 66: 181–185.
11. Currie SR, Wang JL (2004) Chronic back pain and major depression in the general Canadian population. Pain 107: 54–60.
12. Waddell G (2004) Social interactions. In: G Waddell, ed. The Back Pain Revolution. 2nd ed. Edinburgh: Chuechill-Livingstone: 241–63.
13. Matsudaira K, Konishi H, Miyoshi K, Isomura T, Takeshita K, et al. (2012) Potential risk factors for new-onset of back pain disability in Japanese workers: findings from the Japan epidemiological research of occupation-related back pain (JOB) study. Spine (Phila Pa 1976) 37: 1324–1333.
14. Von Korff M, Ormel J, Keefe FJ, Dworkin SF (1992) Grading the severity of chronic pain. Pain 50: 133–149.
15. Muto S, Muto T, Seo A, Yoshida T, Taoda K, et al. (2006) Prevalence of and risk factors for low back pain among staffs in schools for physically and mentally handicapped children. Ind Health 44: 123–127.
16. Kawakami N, Kobayashi Y, Takao S, Tsutsumi A (2005) Effects of web-based supervisor training on supervisor support and psychological distress among workers: a randomized controlled trial. Prev Med 41: 471–478.
17. Kawakami N, Kobayashi F, Araki S, Haratani T, Furui H (1995) Assessment of job stress dimensions based on the job demands- control model of employees of telecommunication and electric power companies in Japan: reliability and validity of the Japanese version of the Job Content Questionnaire. Int J Behav Med 2: 358–375.
18. Haratani T, Kawakami N, Araki S (1993) Reliability and validity of the Japanese version of NIOSH Generic Job Questionnaire. Sangyo Igaku (Jpn J Ind Med) 35(suppl): S214 (in Japanese).
19. Yokoyama K, Araki S, Kawakami N, Takeshita T (1990) Production of the Japanese edition of profile of mood states (POMS): assessment of reliability and validity (in Japanese). Nippon Koshu Eisei Zasshi 37: 913–918.
20. Shima S, Shikano T, Kitamura T, Asai M (1985) New self-rating scales for depression. Clinical Psychiatry 27: 717–723 (in Japanese).
21. Spielberger CD, Gorsuch RL, lushene RE (1970) STAI Manual. Palo Alto: Consulting Psychologist Press.
22. Isaac M, Tacchini G, Janca A (1994) Screener for somatoform disorders (SSD). Geneva: World Health Organization.
23. Ono Y, Yoshimura K, Yamauchi K, Momose T, Mizushima H, et al. (1996) Psychological well-being and ill-being: WHO Subjective Well-being Inventory (SUBI) (in Japanese). Jpn J Stress Sci 10: 273–278.
24. Shimomitsu T, Odagiri Y (2004) The brief job stress questionnaire (in Japanese). Occup Mental Health 12: 25–36.
25. Linton SJ (2001) Occupational Psychological Factors Increase the Risk for Back Pain: a systematic review. J Occup Rehabil 11: 53–66.
26. Bernard BP (1997) Musculoskeletal disorders and workplace factors: a critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity, and low back. Cincinnati: U.S. Department of Health and Human Services. 590p.
27. Hoogendoorn WE, van Poppel MNM, Bongers PM, Koes BW, Bouter LM (2000) Systematic review of psychosocial factors at work and private life as risk factors for back pain. Spine 25: 2114–2125.
28. Matsudaira K, Palmer KT, Reading I, Hirai M, Yoshimura N, et al. (2011) Prevalence and correlates of regional pain and associated disability in Japanese workers. Occup Environ Med 68: 191–196.
29. Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, et al. (2013) Disabling musculoskeletal pain in working populations: is it the job, the person, or the culture? Pain 154: 856–863.
30. Kaplan C, Lipkin M Jr, Gordon GH (1988) Somatization in primary care: patients with unexplained and vexing medical complaints. J Gen Intern Med 3: 177–190.
31. Barsky AJ, Borus JF (1999) Functional somatic syndromes. Ann Intern Med 130: 910–921.
32. Henningsen P, Zipfel S, Herzog W (2007) Management of functional somatic syndromes. Lancet 369: 946–955.
33. Crook J, Milner R, Schultz IZ, Stringer B (2002) Determinants of occupational disability following a low back injury: a critical review of the literature. J Occup Rehabil 12: 277–295.
34. Shaw WS, Pransky G, Fitzgerald TE (2001) Early prognosis for low back disability: intervention strategies for health care providers. Disabil Rehabil 23: 815–828.
35. Osborn J, Derbyshire SW (2010) Pain sensation evoked by observing injury in others. Pain 148: 268–274.
36. Linton SJ (2000) Psychological risk factors for neck and back pain. In: Nachemson AJ, Jonsson E, editors. Neck and Back Pain: The scientific evidence of causes, diagnosis and treatment. Philadelphia: Lippincot Williams & Wilkins. pp. 57–78.
37. Ogino Y, Nemoto H, Inui K, Saito S, Kakigi R, et al. (2007) Inner experience of pain: imagination of pain while viewing images showing painful events forms subjective pain representation in human brain. Cereb Cortex 17:1139–1146.
38. Lynch AM, Kashikar-Zuck S, Goldschneider KR, Jones BA (2006) Psychosocial risks for disability in children with chronic back pain. J Pain 7: 244–251.
39. Evans S, Tsao JC, Lu Q, Myers C, Suresh J, et al. (2008) Parent-child pain relationships from a psychosocial perspective: a review of the literature. J Pain Manag 1: 237–246.
40. Williams RM, Westmorland MG, Lin CA, Schmuck G, Creen M (2007) Effectiveness of workplace rehabilitation interventions in the treatment of work-related low back pain: a systematic review. Disabil Rehabil 29: 607–624.
41. Westman AE, Boersma K, Leppert J, Linton SJ (2011) Fear-avoidance beliefs, catastrophizing, and distress: a longitudinal subgroup analysis on patients with musculoskeletal pain. Clin J Pain. 27: 567–577.
42. Waddell G, Newton M, Henderson I, Somerville D, Main CJ (1993) A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 52: 157–168.
43. Miller RP, Kori SH, Todd DD (1991) The Tampa Scale: a measure of kinisophobia. Clin J Pain 7: 51–52 (Data unpublished).
44. Kori KS, Miller RP, Todd DD (1990) Kinesiophobia: a new view of chronic pain behavior. Pain Manag 3: 35–43.
45. Brinkman GL, Coates O (1963) The effect of bronchitis, smoking and occupation on ventilation. Ann Rev Respir Dis 87: 684–693.



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey

Koji Yamada, Ko Matsudaira, Katsushi Takeshita, Hiroyuki Oka, Nobuhiro Hara & Yasuo Takagi

To cite this article: Koji Yamada, Ko Matsudaira, Katsushi Takeshita, Hiroyuki Oka, Nobuhiro Hara & Yasuo Takagi (2014) Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey, Modern Rheumatology, 24:2, 343-348, DOI: 10.3109/14397595.2013.854067

To link to this article: http://dx.doi.org/10.3109/14397595.2013.854067

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://informahealthcare.com/mor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2014; 24(2): 343–348

DOI: 10.3109/14397595.2013.854067

informa
healthcare

ORIGINAL ARTICLE

Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey

Koji Yamada[1], Ko Matsudaira[2], Katsushi Takeshita[1], Hiroyuki Oka[3], Nobuhiro Hara[1], and Yasuo Takagi[4]

[1]*Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan,* [2]*Clinical Research Center for Occupational Musculoskeletal Disorders, Kanto Rosai Hospital, Kanagawa, Japan,* [3]*Department of Joint Disease Research, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo, Japan, and* [4]*Graduate School of Health Management, Keio University, Kanagawa, Japan*

Abstract

Objectives. This study aimed to estimate the prevalence, magnitude, and direction of the associations among disability, pain intensity, number of pain sites, and health-related quality of life (HRQoL) in patients reporting low back pain (LBP) as their primary pain.
Methods. In January 2009, an Internet survey was performed for randomly selected adults aged 20–79 years who were registered as Internet research volunteers. Of 20 044 respondents, individuals with LBP as the primary pain were analyzed for associations among disability, number of pain sites, and HRQoL. Factors associated with low HRQoL were examined using multiple logistic regression modeling.
Results. Of the 20 044 respondents, 25.2 % ($n = 5060$) reported LBP and 13.5 % ($n = 2696$) reported LBP as their primary pain. Among those with LBP as the primary pain, HRQoL decreased with increase in disability and number of pain sites. In multivariate analyses, disability [adjusted odds ratio (aOR), 2.93–4.58], number of pain sites (aOR, 1.42–6.12), pain intensity ≥7 (aOR, 1.88), and age ≥60 years (aOR, 1.55) were associated with low HRQoL.
Conclusions. Approximately 13.5 % of patients reported LBP as their primary pain. Disability with absence from social activity and ≥7 pain sites were strongly associated with low HRQoL.

Keywords

Disability, EQ-5D, Low back pain, Multisite pain, Sick leave

History

Received 27 December 2012
Accepted 25 March 2013
Published online 10 April 2013

Introduction

Low back pain (LBP) is a common [1], costly [2], and, at times, disabling [3] condition that can lead to disability and sick leave from work or school. Pain at this site often fluctuates over time with frequent recurrences or exacerbations [4, 5]. The prevalence of LBP has been reported to range from 12–33 % [4] due to the methodologic heterogeneity across LBP prevalence studies [6, 7]. LBP is the most frequent and most expensive cause of work-related disability [8] and can affect health-related quality of life (HRQoL). LBP is a part of musculoskeletal pain [9, 10], but only one-sixth to one-third of individuals who suffer from LBP have LBP as their only pain source. Most LBP respondents also have pain at other sites [10]; this pain could be the primary reason for their disability. Moreover, a positive correlation was reported between the number of pain sites and functional problems in a large clinical study [9]. However, the prevalence and the impact of working disability and number of pain sites on HRQoL in those who have LBP as the primary pain have not been well examined.

Correspondence to: Koji Yamada, Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Tel: +81-3-3815-5411. Fax: +81-3-3818-4082. E-mail: forpatients2008@gmail.com

Therefore, the aim of this study was to estimate the prevalence, magnitude, and direction of the associations among disability, pain intensity, number of pain sites, and HRQoL in those reporting LBP as their primary pain in the pain-associated cross-sectional epidemiological (PACE) survey, which covers a large Japanese population.

Materials and methods

Subjects

The PACE survey was a cross-sectional Internet survey designed to evaluate the prevalence and characteristics of musculoskeletal pain in a large Japanese population. The study was performed over 10–18 January, 2009. Respondents were recruited at random from 1 477 585 research volunteers who were registered with an Internet survey company (Rakuten Research Inc., Tokyo, Japan), consistent with the Japanese demographic composition [11]. An invitation to participate in the research was sent through an e-mail containing a link to the survey. Double registration was prevented by checking the e-mail address and disabling the link to the questionnaire once the responder completed the survey. Forms were configured to automatically reject incomplete questionnaires. An additional credit point for Internet shopping was given as a financial incentive to the responders. On 18 January, 2009, the survey was closed when the number of respondents reached 20 063; thus,

the response rate is not relevant in this survey. Individuals whose reported age was <20 years or >79 years were excluded; thus, 20 044 participants were retained. This study was approved by Keio University's institutional review board.

Measures

The questionnaire included questions regarding musculoskeletal pain in the previous month and various individual factors. The respondents were asked about the characteristics of their musculoskeletal pain, such as the pain site(s), pain intensity at each site, site of the primary pain, duration of the primary pain, and disability due to the primary pain. Pain intensity was scored with a numeric rating scale (NRS) comprising 11 points (0 = no pain, 10 = worst pain imaginable). Disability was classified into three categories using a modified graded chronic pain scale (GCPS) [12], based on disability for social activity, such as work, school, and housework. Those with LBP and no disability were classified as modified GCPS grade 1, those with LBP and disability for social activity as modified GCPS grade 2, and those with LBP and disability leading to absence from social activity as modified GCPS grade 3. Respondents were asked about their demographic characteristics, including age, sex, occupational status, and HRQoL. HRQoL was measured using the Japanese EQ-5D instrument [13].

Definition of LBP

LBP was defined as pain experienced (over the previous month) below the costal margin and above the inferior gluteal folds, as described on the full-body manikin (Fig. 1, site 13), excluding those with pain around the anus (Fig. 1, site 21). Chronic LBP was defined as pain lasting ≥3 months.

EQ-5D

The EQ-5D instrument is a standardized general system for describing and valuing HRQoL [14]. It has good reliability and validity, and comprises five dimensions (mobility, self-care, usual activity, pain/discomfort, anxiety/depression) that are rated on three levels (1 = no problem, 2 = some problem, 3 = extreme problem); thus, it generates 243 theoretically possible health states (11111 = full health, 33333 = most extreme state).

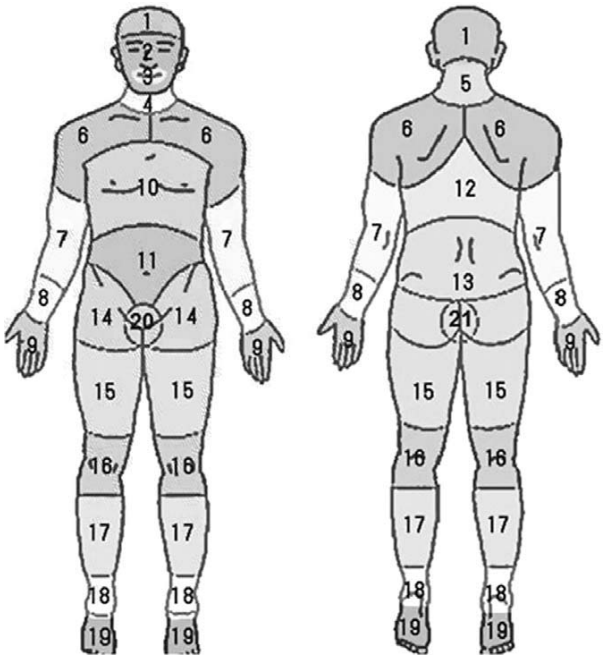


Fig. 1 The full-body manikin used in the pain-associated cross-sectional epidemiological (PACE) survey. Low back pain was defined as pain experienced below the costal margin and above the inferior gluteal folds, described as site number 13, excluding those with pain around the anus (site number 21)

Statistical analysis

First, the 1-month prevalence was calculated for those who had any LBP, LBP as the primary pain, and LBP as the only pain source (localized LBP). Further analyses were performed for those reporting LBP as their primary pain site using SPSS version 18 (IBM Corp., Armonk, N.Y., USA). Spearman's rho correlation coefficient was used to assess the correlations among HRQoL (EQ-5D score), disability, number of pain sites (other than LBP), and pain intensity (NRS score). For logistic regression analysis, the lowest 20 % of the EQ-5D scores in the total study population of the PACE survey was used as the dependent variable. A two-sided 5 % significance level was used in all statistical tests.

Results

LBP prevalence

Of the 20 044 respondents, 9746 (48.6 %) were men, and the overall mean score on the EQ-5D was 0.850 [standard error (SE), 0.001] with a ceiling effect of 45.7 % (9165 respondents; Table 1). The 1-month prevalence of LBP was 25.2 % (5060 respondents), of which only approximately half (2696 respondents; 13.5 % of all respondents) reported LBP as their primary pain and about one-seventh (706 respondents; 3.5 % of all respondents) reported LBP as their only pain source.

HRQoL in those with LBP as the primary pain

Further analyses were conducted for those with LBP as their primary pain. Of the 2696 respondents who reported LBP as the primary pain, 53.8 % (n = 1,424) were men, 78.1 % (n = 2,106) had chronic pain, 55.3 % (n = 1,491) reported LBP and no disability (modified GCPS grade 1), and 44.7 % (n = 1,205) reported disability for social activity with or without absence from social

Table 1. Characteristics of the total study population (n = 20,044)

Characteristic	n (%)
Age group (years)	
20–29	1,981 (9.9)
30–39	3,903 (19.5)
40–49	3,923 (19.6)
50–59	4,328 (21.6)
60–69	4,126 (20.6)
70–79	1,783 (8.9)
Mean ± SD	49.0 ± 14.2
Sex	
Male	9,746 (48.6)
Occupational status	
Worker	10,597 (52.9)
Housework and/or retired	7,655 (38.2)
Other (including student)	1,792 (8.9)
LBP prevalence	
Any LBP[a]	5,060 (25.2)
LBP as primary pain[b]	2,696 (13.5)
Localized LBP[c]	706 (3.5)
EQ5D score, mean ± SE	0.850 ± 0.001
Ceiling effect	9,165 (45.7)

LBP Low back pain, *SE* standard error
[a]Prevalence of respondents with LBP
[b]Prevalence of respondents with LBP as the primary pain site
[c]Prevalence of respondents with LBP as the only pain source

Table 2. Overall and by sex characteristics of respondents with LBP as the primary pain

Characteristic	Overall, *n* (%; *n* = 2696)	Men, *n* (%; *n* = 1424)	Women, *n* (%; *n* = 1272)
Age group (years)			
20–29	196 (7.3)	80 (5.6)	116 (9.1)
30–39	476 (17.7)	229 (16.1)	247 (19.4)
40–49	597 (22.1)	298 (20.9)	299 (23.5)
50–59	596 (22.1)	287 (20.2)	309 (24.3)
60–69	537 (19.9)	295 (20.7)	242 (19.0)
70–79	294 (10.9)	235 (16.5)	59 (4.6)
Mean ± SD	50.2 ± 13.8	52.4 ± 14.3	47.8 ± 12.9
Occupational status			
Worker	1,459 (54.1)	916 (64.3)	543 (42.7)
Housework and/or retired	1,013 (37.6)	395 (27.7)	618 (48.6)
Other (including student)	224 (8.3)	113 (7.9)	111 (8.7)
Duration of LBP			
Acute (<3 months)	526 (19.5)	281 (19.7)	245 (19.3)
Chronic (>3 months)	2,106 (78.1)	1,123 (78.9)	983 (77.3)
Unknown/refused to answer	64 (2.4)	20 (1.4)	44 (3.5)
Disability			
Grade 1[a]	1,491 (55.3)	808 (56.7)	683 (53.7)
Grade 2[b]	876 (32.5)	445 (31.3)	431 (33.9)
Grade 3[c]	329 (12.2)	171 (12.0)	158 (12.4)
NRS score (mean ± SE)	5.0 ± 0.0	4.8 ± 0.1	5.2 ± 0.1
Number of pain sites other than LBP (mean ± SE)	1.8 ± 0.0	1.6 ± 0.0	2.1 ± 0.1
EQ5D score (mean ± SE)	0.776 ± 0.003[d]	0.779 ± 0.004	0.772 ± 0.004

LBP Low back pain, *NRS* numeric rating scale, *SE* standard error
[a]LBP without disability for social activity, such as work, school, and housework
[b]LBP with disability for social activity, such as work, school, and housework
[c]LBP with disability leading to absence from social activity, such as work, school, and housework
[d]EQ5D score was significantly lower than that of the total study population (unpaired *t* test, $P < 0.01$)

activity (Table 2). The mean EQ-5D score was 0.776 (SE, 0.003), which was significantly lower than that of the total study population (unpaired *t* test, $P < 0.01$).

Next, the associations among HRQoL, number of pain sites, and pain intensity according to disability were analyzed (Table 3). We found that HRQoL decreased (Spearman's rank correlation coefficient, −0.371; $P < 0.01$) while pain intensity increased (Spearman's rank correlation coefficient, 0.418; $P < 0.01$) with higher disability. An increase in the number of pain sites was seen only between grade 1 and grade 2 disabilities (Table 3). Based on further evaluation of HRQoL stratified by age, sex, and disability, mean EQ-5D scores generally were lower in those with higher age and higher disability, and in women (Table 4).

Further analyses were conducted to evaluate the association among each variable stratified by the number of pain sites (Table 5). The number of respondents with LBP as a part of multisite pain was approximately 6.2 times larger than the number of those with localized LBP. In this analysis, HRQoL showed a negative correlation with the number of pain sites (Spearman's rank correlation coefficient, −0.256; $P < 0.01$). HRQoL was highest when the pain was localized, and lowest when the number of pain sites was ≥ 7. The proportion of those with disability for social activity (modified GCPS grades 2 and 3) and pain intensity also showed a positive correlation with the number of pain sites (Spearman's rank correlation coefficient, 0.184 and 0.359, respectively; both $P < 0.01$).

Factors associated with low HRQoL

In multivariate analyses adjusted by modified GCPS, number of pain sites, sex, age, and pain intensity, all the variables except sex were positively associated with low HRQoL (Table 6). The odds were higher as both disability and number of pain sites increased. Disability with absence from social activity and number of pain sites ≥ 7 had a strong relationship with low HRQoL. Similar trends were observed in both men and women; however, the impacts of absence from social activity and number of pain sites ≥ 7 were stronger in women than in men.

Discussion

In the present study, the 1-month prevalence of LBP was 25.2 % (5060 respondents), which is similar to that reported by Suzukamo and colleagues [15], who noted 30.6 % as the 1-month prevalence in Japan. Interestingly, of the 5060 respondents, only approximately half (2696 respondents; 13.5 % of all respondents) reported LBP as their primary pain, with the majority reporting chronicity. Recently, LBP has been recognized as a part of widespread musculoskeletal pain. Natvig and colleagues [10] reported that only

Table 3. Mean number of pain sites other than LBP, EQ5D score, and NRS score based on the disability of respondents with LBP as their primary pain

Disability (modified GCPS)	*n*	EQ5D score[d] (mean ± SE)	No. of pain sites other than LBP (mean ± SE)	NRS score[e] (mean ± SE)
Grade 1[a]	1,491	0.817 ± 0.003	1.5 ± 0.0	4.2 ± 0.0
Grade 2[b]	876	0.736 ± 0.004	2.3 ± 0.1	5.8 ± 0.1
Grade 3[c]	329	0.694 ± 0.009	2.3 ± 0.1	6.5 ± 0.1

GCPS Graded chronic pain scale, *LBP* low back pain, *NRS* numeric rating scale, *SE* standard error
[a]LBP without disability for social activity, such as work, school, and housework
[b]LBP with disability for social activity, such as work, school, and housework
[c]LBP with disability leading to absence from social activity, such as work, school, and housework
[d]EQ5D score showed a negative correlation with higher disability (Spearman's rank correlation coefficient, –0.371; $P < 0.01$)
[e]NRS score showed a positive correlation with higher disability (Spearman's rank correlation coefficient, 0.418; $P < 0.01$)

Table 4. Mean EQ5D score based on age, sex, and disability of respondents with LBP as the primary pain

Disability		Total (Grades 1 + 2 + 3)			Grade 1[a]			Grade 2[b]			Grade 3[c]		
Sex	Age (years)	*n*	Mean	SE	*n*	Mean	q	*n*	Mean	SE	*n*	Mean	SE
All	20–29	196	0.797	0.009	110	0.822	0.011	69	0.774	0.015	17	0.732	0.043
	30–39	476	0.785	0.006	236	0.828	0.008	173	0.756	0.009	67	0.706	0.021
	40–49	597	0.789	0.005	311	0.830	0.007	213	0.757	0.009	73	0.712	0.017
	50–59	596	0.777	0.006	360	0.817	0.006	172	0.727	0.010	64	0.686	0.021
	60–69	537	0.770	0.006	320	0.814	0.007	155	0.714	0.010	62	0.683	0.021
	70–79	294	0.729	0.008	154	0.782	0.009	94	0.676	0.010	46	0.659	0.026
	Total	2,696	0.776	0.003	1,491	0.817	0.003	876	0.736	0.004	329	0.694	0.009
Male	20–29	80	0.812	0.015	51	0.822	0.017	24	0.781	0.031	5	0.850	0.062
	30–39	229	0.794	0.009	114	0.837	0.011	80	0.772	0.013	35	0.702	0.033
	40–49	298	0.796	0.008	159	0.828	0.009	109	0.757	0.013	30	0.764	0.027
	50–59	287	0.781	0.008	172	0.820	0.009	81	0.725	0.014	34	0.718	0.024
	60–69	295	0.778	0.008	180	0.817	0.009	80	0.722	0.013	35	0.701	0.022
	70–79	235	0.734	0.008	132	0.781	0.009	71	0.666	0.011	32	0.689	0.034
	Total	1,424	0.779	0.004	808	0.817	0.004	445	0.734	0.006	171	0.718	0.013
Female	20–29	116	0.787	0.012	59	0.822	0.015	45	0.770	0.017	12	0.682	0.050
	30–39	247	0.777	0.008	122	0.820	0.011	93	0.743	0.012	32	0.710	0.024
	40–49	299	0.783	0.008	152	0.832	0.010	104	0.756	0.011	43	0.676	0.021
	50–59	309	0.773	0.008	188	0.814	0.008	91	0.730	0.014	30	0.650	0.034
	60–69	242	0.760	0.010	140	0.809	0.011	75	0.706	0.015	27	0.659	0.040
	70–79	59	0.708	0.020	22	0.787	0.034	23	0.704	0.020	14	0.590	0.035
	Total	1272	0.772	0.004	683	0.818	0.005	431	0.738	0.006	158	0.668	0.013

LBP Low back pain, *SE* standard error
[a]LBP without disability for social activity, such as work, school, and housework
[b]LBP with disability for social activity, such as work, school, and housework
[c]LBP with disability leading to absence from social activity, such as work, school, and housework

25 % of 893 participants who reported LBP during the previous week had localized LBP. In our study, the number of those with LBP as a part of multisite pain was about 6.2 times larger than the number of those with localized LBP. Previous studies [9, 10] have reported that many LBP respondents have pain elsewhere, which could be the primary reason for their disability. Therefore, we focused on LBP respondents reporting LBP as their primary pain for further analyses in this study.

In the present study, the mean EQ-5D score of those with LBP as their primary pain was 0.776 (SE, 0.003), which was significantly lower than that of the total study population [0.850 (SE, 0.001); $P < 0.01$], and slightly lower than the average score of patients with stage 5 chronic kidney disease (CKD) in Japan (0.798; 95 % CI, 0.757–0.839) [16]. Since stage 5 CKD represents established kidney failure, the similar HRQoL obtained in the present study indicates that the HRQoL of those who suffer from LBP could be as low as, or even lower than, those who are candidates for hemodialysis.

Table 5. Proportion of LBP with disability, and mean EQ5D and NRS scores based on number of pain sites other than LBP in respondents with LBP as the primary pain

Number of pain sites other than LBP	*n*	EQ5D score[a] (mean ± SE)	LBP with working disability[b] (%)	NRS score[c] (mean ± SE)
0	706	0.813 ± 0.005	35.7	4.1 ± 0.1
1–3	1,582	0.776 ± 0.003	44.0	5.1 ± 0.1
4–6	325	0.729 ± 0.007	59.7	6.1 ± 0.1
≥7	83	0.644 ± 0.014	75.9	7.1 ± 0.2
Total	2,696	0.776 ± 0.002	44.7	5.0 ± 0.0

LBP Low back pain, *NRS* numeric rating scale, *SE* standard error
[a]EQ5D score showed a negative correlation with the number of pain sites other than LBP (Spearman's rank correlation coefficient, −0.256; $P < 0.01$)
[b]Proportion of those with working disability (modified graded chronic pain scale grade 2 or 3 disability) showed a positive correlation with the number of pain sites other than LBP (Spearman's rank correlation coefficient, 0.184; $P < 0.01$)
[c]NRS score showed a positive correlation with the number of pain sites other than LBP (Spearman's rank correlation coefficient, 0.359; $P < 0.01$)

Generally, lower HRQoL is reported with higher disability in LBP patients [8, 17, 18]. Kovacs and colleagues revealed a negative correlation between the Rolland Morris Disability Questionnaire and the EQ-5D in LBP [8, 18]. In the present study, we used the GCPS [12], a well validated scale for assessing LBP disability, with minor revision. The revision was made to focus on disability and absence from social activity because the impacts of these disabilities on HRQoL have not been well examined. In our study, there was a negative correlation between disability and HRQoL, as in previous studies [8, 17, 18]. The differences in the mean EQ-5D scores between those with and those without disability and absence were 0.08 and 0.04, respectively. Interestingly, the differences were similar to the minimal clinically important difference reported in previous studies (0.033–0.074) [19, 20]. Collectively, these data suggest that the presence of disability for social activity and its severity regarding absence might have a significant meaning for those who suffer from LBP. Therefore, improvement of these disabilities might represent a clinically important difference, which needs further investigation.

In our study, HRQoL decreased as the number of pain sites increased, thus showing a negative correlation, whereas the proportion of disability and pain intensity increased as the pain sites increased. Kamaleri and colleagues [9] revealed that single-site pain did not have a large impact on physical fitness, feelings, or daily and social activities, and that functional problems increased markedly, in an almost linear manner, with increase in number of pain sites. From another study, the widest variation in health-related functioning, such as the items on the short form-36, was observed by the number of pain sites, with lower function seen with increase in number of pain sites [21]. LBP patients also have lower general health, poorer function, and poorer long-term work disability when their LBP is accompanied by multisite pain [10, 22, 23]. Our findings are consistent with those of previous reports, showing a similar relationship among pain intensity, disability, HRQoL, and number of pain sites in LBP responders. The reason why the majority of those with LBP as their primary pain also reported multisite pain could be the generalized hyperalgesia known to exist in

Table 6. Logistic regression analysis (dependent variable = lowest 20 % of EQ5D scores in total study population)

	Total[a]				Male[b]				Female[b]			
		95 % CI				95 % CI				95 % CI		
Variable	Adjusted odds	Lower	Upper	*P* value	Adjusted odds	Lower	Upper	*P* value	Adjusted odds	Lower	Upper	*P* value
Modified GCPS												
Grade 1	1.000				1.000				1.000			
Grade 2	2.930	2.393	3.589	<0.001	3.151	2.377	4.177	<0.001	2.750	2.052	3.686	<0.001
Grade 3	4.580	3.488	6.013	0.001	3.789	2.603	5.517	<0.001	5.642	3.780	8.420	<0.001
No. of pain sites other than LBP												
0	1.000				1.000				1.000			
1–3	1.420	1.128	1.786	0.003	1.173	0.873	1.576	0.290	1.850	1.275	2.685	0.001
4–6	2.367	1.733	3.232	<0.001	2.146	1.365	3.375	0.001	2.856	1.816	4.492	<0.001
≧7	6.124	3.541	10.589	<0.001	4.579	2.010	10.432	<0.001	8.426	3.970	17.882	<0.001
Sex												
F/M	1.044	0.868	1.256	0.644								
Age (years)												
<60	1.000				1.000					1.000		
≧60	1.545	1.271	1.879	<0.001	1.598	1.234	2.068	<0.001	1.485	1.097	2.011	0.010
NRS score												
<7	1.000				1.000					1.000		
≧7	1.883	1.541	2.300	<0.001	2.129	1.608	2.820	<0.001	1.650	1.238	2.200	0.001

CI Confidence interval, *F* female, *GCPS* graded chronic pain scale, *LBP* low back pain, *M* male, *NRS* numeric rating scale
[a]Multivariate analysis adjusted by modified GCPS, number of pain sites other than LBP, sex, age, and NRS score
[b]Multivariate analysis adjusted by modified GCPS, number of pain sites other than LBP, age, and NRS score

LBP patients [24]. Compared with healthy control subjects, LBP patients exhibit significantly lower pressure pain thresholds at all sites [25, 26]. The continuous nociceptive input might initiate central sensitization [27], which could develop widespread pain in those with LBP as their primary pain [24, 27].

In multivariate analyses, after adjusting for all the variables, modified GCPS grade, number of pain sites, age ≥60 years, and pain intensity were found to be associated with low HRQoL. Among these variables, disability with absence from social activity and ≥7 pain sites showed a stronger association than pain intensity (NRS score ≥7) and age ≥60 years. A similar tendency was seen in both men and women, highlighting the importance of multisite pain and disability in those who suffer from LBP. Although our study had limitations (due to its cross-sectional design), we believe the strong relationships seen in our study are noteworthy. Based on our results, occupational management [28, 29] focusing on returning to work, and management of multisite pain might have a more significant effect on HRQoL improvement than the management of pain itself in those who suffer from LBP. Further study is necessary to evaluate the effects of such management.

The strengths of our study include the large size of the population sample used to estimate the prevalence of those with LBP as their primary pain, and the magnitude of the associations among disability, pain intensity, number of pain sites, and HRQoL without any missing data. Some results support the validity of the PACE survey. First, the mean EQ-5D score of the PACE survey was similar to that found in a well-designed general population study (0.835) [30]. Second, the ceiling effect of the EQ-5D seen in the total study population also was similar to that reported in previous studies (42.5–47.0 %) [30–32]. Third, the percentage of those with LBP was similar to that reported previously in Japan [15]. Fourth, the percentage of workers in the total study population (52.8 %) was similar to that announced by the Japanese Ministry of Internal Affairs and Communications in 2009 (56.9 %) [11].

Some limitations in our study are notable, however. First, the selection bias due to the nature of an Internet survey needs to be addressed [33]. Although the study was conducted nationwide, using one of the largest domestic Internet survey companies, the volunteers from whom our sample was drawn were over-representative of people living in large cities, compared with the general population. Since LBP prevalence has geographic differences, with higher rates in urban populations than rural populations [34], caution is needed when interpreting the results of this study. Second, those who participate as Internet research volunteers may differ from the general population, and even from general Internet users. These potential differences could have affected the prevalence of LBP. Third, regarding the type of questionnaire, although a previous study reported that a Web-based questionnaire had adequate reliability compared with the paper-and-pencil version, even for older rural women [35], the mode of administration could affect the nature and rate of response [36]. Fourth, because this was a cross-sectional study, inferences cannot be drawn about causality.

In an Internet-based survey conducted in the United States, more than 27 000 individuals responded with a high response rate (75.7 %). The authors used a nationally representative Web-enabled panel of households that were recruited using a combination of random-digit dialing, landline-telephone recruiting, and address-based sampling [37]. Recruited households that did not have Internet access were provided free access via WebTV. Unlike other Internet-based surveys, the Internet-enabled panel used in the study was not limited to individuals with Internet access, and the sampling methodology was designed to ensure that the demographic characteristics of the panel were similar to those of the United States population. The methods used in this United States study maintain the representativeness of the study, while utilizing the advantages of Internet-based surveys for collecting a large amount of data. Such methodologic improvement might be necessary in our future studies.

Conclusion

Only approximately half of the LBP respondents reported LBP as their primary pain; among them, HRQoL decreased with higher disability and an increase in the number of pain sites. The presence of ≥7 pain sites and disability resulting in absence from social activity were strongly associated with low HRQoL. Occupational management focusing on return to work and management of multisite pain may have a more significant effect on HRQoL improvement than the management of pain itself in individuals with LBP.

Further research should focus on the effectiveness of such management in LBP respondents.

Acknowledgments

This study was supported by the dissemination project on the 13 fields of occupational injuries and illness of the Japan Labor Health and Welfare Organization. All investigations were conducted in conformity with ethical principles of research, and informed consent was obtained for participation in the study.

Conflict of interest

None.

References

1. Jacob T. Low back pain incident episodes: a community-based study. Spine J. 2006;6:306–10.
2. Maniadakis N, Gray A. The economic burden of back pain in the UK. Pain. 2000;84:95–103.
3. Derebery VJ, Giang GM, Saracino G, Fogarty WT. Evaluation of the impact of a low back pain educational intervention on physicians' practice patterns and patients' outcomes. J Occup Environ Med. 2002;44:977–84.
4. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J. 2006;15(Suppl 2): S192–300.
5. van Tulder M, Koes B, Bombardier C. Low back pain. Best Pract Res Clin Rheumatol. 2002;16:761–75.
6. Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976). 2008;33:95–103.
7. Leboeuf-Yde C, Lauritsen JM. The prevalence of low back pain in the literature. A structured review of 26 Nordic studies from 1954 to 1993. Spine (Phila Pa 1976). 1995;20:2112–8.
8. Kovacs FM, Abraira V, Zamora J, Teresa Gil del Real M, Llobera J, Fernandez C, et al. Correlation between pain, disability, and quality of life in patients with common low back pain. Spine (Phila Pa 1976). 2004;29:206–10.
9. Kamaleri Y, Natvig B, Ihlebaek CM, Bruusgaard D, et al. Localized or widespread musculoskeletal pain: does it matter? Pain. 2008; 138:41–6.
10. Natvig B, Bruusgaard D, Eriksen W. Localized low back pain and low back pain as part of widespread musculoskeletal pain: two different disorders? A cross-sectional population study. J Rehabil Med. 2001;33:21–5.
11. Japanese Ministry of Internal Affairs and Communications. Statistics Bureau, Director-General for Policy Planning (Statistical Standards) and Statistical Research and Training Institute. http://www.stat.go.jp/data/jinsui/2007np/index.htm. Accessed Mar 16, 2012; (in Japanese).
12. Van Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain. 1992;50:133–49.
13. Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ. 2002;11:341–53.
14. EuroQol–a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy. 1990;16:199–208.
15. Suzukamo Y, Takahashi N, Konno S, Kikuchi S, Fukuhara S. Outcome study of low back pain. Pharma Medica. 2007;25:9–12 (in Japanese).
16. Tajima R, Kondo M, Kai H, Saito C, Okada M, Takahashi H, et al. Measurement of health-related quality of life in patients with chronic kidney disease in Japan with EuroQol (EQ-5D). Clin Exp Nephrol. 2010;14:340–8.
17. Garratt AM, Klaber Moffett J, Farrin AJ. Responsiveness of generic and specific measures of health outcome in low back pain. Spine (Phila Pa 1976). 2001;26:71–7; (discussion 7).
18. Kovacs FM, Abraira V, Zamora J, Fernandez C; Spanish Back Pain Research Network. The transition from acute to subacute and chronic low back pain: a study based on determinants of quality of life and prediction of chronic disability. Spine (Phila Pa 1976). 2005;30:1786–92.
19. Burstrom K, Johannesson M, Rehnberg C. Deteriorating health status in Stockholm 1998–2002: results from repeated population surveys using the EQ-5D. Qual Life Res. 2007;16: 1547–53.
20. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res. 2005;14:1523–32.
21. Saastamoinen P, Leino-Arjas P, Laaksonen M, Martikainen P, Lahelma E. Pain and health related functioning among employees. J Epidemiol Community Health. 2006;60:793–8.
22. Natvig B, Eriksen W, Bruusgaard D. Low back pain as a predictor of long-term work disability. Scand J Public Health. 2002;30:288–92.
23. Leveille SG, Bean J, Ngo L, McMullen W, Guralnik JM. The pathway from musculoskeletal pain to mobility difficulty in older disabled women. Pain. 2007;128:69–77.
24. Kindler LL, Bennett RM, Jones KD. Central sensitivity syndromes: mounting pathophysiologic evidence to link fibromyalgia with other common chronic pain disorders. Pain Manag Nurs. 2011;12: 15–24.
25. O'Neill S, Manniche C, Graven-Nielsen T, Arendt-Nielsen L. Generalized deep-tissue hyperalgesia in patients with chronic low-back pain. Eur J Pain. 2007;11:415–20.
26. Laursen BS, Bajaj P, Olesen AS, Delmar C, Arendt-Nielsen L. Health related quality of life and quantitative pain measurement in females with chronic non-malignant pain. Eur J Pain. 2005;9:267–75.
27. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. Pain. 2011;152:S2–15.
28. Loisel P, Abenhaim L, Durand P, Esdaile JM, Suissa S, Gosselin L, et al. A population-based, randomized clinical trial on back pain management. Spine (Phila Pa 1976). 1997;22:2911–8.
29. Lemstra M, Olszynski WP. The effectiveness of standard care, early intervention, and occupational management in worker's compensation claims. Spine (Phila Pa 1976). 2003;28:299–304.
30. Saarni SI, Harkanen T, Sintonen H, Suvisaari J, Koskinen S, Aromaa A, et al. The impact of 29 chronic conditions on health-related quality of life: a general population survey in Finland using 15D and EQ-5D. Qual Life Res. 2006;15:1403–14.
31. Burstrom K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. Qual Life Res. 2001;10:621–35.
32. Nordlund A, Ekberg K, Kristenson M. EQ-5D in a general population survey–a description of the most commonly reported EQ-5D health states using the SF-36. Qual Life Res. 2005;14:1099–109.
33. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res. 2002;4:E13.
34. Volinn E. The epidemiology of low back pain in the rest of the world. A review of surveys in low- and middle-income countries. Spine (Phila Pa 1976). 1997;22:1747–54.
35. Boeckner LS, Pullen CH, Walker SN, Abbott GW, Block T. Use and reliability of the World Wide Web version of the Block Health Habits and History Questionnaire with older rural women. J Nutr Educ Behav. 2002;34(Suppl 1):S20–4.
36. Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. Science. 1998;280: 867–73.
37. Johannes CB, Le TK, Zhou X, Johnston JA, Dworkin RH. The prevalence of chronic pain in United States adults: results of an Internet-based survey. J Pain. 2010;11:1230–9.

J Bone Miner Metab (2014) 32:725–731
DOI 10.1007/s00774-013-0552-8

ORIGINAL ARTICLE

Vertebral fractures affect functional status in postmenopausal rheumatoid arthritis patients

Yasunori Omata · Futoshi Hagiwara · Jinju Nishino · Ko Matsudaira · Yuho Kadono · Takuo Juji · Toshihito Mori · Hisanori Nakayama · Yuichi Nagase · Jun Hirose · Tetsuro Yasui · Takumi Matsumoto · Toshihiro Matsui · Shigeto Tohma · Sakae Tanaka

Received: 31 August 2013 / Accepted: 14 November 2013 / Published online: 21 December 2013

Abstract Functional disability is a major concern in patients with rheumatoid arthritis (RA). This retrospective study investigated the risk factors for vertebral fractures (VFs) in postmenopausal RA patients and determined the impact of VFs on functional status. Data from a cohort of 200 postmenopausal RA patients in a single hospital registry were analyzed. Demographic and clinical data, imaging data from spine radiographs, and bone mineral density (BMD) data were collected from the patients at baseline and at the final visit (a mean of 2.9 years after the first visit). Risk factors for incident VFs and their impact on the modified health assessment questionnaire (mHAQ) were analyzed. Twenty-eight patients (14 %) developed new VFs (NVFs). Logistic regression analysis adjusted for age, BMI, and disease duration revealed that daily dose of prednisolone, femoral neck BMD, use of active vitamin D_3, and use of a bisphosphonate at baseline were factors associated with NVF, with odds ratios (95 % confidence interval) of 1.27 (1.05–1.54), 0.94 (0.91–0.97), 0.34 (0.13–0.89), and 0.31 (0.12–0.82), respectively. Patients with NVF exhibited worse mHAQ scores and a greater increase in mHAQ scores from baseline compared with those without NVF. In conclusion, incident VFs were associated with reduced functional status in postmenopausal patients with RA. It is important to prevent VFs to maintain the functional status of RA patients.

Keywords Osteoporosis · Prednisolone · Quality of life · Rheumatoid arthritis · Vertebral fracture

Y. Omata · Y. Kadono · Y. Nagase · J. Hirose · T. Yasui · T. Matsumoto · S. Tanaka (✉)
Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
e-mail: tanakas-ort@h.u-tokyo.ac.jp

F. Hagiwara · H. Nakayama · T. Matsui · S. Tohma
Department of Rheumatology, Sagamihara National Hospital, National Hospital Organization, Kanagawa, Japan

J. Nishino
Nishino Clinic Orthopedics and Rheumatology, Tokyo, Japan

K. Matsudaira
Clinical Research Center for Occupational Musculoskeletal Disorders, Kanto Rosai Hospital, Kanagawa, Japan

T. Juji · T. Mori
Department of Orthopaedic Surgery, Sagamihara National Hospital, National Hospital Organization, Kanagawa, Japan

Introduction

Rheumatoid arthritis (RA) is defined as a chronic systemic disease characterized by proliferative synovitis, and it results in severe joint destruction. The prevalence of RA in Japan is estimated at 1 % of the population, which is almost the same as in Western populations (0.5–1.0 %) [1, 2]. Inflammatory cytokines such as interleukin-1 (IL-1), IL-6, and tumor necrosis factor-α (TNF-α), as well as proteinases such as matrix metalloproteinases, are abundantly expressed in synovial tissues of RA patients and are involved in joint destruction [3]. In severe RA, persistent inflammation leads to generalized osteoporosis as well as local osteoporosis, which increases the risk of fragility fractures.

In RA, functional disability is a major concern because it not only restricts a patient's activities of daily living but also reduces work capability, resulting in a substantial economic burden [4–6]. There are several studies concerning risk factors associated with the functional status of RA patients [7–11]. However, there are few studies assessing the relationship

between fractures and functional status of RA patients. We investigated the influence of vertebral fractures (VFs) on the functional status of postmenopausal RA patients, and determined the risk factors for occurrence of VF.

Materials and methods

Patients

The subjects were selected from a single hospital registry of RA patients (National Hospital Organization Sagamihara Hospital). Two hundred postmenopausal RA patients who had complete records of demographic characteristics, clinical data such as disease activity score in 28 joints (DAS-28) and modified health assessment questionnaire (mHAQ), bone mineral density (BMD), and spine radiographs were eligible for this study. All patients fulfilled the American College of Rheumatology 1987 revised classification criteria for RA. The data were collected at baseline and at the final visit (a mean of 2.9 years after the first visit). All patients gave written informed consent, and the study was approved by the local ethics committee of Sagamihara Hospital.

Assessment of VFs in thoracic and lumbar spine

Lateral view radiographs of thoracic and lumbar spine were obtained from all the patients. VFs were diagnosed using the criteria of the Japanese Society for Bone and Mineral Research: the ratio of vertebral height at the mid and the anterior (or posterior) region was less than 0.8, and/or the ratio of the vertebral height at the anterior and posterior region was less than 0.75. New VFs (NVFs) were defined as follows: the vertebral height at the final assessment decreased by over 15 % or 4 mm compared with baseline. The films were independently investigated and assessed by three expert clinicians (H.F., J.T., and N.J.).

Information on the existence of back pain was also obtained from all patients. Clinical (symptomatic) VF was defined as VF with back pain, and asymptomatic VF as VF without back pain.

Measurements of bone mineral density

BMD measurements of the hip (femoral neck) and the lumbar spine (L1–4, anterior–posterior) were performed by trained technicians using dual-energy x-ray absorptiometry (QDR-4500, Hologic Co., Japan).

Data collection

Data were collected twice at a mean ± standard deviation (SD) interval of 2.9 ± 0.52 years. Baseline data were obtained from 2002 to 2005, and the final assessment was performed from 2005 to 2008. Clinical data included disease duration, age at menopause, length of time after the menopause, prescribed daily dose of prednisolone (PSL), use of bisphosphonates (etidronate disodium, alendronate or risedronate), calcium, and active vitamin D_3 (alfacalcidol or calcitriol). Rheumatology experts examined the tender joint count (28 joints), the swollen joint count (28 joints), and existence of joint replacement (total hip or knee joint replacement), and the patient's and investigator's global assessments of disease activity were measured on a 100-mm visual analog scale. DAS was computed using the 28-joint count. Self-reported physical disability was assessed by mHAQ score.

Statistical analysis

Demographic characteristics, clinical data, and BMD measurements were compared between NVF(−) and NVF(+) groups at baseline using 2-sided *t*-tests for continuous variables and chi-squared tests for categorical variables. The possible predictors of NVF were subsequently entered into a logistic regression analysis. The inclusion criteria for independent variables in the logistic regression analysis were those variables with a p value <0.10 in the univariate analysis to identify the variables that might be associated with NVF. Statistical analyses were performed using SPSS version 14.0 (SPSS Inc., Chicago, IL, USA). To assess which factor was statistically significant, the residuals which represented the difference between the observed value of the variables and the predicted value were analyzed.

Results

Baseline characteristics

Table 1 presents the demographic and clinical characteristics of 200 patients at baseline. The mean age of the patients was 61.5 ± 6.5 years, disease duration was 15.0 ± 10.1 years, and postmenopausal period was 12.1 ± 7.3 years. The proportion of rheumatoid factor-positive patients was 73.0 %. The classification and the stage of RA were as follows: Steinbrocker's classification, class 1, 60 patients (30.0 %); class 2, 117 patients (58.5 %); class 3, 23 patients (11.5 %), and class 4, no patient (0.0 %); Steinbrocker's stage 1, 37 patients (18.5 %); stage 2, 38 patients (19.0 %); stage 3, 37 patients (18.5 %), and stage 4, 88 patients (44 %). Mean mHAQ score was 0.55 ± 0.53 and mean DAS-28 score was 3.82 ± 1.08. Fifty-three patients (26.5 %) had low disease activity (DAS-28 ≤3.2), 124 patients (62.0 %) had

Table 1 Demographic and clinical variables of 200 postmenopausal rheumatoid arthritis patients

Baseline variables	All patients ($N = 200$)	NVF(−) ($N = 172$)	NVF(+) ($N = 28$)	p
Age (years)	61.5 ± 6.5	61.3 ± 6.6	62.4 ± 5.9	0.397
Height (cm)	151.8 ± 5.6	152.1 ± 5.5	150.2 ± 6.3	0.100
Weight (kg)	51.9 ± 8.1	51.7 ± 7.8	53.0 ± 10.0	0.429
BMI (kg/m^2)	22.5 ± 3.5	22.4 ± 3.4	23.5 ± 4.1	0.123
Disease duration (years)	15.0 ± 10.1	14.5 ± 9.9	18.0 ± 11.2	0.090
Postmenopausal duration (years)	12.1 ± 7.3	11.9 ± 7.3	13.4 ± 7.6	0.315
Rheumatoid factor positive (%)	73.0	71.5	82.1	0.240
Corticosteroid user (%)	70.5	68.0	85.7	0.057
Corticosteroid dose (mg/day)[a]	3.1 ± 2.5	2.9 ± 2.5	4.2 ± 2.5	0.011
Bisphosphonate user (%)	51.0	53.5	35.7	0.081
Calcium user (%)	60.0	62.8	42.9	0.059
Active vitamin D_3 user (%)	63.5	66.3	47.4	0.061
Modified HAQ score	0.55 ± 0.53	0.52 ± 0.53	0.73 ± 0.53	0.049
DAS-28[b]	3.82 ± 1.08	3.75 ± 1.09	4.23 ± 0.96	0.029
Global assessment score	3.5 ± 2.4	3.4 ± 2.4	4.2 ± 2.4	0.124
Tender joint count, 28 joints	2.6 ± 3.7	2.5 ± 3.4	3.5 ± 5.0	0.185
Swollen joint count, 28 joints	2.0 ± 2.5	1.8 ± 2.2	3.1 ± 3.8	0.091
CRP (mg/dl)	0.87 ± 1.14	0.83 ± 1.07	1.15 ± 1.48	0.165
ESR (mm/h)	39.6 ± 21.1	39.1 ± 20.3	42.8 ± 25.6	0.396
Serum BAP (U/l)	26.5 ± 10.5	26.2 ± 10.7	28.2 ± 8.7	0.350
Urine NTx (nmol BCE/mmol Cr)	49.0 ± 26.1	48.5 ± 26.7	52.1 ± 22.2	0.490
Joint replacement (%)[c]	16.0	14.5	25.0	0.166
Baseline vertebral fracture (%)	18.0	15.7	32.1	0.036
BMD at lumbar spine (g/cm^2)	0.787 ± 0.152	0.795 ± 0.147	0.737 ± 0.175	0.065
BMD at femoral neck (g/cm^2)	0.576 ± 0.137	0.583 ± 0.137	0.529 ± 0.129	0.050
PR at lumbar spine (%)	77.8 ± 15.1	78.6 ± 14.6	72.9 ± 17.3	0.066
PR at femoral neck (%)	73.1 ± 17.4	74.1 ± 17.4	67.3 ± 16.4	0.053
T-score at lumbar spine	−2.02 ± 1.38	−1.95 ± 1.33	−2.46 ± 1.58	0.070
T-score at femoral neck	−1.95 ± 1.23	−1.89 ± 1.23	−2.37 ± 1.18	0.055

Data of patients with ($N = 28$) and without ($N = 172$) new vertebral fractures (NVF) were compared

Data are expressed as mean ± standard deviation, *N* (%)

BMI body mass index, *HAQ* health assessment questionnaire, *DAS* disease activity score, *CRP* C reactive protein, *ESR* erythrocyte sedimentation rate, *BAP* bone alkaline phosphatase, *NTx* type I collagen cross-linked N-telopeptides, *BMD* bone mineral density, *PR* peak reference

[a] Doses are equivalent to prednisolone

[b] DAS computed using 28-joint counts

[c] Proportion of patients who underwent total knee or hip replacement surgery

moderate disease activity ($3.2 < \text{DAS-28} \leq 5.1$), and 23 patients (11.5 %) had high disease activity (DAS-28 >5.1). The number of patients who had joint replacement surgery in the lower extremities (total hip or knee joint replacement) was 32 (16.0 %).

Mean BMD was 0.787 ± 0.152 g/cm^2 at the lumbar spine, and 0.576 ± 0.137 g/cm^2 at the femoral neck. Osteoporosis, defined as a *T*-score ≤−2.5 SD, was identified at the lumbar spine in 40.5 % of patients and at the femoral neck in 32.5 %. A previous VF was detected in 36 patients (18.0 %) at baseline.

New vertebral fractures

Twenty-eight patients (14 %) had NVFs during the observation period, and the prevalence of VF at the final visit was 27.5 % (55 patients) (Table 1). Of the 28 patients with NVF, 17 patients had one fracture, and 11 patients had two or more fractures. In total, 47 of 3,400 vertebrae (1.4 %) were fractured, 16 (34.0 %) being thoracic fractures and 31 (66.0 %) lumbar fractures. The annual incidence of NVF was 4.8 per 100 patients. Thirteen of 28 (46.4 %) NVF patients had asymptomatic NVF, while 15

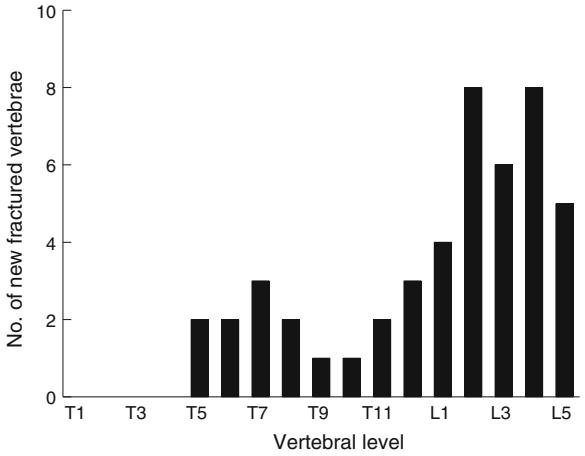


Fig. 1 Distribution of new vertebral fractures identified by lateral radiography in postmenopausal rheumatoid arthritis patients

patients (53.6 %) had clinical (symptomatic) NVF. Thirty-nine of 55 (70.9 %) VF patients at the final assessment had asymptomatic VF, while 16 patients (29.1 %) had clinical VF ($p < 0.001$). In patients treated with PSL ≥5 mg/day, the prevalence of asymptomatic VF was significantly higher than in those with symptomatic VF (30.7 and 8.0 %, respectively, $p < 0.001$). Fractures were most commonly identified in the mid-thoracic and the thoracolumbar regions, and the distribution of fractured vertebrae was similar to that reported in osteoporosis patients (Fig. 1).

The baseline characteristics were compared between NVF(−) and (+) groups (Table 1). There were no statistically significant differences in age, height, weight, BMI, disease duration, postmenopausal duration, rheumatoid factor positivity, proportion of PSL users, proportion of bisphosphonate users, calcium users, active vitamin D_3 (alphacalcidol) users, C-reactive protein level, erythrocyte sedimentation rate, serum bone alkaline phosphatase, urine type I collagen cross-linked N-telopeptides, and the proportion of patients who underwent joint replacement surgery. Significant differences were observed in the daily dose of PSL, mHAQ scores, DAS-28 scores, and prevalent VF. The mean daily doses of PSL were 2.9 ± 2.5 and 4.2 ± 2.5 mg/day, mHAQ scores were 0.52 ± 0.53 and 0.73 ± 0.53, and DAS-28 scores were 3.75 ± 1.09 and 4.23 ± 0.96 in the NVF(−) and (+) groups, respectively. The prevalence of VF at baseline was 15.7 and 32.1 %, respectively. Mean final DAS-28 score was 3.6 ± 1.1 and 4.2 ± 1.2 ($p < 0.05$), and final BMD at the femoral neck was 0.573 ± 0.124 and 0.500 ± 0.111 ($p < 0.01$) in the NVF(−) and (+) groups, respectively.

Logistic regression analysis adjusted for age, BMI, and disease duration using variables whose p-value was less than 0.10 by univariate analysis revealed that dose of PSL (mg/day), femoral neck BMD, use of vitamin D_3, and use of a bisphosphonate at baseline were factors associated with NVF, with odds ratios (95 % confidence interval) of 1.27 (1.05–1.54), 0.94 (0.91–0.97), 0.34 (0.13–0.89), and 0.31 (0.12–0.82), respectively (Table 2). Patients with higher DAS-28 score had a greater tendency to suffer NVF, with no statistical significance (odds ratio 1.45, $p = 0.075$).

Table 2 Logistic regression analysis of risk factors for new vertebral fractures

Variables	Regression coefficient	SE	OR	95 % CI	p
PSL (mg/day)	0.24	0.10	1.27	1.05–1.54*	0.013
DAS-28[a]	0.37	0.21	1.45	0.96–2.18	0.075
BMD, PR at femoral neck (%)	−0.06	0.02	0.94	0.91–0.97*	0.001
Active vitamin D_3 use	−1.07	0.49	0.34	0.13–0.89*	0.028
Bisphosphonate use	−1.18	0.50	0.31	0.12–0.82*	0.019

Multivariate adjusted odds ratio and 95 % confidence intervals and p values for the association are shown

Logistic regression (stepwise) analysis ($N = 200$). Statistical analysis was performed using the variables whose p value was <0.10 in the univariate analysis, and adjusted for age, BMI, and disease duration. * at 95 % CI means $p < 0.05$ in the logistic analysis

SE standard error, *OR* odds ratio, *CI* confidence interval, *PSL* prednisolone

[a] DAS computed using 28-joint counts

When the patients were subdivided into four groups based on the use of PSL and on DAS-28, the incidence of NVF was lowest in patients with low disease activity (DAS-28 <4.0) and receiving lower doses of PSL (<5 mg/day) ($p < 0.01$), while the incidence in patients with lower disease activity (DAS-28 <4.0) receiving higher doses of PSL (≥5 mg/day) was the highest of all the groups ($p < 0.05$). There was no significant difference in the incidence of NVF between patients with high disease activity (DAS-28 ≥4.0) receiving higher doses of PSL (≥5 mg/day) and those receiving lower doses of PSL (<5 mg/day) ($p = 0.426$) (Fig. 2).

Influence of NVFs on functional status

Mean mHAQ scores of NVF(−) and NVF(+) groups at the final assessment were 0.54 ± 0.56 and 0.88 ± 0.56, respectively, with a significant difference ($p = 0.003$). Patients with NVF at L3–5 exhibited significantly poorer final mHAQ scores than NVF(−) patients ($p < 0.05$) (Fig. 3a). Patients with NVF also exhibited a greater increase in mHAQ score from baseline compared with those without NVF ($p < 0.05$) (Fig. 3b). We also examined

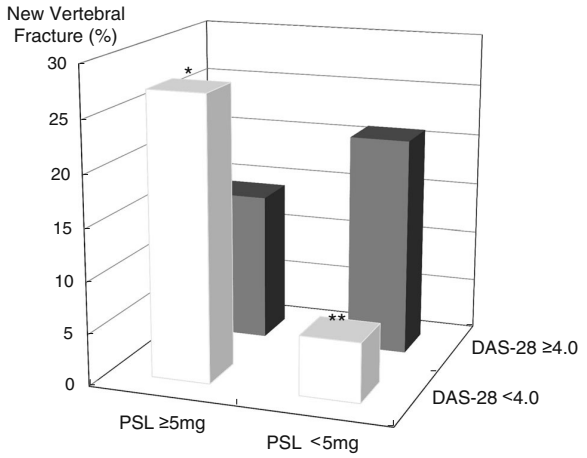


Fig. 2 Comparison of the incidence of new vertebral fractures between subgroups divided by the baseline dose of prednisolone (≥5.0 and <5.0 mg/day) and the DAS-28 score (≥4.0 and <4.0). * and **, significantly different between all groups (residual analysis); $*p < 0.05$, $**p < 0.01$

the influence of NVF in patients without baseline VF, and found that the mHAQ scores in these patients were significantly higher than in NVF(−) patients with no baseline VF (0.86 ± 0.63 and 0.48 ± 0.53, respectively, $p < 0.01$).

The mean mHAQ score in the group with NVF at the L3–5 level (without baseline VF) was significantly higher than in the group with no NVF (0.90 ± 0.69 and 0.48 ± 0.53, respectively, $p < 0.05$). The mean final mHAQ score in the clinical NVF(+) group was significantly higher than in the NVF(−) group (0.91 ± 0.69 and 0.54 ± 0.56, respectively, $p < 0.05$), while the mean mHAQ score (0.79 ± 0.42) in patients with asymptomatic NVF(+) was higher than in the NVF(−) group, with no significant difference (Fig. 3c).

Discussion

Several studies have reported that the risk of VF in RA patients is higher than that in primary osteoporosis patients. The prevalence of VF was reported as 15–36 %, and the annual incidence of NVF was around 4.0 per 100 patients [12–19]. Spector et al. [20] reported that the rate of VF in women with RA was over twice that of controls. Arai et al. [21] reported that the prevalence of VF in Japanese RA patients was 21 %, while that in age-matched healthy women was 5 %. We found that 14 % of the postmenopausal RA patients in our study developed NVF during 2.9 years (mean) of observation, and the prevalence of VFs was 27.5 % in total, which confirmed the high risk of VFs in RA patients, although we did not compare the prevalence with control patients.

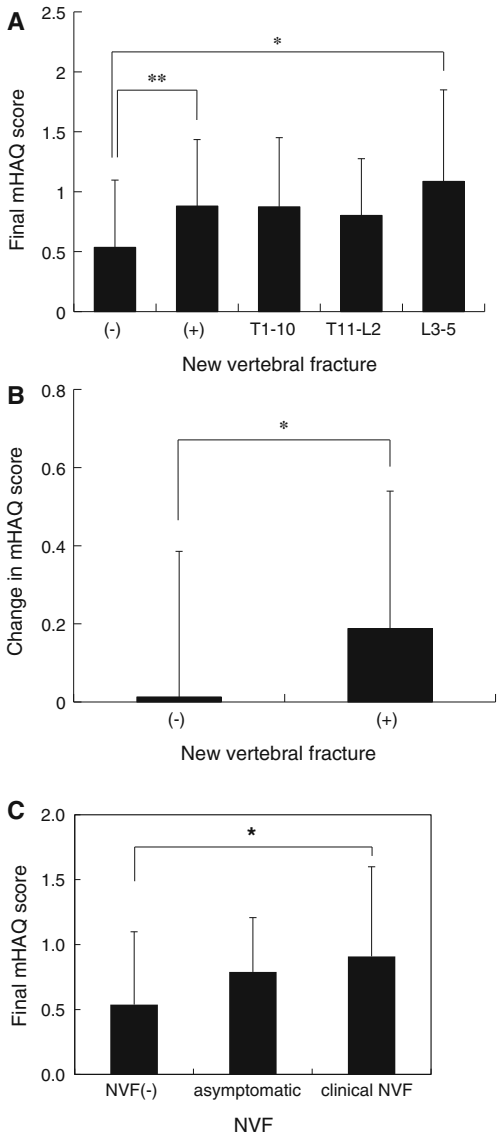


Fig. 3 Impact of vertebral fractures on mHAQ scores. **a** Effect of a new vertebral fracture on final mHAQ score. **b** Effect of a new vertebral fracture on the change in mHAQ score from baseline. **c** Effect of asymptomatic and clinical new vertebral fractures on mHAQ score. * and **, significantly different; $*p < 0.05$, $**p < 0.01$

Although a statistically significant difference was not found for disease activity in the logistic regression analysis, there was a tendency for DAS-28 to be higher in the NVF(+) group, and this also affected the incidence of NVF associated with BMD (existence of osteoporosis). El Maghraoui et al. [13] also reported that the prevalence of VF was increased in patients with higher DAS score. Dirven et al. [12] reported that disease activity over time was higher in patients with VFs. This may be because high levels of proinflammatory cytokines such as TNF-α and IL-6 in RA patients with high

disease activity may lead to deleterious bone quality, resulting in greater susceptibility to VF.

In contrast to previous studies, age was not significantly associated with NVF incidence. When patients were categorized into three age groups ($\leq$59, 60–69, and $\geq$70 years old), the percentage of bisphosphonate users was highest in the $\geq$70-year-old group (46.9, 56.8, and 70.8 %, respectively), which may affect the incidence of NVF.

Although PSL use is a risk factor for NVF, the occurrence of NVF in patients with high disease activity (DAS-28 $\geq$4.0) was lower in PSL $\geq$5 mg treatment group than in PSL <5 mg treatment group, although there was no significant difference (Fig. 2). We speculated that, in spite of its deleterious effects on bone, PSL use has some beneficial effects on preventing VFs in these patients by controlling disease activity. Consistent with our results, Ghazi et al. reported that the prevalence of VF was inversely related to the use of steroids, and Dirven et al. reported that there was no association between VFs and steroid use [12, 17, 21–23].

The efficacy of drug treatment against VF has been well established in osteoporosis patients [24–29]. Although we did not find a significant difference in the occurrence of NVF between the groups with or without bisphosphonate treatment in our study (9.8 and 18.4 %, respectively, $p = 0.08$), there was a tendency for patients using bisphosphonates, in particular alendronate and risedronate, to suffer NVF less frequently than those without bisphosphonate treatment (data not shown). However, the sample size in each category was too small to draw a definite conclusion. In addition, the duration of bisphosphonate treatment was not recorded.

Importantly, the patients with NVF exhibited significantly worse final mHAQ scores and a greater increase in mHAQ compared with those without NVF, indicating that NVFs decrease the functional status of RA patients. Furthermore, we found that a NVF without a baseline VF affected patient functional status similar to a NVF with a previous VF.

There are few studies that have investigated the relationship between VF and functional status of RA patients. Dirven et al. reported that functional ability over time was worse in patients with VFs than in patients without VFs, but they did not examine VFs at baseline. Furuya et al. [30] demonstrated that the risk of VF increased by 2.42 for each 1-point increase in Japanese HAQ scores, which is consistent with our results. These results clearly demonstrate the importance of VF prevention in maintaining good functional status in RA patients. Our data also suggest that fracture of the lower lumbar spine had a greater impact on health-related quality of life (QOL). Previous studies in the general population which investigated the association between the site of vertebral deformity and QOL reported similar results [31–33].

The incidence of asymptomatic VF was significantly higher than that of clinical, symptomatic VF in patients treated with PSL. This result is consistent with a previous study [14]. Functional status was poorer in patients with clinical VF or asymptomatic VF than in patients without VF, but clinical VF affected the mHAQ score more than asymptomatic VF. Because the difference was found in the analysis of both NVF and total VF, it is likely that the clinical symptom itself also contributes to a decrease in functional status.

There are several limitations to our study. First, this was a retrospective cohort study performed in one hospital. In addition, only 200 patients in the registry were eligible for the study, mainly because of the incompleteness of the data. In this study we analyzed NVF regardless of the existence of baseline VF because we thought it clinically essential to include data of previous VF in the investigation of incidental NVF. In addition, a larger number of patients is needed to perform additional sub-analysis. Furthermore, the duration of bisphosphonate treatment was not recorded, as mentioned above. Future prospective studies with more systematic data collection are required to draw definite conclusions.

In conclusion, we demonstrated that the incidence of NVF was a risk factor for worsening functional status in postmenopausal patients with RA. In addition, the dose of PSL, BMD, active vitamin D_3 use, and bisphosphonate use were strongly associated with the occurrence of NVF. Disease activity was also an important factor. To maintain a good QOL in RA patients, it is essential to prevent VFs by treating osteoporosis as well as reducing RA disease activity.

Conflict of interest The authors have no conflicts of interest to declare.

References

1. Firestein GS (2003) Evolving concepts of rheumatoid arthritis. Nature 423:356–361
2. Yamanaka H, Sugiyama N, Inoue E, Taniguchi A, Momohara S (2013) Estimates of the prevalence of and current treatment practices for rheumatoid arthritis in Japan using reimbursement data from health insurance societies and the IORRA cohort (I). Mod Rheumatol. doi:10.1007/s10165-013-0863-6
3. Tanaka S (2007) Signaling axis in osteoclast biology and therapeutic targeting in the RANKL/RANK/OPG system. Am J Nephrol 27:466–478
4. Fautrel B, Guillemin F (2002) Cost of illness studies in rheumatic diseases. Curr Opin Rheumatol 14:121–126
5. Furneri G, Mantovani LG, Belisari A, Mosca M, Cristiani M, Bellelli S, Cortesi PA, Turchetti G (2012) Systematic literature review on economic implications and pharmacoeconomic issues of rheumatoid arthritis. Clin Exp Rheumatol 30:S72–S84
6. Tanaka E, Inoue E, Mannalithara A, Bennett M, Kamitsuji S, Taniguchi A, Momohara S, Hara M, Singh G, Yamanaka H

(2010) Medical care costs of patients with rheumatoid arthritis during the prebiologics period in Japan: a large prospective observational cohort study. Mod Rheumatol 20:46–53
7. Adachi JD, Adami S, Gehlbach S, Anderson FA Jr, Boonen S, Chapurlat RD et al (2010) Impact of prevalent fractures on quality of life: baseline results from the global longitudinal study of osteoporosis in women. Mayo Clin Proc 85:806–813
8. Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LA, van Riel PL (2001) The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 44:2009–2017
9. Sokka T, Kankainen A, Hannonen P (2000) Scores for functional disability in patients with rheumatoid arthritis are correlated at higher levels with pain scores than with radiographic scores. Arthritis Rheum 43:386–389
10. Wolfe F (1999) Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology (Oxford) 38:355–361
11. Breedveld FC, Han C, Bala M, van der Heijde D, Baker D, Kavanaugh AF, Maini RN, Lipsky PE (2005) Association between baseline radiographic damage and improvement in physical function after treatment of patients with rheumatoid arthritis. Ann Rheum Dis 64:52–55
12. Dirven L, van den Broek M, van Groenendael JH, de Beus WM, Kerstens PJ, Huizinga TW, Allaart CF, Lems WF (2012) Prevalence of vertebral fractures in a disease activity steered cohort of patients with early active rheumatoid arthritis. BMC Musculoskelet Disord 13:125
13. El Maghraoui A, Rezqi A, Mounach A, Achemlal L, Bezza A, Ghozlani I (2010) Prevalence and risk factors of vertebral fractures in women with rheumatoid arthritis using vertebral fracture assessment. Rheumatology (Oxford) 49:1303–1310
14. Angeli A, Guglielmi G, Dovio A, Capelli G, de Feo D, Giannini S, Giorgino R, Moro L, Giustina A (2006) High prevalence of asymptomatic vertebral fractures in post-menopausal women receiving chronic glucocorticoid therapy: a cross-sectional outpatient study. Bone 39:253–259
15. Ursum J, Britsemmer K, van Schaardenburg D, Lips PT, Dijkmans BA, Lems W (2009) High prevalence of vertebral deformities in elderly patients with early rheumatoid arthritis. Ann Rheum Dis 68:1512–1513
16. Baskan BM, Sivas F, Alemdaroglu E, Duran S, Ozoran K (2007) Association of bone mineral density and vertebral deformity in patients with rheumatoid arthritis. Rheumatol Int 27:579–584
17. Ghazi M, Kolta S, Briot K, Fechtenbaum J, Paternotte S, Roux C (2012) Prevalence of vertebral fractures in patients with rheumatoid arthritis: revisiting the role of glucocorticoids. Osteoporos Int 23:581–587
18. Orstavik RE, Haugeberg G, Mowinckel P, Hoiseth A, Uhlig T, Falch JA, Halse JI, McCloskey E, Kvien TK (2004) Vertebral deformities in rheumatoid arthritis: a comparison with population-based controls. Arch Intern Med 164:420–425
19. Vis M, Haavardsholm EA, Boyesen P, Haugeberg G, Uhlig T, Hoff M, Woolf A, Dijkmans B, Lems W, Kvien TK (2011) High incidence of vertebral and non-vertebral fractures in the OSTRA cohort study: a 5-year follow-up study in postmenopausal women with rheumatoid arthritis. Osteoporos Int 22:2413–2419
20. Spector TD, Hall GM, McCloskey EV, Kanis JA (1999) Risk of vertebral fracture in women with rheumatoid arthritis. BMJ 306:558
21. Arai K, Hanyu T, Sugitani H, Murai T, Fujisawa J, Nakazono K, Kondo N, Endo N (2006) Risk factors for vertebral fracture in menopausal or postmenopausal Japanese women with rheumatoid arthritis: a cross-sectional and longitudinal study. J Bone Miner Metab 24:118–224
22. Orstavik RE, Haugeberg G, Uhlig T, Falch JA, Halse JI, Hoiseth A, Lilleas F, Kvien TK (2003) Vertebral deformities in 229 female patients with rheumatoid arthritis: associations with clinical variables and bone mineral density. Arthritis Rheum 49:355–360
23. Sinigaglia L, Nervetti A, Mela Q, Bianchi G, Del Puente A, Di Munno O, Frediani B, Cantatore F, Pellerito R, Bartolone S, La Montagna G, Adami B (2000) A multicenter cross sectional study on bone mineral density in rheumatoid arthritis. Italian Study Group on Bone Mass in Rheumatoid Arthritis. J Rheumatol 27:2582–2589
24. Harris ST, Watts NB, Jackson RD, Genant HK, Wasnich RD, Ross P, Miller PD, Licata AA, Chestnut CH III (1993) Four-year study of intermittent cyclic etidronate treatment of postmenopausal osteoporosis: three years of blinded therapy followed by one year of open therapy. Am J Med 95:557–567
25. Liberman UA, Weiss SR, Broll J, Minne HW, Quan H, Bell NH et al (1995) Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The Alendronate Phase III Osteoporosis Treatment Study Group. N Engl J Med 333:1437–1443
26. Black DM, Cummings SR, Karpf DB, Cauley JA, Thompson DE, Nevitt MC et al (1996) Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. Lancet 348: 1535–1541
27. Reginster J, Minne HW, Sorensen OH, Hooper M, Roux C, Brandi ML et al (2000) Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. Osteoporos Int 11:83–91
28. Fujita T, Orimo H, Inoue T, Kaneda K, Sakurai M, Morita R et al (2007) Clinical effect of bisphosphonate and vitamin D on osteoporosis: reappraisal of a multicenter double-blind clinical trial comparing etidronate and alfacalcidol. J Bone Miner Metab 25:130–137
29. Harris ST, Watts NB, Genant HK, McKeever CD, Hangartner T, Keller M et al (1999) Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. Vertebral Efficacy with Risedronate Therapy (VERT) Study Group. JAMA 282: 1344–1352
30. Furuya T, Kotake S, Inoue E, Nanke Y, Yago T, Kobashigawa T et al (2007) Risk factors associated with incident clinical vertebral and nonvertebral fractures in Japanese women with rheumatoid arthritis: a prospective 54-month observational study. J Rheumatol 34:303–310
31. Suzuki N, Ogikubo O, Hansson T (2009) The prognosis for pain, disability, activities of daily living and quality of life after an acute osteoporotic vertebral body fracture: its relation to fracture level, type of fracture and grade of fracture deformation. Eur Spine J 18:77–88
32. Oleksik A, Lips P, Dawson A, Minshall ME, Shen W, Cooper C, Kanis J (2000) Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. J Bone Miner Res 15:1384–1392
33. Cockerill W, Ismail AA, Cooper C, Matthis C, Raspe H, Silman AJ et al (2000) Does location of vertebral deformity within the spine influence back pain and disability? European Vertebral Osteoporosis Study (EVOS) Group. Ann Rheum Dis 59:368–371

PLOS | ONE

RESEARCH ARTICLE

Psychometric Properties of the Japanese Version of the STarT Back Tool in Patients with Low Back Pain

Ko Matsudaira[1]*, Hiroyuki Oka[1], Norimasa Kikuchi[2,3], Yuri Haga[2], Takayuki Sawada[2,3], Sakae Tanaka[4]

1 Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, **2** Clinical Study Support, Inc., Nagoya, Aichi, Japan, **3** Department of Public Health, Aichi Medical University School of Medicine, Nagakute, Aichi, Japan, **4** Department of Orthopaedic Surgery, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

* kohart801@gmail.com

Citation: Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S (2016) Psychometric Properties of the Japanese Version of the STarT Back Tool in Patients with Low Back Pain. PLoS ONE 11(3): e0152019. doi:10.1371/journal.pone.0152019

Editor: Masahiko Sumitani, The University of Tokyo Hospital, JAPAN

Received: October 23, 2015

Accepted: March 8, 2016

Published: March 22, 2016

Data Availability Statement: All relevant data are within the paper.

Funding: The authors received no specific funding for this work. Clinical Study Support, Inc. provided support in the form of salaries for authors NK, YH and TS, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing Interests: The authors have the following interests. NK is a board member of Clinical

Abstract

Background and Objective

The STarT Back Tool uses prognostic indicators to classify patients with low back pain into three risk groups to guide early secondary prevention in primary care. The present study aimed to evaluate the psychometric properties of the Japanese version of the tool (STarT-J).

Methods

An online survey was conducted among Japanese patients with low back pain aged 20–64 years. Reliability was assessed by examining the internal consistency of the overall and psychosocial subscales using Cronbach's alpha coefficients. Spearman's correlation coefficients were used to evaluate the concurrent validity between the STarT-J total score/psychosocial subscore and standard reference questionnaires. Discriminant validity was evaluated by calculating the area under the curves (AUCs) for the total and psychosocial subscale scores against standard reference cases. Known-groups validity was assessed by examining the relationship between low back pain-related disability and STarT-J scores.

Results

The analysis included data for 2000 Japanese patients with low back pain; the mean (standard deviation [SD]) age was 47.7 (9.3) years, and 54.1% were male. The mean (SD) STarT-J score was 2.2 (2.1). The Cronbach's alpha coefficient was 0.75 for the overall scale and 0.66 for the psychosocial subscale. Spearman's correlation coefficients ranged from 0.30 to 0.59, demonstrating moderate to strong concurrent validity. The AUCs for the total score ranged from 0.65 to 0.83, mostly demonstrating acceptable discriminative ability. For known-groups validity, participants with more somatic symptoms had higher total scores. Those in higher STarT-J risk groups had experienced more low back pain-related absences.

Study Support, Inc. and co-authors YH and TS are employed by Clinical Study Support, Inc. HO has received grants to his institution from Pfizer, Inc. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

Conclusions

The overall STarT-J scale was internally consistent and had acceptable concurrent, discriminant, and known-groups validity. The STarT-J can be used with Japanese patients with low back pain.

Introduction

Low back pain (LBP) is a major musculoskeletal problem in the general population from childhood to older adulthood, affecting more than 632 million people worldwide [1]. The 2010 Global Burden of Diseases, Injuries, and Risk Factors Study reported that LBP was the leading cause of disability among 291 diseases and injuries globally, and LBP ranked as the highest global cause of years lived with disability [2]. This highlights the high prevalence of LBP worldwide, and may also reflect the difficulty of successful LBP management. In primary care, approximately 85% of patients with LBP have no specific underlying causes or pathology [3]. Patients with non-specific LBP often experience recurrent pain, and the majority of these patients suffer from chronic pain [4–5]. Recurrent and chronic LBP may result in a serious social and economic burden.

Psychological factors have been widely acknowledged as contributors to the chronicity of LBP [6–7]. These factors include pain catastrophizing, fear-avoidance beliefs, and psychological distress. A number of previous reports suggested an association between psychological factors and poor long-term outcomes [5, 8–9]. In primary care, cognitive behavioral therapy focused on psychological factors is a dominant treatment approach for people with LBP [5]. To provide efficient, targeted care, it is becoming common to stratify patients with LBP according to their risk for poor long-term outcomes [10]. Significant clinical benefits and cost-effectiveness of stratified care compared with non-stratified physiotherapy practice have been demonstrated in a randomized clinical trial [11].

The STarT Back Tool (STarT) has been widely used to stratify patients with LBP according to risk for chronicity (Fig 1). The STarT was originally developed as a screening tool for prognostic indicators of back pain to help primary care clinical decision-making in the UK [12]. The STarT consists of 9 items. Items 1–4 evaluate physical factors, and items 5–9 assess psychosocial factors. The STarT classifies patients into three risk groups: patients with a total score of 0–3 are classified as low-risk; patients with a total score of $\geq$ 4 but a psychosocial subscore of $\leq$ 3 as medium-risk; and patients with a psychosocial subscore of $\geq$ 4 are classified as high-risk [12] (Fig 2). Targeted treatments have been developed for patients in each risk group: a minimal intervention by general practitioners or physiotherapists for the low-risk group, physiotherapy to address pain and disability for the medium-risk group, and psychologically-informed physiotherapy to address pain and disability as well as psychosocial obstacles to recovery for the high-risk group [11, 13, 14].

Although the STarT has been translated into various languages, no validated Japanese version was available. In our previous study, we translated the original English version of the STarT into Japanese (STarT-J) and linguistically validated it [15]. As a next step, we conducted online surveys with Japanese people with LBP to evaluate the psychometric properties of the STarT-J. The present analysis aimed to evaluate the reliability and validity of the STarT-J in a large number of Japanese people with LBP, using cross-sectional data from these surveys.

The Keele STarT Back Screening Tool

Patient name: _____ Date: _____

Thinking about the **last 2 weeks** tick your response to the following questions:

		Disagree 0	Agree 1
1	My back pain has **spread down my leg(s)** at some time in the last 2 weeks	☐	☐
2	I have had pain in the **shoulder** or **neck** at some time in the last 2 weeks	☐	☐
3	I have only **walked short distances** because of my back pain	☐	☐
4	In the last 2 weeks, I have **dressed more slowly** than usual because of back pain	☐	☐
5	It's not really safe for a person with a condition like mine to be physically active	☐	☐
6	**Worrying thoughts** have been going through my mind a lot of the time	☐	☐
7	I feel that **my back pain is terrible** and **it's never going to get any better**	☐	☐
8	In general I have **not enjoyed** all the things I used to enjoy	☐	☐

9. Overall, how **bothersome** has your back pain been in the **last 2 weeks**?

Not at all	Slightly	Moderately	Very much	Extremely
☐	☐	☐	☐	☐
0	0	0	1	1

Total score (all 9): _____ Sub Score (Q5-9):_____

Funded by Arthritis Research UK

Fig 1. STarT Back Tool. Response options for items 1–8 are "disagree" (0 points) or "agree" (1 point). Responses to item 9 are on a scale of 1–5: "not at all," "slightly," "moderately," "very much," or "extremely." The first three options ("not at all," "slightly," and "moderately") are scored as 0, and the remaining two options ("very much" and "extremely") are scored as 1. Items 1–4 constitute the physical subscale. Items 5–9 constitute the psychosocial subscale.

doi:10.1371/journal.pone.0152019.g001

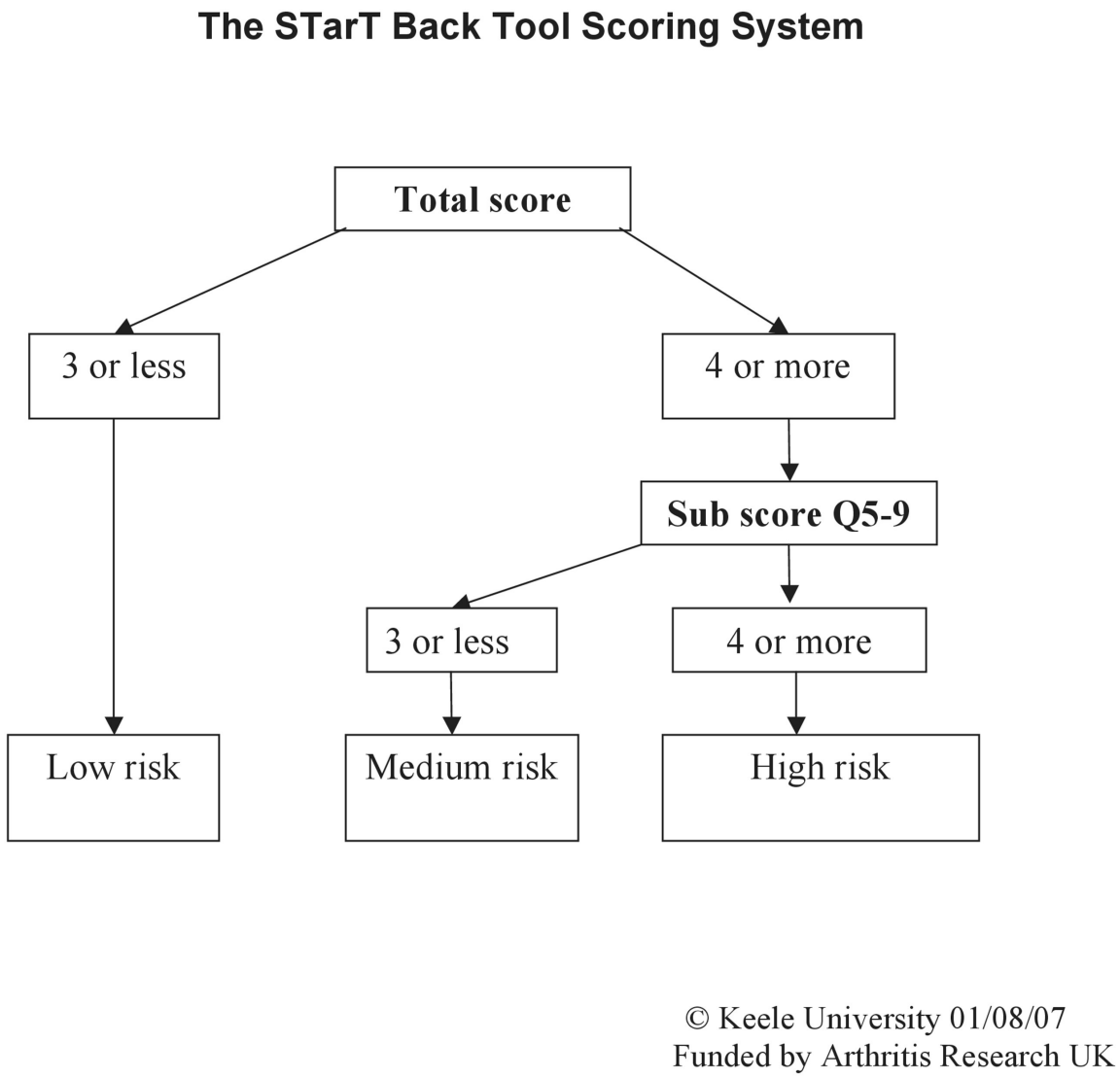


Fig 2. STarT Back Tool risk stratification. Sub score Q5–9: psychosocial subscale.

doi:10.1371/journal.pone.0152019.g002

Materials and Methods

Study population

To assess the psychometric properties of the STarT-J, we conducted online surveys collecting information on LBP in the Japanese population in January and February, 2014. Participants were recruited from an online panel provided by an Internet research company, UNITED, Inc. (Tokyo, Japan), which included approximately 1.25 million individuals aged 20–64 years registered as research volunteers. From these volunteers, 965,919 individuals were randomly selected and invited by e-mail to complete an online questionnaire on health problems associated with pain (first survey). We obtained 52,842 responses by the end of January 2014. Of these initial respondents, those who had LBP in the last 4 weeks were invited to complete

another online questionnaire (secondary survey). LBP was defined as pain in the lower back experienced in the last 4 weeks that lasted for more than 1 day, according to the standard definition of LBP proposed by Dionne et al. [16]. Pain associated with menstruation or pregnancy and pain during a feverish illness were excluded. A diagram showing the lower back area (between the inferior costal margin and gluteal folds) was provided in the questionnaire. The secondary survey closed on 7 February 2014, when the total number of responses reached 2000. The mean (standard deviation [SD]) age of respondents in the secondary survey was 47.7 (9.3) years and 54.1% were male. We conducted two subsequent surveys, 6 and 24 weeks after the secondary survey, to follow up respondents and investigate their LBP condition. In the present analysis, we analyzed secondary survey data to evaluate the psychometric properties of the STarT-J.

We obtained approval from the Medical/Ethics Review Board of the Japan Labour Health and Welfare Organization, Kanto Rosai Hospital (Approval number: 2012–22). Participation was voluntary, and no personal information was collected. Although no written informed consent was obtained, submitting a completed questionnaire was considered as evidence of consent. Potential participants first read an explanation of the aim of the survey and only those who agreed to participate could proceed to the questionnaire. As an incentive, participants received reward points for online shopping from the Internet research company.

Development of the linguistically-validated STarT-J

In our previous study [15], the STarT was translated into Japanese and linguistically validated in a general cross-cultural adaptation process [17–19]. This process occurred in three steps: (1) forward-translation (English to Japanese), (2) back-translation (Japanese to English), and (3) cognitive debriefing. In the third step, we conducted a pilot study to assess if the questions and response scales were understandable and correctly interpreted by Japanese patients. After considering their feedback, and consultation with a specialist as necessary, we published the STarT-J [15].

Measures

We included a number of measures in the online questionnaires.

Pain. The degree of pain associated with LBP during the last 4 weeks was assessed by a numerical rating scale (NRS), ranging from 0 (no pain at all) to 10 (the worst pain imaginable).

Disability caused by LBP. We used the Roland—Morris Disability Questionnaire (RDQ) to assess the LBP-related disability participants experienced in their daily lives. The RDQ comprises 24 Yes/No questions. The total score ranges from 0 to 24, with a higher score indicating greater disability. In this study, we used the Japanese version of the RDQ, for which the reliability and validity have been previously confirmed [20].

Fear-avoidance beliefs. Fear of pain can lead to avoidance of physical activity, an important indicator of a poor long-term LBP prognosis. The Fear-Avoidance Belief Questionnaire (FABQ), consisting of physical activity and work subscales, is widely used to assess fear-avoidance beliefs [21]. We used the FABQ physical activity subscale (FABQ-PA). The FABQ-PA score ranges from 0 to 30; a higher score indicates a stronger fear-avoidance belief. We also used the Tampa Scale of Kinesiophobia (TSK) [22–23], originally developed to measure the fear of movement or injury. The total TSK score sums the scores of 17 items (each rated on a scale of 1–4), and ranges from 17 to 68. A higher score indicates a higher level of kinesiophobia.

Catastrophizing. Pain catastrophizing is also an important indicator of poor LBP prognosis. Catastrophizing was assessed using the Pain Catastrophizing Scale (PCS), originally

developed to measure negative attitudes toward pain involving rumination, helplessness, and magnification. The PCS consists of 13 items. The total score ranges from 0 (no catastrophizing) to 52 (greater catastrophizing). We used the Japanese version of the PCS, for which the reliability and validity have been previously confirmed [24].

Depression and anxiety. A 14-item self-assessment scale, the Hospital Anxiety and Depression Scale (HADS), was used to measure anxiety and depression. The HADS comprises anxiety and depression subscales, each with seven items. The total score ranges from 0 to 21, with a higher score indicating more mental distress. The validity and reliability of the Japanese version of the HADS have been previously confirmed [25].

General health status. The EuroQol 5 Dimension (EQ-5D) [26] is an instrument that provides a simple, descriptive profile and single index value for general health status. The index score is derived from conversion of all responses, and ranges from −0.11 to 1.00. A score of 1 means "perfect health" and a score of 0 denotes "death."

Somatic symptoms. Somatization was assessed using the 7-item somatization subscale from the Brief Symptom Inventory (BSI) [27]. Seven symptoms (faintness or dizziness, pains in the heart or chest, nausea or upset stomach, trouble getting your breath, numbness or tingling in parts of the body, feeling weak in parts of the body, hot or cold spells) are rated on a 5-point scale: "not at all," "a little bit," "moderately," "quite a bit," and "extremely." We used the linguistically validated Japanese version of the BSI-somatization subscale [28].

Data analyses

Participants' demographic and clinical characteristics were summarized using descriptive statistics. To examine floor and ceiling effects, percentages of respondents with total scores of 0 and 9 were calculated. Floor and ceiling effects were considered to exist when more than 15% of respondents had the lowest or highest possible score [29]. To examine the reliability of the STarT-J, we evaluated internal consistency by calculating Cronbach's alpha coefficients for the overall scale and the psychosocial subscale. An alpha index more than 0.70 is considered to indicate satisfactory internal consistency [30].

Concurrent validity was evaluated by measuring correlations between the previously described reference instruments and the STarT-J total score and psychosocial subscore using Spearman's correlation coefficients. Correlation coefficients were evaluated according to the criteria for correlation strength in psychometric validation proposed by Cohen: 0.10 representing a weak, 0.30 a moderate, and 0.50 a strong correlation [31].

To assess discriminant validity, we calculated the area under the curves (AUCs) for the total scores and psychosocial subscores against the reference standards. We defined cases using the following cut-off values: a RDQ score of ≥ 7 for disability, a PCS score of ≥ 20 for catastrophizing, a TSK score of ≥ 41 for fear-avoidance beliefs, and a HADS score of ≥ 8 for depression and anxiety. In addition, a single question was used to determine the presence of referred leg pain within the last 4 weeks. Discriminative ability was interpreted according to the same criteria as used in the original STarT study: 0.70 to < 0.80 indicating acceptable discrimination, 0.80 to < 0.90 indicating excellent discrimination, and ≥ 0.90 indicating outstanding discrimination [12].

For known-groups validity, to test whether the STarT-J scores differentiated participants with known differences, we examined 1) total scores among the groups with a different number of somatic symptoms, and 2) the number of absences due to LBP among the three risk groups (low, medium, and high) using the Jonckheere—Terpstra test. If participants responded "moderately," "quite a bit," or "extremely" to a BSI item, they were considered to have that somatic symptom. Participants were then categorized into three groups according to the number of

somatic symptoms: no symptoms, one symptom, and two or more symptoms. With respect to the number of absences, days on which participants could not perform housework were counted, as well as absences from work. It was hypothesized that participants with more somatic symptoms would have higher total scores, and that participants in the high-risk group would have experienced more LBP-related absences.

All statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA). The level of significance was set at 0.05.

Results

Participant characteristics

The present analysis included data for 2000 Japanese patients with LBP. Table 1 presents a summary of participants' demographic and clinical characteristics. The mean (SD) age was 47.7 (9.3) years; 54.1% of participants were male. More than half (53.7%) of the participants had experienced LBP for more than 1 year. Most participants (92%) experienced recurrent LBP, and more than half (52.9%) reported having LBP 10 times or more.

Scores of the measures

The mean (SD) score for the STarT-J was 2.2 (2.1). No remarkable ceiling effect was observed as 0.9% of participants had the highest score of 9. However, a floor effect was observed as 23.4% of participants had the lowest score of 0. The score distribution for each item is shown in Table 2. Participants were classified into three risk groups according to their STarT-J score: 1557 (77.9%) into the low-risk group, 294 (14.7%) into the medium-risk group, and 149 (7.5%) into the high-risk group.

Reliability

The Cronbach's alpha coefficients were 0.75 for the overall scale and 0.66 for the psychosocial subscale.

Concurrent validity

To examine concurrent validity, Spearman's correlation coefficients were used to measure correlations between the STarT-J total score/psychosocial subscore and the pain NRS, RDQ, FABQ-PA, TSK, PCS, HADS, and the EQ-5D (Table 3). The correlation coefficients for the total score ranged from 0.30 (HADS depression) to 0.59 (RDQ), demonstrating a moderate to strong correlation with these reference standards. Similarly, correlation coefficients for the psychosocial subscore ranged from 0.33 (FABQ-PA) to 0.54 (RDQ), demonstrating a moderate to strong correlation. Both the total score and psychosocial subscore were strongly negatively correlated with the EQ-5D ($\gamma = -0.56$ and $\gamma = -0.53$, $p < 0.0001$). In terms of the correlation with psychosocial measures, the total score was moderately correlated with the TSK ($\gamma = 0.49$), whereas the psychosocial subscore was strongly correlated ($\gamma = 0.53$). Moderate correlation coefficients were observed for both the total score and psychosocial subscore with the PCS ($\gamma = 0.46$ and $\gamma = 0.49$) and the HADS ($\gamma = 0.40$ and $\gamma = 0.45$) ($p < 0.0001$ for all).

Discriminant validity

To assess discriminant validity, AUCs were calculated for the total score and psychosocial subscore against the cases defined by the reference standards (Table 4). The AUCs for the total score were all above 0.70, indicating acceptable to excellent discriminative ability, with the exception of depression and anxiety (0.65). For the psychosocial subscore, the AUCs ranged

Table 1. Participant characteristics: psychometric testing of the STarT-J (n = 2000).

Characteristics	n (%)	Mean (SD)
Sex		
Male	1081 (54.1)	
Female	919 (46.0)	
Age (years)		47.7 (9.3)
BMI $\geq$ 25 (kg/m^2)	506 (25.3)	
Duration of low back pain		
< 2 weeks	350 (17.5)	
$\geq$ 2 weeks, < 1 month	188 (9.4)	
$\geq$ 1, < 3 months	184 (9.2)	
$\geq$ 3, < 6 months	90 (4.5)	
$\geq$ 6 months, < 1 year	115 (5.8)	
$\geq$ 1, < 3 years	200 (10.0)	
$\geq$ 3 years	873 (43.7)	
Number of recurrence		
1	160 (8.0)	
2	135 (6.8)	
3–4	340 (17.0)	
5–9	308 (15.4)	
$\geq$10	1057 (52.9)	
STarT-J score		2.2 (2.1)
RDQ score		4.2 (4.7)
FABQ-PA score		12.9 (4.7)
TSK score		41.0 (6.5)
PCS total score		21.6 (10.0)
PCS rumination		10.6 (4.3)
PCS helplessness		6.2 (4.2)
PCS magnification		4.7 (2.7)
HADS total score		17.2 (6.7)
HADS anxiety		8.7 (3.4)
HADS depression		8.5 (4.1)
EQ-5D index score		0.78 (0.16)
NRS for low back pain		4.2 (1.8)

Values are n (%), or mean (SD).

STarT-J, the Japanese version of the STarT Back Tool; BMI, body mass index; RDQ, Roland—Morris Disability Questionnaire; FABQ-PA, Fear-Avoidance Belief Questionnaire Physical Activity Subscale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale; HADS, Hospital Anxiety and Depression Scale; EQ-5D, EuroQol 5 Dimension; NRS, numerical rating scale.

doi:10.1371/journal.pone.0152019.t001

from 0.67 (depression and anxiety) to 0.79 (disability), indicating poor to acceptable discriminative ability.

Known-groups validity

We examined the STarT-J total scores and risk groups among participants with known-differences. As hypothesized, participants with more somatic symptoms had higher total scores. The mean (SD) score of participants with no somatic symptoms was 1.71 (1.76), one somatic symptom was 2.73 (2.14), and two or more somatic symptoms was 3.76 (2.50) (Fig 3). A linear

Table 2. Score distribution of STarT-J items and risk group distribution (n = 2000).

Item	Number of participants who answered "agree" (1 point) n (%)
1	442 (22.1)
2	1069 (53.5)
3	317 (15.9)
4	264 (13.2)
5	574 (28.7)
6	652 (32.6)
7	425 (21.3)
8	351 (17.6)
9	239 (12.0)
Risk group distribution	
Low-risk	1557 (77.9)
Medium-risk	294 (14.7)
High-risk	149 (7.5)

Values are n (%).

STarT-J: The Japanese version of the STarT Back Tool. For item 9, answers of "very much" and "extremely" were scored as 1 point, and were counted as "agree"; the answers "not at all," "slightly," and "moderately" were scored as 0 points, and were not included.

doi:10.1371/journal.pone.0152019.t002

increasing trend in total score across groups with an increasing number of somatic symptoms was observed (Jonckheere-Terpstra test, $p < 0.0001$). With respect to the associations between risk groups and the number of absences, participants in the high-risk group reported a larger number of absences (Fig 4). The mean (SD) LBP-related absences in the low-risk group was 4.0 (5.4) days, 6.6 (8.3) days in the medium-risk group, and 12.6 (11.1) days in the high-risk group. A linear increasing trend in the number of absences across the risk groups was observed (Jonckheere-Terpstra test, $p < 0.0001$).

Table 3. Spearman's correlation coefficients for the STarT-J and related measures.

Measures	Total score Coefficients (95% CI)	Psychosocial subscore Coefficients (95% CI)
RDQ	0.59 (0.56–0.62)	0.54 (0.51–0.57)
FABQ-PA	0.34 (0.30–0.37)	0.33 (0.29–0.37)
TSK	0.49 (0.45–0.52)	0.53 (0.50–0.56)
PCS total	0.46 (0.42–0.49)	0.49 (0.46–0.52)
PCS rumination	0.43 (0.40–0.47)	0.44 (0.41–0.48)
PCS helplessness	0.39 (0.35–0.43)	0.43 (0.39–0.46)
PCS magnification	0.40 (0.36–0.44)	0.43 (0.39–0.47)
HADS total	0.40 (0.36–0.44)	0.45 (0.41–0.48)
HADS anxiety	0.42 (0.38–0.46)	0.46 (0.42–0.49)
HADS depression	0.30 (0.26–0.34)	0.35 (0.31–0.39)
EQ-5D	–0.56 (–0.59 –-0.52)	–0.53 (–0.56 –-0.50)
NRS for low back pain	0.42 (0.38–0.46)	0.39 (0.35–0.42)

Note: $p < 0.0001$ for all correlation coefficients. STarT-J, the Japanese version of the STarT Back Tool; CI, confidence interval; RDQ, Roland—Morris Disability Questionnaire; FABQ-PA, Fear-Avoidance Belief Questionnaire Physical Activity Subscale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale; HADS, Hospital Anxiety and Depression Scale; EQ-5D, EuroQol 5 Dimension; NRS, numerical rating scale.

doi:10.1371/journal.pone.0152019.t003

Table 4. AUCs for STarT-J total score and psychosocial subscore against reference standards.

Reference standards	Case definition	Total score AUC (95% CI)	Psychosocial subscore AUC (95% CI)
Disability	RDQ score ≥ 7	0.83 (0.81–0.85)	0.79 (0.77–0.82)
Referred leg pain	Yes	0.76 (0.73–0.79)	0.68 (0.65–0.72)
Fear-avoidance belief	PCS score ≥ 20	0.71 (0.69–0.73)	0.72 (0.70–0.74)
Catastrophizing	TSK score ≥ 41	0.74 (0.72–0.76)	0.75 (0.73–0.77)
Depression and anxiety	HADS score ≥ 8	0.65 (0.63–0.68)	0.67 (0.65–0.69)

AUC, area under the curve; STarT-J, the Japanese version of the STarT Back Tool; CI, confidence interval; RDQ, Roland—Morris Disability Questionnaire; PCS, Pain Catastrophizing Scale; TSK, Tampa Scale for Kinesiophobia; HADS, Hospital Anxiety and Depression Scale.

doi:10.1371/journal.pone.0152019.t004

Discussion

In this analysis, we evaluated the psychometric properties of the STarT-J. In summary, the overall scale of the STarT-J was internally consistent, and the STarT-J had acceptable concurrent validity, discriminant validity, and known-groups validity in Japanese patients with LBP.

The Cronbach's alpha coefficient for the overall scale (0.75) demonstrated sufficient internal consistency, and was similar to the original and other language versions: 0.79 for the original [12], 0.74 for the French [32], 0.74 for the Brazilian Portuguese [33], 0.82 for the Iranian [34], and 0.83 for the Persian [35] versions. Although these results could not be compared directly because the study methods varied, the similar values support that the overall scale of the STarT-J is internally consistent and no items are redundant. The Cronbach's alpha coefficient for the psychosocial subscale was 0.66, below the value of 0.70 considered necessary to claim

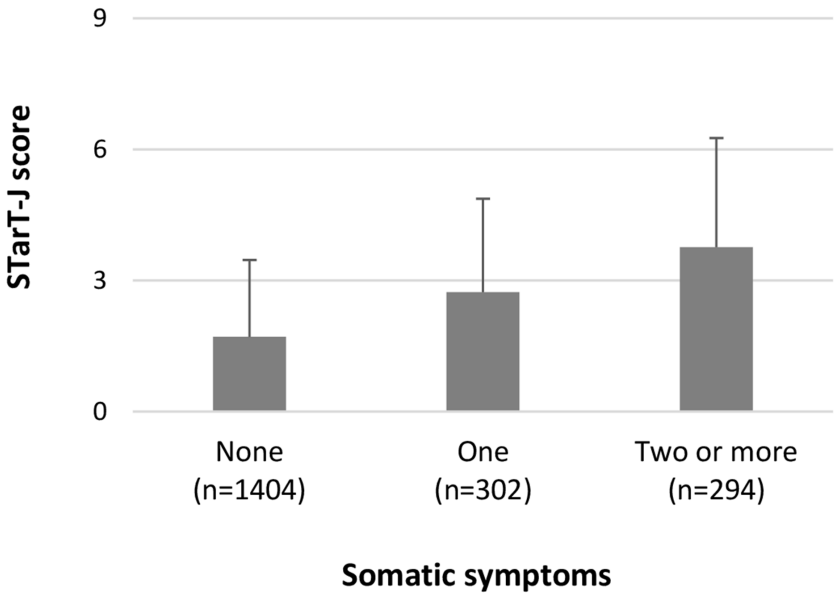


Fig 3. Mean STarT-J scores for participants with different numbers of somatic symptoms. The linear trend was tested using the Jonckheere-Terpstra test ($p < 0.0001$). STarT-J: The Japanese version of the STarT Back Tool. Number of somatic symptoms was assessed by the Brief Symptom Inventory somatization scale: a response of "moderately," "quite a bit," or "extremely" to an item was interpreted as the presence of that somatic symptom, and thus counted.

doi:10.1371/journal.pone.0152019.g003

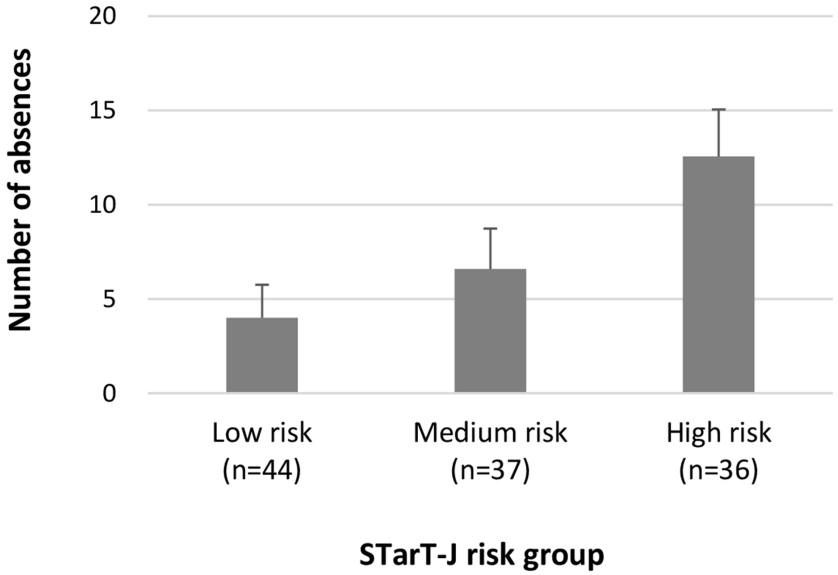


Fig 4. Mean number of absences for the three STarT-J risk groups. The linear trend was tested using the Jonckheere-Terpstra test ($p < 0.0001$). STarT-J: The Japanese version of the STarT Back Tool.

doi:10.1371/journal.pone.0152019.g004

the subscale is internally consistent. However, it should be taken into consideration that the coefficient for the subscale was also lower than for the overall scale in the original version, although it was still 0.74 [12].

To assess concurrent validity, we analyzed the correlations between the STarT-J and reference standards (the pain NRS, RDQ, FABQ-PA, TSK, PCS, HADS, and EQ-5D). Overall, the Spearman's correlation coefficients indicated that both the total score and the psychosocial subscore were moderately to strongly correlated with these existing scales. In particular, the STarT-J total score was strongly correlated with the RDQ ($\gamma = 0.59$). Similar results were observed in the German ($\gamma = 0.55$) [36], French ($\gamma = 0.74$) [32], and Persian ($\gamma = 0.811$) [35] versions. Although a direct comparison cannot easily be made, these similar results reinforce the concurrent validity of the STarT-J.

Discriminant validity was assessed by calculating the AUCs for the total score and the psychosocial subscore. For the total score, the AUCs for disability and referred leg pain were both higher than the AUCs for fear-avoidance beliefs, catastrophizing, and depression and anxiety. This demonstrated that the total score better discriminated cases defined by physical reference standards. However, for the psychosocial subscore, the AUCs for fear-avoidance beliefs, catastrophizing, and depression and anxiety were not remarkably higher than AUCs for the physical reference cases. These AUCs for the psychosocial reference cases were similar to those for the total score, indicating the psychosocial subscale might discriminate cases defined by the psychosocial reference standards at a similar level to the overall scale. A similar trend was observed in the original STarT [12], although overall, the AUCs were higher compared with the STarT-J.

To assess known-groups validity, we investigated relationships between total scores and the number of somatic symptoms, and between risk groups and the number of absences. Participants with more somatic symptoms had higher total scores, and those in the high-risk group had experienced greater LBP-related disability. This demonstrated that the STarT-J can differentiate patients with different levels of LBP-related problems.

The present study has some limitations. First, we did not examine the test—retest reliability. The intra-class, test—retest reliability over specific time intervals should therefore be evaluated in a future study. Second, the analysis might have included patients not targeted by the STarT, that is, patients who had specific causes of LBP. The diagnostic triage for LBP is to classify LBP into one of three categories: LBP with specific pathologic change ("red flag"), LBP with sciatica/radicular syndrome, or non-specific LBP [37]. According to this classification, six of the participants in the present analysis were probable "red flags," 308 had radicular syndrome, and the remaining 1686 participants were considered to have non-specific LBP. As the original study included patients with non-specific LBP who had referred leg pain [12], the STarT is considered applicable to patients with LBP potentially associated with sciatica/radicular syndrome. Therefore, assuming diagnoses were accurate, most participants probably fit into the STarT target group. However, it should be noted that these diagnoses might not necessarily be accurate as they were based on participants' self-report. Third, the study population might not be consistent with the primary care population. Our study included more low-risk participants and less high-risk participants compared with the original study [12]. This might be because we recruited from a general Japanese population registered with an online panel rather than from patients in hospitals. Our study population would therefore represent the general Japanese population with LBP. As the observed floor effect suggests, more patients might have LBP that was not sufficiently severe to require hospital care. Although our study population was broader than the primary care population, the percentage of patients with non-specific LBP was similar to that observed in primary care settings. In our study, 1686 participants (84.3%) probably had non-specific LBP. In primary care, approximately 85% of patients with LBP have non-specific LBP [3]. Therefore, our study population resembled the primary care population in terms of the distribution of non-specific LBP. Fourth, as this was a cross-sectional study, it did not assess the ability of the STarT-J to predict chronicity of LBP. To assess its predictive ability, longitudinal studies will be necessary to investigate associations between risk groups and long-term outcomes of patients with LBP.

In the present analysis, we evaluated the psychometric properties of the STarT-J to enable Japanese clinicians to use the scale in the early stages of LBP. The STarT is a simple and quick tool, and is suitable for use in primary care settings. Stratified care is a dominant approach in the management of LBP [10]. Stratified care based on the STarT risk groups has been shown to be clinically and economically beneficial for patients with LBP [11, 38]. Therefore, we expect that the STarT-J may facilitate early stratified care in primary care settings in Japan. This may alleviate the physical, social, and economical burden of LBP in the Japanese population.

In conclusion, acceptable internal consistency for the overall STarT-J scale demonstrated the reliability of the STarT-J in Japanese patients with LBP; acceptable concurrent validity, discriminant validity, and known-groups validity demonstrated the validity. In a subsequent analysis, the ability of the STarT-J to predict chronicity of LBP will be examined using longitudinal data, to validate its clinical use in Japanese patients.

Acknowledgments

We thank all those who participated in our online surveys.

Author Contributions

Conceived and designed the experiments: KM HO ST. Performed the experiments: KM HO. Analyzed the data: KM HO NK YH TS ST. Contributed reagents/materials/analysis tools: KM HO. Wrote the paper: KM HO NK YH TS ST.

References

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380: 2224–60. doi: 10.1016/S0140-6736(12)61766-8 PMID: 23245609
2. Buchbinder R, Blyth FM, March LM, Brooks P, Woolf AD, Hoy DG. Placing the global burden of low back pain in context. Best Pract Res Clin Rheumatol. 2013; 27: 575–89. doi: 10.1016/j.berh.2013.10.007 PMID: 24315140
3. Deyo RA, Weinstein JN. Low back pain. N Engl J Med. 2001; 344: 363–70. PMID: 11172169
4. Axen I, Leboeuf-Yde C. Trajectories of low back pain. Best Pract Res Clin Rheumatol. 2013; 27: 601–12. doi: 10.1016/j.berh.2013.10.004 PMID: 24315142
5. Dunn KM, Hestbaek L, Cassidy JD. Low back pain across the life course. Best Pract Res Clin Rheumatol. 2013; 27: 591–600. doi: 10.1016/j.berh.2013.09.007 PMID: 24315141
6. Hasenbring M, Hallner D, Klasen B. Psychological mechanisms in the transition from acute to chronic pain: over- or underrated? Schmerz. 2001; 15: 442–7. German. PMID: 11793149
7. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine (Phila Pa 1976). 2002; 27: E109–20.
8. Hill JC, Fritz JM. Psychosocial influences on low back pain, disability, and response to treatment. Phys Ther. 2011; 91: 712–21. doi: 10.2522/ptj.20100280 PMID: 21451093
9. Pincus T, McCracken LM. Psychological factors and treatment opportunities in low back pain. Best Pract Res Clin Rheumatol. 2013; 27: 625–35. doi: 10.1016/j.berh.2013.09.010 PMID: 24315144
10. Foster NE, Hill JC, O'Sullivan P, Hancock M. Stratified models of care. Best Pract Res Clin Rheumatol. 2013; 27: 649–61. doi: 10.1016/j.berh.2013.10.005 PMID: 24315146
11. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet. 2011; 378: 1560–71. doi: 10.1016/S0140-6736(11)60937-9 PMID: 21963002
12. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum. 2008; 59: 632–41. doi: 10.1002/art.23563 PMID: 18438893
13. Main CJ, Sowden G, Hill JC, Watson PJ, Hay EM. Integrating physical and psychological approaches to treatment in low back pain: the development and content of the STarT Back trial's 'high-risk' intervention (StarT Back; ISRCTN 37113406). Physiotherapy. 2012; 98: 110–6. doi: 10.1016/j.physio.2011.03.003 PMID: 22507360
14. Sowden G, Hill JC, Konstantinou K, Khanna M, Main CJ, Salmon P, et al.; IMPaCT Back study team. Targeted treatment in primary care for low back pain: the treatment system and clinical training programmes used in the IMPaCT Back study (ISRCTN 55174281). Fam Pract. 2012; 29: 50–62. doi: 10.1093/fampra/cmr037 PMID: 21708984
15. Matsudaira K, Kikuchi N, Kawaguchi M, Inuzuka K, Arisaka M, Hara N, et al. Development of a Japanese version of the STarT (Subgrouping for Targeted Treatment) Back screening tool: translation and linguistic validation. Journal of Musculoskeletal Pain Research. 2013; 5: 11–19. Japanese.
16. Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976). 2008; 33: 95–103.
17. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993; 46: 1417–32. PMID: 8263569
18. Suzukamo Y, Kumano H. Psychometrics. In: Ikegami N, Fukuhara S, Shimozuma K, Ikeda S, editors. QOL evaluation handbook for clinical diagnosis. Tokyo: Igaku Shoin; 2001. p. 8–13. Japanese.
19. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health. 2005; 8: 94–104. PMID: 15804318
20. Suzukamo Y, Fukuhara S, Kikuchi S, Konno S, Roland M, Iwamoto Y, et al. Validation of the Japanese version of the Roland-Morris Disability Questionnaire. J Orthop Sci. 2003; 8: 543–8. PMID: 12898308
21. Matsudaira K, Kikuchi N, Murakami A, Isomura T. Psychometric properties of the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ). J Orthop Sci. 2014; 19: 26–32. doi: 10.1007/s00776-013-0471-5 PMID: 24091984

22. Matsudaira K, Inuzuka K, Kikuchi N, Sakae C, Arisaka M, Isomura T. Development of a Japanese version of the Tampa Scale for Kinesiophobia (TSK-J): translation and linguistic validation. Seikei Geka (Orthopedic surgery). 2013; 48: 13–9. Japanese.

23. Kikuchi N, Matsudaira K, Sawada T, Oka H. Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury pain and/or low back pain. J Orthop Sci. Epub 2015 Jul 23.

24. Matsuoka H, Sakano Y. Assessment of cognitive aspect of pain: development, reliability, and validation of Japanese version of pain catastrophizing scale. Japanese Journal of Psychosomatic Medicine. 2007; 47: 95–102. Japanese.

25. Higashi A, Yashiro H, Kiyota K, Inokuchi H, Hatta H, Fujita K, et al. Validation of the hospital anxiety and depression scale in a gastro-intestinal clinic. The Japanese journal of gastro-enterology. 1996; 93: 884–92. Japanese. PMID: 8986079

26. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy. 1990; 16: 199–208. PMID: 10109801

27. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med. 1983; 13: 595–605. PMID: 6622612

28. Matsudaira K, Inuzuka K, Kikuchi N, Sakae C, Arisaka M, Isomura T. Development of the Japanese version of the brief symptom inventory-somatization scale: translation and linguistic validation. Seikei Geka (Orthopedic surgery). 2012; 63: 149–53. Japanese.

29. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007; 60: 34–42. PMID: 17161752

30. Nunnally JC. Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.

31. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

32. Bruyere O, Demoulin M, Beaudart C, Hill JC, Maquet D, Genevay S, et al. Validity and reliability of the French version of the STarT Back screening tool for patients with low back pain. Spine (Phila Pa 1976). 2014; 39: E123–8.

33. Pilz B, Vasconcelos RA, Marcondes FB, Lodovichi SS, Mello W, Grossi DB. The Brazilian version of STarT Back Screening Tool—translation, cross-cultural adaptation and reliability. Braz J Phys Ther. 2014; 18: 453–61. PMID: 25372008

34. Azimi P, Shahzadi S, Azhari S, Montazeri A. A validation study of the Iranian version of STarT Back Screening Tool (SBST) in lumbar central canal stenosis patients. J Orthop Sci. 2014; 19: 213–7. doi: 10.1007/s00776-013-0506-y PMID: 24343300

35. Abedi M, Manshadi FD, Khalkhali M, Mousavi SJ, Baghban AA, Montazeri A, et al. Translation and validation of the Persian version of the STarT Back Screening Tool in patients with nonspecific low back pain. Man Ther. Epub 2015 Apr 15.

36. Aebischer B, Hill JC, Hilfiker R, Karstens S. German Translation and Cross-Cultural Adaptation of the STarT Back Screening Tool. PLOS ONE. 2015; 10: e0132068. doi: 10.1371/journal.pone.0132068 PMID: 26161669

37. Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. Spine (Phila Pa 1976). 2001; 26: 2504–13; discussion 2513–4.

38. Foster NE, Mullis R, Hill JC, Lewis M, Whitehurst DG, Doyle C, et al. Effect of stratified care for low back pain in family practice (IMPaCT Back): a prospective population-based sequential comparison. Ann Fam Med. 2014; 12: 102–11. doi: 10.1370/afm.1625 PMID: 24615305

J Bone Miner Metab (2014) 32:524–532
DOI 10.1007/s00774-013-0522-1

ORIGINAL ARTICLE

Prevalence of knee pain, lumbar pain and its coexistence in Japanese men and women: The Longitudinal Cohorts of Motor System Organ (LOCOMO) study

Noriko Yoshimura · Toru Akune · Saeko Fujiwara · Yoko Shimizu · Hideyo Yoshida · Go Omori · Akihiro Sudo · Yuji Nishiwaki · Munehito Yoshida · Hiroshi Shimokata · Takao Suzuki · Shigeyuki Muraki · Hiroyuki Oka · Kozo Nakamura

Received: 9 July 2013 / Accepted: 9 September 2013 / Published online: 9 November 2013

Abstract The Longitudinal Cohorts of Motor System Organ (LOCOMO) study was initiated in 2008 through a grant from the Ministry of Health, Labour, and Welfare of Japan to integrate information from several cohorts established for the prevention of musculoskeletal diseases. We integrated the information of 12,019 participants (3,959 men and 8,060 women) in the cohorts comprising nine communities located in Tokyo (two regions: Tokyo-1 and Tokyo-2), Wakayama [two regions: Wakayama-1 (mountainous region) and Wakayama-2 (seaside region)], Hiroshima, Niigata, Mie, Akita, and Gunma prefectures. The baseline examination of the LOCOMO study consisted of an interviewer-administered questionnaire, anthropometric measurements, medical information recording, X-ray radiography, and bone mineral density measurement. The prevalence of knee pain was 32.7 % (men 27.9 %; women 35.1 %) and that of lumbar pain was 37.7 % (men 34.2 %; women 39.4 %). Among the 9,046 individuals who were surveyed on both knee pain and lumbar pain at the baseline examination in each cohort, we noted that the prevalence of both knee pain and lumbar pain was 12.2 % (men 10.9 %; women 12.8 %). Logistic regression analysis showed that higher age, female sex, higher body mass index (BMI), living in a rural area, and the presence of lumbar pain significantly influenced the presence of knee pain. Similarly, higher age, female sex, higher BMI, living in a rural area, and the presence of knee pain significantly influenced the presence of lumbar pain. Thus, by using the data of the

N. Yoshimura (✉) · H. Oka
Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
e-mail: yoshimuran-ort@h.u-tokyo.ac.jp

T. Akune · S. Muraki
Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

S. Fujiwara
Hiroshima Atomic Bomb Casualty Council, Hiroshima, Japan

Y. Shimizu · H. Yoshida
Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan

G. Omori
Center for Transdisciplinary Research, Institute for Research Promotion, Niigata University, Niigata, Japan

A. Sudo
Department of Orthopaedic Surgery, Mie University Graduate School of Medicine, Mie, Japan

Y. Nishiwaki
Department of Environmental and Occupational Health, School of Medicine, Toho University, Tokyo, Japan

M. Yoshida
Department of Orthopedic Surgery, School of Medicine, Wakayama Medical University, Wakayama, Japan

H. Shimokata
Graduate School of Nutritional Sciences, Nagoya University of Arts and Sciences, Aichi, Japan

T. Suzuki
National Center for Geriatrics and Gerontology, Aichi, Japan

K. Nakamura
National Rehabilitation Center for Persons with Disabilities, Saitama, Japan

LOCOMO study, we clarified the prevalence of knee pain and lumbar pain, their coexistence, and their associated factors.

Keywords Nation-wide population-based cohort study · Epidemiology · Prevalence · Knee pain · Lumbar pain

Introduction

Musculoskeletal diseases, including osteoarthritis (OA) and osteoporosis (OP), are major public health problems among the elderly; these diseases can affect activities of daily living (ADL) and quality of life (QOL), and can lead to increased morbidity and mortality. According to the recent National Livelihood Survey by the Ministry of Health, Labour, and Welfare in Japan, OA is ranked fourth among diseases that cause disabilities and subsequently require support for ADL, whereas falls and osteoporotic fractures are ranked fifth [1]. Studies have reported increased mortality after osteoporotic fractures at the hip and other sites [2]. An estimated 47,000,000 individuals (21,000,000 men and 26,000,000 women) aged ≥40 years will eventually be affected by either OA or OP [3].

Considering that the population of Japan is aging rapidly, a comprehensive and evidence-based prevention strategy for musculoskeletal diseases is urgently needed. However, only a few prospective, longitudinal studies designed to develop such a strategy have been conducted. Therefore, little information is available regarding the incidence of disability and the prevalence and incidence of musculoskeletal disorders, including knee pain, and lumbar pain, and their associated factors in Japan. The absence of such epidemiological data hampers the rational design of clinical and public health approaches for the diagnosis, evaluation, and prevention of musculoskeletal diseases.

Several cohorts have focused on the prevention of OP, knee OA (KOA), lumbar spondylosis (LS) or disability caused by musculoskeletal diseases. However, since the prevalence of the musculoskeletal diseases has been reported to be high [3], the extent of the population at risk after excluding those who had the target disease at the baseline seems to be small. To identify epidemiological indices, especially the incidence of musculoskeletal diseases and/or disability, a large number of subjects is required. In addition, to determine the regional differences in epidemiological indices, we need a survey of cohorts across Japan.

The Longitudinal Cohorts of Motor System Organ (LOCOMO) study was initiated in 2008 by the members of the committee for 'the prevention of knee and back pain and bone fractures in a large cohort of regionally representative residents from across Japan,' through a grant from the Ministry of Health, Labour, and Welfare of Japan (Director, Noriko Yoshimura). This study aimed to integrate the information of several cohorts established for the prevention of musculoskeletal diseases from 2000 onwards, and to initiate a follow-up examination using the unified questionnaire from 2006 onwards in Japan.

In the present paper, by using the integrated information at the baseline of the LOCOMO study, we tried to confirm the prevalence of clinical symptoms of musculoskeletal diseases, such as knee pain and lumbar pain and their characteristics.

Materials and methods

Participants

Participants in the cohorts were residents of nine communities located in Tokyo (two regions: Tokyo-1, principle investigators (PIs): Shigeyuki Muraki, Toru Akune, Noriko Yoshimura, Kozo Nakamura; Tokyo-2, PIs: Yoko Shimizu, Hideyo Yoshida, Takao Suzuki), Wakayama [two regions: Wakayama-1 (mountainous region) and Wakayama-2 (seaside region); PIs: Noriko Yoshimura, Munehito Yoshida], Hiroshima (PI: Saeko Fujiwara), Niigata (PI: Go Omori), Mie (PI: Akihiro Sudo), Akita (PI: Hideyo Yoshida), and Gunma (PI: Yuji Nishiwaki) prefectures [4]. Figure 1 shows the location of each cohort in Japan, and Fig. 2 provides the timeline of the LOCOMO study. Residents of the nine regions were recruited from resident registration lists in the relevant region. Data for the 12,019 participants were collected and registered as an integrated cohort. Numbers of participants in the LOCOMO study classified by regions of each cohort are shown in Table 1. The smallest cohort consisted of 826 individuals in Wakayama-2, and the largest consisted of 2,613 individuals in Hiroshima.

All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo (nos. 1264 and 1326), the Tokyo Metropolitan Institute of Gerontology (no. 5), Wakayama (no. 373), The Radiation Effects Research Foundation (RP03-89), Niigata University (no. 446), Mie University (no. 837 and no. 139), Keio University (no. 16–20), and National Center for Geriatrics and Gerontology (no. 249). Safety of the participants was ensured during the examination and during all other study procedures.

Data collection

The baseline examination of the LOCOMO study consisted of the following: an interviewer-administered questionnaire,

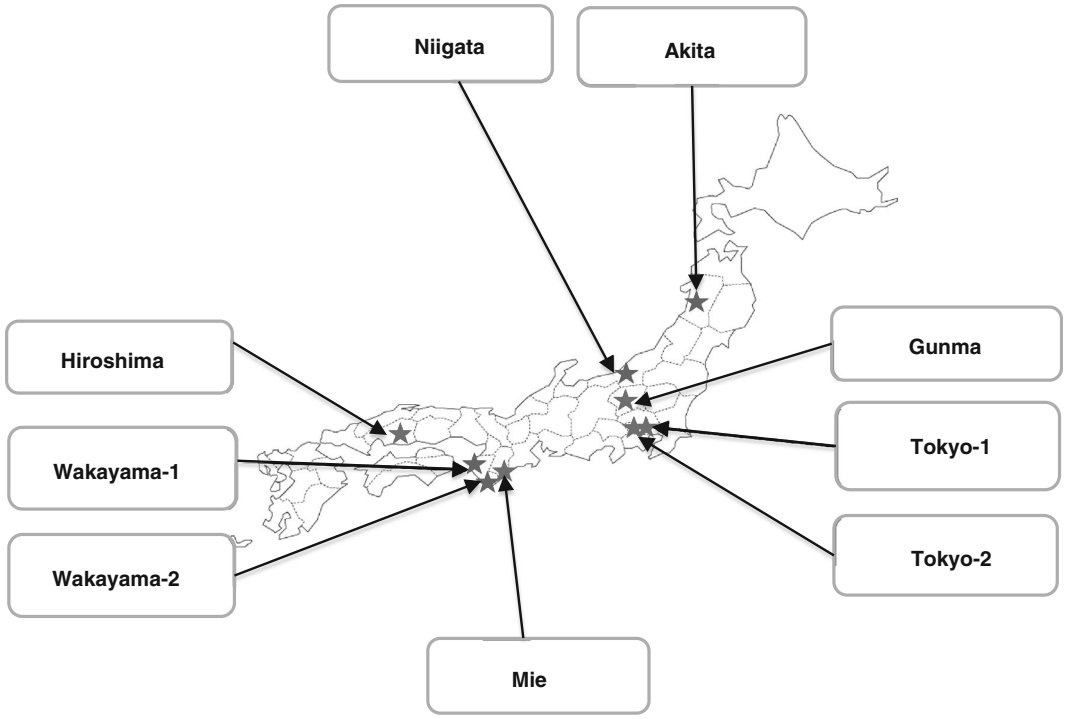


Fig. 1 Locations of the nine different regions from which the study cohorts were derived

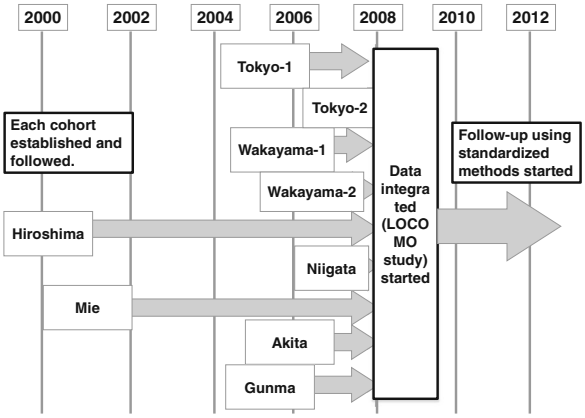


Fig. 2 Timeline of the LOCOMO study

Table 1 Numbers of participants in the LOCOMO study classified by regions of each cohort

Regions of each cohort	Start year	Total	Men	Women
Tokyo-1	2005	1,350	465	885
Tokyo-2	2008	1,453	59	1,394
Wakayama-1 (mountainous)	2005	864	319	545
Wakayama-2 (seaside)	2006	826	277	549
Hiroshima	2000	2,613	794	1,819
Niigata	2007	1,474	628	846
Mie	2001	1,175	423	752
Akita	2006	852	366	486
Gunma	2005	1,412	628	784
Total		12,019	3,959	8,060

anthropometric measurements, medical information recording, radiography, and bone mineral density (BMD) measurement.

Interviewer-administered questionnaire

A questionnaire was prepared by modifying the questionnaire used in the Osteoporotic Fractures in Men Study (MrOS) [5], and some new items were added to the modified questionnaire. Knee symptoms were evaluated using the Western Ontario and McMaster University Osteoarthritis Index (WOMAC) [6]. Health-related QOL was evaluated using the European QOL-5 dimensions instrument (EuroQOL EQ5D) [7] and the Medical Outcomes Study 8-item Short Form (SF-8) [8]. The study staff recorded all the medications administered and their doses.

Anthropometric measurements

Anthropometric factors were measured by well-trained medical nurses. Body mass index [BMI; weight in

kilograms/(height in meters)^2] was calculated on the basis of the current height and weight. Hand grip strength was measured using a Toei Light handgrip dynamometer (Toei Light Co., Ltd., Saitama, Japan). Both hands were tested, and the higher value was used to characterise the maximum muscle strength of the subject. Walking speed was determined by recording the time taken by a subject to walk a determined distance, such as 5 or 6 m, at his/her usual speed. The ability to rise from a chair without using the arms (chair stand) and the ability to perform 5 chair stands was evaluated. The time required to complete the tasks was recorded.

Medical information

Medical information was obtained by experienced medical doctors in each cohort. All participants were questioned about pain in both knees by asking the following questions: 'Have you experienced right knee pain on most days (and continuously on at least one day) in the past month, in addition to the current pain?' and 'Have you experienced left knee pain on most days (and continuously on at least one day) in the past month, in addition to the current pain?' Subjects who answered 'yes' were considered to have knee pain. Lumbar pain was determined by asking the following question: 'Have you experienced lumbar pain on most days (and continuously on at least one day) in the past month, in addition to the current pain?' Subjects who answered 'yes' were considered to have lumbar pain.

In some cohorts (Tokyo-1, Wakayama-1, and Wakayama-2), the participants completed the modified Mini-Mental Status Examination-Japanese version [9] for evaluating cognitive function. Physicians explained any unclear sections of this questionnaire to the participants and assessed the cognitive status on the basis of the completed questionnaire.

Radiography and radiographic assessment

In several cohorts (Tokyo-1, Wakayama-1, Wakayama-2, Hiroshima, Niigata, and Mie), the radiographic examination of knees and/or spine was performed to evaluate the OA or fractures. Plain radiographs were obtained for both knees in the antero-posterior view with weight-bearing and foot map positioning and for the spine in the antero-posterior and lateral views.

The severity of OA was radiographically determined according to the Kellgren-Lawrence (KL) grading system as follows [10]: KL0, normal joint; KL1, slight osteophytes; KL2, definite osteophytes; KL3, narrowing of joint cartilage, and large osteophytes; and KL4, bone sclerosis, narrowing of joint cartilage, and large osteophytes. In the LOCOMO study, joints exhibiting disc-space narrowing alone and no large osteophytes were graded as KL3. In each cohort, radiographs were examined by a single, experienced orthopaedic surgeon who was masked to the clinical status of the participants. If at least one knee joint was graded as KL2 or higher, the participant was diagnosed with radiographic KOA. Similarly, if at least one intervertebral joint of the lumbar spine was graded as KL2 or higher, the participant was diagnosed with radiographic LS.

BMD measurement

In the Wakayama-1, Wakayama-2, and Hiroshima cohorts, BMD of the lumbar spine and proximal femur was measured using dual energy X-ray absorptiometry (DXA) (Hologic Discovery; Hologic, Waltham, MA, USA) during the baseline examination.

OP was defined on the basis of the World Health Organization (WHO) criteria. Specifically, OP was diagnosed when the BMD T scores were lower than the mean lumbar peak bone mass—2.5 SDs [11]. In Japan, the mean BMD of the L2–L4 vertebrae among both young male and female adults has been measured using Hologic DXA [12]. In the present study, lumbar spine BMD < 0.714 g/cm^2 (for both men and women) and femoral neck BMD < 0.546 g/cm^2 (men) or <0.515 g/cm^2 (women) were considered to indicate OP.

Statistical analysis

All statistical analyses were performed using STATA statistical software (STATA Corp., College Station, TX, USA). Differences in proportions were compared using the Chi square test. Differences in continuous variables were tested for significance using analysis of variance for comparisons among multiple groups or Scheffe's least significant difference test for pairs of groups. To test the association between the interaction between the knee pain and lumbar pain, a logistic regression model was used. First, the presence of knee pain was used as an objective variable (0: absence, 1: presence) and age (+1 year), gender (men vs. women), BMI (+1 kg/m^2), regional differences (0: rural areas including Wakayama-1, Wakayama-2, Niigata, Mie, Akita, and Gunma vs. 1: urban areas including Tokyo-1, Tokyo-2, and Hiroshima), and lumbar pain (0: no, 1: yes) were used as explanatory variables. Then, lumbar pain was used as an objective variable, and knee pain was used as an explanatory variable in the identical model. All p values and 95 % confidence intervals (CI) of two-sided analysis are presented.

Results

Table 2 shows the number of participants classified by age and gender. Most participants were aged $\geq$60 years, and

Table 2 Numbers of participants in the LOCOMO study classified by age and gender

Age strata (years)	Total (%)	Men (%)	Women (%)
≤39	125 (1.0)	49 (1.2)	76 (0.9)
40–49	483 (4.0)	183 (4.6)	300 (3.7)
50–59	963 (8.0)	320 (8.1)	643 (8.0)
60–69	3,170 (26.3)	1,161 (29.3)	2,009 (24.9)
70–79	5,041 (41.9)	1,573 (39.7)	3,468 (43.0)
80–89	2,111 (17.6)	627 (15.8)	1,484 (18.4)
≥90	126 (1.1)	46 (1.2)	80 (1.0)
Total	12,019 (100.0)	3,959 (100.0)	8,060 (100.0)

99.0 % of the participants were aged ≥40 years. Two-thirds of the participants were women, and their mean age was 1 year greater than that of the male participants.

Selected characteristics of the study populations, including age, height, weight, BMI, and proportions of participants who smoked and consumed alcohol are shown in Table 3. The participants were considered as smokers and alcohol consumers if they answered 'yes' to the question 'Are you currently smoking/drinking?' in the self-administered questionnaire. The mean values of age and BMI were significantly higher in women than in men ($p < 0.01$). The proportions of both current smokers and alcohol consumers were significantly higher among men than among women ($p < 0.001$).

Table 3 Baseline characteristics of participants in the LOCOMO study classified by age and gender

Variables	Men	Women	*p* Value (men vs. women)
Age (years)	70.0 (10.6)	71.0 (10.3)	<0.001
Height (cm)	161.1 (6.8)	148.5 (6.4)	<0.001
Weight (kg)	59.3 (9.5)	50.8 (8.6)	<0.001
BMI (kg/m^2)	22.8 (3.0)	23.0 (3.5)	0.007
Smoking (%)	34.0	4.8	<0.001
Drinking (%)	52.4	21.1	<0.001

Values are represented as mean (standard deviation)

BMI body mass index

By analysing the data at the baseline examination, we determined the prevalence of knee pain and lumbar pain. Figure 3 shows the age-sex distribution of the prevalence of knee pain and lumbar pain. Overall, the prevalence of knee pain was 32.7 % (27.9 % in men and 35.1 % in women) and that of lumbar pain was 37.7 % (34.2 % in men and 39.4 % in women). The prevalence of pain in both the knee and lumbar region were significantly higher in women than in men ($p < 0.001$). On the basis of the total age and sex distributions derived from the Japanese census in 2010 [13], our results estimate that 18,000,000 people (7,100,000 men and 10,900,000 women) aged ≥40 years would be affected by knee pain and that 27,700,000 people (12,100,000 men and 15,600,000 women) aged ≥40 years would be affected by lumbar pain.

Further, among 9,046 individuals who were surveyed on both knee pain and lumbar pain at the baseline examination in each cohort, the prevalence of both knee pain and lumbar pain was 12.2 % (10.9 % in men and 12.8 % in women). The prevalence of the coexistence of knee and lumbar pain in the participants aged <40, 40–49, 50–59, 60–69, 70–79, and ≥80 years was 4.0, 4.8, 7.4, 13.0, 13.3, and 11.7 %, respectively, (6.1, 5.3, 6.0, 10.0, 11.5, and 13.2 %, respectively, in men and 2.6, 4.6, 8.1, 14.8, 14.2, and 11.0 %, respectively, in women). The prevalence of both knee pain and lumbar pain increased with age in men, whereas that in women reached a plateau at 60–69 and 70–79 years and then declined. On the basis of the total age and sex distributions derived from the Japanese census in 2010 [13], our results estimate that 6,800,000 people

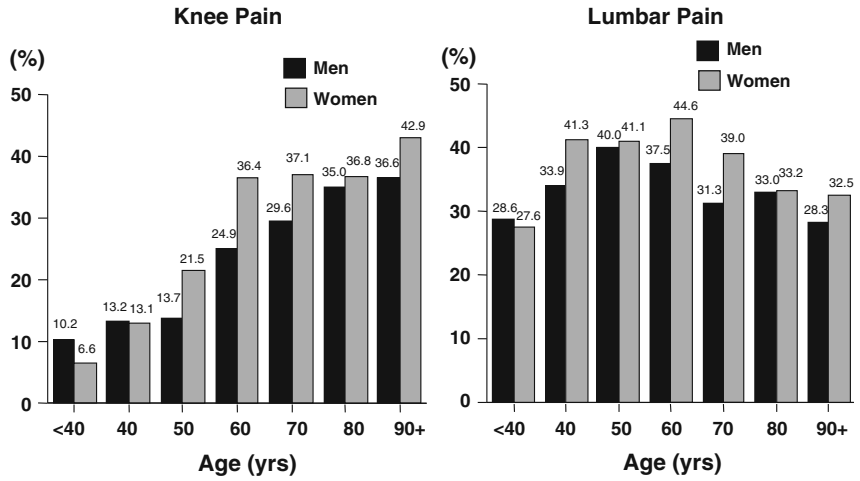


Fig. 3 Prevalence of knee pain and lumbar pain according to age and gender

Table 4 Odds ratios (OR) of potentially associated factors for the presence of knee pain/ lumbar pain vs. absence of pain

Explanatory variables	Reference	OR	95% confident interval	*p*
Knee pain (presence vs. absence)				
Age (years)	+1 year	1.045	1.039–1.051	<0.001***
Gender	0: men, 1: women	1.602	1.441–1.780	<0.001***
Region	0: urban area, 1: rural area	2.419	2.152–2.720	<0.001***
BMI (kg/m^2)	+1 kg/m^2	1.141	1.124–1.158	<0.001***
Lumbar pain	0: absence, 1: presence	1.373	1.243–1.515	<0.001***
Lumbar pain (presence vs. absence)				
Age (years)	+1 year	1.018	1.013–1.023	<0.001***
Gender	0: men, 1: women	1.130	1.023–1.248	0.016*
Region	0: urban area, 1: rural area	2.016	1.801–2.256	<0.001***
BMI (kg/m^2)	+1 kg/m^2	1.020	1.003–1.031	0.021*
Knee pain	0: absence, 1: presence	1.375	1.246–1.518	<0.001***

BMI body mass index

* $p < 0.05$, *** $p < 0.001$

(2,800,000 men and 4,000,000 women) aged ≥40 years would be affected by both knee pain and lumbar pain.

To test the association between the knee pain and lumbar pain, the presence of knee pain was first used as an objective variable (0: absence, 1: presence) and age (+1 year), gender (men vs. women), BMI (+1 kg/m^2), regional differences (0: rural areas including Wakayama-1, Wakayama-2, Niigata, Mie, Akita, and Gunma vs. 1: urban areas including Tokyo-1, Tokyo-2, and Hiroshima), and lumbar pain (0: no, 1: yes) were used as explanatory variables. Then, the presence of lumbar pain was used as an objective variable (0: absence, 1: presence) and age (+1 year), gender (men vs. women), BMI (+1 kg/m^2), regional differences (0: rural areas including Wakayama-1, Wakayama-2, Niigata, Mie, Akita, and Gunma vs. 1: urban areas including Tokyo-1, Tokyo-2, and Hiroshima), and knee pain (0: no, 1: yes) were used as explanatory variables. Table 4 shows the result of the logistic regression analysis. Higher age, female sex, higher BMI, living in a rural area, and the presence of lumbar pain significantly influenced the presence of knee pain. Similarly, higher age, female sex, higher BMI, living in a rural area, and the presence of knee pain significantly influenced the presence of lumbar pain.

Discussion

In the present study, we integrated the information of individual cohorts established for the prevention of musculoskeletal diseases, and created the nationwide large-scale cohorts comprising the LOCOMO study. By using the data of the LOCOMO study, we found that the prevalence of knee pain was 32.7 % and that of lumbar pain was 37.7 %. Both knee pain and lumbar pain were prevalent in 12.2 % of the total population. In the present study, we also clarified that the factors associated with knee or lumbar pain were age, sex, body build, and residential characteristics. In addition, the presence of knee pain affected the lumbar pain, and vice versa. This association remained even after the adjustment for the above-mentioned associated factors. To our knowledge, this is the first study to report the frequency of the knee pain and lumbar pain and to estimate the total number of prevalent subjects, by using a large-scale population-based cohort study in Japan.

With regard to musculoskeletal pain, several population-based epidemiological studies have demonstrated that chronic pain is a highly prevalent condition. Soni et al. [14] reported that the prevalence rates of self-reported knee pain using the baseline data in 1,003 participants from the Chingford Women's Study were 22.97 % in the left knee and 24.80 % in the right knee. The definition of the presence of the knee pain (based on the following two questions: 'Have you had any knee pain in either knee in the last month?' and 'How many days of pain have you experienced in the last month?') was similar but not identical to our definition used in the LOCOMO study, and the subjects' age was younger in the Chinford study than in the LOCOMO study. Therefore, we could not compare the prevalence between the Chinford and LOCOMO studies directly. However, at a glance, the prevalence seems to be higher in the Japanese population. This may be due to the fact that the prevalence of KOA (KL grades ≥2) was higher in the Japanese population than that in the Caucasian population [15]. Verhaak et al. [16] reviewed epidemiological studies on chronic benign pain among adults, including subjects aged between 18 and 75 years, and reported that the prevalence ranged between 2 and 40 % of the population. Coggon et al. did not perform a population-based study, but instead conducted a cross-sectional survey comparing the prevalence of disabling low back pain and disabling wrist/hand pain among groups of workers carrying out similar physical activities in different cultural environments in 18 countries including Japan. They

reported that the 1-month prevalence of disabling low back pain in nurses ranged from 9.6 to 42.6 %, and that of disabling wrist/hand pain in office workers ranged from 2.2 to 31.6 % [17]. We could not compare our results to those of Coggon's results directly because of the difference in the characteristics of the targeted population. However, previous reviews and reports demonstrated that the prevalence of the chronic pain varied in the population surveyed, and therefore, estimating the prevalence and number of patients in pain would require a study that comprises various regions with a large number of subjects. Our LOCOMO study contains 12,019 participants from the cohorts consisting of nine communities in different locations in Japan. Therefore, we believe that our estimation of the prevalence of knee pain and lumbar pain is appropriate, and the number of patients was sufficient.

With regard to the characteristics of subjects with chronic pain, Soni et al. [14] reported that among subjects who could be followed up for 12 years, a higher BMI was predictive of persistent knee pain (odds ratio = 1.14) and incident knee pain (odds ratio = 1.10). Verhaak et al. [16] demonstrated that chronic pain generally increased with age, with some studies reporting a peak prevalence between the ages of 45 and 65 years. These results were not consistent with our results. Moreover, we noted that living in a rural area was associated with the presence of knee pain and lumbar pain, which may be due to the difference of the primary occupation in that area. Muraki et al. [18] reported that the presence of KOA and LS was influenced by the primary occupation of the participants. According to their report, the prevalence of higher K/L grades of KOA and LS was significantly higher among agricultural, forestry, and fishery workers than among clerical workers and technical experts [18]. For occupational activities, sitting on a chair had a significant inverse association with K/L grades $\geq$2 for KOA and LS, whereas standing, walking, climbing and heavy lifting were associated with higher K/L grades for KOA [18]. An association between occupational activities and KOA was also observed in several studies [19–21]. Agricultural, forestry, and fishery workers seemed to be more common in rural areas than in urban areas. In addition, occupational activities, such as sitting on a chair, might be observed more commonly in clerical workers than in agricultural, forestry, and fishery workers. These findings might support the regional differences of pain that were observed in the present study. The main focus of the present study was pain, and not OA; however, the most probable diagnosis underlying knee pain among older people was reported to be OA [22].

There are also several reports regarding the coexistence of pain. The above-mentioned Coggon's investigation indicated that the rates of disabling pain at 2 anatomical sites—the lumbar spine and wrist/hand—covaried ($r = 0.76$) [17]. In their cross-sectional study, Smith et al., examined the presence and sites of chronic pain in 11,797 women. The presence of chronic pain was noted in 38 % of women; among them, the percentage of women experiencing chronic pain at 1, 2, 3, 4, and $\geq$5 sites was 23.2, 24.4, 20.0, 14.3, and 18.2 %, respectively [23]. These results showed that chronic pain coexists at other anatomical sites. In the present study, the prevalence of both knee pain and lumbar pain was 12.2 % (10.9 % in men and 12.8 % in women) among the general population. However, among the subjects with lumbar pain, 37.3 % also had knee pain (39.0 % in men and 36.6 % in women). Unfortunately, in the LOCOMO study, we were unable to collect the data regarding pain at anatomical sites other than knee pain and lumbar pain. Nevertheless, the coexistence of pain was commonly noted, which is inconsistent with previous reports.

There were several limitations in the present study. First, the current subjects do not truly represent the entire Japanese population. We should carefully consider this limitation, especially when determining the generalisability of the results. However, the LOCOMO study is the first large-scale population-based prospective study with more than 12,000 participants. Although it does not comprise the whole population of Japan, the number of participants in the cohorts established for the prevention of the musculoskeletal diseases appears to be biggest worldwide. Second, all the items of our survey in the baseline examination were not recorded in all cohorts. For example, radiographic examination of knees was performed only in Tokyo-1, Wakayama-1, Wakayama-2, Niigata, and Mie prefectures and radiographic examination of the lumbar spine was performed only in Tokyo-1, Wakayama-1, Wakayama-2, Hiroshima, and Mie prefectures. Third, the radiographic findings for OA assessment using KL scales have not been integrated yet, because of the delay in the standardisation of reading methods of the observers. Radiographs should be assessed by a single observer to omit the inter-observer variability, and if this is impossible, then the inter-observer variability among observers should be tested using the standardised criteria. Therefore, in the present study, we could not evaluate the severity of knee/spinal OA or vertebral fractures for assessing knee pain and lumbar pain. After suitable evaluation of intra-observer and inter-observer variability in the assessment of radiography findings and integration of this information, we hope to re-analyse the factors associated with the presence of chronic pain. Moreover, not only OA and fractures, but also rheumatoid arthritis and spondyloarthritis should be considered as parameters for assessing knee pain and lumbar pain. Although collection of the information on the diagnosis may be difficult on a large scale due to the associated cost, it may be possible to obtain this information in at least two cohorts.

In addition, our study has several strengths. First, as mentioned above, the large number of the integrated subjects included in the LOCOMO study is the biggest strength of this study. Moreover, we collected data from nine cohorts across Japan. By using the data of the LOCOMO study, we could compare the regional differences of specific clinical symptoms such as knee pain or lumbar pain, or particular diseases, such as KOA, LS, or OP, as well as its prognosis, such as the incidence of disability or mortality. In particular, we identified regional differences in the prevalences of knee pain and lumbar pain. In addition, we collected a substantial amount of information, via an interviewer-administered questionnaire, dietary assessment, anthropometric measurements, neuromuscular function assessment, biochemical measurements, medical history recording, radiographic assessment, and BMD measurement. However, all items were not recorded in all cohorts and the regional selection bias in each examination should be considered when interpreting the results.

In summary, by using the data of the LOCOMO study, we clarified the prevalence of knee pain and lumbar pain, their coexistence, and their associated factors.

Acknowledgments This work was supported by grants from Grant-in-Aid for H17-Men-eki-009 (Director, Kozo Nakamura), H20-Choujyu-009 (Director, Noriko Yoshimura), H23-Choujyu-002 (Director, Toru Akune), and H-25-Chojyu-007 (Director, Noriko Yoshimura) of the Ministry of Health, Labour and Welfare; and Scientific Research B23390172, B20390182, and Challenging Exploratory Research 24659317 to Noriko Yoshimura, B23390357 and C20591737 to Toru Akune, and B23390356 and C20591774, Challenging Exploratory Research 23659580 to Shigeyuki Muraki, Challenging Exploratory Research 24659666, 21659349 and Young Scientists A18689031 to Hiroyuki Oka, and Collaborating Research with NSF 08033011-00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. The sponsors did not contribute to the study design, data collection, data analysis, data interpretation, or the writing of the report.

Conflict of interest The authors wish to thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, and Mrs. Saeko Sahara for their assistance with data reduction and administration.

References

1. Ministry of Health, Labour and Welfare. Outline of the results of National Livelihood Survey 2010. http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/4-2.html
2. Muraki S, Yamamoto S, Ishibashi H, Nakamura K (2006) Factors associated with mortality following hip fracture in Japan. J Bone Miner Metab 24:100–104
3. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Yamamoto S, Ishibashi H, Kawaguchi H, Nakamura K, Akune T (2009) Prevalence of knee osteoarthritis, LS and osteoporosis in Japanese men and women: the Research on Osteoarthritis/Osteoporosis Against Disability study. J Bone Miner Metab 27:620–628
4. Yoshimura N, Nakamura K, Akune T, Fujiwara S, Shimizu Y, Yoshida H, Omori G, Sudo A, Nishiwaki Y, Yoshida M, Shimokata H (2013) The longitudinal cohorts of motor system organ (LOCOMO) study (In Japanese). Nippon Rinsho 71:642–645
5. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, Lewis C, Cawthon PM, Marcus R, Marshall LM, McGowan J, Phipps K, Sherman S, Stefanick ML, Stone K (2005) Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study: a large observational study of the determinants of fracture in older men. Contemp Clin Trials 26:569–585
6. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW (1998) Validation study of WOMAC. A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheum 15:1833–1840
7. Johnson JA, Coons SJ, Ergo A, Szava-Kovats G (1998) Valuation of EuroQOL (EQ-5D) health states in an adult US sample. Pharmacoeconomics 13:421–433
8. Fukuhara S, Suzukao Y (2004) Manual of the SF-8 Japanese version. Institute for Health Outcomes & Process Evaluation Research, Kyoto
9. Teng EL, Chui HC (1987) The Modified Mini-Mental State (3MS) examination. J Clin Psychiatry 48:314–318
10. Kellgren JH, Lawrence LS (1957) Radiological assessment of osteoarthrosis. Ann Rheum Dis 16:494–502
11. World Health Organization (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. WHO Technical Report Series 843, Geneva
12. Orimo H, Hayashi Y, Fukunaga M, Sone T, Fujiwara S, Shiraki M, Kushida K, Miyamoto S, Soen S, Nishimura J, Oh-Hashi Y, Hosoi T, Gorai I, Tanaka H, Igai T, Kishimoto H, Osteoporosis Diagnostic Criteria Review Committee: Japanese Society for Bone and Mineral Research (2001) Diagnostic criteria for primary osteoporosis: year 2000 revision. J Bone Miner Metab 19:331–337
13. Portal site of Official Statistics of Japan. Population Census 2010. http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&tclassID=000001034991&cycleCode=0&requestSender=search
14. Soni A, Kiran A, Hart DJ, Leyland KM, Goulston L, Cooper C, Javaid MK, Spector TD, Arden NK (2012) Prevalence of reported knee pain over twelve years in a community-based cohort. Arthritis Rheum 64:1145–1152
15. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, Saika A, Suzuki T, Yoshida H, Ishibashi H, Yamamoto S, Nakamura K, Kawaguchi H, Yoshimura N (2009) Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthr Cartil 17:1137–1143
16. Verhaak PF, Kerssens JJ, Dekker J, Sorbi MJ, Bensing JM (1998) Prevalence of chronic benign pain disorder among adults: a review of the literature. Pain 77:231–239
17. Coggon D, Ntani G, Palmer KT, Felli VE, Harari R et al (2013) Disabling musculoskeletal pain in working populations: is it the job, the person, or the culture? Pain 154:856–863
18. Muraki S, Akune T, Oka H, Mabuchi A, En-yo Y, Yoshida M, Saika A, Nakamura K, Kawaguchi H, Yoshimura N (2009) Association of occupational activity with radiographic knee osteoarthritis and LS in elderly patients of population-based cohorts: a large-scale population-based study. Arthr Rheum 61:779–786
19. Coggon D, Croft P, Kellingray S, Barrett D, McLaren M, Cooper C (2000) Occupational physical activities and osteoarthritis of the knee. Arthr Rheum 43:1443–1449

20. Felson DT (2004) An update on the pathogenesis and epidemiology of osteoarthritis. Radiol Clin North Am 42:1–9
21. Jensen LK (2008) Knee osteoarthritis: influence of work involving heavy lifting, kneeling, climbing stairs or ladders, or kneeling/squatting combined with heavy lifting. J Occup Environ Med 65:72–89
22. Duncan R, Peat G, Thomas E, Hay E, McCall I, Croft P (2007) Symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? Ann Rheum Dis 66:86–91
23. Smith BH, Elliott AM, Hannaford PC, Royal College of General Practitioners' Oral Contraception Study (2004) Is chronic pain a distinct diagnosis in primary care? Evidence arising from the Royal College of General Practitioners' Oral Contraception Study. Fam Pract 21:66–74

ELSEVIER

CrossMark

The Spine Journal 14 (2014) 2811–2817

THE SPINE JOURNAL

Clinical Study

The prevalence of cervical myelopathy among subjects with narrow cervical spinal canal in a population-based magnetic resonance imaging study: the Wakayama Spine Study

Keiji Nagata, MD, PhD[a], Noriko Yoshimura, MD, PhD[b], Hiroshi Hashizume, MD, PhD[a,*], Shigeyuki Muraki, MD, PhD[c], Yuyu Ishimoto, MD, PhD[a], Hiroshi Yamada, MD, PhD[a], Noboru Takiguchi, MD, PhD[a], Yukihiro Nakagawa, MD, PhD[a], Akihito Minamide, MD, PhD[a], Hiroyuki Oka, MD[b], Hiroshi Kawaguchi, MD, PhD[d], Kozo Nakamura, MD, PhD[e], Toru Akune, MD, PhD[c], Munehito Yoshida, MD, PhD[a]

[a]*Department of Orthopaedic Surgery, Wakayama Medical University, 811-1, Kimiidera, Wakayama City, 641-8510, Japan*
[b]*Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8654, Japan*
[c]*Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8654, Japan*
[d]*Japan Community Health Care Organization Tokyo Shinjuku Medical Center, 6-1-1 Shinjuku, Shinjuku-ku, Tokyo 160-8402, Japan*
[e]*Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, 1, Namiki 4-chome, Tokorozawa City, Saitama Prefecture 359-8555, Japan*

Received 14 August 2013; revised 26 December 2013; accepted 18 March 2014

Abstract

BACKGROUND CONTEXT: A narrow cervical spinal canal (CSC) is a well-known risk factor for cervical myelopathy (CM). However, no epidemiologic data of the CSC based on a population-based cohort are available.

PURPOSE: The purpose of the study was to investigate the age-related differences in CSC diameters on plain radiographs and to examine the associated magnetic resonance imaging (MRI) abnormalities including cervical cord compression and increased signal intensity (ISI) as well as the clinical CM with the narrow CSC.

STUDY DESIGN/SETTING: This was a cross-sectional study.

PARTICIPANT SAMPLE: Data were obtained from the baseline survey of the Wakayama Spine Study that was performed from 2008 to 2010 in a western part of Japan. Finally, a total of 959 subjects (319 men and 640 women; mean age, 66.4 years) were included.

OUTCOME MEASURES: The outcome measures included in the study were the CSC diameter at C5 level on plain radiographs, cervical cord compression and ISI on sagittal T2-weighted MRI, and physical signs related to CM (eg, the Hoffmann reflex, hyperreflexia of the patellar tendon, the Babinski reflex, sensory and motor function, and bowel/bladder symptoms).

FDA device/drug status: Not applicable.

Author disclosures: ***KNag:*** Grant: Grant No.255 for young investigator from Japan Orthopedics and Traumatology Foundation (B, Paid directly to institution). ***NY:*** Grant: Grant-in-Aid for Scientific Research B23390172 and B20390182 (E, Paid directly to institution), Collaborating Research with NSF 08033011-00262 from the Ministry of Education, Culture, Sports, Science and Technology (E, Paid directly to institution). ***HH:*** Grant: The consigned research fund of Wakayama Prefecture Nos. B-23004 and B-24001 (C, Paid directly to institution). ***SM:*** Grant: Grant-in-Aid for Scientific Research C20591774 (D, Paid directly to institution). ***YI:*** Nothing to disclose. ***HY:*** Grant: Grant-in-Aid for Scientific Research C22591639 and C25462305 (E: Paid directly to institution). ***NT:*** Nothing to disclose. ***YN:*** Nothing to disclose. ***AM:*** Nothing to disclose. ***HO:*** Grant: Grant-in-Aid for Scientific Research for young scientists A18689031 (D, Paid directly to institution). ***HK:*** Nothing to disclose. ***KNak:*** Nothing to disclose. ***TA:*** Grant: Grant-in-Aid for Scientific Research C20591737 and H23-Chojyu-002 from the Ministry of Health, Labour and Welfare in Japan (E, Paid directly to institution). ***MY:*** Nothing to disclose.

The disclosure key can be found on the Table of Contents and at www.TheSpineJournalOnline.com.

Conflicts of interest and source of funding: No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

* Corresponding author. Department of Orthopaedic Surgery, Wakayama Medical University, 811-1 Kimidera, Wakayama City, Wakayama 641-8510, Japan. Tel.: +81-73-441-0645; fax.: +81-73-448-3008.

E-mail address: hashizum@wakayama-med.ac.jp (H. Hashizume)

http://dx.doi.org/10.1016/j.spinee.2014.03.051

METHODS: The age-related differences of CSC diameters in men and women were investigated by descriptive statistics. The prevalence of MRI abnormalities and clinical CM was compared among the groups divided by the CSC diameter (less than 13, 13–15, and 15 mm or more). In addition, a logistic regression analysis was performed to determine the association of the CSC diameter with cervical cord compression/clinical CM after overall adjustment for age, sex, and body mass index.
RESULTS: The CSC diameter was narrower with increasing age in both men and women. The prevalence of cervical cord compression, ISI, and the clinical CM was significantly higher in the narrower CSC group. The prevalence of cervical cord compression, ISI, and CM among subjects with CSC diameter less than 13 mm was 38.0%, 5.4%, and 10.1%, respectively. In the logistic model, the CSC diameter was a significant predictive factor for the clinical CM ($p<.0001$).
CONCLUSIONS: This study firstly confirmed the age-related differences in CSC diameters and the significant association of the narrow CSC diameter with CM in a population-based cohort.

Keywords: Cervical spine; Spinal canal stenosis; Cervical myelopathy; Magnetic resonance imaging; Population-based cohort; Epidemiology

Introduction

In cervical spinal disorders such as cervical myelopathy (CM) and spinal cord injury, developmental cervical spinal canal (CSC) stenosis has been considered as an effective predictor of clinical outcome [1,2]. The spinal cord area should be evaluated after comparing with data obtained from asymptomatic subjects of each age group. Age-dependent data are required because the spinal cord may change with age, just as the cerebrum decreases in size with age in elderly subjects. The spinal canal should also be considered in asymptomatic subjects when treating cervical spinal disorders because patients with a tight spinal canal are more susceptible to spinal cord damage. However, the prevalence of spinal cord disorders and CM among patients with CSC of narrow diameter is not known. To date, few studies have focused on age-related differences in the cervical spinal cord and CSC [3,4]. Recent advances in magnetic resonance imaging (MRI) have made it possible to noninvasively obtain clear images of the cervical spinal cord, thereby making evaluation of traumatic spinal cord injury and cervical cord compression more applicable in routine practice. This study was undertaken to clarify age-related differences in the cervical spinal cord and CSC using magnetic resonance imaging (MRI) to establish the basis for morphometric evaluation of patients with cervical spinal cord disorders. More specifically, the purposes of this study were to investigate age-related changes of the CSC in a population-based cohort in Japan and to examine the associated MRI abnormalities including cervical cord compression and increased signal intensity (ISI) as well as the clinical CM with the narrow CSC diameters.

Participants and methods

Participants

The present study is a part of ''The Wakayama Spine Study: a population-based cohort,'' which was a large-scale population-based MRI study. Because a detailed profile of the Wakayama Spine Study has already been described elsewhere, only a brief summary is provided here [5,6]. The Wakayama Spine Study was conducted between 2008 and 2010 in a mountainous region in Hidakagawa, Wakayama, and a coastal region in Taiji, Wakayama. From inhabitants of the Hidakagawa and Taiji regions, 1,063 potential study subjects were recruited for MRI examinations. Among those 1,063 candidates, 52 declined the examination; therefore, 1,011 inhabitants were registered in the present study. Among those 1,011 participants, individuals with MRI-sensitive implanted devices (such as a pacemaker) and other disqualifiers were excluded. Ultimately, the cervical spine was scanned with MRI in 985 participants. Four participants who had undergone a previous cervical operation were excluded from the analysis, and another four participants whose MRI interpretation was difficult because of poor image quality were also excluded. After these exclusions, the present study had 977 participants. Radiographic evaluation of the cervical spine was also performed in 959 of the subjects. In total, both MRI and radiographic results were available for 959 participants (319 men and 640 women) with an age range of 21 to 97 years (mean, 67.3 years for men and 65.9 years for women). The participants completed an interviewer-administered questionnaire of 400 items that included lifestyle information; and anthropometric and physical performance measurements were taken. All study participants provided informed consent, and the study design was approved by the appropriate ethics review boards.

Anthropometric measurements included height (meter), weight (kilogram), and body mass index (BMI; weight [kilogram]/height^2 [m^2]). Medical information concerning neck pain, sensory disturbances, the Hoffmann reflex, the Babinski reflex, and the deep tendon reflex of the patellar tendon was gathered by an experienced orthopedic surgeon. The Hoffmann reflex was elicited with the hand in a neutral position by flicking the distal phalanx of the middle finger and observing flexion of the distal phalanx of the thumb [7,8].

The Babinski reflex was elicited by firmly sweeping from the lateral part of the sole to the base of the toes with a pointed end of a reflex hammer and observing the hallux extensor response [9,10]. Hyperreflexia of the patellar tendon, a positive Hoffmann reflex, and a positive Babinski reflex were defined as aggravation on both sides. A myelopathic sign was defined as the presence of hyperreflexia of the patellar tendon, Hoffmann reflex, or Babinski reflex.

Measurements of CSC diameter and canal-to-body ratio on radiographs

All subjects also underwent lateral radiography with their neck in the neutral position. They were told by an X-ray technician to look straight ahead in a relaxed position. The radiographic data were scanned and calibrated using the ruler, which was put on the film. The sagittal spinal canal diameter at the C5 level was measured as the shortest distance from the midpoint between the vertebral body's superior and inferior end plates to the spinolaminar line. The canal-to-body ratio (CBR) was obtained by dividing the diameter of the spinal canal by that of the vertebral body to assess the tightness of the spinal canal and also to eliminate the magnification effect of radiographs.

Magnetic resonance imaging

An MRI scan of the cervical spine was obtained for each participant using a 1.5-T Excelart imaging system (Toshiba, Tokyo, Japan). All scans were taken in the supine position, except for participants with a rounded back, who used a triangular pillow under their heads and knees. The imaging protocol included a sagittal T2-weighted fast spin-echo pulse sequence (repetition time: 4,000 ms; echo time: 120 ms; and field of view: 300 × 320 mm) and an axial T2-weighted fast spin-echo pulse sequence (repetition time: 4,000 ms; echo time: 120 ms; and field of view: 180 × 180 mm).

MRI measures

Midsagittal T2-weighted images were assessed by an experienced orthopedic surgeon (Keiji Nagata), who was blinded to participants' clinical status.

Evaluation of cervical cord compression

Cervical cord compression was defined as compression with an anterior and/or a posterior component of the spinal cord [6]. Cervical cord compression was evaluated at each intervertebral level from C2–C3 to C7–T1.

Evaluation of signal intensity of the spinal cord

Increased signal intensity was defined as a high-intensity area in contrast with the adjacent isointensity portion of the spinal cord [11]. The ISI was evaluated in the area from C2 to T1.

Measurement of spinal cord diameter

The spinal cord diameter was measured manually at the midpoint of the C5 vertebral body level using the imaging software OsiriX (http://www.osirix-viewer.com/).

Definition of clinical CM

Myelopathy is defined clinically by the presence of myelopathic signs (eg, the Hoffmann reflex, hyperreflexia of the patellar tendon, and the Babinski reflex), usually accompanied by bilateral sensory deficits or sensory level and bowel/bladder symptoms. Among participants with myelopathic signs, cervical cord compression was the essential condition for diagnosing CM.

Statistical analyses

A comparison of baseline characteristics between sexes was performed using the Student *t* test. Differences in the CSC diameter, vertebral body, spinal cord, and CBR among men and women were determined using the Student *t* test. One-way analysis of variance was used to evaluate the differences in CSC diameter, vertebral body, spinal cord, and CBR among different age groups. The chi-square test was used to assess the presence of ISI among different age groups.

For categorical data, the chi-square test was used to assess the presence of significant differences among different diameters of the CSC. For continuous outcomes, the analysis of variance test was used to assess differences among different diameters of the CSC. In addition, to determine the association of ISI, CSC diameter, and CBR with cervical cord compression and CM, logistic regression analysis was used after overall adjustment for age, sex, and BMI. All statistical tests were performed at a significance level of .05 (two-sided). Data analyses were performed using JMP, version 8 (SAS Institute, Inc, Cary, NC, USA).

Results

Characteristics of the participants

The baseline characteristics of the 977 participants, including data for anthropometric measurements and physical performance, are listed in Table 1. There was no

Table 1
Characteristics of men and women participating in the present study

Characteristic	Men	Women
N	319	640
Age, y	67.3±13.8	65.9±13.3
Height, cm	164.6±7.2**	151.6±7.2
Weight, kg	64.4±11.6**	53.0±9.4
Body mass index, kg/m^2	23.7±3.4*	23.1±3.7
Grip strength, kg	37.9±9.1**	23.9±5.9

Note: Significantly different from women by the Student *t* test (*p<.01; **p<.001).

Values are the mean±standard deviation.

Table 2
Radiographic and MRI measures stratified by gender and age strata

	Radiographic measures		MRI measures	
Age strata	Diameter of cervical spinal canal (mm)	Canal-to-body ratio	Increased signal intensity, N (%)	Diameter of spinal cord (mm)
Men				
Overall	14.8±1.3	0.82±0.12	15 (4.6)	6.9±0.9
<50 y	15.2±1.3	0.86±0.09	2 (5.2)	7.3±0.8
50–59 y	14.8±1.7	0.85±0.13	4 (6.9)	7.1±0.9
60–69 y	14.9±1.2	0.82±0.11	3 (4.5)	6.9±0.7
70–79 y	14.8±1.2	0.82±0.11	2 (2.3)	6.9±0.8
≥80 y	14.4±1.1	0.79±0.12	4 (5.5)	6.6±0.9
Women				
Overall	14.1±1.2	0.92±0.13	11 (1.7)	6.8±0.9
<50 y	14.5±1.3	0.99±0.14	1 (1.1)	6.9±0.9
50–59 y	14.4±1.3	0.96±0.12	1 (0.0)	7.0±0.7
60–69 y	14.1±1.1	0.91±0.12	0 (0)	6.8±0.8
70–79 y	13.9±1.1	0.89±0.12	6 (3.5)	6.8±0.9
≥80 y	13.8±1.0	0.86±0.12	3 (2.5)	6.7±0.9

Note: Otherwise indicated, values are mean±standard deviation for each age strata in men and women.

significant difference in age between sexes. Height, weight, and BMI were significantly higher in men than in women.

Age and sex differences of CSC diameter, CBR, ISI, and spinal cord diameter

Table 2 lists the age-related differences in diameters of the CSC, the CBR on radiograph, ISI, and spinal cord diameter on MRI in men and women among different age groups. The CSC diameter was significantly narrower with age in women ($p<.0001$). In men, the CSC diameter had a tendency to be narrower with age, but it was not significantly different in women. The mean diameter of the CSC was not significantly different between men and women. The diameter of the vertebral body was significantly higher in men and women with increasing age ($p<.0001$). The mean CBR in men and women was 0.82 and 0.92, respectively, and it was significantly higher in women than in men at the C5 vertebral level. The CBR was significantly lower with increasing age in both sexes (men, $p=.0004$; women: $p<.0001$).

The prevalence of ISI in all participants was 2.7% (4.6% in men and 1.7% in women) and was significantly higher in men than in women ($p=.007$). The prevalence of ISI was not significantly different with age between sexes. The diameter of the spinal cord was significantly lower with increasing age in both sexes (men, $p=.0012$; women, $p=.0068$). The mean diameter of the spinal cord was not significantly different between men and women.

Prevalence of MRI measures and CM among different diameters of the CSC

Anthropometric measures such as CSC diameter were found to be significantly different according to age (Table 3). Regarding MRI measures, significant differences between different CSC diameters were found with respect to cervical cord compression ($p<.0001$), ISI ($p<.0001$), and spinal cord diameter ($p<.0001$), except for ISI in women. The prevalence of cervical cord compression, ISI, and CM in subjects with a CSC diameter less than 13 mm was 61.9%, 23.8%, and 4.8% in men, respectively. Meanwhile, the prevalence of cervical cord compression, ISI, and CM among female subjects with a CSC diameter less than 13 mm was 33.3%, 1.9%, and 11.1%, respectively. Multiple logistic regression analysis was performed to estimate the predictive factors for CM in MRI and radiographic measurements after adjustment for age, sex, and BMI (Table 4). As an overall result, ISI, CSC diameter, and CBR were significant predictive factors for CM ($p<.01$). There was a positive association between cervical cord compression and spinal cord diameter, whereas spinal cord diameter itself was not a significant predictive factor for CM.

Discussion

The present study is the first population-based study to clarify the normal value of the diameter of the CSC and its association with cervical cord compression, ISI, and CM in Japanese men and women. We clarified that the CSC diameter was narrower with age in both men and women in the population-based cohort. The prevalence of the clinical CM was significantly higher in the narrower CSC group. Furthermore, in the logistic model, the CSC diameter was a significant predictive factor for clinical CM.

In this study, the CSC and vertebral body diameters were measured using plain radiographs because the posterior longitudinal ligament could not be distinguished from the vertebral body on MRI. There have been several reports on the diameter of the CSC. Porter et al. [12] reported that canal size did not appear to change significantly with biomechanical stress and aging. Meanwhile, Goto et al. [3] and Kato et al. [4] reported that the younger generation (younger than 40 years of age) had a statistically wider CSC. Our result was consistent with the latter reports. Why do younger persons have a wider CSC than elderly persons? There are two possible reasons for the differences in CSC diameter between generations. First, the CSC diameter becomes narrower with aging. A CSC with a small diameter is primarily a developmental and not a degenerative phenomenon. However, Hukuda and Kojima [13] reported that the diameter of the vertebral body was wider in older people compared with younger people. Those morphologic changes of the vertebral component may affect the diameter of the CSC. Second, the changes in Japanese eating habits and physique in the past few decades may have contributed to the changes in the diameter of the CSC. The variation of CSC diameter with different generation may be a limited phenomenon in Japan. However, we believe the results prompt future investigations into the various factors affecting the CSC dimensions, apart from aging.

Table 3
Prevalence of MRI measures and cervical myelopathy among different diameter of cervical spinal canal

Factors	Diameter of cervical spinal canal (mm)			p value
	<13	13–15	≥15	
Men				
N	21	162	136	
Age, y	69.5±12.0	69.1±12.8	64.9±14.9	.027
Height, cm	163.7±5.6	163.6±6.8	165.8±7.7	.025
Weight, kg	67.6±10.5	61.7±10.6	67.0±12.3	.0002
Body mass index, kg/m^2	25.2±3.3	23.0±3.2	24.2±3.4	.0006
MRI measures				
Cervical cord compression, N (%)	13 (61.9)	58 (35.8)	22 (16.2)	<.0001
Increased signal intensity, N (%)	5 (23.8)	9 (5.6)	1 (0.7)	<.0001
Diameter of spinal cord (mm)	6.2±0.5	6.8±0.8	7.2±0.8	<.0001
Cervical myelopathy, N (%)	1 (4.8)	2 (1.2)	0 (0)	.09
Women				
N	108	383	149	
Age, y	68.7±13.3	67.2±12.8	60.4±13.4	<.0001
Height, cm	149.1±7.2	151.4±7.1	153.9±6.7	<.0001
Weight, kg	49.8±8.1	53.2±9.6	54.8±9.4	.0001
Body mass index, kg/m^2	22.4±3.2	23.2±3.7	23.2±3.8	.12
MRI measures				
Cervical cord compression, N (%)	36 (33.3)	92 (24.0)	11 (7.4)	<.0001
Increased signal intensity, N (%)	2 (1.9)	8 (2.1)	0 (0)	.21
Diameter of spinal cord (mm)	6.5±0.9	6.9±0.8	7.0±0.8	<.0001
Cervical myelopathy, N (%)	12 (11.1)	12 (3.1)	0 (0)	<.0001

MRI, magnetic resonance imaging.

Note: For categorical data, the chi-square test was used to assess the presence of significant differences among different diameters of the cervical spinal canal.

For continuous outcomes, comparison was made by the analysis of variance test differences among different diameters of the cervical spinal canal.

Regarding MRI measurements, the prevalence of cervical cord compression and ISI among persons with a CSC diameter less than 13 mm, which was considered to be developmental canal stenosis [14], was 61.9% and 23.8% in men and 33.3% and 1.9% in women, respectively. Above all, a CSC diameter less than 13 mm was observed in more than 10% of the participants. Of those with a CSC diameter less than 13 mm, cervical cord compression (ie, the preliminary step in the development of CM) was also observed in 61.9% of men and 33.3% of women. From these results, the number of people who have a risk for CM was considered quite high in the general population. Countee and Vijayanathan [15] reported that congenital stenosis in men with a cervical canal diameter of 14 mm or less was associated with quadriplegia after trauma. In the present study, we noted that the narrower the diameter of the CSC, the higher the prevalence of ISI. Of note, the distribution of prevalence between men and women was different. Increased signal intensity was seen in approximately 10% of men younger than 60 years, whereas it was seen in only 1% of women younger than 60 years, and was relatively higher in older people. In the present study, the prevalence of the clinical CM was significantly higher in the narrower CSC group. The result may show that patients with a narrowed spinal canal are more likely to develop CM. Further longitudinal studies are needed to clarify the causal

Table 4
The odds ratio and 95% confidence interval of increased signal intensity, diameter of spinal cord, diameter of cervical spinal canal, and canal-to-body ratio for cervical myelopathy

Variables	Cervical cord compression		Cervical myelopathy	
	OR* (95% CI)	p value	OR (95% CI)	p value
Age, y (+10 y)	23.6 (9.62–60.0)	<.0001	11.0 (1.15–133.9)	.047
Women (vs. men)	1.41 (1.03–1.92)	.032	4.33 (1.50–18.4)	.018
Body mass index, kg/m^2 (+1 SD)	2.12 (0.87–5.16)	.095	1.04 (0.94–1.15)	.41
Increased signal intensity positive	18.8 (6.87–66.4)	<.0001	6.32 (1.36–21.8)	.007
Diameter of spinal cord, mm (−1 mm)	1.40 (1.17–1.68)	.0002	1.46 (0.93–2.31)	.11
Diameter of cervical spinal canal, mm (−1 mm)	1.67 (1.45–1.93)	<.0001	2.73 (1.83–4.23)	<.001
Canal-to-body ratio (−10%)	1.85 (1.60–2.16)	<.0001	2.12 (1.47–3.16)	.0001

OR, odds ratio; CI, confidence interval; SD, standard deviation.

* OR was calculated by multiple logistic regression analysis after adjustment for age, gender, and body mass index.

relationship between narrowed spinal canal and CM. In addition, we clarified the positive association between cervical cord compression and spinal cord diameter, whereas the diameter of the spinal cord itself was not a significant predictive factor for CM. This may indicate that the spinal cord can become atrophied in individuals with cervical cord compression or in those who have a congenitally narrow spinal cord.

The present study also clarified the difference in age- and sex-related changes in the CSC diameter and CBR. The CSC has been the focus as a risk factor for CM [14,16]. However, in recent years, the CBR rather than CSC diameter has been reported to be a useful predictor for CM because of a magnification error resulting from the focus-to-film distance and the object-film distance on MRI [17]. However, Blackley et al. [18] showed that there is currently a poor correlation between the CBR and the true sagittal diameter of the spinal canal on computed tomography scans because of the wide normal variations in the diameter of the vertebral body. Therefore, the characteristics of the variations between the CSC diameter and the CBR should be considered. Of note, the present study found the CSC diameter to be higher in men than in women. However, the CBR was higher in women than in men, which is the reason for the increased diameter of the vertebral body in men. Therefore, the differences between the sexes should be taken into account when considering the CBR as a risk factor for CM.

Study limitations

The present study had several limitations. First, although more than 1,000 participants were included in the present study, these participants may not represent the general population because they were recruited from only two areas of Japan. However, anthropometric measurements were compared between the participants of the present study and the general Japanese population [19], and no significant differences in BMI were found between the participants in the present study and the Japanese population at large in both sexes (BMI [standard deviation] in men: 23.71 kg/m^2 [3.41 kg/m^2] and 23.95 kg/m^2 [2.64 kg/m^2], p=.33, respectively; BMI [standard deviation] in women: 23.06 kg/m^2 [3.42 kg/m^2] and 23.50 kg/m^2 [3.69 kg/m^2], p=.07, respectively). Second, the distribution of the CSC diameter applies to only a small portion of the Japanese population and cannot be extrapolated to other populations.

Conclusions

This study confirmed the significant association of the narrow CSC diameter with CM in a population-based cohort. The results prompt future studies to look into the various factors affecting the dimensions of the CSC, apart from aging.

Acknowledgments

The authors thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka of the Gobo Public Health Centre; Dr. Naoki Hirabayashi of the Kawakami Clinic, Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Kumiko Shinou, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, Ms. Ikuyo Ueyama, Mrs. Michiko Mori, Mrs. Hisayo Sugimoto, and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Centre; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mr. Shoichi Shimoichi, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji, and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. The authors also thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, and Mrs. Saeko Sahara for their assistance with data reduction and administration.

References

[1] Murone I. The importance of the sagittal diameters of the cervical spinal canal in relation to spondylosis and myelopathy. J Bone Joint Surg Br 1974;56:30–6.
[2] Eismont FJ, Clifford S, Goldberg M, et al. Cervical sagittal spinal canal size in spine injury. Spine 1984;9:663–6.
[3] Goto S, Umehara J, Aizawa T, et al. Comparison of cervical spinal canal diameter between younger and elder generations of Japanese. J Jap Orthop Ass 2010;15:97–103.
[4] Kato F, Yukawa Y, Suda K, et al. Normal morphology, age-related changes and abnormal findings of the cervical spine. Part II: magnetic resonance imaging of over 1,200 asymptomatic subjects. Eur Spine J 2012;21:1499–507.
[5] Ishimoto Y, Yoshimura N, Muraki S, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012;20:1103–8.
[6] Nagata K, Yoshimura N, Muraki S, et al. Prevalence of cervical cord compression and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Spine 2012;37:1892–8.
[7] Houten JK, Noce LA. Clinical correlations of cervical myelopathy and the Hoffmann sign. J Neurosurg Spine 2008;9:237–42.
[8] Sung RD, Wang JC. Correlation between a positive Hoffmann's reflex and cervical pathology in asymptomatic individuals. Spine 2001;26:67–70.
[9] Rhee JM, Heflin JA, Hamasaki T, et al. Prevalence of physical signs in cervical myelopathy: a prospective, controlled study. Spine 2009;34:890–5.
[10] Harrop JS, Hanna A, Silva MT, et al. Neurological manifestations of cervical spondylosis: an overview of signs, symptoms, and pathophysiology. Neurosurgery 2007;60(1 Supp1 1):S14–20.
[11] Matsumoto M, Toyama Y, Ishikawa M, et al. Increased signal intensity of the spinal cord on magnetic resonance images in cervical compressive myelopathy. Does it predict the outcome of conservative treatment? Spine 2000;25:677–82.
[12] Porter RW, Hibbert C, Wellman P. Backache and the lumbar spinal canal. Spine 1980;5:99–105.
[13] Hukuda S, Kojima Y. Sex discrepancy in the canal/body ratio of the cervical spine implicating the prevalence of cervical myelopathy in men. Spine 2002;27:250–3.

[14] Arnold JG Jr. The clinical manifestations of spondylochondrosis (spondylosis) of the cervical spine. Ann Surg 1955;141:872–89.
[15] Countee RW, Vijayanathan T. Congenital stenosis of the cervical spine: diagnosis and management. J Natl Med Assoc 1979;71: 257–64.
[16] Wolf BS, Khilnani M, Malis L. The sagittal diameter of the bony cervical spinal canal and its significance in cervical spondylosis. J Mt Sinai Hosp NY 1956;23:283–92.
[17] Pavlov H, Torg JS, Robie B, et al. Cervical spinal stenosis: determination with vertebral body ratio method. Radiology 1987;164:771–5.
[18] Blackley HR, Plank LD, Robertson PA. Determining the sagittal dimensions of the canal of the cervical spine. The reliability of ratios of anatomical measurements. J Bone Joint Surg Br 1999;81:110–2.
[19] Ministry of Health, Labour and Welfare. The report of National Health and Nutrition Survey 2005. Available at: www.mhlw.go.jp/bunya/kenkou/eiyou/dl/h20-houkoku-04.pdf. Accessed May 30, 2013.

Osteoarthritis and Cartilage 22 (2014) 104–110

Osteoarthritis and Cartilage

Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study

M. Teraguchi †, N. Yoshimura ‡, H. Hashizume †*, S. Muraki §, H. Yamada †, A. Minamide †, H. Oka ‡, Y. Ishimoto †, K. Nagata †, R. Kagotani †, N. Takiguchi †, T. Akune §, H. Kawaguchi ||, K. Nakamura ¶, M. Yoshida †

† *Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan*
‡ *Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*
§ *Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*
|| *Department of Sensory & Motor System Medicine, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*
¶ *Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, Saitama, Japan*

ARTICLE INFO

Article history:
Received 23 May 2013
Accepted 29 October 2013

Keywords:
Magnetic resonance imaging
Population-based cohort
Intervertebral disc degeneration
Entire spine
Prevalence
Distribution

SUMMARY

Objectives: The purposes of this study were to investigate the prevalence and distribution of intervertebral disc degeneration (DD) over the entire spine using magnetic resonance imaging (MRI), and to examine the factors and symptoms potentially associated with DD.
Design: This study included 975 participants (324 men, mean age of 67.2 years; 651 women, mean age of 66.0 years) with an age range of 21–97 years in the Wakayama Spine Study. DD on MRI was classified into Pfirrmann's system (grades 4 and 5 indicating DD). We assessed the prevalence of DD at each level in the cervical, thoracic, and lumbar regions and the entire spine, and examined DD-associated factors and symptoms.
Results: The prevalence of DD over the entire spine was 71% in men and 77% in women aged <50 years, and >90% in both men and women aged >50 years. The prevalence of an intervertebral space with DD was highest at C5/6 (men: 51.5%, women: 46%), T6/7 (men: 32.4%, women: 37.7%), and L4/5 (men: 69.1%, women: 75.8%). Age and obesity were associated with the presence of DD in all regions. Low back pain was associated with the presence of DD in the lumbar region.
Conclusion: The current study established the baseline data of DD over the entire spine in a large population of elderly individuals. These data provide the foundation for elucidating the causes and mechanisms of DD.

Introduction

Intervertebral disc degeneration (DD) is thought to be the first step in degenerative spinal changes[1], and is typically followed by the gradual formation of osteophytes, disc narrowing, and spinal stenosis[2,3]. Furthermore, DD is considered to be one of the causes of several symptoms (neck pain or low back pain)[4–7]. Therefore, in terms of developing preventive strategies for spinal disorders, it will be important to obtain fundamental data on DD (prevalence, distribution, associated factors, etc.) in a population-based cohort.

We believe that the analysis of DD over the entire spine would provide more useful data than that of DD in the cervical, thoracic, or lumbar regions, separately. In particular, investigations on the extent of DD in these three regions using whole spine magnetic resonance imaging (MRI) could provide useful data concerning intra-individual factors in the development of DD. Several studies have examined degenerative changes in only cervical and lumbar discs because of the high susceptibility to DD in these regions[8–12]. As well, several previous studies have investigated the aging process of the intervertebral discs in the cervical and lumbar regions using MRI in population-based cohorts[13,14]. However, degenerative changes in the thoracic region and correspondingly over the entire spine are poorly understood, because DD in the thoracic region is considered to be an uncommon problem[15,16]. In particular, the stabilization of the thoracic region by the thoracic cage, which

* Address correspondence and reprint requests to: H. Hashizume, Wakayama Medical University, 811-1 Kimiidera, Wakayama City, Wakayama 641-8510, Japan. Tel: 81-73-447-2300; Fax: 81-73-448-3008.
E-mail address: hashizum@wakayama-med.ac.jp (H. Hashizume).

1063-4584/$ – see front matter
http://dx.doi.org/10.1016/j.joca.2013.10.019

reduces the mechanical stress imposed on the intervertebral discs, is believed to reduce the incidence of degenerative diseases in this region[17].

Consistent with the above-mentioned previous studies, a population-based cohort analysis of DD in the different spinal regions using MRI could be used to examine the distribution of DD over the entire spine. However, to our knowledge, no previous studies have performed this type of investigation with a population-based cohort.

From the perspective of discogenic pain, the association between DD and symptoms remains controversial, although several reports have found that DD was a source of low back pain[4,5]. Moreover, reports on the association between the presence of DD in the cervical and thoracic regions and neck pain are rare[6,7]. Further, these studies were not performed with population-based cohorts and did not use whole spine MRI. Thus, no study has assessed neck pain and low back pain within individuals using whole spine MRI. To clarify the points described above, we established a population-based cohort study in which participants underwent whole spine MRI and were examined for symptoms associated with spinal disorders. This is our first report of DD over the entire spine based on a cross-sectional examination of a baseline population.

The aims of this study were to examine (1) the prevalence and distribution of DD over the entire spine using MRI in a population-based cohort, (2) the factors associated with DD (age, gender, and body mass index [BMI]) in the cervical, thoracic, and lumbar regions, and (3) the association between DD and symptoms (neck pain and low back pain).

Methods

Participants

The present study, entitled the Wakayama Spine Study, was performed with a sub-cohort of the second visit of the ROAD (Research on Osteoarthritis/osteoporosis Against Disability) study, which was initiated as a nationwide, prospective study of bone and joint diseases in population-based cohorts; the cohorts were established in three communities with different characteristics (i.e., urban, mountainous, and coastal regions) in Japan. A detailed profile of the ROAD study has already been described elsewhere[18,19]. Here, we briefly summarize the profile of the present study. The second visit of the ROAD study began in 2008 and was completed in 2010. All the participants in the baseline study were invited to participate in the second visit. In addition to the former participants, inhabitants aged 60 years and older in the urban area and those aged 40 years and younger in the mountainous and coastal areas who were willing to participate in the ROAD survey were also included in the second visit (both the mountainous and coastal areas were in Wakayama prefecture). Finally, 2674 individuals (900 men, 1774 women) participated in the second visit of the ROAD study, and comprised 1067 individuals (353 men, 714 women) in the urban area, 742 individuals (265 men, 477 women) in the mountainous area, and 865 individuals (282 men, 583 women) in the coastal area. Among these three communities in the ROAD study, the mountainous and coastal areas from which we invited all 1607 participants (547 men, 1060 women) to the Wakayama Spine Study are located in Wakayama prefecture. Of the 1607 participants, a total of 1011 individuals provided written informed consent and attended the Wakayama Spine Study with MRI examinations[20,21]. Among the 1011 participants, those who had MRI-sensitive implanted devices (e.g., pacemakers) and other disqualifiers were excluded. Consequently, 980 individuals underwent MRI of the whole spine. Furthermore, one participant who had undergone a previous cervical operation and four participants who had undergone a previous posterior lumbar fusion were excluded from the analysis. Finally, whole spine MRI results were available for 975 participants (324 men, 651 women) with an age range of 21–97 years (mean, 67.2 years for men and 66.0 years for women). Table I shows the demographic and baseline characteristics of the 975 participants in the present study.

For the purpose of analysis, the participants were divided into five age groups: (1) under 50 years, (2) 50–59 years, (3) 60–69 years, (4) 70–79 years, and (5) 80 years and over. The anthropometric measurements included height, weight, and BMI (weight [kg]/height^2 [m^2]). BMI was categorized according to the guidelines for Asians proposed by the World Health Organization and was thus defined as follows: underweight, less than 18.5; normal, 18.5–23; overweight, 23–27.5; and obesity, greater than 27.5[22]. Experienced orthopedists also asked all participants the following question regarding neck pain and low back pain: “Have you experienced neck pain on most days during the past month, in addition to now?” and “Have you experienced low back pain on most days during the past month, in addition to now?” Those who answered “yes” were defined as having neck pain or low back pain based on previous studies[23–26].

MRI

A mobile MRI unit (Excelart 1.5 T, Toshiba, Tokyo, Japan) was used in the present study, and whole spine MRI was performed for all participants on the same day as the examination. The participants were supine during the MRI, and those with rounded backs used triangular pillows under their head and knees. The imaging protocol included sagittal T2-weighted fast spin echo (FSE) (repetition time [TR]: 4000 ms/echo, echo time [TE]: 120 ms, field of view [FOV]: 300 × 320 mm), and axial T2-weighted FSE (TR: 4000 ms/echo, TE: 120 ms, FOV: 180 × 180 mm).

Sagittal T2-weighted images were used to assess the intervertebral space from C2/3 to L5/S1. C2/3 to C7/T1, T1/2 to T12/L1, and L1/2 to L5/S1 were defined as the cervical region, thoracic region, and lumbar region, respectively. DD grading was performed by an

Table I
Characteristics of participants

	Overall	Men	Women
No. of participants	**975**	**324**	**651**
Age strata (years)			
<50	125	38	87
50–59	175	59	116
60–69	223	65	158
70–79	261	89	172
≧80	191	73	118
Demographic characteristics			
Age, years	66.4 ± 13.5	67.2 ± 13.9	66.0 ± 13.4
Height, cm	156.4 ± 9.4	164.6 ± 7.2	151.5 ± 7.2
Weight, kg	56.8 ± 11.5	64.5 ± 11.6	53.0 ± 9.4
BMI (kg/m^2)	23.3 ± 3.6	23.6 ± 3.4	23.1 ± 3.7
BMI (WHO-Asian category) (*N*)			
Underweight	61	16	45
Normal	425	124	300
Overweight	361	139	221
Obesity	128	44	84
Baseline characteristics			
Symptoms (%)			
Neck pain	24.9	19.4	27.7
Low back pain	43	36.7	42.1
Life style (%)			
Smoking	10.7	25.2	4.1
Alcohol consumption	31.4	56.8	18.8

BMI category for Asian was based on World Health Organization (WHO) guidelines defining underweight (<18.5), normal (18.5–23), overweight (23–27.5), and obese (>27.5). Values are the means ± standard deviation.

orthopedist (MT) who was blind to the background of the subjects. The degree of DD on MRI was classified into five grades based on Pfirrmann's classification system[27], with grades 4 and 5 indicating DD. As shown in Fig. 1, the signal intensity for grade 4 was intermediate to hypointense to the cerebrospinal fluid (dark gray), while the structure is inhomogeneous. Meanwhile, for grade 5, the signal intensity is hypointense to the cerebrospinal fluid (black), and the structure is likewise inhomogeneous. In addition, the disc space is collapsed. It has been reported that loss of signal intensity is significantly associated with the morphological level of the DD and is also associated with both the water and proteoglycan content in a disc[28]. Therefore, we used a grading based on signal intensity and disc height. For evaluating intraobserver variability, 100 randomly selected magnetic resonance images of the entire spine were rescored by the same observer (MT) more than 1 month after the first reading. Furthermore, to evaluate interobserver variability, 100 other magnetic resonance images were scored by two orthopedists (MT and RK) using the same classification. The intraobserver and interobserver variability for DD, as evaluated by kappa analysis, was 0.94 and 0.94, respectively.

"Prevalence of DD", which was defined as "the proportion of the number of participants who had DD at each intervertebral space or region or over the entire spine divided by the total number of participants", was used to describe the frequency of the presence of DD. In the analysis, to clarify the associated factors using multiple logistic regression analysis, we entered a variable of prevalence state (1, presence; 0, absence) of DD as a dependent variable.

Statistical analysis

Multiple logistic regression analysis was used to estimate the association between the presence of DD in each region (cervical, thoracic, and lumbar) as dependent variables and the age group, gender, and BMI category as nominal independent variables after adjustment for the age group, gender and BMI category, mutually.

Additionally, multiple logistic regression analysis was used to estimate the association between the presence of neck pain or low back pain and the presence of DD in each region after adjustment for age, gender, and BMI. Furthermore, in cases in which the presence of DD was significantly associated with a symptom, we examined as a sub-analysis the association between the presence of neck pain or low back pain and the number of DD (categorized into "0", "1 or 2", "3 or more" for ready assessment) in each region using multiple logistic regression analysis after adjustment for age, gender, and BMI. All statistical analyses were performed using JMP version 8 (SAS Institute Japan, Tokyo, Japan).

Results

As shown in Table II, the prevalence of DD in the cervical and thoracic regions and over the entire spine increased with the elevation of the age strata in both men and women. For both genders, the prevalence of DD in the lumbar region was also increased with the elevation of the age strata up to the 70-year-old age group but decreased in the 80-year-old age group. Table III

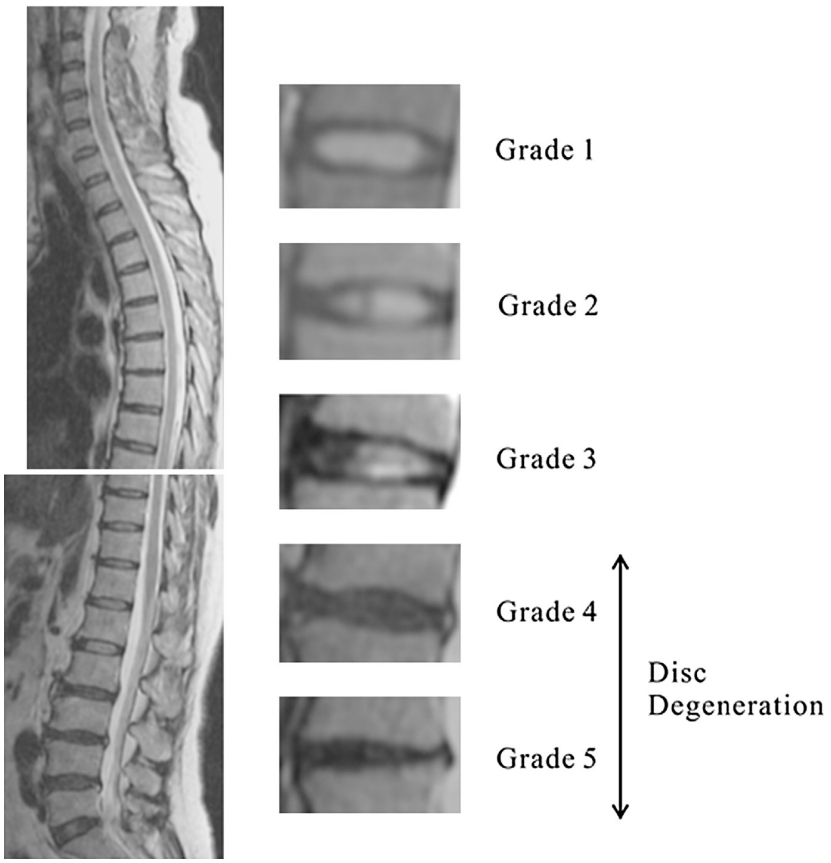


Fig. 1. Mid-sagittal view on T2-weighted images of the whole spine MRI with Pfirrmann classification. The grade is described according to Pfirrmann classification. Grades 4 and 5 were considered degenerated. The signal intensity for grade 4 was intermediate to hypointense to the cerebrospinal fluid (dark gray), while the structure is inhomogeneous. Meanwhile, for grade 5, the signal intensity is hypointense to the cerebrospinal fluid (black), and the structure is also inhomogeneous. Additionally, the disc space is collapsed.

shows the prevalence of intervertebral spaces with DD over the entire spine for the participants in this study. The three highest prevalence levels of DD in the intervertebral spaces in the cervical, thoracic, and lumbar regions were as follows. The prevalence at C5/6 was 51.5% (95% CI: 46.1–56.3) in men and 46% (95% CI: 42.2–49.9) in women, followed by the prevalence at C6/7 of 43.5% in men and 33.3% in women, and at C4/5 of 38.6% in men and 35.8% in women. The prevalence at T6/7 was 32.4% (95% CI: 27.5–37.6) in men and 37.7% (95% CI: 34.1–41.5) in women, followed by the prevalence at T7/8 of 31.8% in men and 36.2% in women, and at T5/6 of 28.4% in men and 35.9% in women. The prevalence at L4/5 was 69.1% (95% CI: 63.9–73.9) in men and 75.8% (95% CI: 72.3–78.9) in women, followed by that at L5/S1 of 66.7% in men and 70.9% in women, and at L3/4 of 59.3% in men and 61.9% in women.

An older age was significantly associated with the presence of DD in each region. Gender was not significantly associated with the presence of DD in each region, although men demonstrated a tendency for a greater number of DD than women in the cervical region. In addition, overweight status (BMI: 23–27.5) was a significantly associated factor in the cervical and thoracic regions, and obesity (BMI: >27.5) was a significantly associated factor in all regions compared with participants of a normal weight (BMI: 18.5–23) (Table IV).

The participants with DD in the cervical region did not significantly differ in terms of the presence of neck pain (OR 0.88, 95% CI: 0.63–1.22, $P = 0.53$). The presence of DD in the thoracic region was not significantly associated with neck pain (OR 0.84, 95% CI: 0.60–1.19, $P = 0.33$) and low back pain (OR 1.08, 95% CI: 0.80–1.47, $P = 0.60$). However, the presence of DD in the lumbar region was significantly associated with low back pain (OR 1.57, 95% CI: 1.02–2.49, $P < 0.05$). Moreover, in a sub-analysis, we investigated the association between low back pain and the number of DD in the lumbar region ("0", "1 or 2", "3 or more"). The presence of low back pain was significantly higher in participants with three or more DD (OR 1.75, 95% CI: 1.11–2.81, $P < 0.05$), but not in those with one or two DD (OR 1.34, 95% CI: 0.84–2.20, $P = 0.22$), as compared with participants without DD.

Discussion

This study is the first to report the prevalence and distribution of DD over the entire spine using whole spine MRI in a population-based cohort. The prevalence of DD over the entire spine and in each of the three spinal regions was higher in older participants. In addition, we noted that the presence of DD was significantly associated with low back pain in the lumbar region but not with neck pain in the cervical region.

Battié *et al.* reviewed the prevalence of DD in the lumbar region and noted that it ranged from 20% to 83%[29]. Consistent with the observations of this review, other reported prevalence levels of DD in the lumbar region have shown wide variation between samples and have often been quite high because the studies had certain drawbacks, including relatively small sample sizes[1,30], narrow age ranges[5,31], and asymptomatic subjects[32]. However, no previous study has assessed the prevalence of DD over the entire spine using whole spine MRI. We noted that the prevalence of DD over the entire spine exceeded 70% in participants less than 50 years of age and was greater than 90% in participants older than 50 years of age.

Little epidemiological data are available concerning DD in the intervertebral space using MRI assessments in a population-based cohort. Matsumoto *et al.*[4] reported that the prevalence of DD in the cervical region was the highest at C5/6 (86% in men and 89% in women over the age of 60 years). In addition, Hanagai *et al.*[33] and Kanayama *et al.*[34] reported that the prevalence of DD in the lumbar region was the highest at L4/5 (67%; mean age 68.4 years) and L5/S1 (49.5%; mean age 39.7 years), respectively. In the present study, the prevalence of DD was the highest at C5/6 (51.5% in men and 46.0% in women) and L4/5 (69.1% in men and 75.8% in women). The prevalence of cervical DD in the previous study by Matsumoto *et al.*[4] was higher than that in the present study. However, the subjects were recruited from volunteers in the hospital rather than a population; thus, the capacity for strict comparisons are limited. Furthermore, few studies have reported age-related DD in the thoracic region. Matsumoto *et al.* reported that the highest prevalence of DD occurred at T7/8 (30.9%; mean age 48.0 y) followed by T6/7 in the thoracic region; however, all 94 participants in this report were asymptomatic[35]. In the present study, we confirmed a high prevalence of DD at T6/7 in the thoracic region. This finding is supported by results from thoracic MRI investigations demonstrating a high prevalence of DD in asymptomatic individuals.

The distribution of prevalence of DD was similar to the alignment of the spine in the sagittal plane, such as cervical lordosis (C3–C7), thoracic kyphosis (T1–T12), and lumbar lordosis (L1–L5)[36]. The high prevalence of DD in the lumbar region can potentially be explained by mechanical stress. Our results support the hypothesis that compressive stress affected DD, since compressive stresses are the highest in the mid-thoracic region of the entire spine[37]. Mechanical stress on the thoracic intervertebral disc is reduced due to stabilization by the thoracic cage, and therefore, the thoracic intervertebral disc may be affected by the detrimental effect of compressive stress caused by posture on the sagittal balance of the spine[38]. This study also provides the first mapping of intervertebral spaces with DD over the entire spine by MRI analysis, which adds to our knowledge of the distribution of prevalence of DD in the cervical, thoracic, and lumbar regions, which has been reported only fragmentarily in previous reports.

Our current results confirmed that age was a significant factor associated with the presence of DD in all three regions. Previous studies reported that the association of DD to factors such as height, weight, and gender was uncertain; however, age, obesity, smoking, and occupation have been suggested to be DD-associated factors[39–42]. The previous studies focused almost entirely on the lumbar region, and the identification of associated factors may be challenging for this region because it is affected to a greater extent by various factors, including mechanical stress. Moreover, it remains unknown what other factors (beyond age) are associated with DD in the cervical and thoracic regions[6,13]. In the present study, overweight and obesity significantly influenced DD in the cervical and thoracic regions (cervical; OR: overweight 1.38 [95% CI 1.00–1.90], obesity 1.60 [95% CI 1.04–2.51], thoracic; OR: overweight 1.64 [95% CI 1.17–2.29], obesity 3.12 [95% CI 1.91–5.19]), and obesity also significantly influenced DD in the lumbar region (OR: 2.56 [95% CI 1.20–6.14]). In a previous study, Samartzis *et al.* reported that DD in the lumbar region was significantly associated with overweight and obesity[39]. However, DD in the cervical and thoracic region did not demonstrate a significant association with BMI, as reported by Okada *et al.*[6] and Matsumoto *et al.*[35]. Of note, the previous studies were

Table II
Prevalence of DD by age strata in men and women

	Entire spine		Cervical		Thoracic		Lumbar	
	Men	Women	Men	Women	Men	Women	Men	Women
Age strata (years)								
<50	71.0	77.0	26.3	27.9	15.7	11.4	55.2	71.2
50–59	91.5	93.1	47.4	49.1	49.1	35.3	86.4	91.3
60–69	98.4	95.5	66.1	54.4	61.5	63.2	96.9	94.3
70–79	95.8	99.4	80.9	72.0	73.0	79.6	96.6	96.5
≧80	93.2	97.4	86.3	85.5	79.4	88.9	82.1	84.5

Values are percentage.

Table III
Prevalence of intervertebral spaces with DD over the entire spine by age strata in men and women

Age strata (years)	C2/3	C3/4	C4/5	C5/6	C6/7	C7/T1	T1/2	T2/3	T3/4	T4/5	T5/6	T6/7	T7/8	T8/9	T9/10	T10/11	T11/12	T12/L1	L1/2	L2/3	L3/4	L4/5	L5/S1
Men																							
Total	28.3	30.2	38.6	51.5	43.5	26.8	20.3	23.4	22.2	24.0	28.4	32.4	31.8	28.7	31.4	25.0	24.0	17.5	30.0	51.5	59.3	69.1	66.7
<50	10.5	10.5	13.1	15.7	13.1	5.2	5.2	7.8	7.8	5.2	10.5	7.8	5.2	2.6	2.6	2.6	0.0	0.0	2.6	10.5	7.8	34.2	47.3
50–59	6.7	11.8	15.2	37.2	27.1	10.1	8.4	6.7	11.8	11.8	16.9	23.7	27.1	16.9	20.3	16.9	13.5	5.1	15.2	35.5	61.0	74.5	50.8
60–69	35.3	36.9	49.2	50.7	40.0	21.0	20.0	24.6	23.0	27.6	27.6	35.3	32.3	36.9	41.5	23.0	24.6	18.4	40.0	60.0	69.0	76.9	75.3
70–79	35.9	35.9	49.4	64.0	51.6	34.8	24.7	26.9	25.8	30.3	33.7	38.2	41.5	35.9	40.4	37.0	31.4	26.9	39.3	69.6	73.0	79.7	79.7
≧80	39.7	42.4	47.9	67.1	65.7	46.5	32.8	39.7	32.8	32.8	41.0	42.4	36.9	35.6	35.6	30.1	35.6	24.6	39.7	56.1	58.9	63.0	65.7
Women																							
Total	21.9	24.8	35.8	46.0	33.3	13.6	15.2	23.1	29.8	31.7	35.9	37.7	36.2	34.2	32.7	28.7	23.8	20.0	31.7	49.7	61.9	75.8	70.9
<50	2.2	3.4	10.3	20.6	10.3	1.1	0.0	1.1	4.5	0.0	1.1	4.5	3.4	5.7	4.5	4.5	1.1	0.0	4.5	12.6	18.3	49.4	56.3
50–59	11.2	9.4	23.2	36.2	23.2	3.4	6.8	12.0	15.5	15.5	16.3	18.1	19.8	12.9	13.7	10.3	6.9	6.9	15.6	35.6	55.6	73.9	70.4
60–69	13.9	20.8	31.0	43.6	29.1	11.3	13.2	18.3	29.7	32.2	37.9	39.8	31.6	32.2	30.3	19.6	15.8	14.5	25.3	55.0	66.4	85.4	75.9
70–79	33.7	34.8	46.5	53.4	42.4	16.2	22.0	34.3	41.2	44.7	50.0	50.0	47.0	45.9	44.7	42.4	34.3	26.1	44.7	64.5	80.2	86.0	81.9
≧80	40.6	46.6	57.6	66.9	52.5	32.2	27.1	40.6	45.7	51.6	57.6	61.0	66.9	61.8	57.6	56.7	52.9	46.1	57.2	62.3	67.5	69.2	58.9

Values are percentage.

conducted with asymptomatic healthy subjects. Therefore, based on our findings, obesity appears to have some influence on the process of DD over the entire spine.

An association between DD in the lumbar region and low back pain was previously demonstrated in a twin study[43]. Moreover, Okada *et al.*[6] reported an association between neck pain and DD in the cervical region, whereas Arana *et al.*[7] found an association between neck pain and DD in the upper thoracic region. Of interest, no agreement has been reached regarding the most appropriate definition of neck pain and low back pain in population cohorts[7]. Nonetheless, we observed a significant association between the presence of DD in the lumbar region and low back pain.

The present study has several limitations. First, it was a cross-sectional study, and therefore, the transition to DD cannot be clarified. Second, the participants included in the present study may not represent the general population, since they were recruited from only two local areas. To confirm whether the participants of the Wakayama Spine Study are representative of the Japanese population, we compared the anthropometric measurements and frequencies of smoking and alcohol consumption between the general Japanese population and the study participants. No significant differences in BMI were observed (men: 24.0 and 23.7, $P = 0.33$; women: 23.5 and 23.1, $P = 0.07$). Further, the proportion of current smokers and those who consumed alcohol (those who regularly smoked or consumed alcohol more than once per month) in men and the proportion of those who consumed alcohol in women were significantly higher in the general Japanese population than in the study population, whereas there was no significant difference in the proportion of current smokers in women (male smokers, 32.6% and 25.2%, $P = 0.015$; female smokers, 4.9% and 4.1%, $P = 0.50$; men who consumed alcohol, 73.9% and 56.8%, $P < 0.0001$; women who consumed alcohol, 28.1% and 18.8%, $P < 0.0001$). These results suggest the likelihood that in this study, participants had healthier lifestyles than those of the general Japanese population[44]. This "healthy" selection bias should be taken into consideration when generalizing the results obtained from the Wakayama Spine Study. Third, the Pfirrmann classification introduced a comprehensive MRI grading system based on the assessment of structure, the distinction of the nucleus and annulus fibrosis, the signal intensity[28], and the height of the intervertebral discs[27]. However, bony endplate alterations, osteophyte changes, spinal stenosis, and disc protrusion are not covered by the Pfirrmann classification. Therefore, it is necessary to perform investigations that include these morphological changes. Finally, the accurate measurement of obesity, such as abdominal obesity and/or body composition, might reveal that obesity has a stronger association with DD; however, the present study examined only BMI as a measurement of obesity. Thus, we plan to examine the girth of the abdomen and body composition using electrical impedance in the assessment of human body composition (the BIA method) in a future study.

In conclusion, this study is the first one to investigate the prevalence of DD over the entire spine in a large population of individuals to establish baseline data for a prospective longitudinal

Table IV
Multiple logistic regression of the association with presence of DD with age, BMI, and gender

	Cervical	Thoracic	Lumbar
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age group (years)			
<50	1	1	1
50–59 (vs <50)	2.45 (1.5–4.06)**	4.60 (2.53–8.76)***	4.47 (2.44–8.48)***
60–69 (vs <50)	3.62 (2.26–5.91)***	12.0 (6.77–22.7)***	9.95 (5.02–21.3)***
70–79 (vs <50)	7.87 (4.86–12.9)***	24.9 (13.8–47.6)***	15.0 (7.26–34.5)***
≧80 (vs <50)	16.9 (9.68–30.5)***	47.0 (24.5–95.6)***	2.94 (1.71–5.13)**
Men (vs women)	1.20 (0.89–1.64)	0.88 (0.64–1.21)	0.70 (0.45–1.09)
BMI (WHO-Asian category)			
Underweight (vs normal)	0.91 (0.49–1.70)	1.36 (0.71–2.67)	0.81 (0.38–1.84)
Normal	1	1	1
Overweight (vs normal)	1.38 (1.00–1.90)*	1.64 (1.17–2.29)*	1.14 (0.71–1.85)
Obesity (vs normal)	1.60 (1.04–2.51)*	3.12 (1.91–5.19)***	2.56 (1.20–6.14)*

BMI category for Asian was based on World Health Organization (WHO) guidelines defining underweight (<18.5), normal (18.5–23), overweight (23–27.5), and obese (>27.5). OR = odds ratio, CI = confidential interval.
*$P < 0.05$, **$P < 0.001$, ***$P < 0.0001$.

study. The prevalence of intervertebral spaces with DD was the highest at C5/6, T6/7, and L4/5 in the cervical, thoracic, and lumbar regions, respectively. DD in the cervical, thoracic, and lumbar regions was significantly associated with age and obesity. A significant positive association was observed between the presence of DD in the lumbar region and low back pain.

Author contributions

All authors worked collectively to develop the protocols and method described in this paper. MT, NY, SM, HO, YI, KN, NT, and TA were principal investigators responsible for the fieldwork in the Wakayama Spine study. MT and SM performed the statistical analysis. All authors contributed to the analysis and interpretation of results. MT wrote the report. All authors read and approved the final manuscript.

Role of the funding source

The sponsors had no role in study design, data collection, data analysis, data interpretation, or in writing of the report.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgments

This study was supported by a Grant-in-Aid for Scientific Research (B20390182, B23390357, C20591737, C20591774), and for Exploratory Research (19659305) from the Japanese Ministry of Education, Culture, Sports, Science and Technology, H17-Men-eki-009, H18-Choujyu-037, and H20-Choujyu-009 from the Ministry of Health, Labour and Welfare, Research Aid from the Japanese Orthopaedic Association, a Grant from the Japanese Orthopaedics and Traumatology Foundation, Inc. (No. 166), and a Grant-in-Aid for Scientific Research, Scientific Research (C22591639) from the Japanese Society for the Promotion of Science.

The authors wish to thank Mrs Tomoko Takijiri and other members of the Public Office in Hidakagawa Town, and Mrs Tamako Tsutsumi, Mrs Kanami Maeda, and other members of the Public Office in Taiji Town for their assistance in the location and scheduling of participants for examinations.

References

1. Boos N, Weissbach S, Rohrbach H, Weiler C, Spratt KF, Nerlich AG. Classification of age-related changes in lumbar intervertebral discs: 2002 Volvo Award in basic science. Spine 2002;27:2631–44.
2. Kirkaldy-Willis WH, Farfan HF. Instability of the lumbar spine. Clin Orthopedics Relat Res 1982;165:110–23.
3. Ahn TJ, Lee SH, Choi G, Ahn Y, Liu WC, Kim HJ, *et al.* Effect of intervertebral disk degeneration on spinal stenosis during magnetic resonance imaging with axial loading. Neurol Med Chir 2009;49:242–7.
4. Sambrook PN, MacGregor AJ, Spector TD. Genetic influences on cervical and lumbar disc degeneration: a magnetic resonance imaging study in twins. Arthritis Rheum 1999;42:366–72.
5. Kjaer P, Laboeuf-Yde C, Korsholm L, Sorensen JS, Bendix T. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. Spine 2005;30:1173–80.
6. Okada E, Matsumoto M, Ichihara D, Chiba K, Toyama Y, Fujiwara H, *et al.* Aging of the cervical spine in healthy volunteers: a 10-year longitudinal magnetic resonance imaging study. Spine 2009;34:706–12.
7. Arana E, Marti-Bonmatí L, Mollá E, Costa S. Upper thoracic-spine disc degeneration in patients with cervical pain. Skeletal Radiol 2004;33:29–33.
8. Hassett G, Hart D, Manek N, Doyle DV, Spector TD. Risk factors for progression of lumbar spine disc degeneration: the Chingford Study. Arthritis Rheum 2003;48:3112–7.
9. Videman T, Battié MC. The influence of occupation on lumbar degeneration. Spine 1999;11:1164–8.
10. Battié MC, Videman T, Gibbons L, Manninen H, Gill K, Pope M, *et al.* Occupational driving and lumbar disc degeneration: a case control study. Lancet 2002;360:1369–74.
11. Videman T, Nurminen M, Troup JD. 1990 Volvo Award in clinical sciences. Lumbar spinal pathology in cadaveric material in relation to history of back pain, occupation, and physical loading. Spine 1990;15:728–40.
12. Adams MA, Roughley PJ. What is intervertebral disc degeneration, and what causes it? Spine 2006;31:2151–61.
13. Matsumoto M, Fujiwara Y, Suzuki N, Nishi Y, Nakayama M, Yabe Y, *et al.* MRI of cervical intervertebral discs in asymptomatic subjects. J Bone Joint Surg Br Vol 1998;80:19–24.
14. Cheung KM, Karppinen J, Chan D, Ho DW, Song YQ, Sham P, *et al.* Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine 2009;34:934–40.
15. Aizawa T, Sato T, Tanaka Y, Ozawa Y, Hoshikawa T, Ishii Y, *et al.* Thoracic myelopathy in Japan: epidemiological retrospective study in Miyagi Prefecture during 15 years. Tohoku J Exp Med 2006;210:199–208.
16. Girard CJ, Schweitzer ME, Morrison WB, Parellade JA, Carrino JA. Thoracic spine disc-related abnormalities: longitudinal MR imaging assessment. Skeletal Radiol 2004;33:216–22.
17. Mclnerney J, Ball PA. The pathophysiology of thoracic disc disease. Neurosurg Focus 2000;9:e1.
18. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study. Int J Epidemiol 2010;39:988–95.
19. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, *et al.* Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 2009;27:620–8.
20. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, *et al.* Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012;20:1103–8.
21. Nagata K, Yoshimura N, Muraki S, Hashizume H, Ishimoto Y, Yamada H, *et al.* Prevalence of cervical cord compression and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Spine 2012;37:1892–8.
22. Choo V. WHO reassesses appropriate body-mass index for Asian populations. Lancet 2002;360:235.
23. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, *et al.* Prevalence of radiographic lumbar spondylosis and its association with low back pain in elderly subjects of population-based cohorts: the ROAD study. Ann Rheum Dis 2009;68:1401–6.
24. Muraki S, Akune T, Oka H, En-yo Y, Yoshida M, Saika A, *et al.* Impact of knee and low back pain on health-related quality of life in Japanese women: the Research on Osteoarthritis Against Disability (ROAD). Mod Rheumatol 2010;20:444–51.
25. Muraki S, Akune T, Oka H, En-yo Y, Yoshida M, Saika A, *et al.* Health-related quality of life in subjects with low back pain

and knee pain in a population-based cohort study of Japanese men: the ROAD study. Spine 2011;36:1312–9.
26. Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, *et al*. Incidence and risk factors for radiographic lumbar spondylosis and lower back pain in Japanese men and women: the ROAD study. Osteoarthritis Cartilage 2012;20:712–8.
27. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine 2001;26:1873–8.
28. Benneker LM, Heini PF, Anderson SE, Alini M, Ito K. Correlation of radiographic and MRI parameters to morphological and biochemical assessment of intervertebral disc degeneration. Eur Spine J 2005;14:27–35.
29. Battié MC, Videman T, Parent E. Lumbar disc degeneration. Epidemiology and genetic influences. Spine 2004;29:2679–90.
30. Elfering A, Semmer N, Birkhofer D, Zanetti M, Hodler J, Boos N. Risk factors for lumbar disc degeneration: a 5-year prospective MRI study in asymptomatic individuals. Spine 2001;27:125–34.
31. Luoma K, Riihimäki H, Luukkonen R, Raininko R, Viikari-Juntura E, Lamminen A. Low back pain in relation to lumbar disc degeneration. Spine 2000;25:487–92.
32. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects: a prospective investigation. J Bone Joint Surg Am Vol 1990;72:403–8.
33. Hanagai M, Kaneoke K, Kuno S, Hinotsu S, Sakane M, Mamizuka N, *et al*. Factors associated with lumbar intervertebral disc degeneration in the elderly. Spine J 2008;8:732–40.
34. Kanayama M, Togawa D, Takahashi C, Terai T, Hashimoto T. Cross-sectional magnetic resonance imaging study of lumbar disc degeneration in 200 healthy individuals. J Neurosurg Spine 2009;11:501–7.
35. Matsumoto M, Okada E, Ichihara D, Watanabe K, Chiba K, Toyama Y, *et al*. Age-related changes of thoracic and cervical intervertebral discs in asymptomatic subjects. Spine 2010;35:1359–64.
36. Lee CS, Chung SS, Kang KC, Park SJ, Shin SK. Normal patterns of sagittal alignment of the spine in young adults radiological analysis in a Korean population. Spine 2011;36:1648–54.
37. Keller T, Colloca CJ, Harrison DE, Harrison DD, Janik TJ. Influence of spine morphology on intervertebral disc loads and stresses in asymptomatic adults: implications for the ideal spine. Spine J 2005;5:297–309.
38. An HS, Wise JJ, Xu R. Anatomy of the cervicothoracic junction: a study of cadaveric dissection, cryomicrotomy and magnetic resonance imaging. J Spinal Disord 1999;12:519–25.
39. Samartzis D, Karppinen J, Chan D, Luk KD, Cheung KM. The association of lumbar intervertebral disc degeneration on magnetic resonance imaging with body mass index in overweight and obese adults. Arthritis Rheum 2012;64:1488–96.
40. Miller JA, Schmatz C, Schultz AB. Lumbar disc degeneration: correlation with age, sex, and spine level in 600 autopsy specimens. Spine 1988;13:173–8.
41. Wang YX, Griffith JF, Ma HT, Kwok AW, Leung JC, Yeung DK, *et al*. Relationship between gender, bone mineral density, and disc degeneration in the lumbar spine: a study in elderly subjects using an eight-level MRI-based disc degeneration grading system. Osteoporos Int 2011;22:91–6.
42. Battié MC, Videman T, Gill K, Moneta GB, Nyman R, Kaprio J, *et al*. 1991 Volvo Award in clinical sciences. Smoking and lumbar intervertebral disc degeneration: an MRI study of identical twins. Spine 1991;16:1015–21.
43. Battié MC, Videman T, Kaprio J, Gibbons LE, Gill K, Manninen H, *et al*. The Twin Spine Study: contributions to a changing view of disc degeneration. Spine J 2009;9:47–59.
44. Ministry of Health, Labour and Welfare. The Outline of the Results of National Livelihood Survey. Available at: http://www.mhlw.go.jp/toukei/list/20-19-1.html; 2007.

Osteoporos Int (2014) 25:1081–1088
DOI 10.1007/s00198-013-2550-z

ORIGINAL ARTICLE

Exercise habits during middle age are associated with lower prevalence of sarcopenia: the ROAD study

T. Akune · S. Muraki · H. Oka · S. Tanaka · H. Kawaguchi · K. Nakamura · N. Yoshimura

Received: 8 July 2013 / Accepted: 3 October 2013 / Published online: 22 October 2013

Abstract
Summary The present cross-sectional study investigated the prevalence of sarcopenia and clarified its associated factors in 1,000 elderly participants of Japanese population-based cohorts. Exercise habit in middle age was associated with low prevalence of sarcopenia in older age, suggesting that it is a protective factor against sarcopenia in older age.
Introduction The present study investigated the prevalence of sarcopenia using the European Working Group on Sarcopenia in Older People (EWGSOP) definition, and clarified the association of sarcopenia with physical performance in the elderly participants of Japanese population-based cohorts of the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.
Methods We enrolled 1,000 participants (aged ≥65 years) from the second visit of the ROAD study who had completed assessment of handgrip strength, gait speed, and skeletal muscle mass measured by bioimpedance analysis. Presence of sarcopenia was determined according to the EWGSOP algorithm. Information collected included exercise habits in middle age.
Results Prevalence of sarcopenia was 13.8 % in men and 12.4 % in women, and tended to be significantly higher according to increasing age in both sexes. Factors associated with sarcopenia, as determined by logistic regression analysis, were chair stand time (odds ratio [OR], 1.09; 95 % confidence interval [CI], 1.04–1.14), one-leg standing time (OR, 0.97; 95 % CI, 0.96–0.99), and exercise habit in middle age (OR, 0.53; 95 % CI, 0.31–0.90). Exercise habit in middle age was associated with low prevalence of sarcopenia in older age. Furthermore, linear regression analysis revealed that exercise habits in middle age were significantly associated with grip strength ($P<.001$), gait speed ($P<.001$), and one-leg standing time ($P=.005$) in older age.
Conclusions This cross-sectional study suggests that exercise habit in middle age is a protective factor against sarcopenia in older age and effective in maintaining muscle strength and physical performance in older age.

Keywords Elderly · Epidemiology · Exercise · Physical performance · Sarcopenia

T. Akune (✉) · S. Muraki
Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, Hongo 7-3-1 Bunkyo-ku, Tokyo 113-8655, Japan
e-mail: akune-ort@h.u-tokyo.ac.jp

H. Oka · N. Yoshimura
Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

S. Tanaka · H. Kawaguchi
Department of Sensory and Motor System Medicine, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

K. Nakamura
National Rehabilitation Center for Persons with Disabilities, Saitama, Japan

Introduction

Sarcopenia is characterized by generalized loss of skeletal muscle mass and muscle strength and/or function in the elderly, causing multiple adverse health outcomes, including physical disability, poor quality of life, and death [1–6]. Although cross-sectional studies have investigated prevalence of sarcopenia [7–13], epidemiologic evidence using population-based samples is insufficient despite the urgent need for strategies to prevent and treat this condition.

Japan is a super-aged society, and the proportion of the aged population is increasing. The percentage of individuals

aged ≥65 years was 23 % in 2010 and is expected to reach 30.1 % in 2024 and 39 % in 2051 [14]. The government of Japan reported that musculoskeletal disorders were present in 22.9 % of the entire population of those who were certified as requiring assistance or long-term care elderly in 2010 and were ranked first among its causes, together with joint diseases, falls, fractures, and spinal cord disorders [15]. For preventing and treating musculoskeletal disorders, there is an urgent need to develop and establish a prevention strategy and treatment programs that are effective in reducing the risk of disability among the elderly, which leads to requirement of assistance or long-term care. Although sarcopenia is a common musculoskeletal disease in the elderly, it is not clearly categorized [15]. There appears to be insufficient recognition of sarcopenia in daily clinical practice and society, leading to the disease being undiagnosed and untreated. One of the reasons may be the lack of a broadly accepted definition of sarcopenia until the European Working Group on Sarcopenia in Older People (EWGSOP) developed a practical clinical definition and consensus diagnostic criteria for this disease in 2010 [4]. There is a growing consensus that sarcopenia should not be defined merely on the basis of muscle mass but also with regard to muscle strength and function [4]. However, few epidemiologic studies have been based on the EWGSOP definition of sarcopenia using population-based samples, and no epidemiologic study has investigated the relationship between exercise habits in middle age and sarcopenia in older age.

The Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study is a prospective cohort study aimed at elucidating the environmental and genetic background of musculoskeletal diseases [16, 17]. The present study investigated the prevalence of sarcopenia using the EWGSOP definition, and clarified the association of sarcopenia with exercise habits in middle age and physical performance in the elderly participants of Japanese population-based cohorts of the ROAD study.

Methods

Participants

From 2005–2007, we began a large-scale population-based cohort study entitled Research on Osteoarthritis/osteoporosis Against Disability consisting of 3,040 participants in three regions (baseline study) [16, 17]. The ROAD study is a prospective cohort study with the aim of elucidating the environmental and genetic backgrounds of musculoskeletal diseases. It is designed to examine the extent to which risk factors for these diseases are related to clinical features of the diseases, laboratory and radiographic findings, bone mass, bone geometry, lifestyle, nutritional factors, anthropometric and neuromuscular measures, and fall propensity. It also aims to determine how these diseases affect activities of daily living and quality of life of Japanese men and women. The subjects were residents of any one of three communities: an urban region in Itabashi, Tokyo; a mountainous region in Hidakagawa, Wakayama; and a coastal region in Taiji, Wakayama. The inclusion criteria were as follows: ability to (1) walk to the clinic where the survey was performed, (2) provide self-reported data, and (3) understand and sign an informed consent form. Participants from the urban region were aged ≥60 years and were recruited from those enrolled in a randomly selected cohort study from the previously established Itabashi Ward residential registration database [18]. Invitation letters were distributed only to inhabitants whose names were listed on this database. Participants from Hidakagawa and Taiji were aged ≥40 years and were recruited from residential registration listings. Residents aged <60 years from Itabashi and <40 years from Hidakagawa and Taiji who were interested in participating in the study were also invited. A total of 99.8, 84.3, and 54.7 % of the participants were aged ≥60 years in Itabashi, Hidakagawa, and Taiji, respectively. The response rates in the groups aged ≥60 years were 75.6 % in Itabashi, 68.4 % in Hidakagawa, and 29.3 % in Taiji. Two-thirds of the 3,040 participants in the baseline survey were women, and their mean age was 1 year less than that of the male participants. No significant difference was observed in body mass index (BMI) between the sexes.

After the baseline study, a second survey was performed in the same communities from 2008 to 2010, in which 2,674 inhabitants (892 men, 1,782 women) aged 21–97 years participated (second visit) [19]. Invitation letters were distributed to the inhabitants whose names were listed on the baseline database of the ROAD study. In addition to the former participants, inhabitants aged ≥60 years from Itabashi and those aged ≥40 years from Hidakagawa and Taiji who were willing to participate in the ROAD survey performed in 2008–2010 were also included in the second visit. In addition, residents aged <60 years from Itabashi and <40 years from Hidakagawa and Taiji who were interested in participating in the study were invited to be examined as well at the baseline. The inclusion criteria were as follows: ability to (1) walk to the clinic where the survey was performed, (2) provide self-reported data, and (3) understand and sign an informed consent form. No other exclusion criteria were used. Thus, 2,674 residents (892 men and 1,782 women) aged 21–97 years participated in the second visit. Of the 2,674 participants, 1,846 individuals aged ≥65 years visited the clinic and underwent an examination at the survey site located in Hidakagawa (504 individuals), Taiji (391 individuals), the University of Tokyo Hospital (132 individuals), or Tokyo Metropolitan Geriatric Hospital (819 individuals). For participants from Itabashi, the survey site was randomly assigned to either the University of Tokyo Hospital or Tokyo

Metropolitan Geriatric Hospital. Since gait speed was not measured at Tokyo Metropolitan Geriatric Hospital, 819 individuals who visited this hospital were removed from the present study. Of 1,846 participants, the remaining 1,019 individuals aged ≥65 years who visited the survey site located in Hidakagawa, Taiji, or at the University of Tokyo Hospital and underwent an examination including gait speed assessment were recruited for the present study. Of the 1,019 individuals, 19 were removed because 1 did not undergo handgrip strength measurement and 18 did not undergo skeletal muscle mass measurement. For the present study, we enrolled 1,000 participants (349 men and 651 women aged ≥65 years) from the second visit who completed assessment of handgrip strength, gait speed, and skeletal muscle mass. The mean age of the participants was 75.7 (SD, 5.9) years in men and 74.4 (SD, 6.1) years in women. All participants provided written informed consent, and the study was conducted with approval from the Ethics Committee of the University of Tokyo.

Participants completed an interviewer-administered questionnaire comprising 400 items regarding lifestyle information such as smoking habits, alcohol consumption, and physical activity. An interviewer asked the following question regarding past physical activity: "During the time you were aged 25–50 years, did you ever practice sports or physical exercise sufficient to produce sweating or shortness of breath?" Possible responses were as follows: never, occasionally, <2 hours per week, and ≥2 hours per week. Those who answered "occasionally, <2 hours per week, or ≥2 hours per week" were defined as having exercise habits in middle age. The following question was asked regarding current physical activity: "Do you practice walking more than 30 minutes every day?" Those who answered "yes" were defined as having a current walking habit.

Anthropometric and physical performance measurements

Anthropometric measurements, including height and weight, were obtained, and body mass index (weight [kg]/height [m^2]) was estimated based on the measured height and weight. Grip strength was measured on the right and left sides using a TOEI LIGHT handgrip dynamometer (TOEI LIGHT CO. LTD, Saitama, Japan), and the highest measurement was used to characterize maximum muscle strength. Subjects were defined as having low grip strength if grip strength was <30 kg in men and <20 kg in women, as reported by Lauretani and colleagues [20].

Skeletal muscle mass was measured by bioimpedance analysis [21–25] using the Body Composition Analyzer MC-190 (Tanita Corp., Tokyo, Japan). The protocol was described by Tanimoto and colleagues [10, 12], and the method has been validated [26]. Appendicular skeletal muscle mass (ASM) was derived as the sum of the muscle mass of the arms and the legs. Absolute ASM was converted to an appendicular muscle mass index (SMI) by dividing by height in meters squared (kg/m^2). Subjects were defined as having low skeletal muscle mass if the SMI was <2 SDs of the young adult mean. We used an SMI of <7.0 kg/m^2 in men and <5.8 kg/m^2 in women as cut-off points for low skeletal muscle mass based on the reference data of SMI measured by the MC-190 in 1,719 healthy young Japanese volunteers aged 18–39 years [10].

To measure physical performance, the time taken to walk 6 m at normal walking speed in a hallway was recorded, and usual gait speed was calculated. Subjects were defined as having low gait speed if usual gait speed was ≤0.8 m/s. The time taken for five consecutive chair rises without the use of hands was also recorded. Timing began with the command "Go" and ended when the buttocks contacted the chair on the fifth landing. One-leg standing time with eyes open was measured on both sides, and the best measurement was used. Participants were asked to stand on one leg while continuing to elevate their contralateral limb. Timing commenced when the participant assumed the correct posture and ended when any body part touched a supporting surface.

Statistical analysis

All statistical analyses were performed using STATA statistical software (STATA, College Station, TX). Differences in the values of the parameters between two groups were tested for significance using the nonpaired Student's *t* test and chi-square test. Trends in values were tested using the Jonckheere-Terpstra trend test. Factors associated with sarcopenia were determined using multivariate logistic regression analysis with sarcopenia as the dependent variable; the odds ratio (OR) and 95 % confidence interval were determined after adjusting for age, sex, and BMI. Factors associated with exercise habits in middle age were determined using multivariate linear regression analysis with exercise habits in middle age as the independent variable; the regression coefficient and 95 % CI were determined after adjusting for age, sex, and BMI.

Results

Table 1 shows the characteristics of the participants according to EWGSOP sarcopenia status. Age was significantly greater, while BMI, ASM, and SMI were significantly lesser in those with sarcopenia than in those without sarcopenia in both men and women. In physical performance, chair stand time was significantly greater and one-leg standing time was significantly lesser in those with sarcopenia than in those without sarcopenia in both men and women. The percentage of individuals with exercise habits in middle age was significantly lower in those with sarcopenia than in those without sarcopenia in both men and women.

Table 1 Characteristics of participants according to EWGSOP sarcopenia status

	Men		Women	
	No sarcopenia	Sarcopenia	No sarcopenia	Sarcopenia
No. of subjects	301	48	570	81
Age, years	75.1 (5.8)	79.9 (5.2)*	73.5 (5.6)	80.8 (5.8)*
Height, cm	161.9 (6.0)	158.5 (5.8)*	148.9 (6.4)	145.6 (6.6)*
Weight, kg	61.2 (9.5)	52.9 (6.5)*	52.4 (8.4)	42.6 (6.3)*
BMI, kg/m^2	23.3 (3.0)	21.0 (2.0)*	23.6 (3.3)	20.0 (2.3)*
ASM, kg	19.8 (3.0)	16.0 (1.7)*	13.8 (1.8)	11.4 (1.2)*
SMI, kg/m^2	7.54 (0.90)	6.36 (0.47)*	6.22 (0.66)	5.35 (0.30)*
Grip strength, kg	36.9 (6.8)	28.0 (4.0)*	23.9 (4.6)	16.8 (3.4)*
Usual gait speed, m/s	1.11 (0.25)	0.85 (0.27)*	1.06 (0.28)	0.82 (0.22)*
Chair stand time, s	9.6 (3.7)	11.9 (4.2)*	9.9 (4.2)	13.4 (5.9)*
One-leg standing time, median (IQR), s	31.0 (10.0–60.0)	8.0 (4.0–16.0)*	26.0 (8.0–60.0)	11.0 (5.0–23.0)*
Smoking, %	15.6	16.7	2.3	6.2
Alcohol consumption, %	58.8	45.8	14.7	18.8
Current walking habits, %	56.5	45.0	55.1	56.5
Exercise habits in middle age, %	69.9	46.2†	43.3	26.1†

Except where indicated otherwise, values are mean (SD)

ASM appendicular skeletal muscle mass, *BMI* body mass index, *EWGSOP* European Working Group on Sarcopenia in Older People, *IQR* interquartile range, *SMI* skeletal muscle mass index

*P < .001 vs. no sarcopenia in the same sex group by unpaired Student's t test; †P < .01 vs. no sarcopenia in the same sex group by chi-square test

Figure 1 shows sex- and age-wise distributions of prevalence of sarcopenia (Fig. 1a), low SMI (Fig. 1b), low grip strength (Fig. 1c), and low gait speed (Fig. 1d). The total prevalence of sarcopenia was 13.8 % in men and 12.4 % in women. Prevalence of sarcopenia (number of cases/subjects) in the age strata of 65–69, 70–74, 75–79, 80–84, and ≥85 years was 1.6 % (1/63), 5.7 % (5/88), 17.8 % (19/107), 23.2 % (16/69), and 31.8 % (7/22) in men and 0.6 % (1/163), 5.5 % (10/182), 13.8 % (22/160), 22.9 % (25/109), and 62.2 % (23/37) in women. Prevalence of sarcopenia tended to be significantly higher according to increasing age (P < .001 for trend) in both men and women. Prevalence of low grip strength and low gait speed also tended to be significantly higher according to increasing age (P < .001 for trend) in both men and women. However, the increasing tendency of prevalence of low SMI (P < .001 for trend) was milder compared with that of sarcopenia, low grip strength, and low gait speed.

Then, we determined the factors associated with sarcopenia by logistic regression analysis; the upper part of Table 2 shows the results using sarcopenia as the dependent variable. In the overall population, age (OR, 1.20; 95 % CI, 1.15–1.24) and BMI (OR, 0.68; 95 % CI, 0.63–0.75) were significantly associated with sarcopenia, whereas sex was not. In physical performance, chair stand time (OR, 1.09; 95 % CI, 1.04–1.14) and one-leg standing time (OR, 0.94; 95 % CI, 0.96–0.99) were significantly associated with sarcopenia in the overall population after adjusting for age, sex, and BMI. Current walking habit (OR, 0.69; 95 % CI, 0.42–1.12) was not significantly associated with sarcopenia. However, exercise habit in middle age (OR, 0.53; 95 % CI, 0.31–0.90) was associated with sarcopenia in the overall population after adjusting for age, sex, and BMI, indicating that exercise habit in middle age was significantly associated with low prevalence of sarcopenia in older age. The significance of the association did not change when current walking habit was added as an explanatory variable in this logistic regression model (OR, 0.53; 95 % CI, 0.32–0.90). In addition, we investigated the association of each category—occasionally, <2 h per week, and ≥2 h per week—with sarcopenia using "never" as a reference, in addition to the association of the presence of exercise habits in middle age with sarcopenia. The associated ORs for the three categories were comparable, but they did not reach significance level (occasionally: OR, 0.63; 95 % CI, 0.34–1.17; <2 h per week: OR, 0.30; 95 % CI, 0.09–1.01; ≥2 h per week: OR, 0.49; 95 % CI, 0.22–1.09).

The lower part of Table 2 shows the results of linear regression analysis using SMI, grip strength, gait speed, chair stand time, or one-leg standing time as the dependent variable and exercise habit in middle age as the independent variable. Exercise habit in middle age was significantly associated with grip strength in older age (P < .001), gait speed in older age (P < .001), and one-leg standing time in older age (P = .005) after adjusting for age, sex, and BMI in the overall population. We conducted the same analyses in men and women separately (Tables 3 and 4) and found results similar to those in the overall population. Some sex differences were observed in the present results. Exercise habit in middle age was significantly associated with grip strength and gait speed in older age in both men and women, whereas it was significantly associated with chair stand time and one-leg standing time only in men; however, the sample size of men was smaller than that of women. In the overall population, exercise habit in middle age was not associated with chair stand time.

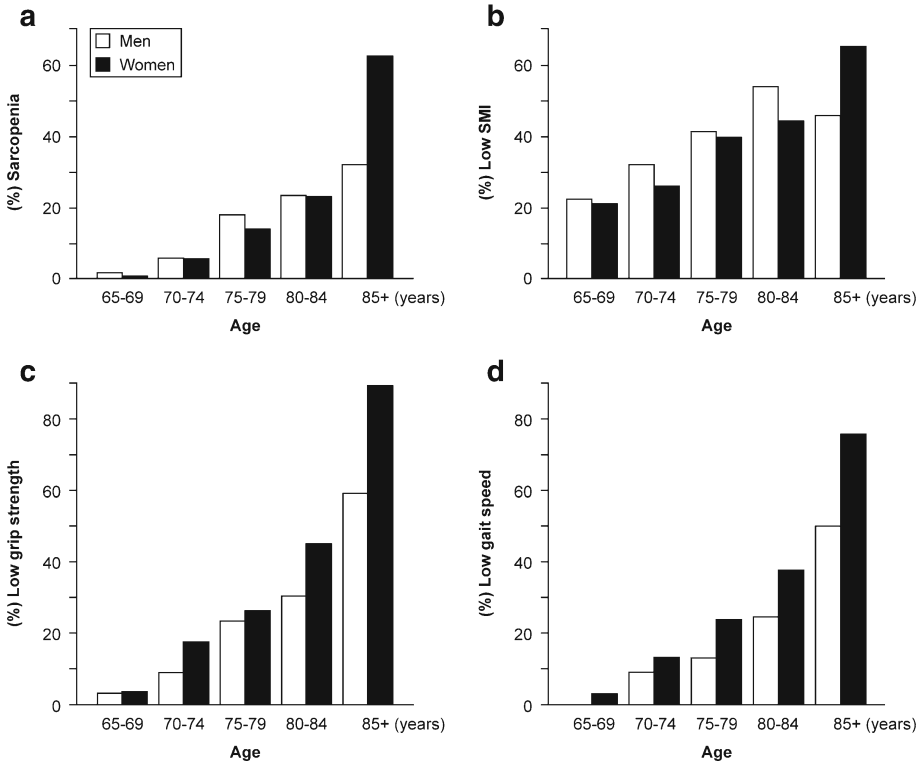


Fig. 1 Percentage of sarcopenia (**a**), low skeletal muscle mass index (*SMI*) (**b**), low grip strength (**c**), and low gait speed (**d**) in men and women in each age stratum (65–69, 70–74, 75–79, 80–85, and ≥85 years). Low SMI was defined as a value of <7.0 kg/m^2 in men and <5.8 kg/m^2 in women. Low grip strength was defined as a value of <30 kg in men and <20 kg in women. Low gait speed was defined as a value of ≤0.8 m/s

Discussion

The present study investigated the prevalence of sarcopenia using the EWGSOP definition in the elderly participants of Japanese population-based cohorts. We determined that age was positively associated with sarcopenia and that BMI was inversely associated, but sex was not. Exercise habit in middle age was associated with increased muscle strength and physical performance and low prevalence of sarcopenia in older age. To the best of our knowledge, this is the first study to show the relationship between exercise habits in middle age and sarcopenia in older age in the elderly participants of population-based cohorts.

Previous studies have reported the prevalence of sarcopenia and its associated factors. For example, Tanimoto and colleagues reported the prevalence of sarcopenia in

Table 2 Factors associated with sarcopenia and exercise habits in middle age in the overall population

Factors associated with sarcopenia	Odds ratio	95 % CI	*P* value
Age (+1 year)	1.20	1.15–1.24	<.001
Sex (women vs. men)	0.98	0.63–1.53	.9
BMI (+1 kg/m^2)	0.68	0.63–0.75	<.001
Chair stand time (+1 s)	1.09[a]	1.04–1.14	.001
One-leg standing time (+1 s)	0.97[a]	0.96–0.99	<.001
Smoking (yes vs. no)	1.86[a]	0.86–4.02	.1
Alcohol consumption (yes vs. no)	1.00[a]	0.60–1.67	.9
Current walking habits (yes vs. no)	0.69[a]	0.42–1.12	.1
Exercise habits in middle age (yes vs. no)	0.53[a]	0.31–0.90	.01
Factors associated with exercise habits in middle age	Regression coefficient	95 % CI	*P* value
SMI	0.09[b]	−0.02–0.19	.1
Grip strength	1.73[c]	1.02–2.44	<.001
Gait speed	0.07[c]	0.04–0.10	<.001
Chair stand time	−0.47[c]	−1.02–0.09	.09
One-leg standing time	4.14[c]	1.26–7.02	.005

BMI body mass index, *CI* confidence interval, *SMI* skeletal muscle mass index

[a] Odds ratio and 95 % CI were calculated by logistic regression analysis after adjusting for age, sex, and BMI

[b] Regression coefficient and 95 % CI were calculated by linear regression analysis after adjusting for age and sex

[c] Regression coefficient and 95 % CI were calculated by linear regression analysis after adjusting for age, sex, and BMI

Table 3 Factors associated with sarcopenia and exercise habits in middle age in men

Factors associated with sarcopenia	Odds ratio	95 % CI	*P* value
Chair stand time (+1 s)	1.09[a]	1.01–1.18	.03
One-leg standing time (+1 s)	0.97[a]	0.95–0.99	.001
Smoking (yes vs. no)	1.49[a]	0.59–3.75	.4
Alcohol consumption (yes vs. no)	0.78[a]	0.40–1.53	.4
Current walking habits (yes vs. no)	0.60[a]	0.28–1.27	.1
Exercise habits in middle age (yes vs. no)	0.48[a]	0.22–1.03	.06
Factors associated with exercise habits in middle age	Regression coefficient	95 % CI	*P* value
SMI	0.16[b]	−0.06 to 0.38	.1
Grip strength	3.17[c]	1.70 to 4.65	<.001
Gait speed	0.10[c]	0.04 to 0.15	.001
Chair stand time	−1.12[c]	−1.95 to −0.28	.009
One-leg standing time	7.81[c]	2.57 to 13.05	.004

CI confidence interval, *SMI* skeletal muscle mass index

[a] Odds ratio and 95 % CI were calculated by logistic regression analysis after adjusting for age and BMI

[b] Regression coefficient and 95 % CI were calculated by linear regression analysis after adjusting for age

[c] Regression coefficient and 95 % CI were calculated by linear regression analysis after adjusting for age and BMI

Japanese community-dwelling elderly individuals based on the EWGSOP definition using bioimpedance analysis (MC-190) [12]. They reported a prevalence of 11.3 % in men and 10.7 % in women [12], which is similar to our results. Although the cut-off value for low SMI was the same in these two studies, the cut-off value used for handgrip strength was different; we used cutoff values of <30 kg in men and <20 kg in women, in accordance with Lauretani and colleagues [20], while they used values of <30.3 kg in men and <19.3 kg in women, based on the lowest quartile of handgrip strength in their study population [12]. In the population of the present study, the lowest quartile of grip strength was 30.5 kg in men and 20.0 kg in women. Considering that these two studies showed similar results, cut-off values of 30 kg in men and 20 kg in women for handgrip strength [20] also may be appropriate for the practical case definition of the EWGSOP algorithm in the Japanese population.

Patel and colleagues reported the prevalence of sarcopenia in Caucasians using the EWGSOP definition, in which low muscle mass is defined as the lowest tertile of lean or fat-free

Table 4 Factors associated with sarcopenia and exercise habits in middle age in women

Factors associated with sarcopenia	Odds ratio	95 % CI	*P* value
Chair stand time (+1 s)	1.08[a]	1.02–1.15	.01
One-leg standing time (+1 s)	0.98[a]	0.96–1.00	.01
Smoking (yes vs. no)	2.44[a]	0.61–9.72	.2
Alcohol consumption (yes vs. no)	1.26[a]	0.58–2.71	.5
Current walking habits (yes vs. no)	0.75[a]	0.39–1.44	.3
Exercise habits in middle age (yes vs. no)	0.55[a]	0.27–1.13	.1
Factors associated with exercise habits in middle age	Regression coefficient	95 % CI	*P* value
SMI	0.06[b]	−0.05 to 0.17	.2
Grip strength	1.03[c]	0.29 to 1.78	.007
Gait speed	0.06[c]	0.01 to 0.10	.01
Chair stand time	−0.12[c]	−0.83 to 0.60	.7
One-leg standing time	2.19[c]	−1.24 to 5.62	.2

CI confidence interval, *SMI* skeletal muscle mass index

[a] Odds ratio and 95 % CI were calculated by logistic regression analysis after adjusting for age and BMI

[b] Regression coefficient and 95 % CI were calculated by linear regression analysis after adjusting for age

[c] Regression coefficient and 95 % CI were calculated by linear regression analysis after adjusting for age and BMI

mass [11]. They recommended use of the lowest tertile of muscle mass as a cut-off value if the reference value of muscle mass in a young healthy population is unavailable. In the population of the present study, the lowest tertile of SMI was 6.92 kg/m^2 in men and 5.80 kg/m^2 in women, which is similar to the cut-off value of <2 SDs of the young adult mean (7.0 kg/m^2 in men and 5.8 kg/m^2 in women) [10]. For evaluating low muscle mass, use of the lowest tertile may be an appropriate alternative method if the reference value of a young healthy population is unavailable.

The present study showed an association between sarcopenia and physical performance, including chair stand time and one-leg standing time, which is consistent with results of previous reports using the EWGSOP definition [11, 13]. However, these were comparisons between sarcopenia and current status of physical performance or exercise habit. Therefore, causal association was unclear whether sarcopenia was caused by decreased physical performance or activity or whether low physical performance or activity was due to sarcopenia. We also revealed that exercise habit in middle age was associated with increased muscle strength and physical performance and low prevalence of sarcopenia in older age. These results suggest that exercise habit in middle age is a protective factor against sarcopenia in older age and effective in maintaining muscle strength and physical performance in older age.

Some sex differences were observed in the present results. Exercise habit in middle age was significantly associated with grip strength and gait speed in older age in both men and women, whereas it was significantly associated with chair stand time and one-leg standing time only in men; however, the sample size of men was smaller than that of women. In the overall population, exercise habit in middle age was not associated with chair stand time; this finding may have been influenced by the fact that the sample size of women was almost twice that of men. The present results suggest that the impact of exercise habit in middle age on physical ability in older age is greater in men than in women.

Since exercise is a modifiable factor, it is a promising finding that exercise habit may be effective in preventing sarcopenia. In the present study, exercise habit was defined as physical activity in the period when the individual was aged 25–50 years, in which subjects practiced sports or physical exercise sufficient to produce sweating or shortness of breath, occasionally or more frequently. Although exercise habit was associated with low prevalence of sarcopenia at the age of ≥65 years, some details remain unclear, including exercise type, intensity, time, and other factors appropriate for prevention of sarcopenia. In addition to the association of the presence of exercise habit in middle age with sarcopenia, we further investigated the association of each category—occasionally, <2 h per week, and ≥2 h per week—with sarcopenia using "never" as a reference. Among the three categories, the analysis could not determine the best frequency and amount of exercise for protection from sarcopenia. The associated ORs for the three categories were comparable, and no dose–response tendency was seen in the relationship between frequency and amount of exercise and prevalence of sarcopenia; the associations also did not reach significance level. The present results suggest that abstaining from exercise during middle age is a risk factor for sarcopenia in older age. Furthermore, the presence of exercise habit in middle age might be much more important than the frequency and amount of exercise. Further studies are necessary to develop intervention programs and to test their effectiveness, along with accumulation of epidemiologic evidence including longitudinal studies.

The present study has several limitations. First, since this was a cross-sectional design, a causal relationship could not be determined. Second, information regarding exercise habits in middle age was obtained by self-report, and there is a possibility of recall bias. Third, the present study included participants who could walk to the survey site and could understand and sign an informed consent form. Since those who did not meet these inclusion criteria were not included in the analyses, the study participants do not truly represent the general population because of health bias. This should be considered when generalizing the results of the present study. Fourth, the results may have been affected by the characteristics of the population, including age and BMI. In the present study, age was positively associated with sarcopenia, whereas BMI was inversely associated with sarcopenia. Therefore, care should be taken when extrapolating the data to other populations with different characteristics, including age and BMI, which may confound the results.

In conclusion, the present study revealed prevalence of sarcopenia in the elderly participants of Japanese population-based cohorts. Exercise habit in middle age was associated with increased muscle strength and physical performance and low prevalence of sarcopenia in older age. These results suggest that exercise habit in middle age is a protective factor against sarcopenia in older age and is effective in maintaining muscle strength and physical performance in older age. Further long-term longitudinal epidemiological studies are necessary to develop effective intervention programs for the prevention and treatment of sarcopenia.

Acknowledgments This study was supported by Grants-in-Aid for Scientific Research (S19109007, B20390182, B23390172, B23390356, and B23390357) from the Japanese Ministry of Education, Culture, Sports, Science and Technology; H17-Men-eki-009, H18-Choujyu-037, H20-Choujyu-009, H21-Chouju-Wakate-011, H22-Chouju-Wakate-007, and H23-Chouju-002 from the Ministry of Health, Labour and Welfare; and Research Aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006–1 and 2010–2).

Conflicts of interest None.

References

1. Rosenberg I (1989) Summary comments: epidemiological and methodological problems in determining nutritional status of older persons. Am J Clin Nutr 50:1231–1233
2. Rosenberg IH (1997) Sarcopenia: origins and clinical relevance. J Nutr 127(5 Suppl):990S–991S
3. Morley JE, Baumgartner RN, Roubenoff R, Mayer J, Nair KS (2001) Sarcopenia. J Lab Clin Med 137(4):231–243
4. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, Martin FC, Michel JP, Rolland Y, Schneider SM, Topinkova E, Vandewoude M, Zamboni M (2010) European Working Group on Sarcopenia in Older People. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on Sarcopenia in Older People. Age Ageing 39(4):412–423
5. Delmonico MJ, Harris TB, Lee JS, Visser M, Nevitt M, Kritchevsky SB, Tylavsky FA, Newman AB, Health, Aging and Body Composition Study (2007) Health, Aging and Body Composition Study. Alternative definitions of sarcopenia, lower extremity performance, and functional impairment with aging in older men and women. J Am Geriatr Soc 55(5):769–774
6. Goodpaster BH, Park SW, Harris TB, Kritchevsky SB, Nevitt M, Schwartz AV, Simonsick EM, Tylavsky FA, Visser M, Newman AB (2006) The loss of skeletal muscle strength, mass, and quality in older adults: the health, aging and body composition study. J Gerontol A Biol Sci Med Sci 61(10):1059–1064
7. Baumgartner RN, Koehler KM, Gallagher D, Romero L, Heymsfield SB, Ross RR, Garry PJ, Lindeman RD (1998) Epidemiology of sarcopenia among the elderly in New Mexico. Am J Epidemiol 147(8):755–763
8. Melton LJ 3rd, Khosla S, Crowson CS, O'Connor MK, O'Fallon WM, Riggs BL (2000) Epidemiology of sarcopenia. J Am Geriatr Soc 48(6):625–630
9. Iannuzzi-Sucich M, Prestwood KM, Kenny AM (2002) Prevalence of sarcopenia and predictors of skeletal muscle mass in healthy, older men and women. J Gerontol A Biol Sci Med Sci 57(12):M772–M777
10. Tanimoto Y, Watanabe M, Sun W, Sugiura Y, Tsuda Y, Kimura M, Hayashida I, Kusabiraki T, Kono K (2012) Association between sarcopenia and higher-level functional capacity in daily living in community-dwelling elderly subjects in Japan. Arch Gerontol Geriatr 55(2):e9–e13
11. Patel HP, Syddall HE, Jameson K, Robinson S, Denison H, Roberts HC, Edwards M, Dennison E, Cooper C, Aihie Sayer A (2013) Prevalence of sarcopenia in community-dwelling older people in the UK using the European Working Group on Sarcopenia in Older People (EWGSOP) definition: findings from the Hertfordshire Cohort Study (HCS). Age Ageing 42(3):378–384
12. Tanimoto Y, Watanabe M, Sun W, Tanimoto K, Shishikura K, Sugiura Y, Kusabiraki T, Kono K (2013) Association of sarcopenia with functional decline in community-dwelling elderly subjects in Japan. Geriatr Gerontol Int. doi:10.1111/ggi.12037
13. Lin CC, Lin WY, Meng NH, Li CI, Liu CS, Lin CH, Chang CK, Lee YD, Lee CC, Li TC (2013) Sarcopenia prevalence and associated factors in an elderly Taiwanese metropolitan population. J Am Geriatr Soc 61(3):459–462
14. National Institute of Population and Society Research. Population projections for Japan (January 2012): 2011 to 2060. http://www.ipss.go.jp/site-ad/index_english/esuikei/gh2401e.asp. Accessed 30 May 2013
15. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2010. http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/4-2.html. Accessed 30 May 2013
16. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T (2010) Cohort profile: research on Osteoarthritis/Osteoporosis Against Disability study. Int J Epidemiol 39(4):988–995
17. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Yamamoto S, Ishibashi H, Kawaguchi H, Nakamura K, Akune T (2009) Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 27(5):620–628
18. Shimada H, Lord SR, Yoshida H, Kim H, Suzuki T (2007) Predictors of cessation of regular leisure-time physical activity in community-dwelling elderly people. Gerontology 53(5):293–297
19. Yoshimura N, Oka H, Muraki S, Akune T, Hirabayashi N, Matsuda S, Nojiri T, Hatanaka K, Ishimoto Y, Nagata K, Yoshida M, Tokimura F, Kawaguchi H, Nakamura K (2011) Reference values for hand grip strength, muscle mass, walking time, and one-leg standing time as indices for locomotive syndrome and associated disability: the second survey of the ROAD study. J Orthop Sci 16(6):768–777
20. Lauretani F, Russo CR, Bandinelli S, Bartali B, Cavazzini C, Di Iorio A, Corsi AM, Rantanen T, Guralnik JM, Ferrucci L (2003) Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. J Appl Physiol 95(5):1851–1860
21. No authors listed (1996) Bioelectrical impedance analysis in body composition measurement: National Institutes of Health Technology Assessment Conference Statement. Am J Clin Nutr 64(3 Suppl):524S-532S
22. Janssen I, Heymsfield SB, Baumgartner RN, Ross R (2000) Estimation of skeletal muscle mass by bioelectrical impedance analysis. J Appl Physiol 89(2):465–471
23. Kyle UG, Genton L, Slosman DO, Pichard C (2001) Fat-free and fat mass percentiles in 5225 healthy subjects aged 15 to 98 years. Nutrition 17(7–8):534–541
24. Kyle UG, Genton L, Karsegard L, Slosman DO, Pichard C (2001) Single prediction equation for bioelectrical impedance analysis in adults aged 20–94 years. Nutrition 17(3):248–253
25. Roubenoff R, Baumgartner RN, Harris TB, Dallal GE, Hannan MT, Economos CD, Stauber PM, Wilson PW, Kiel DP (1997) Application of bioelectrical impedance analysis to elderly populations. J Gerontol A Biol Sci Med Sci 52(3):M129–M136
26. Nemoto M, Yasbushita N, Kim M, Tomoaki M, Satoshi S, Jung S, Hiroyuki S, Kiyoji T (2012) Validity of the bioelectrical impedance method for assessing body composition in non-frail and pre-frail older adults. Int J Body Comps Res 10:225–262

J Orthop Sci (2014) 19:489–496
DOI 10.1007/s00776-014-0537-z

ORIGINAL ARTICLE

Association of physical activities of daily living with the incidence of certified need of care in the long-term care insurance system of Japan: the ROAD study

Toru Akune · Shigeyuki Muraki · Hiroyuki Oka · Sakae Tanaka · Hiroshi Kawaguchi · Fumiaki Tokimura · Hideyo Yoshida · Takao Suzuki · Kozo Nakamura · Noriko Yoshimura

Received: 26 August 2013 / Accepted: 16 January 2014 / Published online: 8 February 2014

Abstract
Background The present study aimed to investigate association of physical activities of daily living with the incidence of certified need of care in the national long-term care insurance (LTCI) system in elderly Japanese population-based cohorts.
Methods Of the 3,040 participants in the baseline examination, we enrolled 1,773 (699 men, 1,074 women) aged 65 years or older who were not certified as in need of care-level elderly at baseline. Participants were followed during an average of 4.0 years for incident certification of need of care in the LTCI system. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was used assess function. Associated factors in the baseline examination with the occurrence were determined by multivariate Cox proportional hazards regression analysis. Receiver operating characteristic curve analysis was performed to evaluate cut-off values for discriminating between the occurrence and the non-occurrence group.
Results All 17 items in the WOMAC function domain were significantly associated with the occurrence of certified need of care in the overall population. Cut-off values of the WOMAC function score that maximized the sum of sensitivity and specificity were around 4–6 in the overall population, in men, and in women. Multivariate Cox hazards regression analysis revealed that a WOMAC function score $\geq$4 was significantly associated with occurrence with the highest hazard ratio (HR) for occurrence after adjusting for confounders in the overall population (HR [95 % confidence interval (CI)] 2.54 [1.76–3.67]) and in women [HR (95 % CI) 3.13 (1.95–5.02)]. A WOMAC function score $\geq$5 was significantly associated with the highest HR for occurrence in men [HR (95 % CI) 1.88 (1.03–3.43)].
Conclusions Physical dysfunction in daily living is a predictor of the occurrence of certified need of care. Elderly men with a WOMAC function score $\geq$5 and women with a score $\geq$4 should undergo early intervention programs to prevent subsequent deterioration.

T. Akune (✉) · S. Muraki
Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
e-mail: akune-ort@h.u-tokyo.ac.jp

H. Oka · N. Yoshimura
Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

S. Tanaka · H. Kawaguchi
Department of Sensory and Motor System Medicine, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

F. Tokimura
Department of Orthopaedic Surgery, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

H. Yoshida
Research Team for Promoting Independence of the Elderly, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan

T. Suzuki
Research Institute, National Center for Geriatrics and Gerontology, Aichi, Japan

K. Nakamura
National Rehabilitation Center for Persons with Disabilities, Saitama, Japan

Introduction

Japan is a super-aged society experiencing an unprecedented aging of the population. The proportion of the population aged 65 years or older was 23 % in 2010, and is expected to reach 30.1 % in 2024 and 39 % in 2051 [1]. This leads to an increasing proportion of disabled elderly requiring support or long-term care, imposing enormous economic and social burdens on the country. The Japanese Government started the national long-term care insurance (LTCI) system in 2000 based on the Long-Term Care Insurance Act [2]. The aim was to certify need of care-level elderly and to provide suitable care services according to the level of care required [7 levels, including requiring support (levels 1 and 2) and requiring long-term care (levels 1–5)]. The total number of certified need of care-level elderly was reported to be 5 million in 2011 [2]. Certification of need of care in the national LTCI system is an important outcome in Japan not only because of its massive social and economic burdens, but also because it is urgently necessary to reduce risk and decrease the number of disabled elderly requiring care in their activities of daily living (ADLs). It is critically important to accumulate epidemiologic evidence, including identification of predictors, to establish evidence-based prevention strategies. However, no studies have determined the association of physical ADLs with the incidence of certified need of care in the national LTCI system using large-scale, population-based cohorts. The objective of the present study was to investigate the association of physical ADLs with the incidence of certified need of care in the national LTCI system and determine its predictors in elderly participants of large-scale, population-based cohorts of the research on osteoarthritis/osteoporosis against disability (ROAD) study.

Subjects and methods

Participants

The analysis was based on data collected from cohorts established in 2005 for the ROAD study. Details of the cohorts have been reported elsewhere [3, 4]. Briefly, a baseline database was created from 2005 to 2007, which included clinical and genetic information on 3,040 residents of Japan (1,061 men, 1,979 women). Participants were recruited from resident registration listings in three communities, namely, an urban region in Itabashi, Tokyo, and rural regions in Hidakagawa and Taiji, Wakayama. Participants in the urban region in Itabashi were recruited from those of a cohort study [5] in which the participants were randomly drawn from the register database of Itabashi ward residents, with a response rate in the age group >60 years of 75.6 %. Participants in the rural regions in Hidakagawa and Taiji were recruited from resident registration lists, with response rates in the groups aged >60 years of 68.4 and 29.3 %, respectively. Inclusion criteria were the ability to (1) walk to the survey site, (2) report data, and (3) understand and sign an informed consent form. For the present study, we enrolled 1,773 participants (699 men, 1,074 women; mean age 75.4 years) aged 65 years or older who were not certified as in need of care-level elderly in the national LTCI system at baseline. All participants provided written informed consent, and the study was conducted with approval from the ethics committees of the participating institutions.

Baseline procedures

Participants completed an interviewer-administered questionnaire containing 400 items that included lifestyle information, such as smoking habits, alcohol consumption, and physical activity. At baseline, anthropometric measurements, including height and weight, were taken, and body mass index (BMI) [weight (kg)/height^2 (m^2)] was estimated based on the measured height and weight.

Assessment of physical ADLs

We used the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) for assessment of physical ADLs. The WOMAC is a health status instrument, consisting of three domains: pain, stiffness, and physical function. We used the WOMAC function domain to evaluate physical ADLs. It consisted of 17 items: assessing difficulties in descending stairs, ascending stairs, rising from sitting, standing, bending to floor, walking on a flat surface, getting in/out of car/bus, going shopping, putting on socks/stockings, rising from bed, taking off socks/stockings, lying in bed, getting into/out of bath, sitting, getting on/off toilet, heavy domestic duties, and light domestic duties. Each item in the domain is graded on either a 5-point Likert scale (scores of 0–4) or a 100-mm visual analog scale [6, 7]. In the present study, we used the Likert scale (version LK 3.0). Items were rated from 0 to 4; 0, no difficulty; 1, mild difficulty; 2, moderate difficulty; 3, severe difficulty; 4, extreme difficulty. The domain score ranges from 0 to 68. Japanese versions of the WOMAC have been validated [8].

Certification of need of care in the LTCI system

The nationally uniform criteria for long-term care need certification was established objectively by the Japanese Government, and certification of need of care-level elderly

is determined based on evaluation results by the Certification Committee for Long-term Care Need in municipalities in accordance with basic guidelines formulated by the Government. The process of eligibility for certification of need of care in the LTCI system was described in detail by Chen et al. [9]. An elderly person who requires help with ADLs or the caregiver contacts the municipal government to request official certification of care needs. After the application, a trained official visits the home to assess the current physical status of the elderly person, including presence or absence of muscle weakness or joint contracture of limbs, and difficulties in sitting-up, standing-up, maintaining sitting or standing position, transferring from one place to another, standing on one leg, walking, bathing, dressing, and other ADLs. Mental status, including dementia, also is assessed. These data are analyzed to calculate a standardized score for determination of the level of care needs (certified support, levels 1–2; or long-term care, levels 1–5). In addition, the primary physician of the applicant assesses physical and mental status, including information on diseases causing ADL disability and the extent of disabilities caused by them. Finally, the Certification Committee for Long-term Care Need reviews the data and determines the certification and its level.

Follow-up and definition of incident certified need of care

After the baseline ROAD survey, participants who were not certified as in need of care-level elderly at baseline were followed for incident certification of need of care in the LTCI system. Incident certified need of care was defined as the incident certified 7 levels, including requiring support (levels 1–2) and requiring long-term care (levels 1–5). Information on the presence or absence of certification of need of care and its date of occurrence were collected by the resident registration listings in three communities every year up to 2010, and were used for analyses in the present study.

Statistical analysis

All statistical analyses were performed using STATA statistical software (STATA, College Station, TX, USA). Differences in values of the parameters between the two groups were tested for significance using the unpaired Student's *t* test, the Mann–Whitney's *U* test, and Chi-square test. We used receiver operating characteristic (ROC) curve analysis to determine a cut-off value of the WOMAC function score for discriminating two distinct groups: an occurrence and a non-occurrence group of certified need of care. Cut-off values were determined that maximized the sum of sensitivity and specificity. Factors associated with the occurrence of certified need of care were determined using Cox proportional hazards regression analysis; hazard ratios (HRs) and 95 % confidence intervals (CIs) were determined after adjusting for region, age, sex, and BMI. Smoking habit and alcohol consumption were not included as confounders because they were not significantly associated with the incidence of certified need of care.

Results

Of the 1,773 participants who were not certified as in need of care-level elderly at baseline, information on

Table 1 Baseline characteristics of population at risk for the certified need of care in the LTCI system

	Men	Women
No. of subjects	699	1,074
Age (years)	75.6 (5.1)	75.2 (5.3)
Height (cm)	160.9 (6.0)	147.9 (6.0)[b]
Weight (kg)	59.4 (9.1)	50.0 (8.3)[b]
BMI (kg/m^2)	22.9 (2.9)	22.8 (3.4)
Smoking (%)	21.0	3.2[c]
Alcohol consumption, %	61.2	23.0[c]
WOMAC function domain		
Descending stairs, pts[a]	0 (0, 0, 1, 1)	0 (0, 0, 1, 2)[d]
Ascending stairs, pts[a]	0 (0, 0, 1, 1)	0 (0, 0, 1, 2)
Rising from sitting, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 1, 1)[d]
Standing, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 1, 1)[d]
Bending to floor, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 1, 1)
Walking on a flat surface, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 0, 1)
Getting in/out of car/bus, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 1, 1)[d]
Going shopping, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 0, 1)[d]
Putting on socks/stockings, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 0, 1)[d]
Rising from bed, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 0, 1)[d]
Taking off socks/stockings, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 0, 1)[d]
Lying in bed, pts[a]	0 (0, 0, 0, 0)	0 (0, 0, 0, 1)[d]
Getting into/out of bath, pts[a]	0 (0, 0, 0, 0)	0 (0, 0, 0, 1)[d]
Sitting, pts[a]	0 (0, 0, 0, 0)	0 (0, 0, 0, 0)[d]
Getting on/off toilet, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 1, 2)[d]
Heavy domestic duties, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 0, 1)[d]
Light domestic duties, pts[a]	0 (0, 0, 0, 1)	0 (0, 0, 0, 1)[d]
Total, pts[a]	1 (0, 0, 5, 12)	2 (0, 0, 8, 17)[d]

Except where indicated otherwise, values are mean (SD)

LTCI long-term care insurance system, *BMI* body mass index, *WOMAC* the Western Ontario and McMaster Universities Arthritis Index

[a] Median (10, 25, 75, and 90 percentile)

[b] $P < 0.05$ vs men by unpaired Student's *t* test

[c] $P < 0.05$ vs men by Chi-square test

[d] $P < 0.05$ vs men by Mann–Whitney *U* test

Table 2 Association of physical activities of daily living with the occurrence of certified need of care in the LTCI system

Physical activity	Overall population		Men		Women	
	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value
Descending stairs, pts	1.47 (1.26, 1.72)	<0.001	1.29 (0.96, 1.74)	0.089	1.56 (1.30, 1.87)	<0.001
Ascending stairs, pts	1.47 (1.25, 1.73)	<0.001	1.29 (0.93, 1.77)	0.123	1.55 (1.29, 1.86)	<0.001
Rising from sitting, pts	1.58 (1.34, 1.88)	<0.001	1.38 (0.95, 1.99)	0.092	1.67 (1.37, 2.03)	<0.001
Standing, pts	1.64 (1.41, 1.91)	<0.001	1.39 (1.02, 1.90)	0.037	1.73 (1.45, 2.06)	<0.001
Bending to floor, pts	1.57 (1.32, 1.85)	<0.001	1.61 (1.15, 2.27)	0.006	1.57 (1.29, 1.90)	<0.001
Walking on a flat surface, pts	1.57 (1.30, 1.90)	<0.001	1.25 (0.88, 1.77)	0.22	1.78 (1.41, 2.23)	<0.001
Getting in/out of car/bus, pts	1.76 (1.47, 2.10)	<0.001	1.60 (1.14, 2.26)	0.007	1.85 (1.50, 2.29)	<0.001
Going shopping, pts	1.72 (1.46, 2.03)	<0.001	1.55 (1.14, 2.11)	0.005	1.81 (1.48, 2.21)	<0.001
Putting on socks/stockings, pts	1.60 (1.33, 1.92)	<0.001	1.41 (0.98, 2.03)	0.065	1.71 (1.37, 2.12)	<0.001
Rising from bed, pts	1.68 (1.40, 2.03)	<0.001	1.41 (0.98, 2.02)	0.066	1.83 (1.47, 2.29)	<0.001
Taking off socks/stockings, pts	1.64 (1.37, 1.98)	<0.001	1.48 (1.01, 2.16)	0.046	1.72 (1.39, 2.13)	<0.001
Lying in bed, pts	1.82 (1.44, 2.30)	<0.001	1.96 (1.13, 3.40)	0.017	1.79 (1.38, 2.32)	<0.001
Getting into/out of bath, pts	1.71 (1.43, 2.04)	<0.001	1.64 (1.15, 2.33)	0.006	1.75 (1.43, 2.15)	<0.001
Sitting, pts	2.21 (1.73, 2.82)	<0.001	1.92 (1.14, 3.22)	0.014	2.32 (1.75, 3.06)	<0.001
Getting on/off toilet, pts	1.87 (1.52, 2.29)	<0.001	1.51 (1.00, 2.27)	0.05	2.09 (1.63, 2.68)	<0.001
Heavy domestic duties, pts	1.27 (1.09, 1.49)	0.003	1.20 (0.89, 1.62)	0.238	1.33 (1.10, 1.60)	0.003
Light domestic duties, pts	1.68 (1.41, 2.01)	<0.001	1.49 (1.07, 2.07)	0.019	1.80 (1.45, 2.24)	<0.001

Hazard ratios (HRs) and 95 % confidence intervals (CIs) were determined by Cox proportional hazards regression analysis after adjusting for age, sex, body mass index, and region in the overall population, and after adjusting for age, body mass index, and region in men and in women, respectively

LTCI long-term care insurance system

certification of need of care could be obtained in 1,760 (99.3 %) during the average 4.0-year follow-up. Fifty-four men and 115 women were certified as in need of care-level elderly in the national LTCI system, whereas, 1,591 remained uncertified during the follow-up period. The average period for the certification was 2.3 years. Among the above 54 men and 115 women, those who were certified as requiring long-term care level 1, 2, 3, 4, and 5 were 7, 9, 2, 4, 3 men, and 12, 17, 9, 4, 4 women, respectively. One hundred and twenty-six participants died and eight moved away. Incidence of certified need of care in the LTCI system was 2.3/100 person-years in the overall population, and 2.0/100 person-years in men and 2.5/100 person-years in women. Table 1 shows the baseline characteristics of the population at risk for occurrence of certified need of care in the LTCI system. The score of each item in the WOMAC function domain was significantly higher in women than in men in almost all items.

We then investigated association of each item in the WOMAC function domain with the occurrence of certified need of care in the LTCI system (Table 2). All 17 items in the WOMAC function domain were significantly associated with the occurrence of the certified need of care in the overall population and in women. In men, standing, bending to floor, getting in/out of car/bus, going shopping, taking off socks/stockings, lying in bed, getting into/out of bath, sitting, and light domestic duties were significantly associated with the occurrence of certified need of care, whereas other ADLs were not. In addition, the value of HR for each item in the association was higher in women than in men in 15 of 17 items.

Next we determined cut-off values of total score of the WOMAC function domain for discriminating two groups: an occurrence and a non-occurrence group of certified need of care using ROC curve analysis. The area under ROC curve was 0.70 in the overall population, 0.61 in men, and 0.74 in women (Fig. 1). The cut-off value of the WOMAC function score that maximized the sum of sensitivity and specificity was 6, 5, and 6 in the overall population, in men, and in women, respectively. In addition, the sensitivity/specificity was 57.3/75.0 % in the overall population, 45.7/75.0 % in men, and 64.4/72.6 % in women, respectively (Table 3). Furthermore, the cut-off value by which the sum was the second largest was 4 in the overall population, 4 in men, and 4 in women, and the sensitivity/specificity was 65.3/66.7 % in the overall population, 50.0/70.0 % in men, and 72.1/64.5 % in women, respectively (Table 3).

Because ROC curve analysis is a univariate analysis, we performed multivariate Cox hazards regression analysis to determine the cut-off value of the WOMAC function score for best discriminating between an occurrence and a non-

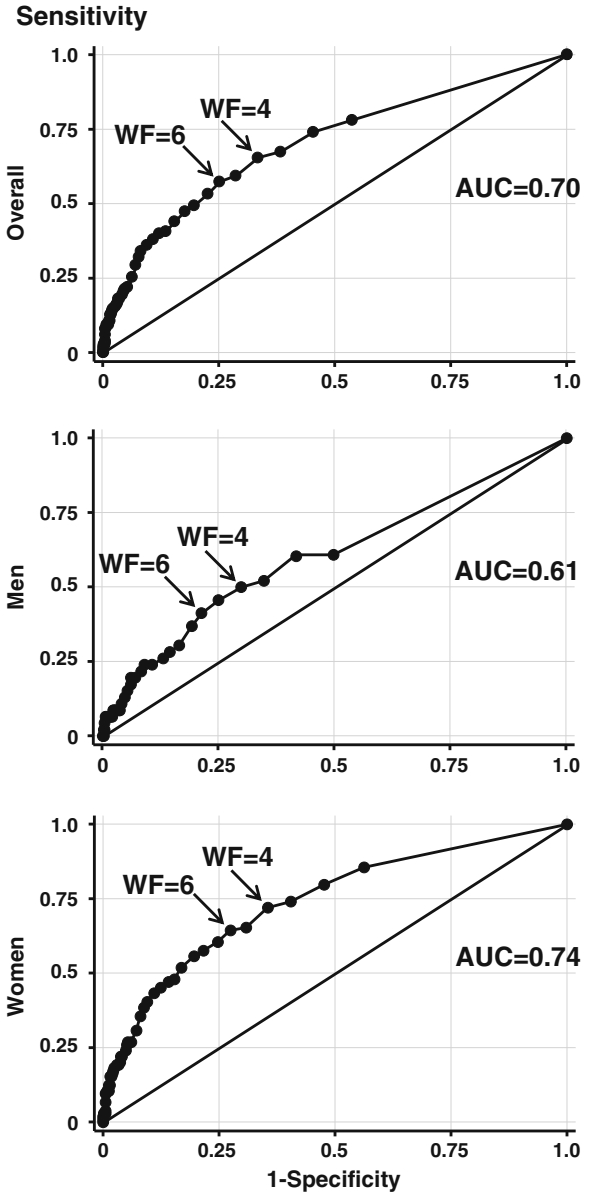


Fig. 1 Receiver operating characteristic (ROC) curve analysis for discriminating the occurrence group of certified need of care in the overall population, in men, and in women. *AUC* area under ROC curve, *WF* WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index) function score

occurrence group of certified need of care after adjusting for age, sex, BMI, and region (Table 4). The group with WOMAC function score ≥4 was significantly associated with the occurrence of certified need of care compared with the group with the score <4 with the highest HR in the overall population [HR 2.54, 95 % CI (1.76–3.67)] and in women [HR 3.13, 95 % CI (1.95–5.02)]. In men, the group with WOMAC function score ≥5 was significantly

Table 3 Sensitivity and specificity of the occurrence of certified need of care determined by the cut-off point of the WOMAC function score

Cut-off point	Overall population			Men			Women		
	Sensitivity (%)	Specificity (%)	Sensitivity + specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity + specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity + specificity (%)
WF = 4pts	65.3	66.7	132.0	50.0	70.0	120.0	72.1	64.5	136.6
WF = 5pts	59.3	71.4	130.7	45.7	75.0	120.7	65.4	69.2	134.6
WF = 6pts	57.3	75.0	132.3	41.3	78.6	119.9	64.4	72.6	137.0

WOMAC the Western Ontario and McMaster Universities Arthritis Index, *WF* WOMAC function score

Table 4 Association of groups divided by the WOMAC function score with the occurrence of certified need of care in the LTCI system

	Overall population		Men		Women	
	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value
WF ≥ 4 pts vs WF < 4 pts	2.54 (1.76, 3.67)	<0.001	1.85 (1.01, 3.39)	0.045	3.13 (1.95, 5.02)	<0.001
WF ≥ 5 pts vs WF < 5 pts	2.35 (1.64, 3.36)	<0.001	1.88 (1.03, 3.43)	0.040	2.71 (1.73, 4.27)	<0.001
WF ≥ 6 pts vs WF < 6 pts	2.50 (1.75, 3.58)	<0.001	1.84 (1.00, 3.39)	0.051	3.03 (1.93, 4.76)	<0.001

Hazard ratios (HRs) and 95 % confidence intervals (CIs) were determined by Cox proportional hazards regression analysis after adjusting for age, sex, body mass index, and region in the overall population, and after adjusting for age, body mass index, and region in men and in women, respectively

WOMAC the Western Ontario and McMaster Universities Arthritis Index, *LTCI* long-term care insurance system, *WF* WOMAC function score

Table 5 Association of the WOMAC function score with the occurrence of different certified need of care levels in the LTCI system

Outcome variable	Overall population		Men		Women	
	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value	HR (95 % CI)	*P* value
RSL1–2 and RCL 1–5	1.05 (1.03, 1.06)	<0.001	1.03 (1.01, 1.06)	0.008	1.05 (1.04, 1.07)	<0.001
RCL 1–5	1.05 (1.03, 1.07)	<0.001	1.04 (1.00, 1.07)	0.046	1.06 (1.03, 1.08)	<0.001
RCL 2–5	1.06 (1.04, 1.08)	<0.001	1.04 (1.01, 1.08)	0.015	1.06 (1.04, 1.09)	<0.001
RCL 3–5	1.05 (1.03, 1.08)	<0.001	1.05 (0.99, 1.10)	0.099	1.06 (1.02, 1.09)	0.001
RCL 4–5	1.04 (1.00, 1.08)	0.048	1.02 (0.95, 1.10)	0.501	1.05 (1.00, 1.10)	0.057
RCL 5	1.01 (0.93, 1.09)	0.830	0.99 (0.82, 1.20)	0.945	1.01 (0.93, 1.11)	0.780

Hazard ratios (HRs) and 95 % confidence intervals (CIs) were determined by Cox proportional hazards regression analysis after adjusting for age, sex, body mass index, and region in the overall population, and after adjusting for age, body mass index, and region in men and in women, respectively

WOMAC the Western Ontario and McMaster Universities Arthritis Index, *LTCI* long-term care insurance system, *RSL* requiring support level, *RCL* requiring long-term care level

associated with the occurrence of certified need of care compared with the group with a score of <5 with the highest HR [HR 1.88, 95 % CI (1.03–3.43)].

Furthermore, we examined association of the WOMAC function domain with the occurrence of different certified need of care levels in the LTCI system (Table 5). When the outcome variable of the occurrence was defined as requiring support level (RSL) 1–2 and requiring long-term care level (RCL) 1–5, RCL 1–5, and RCL 2–5, there were significant associations in the overall population, in men, and in women, respectively. When the outcome variable of the occurrence was defined as RCL 3–5, there were significant associations in the overall population and in women. When the outcome variable of the occurrence was defined as RCL 4–5, there was significant association in the overall population.

Discussion

The present study determined association of physical ADLs with the incidence of certified need of care in the national LTCI system in elderly participants of Japanese population-based cohorts. All 17 items in the WOMAC function domain were significantly associated with the occurrence of certified need of care in the overall population. ROC curve analysis showed that cut-off values of the WOMAC function score of around 4–6 maximized the sum of sensitivity and specificity of the occurrence of certified need of care. Furthermore, multivariate Cox hazards regression analysis revealed that the group with WOMAC function score ≥4 was significantly associated with the occurrence of certified need of care with the highest HR after adjusting for confounders in the overall population and in women, while the group with WOMAC function score ≥5 was significantly associated with the highest HR in men.

In the present study, we could not obtain information on causes of certified need of care in the LTCI system. Therefore, we could not analyze the direct association of each causing condition with the WOMAC function domain. The Government of Japan reported that the top five leading causes of certified need of care were cerebral stroke (21.5 %), dementia (15.3 %), asthenia as a result of older age (13.7 %), joint disease (10.9 %) and fall-related fracture (10.2 %), comprising 71.6 % of all causes in 2010 [10]. Based on these data, most of the causes of incident certification in the present study are inferred to be among the top five leading conditions. Although we could not

know the exact percentage of each causing condition, joint disease and fall-related fracture are inferred to represent approximately 20 % in total causes of incident certification in the present study, and cerebral stroke, dementia, and asthenia as a result of older age are inferred to represent approximately 50 % in total causes of incident certification.

The Government of Japan also reported that the percentage of joint disease and fall-related fracture was 16.7 % for the cause of RCL 1–5 [10]. Furthermore, it was 17.6, 19.8, 14.8, 17.4, and 9.8 % for the cause of RCL 1, 2, 3, 4, and 5, respectively [10]. Although we could not know the exact percentage of joint disease and fall-related fracture for the cause of each RCL in the present study, the percentage for the cause of RCL 1–4 is inferred to be approximately 15 % or more based on the data of the Government of Japan, which may be the reason why the WOMAC domain was significantly associated with the occurrence of certified need of care including RCLs 1–4 in the overall population.

The WOMAC physical function domain assesses difficulties in ADLs, including going up/down stairs, getting in/out of a car and bath, shopping, and household duties. Therefore, results of the present study indicate that the severity of physical dysfunction in ADLs predicts subsequent deterioration in ADLs, leading to the occurrence of certified need of care. Previous studies reported that low physical function was a predictor of subsequent ADL disability in the elderly [11, 12]. Although no previous studies have investigated the association of physical ADLs with the incidence of certified need of care in the national LTCI system in large-scale population-based cohorts, those previous findings are consistent with the present results in that low physical activity predicted subsequent deterioration in ADLs.

All 17 items in the WOMAC domain were significantly associated with the occurrence of certified need of care in women. On the other hand, 9 of 17 items were significantly associated with the occurrence of certified need of care in men. In addition, the HR for each item in the association was higher in women than in men for 15 of 17 items. The sex difference identified in this association may be due to the difference in the prevalence of knee osteoarthritis between the sexes. Muraki et al. [13] reported that prevalence of radiographic knee osteoarthritis determined by the Kellgren–Lawrence grade $\geq$2 was 47.0 % in men and 70.2 % in women, respectively, in subjects aged 60 years and older in Japanese population-based cohorts. Therefore, women are more likely than men to be affected by knee osteoarthritis and have difficulties in physical function of the lower extremities, leading to higher scores on the WOMAC function scale. Another reason for the sex differences may be the weaker muscle strength in women; muscle strength in men is higher than that in women in all decades of life [14], which may obscure the association in men, as muscle strength has been reported to be inversely associated with the WOMAC domains [15].

Functional declines in locomotive organs including physical ADLs usually progress slowly and gradually. As such, it may be difficult for people to recognize this decline in their daily life. Therefore, it is of particular importance to raise awareness of the growing risk caused by such disorders, and to take action to improve and maintain the health of the locomotive organs. The Japanese Orthopaedic Association proposed the concept of "locomotive syndrome" in 2007 for the promotion of preventive healthcare of the locomotive organs [16–18]. Locomotive syndrome refers to conditions under which the elderly have been receiving support or long-term care, or high-risk conditions under which they may soon require support or long-term care, that are caused by musculoskeletal disorders [16–18]. Population approaches, including promotion of the concept of locomotive syndrome to both younger and older generations, are important, in addition to high-risk approaches, including identifying those at risk for certified need of care and practicing intervention programs to reduce the risk of certified need of care.

Because the WOMAC function scale is a self-assessment questionnaire that is easy to conduct and evaluate, it can be used to screen elderly persons at high risk of certified need of care in the LTCI system. Multivariate Cox hazards regression analysis showed that a WOMAC function score of 5 in men and 4 in women best discriminated between the occurrence and the non-occurrence group of certified need of care in this study population. Elderly men with a WOMAC function score $\geq$5 had a 1.88-fold higher risk of occurrence of certified need of care compared with elderly men with a score <5. Elderly women with a WOMAC function score $\geq$4 had a 3.13-fold higher risk of occurrence of certified need of care compared with elderly women with a score <4. Elderly persons screened by these cut-off values should receive early intervention for the prevention of subsequent deterioration in ADLs that could lead to certified need of care. Further studies, along with the accumulation of epidemiologic evidence, are necessary to develop intervention programs that are safe and effective for elderly subjects who are at high risk of certified need of care.

There are some limitations in the present study. First, we could not obtain information on causes of certified need of care in the LTCI system. Therefore, we could not analyze the direct association of each causing condition with measured factors, and could not determine the risk factors for occurrence of certified need of care with respect to each causing condition. The Japanese government reported that the top five leading causes of certified need of care were cerebral stroke, dementia, asthenia, osteoarthritis, and fall-related fracture, comprising 71.6 % of all causes in 2010 [10]. Based on these data, most of the causes of incident certification in the present

study are inferred to be among the top five leading conditions. Additional studies are necessary to identify those direct associations. Second, participants at baseline in the present study were those who could walk to the survey site and could understand and sign an informed consent form. Since those who could not were not included in the analyses, the study participants do not truly represent the general population due to health bias, which should be taken into consideration when generalizing the results of the present study.

In conclusion, the present study determined association of physical ADLs with the occurrence of certified need of care in the LTCI system in elderly participants of Japanese population-based cohorts. The severity of physical dysfunction is a predictor of the occurrence of certified need of care. Further studies are necessary to develop intervention programs that are safe and effective for elderly individuals who are at high risk of certified need of care.

Acknowledgments This study was supported by Grants-in-Aid for Scientific Research (S19109007, B20390182, B23390172, B23390356, and B23390357) from the Japanese Ministry of Education, Culture, Sports, Science and Technology; H17-Men-eki-009, H18-Choujyu-037, H20-Choujyu-009, H21-Chouju-Wakate-011, H22-Chouju-Wakate-007, H23-Chouju-002, and H25-Chouju-007 from the Ministry of Health, Labour and Welfare; and Research Aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 and 2010-2).

Conflict of interest There are no conflicts of interest.

References

1. National Institute of Population and Society Research. Population projections for Japan (January 2012): 2011 to 2060. http://www.ipss.go.jp/site-ad/index_english/esuikei/gh2401e.asp.
2. Ministry of Health, Labour and Welfare. Long-term care, health and welfare services for the elderly. http://www.mhlw.go.jp/english/policy/care-welfare/care-welfare-elderly/index.html.
3. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol. 2010;39:988–95.
4. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Yamamoto S, Ishibashi H, Kawaguchi H, Nakamura K, Akune T. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab. 2009;27:620–8.
5. Shimada H, Lord SR, Yoshida H, Kim H, Suzuki T. Predictors of cessation of regular leisure-time physical activity in community-dwelling elderly people. Gerontology. 2007;53:293–7.
6. Barr S, Bellamy N, Buchanan WW, Chalmers A, Ford PM, Kean WF, Kraag GR, Gerecz-Simon E, Campbell J. A comparative study of signal versus aggregate methods of outcome measurement based on the WOMAC Osteoarthritis Index. Western Ontario and McMaster Universities Osteoarthritis Index. J Rheumatol. 1994;21:2106–12.
7. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988;15:1833–40.
8. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities osteoarthritis index (WOMAC). J Orthop Sci. 2003;8:288–93.
9. Chen W, Fukutomi E, Wada T, Ishimoto Y, Kimura Y, Kasahara Y, Sakamoto R, Okumiya K, Matsubayashi K. Comprehensive geriatric functional analysis of elderly populations in four categories of the long-term care insurance system in a rural, depopulated and aging town in Japan. Geriatr Gerontol Int. 2013;13:63–9.
10. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey. 2010. http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/4-2.html.
11. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med. 1995;332:556–61.
12. Vermeulen J, Neyens JC, van Rossum E, Spreeuwenberg MD, de Witte LP. Predicting ADL disability in community-dwelling elderly people using physical frailty indicators: a systematic review. BMC Geriatr. 2011;11:33.
13. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, Saika A, Suzuki T, Yoshida H, Ishibashi H, Yamamoto S, Nakamura K, Kawaguchi H, Yoshimura N. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthr Cartil. 2009;17:1137–43.
14. Sinaki M, Nwaogwugwu NC, Phillips BE, Mokri MP. Effect of gender, age, and anthropometry on axial and appendicular muscle strength. Am J Phys Med Rehabil. 2001;80:330–8.
15. Muraki S, Akune T, Oka H, En-yo Y, Yoshida M, Saika A, Suzuki T, Yoshida H, Ishibashi H, Tokimura F, Yamamoto S, Nakamura K, Kawaguchi H, Yoshimura N. Association of radiographic and symptomatic knee osteoarthritis with health-related quality of life in a population-based cohort study in Japan: the ROAD study. Osteoarthr Cartil. 2010;18:1227–34.
16. Nakamura K. A "super-aged" society and the "locomotive syndrome". J Orthop Sci. 2008;13:1–2.
17. Nakamura K. Locomotive syndrome: disability-free life expectancy and locomotive organ health in a "super-aged" society. J Orthop Sci. 2009;14:1–2.
18. Nakamura K. The concept and treatment of locomotive syndrome: its acceptance and spread in Japan. J Orthop Sci. 2011;16:489–91.

Geriatrics & Gerontology INTERNATIONAL Medline Indexed

Geriatr Gerontol Int 2014; **14:** *695–701*

ORIGINAL ARTICLE: SOCIAL RESEARCH, PLANNING AND PRACTICE

Incidence of certified need of care in the long-term care insurance system and its risk factors in the elderly of Japanese population-based cohorts: The ROAD study

Toru Akune,[1] Shigeyuki Muraki,[1] Hiroyuki Oka,[2] Sakae Tanaka,[3] Hiroshi Kawaguchi,[3] Fumiaki Tokimura,[4] Hideyo Yoshida,[5] Takao Suzuki,[6] Kozo Nakamura[7] and Noriko Yoshimura[2]

Departments of [1]Clinical Motor System Medicine and [2]Joint Disease Research, 22nd Century Medical & Research Center, Graduate School of Medicine, University of Tokyo, [3]Department of Sensory & Motor System Medicine, Graduate School of Medicine, University of Tokyo, [4]Department of Orthopaedic Surgery, Tokyo Metropolitan Geriatric Hospital, [5]Research Team for Promoting Independence of the Elderly, Tokyo Metropolitan Institute of Gerontology, Tokyo, [6]Research Institute, National Center for Geriatrics and Gerontology, Aichi, and [7]National Rehabilitation Center for Persons with Disabilities, Saitama, Japan

Aim: To examine the incidence of certified need of care in the national long-term care insurance (LTCI) system, and to determine its risk factors in the elderly of Japanese population-based cohorts of the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study.

Methods: Of the 3040 participants in the baseline examination of the ROAD study, we enrolled 1773 (699 men, 1074 women) aged 65 years or older who were not certified as in need of care level elderly at baseline. Participants were followed for incident certification of need of care in the LTCI system. Associated factors in the baseline examination with occurrence were determined by multivariate Cox proportional hazards regression analysis. Muscle dysfunction was defined in accordance with the European Working Group on Sarcopenia in Older People algorithm for screening sarcopenia.

Results: A total of 54 men and 115 women were certified as in need of care level elderly during the average 4.0-year follow up. The incidence was 2.0 and 2.5 per 100 person-years in men and women, respectively. Identified risk factors were region, age, body mass index <18.5 or ≥27.5 kg/m^2, grip strength, knee extension torque, usual gait speed, chair stand time and muscle dysfunction.

Conclusions: Both underweight and obesity, as well as low muscle strength and physical ability, are risk factors for certification of need of care. Considering muscle dysfunction is a risk factor for occurrence, screened individuals are recommended to receive early intervention programs regardless of muscle volume. **Geriatr Gerontol Int 2014; 14: 695–701.**

Keywords: activities of daily living, certification of need of care (*youkaigo-nintei*), disability, long-term care insurance system, prospective cohort study.

Introduction

Japan is a super-aged society experiencing an unprecedented aging of the population. The proportion of the population aged 65 years or older was 23% in 2010, and is expected to reach 30.1% in 2024 and 39% in 2051.[1] This leads to an increasing proportion of disabled elderly requiring support or long-term care, imposing enormous economic and social burdens on the country. The Japanese Government started the national long-term care insurance (LTCI) system in 2000 based on the Long-Term Care Insurance Act.[2] The aim was to certify need of care level elderly, and to provide suitable care services according to the level of care required (seven levels, including requiring support [levels 1 and 2] and requiring long-term care [levels 1–5]). The total number of certified in need of care level elderly was reported to be 5 million in 2011.[2]

Accepted for publication 12 August 2013.

Correspondence: Dr Toru Akune MD PhD, Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. Email: akune-ort@h.u-tokyo.ac.jp

Certification of need of care in the national LTCI system is an important outcome in Japan, not only because of its massive social and economic burdens, but also because it is urgently required to reduce its risk and decrease the number of disabled elderly requiring care in their activities of daily living (ADL). For establishment of an evidence-based prevention strategy, it is critically important to accumulate epidemiological evidence including the incidence of certified need of care and identification of risk factors. However, there have been no studies to clarify the incidence of certified need of care in the LTCI system or its risk factors using large-scale, population-based cohorts.

In 2005, we started a large-scale, population-based cohort study entitled the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study with a total of 3040 participants, which aims to elucidate the environmental and genetic backgrounds of musculoskeletal diseases.[3,4] The present study investigated the incidence of certified need of care in the national LTCI system, and determined its risk factors using a database from the ROAD study.

Methods

Participants

The present analysis was based on data collected from cohorts established in 2005 for the ROAD study. Details of the cohorts have been reported elsewhere.[3,4] Briefly, we created a baseline database from 2005–2007, which included clinical and genetic information on 3040 residents of Japan (1061 men, 1979 women). Participants were recruited from resident registration listings in three communities, namely, an urban region in Itabashi, Tokyo, and rural regions in Hidakagawa and Taiji, Wakayama. Participants in the urban region in Itabashi were recruited from those of a cohort study,[5] in which participants were randomly drawn from the register database of Itabashi ward residents, with a response rate of 75.6% in the group aged >60 years. Participants in the rural regions in Hidakagawa and Taiji were recruited from resident registration lists, with response rates of 68.4% and 29.3%, respectively, in the groups aged >60 years. Inclusion criteria were the ability to: (i) walk to the survey site; (ii) report data; and (iii) understand and sign an informed consent form. For the present study, we enrolled 1773 participants (699 men, 1074 women; mean age 75.4 years) aged 65 years or older who were not certified as need of care level elderly in the national LTCI system at baseline. All participants provided written informed consent, and the study was carried out with approval from the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology.

Baseline procedures

Participants completed an interviewer-administered questionnaire containing 400 items that included lifestyle information, such as smoking habits, alcohol consumption and physical activity. At baseline, anthropometric measurements, including height and weight, were taken, and body mass index (BMI; weight [kg]/height^2 [m^2]) was estimated based on the measured height and weight. Underweight was defined as BMI <18.5 and obesity as BMI ≥27.5, according to the 2004 consensus statement from the WHO regarding appropriate BMI for Asian populations.[6] Grip strength was measured on bilateral sides using a handgrip dynamometer (TOEI LIGHT, Saitama, Japan); the higher measurement was recorded. Isometric peak knee extension torque was measured at a knee flexion angle of 90° using a dynamometer (GT-30; OG GIKEN, Okayama, Japan) twice in participants from the urban regional cohort (Itabashi, Tokyo); the higher measurement was recorded. The time taken to walk 6 m at usual walking speed in a hallway was recorded, and usual gait speed was calculated. Skeletal muscle dysfunction was defined as usual gait speed ≤0.8 m/s or grip strength <30 kg in men and <20 kg in women, according to the algorithm for screening sarcopenia recommended by the European Working Group on Sarcopenia in Older People (EWGSOP).[7,8] The time taken for five consecutive chair rises without the use of hands was recorded in the rural regional cohorts (Hidakagawa and Taiji, Wakayama). Hands were folded in front of the chest with feet flat on the floor. Timing began with the command "Go", and ended when the buttocks contacted the chair on the fifth landing.

Certification of need of care in the LTCI system

The nationally uniform criteria for long-term care need certification was established objectively by the Japanese Government, and certification of need of care level elderly is determined based on evaluation results by the Certification Committee for Long-term Care Need in municipalities in accordance with basic guidelines formulated by the Government. The process of eligibility for certification of need of care in the LTCI system was described in detail by Chen *et al.*[9] An elderly person who requires help with ADL or the caregiver contacts the municipal Government to request official certification of care needs. After the application, a trained official visits the home to assess the current physical status of the elderly person, including the presence or absence of muscle weakness or joint contracture of limbs, and difficulties in sitting-up, standing-up, maintaining sitting or standing position, transferring from one place to another, standing on one leg, walking, bathing, dressing, and other ADL. Mental status, including dementia, is also assessed. These data are analyzed to calculate a

Table 1 Baseline characteristics of population at risk for certified need of care in the long-term care insurance system

	Entire cohort		Urban cohort		Rural cohort	
	Men	Women	Men	Women	Men	Women
No. participants	699	1,074	333	486	366	588
Age (years)	75.6 (5.1)	75.2 (5.3)	77.5 (3.7)	77.3 (3.8)	73.8 (5.5)†	73.5 (5.8)†
Height (cm)	160.9 (6.0)	147.9 (6.0)*	161.0 (5.8)	148.2 (5.4)*	160.8 (6.2)	147.7 (6.5)*
Weight (kg)	59.4 (9.1)	50.0 (8.3)*	59.4 (8.2)	49.8 (7.8)*	59.4 (9.9)	50.1 (8.8)*
BMI (kg/m^2)	22.9 (2.9)	22.8 (3.4)	22.9 (2.7)	22.7 (3.3)	22.9 (3.1)	22.9 (3.5)
BMI <18.5 (%)	6.2	8.0	6.1	7.9	6.3	8.0
BMI ≥27.5 (%)	5.7	9.3**	3.9	8.5**	7.4	9.9
Grip strength (kg)	30.4 (6.8)	19.4 (4.9)*	28.6 (6.1)	18.2 (4.1)*	31.9 (7.0)†	20.3 (5.2)*†
Knee extension torque (kgm)	–	–	79.6 (27.2)	54.8 (17.0)*	–	–
Usual gait speed (m/s)	1.17 (0.31)	1.10 (0.33)*	1.27 (0.24)	1.22 (0.24)*	1.08 (0.34)†	1.00 (0.36)*†
Chair stand time (s)	–	–	–	–	10.8 (3.7)	12.2 (5.4)*
Muscle dysfunction (%)§	48.7	56.0**	52.6	60.0**	45.2	52.6**‡
Smoking (%)	21.0	3.2**	19.2	3.0**	22.6	3.4**
Alcohol consumption (%)	61.2	23.0**	61.0	28.8**	61.3	18.4**‡

Except where indicated otherwise, values are mean (SD). *$P < 0.05$ versus men in the corresponding group of the same cohort by unpaired Student's *t*-test. **$P < 0.05$ versus men in the corresponding group of the same cohort by χ^2-test. †$P < 0.05$ versus urban cohort in the corresponding group of the same sex by unpaired Student's *t*-test. ‡$P < 0.05$ versus urban cohort in the corresponding group of the same sex by χ^2-test. §Muscle dysfunction was defined as usual gait speed ≤0.8 m/s or grip strength <30 kg in men and <20 kg in women. BMI, body mass index; LTCI, long-term care insurance system.

standardized score for determination of the level of care needs (certified support, levels 1–2; or long-term care, levels 1–5). In addition, the primary physician of the applicant assesses physical and mental status, including information on diseases causing ADL disability and the extent of disabilities caused by them. Finally, the Certification Committee for Long-term Care Need reviews the data and determines the certification and its level.

Follow up and definition of incident certified need of care

After the baseline ROAD survey, participants who were not certified as need of care level elderly at baseline were followed for incident certification of need of care in the LTCI system. Incident certified need of care was defined as the incident certified 7 level, including requiring support (levels 1–2) and requiring long-term care (levels 1–5). Information on the presence or absence of certification of need of care and its date of occurrence were collected by the resident registration listings in three communities every year up to 2010, and were used for analyses in the present study.

Statistical analysis

All statistical analyses were carried out using STATA statistical software (STATA, College Station, TX, USA). Differences in the values of the parameters between two groups were tested for significance using the non-paired Student's *t*-test and χ^2-test. Factors associated with occurrence of certified need of care were determined using Cox proportional hazards regression analysis; hazard ratios (HR) and 95% confidence intervals (CI) were determined after adjusting for region, age, sex, and BMI.

Results

Of the 1773 participants who were not certified as in need of care level elderly at baseline, information on certification of need of care could be obtained in 1760 (99.3%) during the average 4.0-year follow up. A total of 54 men and 115 women were certified as in need of care level elderly in the national LTCI system; whereas, 1591 remained uncertified during the follow-up period. A total of 126 participants died, and eight moved away.

Table 1 shows the baseline characteristics of the population at risk for occurrence of certified need of care in the LTCI system. Although BMI was not significantly different between men and women in the entire, urban or rural cohorts, prevalence of obesity (BMI ≥27.5) was significantly higher in women than in men in the entire and urban cohorts. The prevalence of underweight was higher in women than in men in the entire,

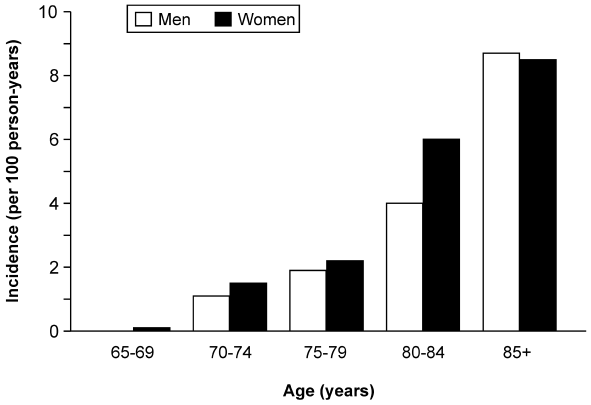


Figure 1 Incidence of certified need of care in the long-term care insurance system in men and women in each age stratum.

urban and rural cohorts; however, there was no significant difference. The prevalence of skeletal muscle dysfunction, determined by gait speed and grip strength, was significantly higher in women than in men in the entire, urban and rural cohorts.

Figure 1 shows sex- and age-distributions of the incidence of certified need of care in the LTCI system. Incidence was 2.3/100 person-years in the overall population of the entire cohort, and 2.0/100 person-years in men and 2.5/100 person-years in women. The incidence was very low in the age-stratum of 65–69 years, whereas, it tended to be markedly higher in the age-strata of 80 years and older in both sexes.

We then determined the risk factors for occurrence of certified need of care in the LTCI system. First, analysis was carried out using region, age, sex and BMI as explanatory variables in the Cox proportional hazards regression model (upper part of Table 2). Rural region and age were found to be risk factors for occurrence of certified need of care in the overall population. Sex and BMI were not significantly different. To further investigate the association between BMI and occurrence, we categorized BMI into three groups. Both underweight (BMI <18.5) and obesity (BMI ≥27.5) were found to be risk factors for occurrence of certified need of care, showing a U-shaped association. As for muscle strength and physical performance, handgrip strength, knee extension torque, usual gait speed, chair stand time and muscle dysfunction were found to be significantly associated with occurrence of certified need of care (lower part of Table 2). We carried out the same analyses in men and women separately (Table 2), and found results similar to those of the overall population.

Discussion

The present study investigated the incidence of certified need of care in the national LTCI system, and

Table 2 Hazard ratios and 95% confidence intervals for occurrence of certified need of care in the long-term care insurance system

	Overall population		Men		Women	
	Crude HR (95% CI)	Adjusted HR (95% CI)	Crude HR (95% CI)	Adjusted HR (95% CI)	Crude HR (95% CI)	Adjusted HR (95% CI)
Region (rural *vs* urban)	1.15 (0.83–1.59)	1.61 (1.17–2.24)[b]	1.13 (0.65–1.96)	1.64 (0.94–2.86)[g]	1.15 (0.77–1.72)	1.59 (1.07–2.38)[g]
Age (+1 year)	1.17 (1.13–1.20)	1.17 (1.14–1.21)[c]	1.19 (1.12–1.26)	1.19 (1.13–1.26)[h]	1.16 (1.12–1.20)	1.16 (1.12–1.21)[h]
Sex (women *vs* men)	1.25 (0.90–1.74)	1.24 (0.89–1.73)[d]	–	–	–	–
BMI (+1 kg/m^2)	0.98 (0.93–1.03)	1.01 (0.96–1.06)[e]	0.93 (0.84–1.02)	0.96 (0.88–1.06)[i]	1.00 (0.94–1.06)	1.02 (0.97–1.08)[i]
≥18.5 or <27.5	1 (Reference)	1 (Reference)	1 (Reference)	1 (Reference)	1 (Reference)	1 (Reference)
BMI <18.5	2.10 (1.31–3.38)	1.77 (1.10–2.84)[e]	2.43 (1.09–5.40)	1.69 (0.75–3.82)[i]	1.93 (1.07–3.48)	1.79 (0.99–3.22)[i]
BMI ≥27.5	1.82 (1.13–2.93)	2.12 (1.32–3.43)[e]	1.39 (0.50–3.87)	1.91 (0.68–5.38)[i]	1.92 (1.12–3.29)	2.18 (1.27–3.75)[i]
Grip strength (+1 kg)	0.93 (0.91–0.95)	0.94 (0.91–0.97)[f]	0.91 (0.87–0.95)	0.94 (0.89–0.99)[j]	0.89 (0.85–0.92)	0.94 (0.89–0.98)[j]
Knee extension torque (+1 kgm)	0.97 (0.96–0.99)	0.97 (0.96–0.99)[f]	0.97 (0.95–0.99)	0.97 (0.95–0.99)[j]	0.97 (0.95–0.99)	0.97 (0.95–1.00)[j]
Usual gait speed (+0.1 m/s)	0.80 (0.77–0.85)	0.84 (0.79–0.90)[f]	0.81 (0.74–0.88)	0.83 (0.74–0.92)[j]	0.80 (0.76–0.85)	0.85 (0.78–0.92)[j]
Chair stand time (+1 s)	1.09 (1.07–1.12)	1.06 (1.03–1.10)[f]	1.18 (1.10–1.27)	1.11 (1.03–1.21)[j]	1.09 (1.06–1.11)	1.06 (1.02–1.09)[j]
Muscle dysfunction (yes *vs* no)[a]	2.91 (2.02–4.19)	1.71 (1.16–2.52)[f]	2.60 (1.45–4.68)	1.68 (0.91–3.09)[j]	3.07 (1.92–4.92)	1.72 (1.04–2.85)[j]
Smoking (yes *vs* no)	0.98 (0.58–1.68)	1.39 (0.79–2.43)[f]	1.18 (0.62–2.26)	1.54 (0.79–3.01)[j]	0.95 (0.30–2.99)	1.09 (0.35–3.47)[j]
Alcohol consumption (yes *vs* no)	0.71 (0.50–0.99)	0.83 (0.58–1.21)[f]	0.78 (0.45–1.35)	0.93 (0.53–1.61)[j]	0.70 (0.42–1.16)	0.76 (0.46–1.27)[j]

[a]Muscle dysfunction was defined as usual gait speed ≤0.8 m/s or grip strength <30 kg in men and <20 kg in women. [b]Adjusted for age, sex and body mass index (BMI). [c]Adjusted for region, sex and BMI. [d]Adjusted for region, age and BMI. [e]Adjusted for region, age and sex. [f]Adjusted for region, age, sex and BMI. [g]Adjusted for age and BMI. [h]Adjusted for region and BMI. [i]Adjusted for region and age. [j]Adjusted for region, age and BMI. Urban region and men were used as references. CI, confidence interval; HR, hazard ratio.

determined its risk factors using Japanese population-based cohorts. Identified risk factors were region, age, underweight, obesity, handgrip strength, knee extension torque, usual gait speed, chair stand time and muscle dysfunction (determined by the EWGSOP algorithm for screening sarcopenia).

In the present study, we could not obtain information on causes of certified need of care in the LTCI system. Therefore, we could not analyze the direct association of each causing condition with such factors as anthropometric and physical performance measurements. The Government of Japan reported that the top five leading causes of certified need of care were cerebral stroke, dementia, asthenia as a result of older age, joint disease and fall-related fracture, comprising 71.6% of all causes in 2010.[10] Based on these data, most of the causes of incident certification in the present study are inferred to be among the top five leading conditions.

Both low and high BMI were found to be risk factors for occurrence of certified need of care, showing an overall U-shaped association. This U-shaped association is similar to that between BMI and risk of death.[11,12] The association between risk of death from cardiovascular disease and other causes, and BMI was reported to be U-shaped in East Asians,[11] whereas the risk of all-cause mortality versus BMI was also found to have a U-shaped association in Western European and North American populations.[12] High BMI is an established risk factor for chronic diseases, including hypertension, dyslipidemia and diabetes mellitus, which increase the risk of cerebral stroke.[13] High BMI is also a major risk factor for knee osteoarthritis,[14–17] which can cause ADL disability in the elderly.[18] In contrast, low BMI is an established risk factor for osteoporosis and related fracture.[19] It also might relate to asthenia, a condition of loss or lack of bodily strength as a result of chronic wasting disease. Underweight as a result of malnutrition or sarcopenia is suggested to be included in this category.

Other identified risk factors were handgrip strength, knee extension torque, usual gait speed, chair stand time and muscle dysfunction (determined by the EWGSOP algorithm for screening sarcopenia). Previous studies have reported that low muscle strength and physical performance were predictors of subsequent ADL disability in the elderly.[20–23] The results of the present study are consistent with these previous reports. As many of the performance tests used in the present study are easy to carry out and evaluate, they can be utilized for screening elderly persons at high risk of certified need of care in the LTCI system. Those who were classified as having muscle dysfunction in the present study were at high risk of sarcopenia as well as certified need of care, regardless of muscle volume. Therefore, elderly persons screened by the EWGSOP algorithm are recommended to receive early intervention programs for prevention of ADL disability and subsequent deterioration leading to certified need of care.

The Japanese Orthopedic Association proposed the concept of "locomotive syndrome" in 2007 for the promotion of preventive health care of locomotive organs.[24–26] Locomotive syndrome refers to conditions under which the elderly have been receiving support or long-term care, or high-risk conditions under which they might soon require support or long-term care, that are caused by musculoskeletal disorders.[24–26] Functional declines in locomotive organs, including muscle strength, walking speed and balancing ability, usually progress slowly and gradually. As such, it might be difficult for people to recognize this decline in their daily life. Therefore, it is of particular importance to raise awareness of the growing risk caused by these disorders, and to take action to improve and maintain the health of locomotive organs. Population approaches, including promotion of the concept of locomotive syndrome to both younger and older generations, are important, in addition to high-risk approaches, including identifying those at risk for certified need of care and practicing intervention programs to reduce the risk of certified need of care.

There were some limitations in the present study. As we could not obtain information on causing conditions, we could not determine the risk factors for occurrence of certified need of care with respect to each causing condition. Additional studies are necessary to identify those direct associations. In the present study, the rural region was at higher risk of incident certified need of care compared with the urban region. The reasons for this could include differences in available public and private transportation or delivery services regarding meals and commodities for the elderly. In addition to these, the threshold between certified and non-certified elderly might be different among municipalities, which could lead to regional differences. Although the Certification Committee for Long-term Care Need in each municipality determines certification in accordance with guidelines formulated by the Government, the Committee also has to consider assessment by the applicant's primary physician and objective evaluation results regarding physical and mental status, which could affect the threshold of certification. Another limitation was health bias. Participants at baseline in the present study were those who could walk to the survey site, and could understand and sign an informed consent form. As those who could not were not included in the analyses, the study participants do not truly represent the general population due to health bias. Therefore, incidence of certified need of care was most likely underestimated, which should be taken into consideration when generalizing the results of the present study.

In conclusion, the present study revealed the incidence of certified need of care in the national LTCI

system, and determined its risk factors using Japanese population-based cohorts. Both underweight and obesity were found to be risk factors for certified need of care, suggesting that maintenance of intermediate BMI is important for prevention. Low muscle strength and physical ability were also shown to be risk factors for certified need of care. Physical performance measures identified as predictors can be used as screening tools to identify high-risk individuals. Considering muscle dysfunction, screened by the EWGSOP algorithm, was a risk factor for occurrence, screened individuals are recommended to receive early intervention programs regardless of muscle volume. Further studies are necessary to develop intervention programs and to test their effectiveness, along with accumulation of epidemiological evidence, to prevent certified need of care and reduce the social and economic burdens associated with this condition.

Acknowledgments

This study was supported by Grants-in-Aid for Scientific Research (S19109007, B20390182, B23390172, B23390356, and B23390357) from the Japanese Ministry of Education, Culture, Sports, Science and Technology; H17-Men-eki-009, H18-Choujyu-037, H20-Choujyu-009, H21-Chouju-Wakate-011, H22-Chouju-Wakate-007, H23-Chouju-002, and H25-Chouju-007 from the Ministry of Health, Labour and Welfare; and Research Aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 and 2010-2).

Disclosure statement

The authors declare no conflict of interest.

References

1 National Institute of Population and Society Research. Population projections for Japan (January 2012): 2011 to 2060. [Cited 18 Jun 2013.] Available from URL: http://www.ipss.go.jp/site-ad/index_english/esuikei/gh2401e.asp
2 Ministry of Health, Labour and Welfare. Long-term care, health and welfare services for the elderly. [Cited 18 Jun 2013.] Available from URL: http://www.mhlw.go.jp/english/policy/care-welfare/care-welfare-elderly/index.html
3 Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on Osteoarthritis/Osteoporosis Against Disability study. *Int J Epidemiol* 2010; **39**: 988–995.
4 Yoshimura N, Muraki S, Oka H *et al*. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. *J Bone Miner Metab* 2009; **27**: 620–628.
5 Shimada H, Lord SR, Yoshida H *et al*. Predictors of cessation of regular leisure-time physical activity in community-dwelling elderly people. *Gerontology* 2007; **53**: 293–297.
6 WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004; **363**: 157–163.
7 Cruz-Jentoft AJ, Baeyens JP, Bauer JM *et al*. European Working Group on Sarcopenia in Older People. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on sarcopenia in older people. *Age Ageing* 2010; **39**: 412–423.
8 Lauretani F, Russo CR, Bandinelli S *et al*. Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. *J Appl Physiol* 2003; **95**: 1851–1860.
9 Chen W, Fukutomi E, Wada T *et al*. Comprehensive geriatric functional analysis of elderly populations in four categories of the long-term care insurance system in a rural, depopulated and aging town in Japan. *Geriatr Gerontol Int* 2013; **13**: 63–69.
10 Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2010. [Cited 18 Jun 2013.] Available from URL: http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/4-2.html
11 Zheng W, McLerran DF, Rolland B *et al*. Association between body-mass index and risk of death in more than 1 million Asians. *N Engl J Med* 2011; **364**: 719–729.
12 Prospective Studies Collaboration, Whitlock G, Lewington S, Sherliker P *et al*. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* 2009; **373**: 1083–1096.
13 Haslam DW, James WP. Obesity. *Lancet* 2005; **366**: 1197–1209.
14 Davis MA, Ettinger WH, Neuhaus JN, Cho SA, Hauck WW. The association of knee injury and obesity with unilateral and bilateral osteoarthritis of the knee. *Am J Epidemiol* 1989; **130**: 278–288.
15 Felson DT, Zhang Y, Hannan MT *et al*. Risk factors for incident radiographic knee osteoarthritis in the elderly: the Framingham Study. *Arthritis Rheum* 1997; **40**: 728–733.
16 Muraki S, Oka H, Akune T *et al*. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. *Osteoarthritis Cartilage* 2009; **17**: 1137–1143.
17 Muraki S, Akune T, Oka H *et al*. Incidence and risk factors for radiographic knee osteoarthritis and knee pain in Japanese men and women: a longitudinal population-based cohort study. *Arthritis Rheum* 2012; **64**: 1447–1456.
18 Sharma L, Kapoor D. Epidemiology of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM, eds. *Osteoarthritis: Diagnosis and Medical/Surgical Management*, 4th edn. Philadelphia, PA: Lippincott Williams & Wilkins, 2007; 3–26.
19 De Laet C, Kanis JA, Odén A *et al*. Body mass index as a predictor of fracture risk: a meta-analysis. *Osteoporos Int* 2005; **16**: 1330–1338.
20 Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med* 1995; **332**: 556–561.
21 Giampaoli S, Ferrucci L, Cecchi F *et al*. Hand-grip strength predicts incident disability in non-disabled older men. *Age Ageing* 1999; **28**: 283–288.
22 Onder G, Penninx BW, Ferrucci L, Fried LP, Guralnik JM, Pahor M. Measures of physical performance and risk for

progressive and catastrophic disability: results from the Women's Health and Aging Study. *J Gerontol A Biol Sci Med Sci* 2005; **60**: 74–79.
23 Vermeulen J, Neyens JC, van Rossum E, Spreeuwenberg MD, de Witte LP. Predicting ADL disability in community-dwelling elderly people using physical frailty indicators: a systematic review. *BMC Geriatr* 2011; **11**: 33.
24 Nakamura KA. "super-aged" society and the "locomotive syndrome". *J Orthop Sci* 2008; **13**: 1–2.
25 Nakamura K. Locomotive syndrome: disability-free life expectancy and locomotive organ health in a "super-aged" society. *J Orthop Sci* 2009; **14**: 1–2.
26 Nakamura K. The concept and treatment of locomotive syndrome: its acceptance and spread in Japan. *J Orthop Sci* 2011; **16**: 489–491.

Osteoporos Int (2014) 25:1089–1098
DOI 10.1007/s00198-013-2489-0

ORIGINAL ARTICLE

Prevalence and progression of radiographic ossification of the posterior longitudinal ligament and associated factors in the Japanese population: a 3-year follow-up of the ROAD study

N. Yoshimura · K. Nagata · S. Muraki · H. Oka · M. Yoshida · Y. Enyo · R. Kagotani · H. Hashizume · H. Yamada · Y. Ishimoto · M. Teraguchi · S. Tanaka · H. Kawaguchi · Y. Toyama · K. Nakamura · T. Akune

Received: 14 May 2013 / Accepted: 5 August 2013 / Published online: 22 August 2013

Abstract
Summary The prevalence of radiographic cervical ossification of the posterior longitudinal ligament (OPLL) in 1,562 Japanese from a population-based cohort was 1.9 %. The presence of OPLL showed a significant association with the femoral neck bone mineral density (BMD), presence of diffuse idiopathic skeletal hyperostosis (DISH) and plasma pentosidine levels. Only one new case of radiographic OPLL was detected, but OPLL progressed in all affected subjects.
Introduction The purpose of this study was to clarify the prevalence and progression of radiographic OPLL and the associated factors, using the population-based cohort Research on Osteoarthritis/osteoporosis Against Disability (ROAD).
Methods In the ROAD study, 1,690 participants underwent X-ray examination of the entire spine and both knees. Radiographic OPLL, lumbar spondylosis, knee osteoarthritis and DISH were diagnosed by a single, well-experienced orthopaedic surgeon. An interviewer-administered questionnaire and tests for anthropometric measurements were administered, and the BMDs of the lumbar spine and proximal femur were determined. A new OPLL case was considered if heterotopic ossification in the posterior longitudinal ligament was absent at baseline but present during follow-up. Progression was defined as an increase in the maximum length or width of the ossification at follow-up over that at baseline.
Results Radiographic OPLL was detected in 30 (17 men, 13 women) of 1,562 individuals who underwent X-ray examination of the cervical spine (prevalence=1.9 %). Its prevalence was significantly higher in men than in women (p =0.007), but no association with age was observed. In a logistic regression analysis, OPLL showed a significant association with the femoral neck BMD, presence of DISH and plasma pentosidine levels. Only one new case of radiographic OPLL was detected, but OPLL progressed in all affected subjects.
Conclusion This population-based study clarified the prevalence of radiographic OPLL in the Japanese population as well as its progression. OPLL showed significant association with plasma pentosidine levels, BMD and DISH.

Keywords Bone mineral density · Diffuse idiopathic skeletal hyperostosis · Ossification of posterior longitudinal ligament of cervical spine · Plasma pentosidine · Prevalence · Progression

Electronic supplementary material The online version of this article (doi:10.1007/s00198-013-2489-0) contains supplementary material, which is available to authorized users.

N. Yoshimura (✉) · H. Oka
Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
e-mail: YOSHIMURAN-ORT@h.u-tokyo.ac.jp

K. Nagata · M. Yoshida · Y. Enyo · R. Kagotani · H. Hashizume · H. Yamada · Y. Ishimoto · M. Teraguchi
Department of Orthopedic Surgery, Wakayama Medical University, 811-1 Kimidera, Wakayama, Wakayama 641-8509, Japan

S. Muraki · T. Akune
Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

S. Tanaka · H. Kawaguchi
Department of Orthopaedic Surgery, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

Y. Toyama
Department of Orthopaedic Surgery, School of Medicine, Keio University, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

K. Nakamura
National Rehabilitation Center for Persons with Disabilities, 1, Namiki 4-chome, Tokorozawa, Saitama 359-8555, Japan

Introduction

Ossification of the posterior longitudinal ligament of the spine (OPLL) is the pathological ectopic ossification of this ligament at the cervical and thoracic spine. It causes myeloradiculopathy as a result of chronic pressure on the spinal cord and nerve roots [1, 2]. Epidemiologic studies have shown a relatively high prevalence of OPLL among the Japanese, a slightly lower prevalence among East Asians and a substantially lower prevalence among whites [3, 4].

In terms of its characteristics, several epidemiological studies have reported that adult-onset obesity and diabetes mellitus (DM) are independent risk factors of OPLL [5, 6]. Further, OPLL often coincides with diffuse idiopathic skeletal hyperostosis (DISH), a systemic disorder of hyperossification. McAfee et al. [7] found that seven (50 %) of 14 patients with OPLL had DISH, and in a Japanese study, DISH was present in 27 (25 %) of 109 patients with OPLL [8].

Besides the coexistence of other disorders such as DM and DISH, little detailed information is available on the profile of OPLL in the general population. These data are important in order to characterise the disease burden. In addition, limited information is available regarding factors associated with OPLL, including biochemical markers of bone turnover, bone mineral density (BMD) values, lifestyle factors, or other coexisting disorders, such as dyslipidaemia, impairment of glucose tolerance, lumbar spondylosis (LS) and knee osteoarthritis (KOA).

Thus, the aims of the present study were to clarify the prevalence of OPLL in the Japanese population and to examine the association of OPLL with biological and environmental factors as well as coexisting disorders. For this, we used a questionnaire survey and the large, population-based cohort Research on Osteoarthritis/osteoporosis Against Disability (ROAD), which included lifestyle factors and nutrition, blood and urinary examinations, BMD measurements and X-ray examinations [9, 10].

Methods

Outline of the ROAD study

We conducted the present study using the cohorts established in 2005 for the ROAD study. The ROAD study is a nationwide, prospective study of OA comprising population-based cohorts from several communities in Japan. The details of the cohort profile have been reported elsewhere [9, 10]. Briefly, in 2005–2007, we created a baseline database that included clinical and genetic information for 3,040 residents of Japan (1,061 men, 1,979 women); the mean age (deviation [SD]) of the participants was 70.3 [11.0]years (71.0 [10.7]years for men and 69.9 [11.2]years for women). The subjects were recruited from resident registration listings in three communities with different characteristics: 1,350 subjects (465 men, 885 women) were from an urban region in Itabashi, Tokyo; 864 subjects (319 men, 545 women) were from a mountainous region in Hidakagawa, Wakayama and 826 subjects (277 men, 549 women) were from a coastal region in Taiji, Wakayama.

The participants completed an interviewer-administered questionnaire of 400 items that included lifestyle information such as occupation, smoking habits and alcohol consumption; family history; medical history; physical activity; reproductive variables and health-related quality of life. A questionnaire was prepared by modifying the one used in the Osteoporotic Fractures in Men Study [11], and some new items were added to the modified questionnaire. The participants were asked whether they took prescription medication daily or nearly every day (0 = no, 1 = yes). If participants did not know the reason for the prescribed medication, they were asked to bring their medications to the medical doctor (NY).

Anthropometric measurements included height (in centimetres), body weight (in kilograms), arm span (in centimetres), bilateral grip strength (in kilograms) and body mass index (BMI; in kilograms per square metre). Experienced orthopaedic surgeons collected medical information on systematic, local and mental status, including information on back, knee and hip pain; swelling and range of motion of the joints and patellar and Achilles tendon reflexes.

In 2008–2010, we attempted to locate and follow up all 3,040 subjects. They were invited for the second survey of the ROAD study, which included a 3-year follow-up of the same examinations as the baseline.

Subjects eligible for the present study

In the present study, we enrolled all 1,690 subjects (men, 596; women, 1,094) from mountainous and coastal areas who had enrolled in the ROAD study. In the ROAD study, X-ray examination of the cervical and thoracic spine had been performed only for these subjects and not for those from the urban region. Further, for all these 1,690 participants, the BMDs for the lumbar spine and the proximal femur had been measured using dual energy X-ray absorptiometry (Hologic Discovery; Hologic, Waltham, MA, USA) during the baseline examination. Additionally, blood and urinary examinations had also been performed for these subjects.

The study participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo (no. 1264 and no. 1326) and the University of Wakayama Medical University (no. 373).

Radiographic assessment

Plain radiographs were obtained for the cervical, thoracic and lumbar spine in the anteroposterior and lateral views and both knees in the anteroposterior view with weight-bearing and foot-map positioning.

Cervical OPLL was diagnosed using plain radiographs of the cervical spine in the lateral view. OPLL was indicated by the presence of heterotopic ossification in the posterior longitudinal ligament on a lateral cervical radiograph. Radiographic OPLL was diagnosed by a single, experienced orthopaedic surgeon (KN) who was blinded to participants' clinical status. OPLL was classified into the following types: continuous, segmental and mixed. In the original OPLL classification by Tsuyama [3], it was categorised into four modes, namely continuous, segmental, mixed and localised. However, here, because of the small number of subjects in the localised category, these subjects were included in the continuous category. If OPLL was observed, the maximum length (continuous and localised type, upper limit to lower limit; segmental and mixed types, upper limit to lower limit of the longest serial region) and width of ossification were measured using the imaging software OsiriX (http://www.osirix-viewer.com/).

In addition, using radiographs of spine and knees, we determined the grade of OA. The severity of radiographic OA was determined according to the Kellgren–Lawrence (KL) grading [12] as follows: KL0, normal; KL1, slight osteophytes; KL2, definite osteophytes; KL3, joint or intervertebral space narrowing with large osteophytes and KL4, bone sclerosis, joint or intervertebral space narrowing and large osteophytes. Radiographs for each site, i.e. the vertebrae and knees, were examined by a single, experienced orthopaedic surgeon (SM) who was blinded to participants' clinical status. In the present study, the subject's KL grade was considered the maximum grade diagnosed for at least one intervertebral level of the lumbar spine or at least one knee joint.

We also investigated the presence of DISH using whole-spine X-ray films. The criterion for the definite diagnosis of DISH was the presence of four or more vertebral bodies with contiguous ligamentous ossification and calcification, which is known as Resnick and Niwayama's criterion [13]. DISH was diagnosed by a single, experienced orthopaedic surgeon (RK) who was blinded to participants' clinical status.

Blood and urine examinations

Samples were collected from the end of October to the middle of January from both mountainous and coastal areas. All blood and urine samples were extracted between 0900 and 1500 hours. The blood samples were centrifuged, and the sera and urine samples were immediately placed on dry ice and transferred to a deep freezer within 24 h. The samples were stored at −80 °C until assayed.

The blood samples were used to measure haemoglobin A1c (HbA1c, Japan Diabetes Society), serum levels of total cholesterol, uric acid and creatinine levels. The analyses were performed at the same laboratory within 24 h of collection (Osaka Kessei Research Laboratories, Inc., Osaka, Japan).

Serum levels of intact parathyroid hormone (iPTH) were measured using an electrochemiluminescence immunoassay (Roche Diagnostics GmbH, Mannheim, Germany). As a marker of bone formation, serum levels of N-terminal propeptide of type I procollagen (PINP) were measured using a radioimmunoassay (Orion Diagnostics, Espoo, Finland). The urinary levels of β-isomerised C-terminal cross-linking telopeptide of type I collagen (β-CTX), a bone resorption marker, were determined using an enzyme-linked immunosorbent assay (Fujirebio, Inc., Tokyo, Japan). Urinary β-CTX values were standardised to urinary creatinine concentrations. Plasma pentosidine levels were detected using a competitive ELISA kit (FSK pentosidine ELISA kit; Fushimi Pharmaceutical, Kagawa, Japan) as previously described [14].

Three-year follow-up and definition of OPLL occurrence and progression

In 2008–2010, the 1,690 subjects were invited to enrol in the second survey of the ROAD study, a 3-year follow-up consisting of examinations identical to those conducted at baseline. Spine and knee radiographs were also obtained at follow-up. All cervical radiographs were read by the same orthopaedic surgeon who read them at the baseline (KN), and he was again blinded to participants' clinical status. He simultaneously compared the X-ray films at the baseline and 3-year follow-up and thereby diagnosed OPLL. A new OPLL case was diagnosed if heterotopic ossification in the posterior longitudinal ligament was absent on the lateral cervical radiograph obtained at baseline but present in that obtained during follow-up. OPLL progression was defined as an increase in the maximum length or width of the heterotopic ossification during follow-up compared to that at baseline.

Statistical analysis

All statistical analyses were performed using STATA statistical software (STATA Corp., College Station, TX, USA). Differences in proportions were compared using the chi-square test. Differences in continuous variables were tested for significance using analysis of variance for multiple groups or Scheffe's least significant difference test for pairs of groups. All p values and 95 % confidence intervals (CI) are two sided.

To test the association between OPLL and potential risk factors, we used logistic regression analysis with the presence or absence of OPLL (0 = absence, 1 = presence) as an objective variable and select potential explanatory variables, in addition to basic characteristics such as age (+1 year), gender (0 = men, 1 = women) and regional differences (0 = mountainous area, 1 = coastal area). The selected associated factors were those that showed a significant ($p < 0.05$) association with OPLL status in a simple linear analysis. To test the association between OPLL progression and associated factors, we used multivariate

regression analysis with the change rate (percent per year) of the maximum length or width as an objective variable and the explanatory variables used in the above-mentioned logistic regression analysis. The explanatory variables in the logistic regression analysis and multivariate regression analysis are described in the "Results" section.

Results

Prevalence of radiographic OPLL

The X-ray radiographs of 1,562 of the 1,690 subjects (92.4 %, 520 men, 1,038 women) showed all parts of the lateral cervical spine, from C1 to C7. Among these 1,562 individuals, 30 (17 men, 13 women) were diagnosed with radiographic OPLL; thus, the prevalence of OPLL was estimated at 1.9 % (men, 3.2 %; women, 1.3 %), and it was significantly higher in men than in women ($p = 0.007$).

Figure 1 shows the prevalence of OPLL classified by age and gender. The prevalence of OPLL was not associated with age in either men or women.

In the 30 subjects with radiographic OPLL, the OPLL was categorised into the continuous type in 13 subjects (six men and seven women, 43.3 %), the segmented type in eight (six men and two women, 26.7 %), the mixed type in seven (four men and three women, 23.3 %) and the localised type in two (one man and one woman, 6.7 %). The largest OPLL region was most commonly observed in C4 (ten individuals; 33.3 %; three men and seven women), followed by C5 (nine individuals; 33.0 %; eight men and one woman), C3 (seven individuals; 23.3 %; four men and three women), C6 (three individuals; 10.0 %; two men and one women) and C2 (one individual; 3.3 %; one woman). The largest OPLL region was not found in C1 or C7 in any subject.

The mean length and width (standard deviation, SD) of the largest region of ossification at the baseline were 27.6 (16.0) and 3.0 (1.5)mm, respectively. The values in men were 26.1 (14.5) and 2.9 (1.4)mm, and those in women were 29.6 (18.1) and 3.2 (1.5)mm, respectively; thus, no significant difference was observed between men and women in this regard.

Factors associated with OPLL

Table 1 shows the baseline characteristics of 1,562 participants with and without OPLL. Overall, subjects with OPLL tended to be taller and heavier than those without OPLL ($p < 0.05$). Further, compared to individuals without OPLL, those with OPLL had higher plasma pentosidine levels and higher BMD values for both the lumbar spine (L2–4) and femoral neck ($p < 0.05$).

Table 1 also shows the prevalence of LS, KOA and DISH on the basis of OPLL status. The prevalence of LS with ≥grade 2 KL and that of DISH was higher in the group with OPLL than in the one without OPLL ($p < 0.05$), although no significant association was observed between the prevalence of KOA and the presence of OPLL.

Logistic regression analysis was performed with the OPLL status as the objective variable (0 = absence, 1 = presence). As explanatory variables, the analysis involved select associated factors that showed a significant ($p < 0.05$) association with OPLL status in the simple linear analysis, namely, height (in centimetres), weight (in kilograms), values of plasma pentosidine (+1 μg/mL), BMD of the femoral neck (+1 SD), presence of LS based on KL grade (0 = KL grade 0 or 1, 1 = KL grade≥2) and DISH (0 = absent, 1 = present), after adjustments were made for age (years) and gender (0 = men, 1 = women). As seen from Table 2, plasma pentosidine levels, BMD of the femoral neck and the presence of DISH were found to be significant associated factors for the presence of OPLL (Table 2). Further, when BMD of the lumbar spine (L2–4) was used instead of that of the femoral neck, this factor was also found to be significantly associated with OPLL (+1 SD; odds ratio (OR), 1.52; 95 % CI, 1.05–2.20; $p = 0.026$), but the

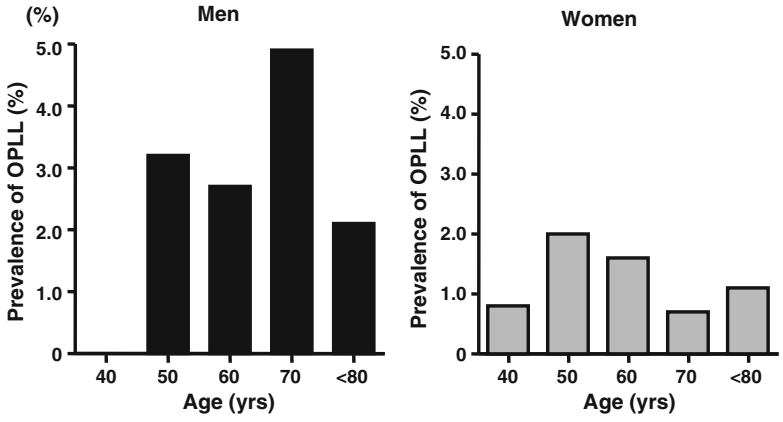


Fig. 1 Prevalence of OPLL classified by age and gender

Table 1 Baseline characteristics of participants classified by the presence or absence of OPLL

	Total (N=1,562)			Men (N=524)			Women (N=1,038)		
	OPLL (−) N=1,532	OPLL (+) N=30	p	OPLL (−) N=507	OPLL (+) N=17	p	OPLL (−) N=1,025	N=1,025 N=13	p
Age distribution (prevalence, %)									
30 years and younger	43	0 (0.0)		12	0 (0.0)		31	0 (0.0)	
40–49 years	141	1 (0.7)		39	0 (0.0)		102	1 (1.0)	
50–59 years	291	7 (2.4)	0.729	92	3 (3.2)	0.604	199	4 (2.0)	0.787
60–69 years	449	9 (2.0)		142	4 (2.7)		307	5 (1.6)	
70–79 years	468	11 (2.3)		175	9 (4.9)		293	2 (0.7)	
80 years and older	140	2 (1.4)		47	1 (2.1)		93	1 (1.1)	
Age (years), mean (SD)	62.9 (12.1)	67.0 (9.3)	0.3495	66.0 (11.7)	70.7 (8.0)	0.0990	64.4 (12.2)	62.2 (9.0)	0.5069
Height (cm), mean (SD)	154.9 (9.1)	159.1 (7.5)	0.0132*	163.3 (7.0)	163.9(5.4)	0.7414	150.8 (6.9)	152.8 (4.6)	0.2945
Weight (kg), mean (SD)	55.0 (10.3)	60.3 (10.1)	0.0053**	61.6 (10.5)	62.7 (8.2)	0.6759	51.7 (8.5)	57.1 (11.7)	0.0219*
BMI (kg/m^2), mean (SD)	22.8 (3.2)	23.8 (3.4)	0.1135	23.0 (3.1)	23.3 (2.1)	0.7434	22.7 (3.3)	24.4 (4.6)	0.0671
Residing in the coastal area (%)	49.4	53.3	0.671	46.4	58.8	0.311	50.9	46.2	0.732
Current smoking habit (regularly, ≥1/month) (%)	12.9	23.3	0.095	31.1	41.2	0.377	3.8	0.0	0.472
Current alcohol consumption (regularly, ≥1/month) (%)	39.1	43.3	0.637	66.1	64.7	0.907	25.8	15.4	0.395
Total cholesterol (mg/dL), mean (SD)	208.8 (34.5)	209.6 (36.2)	0.8954	198.6 (34.1)	204.4 (33.5)	0.4874	213.8 (33.6)	216.4 (39.8)	0.7840
Uric acid (mg/dL), mean (SD)	4.84 (1.30)	5.24 (1.21)	0.0943	5.71 (1.26)	5.71 (1.03)	0.9867	4.42 (1.09)	4.65 (1.21)	0.4528
HbA1c (Japan Diabetes Society) (%), mean (SD)	5.17 (0.70)	5.38 (0.79)	0.1124	5.20 (0.79)	5.44 (0.95)	0.2162	5.16 (0.64)	5.29 (0.56)	0.4595
Serum levels of iPTH (pg/mL), mean (SD)	41.2 (34.4)	41.2 (14.2)	0.9952	42.6 (54.4)	41.1 (13.9)	0.9083	40.5 (17.4)	41.3 (15.1)	0.8748
Serum levels of PINP (μg/L), mean (SD)	57.9 (27.0)	52.6 (29.9)	0.2915	47.5 (22.0)	42.6 (14.9)	0.3619	63.1 (27.8)	65.8 (39.2)	0.7301
Urinary levels of β-CTX (μg/mmol Cr), mean (SD)	187.2 (121.3)	150.4 (79.1)	0.0985	128.4 (78.7)	119.8 (58.3)	0.6529	216.2 (128.0)	190.5 (86.8)	0.4693
Plasma levels of pentosidine (μg/mL), mean (SD)	0.058 (0.037)	0.085 (0.140)	0.0005***	0.061 (0.048)	0.102 (0.184)	0.0042**	0.057 (0.030)	0.062 (0.037)	0.5012
BMD of the lumbar spine L2-4 (g/cm^2), mean (SD)	0.925 (0.205)	1.084 (0.205)	<0.0001***	1.038 (0.203)	1.176 (0.176)	0.0058**	0.868 (0.181)	0.965 (0.181)	0.0575
BMD of the femoral neck (g/cm^2), mean (SD)	0.667 (0.137)	0.747 (0.134)	0.0016**	0.739 (0.132)	0.797 (0.110)	0.0727	0.631 (0.124)	0.681 (0.139)	0.1558
Presence of LS (KL grade≥2) (%)	61.8	83.3	0.016*	76.1	100.0	0.022*	54.7	61.5	0.624
Presence of KOA (KL grade≥2) (%)	49.5	56.7	0.440	41.4	41.2	0.986	53.6	76.9	0.093
Presence of DISH (%)	9.4	33.3	<0.001***	0.7	52.9	0.002**	3.8	7.7	0.469

OPLL ossification of posterior longitudinal ligament, *SD* standard deviation, *BMI* body mass index, *HbA1c* haemoglobin A1c, *iPTH* intact parathyroid hormone, *PINP* N-terminal propeptide of type I procollagen, *β-CTX* β-isomerised C-terminal cross-linking telopeptide of type I collagen, *BMD* bone mineral density, *LS* lumbar spondylosis, *KOA* knee osteoarthritis, *KL grade* Kellgren–Lawrence grade, *DISH* diffuse idiopathic skeletal hyperostosis, *OPLL(−)* absence of OPLL, *OPLL(+)* presence of OPLL

*p<0.05; **p<0.01; ***p<0.001

Table 2 Odds ratios of potential factors associated with the presence of OPLL vs. the absence of OPLL

Explanatory variables	Reference	OR	95 % CI	*p*
Age (years)	+1 year	1.03	0.98–1.07	0.269
Gender	0 = men, 1 = women	1.30	0.39–4.34	0.666
Height (cm)	+1 cm	1.04	0.96–1.12	0.352
Weight (kg)	+1 kg	1.00	0.96–1.05	0.909
Pentosidine (μg/mL)	+0.01 μg/mL	1.05	1.00–1.09	0.038*
BMD (femoral neck) (g/cm2)	+1 SD	1.55	1.04–2.33	0.033*
Presence of LS (KL grade≥2)	0 = no, 1 = yes	1.94	0.67–5.61	0.219
Presence of DISH	0 = no, 1 = yes	2.78	1.11–6.92	0.029*

Logistic regression analysis was performed using the status of OPLL as the objective variable (0 = absence, 1 = presence), and the abovementioned factors were correspondingly adjusted

OPLL ossification of posterior longitudinal ligament, *BMD* bone mineral density, *LS* lumbar spondylosis, *KL grade* Kellgren–Lawrence grade, *DISH* diffuse idiopathic skeletal hyperostosis, *SD* standard deviation, *OR* odds ratios, *95 % CI* 95 % confidence interval

*p <0.05; **p <0.01; ***p <0.001

association of plasma pentosidine levels and DISH weakened (plasma pentosidine +0.01 μg/mL, 1.04, 0.997–1.087, p =0.069; presence of DISH 2.37, 0.94–6.00, p =0.069).

New occurrence or progression of OPLL

During the three study years, 1,380 individuals (88.3 %; 466 men, 914 women) among the 1,562 subjects at baseline returned for follow-up, and their radiographs were available for observation. Among the 30 individuals with radiographic cervical OPLL at baseline, 25 (83.3 %; 14 men and 11 women) participated in the second survey.

The remaining 1,355 individuals who did not have cervical OPLL at baseline and who participated in the initial and second surveys were regarded as members of the population at risk for the occurrence of OPLL. Among them, only one woman was diagnosed with newly developed radiographic OPLL (incidence 2.46/10,000 per year).

At follow-up, the mean length (in millimetres, SD) and width (in millimetres, SD) of the maximum region of ossification among the 25 individuals with OPLL was 28.7 (16.1) and 3.5 (1.5)mm, respectively. Since the mean values of length and width of the maximum region of ossification of these 25 subjects were 27.0 (16.2) and 3.0 (1.5)mm at the baseline, respectively, both the length and width of the maximum region of ossification increased, although a significant difference was not observed.

To clarify the risk factors associated with this increase in the length and width of the ossification, we performed multivariate regression analysis using the rate of change in these parameters as objective variables and the explanatory variables as those used in the logistic regression analysis, namely height (in centimetres), weight (in kilograms), plasma pentosidine levels (+1 μg/mL), BMD of the femoral neck (+1 SD), presence of LS based on the KL grade (0 = KL grade 0 or 1, 1 = KL grade≥2) and DISH (0 = absence, 1 = presence). Adjustments for age (years) and gender (0 = men, 1 = women) were made. However, none of the abovementioned variables was found to be significantly associated with the rate of changes in the length or width.

Discussion

In the present population-based study, we clarified the prevalence of radiographic OPLL in the general Japanese population, and we found that it is significantly associated with high plasma pentosidine levels, high BMD and the presence of DISH. The 3-year follow-up study also showed that new cases were very rare, and the length and width of the maximum region of ossification among the subjects with OPLL tended to increase.

The prevalence of OPLL in Japan has been reported to be 1.9 to 4.3 % among individuals aged 30 years and older [1, 15–17]. In other Asian countries, such as in Korea [18, 19] and Taiwan [20], a similar prevalence was reported, but it was lower in Western countries [21], suggesting that ethnic and/or genetic factor(s) could be associated with the onset of OPLL. In the present study, the prevalence of OPLL was found to be 1.9 %. This is consistent with the value found in previous reports. However, it is difficult to clearly distinguish localised-type OPLL from osteophytic changes, and we included two individuals with localised-type OPLL in the OPLL group. Thus, we may have overestimated the presence of radiographic OPLL. If we exclude individuals with localised-type OPLL from the OPLL group, the prevalence of the OPLL in the present study is 1.8 %.

With regard to the gender difference in OPLL prevalence, the prevalence was previously reported to be three times higher in men than in women [22]. We found that men are 2.5 times more likely to have OPLL than women (men 3.2 %, women

1.3 %), which is consistent with results reported previously among Japanese subjects. In contrast, symptomatic OPLL was reported to be usually observed in the sixth decade of life [22], although we were unable to find a significant association between age and the presence of OPLL. This might be explained by the fact that previous studies on the characteristics of OPLL were performed on the subjects with symptomatic OPLL, i.e. they had been clinically diagnosed with OPLL, while our subjects had radiographic OPLL that had not been clinically diagnosed. If the OPLL in our subjects progresses in the future, the peak age at which the symptoms could be expressed may be their 60s.

With regard to the comorbidities of OPLL, several reports have indicated that obesity and DM might be associated with OPLL [5, 6]. In the present study, the values of BMI tended to be higher in the group with OPLL than in that without OPLL, although this difference was not significant. A similar pattern was found in the values of HbA1c, and this finding could be explained by previous findings that the extent of ossification was significantly associated with the fasting serum insulin level but not with the fasting glucose level or the HbA1c level [23]. However, in the ROAD study, since all subjects could not be requested to fast, we could not confirm the association between fasting serum insulin levels and OPLL.

With regard to the association between biochemical markers of bone turnover and OPLL, Matsui et al. showed that the levels of the bone markers serum procollagen type I carboxyl-terminal peptide and intact osteocalcin were higher in patients with OPLL than in normal subjects [24]. This suggested that OPLL was associated with biochemical markers of bone turnover. In the present study, to evaluate the role of bone metabolism in OPLL, we compared the serum levels of iPTH and PINP as bone formation markers and the urinary levels of β-CTX between the groups with and without OPLL. However, we could not find significant differences between the groups.

Instead, the plasma pentosidine levels of the OPLL group were found to be significantly higher than those of the group without OPLL. This tendency remained after potential associated factors were adjusted for. Pentosidine is an advanced glycation end product, products generated by the sequential nonenzymatic glycosylation of protein amino groups [25] that accumulate in various tissues including kidney and coronary arteries, resulting in the development of diabetic vascular complications [26]. The concentrations of pentosidine in cortical and trabecular bone are reported to be adversely associated with bone strength [27–29]. Yamamoto et al. [30] found that serum pentosidine levels were positively associated with the presence of vertebral fractures in postmenopausal women with type 2 diabetes. Renal insufficiency was reported to be a dominant determinant of serum pentosidine levels [31] because of which serum pentosidine levels are increased in patients with chronic renal failure [32, 33]. However, no report has shown the association between pentosidine levels and the presence of OPLL. On the basis of the abovementioned reports, we performed multivariate logistic regression analysis using the same explanatory factors we had used in the analysis shown in Table 2, along with the estimated glomerular filtration rate. We found that the plasma pentosidine levels were still significantly related to the presence of OPLL (OR, 1.05; 95 % CI, 1.00–1.09; $p = 0.042$). We speculate that the levels of pentosidine might be associated with ectopic ossification, such as vascular calcification in patients with renal dysfunction, or the presence of OPLL, directly or indirectly, although the currently available information is inadequate to prove this hypothesis. One reason for the inadequacy of the information obtained in this study could be that we did not evaluate genetic factors in the present study. Further investigations are needed to clarify whether the observed relationship between pentosidine levels and OPLL remains after analysis of other possible confounders, including genetic factors.

In addition to the biochemical markers, high BMDs have been observed in patients with OPLL [24, 34, 35]. However, Morio et al. reported that the BMD was lower in patients with advanced OPLL [36], suggesting that the disuse atrophy may result during advanced-stage OPLL. Our results also showed that subjects with OPLL had higher BMDs. However, the subjects in the present study all had radiographically determined OPLL but few clinical symptoms, so their condition may not have been in the advanced stage. Therefore, based solely on the results of the present study, we were unable to discuss the association between BMD and advanced-stage OPLL.

Several reports have shown that the coexistence of OPLL and DISH is quite common [4, 7, 8]. The pathogenesis of DISH and OPLL has been speculated to be similar, although the details remain unclear. For example, Havelka et al. analysed intron 6 (−4) polymorphisms in the COL 11 A2 gene in Czech patients with DISH and Japanese patients with OPLL, but they found no agreement between the data of subjects with DISH and OPLL [37]. Additional studies with a broader spectrum of genotyping and a larger cohort of patients may clarify the presence or absence of genetic relations between DISH and OPLL.

Few studies have been reported regarding the incidence of OPLL in the general population because OPLL is relatively rare and based on ethnicity, as noted. Using data collected in a pilot study in the corporation of 360 Japanese hospitals [3], Tsuyama described the incidence of OPLL and found that 2,142 patients were treated in these hospitals and the estimated incidence of OPLL was 19 patients per million persons of the total population [3]. In the present study, only one new case of OPLL was detected, so we could not accurately estimate the incidence of OPLL and compare our results to those of previous reports. In order to confirm the incidence of OPLL, we need to follow this cohort for a longer time.

Several studies have investigated the course of OPLL. Chiba et al. use computer-assisted measurement to examine OPLL

progression, and they found that the rate of OPLL progression was 56.5 % over 2 years, and this rate was most common in younger patients with continuous- and mixed-type OPLL [38]. Murakami et al. followed the case of a 67-year-old man who had had cervical OPLL for more than 26 years, and they found that the rate of OPLL progression was 2.2, 8.8 and 2.0 mm/year from 1–4, 4–8 and 8–10 years after the first visit, respectively [39]. However, to our knowledge, no study has reported the progression of radiographically defined OPLL in the general population. In the present study, we found that both the length and width of the maximum region of ossification increased during the 3 years of the study, although it was not a significant change. A previous report [39] found no evidence of OPLL progression after 10 years. We must carefully examine whether or not radiographically defined OPLL progresses to clinical OPLL.

This study has several limitations. First, although the ROAD study includes a large number of participants, these participants may not truly be representative of the general population. To address this, we compared the anthropometric measurements and the frequencies of smoking and alcohol consumption between the study participants and the general Japanese population. No significant differences were found, with the exception that male ROAD study participants aged 70–74 years were significantly smaller in terms of body structure than men from the overall Japanese population ($p < 0.05$) [10]. This difference should be considered when evaluating potential risk factors for men aged 70–74 years; factors such as body build, particularly weight, are known to be associated with metabolic risk factors and KOA. Therefore, our results may have underestimated the prevalence of these conditions. Second, the total number of subjects with confirmed OPLL was very small, which might make the results somewhat less credible. In the present study, we used logistic regression analysis to adjust for gender differences. When the total number of the objective variable, namely OPLL cases, is small, using the multivariate model to adjust for gender differences may be more useful than using a gender-specific analysis. This is because the total number of cases in a gender-specific analysis will be even smaller, which reduces the statistical power. Although the significant associations between OPLL and the plasma levels of pentosidine and between OPLL and DISH were observed only in men in the simple comparative analysis, the pentosidine levels and DISH remained significant factors associated with the presence of OPLL even in the logistic regression analysis with adjustments for gender. We interpreted this result to mean that the female sex might dilute the strength of the association between OPLL and DISH, but the tendency in both genders remained significant.

To clarify the effect of sex differences in the interaction among OPLL, pentosidine levels and DISH, the logistic regression analysis was performed in men and women separately (Supplementary Table 1). In this logistic regression analysis, the presence of OPLL was significantly associated with the pentosidine levels and femoral neck BMD in men, but the association of OPLL with the presence of DISH was diluted to a marginal association ($p = 0.080$). Further, since all male patients with DISH had radiographic LS, we could not evaluate the association between OPLL and LS. In women, the associations among OPLL, pentosidine levels and DISH were not significant. Although these results may indicate that the significant associated factors were observed only in men, they may even be skewed by the small number of female cases. Under these circumstances, it is difficult to distinguish which model should be used, i.e. logistic regression analysis or the multivariate model. It may be necessary to first include an adequate number of OPLL cases before this can be decided. To compensate for these limitations, we decided to include the urban cohort of the ROAD study in the OPLL survey. Thus, more participants will be included in the third ROAD survey planned from 2012 to 2013, and further detailed investigation regarding the risk factors for the presence, occurrence or exacerbation of OPLL may be possible.

In summary, the present study clarified that the prevalence of radiographic cervical OPLL in 1,562 individuals was 1.9 %, which was significantly higher in men than in women ($p = 0.007$), but no association with age was observed. In logistic regression analysis, OPLL showed a significant association with the femoral neck BMD, presence of DISH and plasma pentosidine levels. Only one new case of radiographic OPLL was detected, but OPLL progressed in all affected subjects.

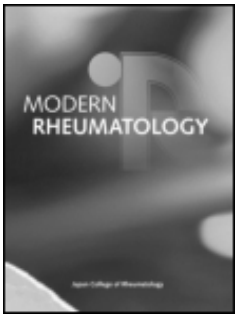
Acknowledgments This work was supported by Grants-in-Aid for Scientific Research B23390172 and B20390182 to NY, C20591737 to TA and C20591774 to SM; grants for Young Scientists A18689031 to HO; Collaborating Research with NSF 08033011-00262 (Director, NY) from the Ministry of Education, Culture, Sports, Science and Technology and H17-Men-eki-009 (Director, KN), H18-Choujyu-037 (Director, TN), H20-Choujyu-009 (Director, NY), H23-Choujyu-002 (Director, TA), H23-Nanchi-Ippan-032 (Director, YT) and H25-Choujyu-007 (Director, NY) from the Ministry of Health, Labour and Welfare in Japan. This study was also supported by grants from the Japan Osteoporosis Society (NY, SM, HO and TA) and research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 and 2010-2, Director, HK). The authors wish to thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka from the Gobo Public Health Centre; Dr. Naoki Hirabayashi of the Kawakami Clinic, Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Kumiko Shinou, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, Ms. Ikuyo Ueyama, Mrs. Michiko Mori, Mrs. Hisayo Sugimoto and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Centre and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mr. Shoichi Shimoichi, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. We also thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai and Mrs. Saeko Sahara for their assistance with data reduction and administration.

Conflicts of interest None.

References

1. Sakou T, Matsunaga S, Koga H (2000) Recent progress in the study of pathogenesis of ossification of the posterior longitudinal ligament. J Orthop Sci 5:310–315
2. Schmidt MH, Quinones-Hinojosa A, Rosenberg WS (2002) Cervical myelopathy associated with degenerative spine disease and ossification of the posterior longitudinal ligament. Semin Neurol 22:143–148
3. Tsuyama N (1984) Ossification of the posterior longitudinal ligament of the spine. Clin Orthop Relat Res 184:71–84
4. Inamasu J, Guiot BH, Sachs DC (2006) Ossification of the posterior longitudinal ligament: an update on its biology, epidemiology, and natural history. Neurosurgery 58:1027–1039
5. Kobashi G, Washio M, Okamoto K, Sasaki S, Yokoyama T, Miyake Y, Sakamoto N, Ohta K, Inaba Y, Tanaka H, Japan Collaborative Epidemiological Study Group for Evaluation of Ossification of the Posterior Longitudinal Ligament of the Spine Risk (2004) High body mass index after age 20 and diabetes mellitus are independent risk factors for ossification of the posterior longitudinal ligament of the spine in Japanese subjects: a case–control study in multiple hospitals. Spine 29:1006–1010
6. Shingyouchi Y, Nagahama A, Niida M (1996) Ligamentous ossification of the cervical spine in the late middle-aged Japanese men. Its relation to body mass index and glucose metabolism. Spine 21:2474–2478
7. McAfee PC, Regan JJ, Bohlman HH (1987) Cervical cord compression from ossification of the posterior longitudinal ligament in non-orientals. J Bone Joint Surg Br 69:569–575
8. Ehara S, Shimamura T, Nakamura R, Yamazaki K (1998) Paravertebral ligamentous ossification: DISH, OPLL and OLF. Eur J Radiol 27:196–205
9. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T (2010) Cohort profile: Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study. Int J Epidemiol 39:988–995
10. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Yamamoto S, Ishibashi H, Kawaguchi H, Nakamura K, Akune T (2009) Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 27:620–628. doi:10.1007/s00774-009-0080-8
11. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, Lewis C, Cawthon PM, Marcus R, Marshall LM, McGowan J, Phipps K, Sherman S, Stefanick ML, Stone K (2005) Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study: a large observational study of the determinants of fracture in older men. Contemp Clin Trials 26:569–585
12. Kellgren JH, Lawrence LS (1957) Radiological assessment of osteo-arthrosis. Ann Rheum Dis 16:494–502
13. Resnick D, Niwayama G (1976) Radiographic and pathologic features of spinal involvement in diffuse idiopathic skeletal hyperostosis (DISH). Radiology 119:559–568
14. Sanaka T, Funaki T, Tanaka T, Hoshi S, Niwayama J, Taitoh T, Nishimura H, Higuchi C (2002) Plasma pentosidine levels measured by a newly developed method using ELISA in patients with chronic renal failure. Nephron 91:64–73
15. Matsunaga S, Yamaguchi M, Hayashi K, Sakou T (1999) Genetic analysis of ossification of the posterior longitudinal ligament. Spine (Phila Pa 1976) 24:937–938
16. Okamoto K, Kobashi G, Washio M, Sasaki S, Yokoyama T, Miyake Y, Sakamoto N, Ohta K, Inaba Y, Tanaka H, Japan Collaborative Epidemiological Study Group for Evaluation of Ossification of the Posterior Longitudinal Ligament of the Spine (OPLL) Risk (2004) Dietary habits and risk of ossification of the posterior longitudinal ligaments of the spine (OPLL); findings from a case–control study in Japan. J Bone Miner Metab 22:612–617
17. Washio M, Kobashi G, Okamoto K, Sasaki S, Yokoyama T, Miyake Y, Sakamoto N, Ohta K, Inaba Y, Tanaka H, Japan Collaborative Epidemiological Study Group for Evaluation of Ossification of the Posterior Longitudinal Ligament of the Spine Risk (2004) Sleeping habit and other life styles in the prime of life and risk for ossification of the posterior longitudinal ligament of the spine (OPLL): a case–control study in Japan. J Epidemiol 14:168–173
18. Jin BH, Kim YS (1991) Ossification of spinal ligaments. J Korean Neurosurg Soc 20:875–884
19. Kim TJ, Bae KW, Uhm WS, Kim TH, Joo KB, Jun JB (2008) Prevalence of ossification of the posterior longitudinal ligament of the cervical spine. Joint Bone Spine 75:471–474, d
20. Wang PN, Chen SS, Liu HC, Fuh JL, Kuo BI, Wang SJ (1999) Ossification of the posterior longitudinal ligament of the spine. A case–control risk factor study. Spine (Phila Pa 1976) 24:142–145
21. Resnick D (1994) Diagnosis of bone and joint disorders. Saunders, London, pp 1496–1507
22. Otsuka K, Terayama K, Yanagihara M (1986) An epidemiological survey on ossification of ligaments in the cervical and thoracic spine in individuals over 50 years of age. J Jpn Orthop Assoc 60:1087–1098
23. Akune T, Ogata N, Seichi A, Ohnishi I, Nakamura K, Kawaguchi H (2001) Insulin secretory response is positively associated with the extent of ossification of the posterior longitudinal ligament of the spine. J Bone Joint Surg Am 183A:1537–1544
24. Matsui H, Yudoh K, Tsuji H (1996) Significance of serum levels of type I procollagen peptide and intact osteocalcin and bone mineral density in patients with ossification of the posterior longitudinal ligaments. Calcif Tissue Int 59:397–400
25. Brownlee M (1995) Advanced protein glycosylation in diabetes and aging. Annu Rev Med 46:223–234
26. Brownlee M, Cerami A, Vlassara H (1988) Advanced glycosylation end products in tissue and the biochemical basis of diabetic complications. N Engl J Med 318:1315–1321
27. Hernandez CJ, Tang SY, Baumbach BM, Hwu PB, Sakkee AN, van der Ham F, DeGroot J, Bank RA, Keaveny TM (2005) Trabecular microfracture and the influence of pyridinium and non-enzymatic glycation-mediated collagen crosslinks. Bone 37:825–832
28. Wang X, Shen X, Li X, Agrawal CM (2002) Age-related changes in the collagen network and toughness of bone. Bone 31:1–7
29. Viguet-Carrin S, Roux JP, Arlot ME, Merabet Z, Leeming DJ, Byrjalsen I, Delmas PD, Bouxsein ML (2006) Contribution of the advanced glycation end product pentosidine and of maturation of type I collagen to compressive biomechanical properties of human lumbar vertebrae. Bone 39:1073–1079
30. Yamamoto M, Yamaguchi T, Yamauchi M, Yano S, Sugimoto T (2008) Serum pentosidine levels are positively associated with the presence of vertebral fractures in postmenopausal women with type 2 diabetes. J Clin Endocrinol Metab 93:1013–1019
31. Hricik DE, Schulak JA, Sell DR, Fogarty JF, Monnier VM (1993) Effects of kidney or kidney–pancreas transplantation on plasma pentosidine. Kidney Int 43:398–403
32. Sugiyama S, Miyata T, Ueda Y, Tanaka H, Maeda K, Kawashima S, Ypersele V, de Strihou C, Kurokawa K (1998) Plasma levels of pentosidine in diabetic patients: an advanced glycation end product. J Am Soc Nephrol 9:1681–1688
33. Miyata T, Ueda Y, Shinzato T, Iida Y, Tanaka S, Kurokawa K, van Ypersele de Strihou C, Maeda K (1996) Accumulation of albumin-linked and free-form pentosidine in the circulation of uremic patients with end-stage renal failure: renal implications in the pathophysiology of pentosidine. J Am Soc Nephrol 7:1198–1206
34. Hirai N, Ikata T, Murase M, Morita T, Katoh S (1995) Bone mineral density of the lumbar spine in patients with ossification of the posterior longitudinal ligament of the cervical spine. J Spinal Disord 8:337–341
35. Yamauchi T, Taketomi E, Matsunaga S, Sakou T (1999) Bone mineral density in patients with ossification of the posterior

longitudinal ligament in the cervical spine. J Bone Miner Metab 17:296–300
36. Morio Y, Yamamoto K, Kishimoto H, Hagino H, Kuranobu K, Kagawa T (1993) Bone mineral density of the radius in patients with ossification of the cervical posterior longitudinal ligament. A longitudinal study. Spine 18:2513–2516
37. Havelka S, Vesela M, Pavelkova A, Ruzickova S, Koga H, Maeda S, Inoue I, Halman L (2001) Are DISH and OPLL genetically related? Ann Rheum Dis 60:902–903
38. Chiba K, Yamamoto I, Hirabayashi H, Iwasaki M, Goto H, Yonenobu K, Toyama Y (2005) Multicenter study investigating the postoperative progression of ossification of the posterior longitudinal ligament in the cervical spine: a new computer-assisted measurement. J Neurosurg Spine 3:17–23
39. Murakami M, Seichi A, Chikuda H, Takeshita K, Nakamura K, Kimura A (2010) Long-term follow-up of the progression of ossification of the posterior longitudinal ligament. Case report. J Neurosurg Spine 12:577–579



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

Development and evaluation of a video exercise program for locomotive syndrome in the elderly

Hiroshi Hashizume, Noriko Yoshimura, Keiji Nagata, Nobuyuki Miyazaki, Yuyu Ishimoto, Ryoko Nishiyama, Hiroyuki Oka, Hiroshi Yamada & Munehito Yoshida

To cite this article: Hiroshi Hashizume, Noriko Yoshimura, Keiji Nagata, Nobuyuki Miyazaki, Yuyu Ishimoto, Ryoko Nishiyama, Hiroyuki Oka, Hiroshi Yamada & Munehito Yoshida (2014) Development and evaluation of a video exercise program for locomotive syndrome in the elderly, Modern Rheumatology, 24:2, 250-257, DOI: 10.3109/14397595.2013.854063

To link to this article: http://dx.doi.org/10.3109/14397595.2013.854063

Published online: 05 Mar 2014.

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://informahealthcare.com/mor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2014; 24(2): 250–257

DOI: 10.3109/14397595.2013.854063

informa
healthcare

ORIGINAL ARTICLE

Development and evaluation of a video exercise program for locomotive syndrome in the elderly

Hiroshi Hashizume[1], Noriko Yoshimura[2], Keiji Nagata[1], Nobuyuki Miyazaki[1], Yuyu Ishimoto[1], Ryoko Nishiyama[1], Hiroyuki Oka[2], Hiroshi Yamada[1], and Munehito Yoshida[1]

[1]*Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama City, Wakayama, Japan, and* [2]*Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan*

Abstract

Objectives. To develop and evaluate an exercise program that the elderly could sustainably perform in the community or at home to recover from locomotive syndrome.
Methods. We produced 2 types of teaching media, video and pamphlet, describing 10 physical and mobility training exercises. The pilot study examined changes in pulse rate, percutaneous oxygen saturation (SpO_2), and the Borg scale rating of perceived exertion in 20 elderly volunteers. Separately, 120 elderly subjects were recruited and divided into 3 groups according to the teaching medium (video, group V; pamphlet, group P; none, group C). Before and 3 months after the intervention, visual analog scale (VAS) scores of low back and knee pain, single-leg standing time, 6-m walking time, Roland–Morris Disability Questionnaire, Oswestry Disability Index, Short Form-8, and 25-question Geriatric Locomotive Function Scale were evaluated.
Results. Pulse-rate changes before and after exercise did not exceed 20 %, and SpO_2 changes were within 4 points in all cases. The Borg scale ranged between 11 and 14. The intergroup comparison revealed the advantage of the video program in improving the VAS of low back pain, left-leg standing time, and 6-m walking time.
Conclusion. A video exercise program can potentially aid recovery from locomotive syndrome in the elderly.

Keywords

DVD video, Elderly, Exercise program, Locomotive syndrome, Low back pain

History

Received 14 November 2012
Accepted 6 March 2013
Published online 30 March 2013

Introduction

Population aging is occurring in countries worldwide, but it is most advanced in highly developed countries. The population of Japan in particular is aging very rapidly. According to an estimate by the Ministry of Internal Affairs and Communications released on October 1, 2007, 29,005,000 people (22.7 % of the total population) in Japan were 65 years old or older [1]. The United Nations has therefore labeled Japan a "superaged" society. Aging is associated with an increased risk of problems related to physical mobility. By 2006, the number of elderly persons in Japan needing nursing care had increased to 4,300,000; orthopedic problems are unquestionably one of the main reasons for this need [2]. According to the Comprehensive Survey of Living Conditions conducted in 2007 by the Japan Ministry of Health, Labour and Welfare, the most frequent symptom in both men and women 65 years old or older was low back pain [3]. Joint disease is also a major cause of the need for long-term care; in a large-scale population-based cohort study, the number of patients in Japan with knee osteoarthritis (KOA) was estimated to be approximately 25 million [4].

To increase society's awareness of this problem, the Japanese Orthopaedic Association (JOA) has proposed the concept of *locomotive syndrome*, a condition in which the elderly receive care services, or are at high risk of soon requiring care services, because of difficulty with physical mobility [2, 5]. The earliest possible intervention is required to prevent the need for long-term care among individuals with locomotive syndrome.

Many reports indicated that exercise is effective for most disorders affecting mobility, including low back pain and knee disease [6, 7]. However, few reports have investigated the efficacy of exercise for the elderly, especially from the point of view of preventing locomotive syndrome. Because the elderly often have multiple diseases that affect their mobility, studies to develop tools for the early detection of locomotive syndrome are ongoing. Investigations of specific methods for teaching exercise to the elderly are also underway.

The purposes of this study are to (1) develop an exercise program that the elderly could sustainably perform in the community or at home, (2) investigate which medium of exercise instruction (video or pamphlet) is superior, and (3) determine the parameter that is most useful for evaluating the effectiveness of this intervention for locomotive syndrome.

Methods

For this study, "elderly" was defined according to the Japan Ministry of Health, Labour and Welfare's definition of 65 years old or older. The study was conducted after approval from the ethics review board and consent from the participants were obtained.

Correspondence to: Hiroshi Hashizume, Department of Orthopaedic Surgery, Wakayama Medical University, 811-1 Kimiidera, Wakayama City, Wakayama 649-6122, Japan. Tel: +073-441-0645. Fax: +073-448-3008. E-mail: hashizum@wakayama-med.ac.jp

Development of the exercise program for the elderly

Exercise program

Ten types of exercises that an elderly person could perform without excess load on the musculoskeletal system or cardiopulmonary function were selected from widely known physical and mobility training exercises. The exercises were combined into a single exercise program comprising 2 distinct parts:

(A) Mobility training for fall prevention (Fig. 1a)

1. Stepping (40 s)
2. Single-leg standing with eyes open (10 s/cycle × 4)
3. Squatting (10 s/cycle × 10)
4. Quadriceps femoris training (5 s/bilateral side × 5)

(B) Muscle training to prevent/improve low back pain (Fig. 1b)

1. Muscle training of the hips and pelvis (10 s/cycle × 8)
2. Exercises for back flexibility (10 s/cycle × 5)
3. Abdominal muscle training (10 s/cycle × 5)
4. Back muscle training (5 s/cycle × 3)
5. Stretching of the lumbar spine (30 s/cycle × 2)
6. Rounding the back like a cat (10 s)

Teaching materials

Two types of teaching media were produced (Fig. 2). One was a DVD video of exercise demonstration by 2 instructors with background music and commentary. The other was a pamphlet including the same content as the DVD video but consisting of photographs and descriptions. In the video, the rhythm was regulated so that 4 beats of the background music lasted for 5 s.

Estimation of the physical load from the DVD video exercise

Twenty volunteers (age range 65–88 years) who used a day hospital service were asked to perform the video exercises for 15 min. Changes in pulse rate and percutaneous oxygen saturation (SpO_2) before (within 5 min of starting) and after (within 60 s of finishing) the exercise were measured. The Borg scale rating of perceived exertion (RPE, range 6–20) was also recorded to measure the subjective intensity level of the physical activity [8].

Evaluation of the efficacy of the exercise program

Participants and sampling in groups

One hundred twenty elderly residents (age range 65–85 years, mean 72.2 years) of Hashimoto City were recruited by open

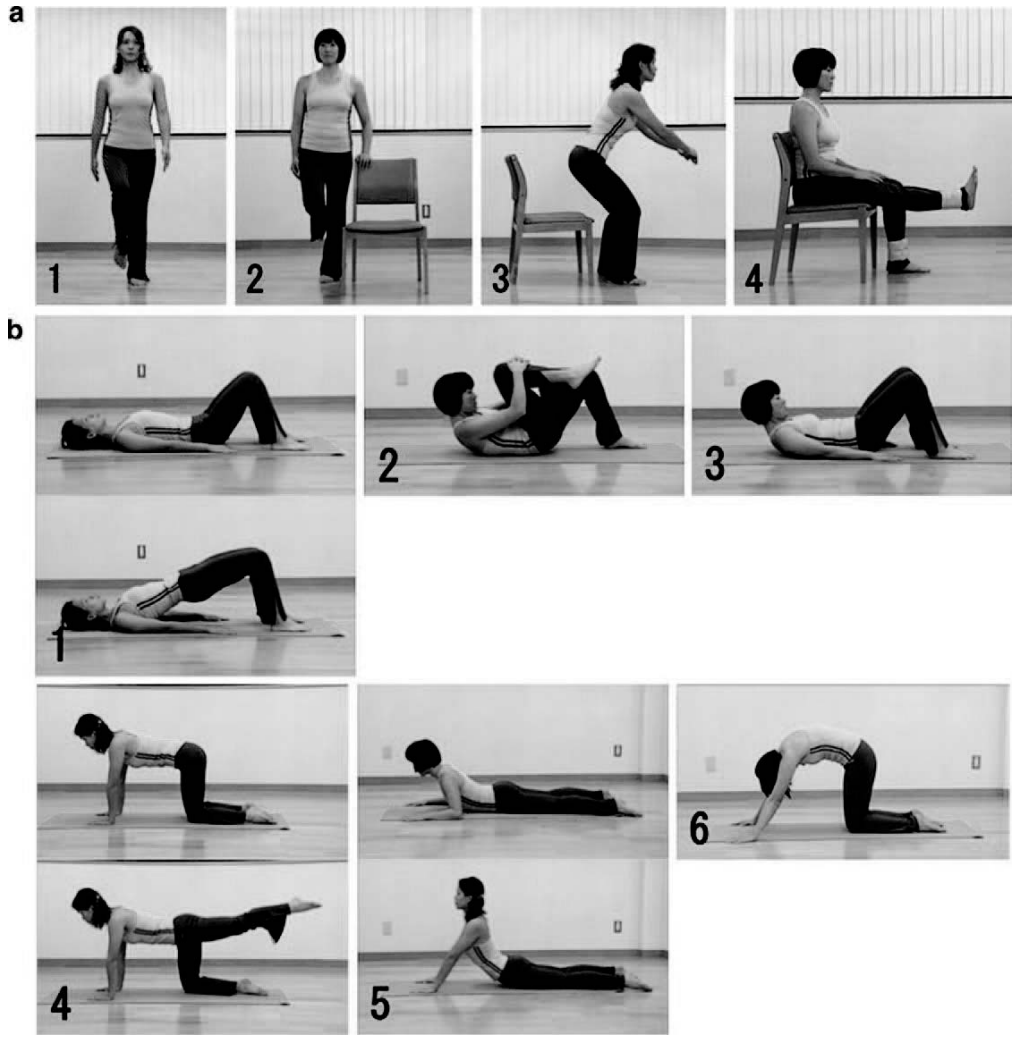


Fig. 1 Exercise program for the elderly. (a) Mobility training for fall prevention: (1) stepping, (2) single-leg standing with eyes open, (3) squatting, and (4) quadriceps femoris training. (b) Muscle training to prevent/improve low back pain: (1) muscle training of the hips and pelvis, (2) exercises for back flexibility, (3) abdominal muscle training, (4) back muscle training, (5) stretching of the lumbar spine, and (6) rounding the back like a cat

Fig. 2 Two types of teaching media produced for the participants. (a) DVD jacket cover. (b) Example of the instructions in the pamphlet: (1) stepping (*top*), (2) single-leg standing with eyes open (*middle*), and (3) squatting (*bottom*)

invitation. Hashimoto City is located between the mountains and a dormitory town of Osaka. In 2008, elderly people constituted 22.3 % of the approximately 69,000-resident population.

The participants were divided randomly into the following 3 groups by a member of the city staff: (1) group V participants performed the exercises while watching the video; (2) group P participants performed the exercises while reading the pamphlet; and (3) group C participants did not perform the exercises. Participants were allowed to change groups when requested because of a family relationship or friendship. Finally, there were 43, 41, and 36 participants in groups V, P, and C, respectively. Participants in groups V and P were instructed to perform all exercises twice a day. The period of intervention was set at 3 months.

Radiographic assessment

All participants underwent radiographic examination at the start of the intervention to assess degenerative changes in their knee joints or lumbar spine. KOA and lumbar spondylosis (LS) were defined as grade ≥2 on the Kellgren–Lawrence scale [9]; i.e., radiographic findings of definite osteophytes and definite narrowing of the joint space or intervertebral space were defined as grade 2. Osteoporosis (OP) of the lumbar spine was defined as sparse or absent longitudinal trabeculae in the vertebral body in accordance with the criteria proposed by the Japanese Society for Bone and Mineral Research [10]. Vertebral fracture was assessed by a quantitative method using lateral radiographs of the lumbar spine (L1–L5), according to the Japanese Society of Bone and Mineral Research criteria [10]. Wedge appearance was defined as a site at which the anterior height of the vertebra was ≤75 % of the posterior height. Biconcave appearance occurred if the height of the central part of the vertebra was ≤80 % of that of the anterior or posterior parts of the vertebra. Crush appearance was indicated if the heights of the anterior, central, and posterior parts of an axial vertebra were all reduced to ≤80 % of the normal values.

Clinical assessment

Clinical assessments of the participants were performed at the start of the intervention. Anthropometric measurements included height, weight, and body mass index [BMI; weight (kg)/height (m^2)]. To evaluate physical performance, the single-leg standing time for each leg was measured using a stopwatch (upper limit, 60 s). Six-meter walking times with normal steps and quick steps were also measured using a stopwatch. These measurements were performed by members of the local government staff who were blinded to the intervention groups. At the same time, the participants completed several types of self-report questionnaires. Presence of pain was assessed by a questionnaire asking the participants if they had experienced low back pain or knee pain lasting more than 24 h within the previous month. The participants were also asked to rate the intensity of their current pain pertaining

to the lower back or knee joints by using a visual analog scale (VAS, range 0–100). For the assessment of functional disability, participants completed the Oswestry Disability Index (ODI, Japanese version; range 0–100) [11]; for the assessment of low back pain, the Roland–Morris Disability Questionnaire (RDQ, Japanese version; range 0–24) [11–13] was used. The participants completed the Short Form-8 health survey (SF-8, Japanese version) [14] for assessment of health-related quality of life. The physical component summary (PCS) and mental component summary (MCS) scores of this instrument were calculated using the norm-based scoring system [14]. The 25-question Geriatric Locomotive Function Scale (GLFS-25, Japanese version) was used to detect locomotive syndrome [15]. The GLFS-25 is a self-administered, relatively comprehensive measure consisting of 25 items. These 25 items are graded on a 5-point scale, ranging from no impairment (0 points) to severe impairment (4 points), and the scores are then added together to produce a total (range 0–100). We set the cutoff score for identifying locomotive syndrome at 16, according to the currently accepted criteria [15].

The same clinical assessments, excluding anthropometric measurements or the presence of pain, were performed at the end (after 12 weeks) of the intervention. In addition, the participants in groups V and P recorded their daily level of exercise during the intervention. When a subject performed all of the exercises once, a score of 2 was given. One point was given if some of the exercises were performed, and a score of 0 was given if the subject did not perform any exercises. Because the participants in the 2 groups were instructed to perform all of the exercises twice daily, the possible range of daily points was 0–4. The exercise achievement summary scale was calculated using the following formula: (sum of the daily points/4 $\times$ number of days) $\times$ 100.

Statistical analyses

JMP9 (SAS Institute Inc., Cary, NC, USA) and IBM SPSS Statistics 18 (International Business Machines Corp., Armonk, NY, USA) statistical software were used for the statistical analyses in this study. To compare the baseline data among the 3 groups, a chi-square test for independence was used for proportional variables (sex and prevalence). One-factor analysis of variance (ANOVA) was performed for parametric variables (age, VAS, single-leg standing time, and 6-m walking time). A Kruskal–Wallis rank test was performed for nonparametric data (RDQ, ODI, PCS, MCS, and GLFS-25). For intragroup comparisons between baseline and 3 months after the intervention, a paired *t* test was used for parametric data (age, VAS, single-leg standing time, and 6-m walking time) and a Wilcoxon signed-rank test was used for nonparametric data (RDQ, ODI, PCS, MCS, and GLFS-25). A Mann–Whitney *U* test was applied to compare the exercise achievement summary scale scores between groups V and P. The intervention effect (the change of each parameter) was also compared. In advance of the comparison, split-plot ANOVA was performed to assess inter- and intragroup differences by using repeated values of VAS; single-leg standing time; 6-m walking time; and PCS, MCS, RDQ, ODI, and GLFS-25 scores as objective factors and group as an explanatory factor. Mauchly's sphericity test was used to validate the equality of the variances for repeated measures of the 3 groups. Subsequently, Scheffé's *F* post hoc pairwise multiple-comparison test was performed to assess the significance of the mean differences between the groups. In cases in which the baseline parameters significantly differed among the groups, analysis of covariance (ANCOVA) was applied in which the baseline value was added as a covariate. A value of $p < 0.05$ was considered significant.

Results

Physical load of the DVD video exercise

All 20 elderly participants with locomotive disability and/or concomitant internal disease completed the 15-min DVD video exercise program. Pulse-rate changes before and after exercise did not exceed 20 %, and SpO_2 changes were within 4 points in all cases. The RPE scores were 11 (light), 12, 13 (somewhat hard), and 14 for 5, 12, 2, and 1 participants, respectively. No participants had a score of 15 (hard) or higher for the intensity of the exercises (Table 1).

Efficacy of the exercise program

Table 2 presents the age, anthropometric measurements, prevalence of bone and joint diseases, and SF-8 summary scores (i.e.,

Table 1. Physical load of the DVD video exercise in 20 volunteers

			Concomitant chronic diseases		Pulse rate (bpm)		SpO_2 (%)		
No.	Age (years)	Sex	Musculoskeletal disease	Internal disease	Before EX	After EX	Before EX	After EX	RPE
1	78	Female	LBP	HT, bronchial asthma	74	78	95	95	12
2	71	Female	LBP	DM	88	84	93	92	12
3	69	Female	LBP	HT	92	91	96	96	13
4	73	Female	LBP		78	76	95	93	11
5	66	Female		Hyperlipemia	88	84	92	92	12
6	74	Female	KOA	HT	84	90	96	94	12
7	66	Male		HT, cerebral infarction	89	92	95	95	11
8	88	Male	LBP	HT, arrhythmia	81	83	94	92	12
9	65	Male	LSS		76	88	95	92	12
10	80	Male		LBP, KOA arrhythmia	78	92	95	91	14
11	83	Female	LBP	Cerebral infarction	77	79	97	97	11
12	82	Male	LBP	HT	80	84	97	96	12
13	80	Male	LBP	Parkinson's disease	80	88	94	96	12
14	69	Male	LSS	HT	85	87	96	96	12
15	73	Male	LBP	DM	72	68	97	95	11
16	78	Female	LBP	HT, arrhythmia	85	89	93	93	12
17	77	Male	LBP	Vertigo	63	72	97	97	11
18	79	Male		HT, bronchitis, angina pectoris	69	72	95	97	12
19	75	Male	LBP	Bronchial asthma	82	96	90	90	13
20	70	Male	LBP, HOA		86	86	97	95	12

Perceived exertion ratings of 12–14 suggest that physical activity was performed at a moderate level of intensity

LBP low back pain, *LSS* lumbar spinal stenosis, *KOA* knee osteoarthritis, *HOA* hip osteoarthritis, *HT* hypertension, *DM* diabetes mellitus, *bpm* beats per minute, *EX* exercise, *SpO_2* percutaneous oxygen saturation, *RPE* Borg rating of perceived exertion scale (range 6–20); RPE 11, "light" intensity; RPE 13, "somewhat hard" intensity; RPE 15, "hard" intensity

Table 2. Characteristics of 120 elderly participants

	Men	Women
Number of subjects	32	88
Age (years)	74.1 ± 5.4	71.6 ± 4.3
Height (cm)	164.5 ± 4.7	151.4 ± 4.9
Body weight (kg)	67.2 ± 9.1	53.1 ± 7.7
Body mass index (kg/m^2)	24.8 ± 2.6	23.1 ± 3.0
Prevalence		
Musculoskeletal pain		
Low back pain	26/31 (83.9 %)	58/86 (67.4 %)
Knee pain	11/31 (35.5 %)	32/86 (37.2 %)
Radiographic findings		
LS	28/31 (90.3 %)	71/86 (82.6 %)
KOA	16/32 (48.4 %)	67/86 (77.9 %)
OP of the lumbar spine	1/31 (3.2 %)	38/86 (44.2 %)
VF of the lumbar spine	1/31 (3.2 %)	11/86 (12.8 %)
Locomotive syndrome	8/31 (25.8 %)	36/88 (40.9 %)
Short Form-8 summary scores		
PCS	46.1 (40.9–49.3)	45.3 (41.2–49.6)
MCS	52.1 (49.6–55.7)	53.0 (48.3–55.3)

PCS and MCS presented as median (25–75 %). Locomotive syndrome was indicated by the 25-question Geriatric Locomotive Function Scale when the total score was ≥16 points

LS lumbar spondylosis, *KOA* knee osteoarthritis, *OP* osteoporosis, *VF* vertebral fracture, *PCS* physical component scale, *MCS* mental component scale

Table 3. Characteristics of the participants in the 3 intervention groups

	Group V	Group P	Group C
Number of subjects	43	41	36
Age (years)	72.9 ± 5.1	70.9 ± 3.9	73.1 ± 4.9
Gender (male:female)	18:25*	7:34	7:29
Body mass index (kg/m^2)	23.5 ± 3.0	23.5 ± 2.6	23.5 ± 3.3
Prevalence			
Musculoskeletal pain			
Low back pain	32/42 (76.2 %)	30/40 (75.0 %)	22/35 (62.9 %)
Knee pain	15/42 (35.7 %)	15/41 (36.6 %)	12/36 (33.3 %)
Radiographic findings			
LS	36/40 (90.0 %)	32/41 (78.0 %)	31/36 (86.1 %)
KOA	25/40 (62.5 %)	30/41 (73.2 %)	27/36 (75.0 %)
OP of the lumbar spine	10/40 (25.0 %)	18/41 (43.9 %)	11/36 (30.6 %)
Locomotive syndrome	17/42 (40.5 %)	15/41 (36.6 %)	12/36 (33.3 %)

Locomotive syndrome was indicated by the 25-question Geriatric Locomotive Function Scale when the total score was ≥16 points

Group V video exercise group, *Group P* pamphlet exercise group, *Group C* control group, *LS* lumbar spondylosis, *KOA* knee osteoarthritis, *OP* osteoporosis

*$p < 0.05$

PCS and MCS) in men and women. These data are used to verify the characteristics of the participants in the "Discussion."

The participants' characteristics by group are presented in Table 3. The mean age and BMI of the participants did not differ significantly among the 3 groups. There was a bias in the sex distribution among the 3 groups ($p = 0.0453$). At the start of the intervention, the prevalence of low back and knee pain did not differ significantly among the groups. No significant differences were observed among the groups regarding the prevalence of LS, KOA, and OP, although there was a tendency of a relatively higher prevalence of OP in group P. In total, 17 of 42 participants in group V, 15 of 41 participants in group P, and 12 of 36 participants in group C were diagnosed with locomotive syndrome. The prevalence of locomotive syndrome was not significantly different among the groups.

The follow-up rates (proportion of the participants who completed the 3-month intervention) were 88.4, 90.2, and 83.3 % for groups V, P, and C, respectively. One woman in group P withdrew from the study because she sustained a vertebral body fracture during her daily activities. Furthermore, 1 woman in group C withdrew because she was awaiting surgery for cervical spondylotic myelopathy. Another 13 participants withdrew from the study for personal reasons. No participants in group V or group P withdrew because of the difficulty of the exercise itself. The exercise achievement summary scale score in group V (median = 74.1 %, 25–75 percentile = 58.0–91.7 %) was significantly higher ($p = 0.0015$) than that in group P (median = 53.2 %, 25–75 percentile = 35.9–73.8 %). There was no significant difference in the exercise achievement summary scale score between men (median = 69.6 %, 25–75 percentile = 40.4–91.1 %) and women (median = 64.7 %, 25–75 percentile = 39.1–81.2 %).

Changes in the evaluation items before and after the intervention are presented in Table 4. Group V included more physically inferior participants than the other 2 groups at baseline. During the 3 months, different responses to the intervention were observed in each group. In group V, significant improvements were observed in the VAS score for low back pain, single-leg standing time (both right and left legs), 6-m walking time (both with normal steps and with quick steps), and PCS score. In group P, significant improvements were observed in the single-leg standing time (left leg) and PCS and RDQ scores. In group C, no significant change was observed in any parameter during the 3 months.

Thereafter, the changes in each parameter were compared among groups V, P, and C (Table 5). Split-plot ANOVA revealed a significant interaction for the VAS of low back pain, single-leg standing time (left leg), and 6-m walking time for both normal and quick steps among the 3 groups. The baseline values of the 6-m walking time (with both normal and quick steps) and PCS score were significantly different among the 3 groups. ANOVA revealed significant differences in changes in the VAS of low back pain, left-leg standing time, 6-m walking time (both with normal steps and with quick steps), and PCS score among the groups. Statistical differences were observed in these parameters excluding the PCS score after adjustment by the covariate (i.e., baseline value) in ANCOVA. Moreover, statistical differences were similarly observed in the same parameters (VAS of low back pain: $p = 0.0471$; left-leg standing time: $p = 0.0205$; 6-m walking time with normal steps: $p = 0.0155$; 6-m walking time with quick steps: $p = 0.0422$) when gender was added as a covariate.

During the 3 months, the numbers of locomotive syndrome participants who withdrew from the study were 0, 2, and 3 in groups V, P, and C, respectively. In total, 6 of 17 participants in group V, 3 of 13 participants in group P, and 0 of 9 participants in group C recovered from locomotive syndrome after the intervention.

Discussion

As stated in the "Introduction" the increasing number of elderly persons who need nursing care is becoming an urgent social issue in many countries. At the beginning of this study, we indicated 3 purposes for conducting an exercise intervention to recover from locomotive syndrome.

The first purpose of this study was to develop an exercise program that the elderly could sustainably perform in the community or at home. It was intended that the video exercise program in this study would be performed without difficulty by the elderly and would thus improve their physical performance and prevent the need for long-term care. Each exercise that we selected has been conventionally used for patients, depending on their condition. The

Table 4. Changes in the evaluation items before and after intervention

	Group V ($n = 38$)		Group P ($n = 37$)		Group C ($n = 30$)	
	Baseline	3 months later	Baseline	3 months later	Baseline	3 months later
Visual analog scales						
Low back pain	38.5 ± 22.3	25.6 ± 18.0**	32.7 ± 21.2	31.5 ± 27.7	28.2 ± 26.4	22.3 ± 24.3
Knee pain	19.5 ± 19.5	21.8 ± 23.2	19.8 ± 17.3	14.8 ± 15.7	21.6 ± 23.1	16.9 ± 17.5
Single-leg standing time						
Right leg (s)	33.7 ± 22.9	42.8 ± 21.8*	42.6 ± 22.7	42.4 ± 22.1	38.2 ± 25.0	38.9 ± 23.4
Left leg (s)	27.0 ± 22.5	39.3 ± 23.5**	37.7 ± 21.9	45.0 ± 19.3*	32.9 ± 23.4	34.9 ± 22.3
6-m walking time						
Normal steps (s)	5.7 ± 0.7	5.4 ± 0.9*	4.9 ± 1.1	5.3 ± 1.0*	5.1 ± 1.1	5.2 ± 1.0
Quick steps (s)	4.4 ± 0.6	4.2 ± 0.6**	3.8 ± 0.7	3.9 ± 0.7	4.1 ± 0.8	3.9 ± 0.6
Short Form-8 summary scores						
PCS	42.5 (38.6–47.9)	44.1 (40.4–49.1)*	45.1 (41.0–48.8)	47.9 (42.8–51.8)*	49.1 (44.8–52.1)	48.6 (43.5–52.7)
MCS	52.9 (49.0–55.7)	53.9 (48.3–56.7)	53.0 (49.3–56.3)	54.3 (51.4–56.4)	52.8 (47.2–55.2)	52.8 (48.0–55.2)
RDQ score	4.0 (2.0–9.0)	4.0 (0.0–11.0)	3.5 (0.0–7.3)	2.0 (0.0–6.5)*	2.0 (0.0–5.5)	1.0 (0.0–5.0)
ODI (% disability)	17.8 (7.2–30.6)	17.8 (6.7–30.6)	17.8 (8.9–23.9)	14.4 (5.0–24.4)	13.3 (3.3–22.2)	11.1 (2.8–21.7)
GLFS-25 score	14.0 (6.0–27.3)	10.0 (5.5–20)	10.0 (5.5–23)	7.0 (4.0–19.5)	10.0 (5.0–16.0)	9.0 (4.0–17.0)

Visual analog scales, single-leg standing time, and 6-m walking time presented as mean ± standard deviation. Short Form-8 summary scores, RDQ score, ODI, and GLFS-25 score presented as median (25–75 %)

Group V video exercise group, *Group P* pamphlet exercise group, *Group C* control group, *PCS* physical component scale, *MCS* mental component scale, *RDQ* Roland–Morris Disability Questionnaire, *ODI* Oswestry Disability Index, *GLFS-25* the 25-question Geriatric Locomotive Function Scale

*$p < 0.05$, **$p < 0.01$

first half of the program consists of quadriceps femoris exercise and fall-prevention exercises, which are reported to be effective for KOA [16]. The latter half of the program consists of exercises for low back pain. Many studies have confirmed the effectiveness of these exercises [6, 17]. The current exercise program was intended to improve general physical performance with the aim of preventing locomotive syndrome because the elderly often have multiple diseases affecting their mobility. During the development of the program, the first consideration was to avoid an excessive burden on the cardiopulmonary function of the participants. The pilot study, which examined the physical load on 20 elderly volunteers, demonstrated that pulse-rate changes before and after exercise did not exceed 20 % and that the SpO_2 change was within 4 points in all the cases. The RPE scores were 11–14 for all participants. These results confirmed that the video exercise program provided a moderate physical load for most of the elderly participants without imposing an excessive cardiopulmonary burden.

The second purpose of this study was to determine which medium of exercise instruction (video or pamphlet) is superior. We investigated the short-term efficacy of the exercise program with regard to physical performance in elderly participants and compared the effectiveness of the different teaching media. Several studies have used a video exercise program for frail elderly individuals [18, 19]. The superiority of video programs over written instructions has been reported for shoulder exercises [20] and for educating candidates for back surgery [21] and total knee arthroplasty [22]. Therefore, we expected that the video exercise program would relieve participants of body pain and improve their physical performance better than the pamphlet exercise program at the beginning of this study. In fact, the exercise achievement summary scale score in group V was significantly higher than that in group P. This finding indicated that the video exercise program provides greater motivation for participants than does the pamphlet exercise program. Moreover, intergroup comparisons of changes in the parameters after 3 months revealed statistically significant differences among the groups.

The third purpose of this study was to determine which parameter is most useful for evaluating the effectiveness of this intervention in preventing locomotive syndrome. We used the single-leg standing and 6-m walking tests as indices of physical performance to evaluate the effect of exercise on the elderly participants. The single-leg standing test has been reported to be a useful index for

Table 5. Intergroup comparison of the change of each parameter

	Interaction Split-plot ANOVA *p* value	Baseline value ANOVA *p* value	Effect of intervention	
			ANOVA *p* value	ANCOVA *p* value
Visual analog scales				
Low back pain	**0.027**	0.250	**0.037**	
Knee pain	0.266	0.892	0.182	
Single-leg standing time				
Right leg	0.150	0.154	0.150	
Left leg	**0.024**	0.069	**0.009**	
6-m walking time				
Normal steps	**<0.001**	**<0.001**	**<0.001**	**0.022**
Quick steps	**0.001**	**0.001**	**0.001**	**0.040**
Short Form-8 summary scores				
PCS	0.088	**0.016**	**0.014**	0.426
MCS	0.798	0.316	0.230	
RDQ score	0.436	0.454	0.099	
ODI (% disability)	0.803	0.669	0.297	
GLFS-25 score	0.347	0.690	0.508	

Values in bold are statistically significant

ANOVA analysis of variance, *ANCOVA* analysis of covariance, *PCS* physical component scale, *MCS* mental component scale, *RDQ* Roland–Morris Disability Questionnaire, *ODI* Oswestry Disability Index, *GLFS-25* the 25-question Geriatric Locomotive Function Scale

examining elderly populations [23]. Moreover, gait velocity has been reported to be sensitive to changes in mobility in frail elderly individuals [23–25]. Although the observation period of this study was short, the values of the 2 measurements revealed significant improvement, at least in group V. These 2 measurements may be useful indices for evaluating short-term effects on physical performance (mobility and static balance) in elderly individuals.

This study had several limitations. First, the grouping of the participants was not perfectly randomized. Group V included more physically inferior participants than the other 2 groups. We permitted group changes as requested by the participants because they were recruited from a community-based population. It was difficult to place couples and friends in different groups without the risk of information leakage. The short period (3 months) of the intervention/observation was another limitation of this study. If the aim of the study is to prevent the need for care in elderly individuals, then a longer observation time is necessary. Further investigation is necessary regarding the long-term effects on society, such as changes in medical costs and the number of elderly individuals requiring nursing care. Regarding the static balance exercise, the significant improvement in the single-leg standing time after 3 months may be surprising especially because our exercise program contains only a short (i.e., 10 s/cycle $\times$ 4) single-leg standing exercise. The JOA and the Japanese Clinical Orthopaedic Association (JCOA) recommend the one-leg standing balance exercise for 1 min to prevent falls and hip fractures [26, 27]. A systematic review concerning falls prevention suggests that greater relative effects are observed in programs that include exercises that challenge balance (exercises conducted while standing in which people aim to stand with their feet closer together or on 1 leg, minimize the use of their hands to assist, and practice controlled movements of the center of gravity), use a higher intensity of exercise, and do not include a walking program [28]. A Cochrane review including 94 studies of balance exercise in the elderly suggests that the more effective programs ran 3 times a week for 3 months and involved dynamic exercise in standing [29]. We believe our results do not contradict the summary of the 2 systematic reviews. Moreover, we speculate that the improvement of our participants in a short time may be due to the mildness of locomotive disability among the participants. Our participants were community-dwelling elderly subjects, whereas the subjects in the JOA and JCOA reports were clinic patients [26, 27]. We may need to verify the possibility that the participants were healthy apart from their locomotive disability because the participants of this study were community-dwelling individuals who were recruited by open invitation. However, the data presented in Table 2 suggest that the participants were not particularly healthy with respect to their musculoskeletal conditions. The mean ages of the male and female participants of the current study were 74.1 and 71.6 years, respectively. In a study of a large-scale population-based cohort in Japan, Yoshimura et al. [4] found that the prevalences of KOA, LS, and lumbar OP in the group aged 70–79 years were 48.2, 85.3, and 3.6 %, respectively, among men and 71.9, 75.1, and 29.8 %, respectively, among women. Muraki et al. [30, 31] reported that, among the radiological osteoarthritis-affected subjects, one-fourth of the male participants and one-third of the female participants experienced pain. The national standard values (median) of PCS and MCS scores were 47.5 and 53.1, respectively, for men, and 47.3 and 53.6, respectively, for women (age range 70–75 years) [14]. Although careful judgment is required when comparing our data with those of previous studies, it may be safely said that our participants constituted a typical group for their age with respect to their mobility, rather than a particularly healthy volunteer group.

In conclusion, this study confirmed the safety of the exercise program we developed and indicated that the video exercise program provides greater motivation to participants than does the pamphlet exercise program. Moreover, this study demonstrated that our exercise program may improve low back pain and functional disability in participants. The single-leg standing and 6-m walking tests are possibly useful indices for evaluating the short-term effects of exercise on balance and mobility in the elderly. Although these results are preliminary, we believe that this study provides fundamental information for future studies.

Acknowledgments

This work was supported by the consigned research fund of Wakayama Prefecture Nos. B-21006 and B-22014. The authors thank Dr. Hiroaki Terashita of Mikimachi Terashita Orthopaedic Clinic for his precise advice in planning this study and Mr. Yoshihisa Kitaoka and other members of the public office in Hashimoto City for their assistance in selecting locations and scheduling participants for examinations.

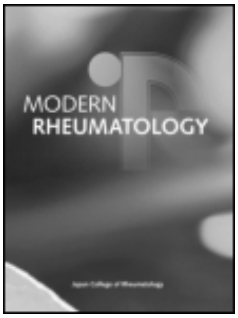
Conflict of interest

The video exercise program in this study was developed with technical cooperation from Wakayama Telecasting Corp (WTV). WTV developed a commercial DVD video containing this exercise program under the editorial supervision of M.Y. and H.H. However, none of the authors received any benefits from the company.

References

1. Statistics Bureau. Current population estimates as of October 1, 2007. Available at: http://www.stat.go.jp/english/data/jinsui/2007np/index.htm. Accessed Mar 13, 2012.
2. Nakamura K. A "super-aged" society and the "locomotive syndrome". J Orthop Sci. 2008;13:1–2.
3. Ministry of Health, Labour and Welfare. Comprehensive survey of living conditions. 14. Rate of Persons with subjective symptoms (per 1000 population) by type of symptoms, sex and age group (multiple answer), 2007. Available at: http://www.mhlw.go.jp/english/database/db-hss/cslc-tables.html. Accessed Mar 13, 2012.
4. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab. 2009;27:620–8.
5. Nakamura K. Locomotive syndrome: disability-free life expectancy and locomotive organ health in a "super-aged" society. J Orthop Sci. 2009;14:1–2.
6. Hayden JA, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. Cochrane Database Syst Rev. 2005;3:CD000335.
7. Richmond J, Hunter D, Irrgang J, Jones MH, Levy B, Marx R, et al. Treatment of osteoarthritis of the knee (nonarthroplasty). J Am Acad Orthop Surg. 2009;17:591–600.
8. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc. 1982;14:377–81.
9. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16:494–502.
10. Orimo H, Hayashi Y, Fukunaga M, Sone T, Fujiwara S, Shiraki M, et al. Diagnostic criteria for primary osteoporosis: year 2000 revision. J Bone Miner Metab. 2001;19:331–7.
11. Fujiwara A, Kobayashi N, Saiki K, Kitagawa T, Tamai K, Saotome K. Association of the Japanese Orthopaedic Association score with the Oswestry Disability Index, Roland-Morris Disability Questionnaire, and short-form 36. Spine. 2003;28:1601–7.
12. Suzukamo Y, Fukuhara S, Kikuchi S, Konno S, Roland M, Iwamoto Y, et al. Validation of the Japanese version of the Roland-Morris Disability Questionnaire. J Orthop Sci. 2003;8:543–8.
13. Nakamura M, Miyamoto K, Shimizu K. Validation of the Japanese version of the Roland-Morris Disability Questionnaire for Japanese patients with lumbar spinal disorders. Spine. 2003;28:241–8.
14. Fukuhara S, Suzukamo Y. Manual of the SF-8 Japanese version: Institute for Health Outcomes and Process Evaluation Research, Kyoto, 2004 (in Japanese).
15. Seichi A, Hoshino Y, Doi T, Akai M, Tobimatsu Y, Iwaya T. Development of a screening tool for risk of locomotive syndrome in the elderly: the 25-question Geriatric Locomotive Function Scale. J Orthop Sci. 2012;17:163–72.

16. Roddy E, Zhang W, Doherty M. Aerobic walking or strengthening exercise for osteoarthritis of the knee? A systematic review. Ann Rheum Dis. 2005;64:544–8.
17. Rainville J, Hartigan C, Martinez E, Limke J, Jouve C, Finno M. Exercise as a treatment for chronic low back pain. Spine J. 2004;4:106–15.
18. Vestergaard S, Kronborg C, Puggaard L. Home-based video exercise intervention for community-dwelling frail older women: a randomized controlled trial. Aging Clin Exp Res. 2008;20:479–86.
19. Haines TP, Russell T, Brauer SG, Erwin S, Lane P, Urry S, et al. Effectiveness of a video-based exercise programme to reduce falls and improve health-related quality of life among older adults discharged from hospital: a pilot randomized controlled trial. Clin Rehabil. 2009;23:973–85.
20. Reo JA, Mercer VS. Effects of live, videotaped, or written instruction on learning an upper-extremity exercise program. Phys Ther. 2004;84:622–33.
21. Phelan EA, Deyo RA, Cherkin DC, Weinstein JN, Ciol MA, Kreuter W, et al. Helping patients decide about back surgery: a randomized trial of an interactive video program. Spine. 2001;26:206–11.
22. Lin PC, Lin LC, Lin JJ. Comparing the effectiveness of different educational programs for patients with total knee arthroplasty. Orthop Nurs. 1997;16:43–9.
23. van Iersel MB, Munneke M, Esselink RA, Benraad CE, Olde Rikkert MG. Gait velocity and the Timed-Up-and-Go test were sensitive to changes in mobility in frail elderly patients. J Clin Epidemiol. 2008;61:186–91.
24. Shinkai S, Watanabe S, Kumagai S, Fujiwara Y, Amano H, Yoshida H, et al. Walking speed as a good predictor for the onset of functional dependence in a Japanese rural community population. Age Ageing. 2000;29:441–6.
25. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med. 1995;332:556–61.
26. Sakamoto K, Nakamura T, Hagino H, Endo N, Mori S, Muto Y, et al. Effects of unipedal standing balance exercise on the prevention of falls and hip fracture among clinically defined high-risk elderly individuals: a randomized controlled trial. J Orthop Sci. 2006;11:467–72.
27. Kita K, Hujino K, Nasu T, Kawahara K, Sunami Y, Japanese Clinical Orthopaedic Association, Committee on Musculoskeletal Rehabilitation. A simple protocol for preventing falls and fractures in elderly individuals with musculoskeletal disease. Osteoporos Int. 2007;18:611–9.
28. Sherrington C, Whitney JC, Lord SR, Herbert RD, Cumming RG, Close JC. Effective exercise for the prevention of falls: a systematic review and meta-analysis. J Am Geriatr Soc. 2008;56:2234–43.
29. Howe TE, Rochester L, Neil F, Skelton DA, Ballinger C. Exercise for improving balance in older people. Cochrane Database Syst Rev. 2011;11:CD004963.
30. Muraki S, Oka H, Akune T, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of radiographic lumbar spondylosis and its association with low back pain in elderly subjects of population-based cohorts: the ROAD study. Ann Rheum Dis. 2009;68:1401–6.
31. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthr Cartil. 2009;17:1137–43.



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

Association of knee osteoarthritis with onset and resolution of pain and physical functional disability: The ROAD study

Shigeyuki Muraki, Toru Akune, Keiji Nagata, Yuyu Ishimoto, Munehito Yoshida, Fumiaki Tokimura, Sakae Tanaka, Hiroyuki Oka, Hiroshi Kawaguchi, Kozo Nakamura & Noriko Yoshimura

To cite this article: Shigeyuki Muraki, Toru Akune, Keiji Nagata, Yuyu Ishimoto, Munehito Yoshida, Fumiaki Tokimura, Sakae Tanaka, Hiroyuki Oka, Hiroshi Kawaguchi, Kozo Nakamura & Noriko Yoshimura (2014) Association of knee osteoarthritis with onset and resolution of pain and physical functional disability: The ROAD study, Modern Rheumatology, 24:6, 966-973, DOI: 10.3109/14397595.2014.883055

To link to this article: http://dx.doi.org/10.3109/14397595.2014.883055

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://informahealthcare.com/mor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2014; 24(6): 966–973

DOI: 10.3109/14397595.2014.883055

informa
healthcare

ORIGINAL ARTICLE

Association of knee osteoarthritis with onset and resolution of pain and physical functional disability: The ROAD study

Shigeyuki Muraki[1], Toru Akune[1], Keiji Nagata[2], Yuyu Ishimoto[2], Munehito Yoshida[2], Fumiaki Tokimura[3], Sakae Tanaka[4], Hiroyuki Oka[5], Hiroshi Kawaguchi[4], Kozo Nakamura[6], and Noriko Yoshimura[5]

[1]*Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Tokyo, Japan,* [2]*Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan,* [3]*Department of Orthopaedic Surgery, Tokyo Metropolitan Geriatric Medical Center, Tokyo, Japan,* [4]*Department of Orthopaedic Surgery, Faculty of Medicine, the University of Tokyo, Tokyo, Japan,* [5]*Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Tokyo, Japan, and* [6]*National Rehabilitation Center for Persons with Disabilities, Saitama, Japan*

Abstract

Objectives. To examine the onset and resolution of pain and physical functional disability using Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and their association with knee osteoarthritis (OA) in the longitudinal large-scale population of the nationwide cohort study, Research on Osteoarthritis/osteoporosis Against Disability (ROAD).

Methods. Subjects from the ROAD study who had been recruited during 2005–2007 were followed up 3 years later. A total of 1,578 subjects completed the WOMAC questionnaire at baseline and follow up, and the onset and resolution rate of pain and physical functional disability were examined. We also examined the association of onset of pain and physical functional disability and their resolution with severity of knee OA as well as age, body–mass index and grip strength.

Results. After a 3.3-year follow-up, the onset rate of pain was 35.0% and 35.3% in men and women, respectively, and the onset rate of physical functional disability was 38% and 40%, respectively. Resolution rate of pain was 20.3% and 26.2% in men and women, respectively, and resolution rate of physical functional disability was 16% and 14% in men and women, respectively. Knee OA was significantly associated with onset and resolution of pain and physical functional disability in women, but there was no significant association of knee OA with onset of pain and resolution of physical functional disability in men.

Conclusions. The present longitudinal study revealed the onset rate of pain and physical functional disability as well as their resolution, and their association with knee OA.

Keywords

Knee joint, Osteoarthritis, Epidemiology, Longitudinal studies

History

Received 9 September 2013
Accepted 10 January 2014
Published online 24 February 2014

Introduction

Knee osteoarthritis (OA), characterized by pathological features including joint space narrowing and osteophytosis, is a major public health issue causing chronic pain and disability among the elderly in most developed countries [1]. The prevalence of radiographic knee OA in Japan is high [2], with 25,300,000 subjects aged 40 years and older estimated to experience radiographic knee OA [3]. According to the recent National Livelihood Survey of the Ministry of Health, Labour and Welfare in Japan, OA is ranked fourth among diseases that cause disabilities that subsequently require support with activities of daily living [4].

The principal clinical symptoms of knee OA are pain and physical functional disability [5], but the correlation of these symptoms with radiographic severity of knee OA is controversial [2,6–8]. Thus it would be interesting to determine whether the impact of radiographic knee OA on pain and physical functional disability differs according to the severity of OA. In terms of disease-specific scales for estimating pain and physical functional disability due to knee OA, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) has been used for Caucasians [9] and Asians [10,11], although these reports were not population-based studies. Furthermore, there is little information on the impact of knee OA on onset of pain and physical functional disability using WOMAC in Japan, although a population survey suggests that the disease pattern differs among races [12–14]. In addition, to the best of our knowledge, although pain and physical functional disability can disappear or improve, there is no information on the impact of knee OA on the resolution of pain and physical functional disability.

Grip strength is a useful marker of muscle function and sarcopenia [15]. There is growing evidence that reduced grip strength is associated with adverse outcomes including morbidity, disability, falls, higher fracture rates, increased length of hospital stay and mortality [16–18]. A previous study also showed that grip strength is related to total muscle strength [19]. Thus, the association of knee OA with pain and physical functional disability may be influenced by grip strength, but again, no studies have examined the association of knee OA and grip strength with onset of pain and disability as well as their resolution simultaneously in the same population using a longitudinal cohort study.

Correspondence to: Shigeyuki Muraki, MD, PhD, Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Tokyo, Japan. Tel: +81-3-5800-9178. Fax: +81-3-5800-9179. E-mail: murakis-ort@h.u-tokyo.ac.jp

The objective of the present study was to clarify the onset and resolution rate of pain and physical functional disability using WOMAC in Japanese men and women who were part of the large-scale, longitudinal, population-based cohort study known as the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study. In addition, we examined the association of body–-mass index (BMI), grip strength and severity of knee OA with onset of pain and physical functional disability as well as their resolution in men and women.

Materials and methods

Subjects

The ROAD study was a nationwide prospective study for bone and joint diseases (with OA and osteoporosis as the representative bone and joint diseases) constituting population-based cohorts established in several communities in Japan. As a detailed profile of the ROAD study has already been described elsewhere [2,3,20], only a brief summary is provided here. During 2005–2007, we created a baseline database that included clinical and genetic information for 3,040 inhabitants (1,061 men; 1,979 women) aged 23–95 years (mean, 70.6 years), recruited from listings of resident registrations in three communities: an urban region in Itabashi, Tokyo; a mountainous region in Hidakagawa, Wakayama; and a coastal region in Taiji, Wakayama. All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology. Participants completed an interviewer-administered questionnaire of 400 items that included lifestyle information such as smoking habit, alcohol consumption, family history, medical history and previous knee injury history. Furthermore, subjects were interviewed by well-experienced orthopedists regarding the treatment for knee OA, such as medication, injections, physical therapy, bracing, etc. between the baseline and follow-up study. Anthropometric measurements included height and weight, from which BMI (weight [kg]/height^2 [m^2]) was calculated. Grip strength was measured on bilateral sides using a TOEI LIGHT handgrip dynamometer (Toei Light Co., Ltd., Saitama, Japan), and the better measurement was used to represent maximum muscle strength. During 2008–2010, we attempted to trace and review all 3,040 subjects; they were invited to attend a follow-up interview. The interviews were conducted by the same trained orthopedists who undertook the baseline study during 2005–2007.

Radiographic assessment

All participants underwent radiographic examination of both knees using an anterior–posterior view with weight-bearing and foot map positioning. Fluoroscopic guidance with a horizontal anterior–posterior X-ray beam was used to properly visualize the joint space. Knee radiographs at baseline and follow-up were read in pairs without knowledge of the participant's clinical status by a single well-experienced orthopedist (S.M.), and the Kellgren Lawrence (KL) grade was defined using the KL radiographic atlas for overall knee radiographic grades [21]. In the KL grading system, radiographs are scored from grade 0 to grade 4, with the higher grades being associated with more severe OA. To evaluate the intraobserver variability of the KL grading, 100 randomly selected radiographs of the knee were scored by the same observer more than 1 month after the first reading. One hundred other radiographs were also scored by two experienced orthopedic surgeons (S.M. & H.O.) using the same atlas for interobserver variability. The intra- and inter variabilities evaluated for KL grades (0–4) were confirmed by kappa analysis to be sufficient for assessment (0.86 and 0.80, respectively).

Instruments

The WOMAC, a 24-item OA-specific index, consists of three domains: pain, stiffness and physical function. Each of these 24 items is graded on either a 5-point Likert scale or a 100-mm visual analog scale [22,9]. In the present study, we used the Likert scale (version LK 3.0). The domain score ranges from 0 to 20 for pain, 0 to 8 for stiffness and 0 to 68 for physical function. Japanese versions of the WOMAC have also been validated [23]. In the present study, onset of pain and physical functional disability were defined as WOMAC pain score $= 0$ at baseline and > 0 at follow up and WOMAC physical function score $= 0$ at baseline and > 0 at follow up, respectively. Resolution of pain and physical functional disability were defined as WOMAC pain score > 0 at baseline and $= 0$ at follow up and WOMAC physical function score > 0 at baseline and $= 0$ at follow up, respectively. Worsening pain and physical functional disability were defined as WOMAC pain and physical function at follow up was worse than at baseline, respectively.

Statistical analysis

The differences in age, height, weight, BMI, grip strength, and WOMAC pain and physical function scores at baseline and follow up between men and women were examined using a non-paired Student's t-test. The prevalence of knee OA was compared between men and women using chi-square test. Tukey's honestly significant difference test after adjustment for age and BMI was used to compare WOMAC pain and physical functional score and differences between baseline and follow up among subjects with KL = 0/1, 2 and 3/4. The non-paired Student's t test was used to compare age, BMI and grip strength between subjects with and without onset of pain and physical functional disability as well as those with and without resolution of pain and physical functional disability. Chi-square test was used to compare prevalence of knee OA between subjects with and without onset of pain and physical functional disability as well as those with and without resolution of pain and physical functional disability. Multiple logistic regression analysis after adjustment for age was also used to determine the association of severity of knee OA with onset of pain and physical functional disability as well as their resolution. In addition, to determine independent association of age, BMI, grip strength and knee OA with onset of pain and physical function as well as their resolution, multiple logistic regression analysis was used with significant variables ($p < 0.01$) in univariate analyses as explanatory variables. Data analyses were performed using SAS version 9.0 (SAS Institute Inc., Cary, NC).

Results

Of the 3,040 subjects in the baseline study during 2005–2007, 125 had died by the time of the review held 3 years later, 123 did not participate in the follow-up study due to bad health, 69 had moved away, 83 declined the invitation to attend the follow-up study, and 155 did not participate in the follow-up study for other reasons. Among the 2,485 subjects who did participate in the follow-up study, we excluded 39 subjects who were younger than 40 years at baseline. Those participating in the follow-up study were younger than those who did not survive or who did not participate for other reasons (responders 68.6 years, non-responders 75.1 years; $p < 0.0001$). The follow-up study participants also were more likely to be women (responders 66.3% women, nonresponders 61.8% women; $P = 0.03$) and were more likely to have knee OA at the baseline examination than either those who did not survive to follow-up or those who did not participate for other reasons (responders 51.5%, nonresponders 60.9%; $P < 0.0001$). Among them, 1,578 subjects provided completed WOMAC questionnaires both at baseline and follow up. We also excluded three subjects

Table 1. Characteristics of subjects.

	Overall	Men	Women	p value
N	1558	553	1005	
Age	67.0 ± 11.0	68.1 ± 10.7	66.5 ± 11.0	0.004
Height	155.2 ± 8.9	163.4 ± 6.5	150.8 ± 6.5	<0.0001
Weight	55.5 ± 10.4	62.2 ± 10.2	51.8 ± 8.5	<0.0001
BMI	22.9 ± 3.3	23.2 ± 3.1	22.8 ± 3.3	0.0043
Grip strength	27.2 ± 9.5	35.4 ± 8.7	22.7 ± 6.4	<0.0001
Knee OA (%)	49.3	38.7	55.2	<0.0001
WOMAC at baseline				
Pain	1.12 ± 2.18	1.02 ± 2.05	1.18 ± 2.25	0.157
Physical function	3.03 ± 6.63	2.56 ± 5.71	3.29 ± 7.07	0.0268
WOMAC at follow up				
Pain	1.82 ± 2.83	1.72 ± 2.67	1.88 ± 2.91	0.291
Physical function	5.59 ± 9.7	4.73 ± 8.30	6.06 ± 10.36	0.0061

Knee OA was defined as Kellgren Lawrence grade 2 or worse at baseline.
BMI, body–mass index; OA, osteoarthritis; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

who did not undergo plain radiography at knee and 17 subjects who underwent total knee arthroplasty before the follow-up study, leaving a total of 1,558 subjects (553 men and 1,005 women). The mean duration between baseline and follow up was 3.3 ± 0.6 years.

The characteristics of the 1,578 participants at baseline in the present study are shown in Table 1. Men were significantly older than women, and BMI was significantly higher in men than in women. The prevalence of knee OA was significantly higher in women than in men at baseline. WOMAC pain score was not significantly different between gender, while, physical function score was significantly worse in women than in men at baseline and follow up. The scores of WOMAC pain and physical function scores worsened at follow up compared with those at baseline in men and women ($p < 0.05$).

The scores of WOMAC pain and physical function scores and their differences between baseline and follow up according to the KL grade are shown in Supplementary Table 1 available online at http://informahealthcare.com/doi/abs/10.3109/14397595.2014.883055. In men, differences in WOMAC physical function scores were significantly larger in subjects with KL 3/4 than those with KL 0/1 after adjustment for age and BMI, while differences in WOMAC pain scores were not. In women, after adjustment for age and BMI, differences in WOMAC pain and physical function scores between baseline and follow up were significantly larger in subjects with KL 3/4 than those with KL 0/1.

Among 366 men and 634 women in subjects without pain at baseline, 128 (35.0%) men and 224 (35.3%) women had onset of pain at follow up (Table 2). In men, subjects with onset of pain tended to be older than those without pain, while BMI and grip strength were not significantly different between them. In women, age and BMI were significantly different between subjects with and without onset of pain, and grip strength tended to be higher in subjects with onset of pain than those without pain. Among 346 men and 601 subjects without physical functional disability at baseline, 132 (38.2%) men and 243 (40.4%) women had onset of physical functional disability at follow up (Table 2). Age and BMI were significantly different between subjects with and without onset of physical functional disability in both men and women, and BMI tended to be higher in subjects with onset of physical functional disability than those without it in women only.

We next examined onset of pain and physical functional disability according to KL grade (Figure 1). There were no significant differences in onset of pain among men with KL 0/1 knee, KL 2 knee OA and KL 3/4 knee OA (33.3%, 36.0% and 46.2%, respectively, $p = 0.4149$ by chi-square test), while there were significant differences in onset of pain among women with KL 0/1 knee, KL 2 knee OA and KL 3/4 knee OA (30.4%, 38.6% and 48.5%, respectively, $p = 0.0082$ by chi-square test). Multiple logistic regression analysis after adjustment for age showed that women with KL 3/4 knee OA had significant higher onset of pain compared with those with KL 0/1. There were significant differences in onset of physical functional disability among subjects with KL 0/1 knee OA, KL 2 knee OA and KL 3/4 knee OA in men and women (men 33.2%, 41.7% and 66.7%, respectively, $p = 0.0023$ by chi-square test, women 35.8%, 43.8% and 53.1%, respectively, $p = 0.0165$ by chi-square test). Multiple logistic regression analysis after adjustment for age showed that men with KL 3/4 knee OA had a significant higher onset of physical functional disability compared with those with KL 0/1.

In addition, we examined the association of age, BMI, grip strength and WOMAC pain and physical function scores at baseline with resolution of pain and physical functional disability (Table 3). Among 187 men and 371 women with WOMAC pain at baseline, pain disappeared in 38 (20.3%) men and 97 (26.2%) women at follow up. In men, WOMAC pain score at baseline was significantly different between subjects with resolution of pain and those with continuous pain. BMI tended to be higher in subjects with continuous pain than in those with resolution of pain. In women, age, BMI, grip strength and WOMAC pain score at baseline were significantly different between subjects with resolution of pain and those with continuous pain. Among 207 men and 404 women with physical functional disability at baseline,

Table 2. Age, BMI, grip strength, and WOMAC pain and physical function score according to onset of pain and physical functional disability in subjects without pain and physical functional disability at baseline.

	Pain $N = 1{,}000$			Physical function $N = 947$		
	Continuous no pain	Onset of pain	p value	Continuous no physical functional disability	Onset of physical functional disability	p value
Men						
N	238	128		214	132	
Age	65.3 ± 11.3	67.6 ± 10.8	0.056	63.3 ± 11.0	68.9 ± 10.2	<0.0001
BMI	23.1 ± 3.1	23.1 ± 2.8	0.7981	23.1 ± 3.0	23.0 ± 3.2	0.8946
Grip strength	37.1 ± 8.8	36.6 ± 9.3	0.6531	37.4 ± 8.6	35.9 ± 9.1	0.0149
Women						
N	410	224		358	243	
Age	62.7 ± 11.0	65.4 ± 9.9	0.0017	60.2 ± 10.4	65.7 ± 10.0	<0.0001
BMI	22.0 ± 3.1	22.7 ± 3.1	0.0023	22.2 ± 3.1	22.6 ± 3.1	0.0823
Grip strength	24.2 ± 6.4	23.3 ± 6.5	0.0948	25.3 ± 6.5	22.8 ± 5.3	<0.0001

Values are the means ± standard deviation.
BMI, body mass index; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

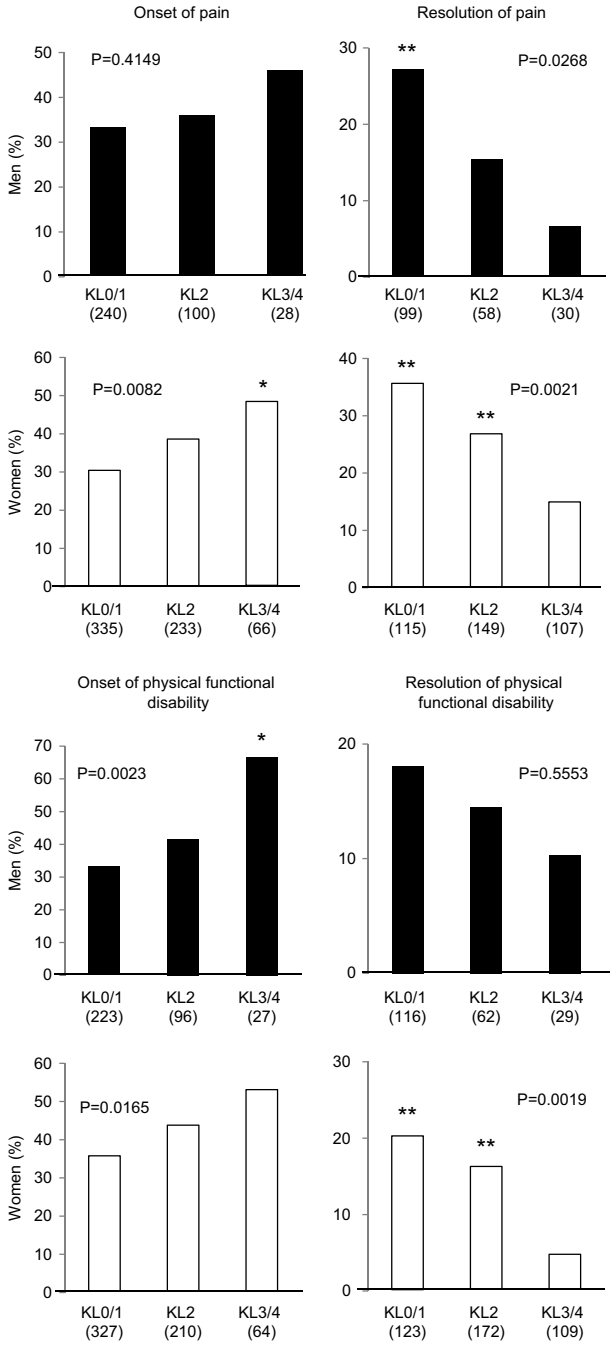


Figure 1. Onset and resolution rate of pain and physical functional disability according to Kellgren Lawrence (KL) grade in men and women. The number of subjects in each subgroup is shown in brackets. Chi-square test was used to determine the association of KL grade with onset of pain and physical functional disability as well as their resolution. $^{*}p < 0.05$ versus KL grade 0/1 by multiple logistic regression analysis after adjustment for age. $^{**}p < 0.05$ versus KL grade 3/4 by multiple logistic regression analysis after adjustment for age.

disability disappeared in 33 (15.9%) men and 58 (14.4%) women at follow up. In men, age and grip strength were significantly different between subjects with resolution of physical functional disability and those with continuous physical functional disability. Age, BMI, grip strength and WOMAC physical function score at baseline were significantly different between subjects with resolution of physical functional disability and those with continuous physical functional disability. In women, age, BMI, grip strength and WOMAC physical function score at baseline were significantly different between subjects with resolution of physical functional disability and those with continuous physical functional disability.

We next examined resolution of pain and physical functional disability according to KL grade (Figure 1). There were significant differences in resolution of pain among subjects with KL 0/1 knee, KL 2 knee OA and KL 3/4 knee OA in men and women (men 27.3%, 15.5% and 6.7%, respectively, $p = 0.0268$ by chi-square test; women 35.7%, 26.8% and 15.0%, respectively, $p = 0.0021$ by chi-square test). Multiple logistic regression analysis after adjustment for age showed that men with KL 3/4 knee OA had a significantly higher onset of pain compared with those with KL 0/1. Regarding resolution of physical functional disability, there were no significant differences among subjects with KL 0/1 knee, KL 2 knee OA and KL 3/4 knee OA in men (18.1%, 14.5% and 10.3%, respectively, $p = 0.5553$ by chi-square test), while there were significant differences subjects with KL 0/1 knee, KL 2 knee OA and KL 3/4 knee OA in women (20.3%, 16.3% and 4.6%, respectively, $p = 0.0019$ by chi-square test). Multiple logistic regression analysis after adjustment for age showed that women with KL 2 and 3/4 knee OA had a significantly higher onset of physical functional disability compared with those with KL 0/1.

To determine the independent association of age, BMI, grip strength and KL grade with onset of pain and physical functional disability, we next used multiple logistic regression analysis with significant variables ($p < 0.01$) by non-paired Student's t test or chi-square test shown in Table 2 and Figure 1 as explanatory variables (Table 4). Regarding onset of pain, there were no significant variables in men; thus, we did not examine the independent association with onset of pain. In women, older age and higher BMI were independently associated with onset of pain. Older age and KL 3/4 knee OA were independent risk factors for onset of physical functional disability in men, whereas older age, higher BMI and weaker grip strength were independent risk factors for onset of physical functional disability in women. The significant association of knee OA with onset of physical functional disability disappeared after adjustment age, BMI and grip strength in women.

We also examined independent associations of age, BMI, grip strength and KL grade with resolution of pain and physical functional disability (Table 5). KL 0/1 knee and lower WOMAC pain score at baseline were significantly associated with resolution of pain in men, whereas lower BMI, higher grip strength and lower WOMAC pain score were significantly associated with resolution of pain in women. Regarding physical function, only age was significantly associated with resolution of physical functional disability in men, whereas higher grip strength, KL 2 knee OA and lower WOMAC physical function score were significantly associated with resolution of physical functional disability in women. KL 01 knee also tended to be associated with resolution of physical functional disability in women. Because treatment for knee OA might affect the resolution of pain and physical functional disability, we further examined the association of treatment for knee OA with the resolution of pain and physical functional disability. Among subjects with pain at baseline, the resolution rate of pain was 36.2% in subjects who underwent treatment for knee OA, and 14.2% in subjects who did not undergo treatment for knee OA. Among subjects with physical functional disability at baseline, the resolution rate of physical functional disability was 19.3% in subjects who underwent treatment for knee OA, while, 7.2% in subjects who did not undergo treatment for knee OA. The resolution rate of pain and physical functional disability was significantly different between subjects who had and had not undergone treatment for knee OA (chi-square test, $p < 0.0001$). Thus, we examined independent associations of age, BMI, grip strength and KL grade with resolution of pain and physical functional disability after adjustment for the treatment for

Table 3. Age, BMI, grip strength, and WOMAC pain and physical function score according to resolution of pain and physical functional disability in subjects with pain and physical functional disability at baseline, respectively.

	Pain $N = 558$			Physical function $N = 611$		
	Resolution of pain	Continuous pain	p value	Resolution of physical functional disability	Continuous physical functional disability	p value
Men						
N	38	149		33	174	
Age	72.3 ± 8.9	71.9 ± 8.5	0.8	67.9 ± 11.6	73.4 ± 7.6	0.0118
BMI	22.8 ± 3.0	23.7 ± 3.3	0.08	23.4 ± 3.2	23.6 ± 3.2	0.8041
Grip strength	32.6 ± 6.4	32.4 ± 7.5	0.8694	34.9 ± 6.7	31.4 ± 7.3	0.0091
WOMAC at baseline						
Pain	1.82 ± 1.20	3.32 ± 2.69	< 0.0001	–	–	–
Physical function	–	–	–	4.85 ± 7.69	7.20 ± 7.58	0.1132
Women						
N	97	274		58	346	
Age	68.1 ± 12.6	72.4 ± 8.6	0.0022	68.1 ± 11.1	73.2 ± 8.2	0.0015
BMI	22.4 ± 3.2	24.0 ± 3.6	< 0.0001	22.3 ± 3.2	23.6 ± 3.6	0.0066
Grip strength	22.9 ± 7.2	19.8 ± 4.9	0.0002	23.7 ± 7.4	19.7 ± 5.4	0.0002
WOMAC at baseline						
Pain	1.84 ± 1.18	3.68 ± 2.90	< 0.0001	–	–	–
Physical function	–	–	–	3.33 ± 4.32	8.99 ± 9.54	< 0.0001

Values are the means ± standard deviation.
BMI, body mass index; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

knee OA. Results were similar to findings without adjustment for treatment of knee OA (Supplementary Table 2 available online at http://informahealthcare.com/doi/abs/10.3109/14397595.2014.883055). In addition, we examined associations of age, BMI, grip strength and severity of knee OA with worsening pain and physical functional disability in subjects with pain and physical functional disability at baseline (Supplementary Table 3 available online at http://informahealthcare.com/doi/abs/10.3109/14397595.2014.883055). Multiple logistic regression analysis showed that weaker grip strength was a risk factor for worsening pain, whereas KL 3/4 knee OA was a risk factor for worsening physical functional disability (Supplementary Table 4 available online at http://informahealthcare.com/doi/abs/10.3109/14397595.2014.883055).

Discussion

This is the first longitudinal population-based study to examine the onset, resolution and worsening of pain and physical functional disability using WOMAC. We also clarified the associations of age, BMI, grip strength and knee OA with the onset, resolution and worsening of pain and physical functional disability.

Our previous study showed that onset of knee pain during 3 years was approximately 20% and 30% in men and women, respectively [24]. The Chingford study also showed that more than 10% women had onset of pain during 2 years [25]. However, in these previous studies, knee pain was defined as present or absent, rather than as an established measure of pain such as WOMAC. In addition, in our previous study, we did not examine resolution of pain. In the present study, we found that 35% of men and women had onset of pain. These values were higher than onset values obtained from questionnaires in our previous study [24], indicating that WOMAC may be more powerful for detecting pain than questionnaires regarding only the presence or absence of pain. We also found that pain disappeared in approximately 20% men and 25% women using WOMAC. The Chingford study previously showed that knee pain disappeared in approximately 40% of Caucasian women during 2 years using a questionnaire on the presence and absence of pain [25], which is higher than the values

Table 4. Association of onset of pain and physical functional disability with age, BMI, grip strength, and KL grade.

	Onset of pain			Onset of physical functional disability		
	Adjusted OR	95% CI	p value	Adjusted OR	95% CI	p value
Men						
Age (+ 1 year)	–	–	–	1.05	1.02–1.08	0.0011
BMI (+ 1kg/m^2)	–	–	–	–	–	–
Grip strength (+ 1kg)	–	–	–	1.01	0.97–1.04	0.628
KL grade						
KL 0/1	–	–	–	1		
KL 2	–	–	–	1.02	0.60–1.72	0.9504
KL 3/4	–	–	–	2.7	1.14–6.69	0.0274
Women						
Age (+ 1 year)	1.02	1.003–1.04	0.023	1.05	1.03–1.07	< 0.0001
BMI (+ 1kg/m^2)	1.08	1.03–1.15	0.0047	1.08	1.02–1.14	0.0141
Grip strength (+ 1kg)	0.99	0.96–1.02	0.4977	0.96	0.92–0.99	0.0152
KL grade						
KL 0/1	1			1		
KL 2	1.09	0.74–1.61	0.6593	0.84	0.56–1.25	0.4035
KL 3/4	1.42	0.79–2.55	0.2337	1	0.54–1.82	0.9894

Multiple logistic regression analysis was used with significant variables ($p < 0.01$) in univariate models as explanatory variables.
BMI, body mass index; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

found in the present study. This discrepancy between our study and the Chingford study may be partly explained by age differences in addition to different estimations for pain and racial differences, because mean age was 52 years in the Chingford study compared with 67 years in the present study. Furthermore, we first found that approximately 40% men and women had onset of physical functional disability and approximately 15% men and women had resolution of physical functional disability. To our knowledge, no other community-based studies have described longitudinal patterns of physical functional disability, and the present study was the first to clarify the onset and resolution of physical functional disability using WOMAC.

Pain is the principal clinical symptom of knee OA [5], but, although much effort has been devoted to defining knee pain, the correlation with radiographic severity of the knee OA is not as strong as one would expect [2,6–8]. In the present study, we examined onset of pain according to KL grade using WOMAC. In men and women without knee OA (KL 0/1), more than 30% subjects had onset of pain. In addition, 50% of men and women with KL 3/4 knee OA had onset of pain, meaning that 50% did not have onset of pain despite having severe radiographic knee OA. In fact, in the present study, radiographic knee OA was not significantly associated with onset of pain in men, and after adjustment, the significant association of knee OA with onset of pain disappeared in women. These findings indicate that pain may arise from a variety of structures other than joint cartilage, such as menisci, synovium, ligaments, bursae, bone and bone marrow [26–30]. In addition, in the present study, the risk for onset of pain was higher with higher BMI rather than knee OA in women, indicating knee pain may be prevented by reducing obesity.

In the present study, we also examined the association of knee OA with the resolution of pain, and found that around 30% of men and women without knee OA had resolution of knee pain, which was a similar rate to onset of pain, and only 7% of men and 15% of women with severe knee OA had resolution of knee pain. These findings indicate that around 90% of subjects with severe knee OA had continuous knee pain. There were significant associations of resolution of pain with KL grade. Considering the results of onset of pain, severe knee OA may lead to difficulties with resolution of pain rather than onset of pain, particularly in men. In addition, after adjustment, resolution of pain was significantly associated with lower BMI and higher grip strength, which is a useful marker of muscle function and sarcopenia [15], rather than radiographic knee OA, indicating that improvement of obesity and performing muscle exercises may help make pain disappear. In addition, the significant association of BMI and grip strength remained after adjustment for treatment of knee OA, indicating that reducing obesity and performing muscle exercises may be as important as treatment to achieve resolution of pain due to knee OA.

We also found that severe knee OA was a risk factor for physical functional disability, particularly in men, despite the finding that severe knee OA was not significantly associated with onset of pain in men. Severe knee OA was not significantly associated with onset of physical functional disability after adjustment for age in women, despite the finding that severe knee OA was significantly associated with onset of pain. This discrepancy between gender may be partly explained by the idea that women are more susceptible to pain. In fact, our previous study showed that the prevalence of knee pain in women with KL 0/1, 2 and 3/4 knee OA was significantly higher than that in men with KL 0/1, 2 and 3/4 knee OA, respectively[2]. In addition, risk factors for onset of physical functional disability were higher BMI and weaker grip strength rather than knee OA in women in the present study. Grip strength is a useful marker of muscle function and sarcopenia [15]. A previous study also showed that grip strength is related to total muscle [19]. Results in the present study indicate that onset of physical functional disability may be prevented by improvement of obesity and muscle exercises.

In the present study, physical functional disability disappeared in 20% of women without knee OA, whereas physical functional disability disappeared only in 5% of women with severe knee OA. The association of knee OA with resolution of physical functional

Table 5. Association of resolution of pain and physical functional disability with age, BMI, grip strength, and KL grade.

	Resolution of pain			Resolution of physical functional disability		
	Adjusted OR	95% CI	p value	Adjusted OR	95% CI	p value
Men						
Age (+ 1 year)	–	–	–	0.95	0.90–0.9985	0.0443
BMI (+ 1kg/m^2)	0.92	0.80–1.04	0.1994	–	–	–
Grip strength (+ 1kg)	–	–	–	1.02	0.96–1.09	0.526
KL grade						
KL 3/4	1			–	–	–
KL 2	2.37	0.52–16.8	0.3042	–	–	–
KL 0/1	5.18	1.32–34.6	0.0378	–	–	–
WOMAC at baseline						
Pain	0.63	0.46–0.80	0.001	–	–	–
Physical function	–	–	–	–	–	–
Women						
Age (+ 1 year)	0.99	0.96–1.02	0.6031	0.98	0.95–1.02	0.4081
BMI (+ 1kg/m^2)	0.88	0.80–0.96	0.0034	0.93	0.84–1.02	0.1358
Grip strength (+ 1kg)	1.08	1.02–1.14	0.014	1.09	1.02–1.16	0.0123
KL grade						
KL 3/4	1			1		
KL 2	1.34	0.66–2.79	0.4312	3.04	1.15–9.62	0.0362
KL 0/1	1.71	0.79–3.77	0.1797	2.52	0.89–8.34	0.0997
WOMAC at baseline						
Pain	0.66	0.53–0.78	<0.0001	–	–	–
Physical function	–	–	–	0.87	0.78–0.93	0.0009

Multiple logistic regression analysis was used with significant variables ($p<0.01$) in univariate model as explanatory variables.

BMI, body mass index; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index; KL, Kellgren Lawrence grade.

disability remained significant after adjustment. This means that in women without knee OA, pain may occur, but it may disappear more easily. In addition, grip strength was also associated with resolution of physical functional disability after adjustment, indicating that muscle exercises may help make physical functional disability disappear.

The present study showed gender differences in the associations of knee OA with pain and physical functional disability. In women, knee OA was significantly associated with onset of pain and physical functional disability as well as their resolution, whereas in men, there were no significant association of knee OA with onset of pain and resolution of physical functional disability. Our previous cross-sectional study also showed that the odds ratio of knee pain for KL 3/4 knee OA was approximately twice as high in women as in men[2]. These findings may be partly explained by the lower muscle mass in women compared with men. In men, muscular strength may obscure the associations of knee OA with pain and physical functional disability.

In conclusion, the present longitudinal study revealed the onset rate of pain and physical functional disability as well as their resolution rate using WOMAC. In addition, severe knee OA was significantly associated with onset of pain and physical functional disability as well as their resolution, particularly in women. Furthermore, we also clarified that BMI and grip strength were associated with onset of pain and physical functional disability as well as their resolution in women.

Acknowledgements

This study was supported by Grants-in-Aid for Scientific Research (S19109007, B20390182, C20591737, C20591774), for Young Scientists (A18689031), and for Exploratory Research (19659305) from the Japanese Ministry of Education, Culture, Sports, Science and Technology, H17-Men-eki-009, H18-Choujyu-037, H20-Choujyu-009, H21-Chouju-Wakate-011 and H22-Chouju-Wakate-007 from the Ministry of Health, Labor and Welfare, Research Aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 & 2010-2), and Grant No.166 from the Japan Orthopaedics and Traumatology Foundation.

The authors thank Dr. Yamamoto, Dr. Ishibashi, Dr. Anamizu and members of Department of Orthopaedics, and Mr. Kutsuma and other members of Department of Radiology at Tokyo Metropolitan Geriatric Medical Center; Mrs. Tomoko Takijiri and other members of the Public Office in Hidakagawa Town; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda and other members of the Public Office in Taiji Town, for their assistance in the location and scheduling of participants for examinations.

Conflict of interest

None.

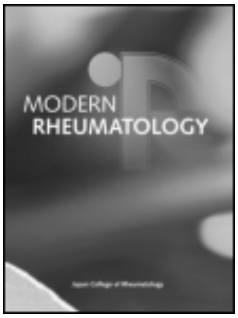
References

1. Sharma L, Kapoor D. Epidemiology of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM, editors. Osteoarthritis: Diagnosis and Medical/Surgical Management. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2007. pp. 3–26.
2. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: The ROAD study. Osteoarthritis Cartilage. 2009;17(9):1137–43.
3. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: The Research on Osteoarthritis/osteoporosis Against Disability (ROAD). J Bone Miner Metab. 2009;27(5):620–8.
4. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2007. August 2013. Available at http://www.mhlw.go.jp/toukei/list/20-19-1.html.
5. Linaker CH, Walker-Bone K, Palmer K, Cooper C. Frequency and impact of regional musculoskeletal disorders. Baillieres Clin Rheumatol. 1999;13(2):197–215.
6. Duncan R, Peat G, Thomas E, Hay E, McCall I, Croft P. Symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? Ann Rheum Dis. 2007;66(1):86–91.
7. Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. J Rheumatol. 2000;27(6):1513–7.
8. Neogi T, Felson D, Niu J, Nevitt M, Lewis CE, Aliabadi P, et al. Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. BMJ. 2009;339:b2844.
9. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988;15(12):1833–40.
10. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario and McMaster University osteoarthritis index in Asians with osteoarthritis in Singapore. Osteoarthritis Cartilage. 2001;9(5):440–6.
11. Woo J, Lau E, Lee P, Kwok T, Lau WC, Chan C, et al. Impact of osteoarthritis on quality of life in a Hong Kong Chinese population. J Rheumatol 2004;31(12):2433–8.
12. Dillon CF, Rasch EK, Gu Q, Hirsch R. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. J Rheumatol. 2006;33(11): 2271–9.
13. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, et al. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol. 2007;34(1):172–80.
14. Yoshida S, Aoyagi K, Felson DT, Aliabadi P, Shindo H, Takemoto T. Comparison of the prevalence of radiographic osteoarthritis of the knee and hand between Japan and the United States. J Rheumatol. 2002;29(7):1454–8.
15. Roubenoff R. Sarcopenia: a major modifiable cause of frailty in the elderly. J Nutr Health Aging. 2000;4(3):4140–2.
16. Bohannon RW. Hand-grip dynamometry predicts future outcomes in aging adults. J Geriatr Phys Ther. 2008;31(1):3–10.
17. Sirola J, Rikkonen T, Tuppurainen M, Jurvelin JS, Kroger H. Association of grip strength change with menopausal bone loss and related fractures: a population based follow-up study. Calcif Tissue Int. 2006;78(4):218–26.
18. Kerr A, Syddall HE, Cooper C, Turner GF, Briggs RS, Sayer AA. Does admission grip strength predict length of stay in hospitalised older patients? Age Ageing 2006;35(1):82–4.
19. Wind AE, Takken T, Helders PJ, Engelbert RH. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? Eur J Pediatr. 2010;169(3):281–7.
20. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: Research on osteoarthritis/osteoporosis against disability (ROAD) study. Int J Epidemiol. 2010;39: 988–95.
21. Kellgren JH, Lawrence JS (eds.). The Epidemiology of Chronic Rheumatism: Atlas of Standard Radiographs of Arthritis. Oxford, England: Blackwell Scientific; 1963.
22. Barr S, Bellamy N, Buchanan WW, Chalmers A, Ford PM, Kean WF, et al. A comparative study of signal versus aggregate methods of outcome measurement based on the WOMAC Osteoarthritis Index. J Rheumatol 1994;21(11):2106–12.
23. Hashimoto H, Hanyu T, Sledge CB, Lingard EA. Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities osteoarthritis index (WOMAC). J Orthop Sci. 2003; 8(3):288–93.
24. Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, et al. Incidence and risk factors for radiographic knee osteoarthritis and knee pain in Japanese men and women: a longitudinal population-based cohort study. Arthritis Rheum. 2012;64(5):1447–56.
25. Soni A, Kiran A, Hart DJ, Leyland KM, Goulston L, Cooper C, et al. Prevalence of reported knee pain over twelve years in a community-based cohort. Arthritis Rheum 2012;64(4):1145–52.
26. Saito T, Koshino T. Distribution of neuropeptides in synovium of the knee with osteoarthritis. Clin Orthop Relat Res. 2000;376: 172–82.

27. Bollet AJ. Edema of the bone marrow can cause pain in osteoarthritis and other diseases of bone and joints. Ann Intern Med. 2001;134(7):591–3.
28. Teichtahl AJ, Wluka AE, Morris ME, Davis SR, Cicuttini FM. The relationship between the knee adduction moment and knee pain in middle-aged women without radiographic osteoarthritis. J Rheumatol. 2006;33(9):1845–8.
29. Thorp LE, Sumner DR, Wimmer MA, Block JA. Relationship between pain and medial knee joint loading in mild radiographic knee osteoarthritis. Arthritis Rheum. 2007;57(7):1254–60.
30. Felson DT, Niu J, Guermazi A, Roemer F, Aliabadi P, Clancy M, et al. Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging. Arthritis Rheum. 2007;56(9):2986–92.

Supplementary material available online

Supplementary Tables 1–4.



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

Association of dietary intake with joint space narrowing and osteophytosis at the knee in Japanese men and women: the ROAD study

Shigeyuki Muraki, Toru Akune, Yoshio En-yo, Munehito Yoshida, Sakae Tanaka, Hiroshi Kawaguchi, Kozo Nakamura, Hiroyuki Oka & Noriko Yoshimura

To cite this article: Shigeyuki Muraki, Toru Akune, Yoshio En-yo, Munehito Yoshida, Sakae Tanaka, Hiroshi Kawaguchi, Kozo Nakamura, Hiroyuki Oka & Noriko Yoshimura (2014) Association of dietary intake with joint space narrowing and osteophytosis at the knee in Japanese men and women: the ROAD study, Modern Rheumatology, 24:2, 236-242, DOI: 10.3109/14397595.2013.854055

To link to this article: http://dx.doi.org/10.3109/14397595.2013.854055

Published online: 05 Mar 2014.

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://informahealthcare.com/mor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2014; 24(2): 236–242

DOI: 10.3109/14397595.2013.854055

informa
healthcare

ORIGINAL ARTICLE

Association of dietary intake with joint space narrowing and osteophytosis at the knee in Japanese men and women: the ROAD study

Shigeyuki Muraki[1], Toru Akune[1], Yoshio En-yo[2], Munehito Yoshida[2], Sakae Tanaka[3], Hiroshi Kawaguchi[3], Kozo Nakamura[4], Hiroyuki Oka[5], and Noriko Yoshimura[5]

[1]*Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan,* [2]*Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan,* [3]*Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan,* [4]*Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, Saitama, Japan, and* [5]*Department of Joint Disease Research, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*

Abstract

Objective. The objective of the present study is to identify dietary nutrients associated with joint space narrowing (JSN) and osteophytosis at the knee in a population-based cohort of the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.
Methods. From the baseline survey of the ROAD study, 827 participants (305 men and 522 women) in a rural cohort were analyzed. Dietary nutrient intakes for the last month were assessed by a self-administered brief diet history questionnaire. Minimum joint space width (mJSW) and osteophyte area (OPA) in the medial compartment of the knee were measured using a knee osteoarthritis (OA) computer-aided diagnostic system.
Results. In men, there were no associations of dietary nutrient intakes with mJSW or OPA. In women, vitamins K, B1, B2, B6, and C were associated with mJSW after adjustment for age, body mass index, and total energy ($p < 0.05$). Vitamins E, K, B1, B2, niacin, and B6 were significantly associated with OPA ($p < 0.05$) in women. Vitamins K, B and C may have a protective role against knee OA in women and might lead to disease-modifying treatments.
Conclusions. The present study revealed that low dietary intake of vitamins K, B, and C are associated with JSN and osteophytosis in women.

Keywords

Osteoarthritis, Knee, Diet, Cohort studies, Epidemiology

History

Received 25 October 2012
Accepted 25 February 2013
Published online 16 March 2013

Introduction

Knee osteoarthritis (OA), characterized by pathological features including joint space narrowing (JSN) and osteophytosis, is a major public health issue causing chronic pain and disability in the elderly in most developed countries [1]. The prevalence of radiographic knee OA is high in Japan [2], with 25,300,000 subjects aged 40 years and older estimated to experience radiographic knee OA [3]. According to the recent National Livelihood Survey of the Ministry of Health, Labour, and Welfare in Japan, OA is ranked fourth among diseases that cause disabilities that subsequently require support with activities of daily living [4]. Despite the urgent need for strategies for the prevention and treatment of this condition, there have been few established risk factors for knee OA except for age, female sex, obesity, previous injury, and occupational activities [5].

Current recommendations for OA include a combination of nonpharmacological interventions and pharmacological treatments [6]. However, considering that nonsteroidal anti-inflammatory drugs (NSAIDs), which may have serious adverse effects with long-term use, remain among the most widely prescribed drugs for OA [7], there is a need for safe and effective alternative strategies for prevention and treatment of this disease. Such strategies could come from dietary nutrition, because dietary factors are modifiable.

There have been several epidemiologic studies on the relationship between nutritional factors and OA [8–15]. Our previous study showed that dietary vitamin K intake was associated with the prevalence of knee OA [14], but disease was defined according to a categorical grade such as the Kellgren–Lawrence (KL) grade [16]. In the Framingham Study, the association of nutrition with JSN and osteophytosis was separately analyzed [8, 9, 12, 13] in Caucasians, but they were also defined by categorical grades. Categorical methods are statistically less powerful than continuous methods. Thus, the association between nutrition and knee OA might have been underestimated in previous studies.

To overcome these problems, joint space width or osteophyte area should be evaluated using a fully automatic system. To the best of our knowledge, there have been no population-based studies to separately measure joint space width or osteophyte area to clarify the association of dietary nutrient intake with JSN and osteophytosis. In the present study, we measured medial minimum joint space width (mJSW) and osteophyte area (OPA) at the knee in the large-scale population-based cohort study called Research on Osteoarthritis/osteoporosis Against Disability (ROAD). The

Correspondence to: Shigeyuki Muraki, Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. Tel: +81-3-58009178. Fax: +81-3-58009179. E-mail: murakis-ort@h.u-tokyo.ac.jp

purpose of the present study is to clarify which nutritional factors were associated with JSN and osteophytosis.

Materials and methods

Subjects

The ROAD study is a nationwide prospective study designed to establish epidemiologic indices for evaluation of clinical evidence for the development of a disease-modifying treatment for bone and joint diseases (OA and osteoporosis are the representative bone and joint diseases, respectively). It consists of population-based cohorts in three communities in Japan. A detailed profile of the ROAD study has been described elsewhere [2, 3, 17]; a brief summary is provided here. To date, we have completed the creation of a baseline database that includes clinical and genetic information for 3,040 subjects (1,061 men and 1,979 women) ranging in age from 23 to 95 years (mean, 70.3 years), who were recruited from resident registration listings in three communities: an urban region in Itabashi, Tokyo; a mountainous region in Hidakagawa, Wakayama; and a coastal region in Taiji, Wakayama.

Residents of these regions were recruited from the resident registration list of the relevant region. Participants in the urban region were recruited from a randomly selected cohort from the Itabashi-ward residents' registration database [18]. The participation rate was 75.6 %. Participants in mountainous and coastal regions were also recruited from the resident registration lists, and the participation rates in these two areas were 56.7 and 31.7 %, respectively. The inclusion criteria, apart from residence in the communities mentioned above, were the ability to (1) walk to the survey site, (2) report data, and (3) understand and sign an informed consent form. The baseline survey of the ROAD study was completed in 2006. All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology.

From the baseline data of 855 subjects aged ≥40 years in the mountainous cohort, we excluded 3 individuals who had undergone knee surgeries. In addition we excluded 18 individuals who had lateral knee OA, defined as being present when a knee had KL grade ≥2 and lateral JSN score ≥1 on a 0–3 scale according to the Osteoarthritis Research Society International (OARSI) atlas [19]. We also excluded 4 who did not complete questionnaires regarding dietary nutrition, and 3 whose radiographic conditions were insufficient for measuring JSN and osteophyte area. Thus, a total of 827 participants (305 men and 522 women) were analyzed in the present study.

Dietary assessment

For the dietary survey, we used a self-administered brief diet history questionnaire (BDHQ) and investigated dietary nutrient intakes for the previous month. A questionnaire was given to each participant with detailed explanation to fill out at home, and was reviewed by well-trained interviewers when the participants visited the clinic. The BDHQ is a 4-page, structured questionnaire that inquires about the consumption frequency of 56 food and beverage items, with specified serving sizes described in terms of a natural portion or the standard weight and volume measurement of servings commonly consumed in general Japanese populations. The BDHQ was developed based on a comprehensive (16-page) version of a validated self-administered diet history questionnaire [20], and is now widely used for dietary survey in Japan [14, 21]. Estimates of dietary intake for the 56 food and beverage items, energy, and selected nutrients were calculated using an ad hoc computer algorithm for the BDHQ, which was based on the Standard Tables of Food Composition in Japan. In the present study, dietary intake levels of total energy and 15 nutrient factors (animal protein, vegetable protein, animal fat, vegetable fat, carbohydrate, vitamin B1, 2, 6, and 12, niacin, vitamins C, D, E, K, and salt) were analyzed.

Radiographic assessment

All participants had radiographic examination of both knees using an anterior–posterior view with weight-bearing and foot map positioning. The beam was positioned parallel to the floor with no angle and aimed at the joint space. To visualize the joint space properly and to make the patella centralized over the lower end of the femur, we used fluoroscopic guidance with an anterior–posterior X-ray beam. The images were downloaded into digital imaging and communication in medicine (DICOM) format files. mJSW (mm) in the medial compartment and OPA (mm^2) at the medial tibia were measured by the KOACAD system, and the knee with lower mJSW was defined as the designated knee of each participant. The KOACAD system has been described in detail elsewhere [22–24], and is summarized here only briefly. The KOACAD system can quantify the major features of knee OA on standard radiographs and allows objective, accurate, simple, and easy assessment of the structural severity of knee OA in general clinical practice. This system was programmed to measure mJSW in the medial and lateral compartments and OPA at the medial tibia using digitized knee radiographs. Initially, correction for radiographic magnification was performed based on the image size of a rectangular metal plate. Next, to determine the region of interest (ROI), the center of the tibiofemoral joint was determined as follows: A vertical neighborhood difference filter that vertically scanned digital images to detect the margins of the tibial and femoral condyles was applied to identify points with high absolute values for differences of scale. Then, the center of all points was calculated and defined as the center of the tibiofemoral joint. Finally, a 480 × 200 pixel rectangle around the center was defined as the ROI. Within the ROI, the outline of the femoral condyle was designated as the upper rim of the joint space by vertical filtering with a 3 × 3 square neighborhood difference filter. Both ends of the upper rim were determined using a Canny filter to remove the noise associated with lines, and vertical lines from the ends were designated as the outside rims of the joint space. Outlines of anterior and posterior margins of the tibial plateau were drawn similarly to that of the femoral condyle, and the middle line between the two outlines was designated as the lower rim of the joint space (Fig. 1a). A straight regression line for the lower rim outline was then drawn, and the intersection of the lower rim outline and the regression line was designated as the inside rim. Medial and lateral joint space areas were determined as areas surrounded by the upper, lower, inside, and outside rims as defined above. Medial and lateral mJSWs were further determined as the minimum vertical distances in the respective joint space area (Fig. 1b). To measure the OPA, medial and lateral outlines of the femur and tibia were drawn. Inflection points for these outlines were then calculated. The medial outline of the tibia from the inflection point was drawn upward to the joint level, and the area that was medially prominent over the smoothly extended outline was designated as the OPA (Fig. 1c). We examined the reproducibility of mJSW and OPA measured on radiographs taken at 2-week intervals for 20 individuals; the reproducibility of both mJSW and OPA were high [intraclass correlation coefficient (ICC) = 0.86 and 0.99, respectively] [22]. In addition, we measured mJSW and OPA by KOACAD more than twice on 1979 radiographs, and confirmed that all parameters were unchanged independent of observer or time measured (all ICC = 1.0) [22]. We have previously published reference values of joint space width and osteophyte area by gender and age strata in Japan using the KOACAD system [25].

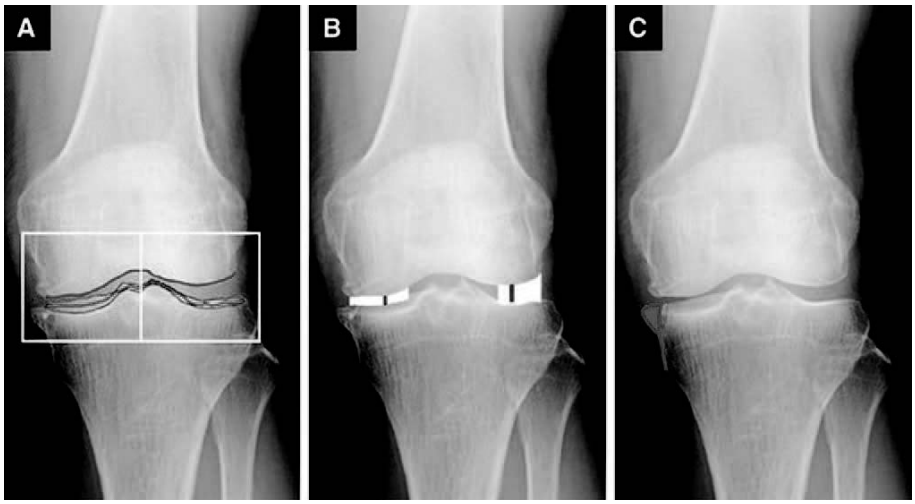


Fig 1. Schema of image processing by KOACAD (cited from Ref. [28]). (a) Outlines of anterior and posterior margins of the tibial plateau. The *middle line* between the two outlines is defined as the lower rim of the joint space. (b) Medial and lateral minimum joint space widths were defined as the minimum vertical distances in the joint space area. (c) Osteophyte area (*red area*) that is medially prominent over the smoothly extended outline of the tibia

Statistical analysis

Differences in age, height, weight, and body mass index (BMI) were examined by nonpaired Student's *t* test. mJSW, OPA, total energy, and dietary nutrient intakes between men and women were examined by Wilcoxon rank-sum test. The distribution of mJSW, OPA, total energy, and dietary nutrient intakes were not normal, thus we applied log transformation to these variables, and multiple regression analysis after adjustment for age, BMI, gender, and total energy was used to determine the association of dietary nutrient intakes with mJSW and OPA in the overall population. Furthermore, multiple regression analysis after adjustment for age, BMI, and total energy was used to determine the association of dietary nutrient intakes with mJSW and OPA in men and women. Data analyses were performed using SAS version 9.0 (SAS Institute Inc., Cary, NC). *p*-Value <0.05 was considered significant.

Results

Characteristics of 827 participants are presented in Table 1. There were no significant differences in BMI between men and women. mJSW was significantly wider in men than women, and OPA was significantly smaller in men than women. Total energy and almost all of dietary nutrient intakes except for vitamins K and C were significantly higher in men than women ($p < 0.01$), whereas vitamin C intake was significantly lower in men than women ($p < 0.0001$) (Table 2). Vitamin K intake was not significantly different between men and women ($p = 0.07$).

Table 1. Characteristics of participants

	Overall	Men	Women	*p* Value
No. of participants	827	305	522	
Age (years)	69.2 ± 9.3	69.6 ± 8.7	68.9 ± 9.6	0.29
Height (cm)	163.0 ± 9.2	161.3 ± 6.7	148.1 ± 6.6	<0.0001
Weight (kg)	54.0 ± 10.2	60.0 ± 10.2	50.5 ± 8.5	<0.0001
BMI (kg/m^2)	23.0 ± 3.2	23.0 ± 3.0	23.0 ± 3.4	0.86
mJSW (mm)	2.43 ± 1.11	2.91 ± 1.01	2.15 ± 1.07	<0.0001
OPA (mm^2)	3.72 ± 8.33	1.72 ± 4.20	4.88 ± 9.79	<0.0001

Data are mean ± standard deviation (SD). Nonpaired Student's *t* test was used to compare age, height, and BMI between men and women. Wilcoxon rank-sum test was used to compare mJSW and OPA between men and women

BMI body mass index, *mJSW* minimum joint space width, *OPA* osteophyte area

We next analyzed the association of dietary nutrient intakes with mJSW and OPA. Overall, after adjustment for age, BMI, gender, and total energy, mJSW was not associated with vitamins D, E, B1 or niacin, but was significantly associated with vitamins K ($R = 0.344$, $p = 0.03$), B2 ($R = 0.343$, $p = 0.04$), and C ($R = 0.345$, $p = 0.02$) (Table 3). OPA was not significantly associated with vitamins D, E, K, B12, C or niacin, but was significantly associated with vitamins B1 ($R = 0.421$, $p = 0.03$), B2 ($R = 0.421$, $p = 0.03$), and B6 ($R = 0.422$, $p = 0.02$) (Table 3). When analyzed in men and women separately, in men, multiple regression analysis after adjustment for age, BMI, and total energy showed that mJSW and OPA were not significantly associated with any nutritional factors (Table 4). In contrast, in women, mJSW was significantly associated with vitamins K ($R = 0.283$, $p = 0.01$), B1 ($R = 0.271$, $p = 0.04$), B2 ($R = 0.270$, $p = 0.04$), B6 ($R = 0.273$, $p = 0.01$), and C ($R = 0.281$, $p = 0.01$) (Table 5), while OPA was significantly associated with vitamins E ($R = 0.426$, $p = 0.04$), K ($R = 0.427$, $p = 0.03$), B1 ($R = 0.436$,

Table 2. Dietary nutrient intakes in men and women

	Overall	Men	Women
Total energy, MJ/day	7.6 (6.3–9.3)	9.5 (8.1–12.1)	6.9* (6.0–7.9)
Dietary nutrients			
Vitamin D, μg/day	17.7 (11.5–25.8)	20.7 (13.3–30.5)	16.4* (10.7–24.2)
Vitamin E, mgα-TE/day	6.9 (5.4–8.8)	7.4 (5.6–9.6)	6.7* (5.3–8.3)
Vitamin K, μg/day	211.0 (146.6–287.9)	224.4 (150.2–313.5)	202.9 (145.3–281.0)
Vitamin B1, mg/day	0.71 (0.58–0.86)	0.79 (0.64–0.97)	0.67* (0.56–0.80)
Vitamin B2, mg/day	0.97 (0.76–1.19)	1.07 (0.82–1.34)	0.92* (0.73–1.12)
Niacin, mgNE/day	14.9 (11.6–19.2)	17.9 (13.9–22.7)	13.6* (10.4–17.1)
Vitamin B6, mg/day	1.1 (0.9–1.4)	1.3 (1.0–1.6)	1.03* (0.86–1.26)
Vitamin B12, μg/day	9.8 (6.8–13.5)	11.0 (7.7–15.8)	8.8* (6.3–12.0)
Vitamin C, mg/day	101.7 (78.3–133.4)	94.0 (71.7–122.0)	108.1* (82.6–137.3)

Values are median (interquartile range)

TE tocopherol equivalent, *NE* niacin equivalent

*$p < 0.01$ versus men by Wilcoxon rank-sum test

Table 3. Association of dietary nutrient intakes with mJSW and OPA overall

	mJSW (mm)			OPA (mm^2)		
	Regression coefficient	95 % CI	*p*-Value	Regression coefficient	95 % CI	*p* Value
Vitamin D, μg/day	0.006	−0.04 to 0.06	0.8044	−0.03	−0.09 to 0.02	0.2000
Vitamin E, mgα-TE/day	0.01	−0.08 to 0.10	0.7613	−0.08	−0.17 to 0.02	0.1114
Vitamin K, μg/day	0.06	0.006 to 0.11*	0.0309	−0.05	−0.11 to 0.004	0.0665
Vitamin B1, mg/day	0.09	−0.05 to 0.23	0.2058	−0.17	−0.32 to 0.02*	0.0271
Vitamin B2, mg/day	0.10	0.004 to 0.20*	0.0418	−0.12	−0.22 to 0.01*	0.0254
Niacin, mgNE/day	0.02	−0.08 to 0.13	0.6422	−0.09	−0.20 to 0.01	0.0877
Vitamin B6, mg/day	0.12	−0.001 to 0.24	0.0526	−0.15	−0.28 to 0.03*	0.0164
Vitamin B12, μg/day	0.04	−0.02 to 0.09	0.2066	−0.03	−0.09 to 0.02	0.2515
Vitamin C, mg/day	0.09	0.01 to 0.16*	0.0179	−0.04	−0.12 to 0.03	0.2640

Log transformation was applied to variables, and multiple regression analysis after adjustment for age, body mass index, gender, and total energy was used to determine the association of nutritional factors with mJSW and OPA

mJSW minimum joint space width, *OPA* osteophyte area, *TE* tocopherol equivalent, *NE* niacin equivalent, *CI* confidence interval

$p = 0.002$), B2 ($R = 0.435$, $p = 0.003$), niacin ($R = 0.428$, $p = 0.02$), and B6 ($R = 0.433$, $p = 0.01$) (Table 5).

Discussion

This is the first population-based cohort study of the relationship between dietary nutrient intakes and JSN and osteophytosis separately in Japanese men and women. In the overall population, vitamins K, B2, and C were significantly associated with mJSW, while vitamins B1, B2, and B6 were significantly associated with OPA. When analyzed in men and women separately, we observed that there were no associations of dietary nutrient intakes with mJSW or OPA in men. In contrast, in women, vitamins K, B1, B2, and B6 were associated with both mJSW and OPA. Vitamin C was associated with mJSW, but not with OPA. Previous studies have already shown that vitamins K and C were associated with knee OA; however, the knee OA was defined by KL grade or other categorical methods in almost all studies [8–15]. KL grade is the most conventional system to grade radiographic severity of knee OA, but in this categorical system, JSN and osteophyte formation are not assessed separately, thus one cannot clarify whether osteophytosis and JSN have distinct risk factors. In addition, a recent cross-sectional study showed that osteophytosis was unrelated to JSN on plain radiographs [26]. Furthermore, our study on an experimental mouse model for OA identified a cartilage-specific molecule, carminerin, that regulates osteophytosis without affecting joint cartilage destruction during OA progression [27, 28]. In addition, there were distinct effects on quality of life (QOL) for JSN and osteophytosis [26]. Such accumulating evidence indicates that JSN and osteophytosis may have distinct etiologic mechanisms and their progression may be neither constant nor proportional. Thus, to examine factors associated with knee OA, these two OA features should be separately assessed. Furthermore, because categorical methods are statistically less powerful than continuous methods, the association between nutrition and knee OA might have been underestimated in previous studies. This study is the first to report that vitamins K, B1, B2, and B6 are significantly associated with both mJSW and OPA, and that vitamin C is significantly associated with mJSW in women. The association of dietary factors with knee OA may be weaker than for gender or obesity, but they are easily modifiable; therefore, these results may contribute to prevent incidence or progression of knee OA, although it is not completely clear what modifications of vitamin intake would be required to achieve clinically meaningful change in mJSW and OPA.

Vitamin K includes vitamin K1, or phylloquinone, which is contained in green leafy vegetables, and vitamin K2, or menaquinone, which is synthesized by bacteria and abundantly contained in a traditional Japanese fermented soybean food called *natto* [29]. Our previous study showed that dietary vitamin K intake was inversely associated with prevalence of knee OA defined by KL grade [14]. However, because of the different etiology that

Table 4. Association of dietary nutrient intakes with mJSW and OPA in men

	mJSW (mm)			OPA (mm^2)		
	Regression coefficient	95 % CI	*p* Value	Regression coefficient	95 % CI	*p* Value
Vitamin D, μg/day	−0.02	−0.10 to 0.06	0.5804	0.04	−0.03 to 0.11	0.2710
Vitamin E, mgα-TE/day	−0.01	−0.14 to 0.11	0.8501	0.03	−0.09 to 0.14	0.6567
Vitamin K, μg/day	0.02	−0.06 to 0.09	0.6626	−0.01	−0.08 to 0.06	0.7939
Vitamin B1, mg/day	−0.01	−0.21 to 0.19	0.8995	0.08	−0.11 to 0.26	0.4275
Vitamin B2, mg/day	0.07	−0.08 to 0.22	0.3515	0.05	−0.09 to 0.19	0.4772
Niacin, mgNE/day	−0.03	−0.18 to 0.12	0.7149	0.06	−0.08 to 0.20	0.4127
Vitamin B6, mg/day	0.04	−0.13 to 0.22	0.6214	−0.005	−0.17 to 0.16	0.9554
Vitamin B12, μg/day	−0.004	−0.09 to 0.09	0.9345	0.06	−0.03 to 0.14	0.1816
Vitamin C, mg/day	0.03	−0.03 to 0.14	0.5079	0.01	−0.08 to 0.11	0.8113

Log transformation was applied to variables, and multiple regression analysis after adjustment for age, body mass index, and total energy was used to determine the association of nutritional factors with mJSW and OPA

mJSW minimum joint space width, *OPA* osteophyte area, *TE* tocopherol equivalent, *NE* niacin equivalent, *CI* confidence interval

Table 5. Association of dietary nutrient intakes with mJSW and OPA in women

	mJSW (mm)			OPA (mm^2)		
	Regression coefficient	95 % CI	*p* Value	Regression coefficient	95 % CI	*p* Value
Vitamin D, μg/day	0.03	−0.03 to 0.09	0.3550	−0.07	−0.14 to 0.004	0.0631
Vitamin E, mgα-TE/day	0.05	−0.08 to 0.18	0.4234	−0.15	−0.29 to −0.008*	0.0383
Vitamin K, μg/day	0.11	0.03 to 0.19*	0.0062	−0.10	−0.18 to −0.009*	0.0302
Vitamin B1, mg/day	0.21	0.01 to 0.41*	0.0366	−0.35	−0.56 to −0.13*	0.0020
Vitamin B2, mg/day	0.13	0.006 to 0.26*	0.0411	−0.22	−0.37 to −0.08*	0.0025
Niacin, mgNE/day	0.08	−0.06 to 0.21	0.2819	−0.18	−0.33 to −0.03*	0.0195
Vitamin B6, mg/day	0.18	0.02 to 0.34*	0.0261	−0.25	−0.42 to −0.07*	0.0053
Vitamin B12, μg/day	0.07	−0.005 to 0.14	0.0679	−0.07	−0.16 to 0.006	0.0699
Vitamin C, mg/day	0.13	0.04 to 0.23*	0.0077	−0.09	−0.20 to 0.02	0.1139

Log transformation was applied to variables, and multiple regression analysis after adjustment for age, body mass index, and total energy was used to determine the association of nutritional factors with mJSW and OPA

mJSW minimum joint space width, *OPA* osteophyte area, *TE* tocopherol equivalent, *NE* niacin equivalent, *CI* confidence interval

may exist between JSN and osteophytosis, these two OA features should be assessed separately to examine factors associated with knee OA. However, the association of these two features with vitamin K cannot be separately analyzed by KL grade. The Framingham Study showed that plasma levels of phylloquinone were inversely associated with osteophytosis in the knee [12], but no population-based study has determined the association of dietary vitamin K intake with mJSW width and OPA separately. In the present study, vitamin K was associated with both JSN and osteophytosis in women, although the results for vitamin K were of borderline significance after adjusting for additional potential confounders, particularly regarding OPA. Several basic studies have shown that vitamin K plays an important role in cartilage metabolism, as an inhibitor of extracellular matrix calcification as well as a promoter of cell survival and proliferation [30–38]. In addition, warfarin, a vitamin K-antagonist anticoagulant, is known to cause warfarin embryopathy characterized by abnormal calcification and decreased growth of cartilage [37, 38]. Habitual low dietary vitamin K intake may exert an inhibitory effect on the vitamin K-dependent MGP and Gas6 functions and modulate the pathogenesis of OA by influencing the process of osteophytosis and cartilage destruction.

Several previous studies have shown that vitamin C intake was inversely associated with knee OA [9, 15], but no population-based study has analyzed the association of vitamin C intake with mJSW and OPA at the same time. In the present study, vitamin C was associated with narrower mJSW in women, but not with OPA. This finding may indicate that vitamin C intake is more strongly associated with JSN than with osteophytosis in women. Damage caused by free radicals has long been thought to be pathogenic, and free radicals play an important role in the progression of many chronic diseases, including OA [9, 11, 39–42]. Vitamin C is an antioxidant, which may partly explain the effect of vitamin C on JSN. This may lead to the logical possibility of using vitamin C supplementation for primary prevention or as a therapeutic intervention for OA.

There have been no studies regarding the association of dietary vitamin B intake with knee OA. In the present study, we found that vitamins B1, B2, and B6 were significantly associated with mJSW in women. Vitamin B is closely involved in the metabolism of homocysteine [43], which has recently been seen to play a role in osteoporosis-related bone damage, and may be linked to its involvement in collagen formation. Homocysteine inhibits the synthesis of insoluble collagen fibrils in vitro by interfering with normal cross-linking [44]. From the perspective of cartilage homeostasis, these changes in matrix organization interfere with chondrocyte-mediated mineralization, potentially altering the function and properties of calcified cartilage [45]. This may be due to homocysteine-mediated inhibition of lysil oxidase, which catalyzes the cross-linking of collagen molecules, a function necessary for its mineralization in bone tissue [46].

In the present study, we found gender differences regarding the association of dietary nutrient intakes with mJSW and OPA. In women, vitamins B and K were significantly associated with both mJSW and OPA, and vitamin C was significantly associated with mJSW, whereas in men, no dietary factors were significantly associated with mJSW or OPA. This difference may be partly explained by muscle strength in men. Because men are known to have greater muscle strength than women at all ages, and muscle strength has a protective effect on knee OA [47–49], it might be that the greater muscle strength obscures the effects of dietary nutrient intakes on knees in men.

There are several limitations to the present study. First, this was a cross-sectional study of baseline data, and thus no causal relationship can be determined. Second, in the present study, we used self-reported measures for dietary assessments; these measurements are prone to bias and measurement error. In addition, the dietary survey in this study investigated dietary habits only for the previous month, which did not necessarily reflect a long habit of several years, despite the fact that OA is a slowly progressing chronic disease. This dietary survey also investigated whether participants had changed their dietary habits. Those who answered "yes" accounted for 9.6 %, whereas 90.4 % of participants answered that they had not changed their dietary habits. Although it is likely that dietary habits in middle-aged and elderly people are usually quite different from those in children and young adults, there is a possibility that most participants in this study had not changed their dietary habits for several years or for a longer time, which may have affected the disease process of OA. Furthermore, the dietary survey in the present study was conducted from autumn to winter although there are four seasons in Japan and diets may vary with the season. Therefore, the present study could suffer from some bias for the effect of season on the nutritional quality of diets. Third, nutritional factors cannot be assumed to be joint location specific, and osteophytes may even be more pronounced in the contralateral tibiofemoral compartment [50]; however, at present, the KOACAD system can only measure medial osteophytes at the tibia. We are now developing a KOACAD system to measure osteophytes at other sites; thus, we may be able to clarify the association between osteophytes at other sites and QOL in the near future. Finally, we clarified the association of vitamins B, C, and K with mJSW and OPA, but did not determine what changes in intake of these vitamins would be needed to achieve clinically meaningful change in mJSW and OPA, because we

have not yet clarified what changes in mJSW and OPA are clinically meaningful. In addition, this is a cross-sectional study, thus causal relationships of vitamins B, C, and K with mJSW and OPA cannot be clarified.

In conclusion, the present cross-sectional study using a population-based cohort revealed that low dietary intakes of vitamins K, B1, B2, and B6 are associated with both JSN and osteophytosis in women. Vitamin C intake was associated with JSN in women, but not with osteophytosis. Further studies, along with longitudinal data from the ROAD study, will elucidate the environmental background of OA and help clarify clinical evidence regarding the development of disease-modifying treatments.

Acknowledgments

The present study was supported by a Grant-in-Aid for Scientific Research (B20390182, C20591737, C20591774), for Young Scientists (A18689031), and for Exploratory Research (19659305) from the Japanese Ministry of Education, Culture, Sports, Science, and Technology, H17-Men-eki-009, H18-Choujyu-037, and H20-Choujyu-009 from the Ministry of Health, Labour, and Welfare, and Research Aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1); and Grant No. 166 from the Japan Orthopaedics and Traumatology Foundation. The authors thank Mrs. Tomoko Takijiri and other members of the Public Office in Hidakagawa Town; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, and other members of the Public Office in Taiji Town, for their assistance in the location and scheduling of participants for examinations.

Conflict of interest

None.

References

1. Sharma L, Kapoor D. Epidemiology of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM. Osteoarthritis: diagnosis and medical/surgical management. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 3–26.
2. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthr Cartil. 2009;17:1137–43.
3. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: the Research on Osteoarthritis/osteoporosis Against Disability (ROAD). J Bone Miner Metab. 2009;27:620–8.
4. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2007. http://www.mhlw.go.jp/toukei/list/20-19-1.html. Accessed 1 Oct 2012.
5. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. Osteoarthr Cartil. 2010;18:24–33.
6. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. Osteoarthr Cartil. 2007;15:981–1000.
7. Simon LS, Strand V. The pharmacologic treatment of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM. Osteoarthritis: diagnosis and medical/surgical management. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 267–86.
8. McAlindon TE, Felson DT, Zhang Y, Hannan MT, Aliabadi P, Weissman B, et al. Relation of dietary intake and serum levels of vitamin D to progression of osteoarthritis of the knee among participants in the Framingham Study. Ann Intern Med. 1996;125:353–9.
9. McAlindon TE, Jacques P, Zhang Y, Hannan MT, Aliabadi P, Weissman B, et al. Do antioxidant micronutrients protect against the development and progression of knee osteoarthritis?. Arthritis Rheum. 1996;39:648–56.
10. McAlindon TE, Biggee BA. Nutritional factors and osteoarthritis: recent developments. Curr Opin Rheumatol. 2005;17:647–52.
11. Ameye LG, Chee WS. Osteoarthritis and nutrition. From nutraceuticals to functional foods: a systematic review of the scientific evidence. Arthritis Res Ther. 2006;8:R127.
12. Neogi T, Booth SL, Zhang YQ, Jacques PF, Terkeltaub R, Aliabadi P, et al. Low vitamin K status is associated with osteoarthritis in the hand and knee. Arthritis Rheum. 2006;54:1255–61.
13. Felson DT, Niu J, Clancy M, Aliabadi P, Sack B, Guermazi A, et al. Low levels of vitamin D and worsening of knee osteoarthritis: results of two longitudinal studies. Arthritis Rheum. 2007;56:129–36.
14. Oka H, Akune T, Muraki S, En-yo Y, Yoshida M, Saika A, et al. Low dietary vitamin K intake is associated with radiographic knee osteoarthritis in the Japanese elderly: dietary survey in a population-based cohort of the ROAD study. J Orthop Sci. 2009;14:687–92.
15. Peregoy J, Wilder FV. The effects of vitamin C supplementation on incident and progressive knee osteoarthritis: a longitudinal study. Pub Health Nutr. 2010;14:709–15.
16. Kellgren JH, Lawrence JS. The epidemiology of chronic rheumatism: atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific; 1963.
17. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: Research on Osteoarthritis/osteoporosis Against Disability (ROAD) Study. Int J Epidemiol. 2010;39:988–95.
18. Shimada H, Lord SR, Yoshida H, Kim H, Suzuki T. Predictors of cessation of regular leisure-time physical activity in community-dwelling elderly people. Gerontology. 2007;53:293–7.
19. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthr Cartil. 2007;15 Suppl A:A1–56.
20. Sasaki S, Yanagibori R, Amano K. Self-administered diet history questionnaire developed for health education: a relative validation of the test-version by comparison with 3-day diet record in women. J Epidemiol. 1998;8:203–15.
21. Murakami K, Mizoue T, Sasaki S, Ohta M, Sato M, Matsushita Y, et al. Dietary intake of folate, other B vitamins, and omega-3 polyunsaturated fatty acids in relation to depressive symptoms in Japanese adults. Nutrition. 2008;24:140–7.
22. Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H, et al. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthr Cartil. 2008;16:1300–6.
23. Muraki S, Oka H, Akune T, En-yo Y, Yoshida M, Nakamura K, et al. Association of occupational activity with joint space narrowing and osteophytosis in the medial compartment of the knee: the ROAD study. Osteoarthr Cartil. 2011;19:840–6.
24. Muraki S, Oka H, Akune T, En-yo Y, Yoshida M, Suzuki T, et al. Independent association of joint space narrowing and osteophyte formation at the knee with health-related quality of life in Japan: a cross-sectional study. Arthritis Rheum. 2011;63:3859–64.
25. Oka H, Muraki S, Akune T, Nakamura K, Kawaguchi H, Yoshimura N. Normal and threshold values of joint space width, joint space area, osteophyte area and fibro-tibial angle using a computer-assisted measuring system (KOACAD) to evaluate knee osteoarthritis: the ROAD study. J Orthop Sci. 2010;15:781–9.
26. Muraki S, Oka H, Akune T, En-yo Y, Yoshida M, Suzuki T, et al. Independent association of joint space narrowing and osteophyte formation at the knee with health-related quality of life in Japan: a cross-sectional study. Arthritis Rheum. 2011;63:3859–64.
27. Yamada T, Kawano H, Koshizuka Y, Fukuda T, Yoshimura K, Kamekura S, et al. Carminerin contributes to chondrocyte calcification during endochondral ossification. Nat Med. 2006;12:665–70.
28. Kamekura S, Kawasaki Y, Hoshi K, Shimoaka T, Chikuda H, Maruyama Z, et al. Contribution of runt-related transcription factor 2 to the pathogenesis of osteoarthritis in mice after induction of knee joint instability. Arthritis Rheum. 2006;54:2462–70.
29. Kaneki M, Hodges SJ, Hosoi T, Fujiwara S, Lyons A, Crean SJ, et al. Japanese fermented soybean food as the major determinant of the large geographic difference in circulating levels of vitamin K2: possible implications for hip-fracture risk. Nutrition. 2001;17:315–21.
30. Furie B, Bouchard BA, Furie BC. Vitamin K-dependent biosynthesis of gamma-carboxyglutamic acid. Blood. 1999;93:1798–808.
31. Schneider C, King RM, Philipson L. Genes specifically expressed at growth arrest of mammalian cells. Cell. 1988;54:787–93.
32. Loeser RF, Varnum BC, Carlson CS, Goldring MB, Liu ET, Sadiev S, et al. Human chondrocyte expression of growth-arrest-specific gene 6 and the tyrosine kinase receptor axl: potential role in autocrine signaling in cartilage. Arthritis Rheum. 1997;40:1455–65.
33. Hafizi S, Dahlbäck B. Gas6 and protein S. Vitamin K-dependent ligands for the Axl receptor tyrosine kinase subfamily. FEBS J. 2006;273:5231–44.

34. Luo G, D'Souza R, Hogue D, Karsenty G. The matrix Gla protein gene is a marker of the chondrogenesis cell lineage during mouse development. J Bone Miner Res. 1995;10:325–34.
35. Munroe PB, Olgunturk RO, Fryns JP, Van Maldergem L, Ziereisen F, Yuksel B, et al. Mutations in the gene encoding the human matrix Gla protein cause Keutel syndrome. Nat Genet. 1999;21:142–4.
36. Luo G, Ducy P, McKee MD, Pinero GJ, Loyer E, Behringer RR, et al. Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. Nature. 1997;386:78–81.
37. Pauli RM, Lian JB, Mosher DF, Suttie JW. Association of congenital deficiency of multiple vitamin K-dependent coagulation factors and the phenotype of the warfarin embryopathy: clues to the mechanism of teratogenicity of coumarin derivatives. Am J Hum Genet. 1987;41:566–83.
38. Howe AM, Lipson AH, de Silva M, Ouvrier R, Webster WS. Severe cervical dysplasia and nasal cartilage calcification following prenatal warfarin exposure. Am J Med Genet. 1997;71:391–6.
39. Felson D, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum. 1998;41:1343–55.
40. Canter PH, Wider B, Ernst E. The antioxidant vitamins A, C, E and selenium in the treatment of arthritis: a systematic review of randomized clinical trials. Rheumatology. 2007;46:1223–33.
41. Kurz B, Jost B, Schünke M. Dietary vitamins and selenium diminish the development of mechanically induced osteoarthritis and increase the expression of antioxidative enzymes in the knee joint of STR/1N mice. Osteoarthr Cartil. 2002;10:119–26.
42. Sowers M, Lachance L. Vitamins and arthritis—the roles of vitamins A, C, D, and E. Rheum Dis Clin North Am. 1999;25:315–32.
43. Selhub J, Jacques PF, Wilson PW, Rush D, Rosenberg IH. Vitamin status and intake as primary determinants of homocysteinemia in an elderly population. JAMA. 1993;270:2693–8.
44. Kang AH, Trelstad RL. A collagen defect in homocystinuria. J Clin Invest. 1973;52:2571–8.
45. Khan M, Yamauchi M, Srisawasdi S, Stiner D, Doty S, Paschalis EP, et al. Homocysteine decreases chondrocyte-mediated matrix mineralization in differentiating chick limb-bud mesenchymal cell micro-mass cultures. Bone. 2001;28:387–98.
46. Lubec B, Fang-Kircher S, Lubec T, Blom HJ, Boers GH. Evidence for McKusick's hypothesis of deficient collagen cross-linking in patients with homocystinuria. Biochim Biophys Acta. 1996;1315: 159–62.
47. Sinaki M, Nwaogwugwu NC, Phillips BE, Mokri MP. Effect of gender, age, and anthropometry on axial and appendicular muscle strength. Am J Phys Med Rehabil. 2001;80:330–8.
48. McAlindon TE, Cooper C, Kirwan JR, Dieppe PA. Determinants of disability in osteoarthritis of the knee. Ann Rheum Dis. 1993;52: 258–62.
49. O'Reilly SC, Jones A, Muir KR, Doherty M. Quadriceps weakness in knee osteoarthritis: the effect on pain and disability. Ann Rheum Dis. 1998;57:588–94.
50. Felson DT, Gale DR, Elon Gale M, Niu J, Hunter DJ, Goggins J, et al. Osteophytes and progression of knee osteoarthritis. Rheumatology (Oxford). 2005;44:100–4.

Taylor & Francis
Taylor & Francis Group

Journal of Manual & Manipulative Therapy

ISSN: 1066-9817 (Print) 2042-6186 (Online) Journal homepage: http://www.tandfonline.com/loi/yjmt20

Can standing back extension exercise improve or prevent low back pain in Japanese care workers?

Ko Matsudaira, Miho Hiroe, Masatomo Kikkawa, Takayuki Sawada, Mari Suzuki, Tatsuya Isomura, Hiroyuki Oka, Kou Hiroe & Ken Hiroe

To cite this article: Ko Matsudaira, Miho Hiroe, Masatomo Kikkawa, Takayuki Sawada, Mari Suzuki, Tatsuya Isomura, Hiroyuki Oka, Kou Hiroe & Ken Hiroe (2015) Can standing back extension exercise improve or prevent low back pain in Japanese care workers?, Journal of Manual & Manipulative Therapy, 23:4, 205-209, DOI: 10.1179/2042618614Y.0000000100

To link to this article: http://dx.doi.org/10.1179/2042618614Y.0000000100

Can standing back extension exercise improve or prevent low back pain in Japanese care workers?

Ko Matsudaira[1,2], Miho Hiroe[3], Masatomo Kikkawa[3], Takayuki Sawada[4], Mari Suzuki[4], Tatsuya Isomura[4,5], Hiroyuki Oka[6], Kou Hiroe[3], Ken Hiroe[3]

[1]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo Hospital, Bunkyo-ku, Japan, [2]Clinical Research Center for Occupational Musculoskeletal Disorders, Kanto Rosai Hospital, Nakahara-ku, Kawasaki, Kanagawa, Japan, [3]Kohoen Social Community Service, Yonago, Tottori, Japan, [4]Clinical Study Support, Inc., Nagoya Life Science Incubator, Chikusa-ku, Aichi, Japan, [5]Institute of Medical Science, Tokyo Medical University, Shinjuku-ku, Japan, [6]Department of Joint Disease Research, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Japan

Background: We suggested a standing back extension exercise 'One Stretch' based on the McKenzie method, to examine the ability to improve or prevent low back pain (LBP) in Japanese care workers.
Methods: We conducted a single-center, non-randomized, controlled study in Japan. Care workers in an intervention group received an exercise manual and a 30-minute seminar on LBP and were encouraged with a group approach, while care workers in a control group were given only the manual. All care workers answered questionnaires at the baseline and end of a 1-year study period. The subjective improvement of LBP and compliance with the exercise were evaluated.
Results: In all, 64 workers in the intervention group and 72 in the control group participated in this study. More care workers in the intervention group exercised regularly and improved or prevented LBP than in the control group ($P=0.003$ and $P<0.0001$, respectively). In the intervention group, none had a first medical consultation or were absent from disability for LBP by the end of the study period.
Conclusion: The exercise 'One Stretch' would be effective to improve or prevent LBP in care workers. Our group approach would lead to better compliance with the exercise.

Keywords: Low back pain, Standing back extension, McKenzie method, Care worker, Population strategy, Prevention

Introduction

Low back pain (LBP) is a major health problem, particularly in industrialized countries, and has affected people's life and social economy in various ways. The Global Burden of Disease Study indicated 'low back pain is one of the leading specific causes of years lived with disability (YLD)',[1] and about 85–90% of LBP has been classified as non-specific LBP.[2–4] Low back pain-associated disability results in loss of work and huge economic impact with substantial direct and indirect social costs.[5–7]

In Japan, as in other industrialized countries, many people suffer from LBP. Recently, a lifetime LBP prevalence of 83% and a 4-week prevalence of 36% were reported.[8] Additionally, LBP was the fifth most common reason for medical consultation among Japanese outpatients,[9] and especially in the health care industry, an increasing number of care workers left the job due to work-related LBP.[10]

Some researchers revealed that physical activity at work, such as lifting and rather keeping forward flexion, sustained forward bending, can be associated with increased back symptoms, further aggravating pain (so-called back injuries).[11] In fact, frequent lifting during working hours greatly impacts non-specific LBP in Japanese workers.[12]

To deal with the socioeconomic problem of LBP, it is important to prevent LBP from developing in people without symptoms. Physical exercises have been recommended in the prevention of LBP, while there is insufficient evidence against any specific type or intensity of exercise.[13] McKenzie, who introduced a subgroup classification method of LBP, recommends extension exercise because posterior displacement of

Correspondence to: Ko Matsudaira, Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo 133-8655, Japan. Email: kohart801@gmail.com

DOI 10.1179/2042618614Y.0000000100

the nucleus by the exercise eliminates or abolishes LBP.[14] The McKenzie method is a system that classifies patients into one of the specific subgroups primarily based on symptomatic and mechanical responses to mechanical loadings. Among the LBP population, the largest subgroup where LBP is improved in a short period of time is by back extension loading strategy. Its theoretical explanation is based on the disk model in which posterior displacement of the nucleus can be reduced by deliberate extension loading strategy. This reduction of the displaced nucleus may result in decreasing or abolishing LBP.

In this study, we used a simple daily standing back extension exercise 'One Stretch', to evaluate the efficacy of this exercise in care workers at risk of developing and aggravating LBP.

Subjects and Methods

Study population

This study was conducted at a health care facility for the elderly, Numbu Kohoen, Japan. Eligible participants were Japanese care workers who worked there on the first and second floors and supported the elderly in need of care. We excluded the workers who had difficulties in participating due to medical causes (e.g. spinal stenosis, rheumatoid arthritis, and ankylosing spondylitis) or other personal reasons. Written informed consent was obtained from all participants.

This study was approved by the medical/ethics review board of Kanto Rosai Hospital. We registered our study (ID: UMIN000004473) in the University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR).

Study design

This was a single-center, non-randomized, controlled study. Participants who worked on the first floor were assigned to the control group, and those on the second floor to the intervention group. We provided an exercise manual for all participants and a 30-minute seminar only for the intervention group. In the exercise manual, we described how to do a standing back extension exercise 'One stretch' (Fig. 1). This exercise is an active extension of the back used as a common technique in physical therapy, and is based on the theory of derangement syndrome proposed by McKenzie and May.[14] We also provided some evidence-based information for treatment and prevention of LBP: self-management and risk factors (e.g. psychosocial factors and fear-avoidance). A 30-minute seminar was given by an orthopedist, the author of this article, where he explained the exercise manual and this exercise.

Participants were asked whether they were willing to do this exercise. In order to promote regular exercise in the intervention group, we took a group approach and routinely monitored participants' motivation for the exercise.

Figure 1 How to do standing back extension 'One Stretch'.

Data collection

At baseline and end of the 1-year study period, data were collected by using a self-administrated questionnaire. The baseline questionnaire contained the following items: age, sex, body mass index (BMI), visit status for medical consultation due to LBP (yes or no), the severity of LBP in the previous 1 month, and psychological factors. The severity of LBP was evaluated by the Von Kroff's grading: (1) no pain, (2) LBP without interfering with work, (3) LBP interfering with work, and (4) LBP interfering with work, leading to sick leave.[15] We defined the pain localized between the costal margin and the inferior gluteal folds[4] as LBP, and illustrated a diagram of the LBP in the questionnaire. Psychological factors were assessed by the mental health score of SF-36 (ver.1.2).[16,17] The questionnaire at the end of the study period assessed the subjective improvement of LBP from baseline (improved, no change, or worse), overall compliance with the exercise during the study period (good or poor), visit status for medical consultation (yes or no), and absence from work due to LBP in the previous 1 year. Participants were asked to record daily exercise to evaluate overall compliance with the exercise during the study period.

Statistical analysis

Values were presented by either means and standard deviations (SDs) or frequencies and percentages. Between-group differences of baseline characteristics were evaluated by using chi-square test for categorical variables and Student's *t*-test for continuous variables. The subjective improvement of LBP and compliance with the exercise were evaluated by using

chi-square test. All statistical tests were two tailed and conducted with a significance level of 0.05.

On medical consultation, we evaluated the change of visit status as the following: (1) improved; participants who had consulted a doctor at baseline, but did not at the end of the study period; (2) no change (−); they had never consulted a doctor; (3) no change (+); they regularly consulted a doctor; and (4) worse; they had not consulted a doctor at baseline, but did at the end of the study period.

Results

A total of 166 care workers participated in this study and were assigned to the intervention group (n=81) or the control group (n=85). The intervention group mean age was 36.8±10.9 years, men (35.8%) and women (64.2%). The control group mean age was 35.9±10.9 years, men (42.3%) and women (57.7%). Thirty care workers were excluded from the analysis because they could not answer the questionnaire at the end of the study period due to moving to other facilities. The analysis population consisted of 64 care workers in the intervention group and 72 care workers in the control group. We took a group approach for the intervention group, where care workers exercise in a group at the daily meeting. This approach was continued to the end of the study period.

Baseline characteristics of the analysis population are shown in Table 1. In all items, including the severity of LBP, mental health score of SF-36, there were no statistically significant differences between the two groups.

The subjective improvement of LBP from baseline and compliance with the exercise were evaluated (Fig. 2). Compared with the control group, the intervention group indicated a higher proportion of care workers who had 'improved' LBP and had 'good' compliance with the exercise, which were statistically significant (P=0.003 and P<0.0001, respectively).

The number of care workers with/without medical consultation and absence from work due to LBP is shown in Table 2. In the intervention group, there were no care workers who had a first medical consultation due to LBP, while six care workers had quit consulting a doctor by the end of the study period. Additionally, no care workers in the intervention group and five care workers in the control group had been absent from work due to LBP at the end of the study period.

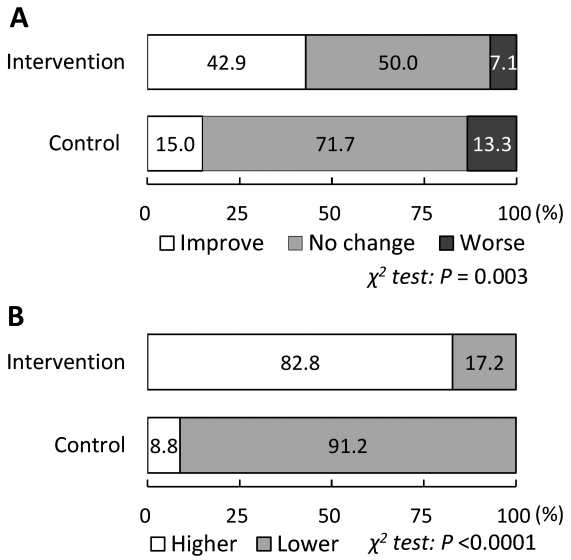


Figure 2 Subjective improvement of low back pain (LBP) and compliance with the exercise. Panel A is the subjective improvement of LBP from baseline to the end of the 1-year study period. Intervention group of 56 care workers and control group of 60 care workers were included due to missing data. Panel B is the overall compliance with the standing back extension exercise during the study period. Intervention group of 64 workers and control group of 68 workers were included due to missing data.

Discussion

The findings of this study suggest that standing back extension exercise 'One Stretch' is effective to prevent care workers from developing and aggravating LBP. We considered that daily practice of this exercise would not only improve LBP but also decrease the number of care workers needing medical consultation or leaving work due to LBP. Furthermore, our group

Table 1 Baseline characteristics in intervention and control groups

Variable	Intervention (n=64)	Control (n=72)	P value
Age, year	38±11	36±11	0.39
Men	23 (35.9)	31 (43.1)	0.39
Women	41 (64.1)	41 (56.9)	
BMI	22.3 (3.5)	21.9 (2.9)	0.49
Medical consultation (+)	7 (10.9)	5 (6.9)	0.41
Severity of LBP in the previous 1 month			
No pain	21 (32.8)	25 (34.7)	0.47
LBP without interfering with work	40 (62.5)	40 (55.6)	
LBP interfering with work	3 (4.7)	7 (9.7)	
Mental health score of SF-36	61.4±19.9	61.3±17.9	0.97

Data were shown as mean±SD or number of participants (%).
LBP: low back pain; BMI: body mass index.

approach would encourage better compliance with the exercise.

Several other studies have supported the use of extension exercises. Long *et al.* found that patients randomized to favorable directional preference exercises, consisting mostly of extension exercises, made significant improvements in LBP compared to those randomized to opposite or mid-range movements.[18] Furthermore, a novel study of kinematic magnetic resonance imaging (kMRI) demonstrated evidence that slightly degenerated intervertebral disks moved in a posterior direction during flexion and in an anterior direction during extension.[19] This may be the mechanism for clinical improvements seen in our study.

In a randomized controlled trial in which military conscripts were randomized to extension in lying exercises or a control group, the intervention group saw a significantly lower prevalence of LBP and care seeking for LBP compared to the control group.[20] The extension approach inhibited developing back problems in young men. This is similar to our study, even if there were differences in age, sex, and an exact posture of extensions.

In this study, there were no significant differences in the baseline characteristics, including the mental health score of SF-36 and the physical activity subscale of FABQ between the intervention and control groups. Previous studies have shown that depression is a risk factor for LBP,[21–23] but depression, as noted by the SF-36 mental score, did not seem to affect our results.

The intervention group showed a higher improvement of LBP and had better compliance with the exercise than the control group. We also noticed that the subjective improvement of 'no change' included both care workers with and without LBP due to the nature of this study design. Indeed, those care workers remained healthy so that none in the intervention group had a first medical consultation or were absent from work by the end of the study period.

On the other hand, the study results suggest that a group approach may improve adherence.[24] Generally, a population approach is considered to be a powerful preventive strategy that affects causal behavior in health care activity.[25] For instance, recommending group exercises for prevention may reduce the prevalence of LBP and save more socioeconomic costs than just treating sick individuals. However, some individuals, such as those having multiple risk factors of LBP or having complaints against small preventive benefits, need an individual approach to preventive behaviors. Both population and individual approaches are required to complement each other.[24]

There were several limitations to this study. First, the questionnaire contained retrospective questions and the participants assessed their condition of LBP 1 year after, and so the possibility for recall bias should be kept in mind. Second, we examined a small sample size and a single population. Owing to the nature of the study, cluster randomized trials with adequate sample size are needed for evaluating intervention. Thus, the generalizability of findings is limited, and the findings should be interpreted with caution. We will perform further examinations through large-scale randomized controlled trials.

Conclusion

Our results suggest that the active exercise 'One Stretch' is effective to control LBP in care workers. In Japan, in addition to the inadequate number of care workers and poor working environment, an increasing number of care workers with LBP disability is a serious problem. Hence, daily practice of this simple exercise would benefit our society, especially in industrial health.

Acknowledgement

We would like to thank Yoshihiro Iwasada (PT, MS, Dip. MDT, the McKenzie Institute Japan) for his valuable advice.

Disclaimer Statements

Contributors Ko Matsudaira and Hiroyuki Oka designed the study. Miho Hiroe, Masatomo Kikkawa, Kou Hiroe, and Ken Hiroe coordinated and supervised data collection at the site. Mari Suzuki, Takayuki Sawada, and Tatsuya Isomura carried out data analyses and drafted the manuscript.

Table 2 The number of workers with/without medical consultation and absence from work due to low back pain (LBP)

	Status (baseline/end of study period)	Intervention ($n=64$)	Control ($n=72$)
Medical consultation			
Improve	±	6 (9.4)	4 (5.6)
No change (−)	−/−	57 (89.1)	65 (90.3)
No change (+)	+/+	1 (1.6)	1 (1.4)
Worse	−/+	0 (0.0)	2 (2.8)
Absence from work in the previous 1 year			
Baseline	+	0 (0.0)	3 (4.2)
End of study period	+	0 (0.0)	5 (6.9)

Data were shown as number of participants (%).

Ko Matsudaira approved the final manuscript as submitted. Ko Matsudaira is the guarantor.

Funding This study was supported as a dissemination project on the 13 fields of occupational injuries and illnesses by the Japan Labour Health and Welfare Organization.

Conflicts of interest All authors disclose no conflicts of interest.

Ethics approval This study was approved by the medical/ethics review board of Kanto Rosai Hospital.

References

1 Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, *et al*. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990 – 2010: a systematic analysis for the global burden of disease study 2010. Lancet. 2012;380:2163–96.
2 Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? JAMA. 1992;268:760–5.
3 Deyo RA, Weinstein JN. Low back pain. N Engl J Med. 2001;344:363–70.
4 Krismer M, van Tulder M; Low Back Pain Group of the Bone; Joint Health Strategies for Europe Project. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best Pract Res Clin Rheumatol. 2007;21:77–91.
5 Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J. 2008;8:8–20.
6 Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA; American Society of Interventional Pain Physicians. Comprehensive review of epidemiology, scope, and impact of spinal pain. Pain Physician. 2009;12:E35–70.
7 Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R. Lost productive time and cost due to common pain conditions in the US workforce. JAMA. 2003;290:2443–54.
8 Fujii T, Matsudaira K. Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. Eur Spine J. 2013;22:432–8.
9 Ministry of Health, Labor and Welfare [Internet]. Comprehensive survey of living conditions 2010. 2010 [cited 2014 April 4]. Available from: http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/3-1.html.
10 Statistics Bureau Ministry of Internal Affairs and Communication [Internet]. Population census and labourforce survey 2012. 2012 [cited 2014 April 4]. Available from: http://www.mhlw.go.jp/bunya/roudoukijun/anzeneisei11/h24.html.
11 Waddell G, Burton AK. Occupational health guidelines for the management of low back pain at work: evidence review. Occup Med (Lond). 2001;51:124–35.
12 Matsudaira K, Konishi H, Miyoshi K, Isomura T, Takeshita K, Hara N, *et al*. Potential risk factors for new onset of back pain disability in Japanese workers: findings from the Japan epidemiological research of occupation-related back pain study. Spine (Phila Pa 1976). 2012;37:1324–33.
13 Burton AK, Balague F, Cardon G, Eriksen HR, Henrotin Y, Lahad A, *et al*. Chapter 2. European guidelines for prevention in low back pain: November 2004. Eur Spine J. 2006;15:S136–68.
14 McKenzie R, May S. Mechanical diagnosis and therapy, 2nd edn. Waikanae, New Zealand: Spinal Publications New Zealand Ltd; 2003.
15 Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain. 1992;50:133–49.
16 Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 health survey for use in Japan. J Clin Epidemiol. 1998;51:1037–44.
17 Fukuhara S, Ware JE Jr., Kosinski M, Wada S, Gandek B. Psychometric and clinical tests of validity of the Japanese SF-36 health survey. J Clin Epidemiol. 1998;51:1045–53.
18 Long A, Donelson R, Fung T. Does it matter which exercise? A randomized control trial of exercise for low back pain. Spine (Phila Pa 1976). 2004;29:2593–602.
19 Zou J, Yang H, Miyazaki M, Morishita Y, Wei F, McGovern S, *et al*. Dynamic bulging of intervertebral discs in the degenerative lumbar spine. Spine (Phila Pa 1976). 2009 34:2545–50.
20 Larsen K, Weidick F, Leboeuf-Yde C. Can passive prone extensions of the back prevent back problems? A randomized, controlled intervention trial of 314 military conscripts. Spine (Phila Pa 1976). 2002;27:2747–52.
21 Currie SR, Wang J. More data on major depression as an antecedent risk factor for first onset of chronic back pain. Psychol Med. 2005;35:1275–82.
22 Hartvigsen J, Frederiksen H, Christensen K. Physical and mental function and incident low back pain in seniors: a population-based two-year prospective study of 1387 Danish Twins aged 70 to 100 years. Spine (Phila Pa 1976). 2006;31:1628–32.
23 Meyer T, Cooper J, Raspe H. Disabling low back pain and depressive symptoms in the community-dwelling elderly: a prospective study. Spine (Phila Pa 1976). 2007;32:2380–6.
24 Doyle YG, Furey A, Flowers J. Sick individuals and sick populations: 20 years later. J Epidemiol Community Health. 2006;60:396–8.
25 Rose G. Sick individuals and sick populations. Int J Epidemiol. 1985;14:32–8.

Clinical Interventions in Aging

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Efficacy of a trunk orthosis with joints providing resistive force on low-back load in elderly persons during static standing

This article was published in the following Dove Press journal:
Clinical Interventions in Aging
1 September 2015
Number of times this article has been viewed

Junji Katsuhira[1]
Ko Matsudaira[2]
Tadashi Yasui[3]
Shinno Iijima[4]
Akihiro Ito[4]

[1]Department of Nursing and Rehabilitation Science at Odawara, International University of Health and Welfare, Odawara, Kanagawa, [2]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo, [3]Kawamura-Gishi Company, Ltd., Daito-shi, Osaka, [4]Graduate School of International University of Health and Welfare, Otawara, Tochigi, Japan

Purpose: Postural alignment of elderly people becomes poor due to aging, possibly leading to low-back pain and spinal deformity. Although there are several interventions for treating these conditions, no previous study has reported the effectiveness of a spinal orthosis or lumbosacral orthosis (LSO) in healthy elderly people without specific spinal deformity. We therefore developed a trunk orthosis to decrease low-back muscle activity while training good postural alignment through resistive force provided by joints with springs (here, called the ORF, which stands for orthosis with joints providing resistive force) as a preventive method against abnormal posture and low-back pain in healthy elderly persons.

Patients and methods: Fifteen community-dwelling elderly men participated in this study. Participants stood freely for 10 seconds in a laboratory setting under three conditions: without an orthosis, with the ORF, and with an LSO. The Damen corset LSO was selected as it is frequently prescribed for patients with low-back pain. Postural alignment during static standing was recorded using a three-dimensional motion capture system employing infrared cameras. Two force plates were used to record center of pressure. Electromyograms were obtained for bilateral erector spinae (ES), left internal abdominal oblique, and right gluteus medius muscles.

Results: Pelvis forward tilt angle tended to increase while wearing the ORF and decrease while wearing the LSO, but these results were not significant compared to no orthosis. Thorax extension angle and thorax angle on pelvis coordinate system significantly increased while wearing the ORF compared to the other two conditions. ES activity significantly decreased while wearing the ORF compared to the other two conditions. Internal oblique activity was significantly smaller while wearing the LSO than with no orthosis. Center of pressure did not significantly differ among the conditions.

Conclusion: The ORF significantly improved trunk alignment and decreased ES activity in healthy elderly subjects during static standing.

Keywords: muscular activity, center of pressure, standing posture, spine

Introduction

Postural alignment worsens gradually over the course of aging.[1] This poor postural alignment, which manifests as spinal kyphosis, can result in irreversible degeneration of the intervertebral disks and ligaments. It also often causes postural instability and leads to vertebral bone fracture and increased risk of falling.[2,3]

Due to these problems, various interventions are used for elderly persons who have spinal deformity to improve postural alignment. Battaglia et al[4] reported that exercise improved spinal flexibility, and Imagama et al[5] reported the improvement of lumbar lordosis angle, sagittal balance, and back muscle strength in elderly patients

Correspondence: Junji Katsuhira
Department of Nursing and Rehabilitation Science at Odawara, International University of Health and Welfare, 1-2-25 Shiroyama, Odawara, Kanagawa 250-8588, Japan
Tel +81 465 21 6704
Fax +81 465 21 6745
Email katsuhir@iuhw.ac.jp

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/CIA.S85294

through a training program that included muscle strength and spinal range of motion exercises.

Spinal orthoses provide a way to directly modify the posture of elderly persons with spinal misalignment. Piffer et al[6] reported that use of a newly designed spinal orthosis, the Spinomed® (medi GmbH & Co. KG, Bayreuth, Germany), for 6 months improved muscle strength, body balance, kyphosis angle, and vital capacity in elderly patients with osteoporosis. Ishida et al[7] reported that a rucksack-type orthosis instantly modified spine alignment and decreased erector spinae (ES) activity in elderly patients with kyphosis.

As mentioned earlier, previous studies have reported that exercise and orthotic therapy are effective in treating the elderly with spinal misalignment. It is, however, also important to offer preventive intervention for the healthy elderly. Costantino et al[8] reported that chronic low-back pain (LBP) in elderly people without specific spinal deformity could be effectively treated using a back school program, including exercise therapy for rehabilitation. However, to our knowledge, no previous studies have reported the effectiveness of a spinal orthosis or lumbosacral orthosis (LSO) in healthy elderly people without specific spinal deformity. A review of data from the Cochrane Database also found no evidence for the efficacy of lumbar support for decreasing low-back load.[9] In addition, Rostami et al[10] reported that use of an LSO for 4 weeks resulted in decreased trunk core muscle volume, and therefore, the drawback of long-term corset use may exceed the benefits.

In a previous study, to provide a preventive method against abnormal posture and LBP in elderly persons, we designed a trunk orthosis to address these issues by training good postural alignment while decreasing low-back load via resistive force provided by joints with springs (Figure 1).[11] This orthosis with joints providing resistive force (hereafter, the ORF) produces a resistive moment that rotates the trunk backward and the pelvis forward (Figure 2). Resistive moment applied to the trunk can not only rotate it backward to shift the center of gravity of the head, arms, and trunk to the L4/L5 joint but also directly decrease low-back extension moment because the resistive moment work is in the same direction as that produced by ES activity. In addition, reaction moment can promote forward rotation of the pelvis, and this effect can also facilitate extension of the thorax. The ORF may therefore improve alignment and decrease ES activity in elderly people. Indeed, our recent studies reported that the ORF improved trunk alignment and gait performance of hemiparetic patients[11] and elderly persons during level walking.[12] However, we did not examine the effect of the ORF on ES activity and spinal alignment during static standing.

In light of the finding of no decrease in low-back muscle activity using LSOs,[9] this study aimed to examine, through

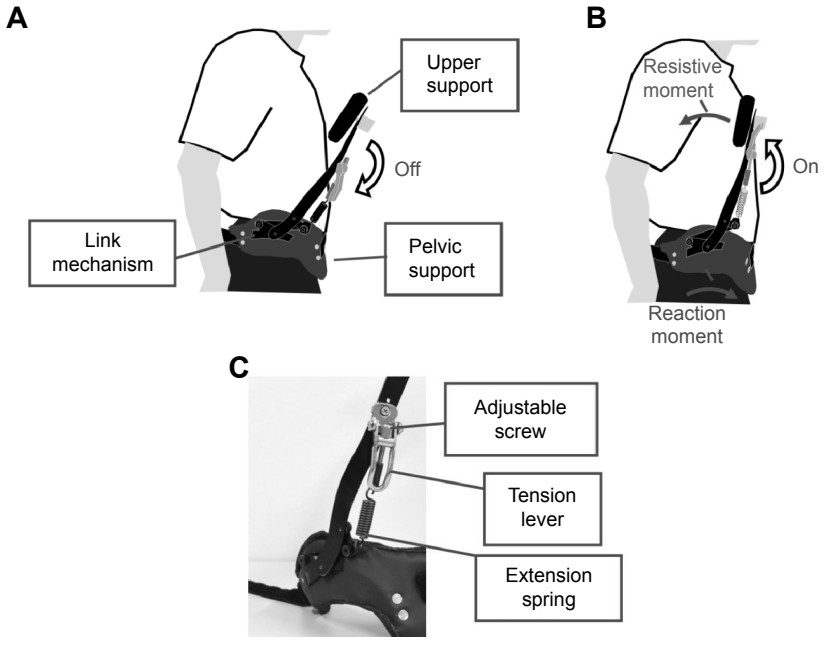


Figure 1 Trunk ORF.
Notes: (**A**) Wearing ORF without resistive force on the chest, (**B**) wearing ORF with resistive force on the chest, (**C**) detail of link mechanism.
Abbreviation: ORF, orthosis with joints providing resistive force.

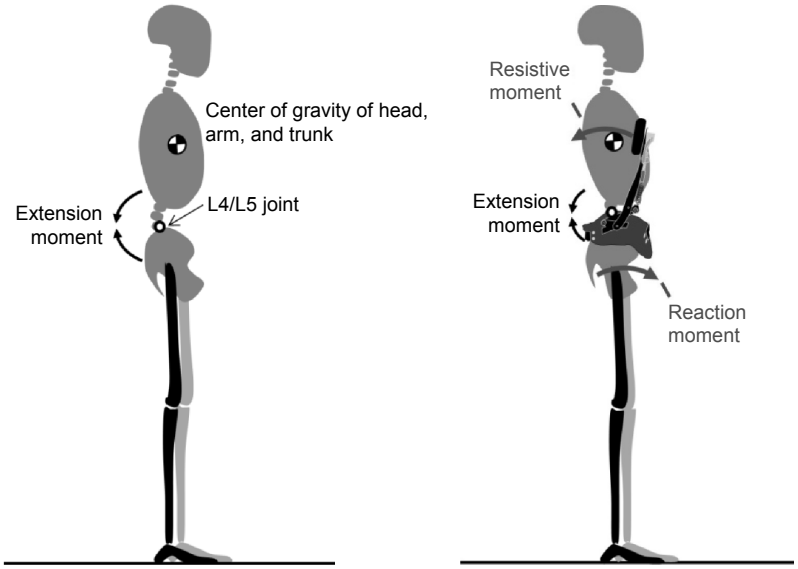


Figure 2 Biomechanical effect of ORF while static standing.
Abbreviation: ORF, orthosis with joints providing resistive force.

biomechanical analysis, the effects of the ORF on healthy elderly people during static standing and to compare the effects with those of an LSO and no orthosis. We hypothesized that wearing the ORF would effectively decrease ES activity and modify trunk alignment in healthy elderly people during static standing.

Material and methods

Participants

We enrolled 15 healthy community-dwelling elderly men (mean age, 67.7±6.1 years; mean height, 162.4±5.7 cm; mean weight, 62.3±7.8 kg) from a group of 31 candidates. We excluded those with neurological disease, pain, history of orthopedic surgery, history of orthopedic treatment within the past 5 years, and history of LBP within the past 1 year. In the first phase of recruitment, we identified only two female subjects willing to participate in our study; however, due to a history of orthopedic treatment within the past 5 years, they were excluded. Therefore, we decided not to include any female subjects to mitigate the potential effects of sex and orthopedic disease. All participants provided written informed consent, and the study was approved by the ethics committees of the participating institutions, International University of Health and Welfare.

Features of the ORF

The features of the ORF (Figures 1 and 2) were described in our previous report.[11] Briefly, the ORF weighs 0.99 kg and has a 40° range of motion. Pelvic and upper supports are placed on the ileum and sternum, respectively. Stainless steel joints are connected to the upper support, with a nylon pad, and also to the pelvic support. These joints employ extension springs to produce tension, which is translated by a link mechanism into a resistive moment on the chest and a reaction moment on the posterior pelvis. The upper support initially inclines backward to exert resistive force on the chest and is then released via a mechanism that pulls tension levers downward. The resistive force can be increased or decreased via adjustment screws. The ORF is currently an investigational product that is not FDA-approved or approved by the corresponding national agency for the indication described herein.

Experimental conditions

Participants freely stood for 10 seconds under three conditions in the laboratory setting: with no orthosis, with the ORF, and with an LSO (Damen Corset, Pacific Supply, Osaka, Japan; Figure 3). As the Damen corset is frequently prescribed for patients with LBP, it was selected for use in this study. Our previous study reported carry-over effects of the ORF on body alignment of hemiparetic patients after removal while level walking.[11] We therefore decided to measure two trials without intervention (no orthosis) first, after which subjects completed the two orthosis trials in random order. Participants were given 5 minutes to become accustomed to wearing the orthoses, and there was a minimum rest interval of 5 minutes between conditions. Resistive force exerted on the chest was

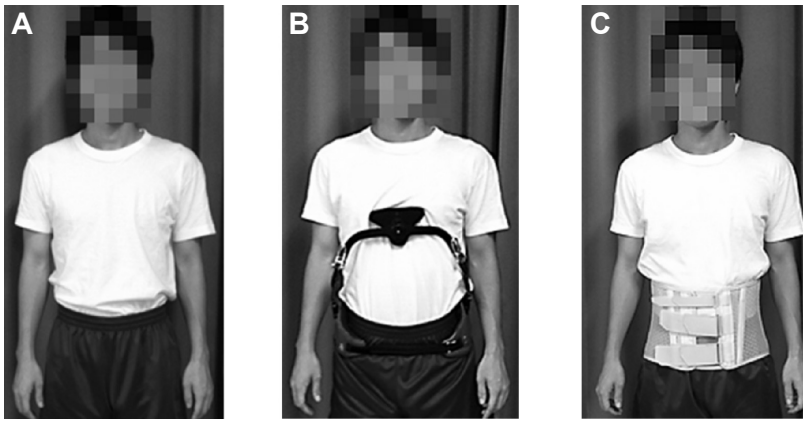


Figure 3 Experimental conditions.
Notes: (**A**) Without an orthosis, (**B**) with the LSO, (**C**) with the ORF.
Abbreviations: ORF, orthosis with joints providing resistive force; LSO, lumbosacral orthosis.

measured in real time with a strain gauge (Compression load cell LCN-A, Kyowa, Tokyo, Japan). The force data were transferred to a laptop via Bluetooth. Force was set to 20–25 N during static standing, and pressure between the corset and abdomen was set to 10 mmHg in all measurement conditions.[13] To obtain maximal voluntary contraction (MVC) values for normalization of individual muscle activities, participants performed maximal isometric contractions against gravity, while the experimenter (a licensed therapist) applied maximum resistance.[14] After practicing stable isometric contraction efforts, single maximal contractions of each muscle were recorded according to *Daniels and Worthingham's Muscle Testing*. Subjects performed contractions against gravity, with maximum resistance applied by the examiner in the supine position to obtain MVC of the left internal abdominal oblique (IO) muscle (lifting head and shoulders from the table with right elbow toward left knee against imposed resistance to the right shoulder region), in the prone position to obtain MVC of the bilateral ES (back extension with hands resting on head against imposed resistance to the scapular region), and in the side-lying position with test leg elevated to obtain MVC of the right GM muscle (abduction with limb slightly extended beyond the midline and the pelvis rotated slightly forward while imposing resistance to the lateral surface of the knee).

Experimental setup

Static standing was recorded with a three-dimensional (3D) motion capture system (Vicon 612, Vicon, Oxford, UK) that employed two force plates (AMTI, Watertown, MA, USA), 12 infrared cameras (sampling rate, 120 Hz), and 13 infrared-reflective markers (diameter, 14 mm) attached to the C7 spinous process, T12 spinous process, L5 spinous process, manubrium sterni, second sacral vertebra and bilateral acromion process, bilateral anterior and posterior superior iliac spine, and bilateral iliac crest (Figure 4). All markers were captured in a reference static standing position, and then, the bilateral anterior superior iliac spine and iliac crest markers were removed before initiating measurements because they interfered with wearing of the orthoses. The positions of these removed markers were then interpolated using the reference static trial. To measure low-back muscle activity during static standing, electromyograms employing active electrodes to decrease noise (Biometrics, Newport, UK) were obtained for bilateral ES (2 cm to the side between L4 and L5 vertebrae),[15] left IO (2 cm below the anterior superior iliac spine [ASIS] aligned approximately 6° from the line between bilateral ASIS),[16] and right GM (2.5 cm below the line between the iliac crest and greater trochanter)[17] (Figure 3). The target muscle to confirm the effect of the ORF orthosis was the bilateral ES. The IO is classified as a core abdominal muscle, and Rostami et al[10] reported that IO volume decreased after long-term use of a corset. Also, hip abduction moment increased when elderly subjects wore the ORF while level walking in our previous study.[12] Thus, unilateral IO and GM activities were included as supplementary measures in this study.

Electromyography (EMG) signals were measured at 1,080 Hz because the acquisition frequency should be a whole-number multiple of the sampling frequency of the Vicon system (Vicon) (120 Hz).

Data analysis

All signals, including marker displacements and analog EMG signals, were acquired by the Vicon Datastation (Vicon) and then synchronized by correcting for the difference in

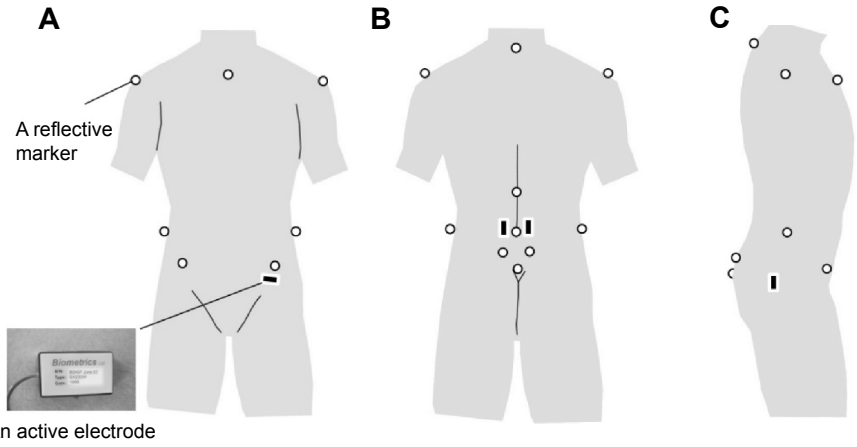


Figure 4 Positions of reflective markers and electrodes for electromyogram recording.
Notes: (**A**) Anterior surface, (**B**) posterior surface, (**C**) lateral surface.

sampling frequency (120 vs 1,080 Hz) using Vicon Workstation software (Vicon). Band pass filter (20–420 Hz) was used during acquisition to decrease noise,[18] and the data were then imported into Visual 3D analytical software (C-motion, Inc., Germantown, MD, USA) for kinetic and kinematic data analysis. Electromyograms were normalized to individual MVCs during isometric contraction (%MVC). Root mean squares (RMS) for a 50 ms window were calculated, and integral values for these muscles (IEMG) were calculated. Consequently, Equation 1 is given as follows:

$$\mathrm{IEMG} = \int_0^t [\mathrm{EMG}_{\mathrm{RMS}}](t)\,\mathrm{d}\,t \qquad (1)$$

where $\mathrm{EMG}_{\mathrm{RMS}}$ denotes EMG after conducting RMS for a 50 ms window and t denotes 10 seconds, indicating the end time of static standing.

The physical coordinates and ground reaction force data were low-pass filtered with a second-order recursive Butterworth filter (cutoff frequencies 6 and 18 Hz, respectively) according to Winter's technique.[19] Also, the link segment model consisted of a trunk segment and a pelvis segment, and markers on each segment were used to calculate 3D trunk and pelvic angles using coordinate systems and the Eulerian method.

Center of pressure (COP) displacements are commonly used to evaluate balance performance in the elderly,[20] and COP path length was validated as a performance outcome measure in a large group study.[21] We therefore calculated COP path length to evaluate balance in the elderly subjects. The COP of the force vector of bilateral feet and COP total trajectory length were calculated for 10 seconds of standing.

Statistical analysis

Average kinetic, kinematic, and EMG values, as well as IEMG were acquired for 10 seconds of static standing. Mean average values of 3D trunk and pelvic angles, average EMG, and IEMG were calculated from two trials, per condition for analysis. Variables were compared by repeated measures analysis of variance (RT-ANOVA) followed by multiple pair-wise comparisons to Bonferroni correction. Statistical analysis was performed with SPSS20 (SPSS Inc., Chicago, IL, USA). Statistical significance was set at $P<0.05$.

Results

Average pelvis forward tilt angle, thorax extension angle, and thorax angle on the pelvis coordinate system are shown in Table 1 for the three static standing conditions: no orthosis, ORF, and LSO. One-way RT-ANOVA revealed a significant main effect of condition for all three angles. Pelvis forward tilt angle tended to increase while wearing the ORF and decrease while wearing the LSO, but these results were not significant compared to those without an orthosis. Thorax extension angle and thorax angle on the pelvis coordinate system were significantly increased while wearing the ORF compared to the other two conditions.

Integral and average muscular activities are shown in Table 2. One-way RT-ANOVA indicated significant main effect of condition on integral and average muscular activities of the bilateral ES and left IO, but not on the GM. Integral and average muscular activities of the bilateral ES were significantly smaller while wearing the ORF compared to no orthosis. For the left ES, integral and average muscular activities were significantly less while wearing the ORF compared to the LSO. The same trend was seen for

Table 1 Mean and standard deviation for kinematic parameters in three standing conditions (N=15) and results of statistical analysis

	Without ORF	ORF mean (SD)	LSO	*P*-value	Without ORF, ORF	Without ORF, LSO	ORF, LSO
Pelvis forward tilt angle (°)	6.13 (5.07)	8.13 (6.55)	4.46 (5.77)	**0.010**	0.281	0.165	0.050
Thorax extension angle (°)	4.78 (3.66)	6.56 (4.12)	5.37 (3.52)	**<0.001**	**0.004**	0.172	**0.021**
Thorax angle on pelvis coordinate system: extension + (°)	−2.40 (7.02)	1.38 (7.62)	−3.49 (7.06)	**0.001**	**0.022**	0.619	**0.013**

Note: Data in bold are statistically significant, statistical significance was set at $P<0.05$.
Abbreviations: LSO, lumbosacral orthosis; ORF, orthosis with joints providing resistive force; SD, standard deviation.

the right ES, but it was not significant. Integral and average muscular activities of the left IO were significantly dimished while wearing the LSO compared to no orthosis, but there was no significant difference between ORF and no orthosis conditions.

COP trajectory is shown in Table 3. No significant difference was observed among the three conditions.

Discussion

Incidence of LBP increases with age.[22] Age-related spinal deformity as well as ES hyperactivity are among the major causes of LBP.[18] We previously developed the ORF to modify trunk and pelvis alignment and decrease ES activity. We examined the effects of this orthosis on healthy elderly participants and found that the ORF could effectively modify trunk alignment while decreasing ES activity compared to an LSO or no orthosis. In fact, no significant positive effects were observed while wearing the LSO.

Coskun Benlidayi and Basaran[23] reported that lumbar lordosis is significantly smaller in elderly than in young subjects. Wearing of the ORF in this study significantly extended the upper trunk and tended to tilt the pelvis forward, which would contribute to increasing lumbar lordosis. Lee et al[13] reported the effectiveness of a lumbar belt and pelvic belt to modify pelvic and spinal alignment on posture in healthy young participants. Additionally, Piffer et al[6] reported that use of the newly designed Spinomed® (medi GmbH & Co. KG) orthosis for 6 months improved kyphosis in elderly patients with osteoporosis. However, the present study appears to be the first to report on an orthosis that can modify alignment in healthy elderly participants without spinal deformity or chronic LBP.

Most spinal orthoses were developed to treat LBP and abnormal spinal alignment. Based on the findings of this study, the ORF shows promise for the modification of spinal alignment in elderly people without any specific deformity or syndromes related to LBP. Previous studies reported that interventions teaching awareness of posture, such as lessons on the Alexander technique, could be beneficial in treating LBP, as well as cost-effective.[24,25] Back school programs that include exercise therapy could also be effective in elderly people with chronic LBP with no specific spinal deformities.[8] Therefore, the ORF's ability to modify trunk alignment in elderly people might be useful to increase awareness of

Table 2 Mean and standard deviation for muscular activities in three standing conditions (N=15) and results of statistical analysis

Muscle activity	Without ORF	ORF mean (SD)	LSO	*P*-value	Without ORF, ORF	Without ORF, LSO	ORF, LSO
Percent IEMG							
Right ES	57.75 (43.97)	45.48 (32.77)	53.58 (38.77)	**0.003**	**0.024**	0.433	0.055
Left ES	81.27 (51.07)	60.81 (51.15)	78.87 (50.29)	**<0.001**	**0.001**	1.000	**0.004**
Right GM	59.17 (53.07)	53.30 (60.42)	67.14 (79.43)	0.132			
Left IO	152.23 (95.66)	132.09 (91.53)	127.77 (82.55)	**0.010**	0.117	**0.037**	1.000
Percent MVC							
Right ES	5.86 (4.53)	4.60 (3.33)	5.43 (3.96)	**0.003**	**0.027**	0.435	0.053
Left ES	8.17 (5.09)	6.10 (5.10)	7.92 (5.01)	**<0.001**	**0.001**	1.000	**0.004**
Right GM	5.92 (5.30)	5.33 (6.40)	6.72 (7.94)	0.132			
Left IO	15.27 (9.52)	13.25 (9.12)	12.83 (8.22)	**0.009**	0.114	**0.037**	1.000

Notes: %IEMG, integral of the EMG over 10 seconds relative to maximum. %MVC, average EMG over 10 seconds relative to maximum. Data in bold are statistically significant, statistical significance was set at $P<0.05$.
Abbreviations: ES, erector spinae; GM, gluteus media; IEMG, integral electromyography; IO, internal oblique; MVC, maximum voluntary contraction; LSO, lumbosacral orthosis; ORF, orthosis with joints providing resistive force; SD, standard deviation.

Table 3 Mean and standard deviation for COP trajectory in three standing conditions (N=15) and results of statistical analysis

	Without ORF	ORF mean (SD)	LSO	*P*-value
Trajectory of COP (m)	0.1285 (0.0610)	0.1314 (0.0454)	0.1304 (0.0481)	0.924

Abbreviations: COP, center of pressure; LSO, lumbosacral orthosis; ORF, orthosis with joints providing resistive force; SD, standard deviation.

proper posture and help prevent misalignment and spinal deformity.

The most significant effect of wearing the ORF in this study was to decrease ES activity. A systematic review reported that wearing an LSO alone could not decrease low-back load.[9] However, Cholewicki et al[18] reported that wearing an LSO could decrease ES activity during a postural control task such as sitting on an unstable seat. In this study, ES activity decreased slightly by 1%–2% MVC while wearing the ORF. Despite this small decrease, a modeling study showed that adding a 32 kg mass to the trunk required an increase in trunk muscle cocontraction of approximately 1%–2% MVC above the level normally necessary to maintain a stable upright position of the spine around the neutral posture.[26] Furthermore, previous studies have reported that maintaining muscular contraction above 5% MVC may cause back-muscle fatigue and pain.[27] In the present study, average right ES activity was below 5% MVC while wearing the ORF, but was higher for the other two conditions. In addition, average left ES activity decreased within the 6% MVC range while wearing the ORF, but was approximately 8% for the other two conditions. These results indicate that the ORF was effective in decreasing the activity of low-back muscles, which was not observed while wearing the LSO. In a postural control task, trunk muscle activity does not usually exceed 3% MVC in young participants.[27] However, muscular activity of elderly people during MVC is lower, and therefore greater relative effort would be needed to maintain an upright standing position. Rostami et al[10] reported that wearing an LSO led to decreased muscle volume in the abdominal side muscles, including the IO. Wearing the ORF in the present study did not significantly decrease IO activity, whereas wearing the LSO significantly decreased it, compared to no orthosis. These results support those of Rostami et al[10] and suggest that resistive force might not decrease low-back muscle activity without a decrease in side abdominal muscle volume.

No significant changes in COP trajectory were observed while wearing either orthosis in this study. Cholewicki et al[18] reported that COP displacement during an unstable sitting task did not significantly differ from the control condition (no orthosis) while wearing an LSO. Furthermore, Chen et al[28] reported that their insole served to improve the stability index, as calculated using COP displacement in elderly participants. COP is mainly controlled by the ankle plantar flexors, and therefore intervention using foot-based orthotic devices may be more useful than trunk-based devices.

Wearing the ORF during static standing served to decrease ES activity and modify trunk alignment, which may be effective for the prevention and treatment of LBP and spinal deformity in elderly people. However, this study has several limitations. First, we did not confirm the effects of long-term ORF use, as wearing the ORF or LSO for lengthy periods might adversely affect muscle control. Second, participants were healthy elderly men only and this was a within-subject trial. Healthy elderly women and participants with LBP or low-back disorders should be included in future studies, including randomized controlled trials. Third, we confirmed reduction of ES activity, but it might not be the exact cause. Resistive moment generated by the ORF joints may have decreased extension moment exerted by the ES muscles, or greater thorax extension angle could have decreased ES activity by reducing the lever arm from the L4/L5 joint to gravitational force on the center of gravity of the upper trunk. Moreover, ES is a surface muscle. In a future study, we must confirm these changes in spinal loading using a more detailed biomechanical model including the deep back muscles, such as bilateral multifidus muscle.

Conclusion

This study revealed that the ORF significantly improved trunk alignment and decreased ES activity in healthy elderly participants. These findings suggest that the ORF may help prevent LBP and spinal deformity in elderly people. Further studies are needed to examine the use of the ORF in patients with LBP and spine deformity.

Acknowledgment

This study was supported by the dissemination project of Clinical Research for Occupational Injuries and Illness from Ministry of Health, Labour and Welfare of Japan.

Author contributions

Provided substantial contribution to conception and design: JK, KM, TY. Involved with data collection and analysis: JK, SI, AI. Involved with interpretation of data: JK, SI, AI. Drafted the article: JK, KM. Revised the manuscript critically for important intellectual content: JK, KM, TY, SI, AI. Provided approval of the version to be published: JK, KM, TY, SI, AI.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Gelb DE, Lenke LG, Bridwell KH, Blanke K, McEnery KW. An analysis of sagittal spinal alignment in 100 asymptomatic middle and older aged volunteers. *Spine*. 1995;20:1351–1358.
2. Nguyen T, Sambrook P, Kelly P, et al. Prediction of osteoporotic fractures by postural instability and bone density. *BMJ*. 1993;307:1111–1115.
3. Kasukawa Y, Miyakoshi N, Hongo M, et al. Relationships between falls, spinal curvature, spinal mobility and back extensor strength in elderly people. *J Bone Miner Metab*. 2010;28:82–87.
4. Battaglia G, Bellafiore M, Caramazza G, Paoli A, Bianco A, Palma A. Changes in spinal range of motion after a flexibility training program in elderly women. *Clin Interv Aging*. 2014;9:653–660.
5. Imagama S, Matsuyama Y, Hasegawa Y, et al. Back muscle strength and spinal mobility are predictors of quality of life in middle-aged and elderly males. *Eur Spine J*. 2011;20:954–961.
6. Pfeifer M, Kohlwey L, Begerow B, Minne HW. Effects of two newly developed spinal orthoses on trunk muscle strength, posture, and quality-of-life in women with postmenopausal osteoporosis: a randomized trial. *Am J Phys Med Rehabil*. 2011;90:805–815.
7. Ishida H, Watanabe S, Yanagawa H, Kawasaki M, Kobayashi Y, Amano Y. Immediate effects of a rucksack type orthosis on the elderly with decreased lumbar lordosis during standing and walking. *Electromyogr Clin Neurophysiol*. 2008;48:53–61.
8. Costantino C, Romiti D. Effectiveness of Back School program versus hydrotherapy in elderly patients with chronic non-specific low back pain: a randomized clinical trial. *Acta Biomed*. 2014;85:52–61.
9. van Duijvenbode I, Jellema P, van Poppel M, van Tulder MW. Lumbar supports for prevention and treatment of low back pain. *Cochrane Database Syst Rev*. 2008:Cd001823.
10. Rostami M, Noormohammadpour P, Sadeghian AH, Mansournia MA, Kordi R. The effect of lumbar support on the ultrasound measurements of trunk muscles: a single-blinded randomized controlled trial. *PM R*. 2014;6:302–308; quiz 308.
11. Katsuhira J, Miura N, Yasui T, Takane M, Sumiko Y. Efficacy of a newly designed trunk orthosis with joints providing resistive force in adults with post-stroke hemiparesis. *Prosthet Orthot Int*. Epub August 18, 2014.
12. Iijima S, Katsuhira J, Ito A, Nomura T, Maruyama H. Effects of a trunk brace with joints that provides a resistive force to modify pelvic alignment during level walking in the elderly. *Phys Ther Japan*. 2014;41:355–363.
13. Lee ES, Ko CW, Suh SW, Kumar S, Kang IK, Yang JH. The effect of age on sagittal plane profile of the lumbar spine according to standing, supine, and various sitting positions. *J Orthop Surg Res*. 2014;9:11.
14. Montgomery J, Hislop H, Connelly B. *Daniels and Worthingham's Muscle Testing: Techniques of Manual Examination*. 8th ed. Maryland Heights, MO: Saunders/Elsevier; 2007.
15. De Foa JL, Forrest W, Biedermann H. Muscle fibre direction of longissimus, iliocostalis and multifidus: landmark-derived reference lines. *J Anat*. 1989;163:243.
16. Ng JK, Kippers V, Richardson CA. Muscle fibre orientation of abdominal muscles and suggested surface EMG electrode positions. *Electromyogr Clin Neurophysiol*. 1998;38:51–58.
17. Perotto A, Delagi EF, Iazzetti J, Morrison D. *Anatomical Guide for The Electromyographer: The Limbs and Trunk*. Springfield, IL: Charles C Thomas; 2005.
18. Cholewicki J, Reeves NP, Everding VQ, Morrisette DC. Lumbosacral orthoses reduce trunk muscle activity in a postural control task. *J Biomech*. 2007;40:1731–1736.
19. Winter DA. *Biomechanics and Motor Control of Human Movement*. Hoboken, NJ: John Wiley & Sons; 2009.
20. Kalisch T, Kattenstroth JC, Noth S, Tegenthoff M, Dinse HR. Rapid assessment of age-related differences in standing balance. *J Aging Res*. 2011;2011:160490.
21. Donath L, Roth R, Zahner L, Faude O. Testing single and double limb standing balance performance: comparison of COP path length evaluation between two devices. *Gait Posture*. 2012;36:439−443.
22. Dunn KM, Hestbaek L, Cassidy JD. Low back pain across the life course. *Best Pract Res Clin Rheumatol*. 2013;27:591–600.
23. Coskun Benlidayi I, Basaran S. Comparative study of lumbosacral alignment in elderly versus young adults: data on patients with low back pain. *Aging Clin Exp Res*. 2015;27:297–302.
24. Little P, Lewith G, Webley F, et al. Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain. *BMJ*. 2008;337:a884.
25. Hollinghurst S, Sharp D, Ballard K, et al. Randomised controlled trial of Alexander technique lessons, exercise, and massage (ATEAM) for chronic and recurrent back pain: economic evaluation. *BMJ*. 2008;337:a2656.
26. Cholewicki J, Panjabi MM, Khachatryan A. Stabilizing function of trunk flexor-extensor muscles around a neutral spine posture. *Spine*. 1997;22:2207–2212.
27. Bjorksten M, Jonsson B. Endurance limit of force in long-term intermittent static contractions. *Scand J Work Environ Health*. 1977;3:23–27.
28. Chen TH, Chou LW, Tsai MW, Lo MJ, Kao MJ. Effectiveness of a heel cup with an arch support insole on the standing balance of the elderly. *Clin Interv Aging*. 2014;9:351–356.

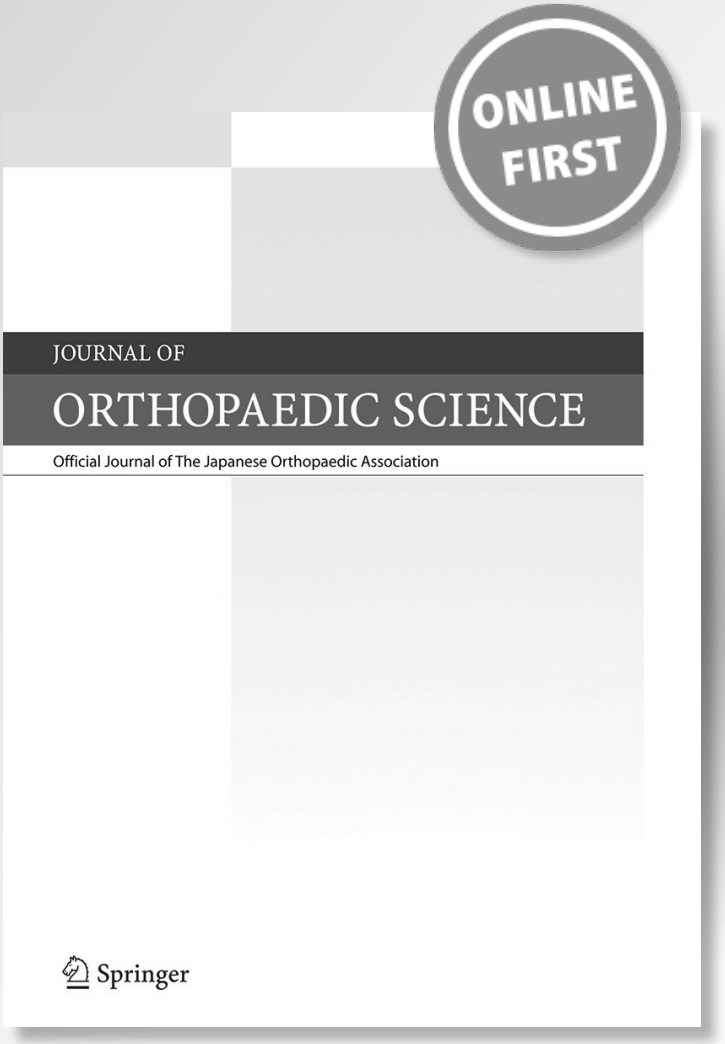
Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury pain and/or low back pain

Norimasa Kikuchi, Ko Matsudaira, Takayuki Sawada & Hiroyuki Oka

Journal of Orthopaedic Science
Official Journal of the Japanese Orthopaedic Association

ISSN 0949-2658

J Orthop Sci
DOI 10.1007/s00776-015-0751-3



Springer

J Orthop Sci
DOI 10.1007/s00776-015-0751-3

ORIGINAL ARTICLE

Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury pain and/or low back pain

Norimasa Kikuchi[1,2] · Ko Matsudaira[3] · Takayuki Sawada[1,2] · Hiroyuki Oka[3]

Received: 17 February 2015 / Accepted: 30 June 2015

Abstract
Background Although the Tampa Scale for Kinesiophobia (TSK) is useful for measuring fear of movement in patients with musculoskeletal pain, no psychometrically validated Japanese version is available. We evaluated the reliability and validity of the Japanese version of the TSK-J (original 17-item version and shorter 11-item version).
Methods The data subset used in this psychometric testing was derived from a survey previously conducted to collect information on musculoskeletal pain due to motor vehicle accident. For reliability, internal consistency was assessed via Cronbach's alpha coefficient. For concurrent validity, Pearson correlation coefficients of the TSK-J with the pain catastrophizing scale (PCS), euroqol 5 dimension (EQ-5D), and numerical rating scales (NRSs) for pain in the neck and back were calculated. For known-group validity, the relationship between variables (e.g., depression, somatic symptoms, treatment period) and the TSK-J score was examined.
Results Data from 956 persons who had suffered from a motor vehicle accident were used in this analysis. For reliability, internal consistency was demonstrated, with Cronbach's alpha statistics of 0.850 (TSK-J17) and 0.919 (TSK-J11). For concurrent validity, significantly strong correlations were demonstrated between the TSK-J versions and PCS total score and subscales ($r = 0.602$–0.680). For known-group validity, as hypothesized, significantly higher TSK-J scores were observed in persons with depressive mood, somatic symptoms, and longer treatment period.
Conclusions The present analysis showed that the Japanese versions of the TSK-J17 and TSK-J11 were psychometrically reliable and valid for detecting fear of movement in the Japanese population suffering from neck to back pain due to a motor vehicle accident.

Introduction

A high level of musculoskeletal pain may evoke the perception of fear of future pain. People may avoid movements or physical activities due to exaggerated fears that pain will result in additional functional restriction [1]. Avoidance of physical activities based on fear of movement (Kinesiophobia) leads to further avoidance [2]. Furthermore, avoidance of pain-inducing activities can result in a reduction of muscle strength and flexibility, which may partly contribute to a delay in recovery. This repeating cycle of fear of movement and avoidance behaviors may perpetuate the chronicity of the condition, resulting in disability. The contributing role of fear avoidance beliefs in the development of long-term disability has been widely recognized [3], and a low level of fear avoidance was reported to be the most useful item for predicting an earlier recovery in patients with acute low back pain [4]. Catastrophizing and somatic symptoms are additional major factors associated with chronicity in patients with whiplash injury [5].

The Tampa Scale of Kinesiophobia (TSK), a 17-item self-reported measure originally developed to discriminate between non-excessive fear and phobia among patients

✉ Norimasa Kikuchi
norimasa_kikuchi@jp-css.com

[1] Clinical Study Support, Inc, Daiei Bldg, 2F 1-11-20 Nishiki, Naka-ku, Nagoya 460-0858, Japan

[2] Department of Public Health, Aichi Medical University School of Medicine, Nagakute, Japan

[3] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

Published online: 23 July 2015

with chronic musculoskeletal pain (Miller RP, Kori SH, Todd DD. The Tampa Scale. Unpublished report 1991), is widely used to assess pain-related fear of movement or re-injury in patients with musculoskeletal complaints. The TSK employs a 4-point Likert scale, with scoring options ranging from 1 (strongly disagree) to 4 (strongly agree). A total score is calculated following inversion of the individual scores of items 4, 8, 12, and 16. The total score of the original 17-item version ranges between 17 and 68, with a higher score indicating a higher degree of Kinesiophobia. The TSK was developed in English, and has thus far been translated into various languages. The psychometric properties of both the original English version and other language versions have been assessed in several patient populations, including patients with chronic musculoskeletal pain [6], low back pain (LBP) [7, 8], whiplash injury pain [9], shoulder pain [10], temporomandibular disorder [11], sciatica [12], and fibromyalgia [13]. Based on the results of exploratory or confirmatory factor analysis in these studies, several factor-structured models with a different number of items (e.g., 17, 13, 12, or 11) have been proposed. Among these versions, an 11-item version that excluded the six psychometrically poor items (4, 8, 9, 12, 14, and 16) is the most widely-used short version. This 11-item version was reported to possess psychometric properties that are similar to the original TSK, and offers the advantage of brevity [14].

In previous work, Matsudaira et al. translated the original English version into Japanese and linguistically validated it, with the aim of introducing the TSK in Japan [15]. In this study, we evaluated the reliability and validity of the Japanese version of the TSK [both the 17-item original version (TSK-J17) and the 11-item shorter version (TSK-J11)] in people with spinal pain due to a motor vehicle accident, including neck pain as a whiplash-associated disorder and LBP.

Methods

Study population

To assess the psychometric properties of the TSK-J, we used a data subset derived from an online survey we had previously conducted in 2012 to collect information on musculoskeletal disorders related to a motor vehicle accident among the general Japanese population. Potential participants were recruited through an Internet panel provided by an Internet research company, including approximately 1.8 million individuals aged from 20 to 79 years as research volunteers. The company's volunteers were consistent with the general Japanese population, and were stratified by sex and age. From these volunteers, 1,063,083 individuals were randomly selected, contacted by e-mail, and invited to complete an online questionnaire regarding a motor vehicle accident experienced in the past 12 months (first survey). Among these individuals, 227,853 were considered effective users, as the research company was unable to exclude non-users from invitations for technical reasons. The first survey was closed when the number of participants reached 127,956 [mean (SD) age 47.7 (10.8), male 63.6 %]. For this reason, the response rate was not relevant to this survey. Of these, 1,639 (1.3 %) individuals who responded that they had suffered from whiplash injury and/or LBP due to a motor vehicle accident in the past 12 months were screened and again invited to complete the online questionnaire (second survey), in order to investigate the impact of the motor vehicle accident on the physical and psychosocial aspects of their lives. Responses from 974 individuals (response rate 59.4 %) were obtained. After excluding data from 18 individuals due to inconsistent responses, data from 956 individuals was included into the analysis. Note that participants received points for online shopping as an incentive for participating in the survey. Double registration was prevented by checking e-mail address duplication and by blocking access to the questionnaire once a responder had completed the survey.

The TSK-J was translated and linguistically validated, according to the general cross-cultural adaptation process: (1) forward-translation (English to Japanese), (2) back-translation (Japanese to English), and 3) cognitive debriefing. Cognitive debriefing interviews of 6 Japanese adult respondents (three male, three female) were conducted to assess their comprehension of the questions and response scales.

This survey was approved by the medical/ethics review board of the Japan Labor Health and Welfare Organization. Personally identifiable information, including name, phone number, and permanent address, were not collected. Due to the nature of this study (an online survey), no written informed consent was obtained; however, receiving an answered questionnaire was considered evidence of consent.

Measures

Whiplash injury and LBP

Whiplash neck injury (cervical sprain and traumatic cervical syndrome) was defined as an injury in the neck, upper back, and shoulder area due to a motor vehicle accident. LBP was defined as pain localized between the costal margin and the inferior gluteal folds that persisted for more than a day at any time, based on the consensus approach for back pain definition proposed by Dionne et al. [16].

Table 1 Items in the Tampa Scales for Kinesiophobia

Item and description	
1	I am afraid that I might injure myself if I exercise
2	If I were to try to overcome it, my pain would increase
3	My body is telling me I have something dangerously wrong
4	My pain would probably be relieved if I were to exercise
5	People are not taking my medical condition seriously enough
6	My accident has put my body at risk for the rest of my life
7	Pain always means I have injured my body
8	Just because something aggravates my pain does not mean it is dangerous
9	I am afraid that I might injure myself accidentally
10	Simply being careful that I do not make any unnecessary movements is the safest thing I can do to prevent my pain from worsening
11	I would not have this much pain if there were not something potentially dangerous going on in my body
12	Although my condition is painful, I would be better off if I were physically active
13	Pain lets me know when to stop exercising so that I do not injure myself
14	It is really not safe for a person with a condition like mine to be physically active
15	I can not do all the things normal people do because it is too easy for me to get injured
16	Even though something is causing me a lot of pain, I do not think it is actually dangerous
17	No one should have to exercise when he/she is in pain

Items 1, 2, 3, 5, 6, 7, 10, 11, 13, 15, and 17 are TSK-11 items

Response choices: 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree

Pain associated only with menstrual periods, pregnancy, or during the course of a feverish illness was excluded. A diagram of affected areas by a whiplash injury and LBP was provided within the questionnaire.

The degree of the experienced pain associated with the whiplash injury or LBP was assessed using an 11-point numerical rating scale (NRS). Scores ranged from 0 (no pain) to 10 (worst pain imaginable), with a higher score indicating greater pain.

Catastrophizing

Pain catastrophizing, which is a maladaptive perception of pain, is an important predictor of future disability. Catastrophizing was assessed by the Japanese version of the pain catastrophizing scale (PCS) [17], a 13-item scale used to measure negative attitudes toward pain, involving rumination, helplessness, and magnification. The reliability and validity of the Japanese version were previously confirmed [17]. The total PCS score ranges from 0 (no catastrophizing) to 52 (severe catastrophizing).

Depressive mood

The presence of depressive mood was assessed using the mental health (MH) domain of the short-form health survey with 36 questions (SF-36) [18].

Somatic symptoms

Somatization was assessed using a subset of items from the brief symptom inventory (BSI). The Japanese version of the BSI-somatization scale was linguistically validated [19]. Seven somatic symptoms (faintness or dizziness, pains in the heart or chest, nausea or upset stomach, breathing difficulty, numbness or tingling in parts of the body, feeling weak in parts of the body, and hot or cold spells) were assessed on a 5-point scale (0, not at all; 1, a little bit; 2, moderate; 3, quite a bit; and 4, extreme).

General health status

The euroqol 5 dimension (EQ-5D) [20], which is a generic measure of health status that provides a simple descriptive profile and a single index value, was included in the questionnaire. The EQ-5D is a universally used tool to describe respondent's perception of his/her own health status. The index score derived from conversion of all responses ranges from −0.11 to 1.00, with a score of 1 denoting "perfect health" and a score of 0 denoting "death".

Data analysis

Demographic and clinical characteristics of the participants were summarized with descriptive statistics. Psychometric

properties were assessed with respect to both versions: TSK-J17 and TSK-J11 (Table 1). With regard to internal consistency, the homogeneity of the items in the TSK-J versions was evaluated using Cronbach's alpha statistics. A Cronbach's alpha coefficient of 0.7 or higher is required to claim that the TSK-J versions are internally consistent [21]. Concurrent validity was evaluated using the Pearson correlation coefficient with the PCS, EQ-5D, and pain NRS. Note that the Pearson correlation coefficient was used because the TSK employed a 4-point (1–4) Likert scale, under the assumption of an equally spaced distance between response choices. According to the criterion for correlation strength in the psychometric validation proposed by Cohen, the correlation coefficient was judged as follows: 0.1, weak correlation; 0.3, medium correlation; and 0.5, strong correlation [22]. For the known-group validity, relationships between selected variables and the subscale scores were examined using the *t* test or one-way analysis of variance (ANOVA). If one-way ANOVA showed there was a significant difference between groups, all pairwise comparisons between groups were conducted. Multiplicity of statistical tests was adjusted by the Tukey–Kramer method. We hypothesized that persons who met the following attributes would obtain higher TSK-J scores: (1) individuals with depressive mood, (2) individuals with more somatic symptoms, and (3) individuals with longer treatment periods. If an individual obtained a score of 52 or lower on the SF-36 Mental Health scale, he/she was considered to exhibit a "depressive mood" (score range 0–100, with lower scores indicating more psychological distress) [23]. With regard to somatic symptoms, if an individual answered 'moderate', 'quite a bit', or 'extremely' on a selected item of the BSI-somatization subscale, he/she was considered to have the somatic symptom described in the item. The number of somatic symptoms was divided into three categories: no symptom, one symptom, and two or more symptoms. The treatment period was divided into three categories: 3 months or less, 3–6 months, and 6 months or longer.

All statistical tests were two-tailed, and the level of significance was set at 0.05. Statistical calculations were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA).

Results

Characteristics of participants

Data from a total of 956 Japanese individuals who experienced a motor vehicle accident and an accompanying subsequent whiplash injury and/or LBP in the past 12 months were included in this analysis. The characteristics of the participants are shown in Table 2. The mean (SD) age was 45.4 (10.4) years; 71.0 % were male. The mean scores obtained on the TSK-J17 and the TSK-J11 were 41.1 (7.7) and 23.2 (6.6), respectively. Neither floor nor ceiling effect was observed. The mean total score for the PCS was 24.0

Table 2 Characteristics of the participants in the psychometric testing of the Japanese version of Tampa scale for Kinesiophobia ($N = 956$)

Characteristics	*n* (%)	Mean (SD)
Sex (*n*, %)		
Male	679 (71.0 %)	
Female	277 (29.0 %)	
Age, years		45.4 (10.4)
BMI (kg/m^2)		
Male		23.8 (3.6)
Female		21.3 (3.4)
Residual symptoms		
Yes	436 (45.6 %)	
No	520 (54.4 %)	
Duration to recovery ($n = 436$)		
Less than 4 weeks	230 (52.8 %)	
4–12 weeks	115 (26.4 %)	
12–24 weeks	65 (14.9 %)	
24 weeks or longer	25 (6.0 %)	
Work missed (*n*, %)		
None	321 (33.6 %)	
Less than 1 week	401 (40.9 %)	
1–4 weeks	118 (12.4 %)	
4–12 weeks	65 (6.8 %)	
12 weeks or longer	51 (5.3 %)	
TSK-J17		41.1 (7.7)
TSK-J11		23.2 (6.6)
PCS total score		24.0 (11.8)
Rumination		11.4 (4.9)
Helplessness		7.4 (5.0)
Magnification		5.2 (3.0)
EQ-5D		0.82 (0.18)
MH subscale score of SF-36		56.9 (19.7)
Scores of 52 or lower	420 (43.9 %)	
NRS for whiplash neck injury pain		6.1 (2.5)
NRS for LBP		4.8 (2.9)

TSK Tampa Scales for Kinesiophobia (score range 17–68 for the TSK-17 and 11–44 for the TSK-11, a higher score indicates stronger fear avoidance beliefs or behaviors), *PCS* Pain Catastrophizing scale (score range 0–52, a higher score indicates stronger catastrophizing), *EQ-5D* Euroqol 5 Dimension (score range −0.11 to 1.0 on a scale where 0.0 = death and 1.0 = perfect health), *MH* Mental Health (score range: 0–100, a lower score indicates more psychological distress), *SF-36* Short-Form Health Survey with 36 questions, *LBP* Low back pain, *NRS* Numerical rating scale (score range 0–10, a higher score indicates greater pain)

Values are *n* (%) or mean (SD)

Table 3 Pearson's correlation coefficients between the Tampa Scale for Kinesiophobia and other related variables

	PCS					NRS for whiplash injury pain	NRS for LBP
	Total score	Rumination	Helplessness	Magnification	EQ-5D		
TSK-J17	0.674	0.616	0.607	0.613	−0.583	0.380	0.393
TSK-J11	0.680	0.635	0.602	0.610	−0.570	0.394	0.401

All correlation coefficients are $p < 0.0001$

PCS Pain Catastrophizing Scale, *EQ-5D* Euroqol 5 Dimension, *NRS* Numerical rating scale, *LBP* Low back pain

(11.8). The mean score for the MH domain of the SF-36 was 56.9 (19.7), and scores of 52 or lower were observed in 43.9 % ($n = 420$) of individuals. The mean EQ-5D score was 0.82 (0.18). The mean NRSs for whiplash injury and LBP were 6.1 (2.5) and 4.8 (2.9), respectively. Absence of work or housework due to whiplash injury or LBP was observed in 66.4 % ($n = 635$) of individuals. Of these, 36.9 % ($n = 234$) had to miss work more than once per week.

Reliability

Cronbach's alpha coefficient was 0.850 for the TSK-J17 and 0.919 for the TSK-J11, indicating sufficient internal consistency.

Concurrent validity

The correlations of the TSK-J versions with the PCS, EQ-5D, and whiplash and LBP NRSs were calculated to examine concurrent validity. Both the TSK-J17 and TSK-J11 correlated strongly with the PCS total score, rumination, helplessness, and magnification subscales ($r = 0.674$, 0.616, 0.607, and 0.613 for the TSK-J17, respectively; $r = 0.680$, 0.635, 0.602, and 0.610 for the TSK-J11, respectively; $P < 0.0001$, for all) (Table 3).

Both the TSK-J17 and TSK-J11 negatively correlated moderately with the EQ-5D ($r = -0.583$ and -0.570, respectively; $p < 0.0001$). Both the TSK-J17 and TSK-J11 correlated moderately with the NRS for whiplash injury pain ($r = 0.380$ and 0.394, respectively; $p < 0.0001$) and NRS for LBP ($r = 0.393$ and 0.401, respectively; $p < 0.0001$).

Known-group validity

The relationship between variables that may affect the TSK-J score was examined. As hypothesized, significantly higher TSK-J scores were observed in persons with depressive mood, more somatic symptom(s), and longer treatment periods (Fig. 1). for depressive mood, the TSK-J17 score was 38.3 (6.7) for the less depressive group and 44.7 (6.5) for the more depressive group ($p < 0.0001$). The TSK-J11 score was 21.0 (6.0) for the less depressive group and 26.0 (6.2) for the more depressive group ($p < 0.0001$).

With respect to the number of somatic symptoms, persons who had more somatic symptoms had significantly higher TSK-J scores. The TSK-J17 scores in persons with no somatic symptom, one somatic symptom, and two or more somatic symptoms were 35.3 (6.1), 38.7 (6.0), and 43.7 (7.3), respectively, with significant differences between groups ($p < 0.0001$ for all). The corresponding TSK-J11 scores were 18.0 (5.8), 21.3 (5.2), and 25.4 (5.8), respectively, with significant differences between groups ($p < 0.0001$ for all).

Individuals with a longer treatment period had significantly higher TSK-J scores. The TSK-J17 scores in persons with treatment periods shorter than 3 months, 3–6 months, and 6 months or longer were 37.6, 41.7, and 46.0, respectively. The corresponding TSK-J11 scores were 20.3, 23.8, and 26.9, respectively, with significant differences between groups ($p < 0.0001$ for all).

Discussion

Matsudaira et al. proposed a linguistically-validated Japanese version of the TSK [15], the linguistic validity of which was established by ensuring the conceptual equivalence between the original and its translation by following a standardized method for developing a translated questionnaire [24]. In the present study, we assessed its psychometric properties with regard to 956 Japanese individuals who had whiplash injury pain or LBP due to a motor vehicle accident. Based on the results for internal consistency, concurrent validity, and known-group validity, the Japanese version of the TSK-J17 and TSK-J11 is considered to be reliable and valid as a measure for assessing fear of movement for (re)injury.

As an index to assess reliability, a highly sufficient internal consistency, with a Cronbach's alpha statistic of 0.850 for the TSK-J17 and 0.919 for the TSK-J11, was demonstrated. Although a direct comparison is not appropriate due to the different characteristics of the adopted study populations, the Cronbach's alpha coefficients obtained in this study are higher, relative to results obtained from

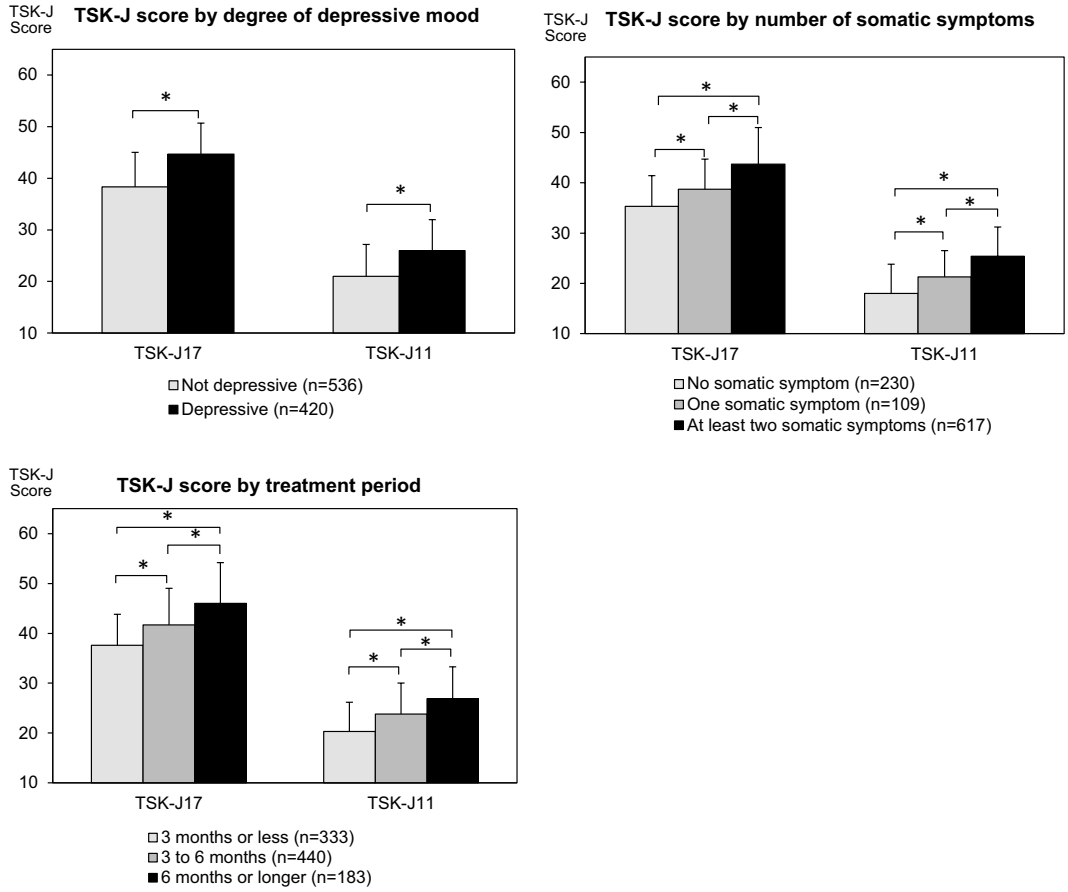


Fig. 1 Known-group validity: Tampa scale for Kinesiophobia scores and associated variables. *p* values were calculated by *t* test for depressive mood, and the Turkey—Kramer method was used to evaluate the number of somatic symptoms and treatment period. *TSK* Tampa Scale for Kinesiophobia

the psychometric testing of other language versions of the TSK-17 or -11, including the Persian (TSK-17; $\alpha = 0.77$–0.78 in acute chronic pain) [25], Chinese (TSK-17; $\alpha = 0.67$ in chronic pain) [26], Brazilian/Portuguese (TSK-17; $\alpha = 0.82$ in acute/subacute and chronic LBP) [27], German (TSK-11; $\alpha = 0.73$ in LBP) [12], Swedish (TSK-11; $\alpha = 0.74$–0.87 in chronic pain) [28], and Dutch (TSK-11; $\alpha = 0.68$–0.80 in acute and chronic LBP) [8] versions. Reasons remain unknown about the higher Cronbach's alpha coefficients obtained in this study; however, as the Cronbach's alpha is an indicator of the internal consistency of items, the results may indicate that the Japanese version was translated more successfully compared to other language versions. As another result to be noted, the TSK-J11 had higher alpha statistic than the TSK-J17 in this study. In general, a superior Cronbach's alpha statistic is obtained with inclusion of an increasing number of questions in the questionnaire. In this respect, the TSK-J11 presented not only the advantage of shorter length, but also that of higher internal consistency, which describes the extent to which all of the items in the test measure the same concept or construct, and hence is connected to the level of inter-relatedness of the items in the test.

Compared with previous studies [25, 29], both the TSK-J17 and TSK-J11 scores showed a stronger positive association with PCS, with a Pearson correlation coefficient above 0.6 for the PCS total and domain scores. There were also moderate associations with pain NRS in the affected area (0.380–0.401). The results were similar in TSK-J17 and TSK-J11. The obtained result of higher correlations with PCS, compared to NRSs may reflect that an individual's psychological perception toward pain, rather than degree of pain itself, may contribute to the development of a fear avoidance belief. For known-group validity, as hypothesized, relevance was exhibited between the TSK-J score and the variables that might affect the scores, including the presence of depressive mood, presence of somatic symptom(s), and duration of the treatment period. It should

be noted that these results do not necessarily imply a causal relationship between fear avoidance belief and the variables.

Fear avoidance behavior was reported to be an important risk factor for chronicity of pain and subsequent disability. In recent guidelines for the management of nonspecific acute LBP, continuing normal daily activities is recommended and bed rest is discouraged [30]. To help reduce pain-related fear, it is important not to focus on imaging findings that could lead to the development of fear avoidance behavior in patients, but to instruct them that pain is a common condition and is self-manageable, along with gradual exposure to activities. For this reason, detecting patient fear avoidance beliefs and encouraging them to change their beliefs and behaviors is of vital importance in the management of musculoskeletal pain, to achieve a better outcome. The TSK-J enables clinicians to detect a patient's fear avoidance beliefs, and helps to establish an effective management program to prevent chronic pain on an individual basis. In this study, the results of concurrent validity and known-group validity were similar for the TSK-J17 and TSK-J11; however, Cronbach's alpha coefficient was higher for the TSK-J11. This result may partly support the sufficiency of using the TSK-J11, in place of the TSK-J17. Moreover, due to its fewer number of questions, the TSK-J11 is more convenient for use in clinical settings, enabling shorter response times and a lower psychological burden on the patients.

There are several study limitations that should be noted. Our results were obtained in individuals who suffered from a motor vehicle accident; accordingly, findings may not be generalizable to other populations. For instance, suffering from a motor vehicle accident may have had a strong psychological impact on the painful experience of these individuals, possibly enhancing the development of fear avoidance beliefs. In addition, the use of an Internet panel to recruit participants could have contributed to a selection bias, although the large sample size collected, throughout the nation is a major strength of this study. Our strategy of using the Internet may invite criticism regarding the representativeness of the sample; however, taking into account both cost and feasibility, we decided to recruit participants via the Internet. As another limitation, it should be noted that factor structure was not analyzed in this study. The original TSK-J17 and the TSK-J11 are frequently used versions; however, we are concerned that different factor solutions were proposed in different language versions and differently targeted populations, potentially making it difficult to compare international data derived from different translated versions. In addition, test–retest reliability over certain time intervals remains unknown. Responsiveness cannot be assessed in the present study due to the cross-sectional nature of the data. Accordingly, future studies are necessary to address these issues.

In conclusion, the present psychometric analyses demonstrated that the Japanese version of the TSK is psychometrically reliable and valid as a measure of fear for movement in a Japanese population who had whiplash injury pain and/or LBP due to a motor vehicle accident. As the TSK-J11, a shorter version of the TSK-J17, showed better internal reliability and similar construction and known-group validity compared to the 17-item version, it may be more useful in routine clinical care, given a limited time for assessment.

Acknowledgments This study was supported by grants from JA Kyosai Research Institute and the Japan Labor Health and Welfare Organization.

Compliance with ethical standards

Conflict of interest NK is a board member of Clinical Study Support, Inc. and received grants to his institution from JA Kyosai Research Institute for this study. KM has no conflict of interests. TS has received grants to his institution from JA Kyosai Research Institute for this study. HO has received grants/grants pending to his institution from Pfizer Inc.

References

1. Leeuw M, Goossens ME, Linton SJ, Crombez G, Boersma K, Vlaeyen JW. The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. J Behav Med. 2007;30(1):77–94.
2. Vowles KE, Gross RT. Work-related beliefs about injury and physical capability for work in individuals with chronic pain. Pain. 2003;101(3):291–8.
3. Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. Pain. 2001;94(1):7–15.
4. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? JAMA. 2010;303(13):1295–302.
5. Walton DM, Pretty J, MacDermid JC, Teasell RW. Risk factors for persistent problems following whiplash injury: results of a systematic review and meta-analysis. J Orthop Sports Phys Ther. 2009;39(5):334–50.
6. Koho P, Aho S, Kautiainen H, Pohjolainen T, Hurri H. Test-retest reliability and comparability of paper and computer questionnaires for the Finnish version of the Tampa Scale of Kinesiophobia. Physiotherapy. 2014;100(4):356–62.
7. Rusu AC, Kreddig N, Hallner D, Hülsebusch J, Hasenbring MI. Fear of movement/(Re)injury in low back pain: confirmatory validation of a German version of the Tampa Scale for Kinesiophobia. BMC Musculoskelet Disord. 2014;19(15):280.
8. Swinkels-Meewisse EJ, Swinkels RA, Verbeek AL, Vlaeyen JW, Oostendorp RA. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. Man Ther. 2003;8(1):29–36.
9. Bunketorp L, Carlsson J, Kowalski J, Stener-Victorin E. Evaluating the reliability of multi-item scales: a non-parametric approach to the ordered categorical structure of data collected with the Swedish version of the Tampa Scale for Kinesiophobia and the Self-Efficacy Scale. J Rehabil Med. 2005;37(5):330–4.

10. Mintken PE, Cleland JA, Whitman JM, George SZ. Psychometric properties of the fear-avoidance beliefs questionnaire and Tampa Scale of Kinesiophobia in patients with shoulder pain. Arch Phys Med Rehabil. 2010;91(7):1128–36.
11. Visscher CM, Ohrbach R, van Wijk AJ, Wilkosz M, Naeije M. The Tampa Scale for Kinesiophobia for temporomandibular disorders (TSK-TMD). Pain. 2010;150(3):492–500.
12. Haugen AJ, Grøvle L, Keller A, Grotle M. Cross-cultural adaptation and validation of the Norwegian version of the Tampa scale for kinesiophobia. Spine (Phila Pa 1976). 2008;33(17):E595–601.
13. Burwinkle T, Robinson JP, Turk DC. Fear of movement: factor structure of the tampa scale of kinesiophobia in patients with fibromyalgia syndrome. J Pain. 2005;6(6):384–91.
14. Woby SR, Roach NK, Urmston M, Watson PJ. Psychometric properties of the TSK-11: a shortened version of the Tampa Scale for Kinesiophobia. Pain. 2005;117(1–2):137–44.
15. Matsudaira K, Inuzuka K, Kikuchi N, Sakae C, Arisaka M, Isomura T. Development of a Japanese version of the Tampa Scale for Kinesiophobia (TSK-J): translation and linguistic validation. Seikei Geka (Orthop surg). 2013;48(1):13–9 **(in Japanese)**.
16. Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, Wyatt M, Cassidy JD, Rossignol M, Leboeuf-Yde C, Hartvigsen J, Leino-Arjas P, Latza U, Reis S, Gil Del Real MT, Kovacs FM, Oberg B, Cedraschi C, Bouter LM, Koes BW, Picavet HS, van Tulder MW, Burton K, Foster NE, Macfarlane GJ, Thomas E, Underwood M, Waddell G, Shekelle P, Volinn E, Von Korff M. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976). 2008;33(1):95–103.
17. Matsuoka H, Sakano Y. Assessment of cognitive aspect of pain: development, reliability, and validation of japanese version of pain catastrophizing scale. Shinshin igaku (Japanese Journal of Psychosomatic Medicine). 2007;47(2):95–102 **(in Japanese)**.
18. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. J Clin Epidemiol. 1998;51(11):1037–44.
19. Matsudaira K, Inuzuka K, Kikuchi N, Sakae C, Arisaka M, Isomura T. Development of the Japanese version of the brief symptom inventory-somatization scale: translation and linguistic validation. Seikei Geka (Orthopedic surgery). 2012;63(2):149–53 **(in Japanese)**.
20. EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199–208.
21. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297–334.
22. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
23. Yamazaki S, Fukuhara S, Green J. Usefulness of five-item and three-item mental health Inventories to screen for depressive symptoms in the general population of Japan. Health Qual Life Outcomes. 2005;8(3):48.
24. Devellis RF. Scale development theory and applications. Newbury Park, CA: Sage Publications, Inc.; 1991.
25. Askary-Ashtiani A, Ebrahimi-Takamejani I, Torkaman G, Amiri M, Mousavi SJ. Reliability and validity of the Persian versions of the fear avoidance beliefs questionnaire and Tampa Scale of Kinesiophobia in patients with neck pain. Spine (Phila Pa 1976). 2014;39(18):E1095–102.
26. Wong WS, Kwok HY, Luk KD, Chow YF, Mak KH, Tam BK, Wong ET, Fielding R. Fear of movement/(re)injury in Chinese patients with chronic pain: factorial validity of the Chinese version of the Tampa Scale for Kinesiophobia. J Rehabil Med. 2010;42(7):620–9.
27. de Souza FS, Marinho Cda S, Siqueira FB, Maher CG, Costa LO. Psychometric testing confirms that the Brazilian-Portuguese adaptations, the original versions of the fear-avoidance beliefs questionnaire, and the Tampa Scale of Kinesiophobia have similar measurement properties. Spine (Phila Pa 1976). 2008;33(9):1028–33.
28. Larsson C, Hansson EE, Sundquist K, Jakobsson U. Psychometric properties of the Tampa Scale of Kinesiophobia (TSK-11) among older people with chronic pain. Physiother Theory Pract. 2014;30(6):421–8.
29. Hapidou EG, O'Brien MA, Pierrynowski MR, de Las Heras E, Patel M, Patla T. Fear and avoidance of movement in people with chronic pain: psychometric properties of the 11-item Tampa Scale for Kinesiophobia (TSK-11). Physiother Can. 2012 Summer;64(3):235–41.
30. Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. Eur Spine J. 2010;19(12):2075–94.

PLOS | ONE

RESEARCH ARTICLE

Risk Factors for Prolonged Treatment of Whiplash-Associated Disorders

Hiroyuki Oka[1]*, Ko Matsudaira[1], Tomoko Fujii[2], Hiroshi Okazaki[2], Yukari Shinkai[3], Yutaka Tsuji[3], Sakae Tanaka[4], Ryuichi Kato[3]

1 Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, University of Tokyo, Tokyo, Japan, **2** Department of Orthopaedic Surgery, Japan Labour, Health and Welfare Organization, Kanto Rosai Hospital, Kanagawa, Japan, **3** Medical Research Center, JA Kyosai Research Institute, Tokyo, Japan, **4** Department of Orthopaedic Surgery, University of Tokyo, Tokyo, Japan

* okah-tky@umin.ac.jp

Citation: Oka H, Matsudaira K, Fujii T, Okazaki H, Shinkai Y, Tsuji Y, et al. (2015) Risk Factors for Prolonged Treatment of Whiplash-Associated Disorders. PLoS ONE 10(7): e0132191. doi:10.1371/journal.pone.0132191

Editor: Antonio Verdejo-García, University of Granada, SPAIN

Received: March 4, 2015

Accepted: June 10, 2015

Published: July 6, 2015

Data Availability Statement: All relevant data are within the paper.

Funding: The authors received no specific funding for this work. Co-authors YS, YT and RK are employed by JA Kyosai Research Institute. JA Kyosai Research Institute provided support in the form of salaries for authors YS, YT and RK, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Abstract

Objectives

Whiplash-associated disorders (WAD) are the most common injuries that are associated with car collisions in Japan and many Western countries. However, there is no clear evidence regarding the potential risk factors for poor recovery from WAD. Therefore, we used an online survey of the Japanese population to examine the association between potential risk factors and the persistence of symptoms in individuals with WAD.

Materials and Methods

An online survey was completed by 127,956 participants, including 4,164 participants who had been involved in a traffic collision. A random sample of the collision participants (n = 1,698) were provided with a secondary questionnaire. From among the 974 (57.4%) respondents to the secondary questionnaire, we selected 183 cases (intractable neck pain that was treated over a period of 6 months) and 333 controls (minor neck pain that was treated within 3 months). Multivariable logistic regression analysis was used to evaluate the potential risk factors for prolonged treatment of WAD.

Results

Female sex, the severity of the collision, poor expectations of recovery, victim mentality, dizziness, numbness or pain in the arms, and lower back pain were associated with a poor recovery from WAD.

Conclusions

In the present study, the baseline symptoms (dizziness, numbness or pain in the arms, and lower back pain) had the strongest associations with prolonged treatment for WAD, although the psychological and behavioral factors were also important. These risk

Competing Interests: Co-authors YS, YT and RK are employed by JA Kyosai Research Institute. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

factors should be considered when evaluating patients who may have the potential for poor outcomes.

Introduction

Whiplash-associated disorders (WAD) are the most common injuries that are associated with car collisions in many Western countries [1] and in Japan [2]. Although the prognosis for WAD is generally favorable, previous studies have found that up to about 20% of patients experience persistent neck pain at 6 months after their injury [3,4]. Unfortunately, this lack of recovery creates personal, economic, and social burdens [1]. To reduce this burden, the number of individuals who develop chronic WAD must be reduced, although it is difficult to predict which patients will experience persistence of their symptoms. However, several prognostic factors have been identified, including sex [5,6], a low level of education [5,6], the severity of the collision [7], expectations of recovery [8], a no-fault claim [7], the presence of dizziness [9], upper extremity numbness or pain [10], and lower back pain [11–13]. Unfortunately, there is no clear evidence regarding the potential risk factors for poor recovery from WAD in the Japanese population. Based on this absence of suitable data, we conducted an online survey of the general Japanese population to identify individuals who had been in a car collision. Using the data from that survey, we examined the associations between the potential risk factors and the persistence of symptoms in individuals with WAD.

Materials and Methods

Sources of data

In 2012, we conducted an online survey to assess the prevalence of WAD in the general population. The participants were recruited through an internet research company that has approximately 1.8 million registered Japanese adult volunteers (20–79 years old). The company's volunteers are representative of the general Japanese population, and were stratified according to sex and age. From among these volunteers, 1,063,083 individuals were randomly selected and invited to participate in this study via an email that contained a unique link to the survey (dated July 1, 2012). Among these invited individuals, only 227,853 were considered effective users, as the research company was unable to exclude the non-users from the invitations due to technical reasons. The participants received points for online shopping as an incentive, and double registration was prevented by reviewing the participant's e-mail address at the beginning of the survey and disabling the link to the questionnaire at the conclusion of the survey. The initial survey was closed when the number of participants reached 127,956 (July 17, 2012). Thus, the response rate for the invitations was not relevant to this survey. This study's design was approved by the ethics review board of Kanto Rosai Hospital.

All participants completed the original questionnaire, which included items regarding their demographical and social characteristics, as well as any traffic collisions that they had experienced. However, for our analysis we only evaluated the questionnaires from participants who had been in a traffic collision (n = 4,164). From among this sample, 1,698 participants were randomly selected to participate in a secondary survey. Among the 974 respondents (57.4%) for the secondary questionnaire, we excluded 44 participants who were not wearing a seatbelt when the collision occurred, as these participants were likely to have sustained serious injuries. From the 930 remaining subjects, we included 183 participants in the cases group (neck pain that was treated over a period of 6 months) and 333 participants in the control group (minor

neck pain that was treated within 3 months) (Fig 1). We defined the self-reported presence of WAD in this study as a response to the internet questionnaire that indicated 1) an obvious instance of an injury that was sustained during a rear-end collision, or 2) an established diagnosis of WAD by a medical doctor.

Assessment

The questionnaire evaluated socio-demographic data, age, sex, weight, height, smoking, education level (not college, or college), the severity of the collision (high, or other; high severity was defined as the vehicle's bumpers exhibiting extensive damage after a rear-end collision). Body mass index (BMI; k/m^2) was calculated using the participant's self-reported weight and height. Expectations of recovery were evaluated by asking "Do you expect that your neck pain will be a problem in the next 3 months?", using response categories of "No", "Possibly", "Probably", and "Definitely". Poor expectations of recovery were defined as answers of "Probably" or "Definitely". We also used the question "Did you have any fault in this accident?" to identify participants with a "victim mentality" (i.e., an answer of "no"). The presence of dizziness (yes/no) was evaluated using the question "Did you have any dizziness in the week after this accident?", and numbness or pain in the arms was evaluated using the question "Did you have any numbness or pain in your arms in the week after this accident?" Lower back pain was defined as pain that lasted for >1 day in the area between the lower costal margin and the gluteal folds, regardless of any accompanying radiating pain, and that was not associated with febrile illness, menstruation, or pregnancy [14].

Statistical analysis

The preliminary survey was administered to 10,000 participants for sample size estimation. Our preliminary study revealed that 16 of the 10,000 participants were assigned to the case group. 2) As our dependent variable was binary, we decided to use logistic regression analysis, because we needed a 1:2 case:control ratio. One guideline has suggested that the accurate estimation of discriminant function parameters requires a sample size with at least 20 cases for each independent variable in the logistic regression [15]. Therefore, based on this guideline and our 10 predictor variables, we required 200 cases for our analysis. Thus, the survey was closed at approximately 125,000 participants, although slightly more than 125,000 participants were included, due to technical reasons.

We compared the characteristics of the cases and controls using the chi-square test for categorical variables, and the one-factor analysis of variance for numerical variables. Age, sex, BMI, smoking, education level, severity of collision, poor expectation of recovery, victim mentality, dizziness, numbness or pain in arms, and lower back pain have previously been identified as risk factors for a poor recovery from WAD [5–13]. Therefore, we entered these variables into the multivariable logistic regression model, in order to adjust for potential confounding. The Variance Inflation Factor (VIF) was used to check for multicollinearity in the model. All statistical tests were performed at a significance level of 0.05 (two-sided), and were not adjusted for multiple testing. All data analyses were performed using SAS software (version 9.1.3, SAS Institute Inc., Cary, NC).

Results

Table 1 shows the demographic characteristics of the participants. When we compared the case and control groups, we observed significant differences in the severity of the collision, poor expectations of recovery, dizziness, upper extremity numbness or pain, and lower back pain.

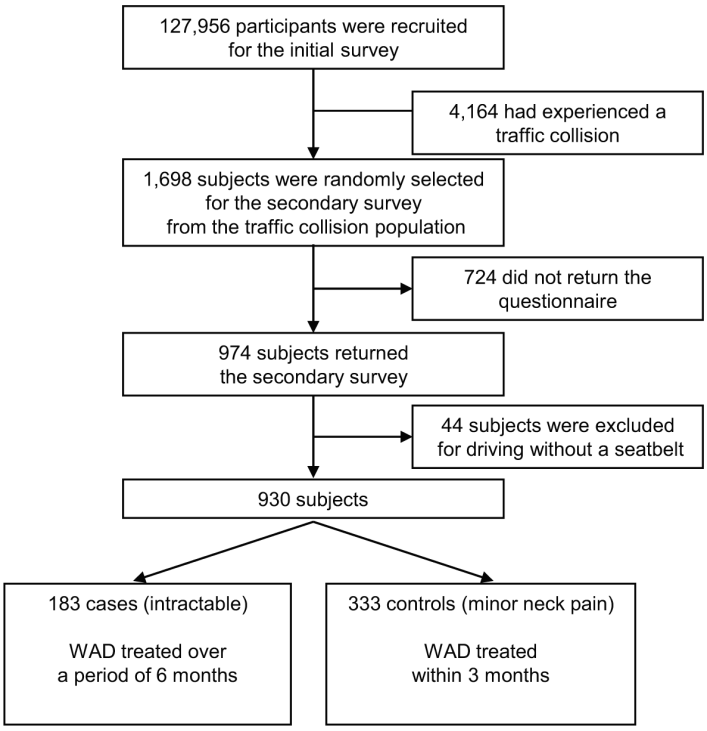


Fig 1. Study flow chart. WAD, whiplash-associated disorders.

doi:10.1371/journal.pone.0132191.g001

However, no significant differences were observed for age, sex, BMI, smoking, and a low level of education.

Table 2 shows the results from the univariate logistic regression analysis for a poor recovery from WAD. Based on the results of this analysis, we found that female sex, the severity of the collision, poor expectations of recovery, victim mentality, dizziness, numbness or pain in the

Table 1.

	Cases (n = 183)	Controls (n = 333)	p-value
Age, years	44.8 ± 10.3	45.3 ± 11.7	0.6218
Sex, male/female	124/59	242/91	0.2397
BMI, kg/m^2	23.4 ± 4.0	23.0±3.7	0.1971
Smoking (%)	74 (36.6)	128 (38.4)	0.6563
Education level: not college (%)	57 (31.2)	99 (29.7)	0.7373
Severity of collision: high (%)	131 (71.6)	159 (47.9)	<0.0001
Poor expectation of recovery (%)	90 (49.2)	41 (12.3)	<0.0001
Victim mentality (%)	150 (83.0)	253 (76.0)	0.1154
Dizziness (%)	120 (65.6)	94 (28.2)	<0.0001
Numbness or pain in arm (%)	149 (81.4)	170 (51.1)	<0.0001
Low back pain (%)	113 (61.2)	74 (22.2)	<0.0001

BMI, body mass index.

doi:10.1371/journal.pone.0132191.t001

Table 2.

	Odds ratio	95% CI	p-value
Age, +1 year	1	0.99–1.02	0.6209
Female (vs. male)	1.26	0.85–1.87	0.2417
BMI (+1 kg/m^2)	0.97	0.92–1.02	0.1983
Smoking	0.92	0.64–1.33	0.6566
Education level: not college	1.06	0.72–1.58	0.7376
Severity of collision: high	2.76	1.88–4.08	<0.0001
Poor expectation of recovery	6.89	4.48–10.76	<0.0001
Victim mentality	1.44	0.92–2.28	0.1114
Dizziness	4.84	3.30–7.17	<0.0001
Numbness or pain in arms	4.2	2.76–6.54	<0.0001
Lower back pain	5.65	3.82–4.82	<0.0001

CI, confidence interval; BMI, body mass index.

doi:10.1371/journal.pone.0132191.t002

arms, and lower back pain were significantly associated with a poor recovery from WAD. Table 3 shows the results from the multivariable logistic regression analysis, after adjusting for the various confounding factors. The VIF values for age, sex, BMI, smoking, education level, severity of collision, poor expectation of recovery, victim mentality, dizziness, numbness or pain in arms, and lower back pain were 1.12, 1.12, 1.14, 1.03, 1.19, 1.17, 1.16, 1.26, 1.23, and 1.24, respectively. However, none of the VIF values exceeded 10, which indicates that there was no collinearity in the model [16]. Based on the results of this model, we found that female sex, the severity of the collision, poor expectations of recovery, victim mentality, dizziness, numbness or pain in the arms, and lower back pain were significantly associated with a poor recovery from WAD.

Discussion

To the best of our knowledge, this is the first study to examine the risk factors that are associated with a prolonged recovery among Japanese patients with WAD. Our final model identified seven risk factors (female sex, the severity of the collision, poor expectations of recovery, victim mentality, presence of dizziness, numbness or pain in the arms, and lower back pain); all of these factors have previously been reported to be independent prognostic factors for recovery from WAD [5–13].

Interestingly, it is not clear which sex is an independent risk factor for poor recovery from WAD, as several studies have reported that female sex was an independent predictor, while others have reported that male sex was an independent predictor. In addition, previous studies have reported that a low level of education was significantly related to a poor recovery [5,6]. However, in the present study, education level was not a significant risk factor for a poor recovery from WAD. Unfortunately, the reasons for these discrepancies between our findings and those of the previous studies are not clear, although they may be related to differences in the populations that were studied.

We also observed that the severity of the collision was an important risk factor for poor recovery from WAD. In this context, a whiplash injury occurs when the force of a rear-end collision "whips" the cervical spine beyond its normal range of motion. Therefore, it is logical that severe car crashes can cause serious damage to the musculoskeletal system, which can result in a poor recovery.

Table 3.

	Odds ratio	95% CI	p-value
Age, +1 year	1	0.98–1.03	0.7577
Female (vs. male)	1.83	1.07–3.17	0.0283
BMI (+1 kg/m^2)	1.07	0.99–1.14	0.0576
Smoking	0.95	0.58–1.57	0.8515
Education level: not college	1.11	0.67–1.85	0.6819
Severity of collision: high	1.97	1.19–3.30	0.0086
Poor expectation of recovery	4.47	2.68–7.53	<0.0001
Victim mentality	3.37	1.76–6.67	0.0002
Dizziness	3.12	1.93–6.00	<0.0001
Numbness or pain in arms	2.56	1.51–4.40	0.0004
Lower back pain	4.77	2.91–7.94	<0.0001

CI, confidence interval; BMI, body mass index.

doi:10.1371/journal.pone.0132191.t003

After adjusting for the relevant confounders, such as socio-demographic characteristics and symptoms, we observed that poor expectations of recovery and victim mentality were significant risk factors for a poor recovery. Similarly, previous studies have reported that expectations for recovery were an important factor in the prognosis for WAD recovery [5]. Therefore, in addition to understanding these injuries and their clinical symptoms, it is also important to understand the patient's perception of recovery, in order to adequately treat WAD. Furthermore, victim mentality is an aspect of the patient's perception, and may affect their expectations for recovery. This finding indicates that psychological factors have prognostic value for evaluating the risk of prolonged recovery from WAD.

A previous study has reported that dizziness, numbness in the arms, and lower back pain did not decrease within 6 months after the accident, although many other symptoms were transient [13]. Similarly, we observed that these symptoms (dizziness, numbness, and lower back pain) were independent risk factors for a prolonged recovery from WAD. Therefore, it appears that these symptoms are more common in severe cases, which are less likely to experience recovery within 6 months. Furthermore, dizziness, numbness, and lower back pain are known as somatic symptom, and patients who have chronic whiplash also report elevated levels of somatic symptoms in body areas that were not affected by their neck trauma [17, 18]. In this context, the symptoms of functional somatic syndromes are very similar to those of somatization disorder, and the two conditions are thought to be closely related [19–21]. Thus, it is important to consider these signs and symptoms when following-up patients who have experienced whiplash. Furthermore, although the baseline symptoms (dizziness, numbness, and lower back pain) had the strongest associations with prolonged treatment for WAD, the psychological and behavioral factors were also important, and these risk factors should also be considered when evaluating patients who have experienced whiplash.

This study has several limitations. First, due to the cross-sectional design, inferences cannot be made regarding the causality of the relationships. Second, the sample was selected from among internet research volunteers, who may not be representative of the general population of internet users. Third, compared to the general population, our sample contained a higher proportion of people who were living in large cities and who had completed university-level or graduate-level education [22]. Fourth, we surveyed the respondents after their traffic collisions, and it is plausible that some reported symptoms may have been preexisting, rather than caused

by the traffic collision. Furthermore, there are other important factors that can affect recovery from WAD, such as coping styles, previous traffic injuries, comorbidities, somatic and psychological pre-injury health, pain intensity and disability, injustice perception, depression and pain-related emotions, social support, personality traits, and post-traumatic stress symptoms. However, these factors were not included because we needed to evaluate the information from at the time of injury as a prognostic factor. Therefore, recall bias may be present, given the interval between the injury and the administration of the validated questionnaires. In addition, we attempted to ensure that the full questionnaire could be completed in 10 min, in order to obtain complete data from the respondents. Unfortunately, the effect of this selection bias on our findings would be difficult to address. Despite these limitations, this study provides useful insight for medical and public health practitioners who treat patients who have experienced whiplash.

Author Contributions

Conceived and designed the experiments: H. Oka KM ST. Performed the experiments: TF H. Okazaki. Analyzed the data: H. Oka KM. Contributed reagents/materials/analysis tools: YS YT RK. Wrote the paper: H. Oka KM.

References

1. Cassidy JD, Carroll LJ, Côté P, Lemstra M, Berglund A, Nygren A. Effect of eliminating compensation for pain and suffering on the outcome of insurance claims for whiplash injury. N Engl J Med 2000; 342: 1179–1186. PMID: 10770984
2. Yayama T, Kokubo Y, Uchida K, et al. Pathophysiology of the traumatic cervical spine syndrome. Seikeigeka 2012; 63: 797–801. [Japanese]
3. Radanov BP, Sturzenegger M, Di Stefano G, Schnidrig A, Aljinovic M. Factors influencing recovery from headache after common whiplash. BMJ 1993; 307: 652–655. PMID: 8401050
4. Radanov BP, Begres, Sturzenegger M, Augustiny KF. Course of psychological variables in whiplash injury—a 2-year follow-up with age, gender and education pair-matched patients. Pain 1996; 64: 429–434. PMID: 8783306
5. Ozegovic D, Carroll LJ, Cassidy JD. Factors associated with recovery expectations following vehicle collision: a population-based study. J Rehabil Med 2010; 42: 66–73. doi: 10.2340/16501977-0466 PMID: 20111847
6. Ozegovic D, Carroll LJ, Cassidy JD. What influences positive return to work expectation? Examining associated factors in a population-based cohort of whiplash-associated disorders. Spine 2010; 35: E708–E713. doi: 10.1097/BRS.0b013e3181d12432 PMID: 20535047
7. Wiles NJ, Jones GT, Silman AJ, Macfarlane GJ. Onset of neck pain after a motor vehicle accident: a case-control study. J Rheumatol 2005; 32: 1576–1583. PMID: 16078337
8. Carroll LJ. Beliefs and expectations for recovery, coping, and depression in whiplash-associated disorders: lessening the transition to chronicity. Spine (Phila Pa 1976) 2011; 36: S250–S256.
9. Treleaven J. Dizziness, unsteadiness, visual disturbances, and postural control: implications for the transition to chronic symptoms after a whiplash trauma. Spine (Phila Pa 1976) 2011; 36: S211–S217.
10. Scott D, Jull G, Sterling M. Widespread sensory hypersensitivity is a feature of chronic whiplash-associated disorder but not chronic idiopathic neck pain. Clin J Pain 2005; 21: 175–181. PMID: 15722811
11. Bohman T, Côté P, Boyle E, Cassidy JD, Carroll LJ, Skillgate E. Prognosis of patients with whiplash-associated disorders consulting physiotherapy: development of a predictive model for recovery. BMC Musculoskelet Disord 2012; 13: 264. doi: 10.1186/1471-2474-13-264 PMID: 23273330
12. Crutebo S, Nilsson C, Skillgate E, Holm LW. The course of symptoms for whiplash-associated disorders in Sweden: 6-month followup study. J Rheumatol 2010; 37: 1527–1533. doi: 10.3899/jrheum.091321 PMID: 20472922
13. Walton DM, Macdermid JC, Giorgianni AA, Mascarenhas JC, West SC, Zammit CA. Risk factors for persistent problems following acute whiplash injury: update of a systematic review and meta-analysis. J Orthop Sports Phys Ther 2013; 43: 31–43. doi: 10.2519/jospt.2013.4507 PMID: 23322093

14. Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976) 2008; 33: 95–103.
15. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis with Readings. 5th edition. Englewood Cliffs, NJ: Prentice-Hall; 1998.
16. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis. 3rd edition. New York: Macmillan; 1995.
17. Berglund A, Alfredsson L, Jensen I, Cassidy JD, Nygren A. The association between exposure to a rear-end collision and future health complaints. J Clin Epidemiol 2001; 54: 851–856. PMID: 11470396
18. Miettinen T, Lindgren KA, Airaksinen O, Leino E. Whiplash injuries in Finland: A prospective 1-year follow-up study. Clin Exp Rheumatol 2002; 20: 399–402. PMID: 12102479
19. Barsky AJ, Borus JF. Functional somatic syndromes. Ann Intern Med 1999; 130: 910–921. PMID: 10375340
20. World Health Organization: The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research. Geneva: World Health Organization; 1993.
21. De Gucht V, Fischler B, Heiser W. Job stress, personality, and psychological distress as determinants of somatization and functional somatic syndromes in a population of nurses. Stress Heal 2003; 19: 195–204.
22. Japan's Population Census and Labour Force Survey. 2007. Available: http://www.stat.go.jp/data/index.htm. Accessed 4 October 2011. [Japanese]

AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 58:1300–1310 (2015)

Disabling Low Back Pain Associated With Night Shift Duration: Sleep Problems as a Potentiator

Masaya Takahashi, PhD,[1]* Ko Matsudaira, MD, PhD,[2,3] and Akihito Shimazu, PhD[4]

Background *We investigated how night shift duration and sleep problems were jointly associated with disabling low back pain (LBP) among workers in different occupations.*
Methods *An online-survey was conducted regarding work schedules, disabling LBP, sleep problems, and other relevant factors in 5,008 workers who were randomly selected from a market research panel. Multiple logistic regression analyses determined the joint associations of night shift duration (0 [permanent day shift], <8, 8–9.9, 10–15.9, ≥16 hr) and sleep problems (no, yes) with disabling LBP adjusted for potential confounders.*
Results *A night shift ≥16 hr was associated with a significant increase in the likelihood of disabling LBP. The magnitude of this association was elevated when participants perceived sleep problems including both sleep duration and quality.*
Conclusion *Associations between extended night shifts and disabling LBP became stronger in the presence of short or poor quality sleep.* Am. J. Ind. Med. 58:1300–1310, 2015. © 2015 Wiley Periodicals, Inc.

KEY WORDS: musculoskeletal disorders; shift schedules; sleep duration; insomnia symptoms

INTRODUCTION

Low back pain (LBP) represents a major health and safety problem in workplaces worldwide [Driscoll et al., 2014]. The situation is serious in Japan as well, where LBP accounts for approximately 60% of occupational injuries requiring absences of 4 days or more among Japanese workers [Japanese Ministry of Health, Labour and Welfare 2013]. Importantly, this problem affects a wide range of industries: 30% of occupational LBP cases are identified in health care, 19% in commerce, and financial advertising, 15% in manufacturing, 14% in transportation and traffic, 6% in customer entertainment, and 5% in construction.

A number of occupational variables have been found to act as causal or exacerbating factors in LBP [Yassi et al., 2013; Matsudaira et al., 2014]. While the two dominant factors, heavy physical work and high psychosocial demands, have been well recognized, evidence for the effects of other occupational factors is largely limited [da Costa et al., 2010]. Recent research highlights the essential role of work schedules in musculoskeletal disorders (MSDs) [Caruso et al., 2008]. Shift work involving night shifts, in particular, is shown to be a target factor in many cases [Eriksen et al., 2004; Takahashi et al., 2008]. The principal component of its burden relates to night shift duration [Rosa et al., 1997; Ferguson et al., 2012]. Previous studies compared low back problems between 8- and 12-hr night shifts [Yamada et al., 2001; He et al., 2011], or between 8- and 16-hr night shifts [Takahashi et al., 1999] for some selected occupations, mainly health care professionals. Preliminary evidence clearly requires that a more detailed

[1]National Institute of Occupational Safety and Health, Tama-ku, Kawasaki, Japan
[2]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan
[3]Clinical Research Center for Occupational Musculoskeletal Disorders, Kanto Rosai Hospital, Nakahara-ku, Kawasaki, Japan
[4]Department of Mental Health, The University of Tokyo, Graduate School of Medicine, Bunkyo-ku, Tokyo, Japan
Contract grant sponsor: Japan Labor Health and Welfare Organization; Contract grant sponsor: National Institute of Occupational Safety and Health, Japan; Contract grant number: P20-05.
*Correspondence to: Masaya Takahashi, PhD, National Institute of Occupational Safety and Health, 6-21-1, Nagao, Tama-ku, Kawasaki 214-8585, Japan.
E-mail: takaham@h.jniosh.go.jp

Accepted 27 May 2015
DOI 10.1002/ajim.22493. Published online 29 June 2015 in Wiley Online Library (wileyonlinelibrary.com).

investigation into the association between night shift duration and LBP be undertaken in a more systematic manner, examining a variety of occupations. Increased knowledge about the effects of night shift duration on LBP can facilitate the optimization of shift work so as to minimize issues related to LBP.

Working at night has been linked with unfavorable consequences in the health, safety, and well-being of workers [Caruso, 2014]. Notably, shift work involving night shifts can disturb sleep [Wright et al., 2013; Takahashi 2014]. Recently published findings indicate the close association between sleep problems and pain [Buxton et al., 2012; Finan et al., 2013a]. Musculoskeletal pain, including LBP, can be a source or predictor of insomnia [Tang et al., 2012; Tang et al., 2015]. Conversely, poor quality of sleep is reported to predict the incidence of LBP in healthy workers [Agmon et al., 2014] and to be associated with a subsequent increase in LBP intensity in patients [Alsaadi et al., 2014b]. Prospective evidence demonstrates that disturbed sleep is significantly associated with an elevated risk of sickness absence and disability retirement due to MSDs [Salo et al., 2012; Ropponen et al., 2013]. Moreover, recent efforts have been devoted to clarification of the brain circuits shared by both sleep and pain [Finan et al., 2013b; Koh et al., 2015].

Associations among the three variables—LBP, sleep problems, and night shift duration—are highly complicated, as demonstrated in previous studies on LBP, sleep problems, and a third factor (e.g., job strain and physical activity) [Canivet et al., 2008; Sorensen et al., 2011]. One common approach to clarifying the associations is to determine if sleep problems lie in the causal pathway between night shift duration and LBP. Results obtained will be meaningful in understanding the potential mechanisms for these three factors. Another possible strategy is to examine if sleep problems have moderating effects on the association between night shift duration and LBP. If this association is modified according to the conditions of sleep, such a finding would provide us with novel information about sleep-related options for LBP prevention among night shift workers.

The present study examined how night shift duration and sleep problems were jointly associated with LBP in a sample of workers in different occupations. We hypothesized that a longer night shift would be associated with an increased likelihood of LBP and that this association would be strengthened with sleep problems. Testing these hypotheses has scientific merit, because little data are available for a dose-response relation of night shift duration with LBP, and because interventions to achieve better sleep may be possible for reducing or preventing LBP. Our research also focused on differences in the associations according to subtypes of sleep problems.

METHODS

Participants

The details of participant recruitment have been reported in a previous paper [Matsudaira et al., 2013]. Briefly, potential participants were selected randomly from a market research panel according to the inclusion criteria: age (20–69 years old) and residential area (23 wards of Tokyo, the City of Osaka, and the City of Nagoya). A total of 5,917 workers completed a web-based questionnaire. The final sample was 5,856 participants after exclusion of those who reported age of either below 20 or beyond 65 years old and those who reported working in primary and secondary industries. This study included 5,008 participants who provided their work schedules. The medical/ethics review board of the Japan Labour Health and Welfare Organization reviewed and approved this study.

Measures

Work schedules

Participants were asked if they engaged in permanent day work, rotating shift work involving night shifts, or other shifts. Participants with rotating shift work also responded to a question about the duration of the night shift: <8, 8–9.9, 10–11.9, 12–13.9, 14–15.9, or 16 hr or longer.

Disabling LBP

LBP was assessed with the question, "How would you describe your LBP in the past year?" Response options included (1) no LBP, (2) LBP that did not interfere with work, (3) LBP that interfered with work but no absence from work, and (4) LBP that interfered with work, leading to sick-leave. A diagram with a shaded area was presented to help participants correctly understand the site of the low back. LBP was defined as pain occurring in the area between the lower costal margin and the gluteal folds. LBP must also have lasted more than one day, and occurred regardless of accompanying radiating pain, but it must not be associated merely with febrile illness, menstrual periods, or pregnancy [Dionne et al., 2008]. LBP was classified as disabling if it caused disruption to the job regardless of absence from work (i.e., positive response to the option 3 or 4) [Von Korff et al., 1992]. Disabling LBP was the outcome of interest in this study.

Sleep problems

Sleep problems were evaluated using questions about the quantity and quality of sleep in the past month [Nakata et al., 2005; Takahashi et al., 2008]. Short sleep duration

was defined as sleep duration of less than 6 hr. Difficulty initiating sleep was defined as taking more than 30 min to fall asleep. Difficulty maintaining sleep and early morning awakening were defined as nocturnal awakenings or early morning awakenings occurring 3 times or more per week. Insomnia symptoms were considered to be present if the participants reported any of the 3 symptoms of insomnia above.

Covariates

We collected self-reported data on age, gender, employment (permanent, other), occupation (white-collar [managers, professionals, clerical workers, sales workers], blue-collar [service, production, security, transportation, and communications workers], other), main work contents (work with video display terminals [VDT], physically repetitive work, neither), weekly work hours (<40, 40–49, 50–59, ≥60 hr), education (high school or lower, university, or higher), regular exercise (no, yes), smoking status (never smoker, former smoker, current smoker), chronic conditions requiring doctor visits (present, not present), height, and weight. The questionnaire also measured psychosocial work characteristics with the Brief Job Stress Questionnaire [Shimomitsu et al., 2000] for job demand, job control, and worksite (supervisor and coworker) social support.

Statistical Analysis

The duration of a night shift was re-classified as 0 (permanent day work; n = 4,691), <8 (n = 100), 8–9.9 (n = 90), 10–15.9 (n = 82), and ≥16 (n = 45) hr according to its distribution. Associations between night shift duration and the study variables were examined using a chi-square test and analysis of variance. Joint associations of night shift duration and sleep problems with disabling LBP were analyzed using logistic regression models with a reference group of permanent day workers without sleep problems. The first model provided crude odds ratios (ORs) and 95% confidence intervals (CIs) for the joint associations. The second model adjusted for age, gender, employment, occupation, main work contents, weekly work hours, education, and smoking status. The third model further adjusted for psychosocial work characteristics. In addition, tests for linear trend were conducted to examine the dose-response relationship between the categories of night shift duration and disabling LBP. Given the small sample size in each group of shift workers, those four groups were collapsed into a single, shift-working group. Data from the shift work group have been listed in parallel. All statistical analyses were conducted using IBM SPSS Statistics version 20 (IBM Corporation, New York).

RESULTS

Characteristics of Study Sample

As summarized in Table I, both the permanent day workers and the shift workers showed a similar gender ratio, with the majority of men among the groups with night shifts of 8–10 and 10–16 hr. The shift workers engaged in more blue-collar jobs with a higher degree of physically repetitive work compared to the permanent day workers. The shift workers were also younger than the permanent day workers, except for those with a night shift of less than 8 hr.

It should be noted that the 16 hr or longer night shift group reported disabling LBP at a rate almost double (42%) that of the other groups (18–23%; $P < 0.01$). Although the percentage reporting sleep duration of less than 6 hr was comparable among the five groups, the shift groups reported insomnia symptoms more often than the permanent day group. Similar differences were observed for each subtype of insomnia symptoms. The shift group working a night shift of 16 hr or more reported higher job demand ($P < 0.001$) and lower job control ($P < 0.001$) compared to the other groups, while they showed a greater level of worksite social support ($P = 0.019$).

Comparisons between the permanent day group and the shift work group revealed results similar to those obtained from comparisons between the permanent day group and the four groups of shift workers. These two groups, however, showed no significant differences in disabling LBP, early morning awakening, or social support at work.

Disabling LBP Associated With Night Shift Duration by Sleep Problem

Table II indicates that the 16 hr or longer night shift group with short sleep duration was more likely to report disabling LBP. This significant association was observed even after adjusting for several confounding factors (Model 3: OR 4.90, 95%CI 2.18–11.03). The permanent day workers also reported more disabling LBP if they experienced short duration of sleep. However, this association was not statistically significant after adjusting for psychosocial work characteristics. The corresponding tests for linear trend became statistically non-significant in Model 3 ($P = 0.135$). No significant associations were observed between working shifts and disabling LBP in the cases of sleep duration of less than 6 hr or greater than 6 hr.

When insomnia symptoms were present, both the permanent day workers (Model 3: 1.42, 1.20–1.68) and the shift workers who worked at night for 16 hr or longer (6.59, 2.35–18.49) showed statistically significant ORs (Table III, P for linear trend = 0.140). The shift work group also produced a significant association (1.78, 1.13–2.80).

TABLE I. Characteristics of Study Participants

		Duration of a night shift						
	Day	<8h	8–10 h	10–16 h	≥16 h		Shift	
	n (%)	n (%)	n (%)	n (%)	n (%)	P[a]	n (%)	P[b]
Gender								
Men	2378 (51)	51 (51)	66 (73)	60 (73)	26 (58)	0.001	203 (64)	0.001
Women	2313 (49)	49 (49)	24 (27)	22 (27)	19 (42)		114 (36)	
Employment								
Permanent	2410 (51)	32 (32)	55 (61)	55 (67)	39 (87)	0.001	181 (57)	0.048
Others	2281 (49)	68 (68)	35 (39)	27 (33)	6 (13)		136 (43)	
Occupation								
White-collar	3336 (71)	30 (30)	33 (37)	31 (38)	17 (38)	0.001	111 (35)	0.001
Blue-collar	923 (20)	52 (52)	49 (54)	42 (51)	20 (44)		163 (51)	
Others	432 (9)	18 (18)	8 (9)	9 (11)	8 (18)		43 (14)	
Main work contents								
VDT work	2498 (53)	18 (18)	18 (20)	22 (27)	11 (24)	0.001	69 (22)	0.001
Physically repetitive work	885 (19)	44 (44)	49 (54)	38 (46)	20 (44)		151 (48)	
Neither	1308 (28)	38 (38)	23 (26)	22 (27)	14 (31)		97 (31)	
Weekly work hours (hours)								
−40 h	1642 (35)	51 (51)	12 (13)	14 (17)	6 (13)	0.001	83 (26)	0.001
40–49 h	1989 (42)	43 (43)	63 (70)	43 (52)	25 (56)		174 (55)	
50–59 h	655 (14)	4 (4)	13 (14)	12 (15)	8 (18)		37 (12)	
≥60 h	405 (9)	2 (2)	2 (2)	13 (16)	6 (13)		23 (7)	
Education								
High school or lower	2676 (57)	80 (80)	60 (67)	59 (72)	30 (67)	0.001	229 (72)	0.001
University or higher	2003 (43)	20 (20)	30 (33)	23 (28)	15 (33)		88 (28)	
Regular exercise								
No	3610 (77)	78 (78)	71 (79)	62 (76)	29 (64)	0.370	240 (76)	0.611
Yes	1081 (23)	22 (22)	19 (21)	20 (24)	16 (36)		77 (24)	
Smoking								
Non-smoke	2687 (57)	64 (64)	44 (49)	44 (54)	21 (47)	0.029	173 (55)	0.013
Past smoke	845 (18)	12 (12)	12 (13)	14 (17)	6 (13)		44 (14)	
Current smoke	1159 (25)	24 (24)	34 (38)	24 (29)	18 (40)		100 (32)	
Chronic conditions that require doctor visits								
Present	1297 (28)	31 (31)	20 (22)	20 (24)	13 (29)	0.672	84 (26)	0.657
Disabling low back pain	915 (20)	18 (18)	21 (23)	16 (20)	19 (42)	0.004	74 (23)	0.097
Sleep problems								
Sleep duration <6 hours	2062 (44)	44 (44)	39 (43)	42 (51)	25 (56)	0.389	150 (47)	0.243
Insomnia symptoms	1087 (23)	34 (34)	26 (29)	20 (24)	16 (36)	0.001	96 (30)	0.004
Difficulty initiating sleep	827 (18)	23 (23)	23 (26)	15 (18)	14 (31)	0.022	75 (24)	0.007
Difficulty maintaining sleep	328 (7)	17 (17)	10 (11)	7 (9)	5 (11)	0.001	39 (12)	0.004
Early morning awakening	316 (7)	14 (14)	6 (7)	5 (6)	4 (9)	0.028	29 (9)	0.101
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)		Mean (SD)	
Age (years)	44.8 (12.5)	44.8 (14.0)	39.0 (11.6)	40.6 (12.4)	40.0 (11.5)	0.001	41.4 (12.8)	0.001
BMI	22.6 (3.5)	21.7 (3.3)	22.6 (3.6)	22.6 (3.9)	23.3 (3.2)	0.092	22.4 (3.6)	0.495
Job demand	7.7 (2.3)	7.0 (2.4)	8.5 (2.0)	8.7 (2.0)	9.1 (2.3)	0.001	8.2 (2.3)	0.001
Job control	8.0 (2.2)	7.1 (2.2)	7.0 (2.2)	6.8 (2.0)	6.8 (2.3)	0.001	6.9 (2.1)	0.001
Worksite social support	15.1 (4.2)	15.4 (3.5)	14.1 (3.6)	14.4 (3.7)	16.4 (4.1)	0.019	14.9 (3.7)	0.608

Day: Permanent day workers. Shift: All shift workers. BMI: Body mass index.
a. Compared between permanent day workers and 4 groups of shift workers.
b. Compared between permanent day workers and all shift workers.
Statistical significance was tested using chi-square test for categorical data and using analysis of variance for continuous data.

TABLE II. Joint Associations of Night Shift Duration/Shift Work and Sleep Duration of Less Than 6 hr (no, yes) With Disabling Low Back Pain (N = 5,008)

	Disabling LBP		Model 1		Model 2		Model 3	
	n	(%)	OR	95%CI	OR	95%CI	OR	95%CI
Day, n	481	(18.3)	1.00		1.00		1.00	
<8h, n	10	(17.9)	0.97	0.49–1.94	0.96	0.48–1.93	1.01	0.50–2.03
−10h, n	11	(21.6)	1.23	0.63–2.42	1.16	0.59–2.30	1.08	0.55–2.15
−16h, n	10	(25.0)	1.49	0.72–3.07	1.34	0.64–2.78	1.28	0.61–2.66
≥16h, n	5	(25.0)	1.49	0.54–4.12	1.37	0.49–3.81	1.20	0.43–3.39
P for linear trend				0.172		0.287		0.437
Shift, n	36	(21.6)	1.23	0.84–1.80	1.15	0.78–1.70	1.12	0.75–1.65
Day, y	434	(21.0)	1.19	1.03–1.37	1.16	1.00–1.35	1.12	0.97–1.31
<8 h, y	8	(18.2)	0.99	0.46–2.15	0.95	0.44–2.08	0.95	0.43–2.09
−10 h, y	10	(25.6)	1.54	0.75–3.19	1.32	0.63–2.75	1.18	0.56–2.46
−16 h, y	6	(14.3)	0.75	0.31–1.78	0.66	0.27–1.58	0.58	0.24–1.40
≥16 h, y	14	(56.0)	5.69	2.57–12.62	5.13	2.29–11.49	4.90	2.18–11.03
P for linear trend				0.025		0.097		0.135
Shift, y	38	(25.3)	1.52	1.04–2.22	1.36	0.92–2.01	1.27	0.86–1.87

n. sleep duration ≥6 h, y. sleep duration <6 h.
Day: Permanent day workers. Shift: All shift workers.
Model 1. Crude.
Model 2. Adjusted for age, gender, employment, occupation, main work contents, weekly work hours, education, and smoking status.
Model 3. Adjusted for Model 2+job demand, job control, and workplace social support

TABLE III. Joint Associations of Night Shift Duration/Shift Work and Insomnia Symptoms (no, yes) With Disabling Low Back Pain (N = 5,008)

	Disabling LBP		Model 1		Model 2		Model 3	
	n	(%)	OR	95%CI	OR	95%CI	OR	95%CI
Day, n	648	(18.0)	1.00		1.00		1.00	
<8 h, n	8	(12.1)	0.63	0.30–1.33	0.64	0.30–1.35	0.67	0.31–1.42
−10 h, n	14	(21.9)	1.28	0.70–2.33	1.20	0.65–2.19	1.10	0.60–2.01
−16 h, n	13	(21.0)	1.22	0.66–2.25	1.12	0.60–2.09	1.06	0.56–1.98
≥16 h, n	9	(31.0)	2.06	0.93–4.55	1.89	0.85–4.21	1.72	0.77–3.86
P for linear trend				0.119		0.281		0.408
Shift, n	44	(19.9)	1.14	0.81–1.60	1.08	0.76–1.53	1.03	0.73–1.47
Day, y	267	(24.6)	1.49	1.26–1.75	1.46	1.24–1.72	1.42	1.20–1.68
<8 h, y	10	(29.4)	1.91	0.91–4.01	1.81	0.86–3.84	1.83	0.86–3.88
−10 h, y	7	(26.9)	1.69	0.71–4.03	1.49	0.62–3.60	1.41	0.58–3.41
−16 h, y	3	(15.0)	0.81	0.24–2.77	0.66	0.19–2.29	0.59	0.17–2.04
≥16 h, y	10	(62.5)	7.63	2.76–21.08	6.80	2.43–18.97	6.59	2.35–18.49
P for linear trend				0.045		0.078		0.140
Shift, y	30	(31.3)	2.08	1.34–3.23	1.86	1.19–2.92	1.78	1.13–2.80

n. insomnia symptoms were not present, y. insomnia symptoms were present.
Day: Permanent day workers. Shift: All shift workers.
Model 1. Crude.
Model 2. Adjusted for age, gender, employment, occupation, main work contents, weekly work hours, education, and smoking status.
Model 3. Adjusted for Model 2+job demand, job control, and workplace social support.

As listed in Table IV, the OR for reporting disabling LBP was greater in the 16 hours or more night shift group with difficulty initiating sleep (Model 3: OR 5.35, 95%CI 1.82–15.68, P for linear trend = 0.091) than in those without difficulty initiating sleep (1.99, 0.93–4.24). Even the permanent day group showed a significantly increased OR with difficulty initiating sleep.

Having difficulty maintaining sleep was significantly associated with increased likelihood of reporting disabling LBP among both the permanent day workers and the shift workers, except the 10–16 hr night shift group (Table V). A clear contrast was observed when comparing between the 16 hr or longer night shift group with (Model 3: OR 13.85, P for linear trend = 0.179) and without difficulty maintaining sleep (OR 2.28), though the former's 95%CI was large.

Results for the joint association between night shift duration and early morning awakening (Table VI) were similar to results in terms of difficulty initiating sleep. Specifically, the OR approached significance for disabling LBP in the 16 hr or longer night shift group who experienced early morning awakening (Model 3: OR 9.75, P for linear trend = 0.077), which far exceeded the OR in their counterpart who did not experience early morning awakening (2.36). The permanent day workers who experienced early morning awakening were significantly more likely to report disabling LBP.

If any of the subtypes of insomnia symptoms were present, shift work was significantly associated with disabling LBP (ORs: 1.70–3.49); otherwise, no associations were found to be significant (Tables IV to VI).

DISCUSSION

The present study indicated that an extended night shift, particularly beyond 16 hr, was associated with a significant increase in the likelihood of disabling LBP. The magnitude of this association increased when participants perceived sleep problems. Indeed, in terms of sleep duration, a significantly greater OR for disabling LBP was obtained with sleep of less than 6 hr. In terms of sleep quality, a significantly increased OR was found for cases where insomnia symptoms were reported. Analyses for subtypes of insomnia symptoms revealed varying results. The ORs for disabling LBP were greater among the permanent day group and the 16 or longer hours of night shift group when those groups had difficulty initiating sleep. Similar results were observed for early morning awakening. However, overall increases in the ORs for disabling LBP were found when difficulty maintaining sleep was present. Our data for the shift work group, if they experienced sleep problems, principally reflected the significant results of the 16 hours or longer night shift group, with the exception of short sleep duration.

The present finding of increased disabling LBP associated with a night shift of 16 hr or longer calls for an appropriate design of shift schedules. We have to consider

TABLE IV. Joint Associations of Night Shift Duration / Shift Work and Difficulty Initiating Sleep (no, yes) With Disabling Low Back Pain (N = 5,008)

	Disabling LBP		Model 1		Model 2		Model 3	
	n	(%)	OR	95%CI	OR	95%CI	OR	95%CI
Day, n	720	(18.6)	1.00		1.00		1.00	
<8 h, n	13	(16.9)	0.89	0.49–1.62	0.89	0.48–1.64	0.94	0.51–1.73
−10 h, n	14	(20.9)	1.16	0.64–2.10	1.07	0.59–1.96	1.00	0.55–1.83
−16 h, n	13	(19.4)	1.06	0.57–1.94	0.96	0.52–1.77	0.89	0.48–1.65
≥16 h, n	11	(35.5)	2.41	1.15–5.05	2.22	1.05–4.69	1.99	0.93–4.24
P for linear trend			0.104		0.290		0.453	
Shift, n	51	(21.1)	1.17	0.85–1.61	1.10	0.79–1.53	1.06	0.76–1.47
Day, y	195	(23.6)	1.35	1.13–1.61	1.35	1.12–1.62	1.32	1.10–1.59
<8 h, y	5	(21.7)	1.22	0.45–3.29	1.19	0.44–3.24	1.18	0.43–3.22
−10 h, y	7	(30.4)	1.92	0.79–4.68	1.75	0.71–4.30	1.56	0.63–3.86
−16 h, y	3	(20.0)	1.10	0.31–3.89	0.94	0.26–3.38	0.89	0.25–3.18
≥16h, y	8	(57.1)	5.84	2.02–16.89	5.26	1.80–15.39	5.35	1.82–15.68
P for linear trend			0.038		0.048		0.091	
Shift, y	23	(30.7)	1.94	1.18–3.19	1.78	1.07–2.95	1.70	1.02–2.82

n. difficulty initiating sleep was not present, y. difficulty initiating sleep was present.
Day: Permanent day workers. Shift: All shift workers.
Model 1. Crude.
Model 2. Adjusted for age, gender, employment, occupation, main work contents, weekly work hours, education, and smoking status.
Model 3. Adjusted for Model 2 + job demand, job control, and workplace social support.

TABLE V. Joint Associations of Night Shift Duration/Shift Work and Difficulty Maintaining Sleep (no, yes) With Disabling Low Back Pain (N = 5,008)

	Disabling LBP		Model 1		Model 2		Model 3	
	n	(%)	OR	95%CI	OR	95%CI	OR	95%CI
Day, n	808	(18.5)	1.00		1.00		1.00	
<8 h, n	10	(12.0)	0.61	0.31–1.18	0.60	0.31–1.18	0.62	0.31–1.21
−10 h, n	16	(20.0)	1.10	0.64–1.92	1.02	0.58–1.78	0.94	0.54–1.65
−16 h, n	14	(18.7)	1.01	0.56–1.82	0.92	0.51–1.67	0.88	0.49–1.60
≥16 h, n	15	(37.5)	2.65	1.39–5.05	2.42	1.26–4.66	2.28	1.18–4.41
P for linear trend				0.080		0.223		0.381
Shift, n	55	(19.8)	1.09	0.80–1.48	1.02	0.75–1.39	0.98	0.72–1.35
Day, y	107	(32.6)	2.14	1.68–2.73	2.02	1.58-2.58	1.92	1.50–2.46
<8 h, y	8	(47.1)	3.93	1.51–10.21	3.70	1.41-9.73	3.96	1.50–10.43
−10 h, y	5	(50.0)	4.42	1.28–15.30	3.59	1.03-12.56	3.59	1.01-12.68
−16 h, y	2	(28.6)	1.77	0.34–9.13	1.32	0.25–6.89	1.00	0.19–5.26
≥16 h, y	4	(80.0)	17.67	1.97–158.34	14.61	1.62–131.67	13.85	1.50–127.83
P for linear trend				0.046		0.171		0.179
Shift, y	19	(48.7)	4.20	2.23–7.90	3.60	1.90–6.83	3.49	1.83–6.66

n. difficulty maintaining sleep was not present, y. difficulty maintaining sleep was present.
Day: Permanent day workers. Shift: All shift workers.
Model 1. Crude.
Model 2. Adjusted for age, gender, employment, occupation, main work contents, weekly work hours, education, and smoking status.
Model 3. Adjusted for Model 2+job demand, job control, and workplace social support.

multiple characteristics of shift schedules, in addition to shift duration [Ferguson et al., 2012; Harris et al., 2015]. Extension of a night shift would be allowed as long as a variety of conditions inside and outside the workplace are optimized [Knauth, 2007]. Challenges have recently been proposed to the use of a night shift of more than 8 hr in light of workers' health and productivity [Hopcia et al., 2012; Sallinen et al., 2010; Griffiths et al., 2014]. Our data reported

TABLE VI. Joint Associations of Night Shift Duration/Shift Work and Early Morning Awakening (no, yes) With Disabling Low Back Pain (N = 5,008)

	Disabling LBP		Model 1		Model 2		Model 3	
	n	(%)	OR	95%CI	OR	95%CI	OR	95%CI
Day, n	823	(18.8)	1.00		1.00		1.00	
<8 h, n	13	(15.1)	0.77	0.43–1.40	0.76	0.42–1.39	0.78	0.43–1.43
−10 h, n	18	(21.4)	1.18	0.70–2.00	1.08	0.64–1.85	1.00	0.59–1.71
−16 h, n	14	(18.2)	0.96	0.54–1.72	0.86	0.47–1.55	0.81	0.45–1.47
≥16 h, n	16	(39.0)	2.77	1.47–5.21	2.52	1.33–4.79	2.36	1.24–4.51
P for linear trend				0.046		0.178		0.313
Shift, n	61	(21.2)	1.16	0.87–1.56	1.08	0.80–1.46	1.04	0.76–1.40
Day, y	92	(29.1)	1.76	1.36–2.27	1.65	1.27–2.13	1.59	1.23–2.07
<8h, y	5	(35.7)	2.40	0.80–7.19	2.26	0.75–6.82	2.43	0.80–7.38
−10h, y	3	(50.0)	4.33	0.87–21.48	3.17	0.63–15.88	3.11	0.61–15.93
−16h, y	2	(40.0)	2.89	0.48–17.30	2.36	0.39–14.20	1.80	0.30–10.90
≥16h, y	3	(75.0)	12.99	1.35–124.99	10.36	1.07–100.33	9.75	0.99–95.57
P for linear trend				0.038		0.058		0.077
Shift, y	13	(44.8)	3.52	1.69–7.34	3.00	1.42–6.30	2.92	1.38v6.19

n. early morning awakening was not present, y. early morning awakening was present.
Day: Permanent day workers. Shift: All shift workers.
Model 1. Crude.
Model 2. Adjusted for age, gender, employment, occupation, main work contents, weekly work hours, education, and smoking status.
Model 3. Adjusted for Model 2+job demand, job control, and workplace social support.

here support this view, requiring critical evaluation of working at night for 16 hours or more to prevent disabling LBP.

The present data highlight the important role of sleep problems in disabling LBP linked to the duration of a night shift. When sleep was short (<6 hr), there was a significantly higher likelihood of disabling LBP. This finding may be due in part to a reduced threshold of pain consequent to sleep restriction, as shown in experimental studies [Ødegård et al., 2015; Roehrs et al., 2012]. If insomnia symptoms existed, similar potentiating effects were evident. These results are consistent with our hypothesis and in turn provide novel insight into the triad: night shift duration, disabling LBP, and sleep problems, given the previously reported dyad of LBP and sleep problems [Finan et al., 2013a; Kelly et al., 2011].

Because of limited evidence on the triad mentioned above and the nature of the present study design, it is difficult to describe how those three variables are associated with each other. Nevertheless, at least two hypotheses can be presented. First, a long night shift may interfere with sleep [Takahashi et al., 2008]; the problems in sleep may then translate into greater LBP. This association is possible, since sleep disturbance can affect the autonomic, neuroendocrine, and neuroimmunologic systems to provoke inflammatory response, delayed recovery of tissue damage, and increased pain sensitivity [Heffner et al., 2011; Garland, 2012; Roehrs et al., 2012; Mertens et al., 2015; Ødegård et al., 2015]. Second, an extended night shift may be associated with higher LBP via increased/prolonged exposure to physical (mechanical) and mental workload during the long period of shift [Katsuhira et al., 2013; Sterud et al., 2013; Coenen et al., 2014]; the elevated level of LBP is likely to impair subsequent sleep. Further research is warranted for better understanding of the complex relationship among night shift duration, disabling LBP, and sleep problems. A cohort study of workers with different lengths of a night shift (e.g., 8, 12, 16 hr) will be needed to test the hypothesis that night shift workers without sleep problems at baseline show night-shift dose-dependent increases in disabling LBP at follow-up if they suffer from sleep problems during the follow-up period. In contrast to this observational strategy, an intervention study can be conducted to examine the hypothesis that participants with different night shift durations who report sleep problems at baseline exhibit dose-dependent decreases in disabling LBP at follow-up after treatments for their sleep problems, compared to after placebo treatments.

When examining differences in the three subtypes of insomnia symptoms, the 16 hr or longer night shift group consistently showed greater ORs for disabling LBP if they experienced any difficulty initiating or maintaining sleep or waking too early from sleep. However, significantly higher ORs were also found in the shift groups working at night for less than 10 hr who reported frequent nocturnal awakenings (OR = 3.59–3.96). The current observation suggests that difficulty maintaining sleep may serve as a potentiating factor for disabling LBP among shift workers. Alternatively, frequent awakenings during sleep could be a target when addressing disabling LBP associated with night shifts.

In the present study, participants working a night shift of 16 hr or more (16+h group) reported higher job demand, lower job control, and higher social support at work compared to the other groups (Table I). The fact that the 16+h group showed greater LBP despite increased social support is contradictory to the common notion of increased LBP with low social support [Lang et al., 2012; Lundberg, 2015]. Possibly, the 16+h group may have managed the longer night shift while receiving more support from their supervisors and colleagues.

Our present results need to be considered in view of the study's limitations. As an initial stage of investigation, a cross-sectional design was used in this study to examine the associations among night shift duration, sleep problems, and disabling LBP. Thus, it was not possible to test the temporal relationship of the study variables. The study sample, derived from workers registered with a market research company, was not representative of the general working population in Japan. Hence, particular caution is required in generalizing our findings. Data collection by an online survey also may have caused several types of bias. Clearly, workers without access to the internet are never able to become participants. Potential candidates are less likely to participate in the survey if they suffer from severe LBP. Individuals with long and/or demanding work also tend to miss the opportunity to respond. This sampling bias may have been reflected in the smaller proportion of shift workers (6%, n = 317) in the current study than that reported by the national survey (18%) [Japanese Ministry of Health, Labour and Welfare, 2007]. In particular, the shift group working a night shift of 16 hr or more consisted of only 45 workers. The low number of shift-work participants may have resulted in missed clear dose-response relationships between night shift duration (four categories) and disabling LBP. Tests for linear trend were not found to be statistically significant among the groups with sleep problems ($P = 0.077$–0.179 in Model 3), although consistently increased ORs of disabling LBP were observed for the 16+h group with sleep problems. It was unable to identify the duration of night shift shorter than 16 hours at which the risk of disabling LBP increased in our study. Given the bias and confounding due to the small sample sizes, the joint associations of night shift duration and sleep problems with disabling LBP found here can be taken as selected and preliminary, and less reliable (i.e., wide confidence intervals for the 16+h group). Furthermore, recall time frames were different between sleep problems (the past month) and disabling LBP (the past year). Prevalence has been reported to be higher at 1 year than at one month for both variables [Matsudaira et al., 2011; Morin et al., 2014]. The longer time

frame of recall on the sleep scale may have identified more cases, which in turn may have resulted in increased stability of estimates. At the same time, however, an increased possibility of recall error has to be traded off. Alternatively, if disabling LBP was asked to be reported in the frame of the same past month as sleep problems, the risk estimates of the sleep problems might have been attenuated due to fewer cases of disabling LBP than the present ones. It is thus preferable to use the same time frames of recall among the scales employed in the future study. All data were self-reported, and it is desirable to assess sleep using an objective method such as wrist actigraphy, even in a subsample [Schuh-Hofer et al., 2013; Alsaadi et al., 2014a]. Additionally, while our previous study excluded participants who had worked in their current job for less than one year [Matsudaira et al., 2013], the present study did not. The associations reported here were found to be consistent with those obtained from the dataset when we excluded the participants with less than one year of work experience (n = 4,222, data not shown). This fact implies that the current results were not confounded by the effects of those specific participants.

The current findings inform us of some practical implications. First of all, adequate scheduling for shift work involving night shifts should be a priority for LBP protection. There is a debate regarding the cost and benefit of extended night shifts [Smith et al., 1998; Lockley et al., 2004; Ferguson et al., 2012]. However, special care should be exercised when implementing a night shift longer than 8 hours, particularly for occupations characterized by greater levels of biomechanical and psychosocial demands (e.g., health care workers). This recommendation is still valid, assuming that in the present study, shift workers with an 8- to 16-h night shift did not show significant ORs for disabling LBP due to small sample sizes. In addition, various sources of workload need to be reduced as much as possible by active use of ergonomic devices, adequate staffing, and planned napping during the night shift. For health care practices in the workplace, treating sleep problems or promoting sleep health can be a promising strategy to reduce the level of LBP and to protect against the new onset of LBP [Eadie et al., 2013; Finan et al., 2014]. At an organizational level, employers are expected to revise multiple aspects of the work environment to ensure adequate sleep of their employees [Takahashi, 2012].

In conclusion, the present findings suggest that a night shift of 16 hr or longer was associated with a greater risk of disabling LBP and that the increased risk was further elevated among workers experiencing short or poor quality sleep.

ACKNOWLEDGMENTS

This work was supported by a dissemination project on the 13 fields of occupational injuries and illness of the Japan Labor Health and Welfare Organization and by the National Institute of Occupational Safety and Health, Japan (P20-05).

AUTHORS CONTRIBUTION

Conception, design, and data acquisition of this study: Ko Matsudaira; Analysis and interpretation of data: Masaya Takahashi and Ko Matsudaira; Drafting of manuscript: Masaya Takahashi; Critical revision: Masaya Takahashi, Ko Matsudaira, Akihito Shimazu; Final approval of the version to be published: All authors.

REFERENCES

Agmon M, Armon G. 2014. Increased insomnia symptoms predict the onset of back pain among employed adults. PLoS One 9:e103591.

Alsaadi SM, McAuley JH, Hush JM, Bartlett DJ, McKeough ZM, Grunstein RR, Dungan GC, 2nd, Maher CG. 2014a. Assessing sleep disturbance in low back pain: The validity of portable instruments. PLoS One 9:e95824.

Alsaadi SM, McAuley JH, Hush JM, Lo S, Lin CW, Williams CM, Maher CG. 2014b. Poor sleep quality is strongly associated with subsequent pain intensity in patients with acute low back pain. Arthritis Rheumatol 66:1388–1394.

Buxton OM, Hopcia K, Sembajwe G, Porter JH, Dennerlein JT, Kenwood C, Stoddard AM, Hashimoto D, Sorensen G. 2012. Relationship of sleep deficiency to perceived pain and functional limitations in hospital patient care workers. J Occup Environ Med 54:851–858.

Canivet C, Ostergren PO, Choi B, Nilsson P, af Sillén U, Moghadassi M, Karasek R, Isacsson SO. 2008. Sleeping problems as a risk factor for subsequent musculoskeletal pain and the role of job strain: Results from a one-year follow-up of the Malmo Shoulder Neck Study Cohort. Int J Behav Med 15:254–262.

Caruso CC 2014. Negative impacts of shiftwork and long work hours. Rehabil Nurs 39:16–25.

Caruso CC, Waters TR. 2008. A review of work schedule issues and musculoskeletal disorders with an emphasis on the healthcare sector. Ind Health 46:523–534.

Coenen P, Gouttebarge V, van der Burght AS, van Dieen JH, Frings-Dresen MH, van der Beek AJ, Burdorf A. 2014. The effect of lifting during work on low back pain: A health impact assessment based on a meta-analysis. Occup Environ Med 71:871–877.

da Costa BR, Vieira ER. 2010. Risk factors for work-related musculoskeletal disorders: A systematic review of recent longitudinal studies. Am J Ind Med 53:285–323.

Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, Wyatt M, Cassidy JD, Rossignol M, Leboeuf-Yde C, et al. 2008. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976) 33:95–103.

Driscoll T, Jacklyn G, Orchard J, Passmore E, Vos T, Freedman G, Lim S, Punnett L. 2014. The global burden of occupationally related low back pain: Estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 73:975–981.

Eadie J, van de Water AT, Lonsdale C, Tully MA, van Mechelen W, Boreham CA, Daly L, McDonough SM, Hurley DA. 2013. Physiotherapy for sleep disturbance in people with chronic low back

pain: Results of a feasibility randomized controlled trial. Arch Phys Med Rehabil 94:2083–2092.

Eriksen W, Bruusgaard D, Knardahl S. 2004. Work factors as predictors of intense or disabling low back pain; a prospective study of nurses' aides. Occup Environ Med 61:398–404.

Ferguson SA, Dawson D. 2012. 12-h or 8-h shifts? It depends. Sleep Med Rev 16:519–528.

Finan PH, Buenaver LF, Runko VT, Smith MT. 2014. Cognitive-behavioral therapy for comorbid insomnia and chronic pain. Sleep Med Clin 9:261–274.

Finan PH, Goodin BR, Smith MT. 2013a. The association of sleep and pain: An update and a path forward. J Pain 14:1539–1552.

Finan PH, Smith MT. 2013b. The comorbidity of insomnia, chronic pain, and depression: Dopamine as a putative mechanism. Sleep Med Rev 17:173–183.

Garland EL 2012. Pain processing in the human nervous system: A selective review of nociceptive and biobehavioral pathways. Prim Care 39:561–571.

Griffiths P, Dall'Ora C, Simon M, Ball J, Lindqvist R, Rafferty AM, Schoonhoven L, Tishelman C, Aiken LH. 2014. Nurses' shift length and overtime working in 12 European countries: The association with perceived quality of care and patient safety. Med Care 52: 975–981.

Harris R, Sims S, Parr J, Davies N. 2015. Impact of 12h shift patterns in nursing: A scoping review. Int J Nurs Stud 52:605–634.

He C, Davis KG. 2011. Impact of shift work on physical and postural demands among nursing aides in long-term health care facilities. Proc Hum Factors Ergon Soc Annu Meet 55:1007–1011.

Heffner KL, France CR, Trost Z, Ng HM, Pigeon WR. 2011. Chronic low back pain, sleep disturbance, and interleukin-6. Clin J Pain 27: 35–41.

Hopcia K, Dennerlein JT, Hashimoto D, Orechia T, Sorensen G. 2012. Occupational injuries for consecutive and cumulative shifts among hospital registered nurses and patient care associates: A case-control study. Workplace Health Saf 60:437–444.

Japanese Ministry of Health, Labour and Welfare. 2007. Survey on the State of Employees' Health (in Japanese).

Japanese Ministry of Health, Labour and Welfare. 2013. Survey on occupational injuries and illnesses (in Japanese).

Katsuhira J, Matsudaira K, Iwakiri K, Kimura Y, Ohashi T, Ono R, Sugita S, Fukuda K, Abe S, Maruyama H. 2013. Effect of mental processing on low back load while lifting an object. Spine (Phila Pa 1976) 38:E832–E839.

Kelly GA, Blake C, Power CK, O'Keeffe D, Fullen BM. 2011. The association between chronic low back pain and sleep: A systematic review. Clin J Pain 27:169–181.

Knauth P 2007. Extended work periods. Ind Health 45:125-136 (Erratum: Ind Health 45:375).

Koh K, Hamada A, Hamada Y, Yanase M, Sakaki M, Someya K, Narita M, Kuzumaki N, Ikegami D, Sakai H, et al. 2015. Possible involvement of activated locus coeruleus-noradrenergic neurons in pain-related sleep disorders. Neurosci Lett 589:200–206.

Lang J, Ochsmann E, Kraus T, Lang JW. 2012. Psychosocial work stressors as antecedents of musculoskeletal problems: A systematic review and meta-analysis of stability-adjusted longitudinal studies. Soc Sci Med 75:1163–1174.

Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CJ, Landrigan CP, Rothschild JM, Katz JT, Lilly CM, Stone PH, et al. 2004. Effect of reducing interns' weekly work hours on sleep and attentional failures. N Engl J Med 351:1829–1837.

Lundberg U, 2015. Work conditions and back pain problems. Stress Health 31:1–4.

Matsudaira K, Konishi H, Miyoshi K, Isomura T, Inuzuka K. 2014. Potential risk factors of persistent low back pain developing from mild low back pain in urban Japanese workers. PLoS One 9:e93924.

Matsudaira K, Palmer KT, Reading I, Hirai M, Yoshimura N, Coggon D. 2011. Prevalence and correlates of regional pain and associated disability in Japanese workers. Occup Environ Med 68:191–196.

Matsudaira K, Shimazu A, Fujii T, Kubota K, Sawada T, Kikuchi N, Takahashi M. 2013. Workaholism as a risk factor for depressive mood, disabling back pain, and sickness absence. PLoS One 8: e 75140.

Mertens P, Blond S, David R, Rigoard P. 2015. Anatomy, physiology and neurobiology of the nociception: A focus on low back pain (part A). Neurochirurgie 61:S22–S34.

Morin CM, Leblanc M, Ivers H, Belanger L, Merette C, Savard J, Jarrin DC. 2014. Monthly fluctuations of insomnia symptoms in a population-based sample. Sleep 37:319–326.

Nakata A, Ikeda T, Takahashi M, Haratani T, Fujioka Y, Fukui S, Swanson NG, Hojou M, Araki S. 2005. Sleep-related risk of occupational injuries in Japanese small and medium-scale enterprises. Ind Health 43:89–97.

Ødegård SS, Omland PM, Nilsen KB, Stjern M, Gravdahl GB, Sand T. 2015. The effect of sleep restriction on laser evoked potentials, thermal sensory and pain thresholds and suprathreshold pain in healthy subjects. Clin Neurophysiol 126:1979–1987.

Roehrs TA, Harris E, Randall S, Roth T. 2012. Pain sensitivity and recovery from mild chronic sleep loss. Sleep 35:1667–1672.

Ropponen A, Silventoinen K, Hublin C, Svedberg P, Koskenvuo M, Kaprio J. 2013. Sleep patterns as predictors for disability pension due to low back diagnoses: A 23-year longitudinal study of Finnish twins. Sleep 36:891–897.

Rosa RR, Colligan MJ. 1997. Plain language about shiftwork. DHHS (NIOSH) Publication No. 97–145. Cincinnati: US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health.

Sallinen M, Kecklund G. 2010. Shift work, sleep, and sleepiness—differences between shift schedules and systems. Scand J Work Environ Health 36:121–133.

Salo P, Vahtera J, Hall M, Rod NH, Virtanen M, Pentti J, Sjosten N, Oksanen T, Kivimäki M. 2012. Using repeated measures of sleep disturbances to predict future diagnosis-specific work disability: A cohort study. Sleep 35:559–569.

Schuh-Hofer S, Wodarski R, Pfau DB, Caspani O, Magerl W, Kennedy JD, Treede RD. 2013. One night of total sleep deprivation promotes a state of generalized hyperalgesia: A surrogate pain model to study the relationship of insomnia and pain. Pain 154:1613–1621.

Shimomitsu T, Haratani T, Nakamura K, Kawakami N, Hayashi T, Hiro H, Arai M, Miyazaki S, Furuki K, Ohya Y. et al. 2000. The Ministry of Labour sponsored grant for the prevention of work-related diseases: The 1999 report. In: Kato M, editor. The final development of the Brief Job Stress Questionnaire mainly used for assessment of the individuals. In: Tokyo: Tokyo Medical College. p 126–138 (in Japanese).

Smith L, Folkard S, Tucker P, Macdonald I. 1998. Work shift duration: A review comparing eight hour and 12 hour shift systems. Occup Environ Med 55:217–229.

Sorensen G, Stoddard AM, Stoffel S, Buxton O, Sembajwe G, Hashimoto D, Dennerlein JT, Hopcia K. 2011. The role of the work

context in multiple wellness outcomes for hospital patient care workers. J Occup Environ Med 53:899–910.

Sterud T, Tynes T. 2013. Work-related psychosocial and mechanical risk factors for low back pain: A 3-year follow-up study of the general working population in Norway. Occup Environ Med 70: 296–302.

Takahashi M. 2014. Assisting shift workers through sleep and circadian research. Sleep Biol Rhythms 12:85–95.

Takahashi M. 2012. Prioritizing sleep for healthy work schedules. J Physiol Anthropol 31:1–8.

Takahashi M, Fukuda H, Miki K, Haratani T, Kurabayashi L, Hisanaga N, Arito H, Takahashi H, Egoshi M, Sakurai M. 1999. Shift work-related problems in 16-h night shift nurses (2): Effects on subjective symptoms, physical activity, heart rate, and sleep. Ind Health 37: 228–236.

Takahashi M, Iwakiri K, Sotoyama M, Higuchi S, Kiguchi M, Hirata M, Hisanaga N, Kitahara T, Taoda K, Nishiyama K. 2008. Work schedule differences in sleep problems of nursing home caregivers. Appl Ergon 39:597–604.

Tang NK, Goodchild CE, Hester J, Salkovskis PM. 2012. Pain-related insomnia versus primary insomnia: A comparison study of sleep pattern, psychological characteristics, and cognitive-behavioral processes. Clin J Pain 28:428–436.

Tang NK, McBeth J, Jordan KP, Blagojevic-Bucknall M, Croft P, Wilkie R. 2015. Impact of musculoskeletal pain on insomnia onset: A prospective cohort study. Rheumatology (Oxford) 54:248–256.

Von Korff M, Ormel J, Keefe FJ, Dworkin SF. 1992. Grading the severity of chronic pain. Pain 50:133–149.

Wright KP, Jr., Bogan RK, Wyatt JK. 2013. Shift work and the assessment and management of shift work disorder (SWD). Sleep Med Rev 17:41–54.

Yamada Y, Kameda M, Noborisaka Y, Suzuki H, Honda M, Yamada S. 2001. Excessive fatigue and weight gain among cleanroom workers after changing from an 8-hour to a 12-hour shift. Scand J Work Environ Health 27:318–326.

Yassi A, Lockhart K. 2013. Work-relatedness of low back pain in nursing personnel: A systematic review. Int J Occup Environ Health 19:223–244.

Disclosure Statement: The authors report no conflicts of interests.

Industrial Health 2015, **53**, 368–377

Original Article

Assessment of psychosocial risk factors for the development of non-specific chronic disabling low back pain in Japanese workers—findings from the Japan Epidemiological Research of Occupation-related Back Pain (JOB) study

Ko MATSUDAIRA[1, 2]*, Mika KAWAGUCHI[3], Tatsuya ISOMURA[3, 4], Kyoko INUZUKA[3], Tadashi KOGA[3], Kota MIYOSHI[5] and Hiroaki KONISHI[6]

[1]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Japan
[2]Clinical Research Center for Occupational Musculoskeletal Disorders, Kanto Rosai Hospital, Japan
[3]Clinical Study Support, Inc., Japan
[4]Institute of Medical Science, Tokyo Medical University, Japan
[5]Department of Orthopaedic and Spine Surgery, Yokohama Rosai Hospital, Japan
[6]Department of Orthopaedic Surgery, Nagasaki Rosai Hospital, Japan

Received December 26, 2014 and accepted March 30, 2015
Published online in J-STAGE June 6, 2015

Abstract: To investigate the associations between psychosocial factors and the development of chronic disabling low back pain (LBP) in Japanese workers. A 1 yr prospective cohort of the Japan Epidemiological Research of Occupation-related Back Pain (JOB) study was used. The participants were office workers, nurses, sales/marketing personnel, and manufacturing engineers. Self-administered questionnaires were distributed twice: at baseline and 1 yr after baseline. The outcome of interest was the development of chronic disabling LBP during the 1 yr follow-up period. Incidence was calculated for the participants who experienced disabling LBP during the month prior to baseline. Logistic regression was used to assess risk factors for chronic disabling LBP. Of 5,310 participants responding at baseline (response rate: 86.5%), 3,811 completed the questionnaire at follow-up. Among 171 eligible participants who experienced disabling back pain during the month prior to baseline, 29 (17.0%) developed chronic disabling LBP during the follow-up period. Multivariate logistic regression analysis implied reward to work (not feeling rewarded, OR: 3.62, 95%CI: 1.17–11.19), anxiety (anxious, OR: 2.89, 95%CI: 0.97–8.57), and daily-life satisfaction (not satisfied, ORs: 4.14, 95%CI: 1.18–14.58) were significant. Psychosocial factors are key to the development of chronic disabling LBP in Japanese workers. Psychosocial interventions may reduce the impact of LBP in the workplace.

Key words: Chronic disabling low back pain, Nonspecific low back pain, Psychosocial factors, Risk factors, Japanese workers

*To whom correspondence should be addressed.
E-mail: kohart801@gmail.com

Introduction

Individuals commonly experience low back pain (LBP) at some stage during their life. Most LBP cases are classified as non-specific[1], which is not attributable to any identifiable pathology in the spine[2]. It is well-acknowledged that those who had LBP once tend to have subsequent episodes within a year[3–6], while each LBP episode can be resolved within a few weeks to 3 months[7, 8]. Despite the resolving nature of LBP, a small proportion of individuals with LBP (2–7%) develop chronic pain[8] which persists for 12 wk or longer[2]. In fact, LBP was found to be the leading specific cause of years lived with disability[9]. Not surprisingly, Western research has indicated that LBP, especially chronic LBP entailing disability, accounts for substantial economic loss at the workplace as well as in the healthcare system[2, 10].

An earlier Japanese study reported a lifetime LBP prevalence of over 80%[11]. Not surprisingly, the Ministry of Health, Labour and Welfare of Japan (MHLW) reported that LBP is the first and second most common health complaint in 2013 among Japanese men and women, respectively[12]. Since LBP is common in the Japanese population, the economic loss caused at the workplace and in the healthcare system is presumably as large as in Western countries.

In previous research, individual factors as well as ergonomic factors related to work have been well-investigated. In recent decades, an increasing body of evidence, however, has revealed that psychosocial factors play an important role in chronic non-specific LBP. In particular, distress (i.e., psychological distress, depressive mood, and depressive symptoms)[13, 14], low job satisfaction[14–16], emotional trauma in childhood such as abuse[17], and pain level[18] affect the development of chronic LBP.

Although the proportion of individuals suffering from chronic LBP is small according to Western studies, it is important to identify potential risk factors since the small proportion accounts for large loss. Little, however, is known concerning chronic disabling LBP in relation to psychosocial factors in Japanese workers. The objective of the present study was to investigate the associations between psychosocial factors and the development of chronic disabling LBP in Japanese workers.

Subjects and Methods

Data source

Data were drawn from a 1-yr prospective cohort of the Japan Epidemiological Research of Occupation-related Back Pain (JOB) study. Ethical approval was obtained from the review board of the MLHW. Participants for the JOB study were recruited at 16 local offices of the participating organizations in or near Tokyo. The occupations of the participating workers were diverse (e.g., office workers, nurses, sales/marketing personnel, and manufacturing engineers). Baseline questionnaires were distributed to employees by the board of each participating organization. Participants provided written informed consent and returned completed self-administered questionnaires with their name and mailing address for the purpose of follow-up directly to the study administration office. At a year after the baseline assessment, the follow-up questionnaire was distributed to the participants.

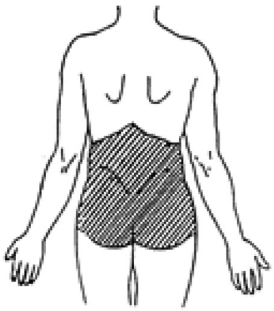


Fig. 1. Diagram showing pain area for low back provided in the baseline and follow-up questionnaires.

The baseline questionnaires contained questions on the presence of LBP, severity of LBP, individual characteristics (e.g., gender, age, obesity, smoking habit), ergonomic work demands (e.g., manual handling at work, frequency of bending, twisting), and work-related psychosocial factors (e.g., interpersonal stress at work, job control, reward to work, depression, somatization). LBP was defined in the questionnaire as pain localized between the costal margin and the inferior gluteal folds[10]. A diagram showing these areas was provided in the questionnaire to facilitate workers' understanding of the LBP area (Fig. 1). To evaluate the severity of LBP, Von Korff's grading[19] was used in the following manner: grade 0 was defined as no LBP; grade 1 as LBP that does not interfere with work; grade 2 as LBP that interferes with work but no absence from work; and grade 3 as LBP that interferes with work, leading to sick-leave. For the assessment of the psychosocial factors, the Brief Job Stress Questionnaire (BJSQ) developed by the MLHW[20, 21] was used. The BJSQ contains 57 ques-

tions and assesses 19 work-related stress factors: mental workload both quantitative- and qualitative-wise, physical workload, interpersonal stress at work, workplace environment stress, job control, utilization of skills and expertise, job fitness, reward to work, vigor, anger, fatigue, anxiety, depressed mood, somatic symptoms, supports by supervisors, supports by coworkers, supports by family or friends, and daily-life (work and life) satisfaction. These work-related factors were rated on a 5-point Likert scale ranging from the lowest score of 1 to the highest of 5.

The BJSQ incorporates questions from various standard questionnaires such as the Job Content Questionnaire (JCQ)[22], the National Institute for Occupational Safety and Health (NIOSH)[23], the Profile of Mood States (POMS)[24], the Center for Epidemiologic Studies Depression Scale (CES-D)[25], the State-trait Anxiety Inventory (STAI)[26], the Screener for Somatoform Disorders (SSD)[27] and the Subjective Well-being Inventory (SUBI)[28]. Standardized scores were developed for the 19 individual factors based on the sample of approximately 10,000 Japanese workers. The BJSQ has been shown to have internal consistency reliability and criterion validity with respect to the JCQ and NIOSH[29].

The follow-up questionnaire contained questions about the severity of LBP during the previous year, length of sick-leave because of LBP, medical care seeking, pain duration, and onset pattern. LBP severity was assessed using Von Korff's grading in the same manner as baseline.

Data analysis

The outcome of our interest was the development of chronic disabling LBP during the 1-yr follow-up period. In the present study, chronic disabling LBP was defined if a participant experienced LBP that interfered with work, with or without sick-leave due to LBP, corresponding to grade 2 or 3 in Von Korff's grading, during the month prior to baseline and experienced LBP with the same grades for 3 months or longer during the 1-yr follow-up period. Absence from work is often used as the outcome measurement for disability in Western studies. The present study, however, defined chronic disabling LBP as LBP that interfered with work for 3 months or longer, regardless of sick leave because our early international epidemiological study indicated that the proportion of Japanese workers who both took time off work and did not due to musculoskeletal disorders is almost equal to that of British workers who took time off work from the same reason[30]. This finding may be a result of cultural differences in attitude toward one's work. For this reason, the present study defined chronic disabling LBP as LBP that interfered with work for 3 months or longer, regardless of sick leave.

Incidence was calculated for the participants who experienced disabling LBP (grade 2 or 3) during the month prior to the baseline survey. Participants were excluded from the analysis if they changed their job for reasons other than LBP or developed LBP due to accident, a tumor, including metastasis, infection, or fracture.

For data analysis, the following factors were initially included: (1) individual characteristics, (2) ergonomic work demands, and (3) work-related psychosocial factors. Individual characteristics included age, sex, obesity (body mass index: BMI $\geq$25 kg/m^2), smoking habit (Brinkman index $\geq$400), education, flexibility, hours of sleep, experience at current job, working hours per wk ($\geq$60 h per week of uncontrolled overtime), work shift, emotional trauma in childhood, and pain level (NRS $\geq$8 as painful). Ergonomic work demands included manual handling at work; bending, twisting ($\geq$half of the day as frequent); and hours of desk work ($\geq$half of the day as frequent). Psychosocial factors were assessed with BJSQ. The 5-point Likert scale was reclassified into 2 categories: the "not feeling stressed" category, where low, slightly low, and moderate were combined, and the "feeling stressed" category, where slightly high and high were combined. Pain level was scaled on the Numerical Rating Scale, ranging from 0 to 11.

To assess smoking habit, the Brinkman Index was calculated based on the total number of cigarettes smoked per day multiplied by duration of smoking in year[31]. A Brinkman Index value of 400 or higher indicated that a respondent was a heavy smoker, whereas a value of less than 400 indicated that a respondent was a non-heavy smoker. Workers were defined as flexible if their wrists could reach beyond their knees but without their fingertips touching their ankles, and not flexible if their wrists could not reach beyond their knees[32].

In addition to descriptive statistics, univariate and multivariate logistic regression analyses were conducted to examine the associations between risk factors and the development of chronic disabling LBP. Results of logistic regression analyses were summarized by odds ratios (ORs) and the respective 95% confidence intervals (CI). To assess potential risk factors, crude ORs were initially computed. Subsequently, all factors with $p<0.1$ in univariate logistic regression analyses were entered into the multivariate logistic regression model, significance levels of $p<0.05$ for entry and $p>0.1$ for removal. The stepwise method was used to select variables with statistical significance at $p<0.05$. All tests were 2-tailed. The software

package STATA 9.0 (StataCorp, LP, College Station, TX) was used for all statistical analyses.

Results

Baseline characteristics of the follow-up vs. drop-out group

The baseline questionnaire was distributed to 6,140 workers and had a response rate of 86.5% (5,310 workers). Of these participants, 3,811 workers successfully completed and returned 1-yr follow-up questionnaires (follow-up rate: 71.8%).

The characteristics of the 3,811 participants who provided follow-up data (follow-up group) did not appear to be much different from those who did not (drop-out group). The mean [standard deviation (SD)] age of the follow-up group was 42.9 (10.1) yr, compared to 38.0 (10.2) yr in the drop-out group. The majority were men in both groups (80.6% and 82.8%, respectively). The mean (SD) BMI of the follow-up group and drop-out group were similar [23.1 (3.3) and 22.9 (4.1), respectively]. In the follow-up group, 78.6% of the participants engaged in the manual handling of objects <20 kg, or not manually handling any objects at work, 17.8% engaged in manually handling objects ≥20 kg or worked as a caregiver, and data was missing for 3.6%. The respective values for the drop-out group were 75.5%, 18.9%, and 5.6%. In both the follow-up and drop-out groups, the most common occupational fields were office workers engaging in the manual handling of objects <20 kg or not manually handling any objects and nurse engaging in manual handling of objects ≥20 kg or caregiver.

Baseline characteristics of the study participants

Of the 3,811 workers, 171 reported LBP and experiencing work interferences with or without sick-leave during a month prior to baseline (Fig. 2). The mean (SD) age of 171 participants was 41.5 (10.2) yr and the majority were men (n=122; 71.4%). The mean (SD) BMI of the participants was 23.0 (3.6; n=170) kg/m^2. About half of the participants did not engage in manually handling heavy objects at work (n=79; 48.8%). Those workers who manually handled objects of less than 20 kg accounted for 17.9% (n=29) and those who manually handled heavy objects 20 kg or heavier or worked as a caregiver accounted for 33.3% (n=54). Desk work and sales, manufacturing, and nurses were the major occupations in the categories of non-manually handling work, manually handling work of less than 20 kg, and manually handling work of 20 kg or heavier, respectively.

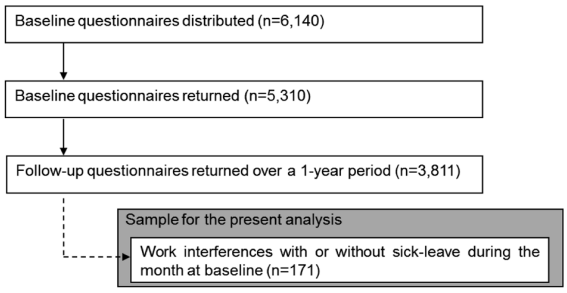


Fig. 2. Flow chart of the sample selection for the present analysis.

Incidence of chronic disabling LBP

Of a total of 171 eligible participants, 29 (17.0%) developed chronic disabling LBP during a year prior to the follow-up period (5 missing cases).

Association between chronic disabling LBP and potential risk factors

Crude and adjusted ORs for the development of chronic disabling LBP and their 95% CIs are shown in Tables 1 and 2. The univariate logistic regression analysis showed that job fitness, reward to work, vigor, anger, fatigue, anxiety, depressed mood, supports by supervisors, daily-life satisfaction, work shift, emotional trauma in childhood, and pain level were potentially associated with the development of chronic disabling LBP (ORs of 2.00–7.93; p<0.1 for all) (Table 1). In the multivariate logistic regression analysis, these 12 factors were entered into the model. As a result, 3 psychosocial factors were selected: reward to work (OR: 3.62, 95%CI: 1.17–11.19), anxiety (OR: 2.89, 95%CI: 0.97–8.57), and daily-life satisfaction (OR: 4.14, 95%CI: 1.18–14.58) (Table 2), indicating that a combination of psychosocial factors can play a key role in the development of chronic disabling LBP. A supplemental analysis was conducted to examine a combination effect of psychosocial factors: reward to work and daily-life satisfaction, which were at p<0.05 in the multiple logistic regression model (Table 3). Consequently, ORs increased with the level of dissatisfaction in a combination of daily-life satisfaction and reward to work. The results suggested that when both daily-life satisfaction and reward to work were not satisfied with an approximately 8-fold higher risk of developing chronic disabling LBP.

Discussion

Results suggest that exposure to multiple psychosocial factors potentially predisposes the development of

Table 1. Crude odds ratios of baseline factors for chronic disabling LBP

Risk factor	n	%	Odds ratio	95%CI	*p* value
Age (yr)	171				
<40	78	45.6	1.00		
40–49	51	29.8	0.95	0.36–2.48	0.909
≥50	42	24.6	1.17	0.44–3.12	0.746
Sex	171				
Male	122	71.4	1.00		
Female	49	28.7	1.26	0.53–3.03	0.601
Obesity[a]	169				
< BMI 25 kg/m^2	129	76.3	1.00		
≥ BMI 25 kg/m^2 (obesity)	40	23.7	0.85	0.32–2.28	0.748
Smoking habit	153				
Heavy smoker	112	73.2	1.00		
Not heavy smoker	41	26.8	1.80	0.72–4.52	0.211
Education	165				
College/Junior college	105	63.6	1.00		
High school/Junior high school	60	36.4	0.44	0.17–1.18	0.103
Flexibility	162				
Flexibility	98	60.5	1.00		
Not flexible	64	39.5	0.57	0.23–1.41	0.225
Manual handling at work	162				
No manual handling (desk work)	79	48.8	1.00		
Manual handling of <20-kg objects	29	17.9	1.40	0.43–4.50	0.577
Manual handling of ≥20-kg objects or working as a caregiver	54	33.3	1.84	0.72–4.72	0.203
Bending	169				
Not frequent	121	71.6	1.00		
Frequent	48	28.4	1.40	0.58–3.40	0.454
Twisting	168				
Not frequent	140	83.3	1.00		
Frequent	28	16.7	1.24	0.42–3.65	0.690
Hours of desk work	167				
Not frequent	111	66.5	1.00		
Frequent	56	33.5	0.74	0.30–1.81	0.510
Mental workload (quantitative aspect)	170				
Not stressed	66	38.8	1.00		
Stressed	104	61.2	1.08	0.47–2.46	0.859
Mental workload (qualitative aspect)	170				
Not stressed	71	41.8	1.00		
Stressed	99	58.2	0.63	0.28–1.42	0.267
Physical workload	171				
Not stressed	75	43.9	1.00		
Stressed	96	56.1	1.62	0.70–3.73	0.260
Interpersonal stress at work	171				
Not stressed	118	69.0	1.00		
Stressed	53	31.0	1.15	0.49–2.68	0.745
Workplace environment stress	171				
Not stressed	102	59.7	1.00		
Stressed	69	40.4	1.95	0.87–4.38	0.105
Job control	169				
Controlled	4	32.0	1.00		
Not controlled	115	68.1	1.81	0.69–4.79	0.230
Utilization of skills and expertise	170				
Utilization of skills and expertise	131	77.1	1.00		
No utilization of skills and expertise	9	22.9	1.59	0.66–3.85	0.304
Job fitness	171				
Feeling fit	114	66.7	1.00		
Not feeling fit	7	33.3	2.04	0.91–4.60	0.086

Table 1. Continued

Risk factor	n	%	Odds ratio	95%CI	*p* value
Reward to work	171				
Feel rewarded	120	70.2	1.00		
Not feeling rewarded	51	29.8	3.59	1.57–8.20	0.002
Vigor	170				
Vigorous	123	72.4	1.00		
Not vigorous	47	27.7	2.12	0.92–4.88	0.078
Anger	170				
Not angry	75	44.1	1.00		
Angry	95	55.9	2.79	1.12–6.97	0.028
Fatigue	171				
No fatigue	69	40.4	1.00		
Fatigue	102	59.7	2.45	0.98–6.11	0.055
Anxiety	171				
Not anxious	95	55.6	1.00		
Anxious	76	44.4	2.75	1.19–6.35	0.018
Depressed mood	169				
Not feeling depressed	79	46.8	1.00		
Depressed	90	53.3	2.16	0.92–5.08	0.078
Somatic symptoms	168				
Not somatic symptoms	58	34.5	1.00		
Somatic symptoms	110	65.5	1.81	0.72–4.55	0.206
Supports by supervisors	167				
Supported	103	61.7	1.00		
Not supported	64	38.3	2.00	0.88–4.55	0.098
Supports by coworkers	168				
Supported	93	55.4	1.00		
Not supported	75	44.6	0.97	0.43–2.18	0.946
Supports by family or friends	169				
Supported	128	75.7	1.00		
Not supported	41	24.3	1.13	0.44–2.90	0.801
Daily-life satisfaction	171				
Satisfied	96	56.1	1.00		
Not satisfied	75	43.9	4.98	1.99–12.47	0.001
Hours of sleep	168				
≤5 h	151	89.9	1.00		
>5 h	17	10.1	1.56	0.47–5.21	0.466
Experience of current job	171				
<5 yr	55	32.2	1.00		
≥5 yr	116	67.8	1.02	0.43–2.42	0.970
Working hours per wk	171				
<60 h	131	76.6	1.00		
≥60 h	40	23.4	0.63	0.22–1.78	0.385
Work shift	171				
Daytime shift	115	67.3	1.00		
Nighttime shift	56	32.8	2.90	1.28–6.58	0.011
Emotional trauma in childhood	143				
No	136	95.1	1.00		
Yes	7	4.9	7.93	1.64–38.26	0.010
Pain level	155				
Not painful (NRS >8)	140	90.3	1.00		
Painful (NRS ≤8)	15	9.7	4.11	1.31–12.85	0.015

LBP: low back pain; CI: confidence interval; BMI: body mass index; NRS: numerical rating scale.
BMI ≥25 kg/m^2 is defined as obesity in Japan

Table 2. Stepwise logistic regression results of baseline factors for chronic disabling LBP

Risk factor	Odds ratio	95%CI	*p* value
Reward to work			
Feel rewarded	1.00		
Not feeling rewarded	3.62	1.17–11.2	0.025
Anxiety			
Not anxious	1.00		
Anxious	2.89	0.97–8.57	0.056
Daily-life satisfaction			
Satisfied	1.00		
Not satisfied	4.14	1.18–14.58	0.027

LBP: low back pain; CI: confidence interval; BMI: body mass index.

chronic disabling LBP in Japanese workers, especially office workers, nurses, sales/marketing personnel, and manufacturing engineers. Similarly, an increasing body of evidence, mostly in Western countries, has indicated that psychosocial factors affect the development of chronic disabling LBP[13–17].

The present study suggests that exposure to not one, but a combination of psychosocial factors, such as daily-life satisfaction and reward to work, may trigger the development of chronic disabling LBP with an 8-fold increased risk, compared to those who were satisfied with psychosocial aspects. Given that daily-life satisfaction in the BJSQ consists of the extent of being content with not only life, but also work, the results in the present study are consistent with Western studies indicating that job dissatisfaction predisposes the development of chronic disabling LBP[14–16, 33–35]. Another psychosocial factor, reward to work, can also be considered to be relevant to the magnitude in job satisfaction. The association between chronic disabling LBP and a combination of such psychosocial factors may possibly be explained by dysfunction in mesolimbic dopaminergic activity. In recent years, there has been an assumption that exposure to chronic, rather than acute, stress could result in a state of hyperalgesia in the body due to the inhabitation of mesolimbic dopaminergic mechanisms where both pain and pleasure are controlled[36, 37]. Hyperalgesia resulting from chronic stress due to not being content with life and work, for example, may lead to the development of chronic disabling LBP.

In the past, the occupational health of the Japanese worker has mainly focused on an ergonomic approach in the management and prevention of LBP. Consistent with Western studies, the present study suggests, however, that we should be more alert to a psychosocial approach to reduce the risk of developing chronic disabling LBP. Although our earlier prospective study indicated that both ergonomic and work-related psychosocial factors were associated with new-onset of disabling LBP in symptom-free Japanese workers[38], no ergonomic factors seemingly affect the development of chronic disabling LBP in the present study probably because workers who already experienced disabling LBP at baseline were the focus of the present study. The results are consistent with the guidelines stating that the development of chronic pain and disability results more from work-related psychosocial issues than from physical features[34].

There are several limitations to the study. First, generalization of the results of the present study is limited. The majority of the study participants were males. The study cohort was also not a representative sample of all Japanese workers in terms of area as well as range of occupations. Second, the sample size for the present analysis is small. Future research with a larger sample size should be conducted for further identification of potential risk factors of chronic disabling LBP. Third, the context of cognitive and emotional aspects, such as fear-avoidance belief and physician's attitudes, was not considered in the present study despite being known to affect the development of serious disability. As of the time of data collection, scales measuring fear avoidance were not available in the Japanese language. Since the author developed the Japanese versions of the Fear-Avoidance Beliefs Questionnaire (FABQ)[39]

Table 3. Odds ratios for chronic disabling LBP in relation with a combination of daily-life satisfaction and reward to work

Risk factor		Chronic disabling LBP		Odds ratio	95%CI
Daily-life satisfaction	Reward to work	Yes (%)	No (%)		
Satisfied	Feel rewarded	6 (7.7%)	72 (92.3%)	-	-
	Not feeling rewarded	1 (7.7%)	12 (92.3%)	1.00	0.11–9.06
Not satisfied	Feel rewarded	7 (18.9%)	30 (81.1%)	2.80	0.87–9.03
	Not feeling rewarded	15 (39.5%)	23 (60.5%)	7.83	2.72–22.52

LBP: low back pain; CI: confidence interval.

and the Tampa Scale of Kinesiophobia (TSK)[40, 41] after the JOB survey, both are currently available. These scales should also be included in future research. Fourth, misclassification, to some extent, is inevitable. Responses that rely on subjective measurement may be distorted and missing values cannot be avoided due to the nature of a self-assessment survey. Moreover, the possibility for recall bias towards retrospective questions should be kept in mind. Fifth, the present study focuses on the baseline factors affecting the development of chronic disabling LBP under the assumption that workers retained the same status quo as the baseline during the follow-up period. The status in some factors could possibly fluctuate during the period. Such fluctuation in factors was not taken into consideration in the present study. Finally, there may be alternative methods for the selection of potential risk factors prior to conducting multivariate analysis. It should be noted that a more complicated model may offer a better explanation of the data although the results are consistent with Western studies. Further research is needed to identify a full range of potential risk factors for inclusion in future studies.

In conclusion, the present study suggests that psychosocial factors could play a key role in the development of chronic disabling LBP in Japanese workers. Therefore, the occupational health of the Japanese worker should be focused not only on ergonomic interventions but also on psychosocial ones to reduce the impact on the workplace from the repercussions of developing chronic disabling LBP.

References

1) van Tulder M, Koes B, Bombardier C (2002) Low back pain. Best Pract Res Clin Rheumatol **16**, 761–75. [Medline] [CrossRef]
2) Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G; COST B13 Working Group on Guidelines for Chronic Low Back Pain (2006) Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J **15**, Suppl 2; S192–300.
3) Carey TS, Garrett JM, Jackman A, Hadler N (1999) Recurrence and care seeking after acute back pain: results of a long-term follow-up study. North Carolina Back Pain Project. Med Care **37**, 157–64. [Medline] [CrossRef]
4) Pengel LH, Herbert RD, Maher CG, Refshauge KM (2003) Acute low back pain: systematic review of its prognosis. BMJ **327**, 323–7. [Medline] [CrossRef]
5) Von Korff M (1994) Studying the natural history of back pain. Spine **19** Suppl, 2041S–6S. [Medline] [CrossRef]
6) Von Korff M, Deyo RA, Cherkin D, Barlow W (1993) Back pain in primary care: outcomes at 1 year. Spine **18**, 855–62. [Medline] [CrossRef]
7) Croft PR, Macfarlane GJ, Papageorgiou AC, Thomas E, Silman AJ (1998) Outcome of low back pain in general practice: a prospective study. BMJ **316**, 1356–9. [Medline] [CrossRef]
8) Nachemson AL, Waddell G, Norlund AI (2000) Epidemiology of neck and low back pain. In: Neck and back pain: The scientific evidence of causes, diagnosis and treatment, Nachemson AL, Jonsson E (Eds.), 165–88, Lippincott Williams & Wilkins, Philadelphia.
9) Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Bahalim AN, Barker-Collo S, Barrero LH, Bartels DH, Basáñez MG, Baxter A, Bell ML, Benjamin EJ, Bennett D, Bernabé E, Bhalla K, Bhandari B, Bikbov B, Bin Abdulhak A, Birbeck G, Black JA, Blencowe H, Blore JD, Blyth F, Bolliger I, Bonaventure A, Boufous S, Bourne R, Boussinesq M, Braithwaite T, Brayne C, Bridgett L, Brooker S, Brooks P, Brugha TS, Bryan-Hancock C, Bucello C, Buchbinder R, Buckle G, Budke CM, Burch M, Burney P, Burstein R, Calabria B, Campbell B, Canter CE, Carabin H, Carapetis J, Carmona L, Cella C, Charlson F, Chen H, Cheng AT, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahiya M, Dahodwala N, Damsere-Derry J, Danaei G, Davis A, De Leo D, Degenhardt L, Dellavalle R, Delossantos A, Denenberg J, Derrett S, Des Jarlais DC, Dharmaratne SD, Dherani M, Diaz-Torne C, Dolk H, Dorsey ER, Driscoll T, Duber H, Ebel B, Edmond K, Elbaz A, Ali SE, Erskine H, Erwin PJ, Espindola P, Ewoigbokhan SE, Farzadfar F, Feigin V, Felson DT, Ferrari A, Ferri CP, Fèvre EM, Finucane MM, Flaxman S, Flood L, Foreman K, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabbe BJ, Gabriel SE, Gakidou E, Ganatra HA, Garcia B, Gaspari F, Gillum RF, Gmel G, Gosselin R, Grainger R, Groeger J, Guillemin F, Gunnell D, Gupta R, Haagsma J, Hagan H, Halasa YA, Hall W, Haring D, Haro JM, Harrison JE, Havmoeller R, Hay RJ, Higashi H, Hill C, Hoen B, Hoffman H, Hotez PJ, Hoy D, Huang JJ, Ibeanusi SE, Jacobsen KH, James SL, Jarvis D, Jasrasaria R, Jayaraman S, Johns N, Jonas JB, Karthikeyan G, Kassebaum N, Kawakami N, Keren A, Khoo JP, King CH, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lalloo R, Laslett LL, Lathlean T, Leasher JL, Lee YY, Leigh J, Lim SS, Limb E, Lin JK, Lipnick M, Lipshultz SE, Liu W, Loane M, Ohno SL, Lyons R, Ma J, Mabweijano J, MacIntyre MF, Malekzadeh R, Mallinger L, Manivannan S, Marcenes W, March L, Margolis DJ, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGill N, McGrath J, Medina-Mora

ME, Meltzer M, Mensah GA, Merriman TR, Meyer AC, Miglioli V, Miller M, Miller TR, Mitchell PB, Mocumbi AO, Moffitt TE, Mokdad AA, Monasta L, Montico M, Moradi-Lakeh M, Moran A, Morawska L, Mori R, Murdoch ME, Mwaniki MK, Naidoo K, Nair MN, Naldi L, Narayan KM, Nelson PK, Nelson RG, Nevitt MC, Newton CR, Nolte S, Norman P, Norman R, O'Donnell M, O'Hanlon S, Olives C, Omer SB, Ortblad K, Osborne R, Ozgediz D, Page A, Pahari B, Pandian JD, Rivero AP, Patten SB, Pearce N, Padilla RP, Perez-Ruiz F, Perico N, Pesudovs K, Phillips D, Phillips MR, Pierce K, Pion S, Polanczyk GV, Polinder S, Pope CA 3rd, Popova S, Porrini E, Pourmalek F, Prince M, Pullan RL, Ramaiah KD, Ranganathan D, Razavi H, Regan M, Rehm JT, Rein DB, Remuzzi G, Richardson K, Rivara FP, Roberts T, Robinson C, De Leòn FR, Ronfani L, Room R, Rosenfeld LC, Rushton L, Sacco RL, Saha S, Sampson U, Sanchez-Riera L, Sanman E, Schwebel DC, Scott JG, Segui-Gomez M, Shahraz S, Shepard DS, Shin H, Shivakoti R, Singh D, Singh GM, Singh JA, Singleton J, Sleet DA, Sliwa K, Smith E, Smith JL, Stapelberg NJ, Steer A, Steiner T, Stolk WA, Stovner LJ, Sudfeld C, Syed S, Tamburlini G, Tavakkoli M, Taylor HR, Taylor JA, Taylor WJ, Thomas B, Thomson WM, Thurston GD, Tleyjeh IM, Tonelli M, Towbin JA, Truelsen T, Tsilimbaris MK, Ubeda C, Undurraga EA, van der Werf MJ, van Os J, Vavilala MS, Venketasubramanian N, Wang M, Wang W, Watt K, Weatherall DJ, Weinstock MA, Weintraub R, Weisskopf MG, Weissman MM, White RA, Whiteford H, Wiersma ST, Wilkinson JD, Williams HC, Williams SR, Witt E, Wolfe F, Woolf AD, Wulf S, Yeh PH, Zaidi AK, Zheng ZJ, Zonies D, Lopez AD, Murray CJ, AlMazroa MA, Memish ZA (2012) Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet **380**, 2163–96 (Erratum in: 2013). Lancet **381**, 628.

10) Krismer M, van Tulder M; Low Back Pain Group of the Bone and Joint Health Strategies for Europe Project (2007) Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best Pract Res Clin Rheumatol **21**, 77–91. [Medline] [CrossRef]
11) Fujii T, Matsudaira K (2013) Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. Eur Spine J **22**, 432–8. [Medline] [CrossRef]
12) The Japan Labour Health and Welfare Organization The 2013 National Livelihood Survey (Kokumin Seikatsu Kiso Chousa) (in Japanese). The Japan Health and Welfare Organization: Tokyo, Japan. http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa13/dl/04.pdf Accessed September 16, 2014.
13) Pincus T, Burton AK, Vogel S, Field AP (2002) A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine **27**, E109–20. [Medline] [CrossRef]
14) Thomas E, Silman AJ, Croft PR, Papageorgiou AC, Jayson MI, Macfarlane GJ (1999) Predicting who develops chronic low back pain in primary care: a prospective study. BMJ **318**, 1662–7. [Medline] [CrossRef]
15) Turk DC, Rudy TE (1992) Classification logic and strategies in chronic pain. In: Handbook of Pain Assessment. Turk DC, Melzack R (Eds.), 409–28, NY, Guildford.
16) Williams RA, Pruitt SD, Doctor JN, Epping-Jordan JE, Wahlgren DR, Grant I, Patterson TL, Webster JS, Slater MA, Atkinson JH (1998) The contribution of job satisfaction to the transition from acute to chronic low back pain. Arch Phys Med Rehabil **79**, 366–74. [Medline] [CrossRef]
17) McMahon MJ, Gatchel RJ, Polatin PB, Mayer TG (1997) Early childhood abuse in chronic spinal disorder patients. A major barrier to treatment success. Spine **22**, 2408–15. [Medline] [CrossRef]
18) Nakamura M, Nishiwaki Y, Ushida T, Toyama Y (2014) Prevalence and characteristics of chronic musculoskeletal pain in Japan: a second survey of people with or without chronic pain. J Orthop Sci **19**, 339–50. [Medline] [CrossRef]
19) Von Korff M, Ormel J, Keefe FJ, Dworkin SF (1992) Grading the severity of chronic pain. Pain **50**, 133–49. [Medline] [CrossRef]
20) Muto S, Muto T, Seo A, Yoshida T, Taoda K, Watanabe M (2006) Prevalence of and risk factors for low back pain among staffs in schools for physically and mentally handicapped children. Ind Health **44**, 123–7. [Medline] [CrossRef]
21) Kawakami N, Kobayashi Y, Takao S, Tsutsumi A (2005) Effects of web-based supervisor training on supervisor support and psychological distress among workers: a randomized controlled trial. Prev Med **41**, 471–8. [Medline] [CrossRef]
22) Kawakami N, Kobayashi F, Araki S, Haratani T, Furui H (1995) Assessment of job stress dimensions based on the job demands—control model of employees of telecommunication and electric power companies in Japan: reliability and validity of the Japanese version of the Job Content Questionnaire. Int J Behav Med **2**, 358–75. [Medline] [CrossRef]
23) Haratani T, Kawakami N, Araki S (1993) Reliability and validity of the Japanese version of NIOSH Generic Job Questionnaire. Sangyo Igaku **35** suppl, S214 (Jpn J Ind Med) (in Japanese). [CrossRef]
24) Yokoyama K, Araki S, Kawakami N, Tkakeshita T (1990) Production of the Japanese edition of profile of mood states (POMS): assessment of reliability and validity. Nippon Koshu Eisei Zasshi **37**, 913–8 (in Japanese).
25) Shima S, Shikano T, Kitamura T, Asai M (1985) New self-rating scales for depression. Clin Psychiatry **27**, 717–23 (in Japanese).
26) Spielberger CD, Gorsuch RL, Lushene RE (1970) Manual for the State-Trait Anxiety Inventory. 23–49, Consulting Psychologists Press, Palo Alto.

27) Isaac M, Tacchini G, Janca A (1994) Screener for somatoform disorders (SSD). World Health Organization, Geneva.
28) Ono Y, Yoshimura K, Yamauchi K, Momose T, Mizushima H, Asai M (1996) Psychological well-being and ill-being: WHO Subjective Well-being Inventory (SUBI). Jpn J Stress Sci **10**, 273–8.
29) Shimomitsu T, Odagiri Y (2004) The brief job stress questionnaire. Occup Ment Health **12**, 25–36 (in Japanese).
30) Matsudaira K, Palmer KT, Reading I, Hirai M, Yoshimura N, Coggon D (2011) Prevalence and correlates of regional pain and associated disability in Japanese workers. Occup Environ Med **68**, 191–6. [Medline] [CrossRef]
31) Brinkman GL, Coates EO Jr (1963) The effect of bronchitis, smoking, and occupation on ventilation. Am Rev Respir Dis **87**, 684–93. [Medline]
32) Akaha H, Matsudaira K, Takeshita K, Oka H, Hara N, Nakamura K (2008) Modified measurement of finger-floor distance—Self assessment bending scale—. J Lumbar Spine Disord **14**, 164–9. [CrossRef]
33) Williams RA, Pruitt SD, Doctor JN, Epping-Jordan JE, Wahlgren DR, Grant I, Patterson TL, Webster JS, Slater MA, Atkinson JH (1998) The contribution of job satisfaction to the transition from acute to chronic low back pain. Arch Phys Med Rehabil **79**, 366–74. [Medline] [CrossRef]
34) Waddell G, Burton AK (2001) Occupational health guidelines for the management of low back pain at work: evidence review. Occup Med (Lond) **51**, 124–35. [Medline] [CrossRef]
35) Heymans MW, Anema JR, van Buuren S, Knol DL, van Mechelen W, de Vet HC (2009) Return to work in a cohort of low back pain patients: development and validation of a clinical prediction rule. J Occup Rehabil **19**, 155–65. [Medline] [CrossRef]
36) Wood PB (2006) Mesolimbic dopaminergic mechanisms and pain control. Pain **120**, 230–4. [Medline] [CrossRef]
37) Leknes S, Tracey I (2008) A common neurobiology for pain and pleasure. Nat Rev Neurosci **9**, 314–20. [Medline] [CrossRef]
38) Matsudaira K, Konishi H, Miyoshi K, Isomura T, Takeshita K, Hara N, Yamada K, Machida H (2012) Potential risk factors for new onset of back pain disability in Japanese workers: findings from the Japan epidemiological research of occupation-related back pain study. Spine **37**, 1324–33. [Medline] [CrossRef]
39) Waddell G, Newton M, Henderson I, Somerville D, Main CJ (1993) A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain **52**, 157–68. [Medline] [CrossRef]
40) Miller RP, Kori SH, Todd DD (1991) The Tampa Scale: a measure of kinisophobia. Clin J Pain **7**, 51–2 (Data unpublished). [CrossRef]
41) Kori KS, Miller RP, Todd DD (1990) Kinesiophobia: a new view of chronic pain behavior. Pain Manag **3**, 35–43.

J Bone Miner Metab (2015) 33:186–191
DOI 10.1007/s00774-014-0573-y

ORIGINAL ARTICLE

Incidence of disability and its associated factors in Japanese men and women: the Longitudinal Cohorts of Motor System Organ (LOCOMO) study

Noriko Yoshimura · Toru Akune · Saeko Fujiwara · Yoko Shimizu · Hideyo Yoshida · Yuji Nishiwaki · Akihiro Sudo · Go Omori · Munehito Yoshida · Hiroshi Shimokata · Takao Suzuki · Shigeyuki Muraki · Hiroyuki Oka · Kozo Nakamura

Received: 27 August 2013 / Accepted: 19 January 2014 / Published online: 27 April 2014

Abstract We investigated the incidence of disability and its risk factors in older Japanese adults to establish an evidence-based disability prevention strategy for this population. For this purpose, we used data from the Longitudinal Cohorts of Motor System Organ (LOCOMO) study, initiated in 2008 to integrate information from cohorts in nine communities across Japan: Tokyo (two regions), Wakayama (two regions), Hiroshima, Niigata, Mie, Akita, and Gunma prefectures. We examined the annual occurrence of disability from 8,454 individuals (2,705 men and 5,749 women) aged ≥65 years. The estimated incidence of disability was 3.58/100 person-years (p-y) (men: 3.17/100 p-y; women: 3.78/100 p-y). To determine factors associated with disability, Cox's proportional hazard model was used, with the occurrence of disability as an objective variable and age (+1 year), gender (vs. women), body build (0: normal/overweight range, BMI 18.5–27.5 kg/m^2; 1: emaciation, BMI <18.5 kg/m^2; 2: obesity, BMI >27.5 kg/m^2), and regional differences (0: rural areas including Wakayama, Niigata, Mie, Akita, and Gunma vs. 1: urban areas including Tokyo and Hiroshima) as explanatory variables. Age, body build, and regional difference significantly influenced the occurrence of disability (age, +1 year: hazard ratio 1.13, 95 % confidence interval 1.12–1.15, $p < 0.001$; body build, vs. emaciation: 1.24, 1.01–1.53, $p = 0.041$; body build, vs. obesity: 1.36, 1.08–1.71, $p = 0.009$; residence, vs. living in rural areas: 1.59, 1.37–1.85, $p < 0.001$). We concluded that higher age,

N. Yoshimura (✉) · H. Oka
Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
e-mail: yoshimuran-ort@h.u-tokyo.ac.jp

T. Akune · S. Muraki
Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

S. Fujiwara
Hiroshima Atomic Bomb Causality Council, Hiroshima, Japan

Y. Shimizu · H. Yoshida
Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan

Y. Nishiwaki
Department of Environmental and Occupational Health, School of Medicine, Toho University, Tokyo, Japan

A. Sudo
Department of Orthopaedic Surgery, Mie University Graduate School of Medicine, Mie, Japan

G. Omori
Center for Transdisciplinary Research, Institute for Research Promotion, Niigata University, Niigata, Japan

M. Yoshida
Department of Orthopedic Surgery, School of Medicine, Wakayama Medical University, Wakayama, Japan

H. Shimokata
Graduate School of Nutritional Sciences, Nagoya University of Arts and Sciences, Aichi, Japan

T. Suzuki
National Center for Geriatrics and Gerontology, Aichi, Japan

K. Nakamura
National Rehabilitation Center for Persons with Disabilities, Saitama, Japan

both emaciation and obesity, and living in rural areas would be risk factors for the occurrence of disability.

Keywords Nation-wide population-based cohort study · Epidemiology · Incidence · Disability · Body build

Introduction

In Japan, the proportion of the population aged 65 years or older has increased rapidly over the years. In 1950, 1985, 2005, and 2010, this proportion was 4.9, 10.3, 19.9, and 23.0 %, respectively [1]. Further, this proportion is estimated to reach 30.1 % in 2024 and 39.0 % in 2051 [2]. The rapid aging of Japanese society, unprecedented in world history, has led to an increase in the number of disabled elderly individuals requiring support or long-term care. The Japanese government initiated the national long-term care insurance system in April 2000 in adherence with the Long-Term Care Insurance Act [3]. The aim of the national long-term care insurance system was to certify the level of care needed by elderly adults and to provide suitable care services to them according to the levels of their long-term care needs. According to the recent National Livelihood Survey by the Ministry of Health, Labour and Welfare in Japan, the number of elderly individuals certified as needing care services increases annually, having reached 5 million in 2011 [4].

However, few prospective, longitudinal, and cross-national studies have been carried out to inform the development of a prevention strategy against disability. To establish evidence-based prevention strategies, it is critically important to accumulate epidemiologic evidence, including the incidence of disability, and identify its risk factors. However, few studies have attempted to estimate the incidence of the disability and its risk factors by using population-based cohorts. In addition, to identify the incidence of disability, a study should have a large number of subjects. Further, to determine regional differences in epidemiological indices, a survey of cohorts across Japan is required.

The Longitudinal Cohorts of Motor System Organ (LOCOMO) study was initiated in 2008, through a grant from Japan's Ministry of Health, Labour and Welfare, for the prevention of knee pain, back pain, bone fractures, and subsequent disability. It aimed to integrate data gathered from cohorts from 2000 onwards and follow-up surveys from 2006 onwards, using a unified questionnaire, with an ultimate goal being the prevention of musculoskeletal diseases. The present study specifically aims at using LOCOMO data, which is based on the long-term care insurance system, to investigate the occurrence of disability in order to clarify its incidence and risk factors, especially in terms of body build and regional differences.

Materials and methods

Participants were residents of nine communities located in Tokyo (two regions: Tokyo-1, principal investigators (PIs): Shigeyuki Muraki, Toru Akune, Noriko Yoshimura, Kozo Nakamura; Tokyo-2, PIs: Yoko Shimizu, Hideyo Yoshida, Takao Suzuki), Wakayama [two regions: Wakayama-1 (mountainous region) and Wakayama-2 (coastal region), PIs: Noriko Yoshimura, Munehito Yoshida], Hiroshima (PI: Saeko Fujiwara), Niigata (PI: Go Omori), Mie (PI: Akihiro Sudo), Akita (PI: Hideyo Yoshida), and Gunma (PI: Yuji Nishiwaki) prefectures [5]. Figure 1 shows the location of each cohort in Japan.

Disability in the present study was defined as 'cases requiring long-term care', as determined by the long-term care insurance system. The procedure for identifying these cases is as follows: (1) each municipality establishes a long-term care approval board consisting of clinical experts, physicians, and specialists at the Division of Health and Welfare in each municipal office; (2) The long-term care approval board investigates the insured person by using an interviewer-administered questionnaire consisting of 82 items regarding mental and physical conditions, and makes a screening judgement based on the opinion of a regular doctor; (3) 'Cases requiring long-term care' are determined according to standards for long-term care certification that are uniformly and objectively applied nationwide [6].

In order to identify the incidence of disability, data were collected from participants aged 65 years and older within the above-mentioned cohorts. In Japan, most individuals certified as 'cases requiring long-term care' are 65 years and older. Table 1 shows the number of subjects per region, as well as the data obtained within the first year of the observation. The smallest cohort consisted of 239 subjects, residing in Mie, while the largest consisted of 1,758, who resided in Gunma.

The earliest baseline data were collected in 2000 in Hiroshima, while the latest were obtained in 2008 in Tokyo-2. The cohorts were subsequently followed until 2012. Data regarding participants' deaths, changes of residence, and occurrence or non-occurrence of certified disability were gathered annually from public health centres of the participating municipalities. As an index of body build, baseline data on participants' height and weight were collected, and used to calculate body mass index (BMI, kg/m^2). Participants were classified as follows: normal or overweight (BMI = 18.5–27.5), obese (BMI >27.5), or emaciated (BMI <18.5). These cut-off points were determined according to a WHO report [7]. From 2008 onwards, follow-up data was obtained using the unified questionnaire.

All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo (nos. 1264 and 1326), the Tokyo Metropolitan Institute of Gerontology

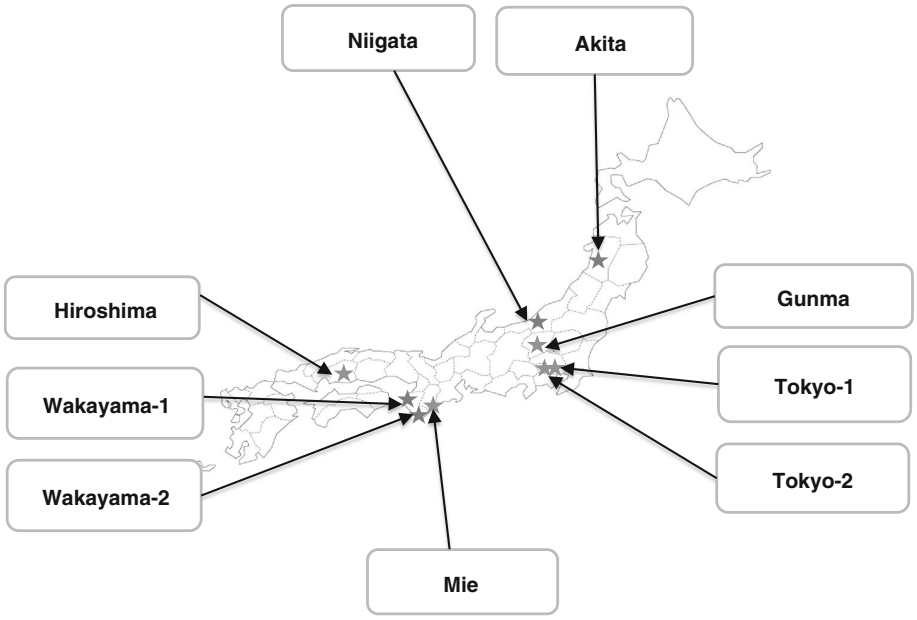


Fig. 1 Location of nine regions from which the study cohorts were selected

Table 1 Number of subjects classified by regions of each cohort

Region	Start year	Total	Men	Women
Tokyo-1	2005	1,332	461	871
Tokyo-2	2008	1,453	59	1,394
Wakayama-1 (Mountainous)	2005	610	239	371
Wakayama-2 (Coastal)	2006	357	129	228
Hiroshima	2000	1,341	351	990
Niigata	2007	805	343	462
Mie	2001	239	95	144
Akita	2006	559	223	336
Gunma	2005	1,758	805	953
Total		8,454	2,705	5,749

(no. 5), Wakayama (no. 373), the Radiation Effects Research Foundation (RP 03-89), Niigata University (no. 446), Mie University (nos. 837 and 139), Keio University (no. 16–20), and the National Center for Geriatrics and Gerontology (no. 249). Careful consideration was given to ensure the safety of the participants during all of the study procedures.

Statistical analysis

All statistical analyses were performed using STATA (STATA Corp., College Station, Texas, USA). Differences in proportions were compared using the chi-squared test. Differences in continuous variables were tested using an analysis of variance (ANOVA) with Scheffe's least significant difference test for post-hoc pairwise comparisons. To test the association between the occurrence of disability and other variables, Cox's proportional hazard regression analysis was used. Hazard ratios (HRs) were estimated using the occurrence of disability as an objective variable (0: non-occurrence, 1: occurrence) and the following explanatory variables: age (±1 year), gender (vs. female), body build (0: normal and overweight vs. 1: emaciation vs. 2: obesity), and regional differences (0: rural areas, including Wakayama-1, Wakayama-2, Niigata, Mie, Akita, and Gunma vs. 1: urban areas, including Tokyo-1, Tokyo-2, and Hiroshima). All p values and 95 % confidence intervals (CI) of two-sided analyses are presented.

Results

Table 2 shows the number of participants classified by age and gender. The majority of participants were 75–79 years old; two-thirds of the participants were women.

Selected characteristics of the study population, including age, height, weight, and BMI, are shown in Table 3. The mean values of age, height, and weight were significantly greater in women than in men ($p < 0.001$), but BMI did not significantly differ between men and women ($p = 0.479$).

The estimated incidence of disability is shown in Fig. 2. In total, the incidence of disability among individuals aged 65 years and older was 3.58/100 person-years (p-y) (p-y; men: 3.17/100 p-y; women: 3.78/100 p-y). The incidence of disability was 0.83/100 p-y, 1.70/100 p-y, 3.00/100 p-y,

Table 2 Number of subjects classified by age and gender

Age strata (years)	Total (%)	Men (%)	Women (%)
65–69	1,390 (16.4)	555 (20.5)	835 (14.5)
70–74	1,704 (20.2)	668 (24.7)	1,036 (18.0)
75–79	2,923 (34.6)	812 (30.0)	2,111 (36.7)
80–84	1,810 (21.4)	463 (17.1)	1,347 (23.4)
≥85	627 (7.4)	207 (7.7)	420 (7.3)
Total	8,454 (100.0)	2,705 (100.0)	5,749 (100.0)

Table 3 Baseline characteristics of subjects classified by age and gender

Variables	Men	Women	*p* (men vs. women)
Age (years)	75.3 (6.4)	76.5 (6.0)	<0.001
Height (cm)	160.5 (6.5)	147.7 (6.1)	<0.001
Weight (kg)	58.7 (9.1)	49.8 (8.4)	<0.001
BMI (kg/m^2)	22.7 (2.9)	22.8 (3.5)	0.479
Living in rural area (%)	84.8	58.5	<0.001

Values are represented as mean (standard deviation)

BMI body mass index

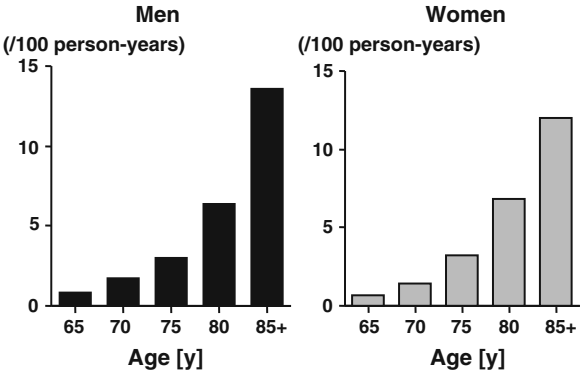


Fig. 2 Incidence of disability according to age and gender

6.36/100 p-y, and 13.54/100 p-y in 65–69-, 70–74-, 75–79-, 80–84-, and ≥85-year-old men, respectively. In women, the incidence of disability was 0.71/100 p-y, 1.40/100 p-y, 3.25/100 p-y, 6.85/100 p-y, and 12.01/100 p-y in the age ranges of 65–69, 70–74, 75–79, 80–84, and 85 or more years, respectively (Table 4).

Cox's proportional hazard regression analysis showed that occurrence of disability was significantly influenced by age, body build, and regional differences, but not gender (age, +1 years: hazard ratio 1.13, 95 % confidence interval 1.12–1.15, $p < 0.001$; sex, vs. female: 1.13, 0.97–1.31, $p = 0.125$; body build: emaciation: 1.24, 1.01–1.53, $p = 0.041$; body build; obesity: 1.36, 1.08–1.71, $p = 0.009$; residence, vs. living in rural areas: 1.59, 1.37–1.85, $p < 0.001$).

Discussion

Using the data of the LOCOMO study, we determined the incidence of disability and identified age, emaciation, obesity, and residence in rural areas as risk factors for the occurrence of disability. More specifically, we integrated data collected from subjects aged 65 and older in individual cohorts established in nine regions across Japan to determine the incidence of disability in the specified regions. We found an association between various risk factors and disability; these include age, emaciation, and obesity, as well as residence in rural areas.

The LOCOMO study was the first nation-wide prospective study to track a large number of the subjects from several population-based cohorts. The LOCOMO study aimed to integrate information from these cohorts, to prevent musculoskeletal diseases and subsequent disability. The data shed light on the prevalence and characteristics of targeted clinical symptoms such as knee pain or lumbar pain, or defined diseases such as knee osteoarthritis (KOA), lumbar spondylosis (LS), and osteoporosis (OP), as well as their prognosis in reference to either mortality or chances of developing a disability. In the present study, we also

Table 4 Hazard ratios (HRs) of potential risk factors for the occurrence and non-occurrence of disability

Disability (occurrence vs. non-occurrence)				
Explanatory variable	Reference	HR	95 % confidence interval	*p*
Age (years)	+1 year	1.13	1.12–1.15	<0.001***
Gender	0: men, 1: women	1.13	0.97–1.31	0.125
Body build	0: 18.5 ≤ BMI ≤ 27.5, 1: BMI < 18.5	1.24	1.01–1.53	0.041*
	0: 18.5 ≤ BMI ≤ 27.5, 2: BMI >27.5	1.36	1.08–1.71	0.009**
Type of residential area	0: urban area, 1: rural area	1.59	1.37–1.85	<0.001***

BMI body mass index

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

compared the above-mentioned symptoms, diseases, and prognoses between regions.

The overall incidence of disability among individuals aged 65 years and older was 3.58/100 person-years. When results from the present study are applied to the total age-sex distribution derived from the Japanese census in 2010 [1], it could be assumed that 1,110,000 people (410,000 men and 700,000 women) aged 65 years and older are newly affected by disability and require support. It has been reported that the total number of subjects who were certified as needing care increases annually [4]; however, few of these reports estimate the number of newly certified cases through a population-based cohort. Clarifying the incidence of disability and its risk factors was viewed as the first step toward preventing its occurrence.

Emaciation and obesity were both identified as risk factors for disability; thus, there appears to be a U-shaped association between BMI and disability as well as between BMI and mortality [8, 9]. According to the recent National Livelihood Survey, the leading cause of disabilities that require support and long-term care is cardiovascular disease (CVD), followed by dementia, senility, osteoarthrosis, and fractures [4]. Obesity is an established risk factor for chronic diseases, including hypertension, dyslipidemia, and diabetes mellitus, which increase the risk for CVD [10]; in turn, CVD causes ADL-related disabilities in older adults. In addition, numerous reports have shown an association between overweight or obesity and KOA [11–17]. In previous reports, we found a significant association between BMI and not only the presence of KOA, but also the occurrence and progression of KOA [18, 19]. In addition, emaciation is an established risk factor for OP and OP-related fractures [20]. OP might be related to low nutrition due to chronic wasting diseases.

The current study also found an association between living in a rural area and the occurrence of disability. There have been reports of regional differences in the certification rate of disability in Japan. For instance, Kobayashi reported a prefectural difference in the certification rate of disability, which was particularly prominent among individuals aged 75 years and older at lower nursing care levels in the long-term care insurance system [21]. In addition, Shimizutani et al. [22] pointed out that the financial condition of the insurer influenced the certification rate of disability. Further, Nakamura found that the certification of lower care levels was influenced by social and/or individual factors, such as the type of service provider, the application rate, and number of medical treatment recipients. However, certification of advanced nursing care levels was influenced by CVD and lifestyle-related diseases [23].

Other than differences in the social backgrounds of individuals in each prefecture, we posited that regional differences (rural or urban) in the occurrence of disability might be due to differences in the frequency of diseases and ailments that cause disability in each area. The prevalence of musculoskeletal diseases, such as KOA and LS, differs among mountainous, coastal, and urban areas [24]. Evidence also exists for regional differences in the incidence of hip fractures [25–27]. It was also found that mortality and incidence of ischemic stroke, which is related to CVD, was higher in the northeastern than in the southwestern part of Japan [28]. However, there is currently no information on regional differences in dementia prevalence and incidence in Japan. In general, differences in the frequency of diseases causing disability might influence regional differences in disability rates. In relation to this, in a future study on follow-up data from the LOCOMO study, it might be necessary to collect information on the prevalence and frequency of diseases that cause disability, such as musculoskeletal diseases, CVD, and dementia. This future study should also attempt to clarify mutual associations among risk factors for disability, so as to inform the development of measures for its primary prevention.

Despite its contribution to existing knowledge, the present study has several limitations. First, its sample does not truly represent the entire Japanese population, because our cohorts were not drawn from the northernmost and southernmost parts of Japan (e.g., Okinawa prefecture or Hokkaido prefecture). This limitation must be taken into consideration, especially when determining the generalisability of the results. However, the LOCOMO study is the first large-scale, population-based prospective study with approximately 9,000 participants aged 65 years and older. Second, data collected from the cohorts were not uniform, as certain information was obtained from some participants, but not others. For example, the X-ray examinations of subjects' knees were performed in Tokyo-1, Wakayama-1, Wakayama-2, Niigata, and Mie; lumbar spine X-ray examinations were performed in Tokyo-1, Wakayama-1, Wakayama-2, Hiroshima, and Mie. Therefore, we could not evaluate the presence or absence of KOA, LS, or OP as a possible cause of disability by using the data of the entire LOCOMO study. Further investigation following the integration of information on musculoskeletal disorders would enable us to evaluate all the factors that are associated with disability.

Nevertheless, our study has several strengths. As mentioned above, the large sample size is the study's biggest strength. The second strength is that we collected data from nine cohorts across Japan, which enabled us to compare regional differences in the incidence of disability. In addition, the variety of measures and assessments used in this study enabled us to collect a substantial amount of detailed information. However, given the fact that not all of the measures were administered in all cohorts, regional selection bias in the analysis should be considered when interpreting the results.

Acknowledgments This work was supported by grants from Grant-in-Aid for H17-Men-eki-009 (Director, Kozo Nakamura), H20-Choujyu-009 (Director, Noriko Yoshimura), H23-Choujyu-002 (Director, Toru Akune), and H-25-Chojyu-007 (Director, Noriko Yoshimura) of the Ministry of Health, Labour and Welfare. Further grants were provided by Scientific Research B24659317, B23390172, B20390182, and Challenging Exploratory Research 24659317 to Noriko Yoshimura; B23390357 and C20591737 to Toru Akune; B23390356, C20591774, and Challenging Exploratory Research 23659580 to Shigeyuki Muraki; Challenging Exploratory Research 24659666, 21659349, and Young Scientists A18689031 to Hiroyuki Oka, and collaborative research with NSF 08033011-00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. The sponsors did not contribute to the study design, data collection, data analysis, data interpretation, or the writing of the manuscript.

The authors wish to thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, and Mrs. Saeko Sahara for their assistance with data consolidation and administration.

Conflict of interest All authors declare no conflicts of interest.

References

1. Statistical Bureau, Ministry of Internal Affairs and Communication. Population Count based on the 2010 Census Released. http://www.stat.go.jp/english/data/kokusei/pdf/20111026.pdf. Accessed 26 Feb 2014
2. National Institute of Population and Society Research (2012) Population projections for Japan (January 2012): 2011–2060. http://www.ipss.go.jp/site-ad/index_english/esuikei/ppfj2012.pdf. Accessed 26 Feb 2014
3. Long-Term Care Insurance Act. http://www.japaneselawtranslation.go.jp/law/detail_main?id=94&vm=4&re=. Accessed 26 Feb 2014
4. Ministry of Health, Labour and Welfare (2010) Outline of the results of National Livelihood Survey. http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/4-2.html. Accessed 26 Feb 2014 (in Japanese)
5. Yoshimura N, Nakamura K, Akune T, Fujiwara S, Shimizu Y, Yoshida H, Omori G, Sudo N, Nishiwaki Y, Yoshida M, Shimokata H (2013) The longitudinal cohorts of motor system organ (LOCOMO) study. Nippon Rinsho 71:642–645 (in Japanese)
6. Ministry of Health, Labour and Welfare. Long-term care insurance in Japan. http://www.mhlw.go.jp/english/topics/elderly/care/index.html. Accessed 26 Feb 2014
7. WHO Expert Consultation (2004) Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 363:157–163
8. Zheng W, McLerran DF, Rolland B, Zhang X, Inoue M et al (2011) Association between body-mass index and risk of death in more than 1 million Asians. N Engl J Med 364:719–729
9. Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey J, Qizilbash N, Collins R, Peto R, Prospective Studies Collaboration (2009) Body-mass index and cause-specific mortality in 900,000 adults: collaborative analyses of 57 prospective studies. Lancet 373:1083–1096
10. Haslam DW, James WP (2005) Obesity. Lancet 366:1197–1209
11. Felson DT, Anderson JJ, Naimark A, Walker WM, Meenan RF (1988) Obesity and knee osteoarthritis: the Framingham study. Ann Intern Med 109:18–24
12. Hart DJ, Spector TD (1993) The relationship of obesity, fat distribution and osteoarthritis in the general population: the Chingford study. J Rheumatol 20:331–335
13. Van Saase JL, Vandenbroucke JP, Van Romunde LK, Valkenburg HA (1998) Osteoarthritis and obesity in the general population. A relationship calling for an explanation. J Rheumatol 15:1152–1158
14. Magliano M (2008) Obesity and arthritis. Menopause Int 14:149–154
15. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N et al (2008) OARSI recommendations for the management of hip and knee osteoarthritis, part II: OARSI evidence-based, expert consensus guidelines. Osteoarthr Cartil 16:137–162
16. Muraki S, Akune T, Oka H, Mabuchi A, En-yo Y, Yoshida M et al (2009) Association of occupational activity with radiographic knee osteoarthritis and lumbar spondylosis in elderly patients of population-based cohorts: a large-scale population-based study. Arthr Rheum 61:779–786
17. Lohmander LS, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Engstrom G (2009) Incidence of severe knee and hip osteoarthritis in relation to different measures of body mass: a population-based prospective cohort study. Ann Rheum Dis 68:490–496
18. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T (2011) Association of knee osteoarthritis with the accumulation of metabolic risk factors such as overweight, hypertension, dyslipidaemia, and impaired glucose tolerance in Japanese men and women: the ROAD study. J Rheumtol 38:921–930
19. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Akune T (2012) Accumulation of metabolic risk factors such as overweight, hypertension, dyslipidaemia, and impaired glucose tolerance raises the risk of occurrence and progression of knee osteoarthritis: a 3-year follow-up of the ROAD study. Osteoarthr Cartil 20:1217–1226
20. De Laet C, Kanis JA, Odén A, Johanson H, Johnell O, Delmas P, Eisman JA, Kroger H, Fujiwara S, Garnero P, McCloskey EV, Mellstrom D, Melton LJ 3rd, Meunier PJ, Pols HA, Reeve J, Silman A, Tenenhouse A (2005) Body mass index as a predictor of fracture risk: a meta-analysis. Osteoporos Int 16:1330–1338
21. Kobayashi T (2011) The relationship between variation in the requirement certification rate in prefectures and nursing care level in long-term care insurance. Otsuma Women's Univ Bull Fac Hum Relat 13:117–128 (in Japanese)
22. Shimitutani S, Inakura N (2007) Japan's public long-term care insurance and the financial condition of insurers: evidence from municipality-level data. Gov Audit Rev 14:27–40
23. Nakamura H (2006) Effect of received condition of long-term care insurance on the regional difference of the certification rate of the disability. J Health Welf Stat 53:1–7 (in Japanese)
24. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Yamamoto S, Ishibashi H, Kawaguchi H, Nakamura K, Akune T (2009) Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 27:620–628
25. Orimo H, Hashimoto T, Sakata K, Yoshimura N, Suzuki T, Hosoi T (2000) Trends in the incidence of hip fracture in Japan, 1987–1997: the third nationwide survey. J Bone Miner Metab 18:126–131
26. Yoshimura N, Suzuki T, Hosoi T, Orimo H (2005) Epidemiology of hip fracture in Japan: incidence and risk factors. J Bone Miner Metab 23:78–80
27. Orimo H, Yaegashi Y, Onoda T, Fukushima Y, Hosoi T, Sakata K (2009) Hip fracture incidence in Japan: estimates of new patients in 2007 and 20-year trends. Arch Osteoporos 4:71–77
28. Ueshima H, Ohsaka T, Asakura S (1986) Regional differences in stroke mortality and alcohol consumption in Japan. Stroke 17:19–24

Clin Rheumatol (2015) 34:1589–1597
DOI 10.1007/s10067-014-2687-y

ORIGINAL ARTICLE

Does osteophytosis at the knee predict health-related quality of life decline? A 3-year follow-up of the ROAD study

Shigeyuki Muraki · Toru Akune · Keiji Nagata · Yuyu Ishimoto · Munehito Yoshida · Fumiaki Tokimura · Sakae Tanaka · Hiroshi Kawaguchi · Kozo Nakamura · Hiroyuki Oka · Noriko Yoshimura

Received: 3 February 2014 / Revised: 18 May 2014 / Accepted: 20 May 2014 / Published online: 30 May 2014

Abstract The objective of the present longitudinal study was to clarify whether osteophytosis and joint space narrowing predict quality of life (QOL) decline using a longitudinal population-based cohort of the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study. The present study analyzed 1,525 participants who completed the radiographic examination at baseline and questionnaires regarding QOL at a 3-year follow-up (546 men and 979 women; mean age, 67.0±11.0 years). This study examined the associations of osteophyte area (OPA) and minimum joint space width (mJSW) in the medial compartment of the knee at baseline with pain and physical functional disability measured by the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). OPA and mJSW in the medial compartment of the knee were measured using a knee osteoarthritis (OA) computer-aided diagnosis system. Overall, OPA independently predicted physical functional disability after 3 years of follow-up. When analyzed in men and women separately, OPA, rather than mJSW, was an independent predictor for pain and physical functional disability after 3 years of follow-up in men. OPA, rather than mJSW, also predicted worsening of pain in men during the 3-year follow-up, whereas in women, mJSW, rather than OPA, predicted worsening of pain. In conclusion, the present longitudinal study using a large-scale population from the ROAD study found gender differences in the association of osteophytosis and joint space narrowing with pain and physical functional disability.

Keywords Epidemiology · Longitudinal Studies · Osteoarthritis · Pain · WOMAC

Electronic supplementary material The online version of this article (doi:10.1007/s10067-014-2687-y) contains supplementary material, which is available to authorized users.

S. Muraki (✉) · T. Akune
Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan
e-mail: murakis-ort@h.u-tokyo.ac.jp

K. Nagata · Y. Ishimoto · M. Yoshida
Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan

F. Tokimura
Department of Orthopaedic Surgery, Tokyo Metropolitan Geriatric Medical Center, Tokyo, Japan

S. Tanaka · H. Kawaguchi
Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

K. Nakamura
National Rehabilitation Center for Persons with Disabilities, Saitama, Japan

H. Oka · N. Yoshimura
Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

Introduction

Knee osteoarthritis (OA) is a major public health issue causing chronic pain and disability [1–3]. The prevalence of radiographically confirmed knee OA is high in Japan [4], with 25,300,000 persons aged 40 years and older estimated to experience radiographic knee OA [5]. According to the recent Japanese National Livelihood Survey of the Ministry of Health, Labour and Welfare, osteoarthritis is ranked fourth among diseases that cause disabilities that subsequently require support with activities of daily living [6].

The principal clinical symptoms of knee OA are pain and physical functional disability [7], but the correlation of these symptoms with radiographic severity of knee OA is controversial [4, 8–10]. In terms of disease-specific scales for

estimating pain and physical functional disability due to knee OA, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) has been used for Caucasians [11] and Asians [12, 13], although these reports were not population-based studies. Furthermore, there is little information on the impact of knee OA on incident pain and physical functional disability using WOMAC in Japan, although reports from a population survey suggest that the disease pattern differs among races [14–16].

Knee OA is characterized by the pathological features of osteophytosis and joint space narrowing, but there is controversy regarding the importance of osteophytes. Nevertheless, hand and hip joint researchers and clinicians have argued that separate radiographic features should be recorded and may be more meaningful than overall composite scores such as the Kellgren-Lawrence (KL) scale [17]. Furthermore, a previous study showed that osteophytes performed better as a primary diagnostic feature than joint space narrowing in cross-sectional knee OA epidemiologic studies [18]. However, most conventional systems for grading radiographic severity have been categorical grades, such as KL grading [19], which is unable to assess osteophytosis and joint space narrowing individually. Several studies have shown that knee OA had a strong effect on quality of life (QOL) [13, 20–22], but in these studies, knee OA was defined by categorical grades such as KL score or American College of Rheumatology grade, total knee arthroplasty, and self-administered questionnaires. In addition, osteophytosis and joint space narrowing were separately evaluated using a radiographic atlas of individual features published by the Osteoarthritis Research Society International in 1995 [23] and revised in 2007 [24]. However, the grading is still limited in reproducibility and sensitivity due to the subjective judgment of individual observers and the categorical classification into four grade scales (0–3). To overcome this problem, osteophyte area (OPA) or joint space width should be evaluated using a fully automatic system [25].

The objective of this study was to clarify whether osteophytosis and joint space narrowing at the knee independently predict decline of QOL measured by WOMAC pain and physical function score during a 3-year follow-up among Japanese men and women using a fully automatic system to measure OPA and joint space width in the longitudinal, population-based cohort from the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.

Materials and methods

Study sample The ROAD study is a nationwide prospective study designed to establish epidemiologic indices for the evaluation of clinical evidence to allow for the development of disease-modifying treatments for bone and joint disorders (with OA and osteoporosis as the representative bone and joint diseases). The ROAD study consists of population-based cohorts in several Japanese communities. A detailed profile of the ROAD study has been published previously [4, 5, 26]; therefore, only a brief summary is provided here. To date, the ROAD study has completed the creation of a baseline database including clinical and genetic information for 3,040 inhabitants (1,061 men and 1,979 women) ranging in age from 23 to 95 years (mean, 70.6 years). Participants were recruited from resident registration listings in three communities: an urban region in Itabashi, Tokyo; a mountainous region in Hidakagawa, Wakayama; and a seacoast region in Taiji, Wakayama. All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology. Anthropometric measurements, including height and weight and body mass index (BMI) (weight [kg]/height^2 [m^2]), were calculated. Grip strength was measured on bilateral sides using a TOEI LIGHT handgrip dynamometer (TOEI LIGHT CO., LTD, Saitama, Japan), and the better measurement was used to characterize maximum muscle strength.

Radiographic assessment All participants underwent radiographic examination of both knees using an anterior-posterior view with weight-bearing and foot map positioning by experienced radiological technologists. The beam was positioned parallel to the floor with no angle and aimed at the joint space. We used fluoroscopic guidance with an anterior-posterior X-ray beam to properly visualize the joint space and to centralize the patella over the lower end of the femur. The images were downloaded into Digital Imaging and Communication in Medicine (DICOM) format files. A single experienced orthopedist (S.M.) read the knee radiographs without knowledge of participant clinical status using the KL radiographic atlas for overall knee radiographic grades [19]. Knee OA was defined as KL ≥2. Medial compartment minimum joint space width (mJSW) and medial tibial OPA were measured with the knee osteoarthritis computer-aided diagnosis (KOACAD) system bilaterally. The knee with the least mJSW was defined as the designated knee for each participant. The KOACAD system has been previously described in detail [25, 27, 28]. The KOACAD system is a fully automatic system capable of quantifying the major features of knee OA on standard radiographs. This system allows for objective, accurate, and simple assessment of the structural severity of knee OA in general clinical practice. The KOACAD system was programmed to measure OPA at the medial tibia and mJSW in the medial and lateral compartments using digitized knee radiographs. The KOACAD system was applied to the DICOM data by the experienced orthopedist who developed this system (H.O.), and there is strong reliability for this measurement [25]. Reference values for OPA and mJSW by gender and age strata in Japan using the KOACAD

system have been published previously [28]. Lateral OA was defined as being present when a knee had a KL grade ≥2 [19] and lateral joint space narrowing score ≥1 on a 0–3 scale according to the Osteoarthritis Research Society International atlas [24].

Instruments All 3,040 subjects were invited to attend a follow-up interview between 2008 and 2010. We used the WOMAC at the follow-up study to evaluate QOL. The WOMAC is a 24-item OA-specific index consisting of three domains: pain, stiffness, and physical function. Each of these 24 items is graded on either a 5-point Likert scale or a 100-mm visual analog scale [11, 29]. The Likert scale (version LK 3.0) was used in the present study. The domain score ranges from 0 to 20 for pain, 0 to 8 for stiffness, and 0 to 68 for physical function. Japanese versions of the WOMAC have also been validated [30].

Statistical analysis Differences in age, height, weight, BMI, grip strength, OPA, mJSW, and WOMAC scores between men and women were examined using a non-paired student *t* test. The associations of mJSW and OPA with pain and physical functional disability after 3 years were determined by using multiple regression analysis after adjustment for age, BMI, gender, grip strength, and pain score at baseline; after adjustment for age, BMI, gender, grip strength, and physical function score at baseline, respectively, in the overall population; and after adjustment for age, BMI, grip strength, and pain score at baseline and after adjustment for age, BMI, grip strength, and physical function score at baseline, respectively, in men and women. In addition, to determine the independent association of OPA and mJSW with pain and physical function scores, multiple regression analysis was used with age, BMI, gender, grip strength, pain score at baseline, OPA, and mJSW and with age, BMI, gender, grip strength, physical function score at baseline, OPA, and mJSW, respectively, as explanatory variables in the overall population, and with age, BMI, grip strength, pain score at baseline, OPA, and mJSW and with age, BMI, grip strength, physical function score at baseline, OPA, and mJSW, respectively, as explanatory variables in men and women. We classified men and women separately into three groups based on grip strength: <20, ≥20 to <30, and ≥30 and examined the associations of BMI, OPA, and mJSW with pain, using multiple regression analysis with age, BMI, OPA, mJSW, and pain score at baseline as explanatory variables. We also calculated changes of scores as follows: "scores at follow-up−scores at baseline" and determined the association of OPA and mJSW with changes of pain and physical function scores after adjustment for age, BMI, gender, grip strength, and pain score at baseline; after adjustment for age, BMI, gender, grip strength, and physical function score at baseline, respectively, in the overall population; and after adjustment for age, BMI, grip strength, and pain score at baseline and after adjustment for age, BMI, grip strength, and physical function score at baseline, respectively, in men and women. In addition, to determine independent associations of OPA and mJSW with changes of pain and physical function scores, multiple regression analysis was used with age, BMI, gender, grip strength, pain score at baseline, OPA, and mJSW and with age, BMI, gender, grip strength, physical function score at baseline, OPA, and mJSW, respectively, as explanatory variables in the overall population and with age, BMI, grip strength, pain score at baseline, OPA, and mJSW and with age, BMI, grip strength, physical function score at baseline, OPA, and mJSW, respectively, as explanatory variables in men and women. Data analyses were performed using SAS version 9.0 (SAS Institute Inc., Cary, NC).

Results

Of the 3,040 subjects in the baseline study in 2005–2007, 125 had died by the time of the review 3 years later, 123 did not participate in the follow-up study due to bad health, 69 had moved away, 83 declined the invitation to attend the follow-up study, and 155 did not participate in the follow-up study for other reasons. Among the 2,485 subjects who did participate in the follow-up study, we excluded 39 subjects younger than 40 years at baseline. Those participating in the follow-up study were younger than those who did not survive or who did not participate for other reasons (responders 68.6 years, non-responders 75.1 years; $P<0.0001$). The follow-up study participants also were significantly more likely to be women (responders 66.3 % women, non-responders 61.8 % women; $P=0.03$) and were significantly more likely to have knee OA at the baseline examination than either those who did not survive to follow-up or those who did not participate for other reasons (responders 51.5 %, non-responders 60.9 %; $P<0.0001$). Among them, 1,578 subjects provided complete questionnaires of WOMAC both at baseline and follow-up. We excluded 3 subjects who did not undergo plain radiography at the knee and 17 subjects who underwent total knee arthroplasty before the follow-up study. We also excluded 12 subjects whose X-rays were too obscure to measure mJSW and OPA and 21 subjects who had lateral knee OA, leaving a total of 1,525 subjects (546 men and 979 women). The mean duration between baseline and follow-up was 3.3±0.6 years.

Characteristics of the 1,525 participants in the present study are shown in Table 1. BMI was higher in men than women. The prevalence of knee OA was significantly higher in women than men. The OPA was significantly larger and mJSW was significantly narrower in women than men. The WOMAC pain score was similar in men and women, whereas the WOMAC physical function score was worse in women than men, both at baseline and follow-up.

Table 1 Characteristics of subjects

	Overall	Men	Women	*p* value
N	1,525	546	979	
Age (years)	67.0±11.0	68.2±10.7	66.3±11.1	0.001
Height (cm)	155.3±8.8	163.3±6.4	150.8±6.4	<0.0001
Weight (kg)	55.5±10.4	62.2±10.3	51.8±8.5	<0.0001
BMI (kg/m^2)	22.9±3.3	23.3±3.1	22.7±3.3	0.0027
Grip strength (kg)	27.2±9.4	35.4±8.7	22.7±6.4	<0.0001
Knee OA (%)	48.8	38.5	54.5	<0.0001
OPA (mm^2)	3.56±8.43	1.79±5.47	4.54±9.56	<0.0001
mJSW (mm)	2.67±0.94	2.99±0.88	2.50±0.92	<0.0001
WOMAC at baseline				
Pain	1.13±2.20	1.03±2.06	1.18±2.27	0.1753
Physical function	3.05±6.68	2.59±5.74	3.30±7.14	0.0328
WOMAC at follow-up				
Pain	1.82±2.81	1.74±2.69	1.87±2.88	0.3881
Physical function	5.56±9.61	4.79±8.34	5.99±10.22	0.0137

Knee OA was defined as Kellgren-Lawrence grade ≥2 at baseline; except where otherwise indicated, the values at baseline was shown

BMI body mass index, *OA* osteoarthritis, *OPA* osteophyte area, *mJSW* minimum joint space width, *WOMAC* Western Ontario and McMaster Universities Osteoarthritis Index

First, we analyzed the associations of age, BMI, and grip strength with WOMAC pain and physical function scores in men and women (Table 2). Age and grip strength were significantly associated with pain and physical function in men and women, while BMI was significantly associated with pain and physical function in women, but not in men.

Multiple regression analysis after adjustment for age, BMI, grip strength, and pain score at baseline showed that, overall, OPA and mJSW were significant predictors for pain (Table 3). To assess whether OPA and mJSW independently predicted pain, we used multiple regression analysis with age, BMI, grip strength, pain score at baseline, OPA, and mJSW as explanatory variables and found that the association of OPA with pain score after 3 years disappeared, whereas mJSW was an independent predictor for pain after 3 years. When analyzed in men and women, separately, OPA was an independent predictor for pain in men, but mJSW was not. In women, mJSW was an independent predictor for pain, but OPA was not.

In terms of physical function, multiple regression analysis after adjustment for age, BMI, grip strength, and physical function score at baseline showed that OPA and mJSW were significant predictors for physical functional disability (Table 4). To assess whether OPA and mJSW independently predicted physical functional disability, we used multiple regression analysis with age, BMI, grip strength, physical function score at baseline, OPA, and mJSW as explanatory variables and found that OPA and mJSW were independent predictors for physical functional disability. When analyzed in men and women separately, OPA was an independent predictor for physical functional disability in men, but mJSW was not. In women, mJSW was an independent predictor for physical functional disability, but OPA was not.

To clarify the association of OPA, mJSW, and BMI with pain according to muscle strength, men and women were separated into three groups based on grip strength: <20, ≥20 to <30, and ≥30 and the associations of BMI, OPA, and mJSW with pain were examined, using multiple regression analysis with age, BMI, OPA, mJSW, and pain score at baseline as explanatory variables (Supplementary Table 1). In women with grip strength <20, mJSW was significantly associated with pain and BMI tended to be associated with pain, but OPA was not. In men with grip strength <20, BMI, OPA, and mJSW were not significantly associated with pain, likely because only nine men had a grip strength <20. In women with grip strength ≥20 to <30, mJSW and BMI was significantly associated with pain, while OPA was not. In men with grip strength ≥20 to <30, BMI was significantly associated with pain, while OPA and mJSW were not. In men and women with grip strength >30, OPA was significantly associated with pain, while mJSW and BMI were not. We also

Table 2 Effect of age, BMI, and grip strength at baseline on WOMAC pain and physical function scores after 3 years

	Pain		Physical function	
	Regression coefficient (95 % CI)	*P* value	Regression coefficient (95 % CI)	*P* value
Men				
Age (years)	0.05 (0.03 to 0.07)	<0.0001	0.23 (0.17 to 0.29)	<0.0001
BMI (kg/m^2)	0.05 (−0.02 to 0.12)	0.1616	0.17 (−0.06 to 0.39)	0.1459
Grip strength (kg)	−0.05 (−0.07 to −0.02)	0.0003	−0.26 (−0.34 to −0.19)	<0.0001
Women				
Age (years)	0.06 (0.05 to 0.08)	<0.0001	0.33 (0.28 to 0.39)	<0.0001
BMI (kg/m^2)	0.20 (0.14 to 0.25)	<0.0001	0.66 (0.47 to 0.85)	<0.0001
Grip strength (kg)	−0.10 (−0.12 to −0.07)	<0.0001	−0.44 (−0.54 to −0.35)	<0.0001

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, *CI* confidence interval, *BMI* body mass index

Table 3 Effect of OPA and mJSW at baseline on WOMAC pain scores after 3 years

	Crude regression coefficient[b] (95 % CI)	P value	Adjusted regression coefficient[a] (95 % CI)	*P* value	Adjusted regression coefficient[b] (95 % CI)	*P* value	Standardized beta
Overall							
OPA (mm^2)	0.08 (0.06 to 0.09)	<0.0001	0.02 (0.006 to 0.04)	0.0051	0.01 (−0.003 to 0.03)	0.1036	0.04
mJSW (mm)	−0.76 (−0.90 to −0.61)	<0.0001	−0.30 (−0.44 to −0.16)	<0.0001	−0.26 (−0.41 to −0.12)	0.0005	−0.09
Men							
OPA (mm^2)	0.09 (0.04 to 0.13)	<0.0001	0.05 (0.01 to 0.08)	0.0078	0.05 (0.01 to 0.09)	0.0127	0.1
mJSW (mm)	−0.45 (−0.71 to −0.20)	0.0005	−0.11 (−0.33 to 0.12)	0.3466	0.02 (−0.22 to 0.27)	0.8574	0.007
Women							
OPA (mm^2)	0.08 (0.06 to 0.09)	<0.0001	0.02 (−0.0008 to 0.03)	0.0623	0.006 (−0.01 to 0.02)	0.4789	0.02
mJSW (mm)	−0.96 (−1.15 to −0.78)	<0.0001	−0.41 (−0.58 to −0.23)	<0.0001	−0.39 (−0.57 to −0.20)	<0.0001	−0.12

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, *CI* confidence interval, *OPA* osteophyte area, *mJSW* minimum joint space width

[a] Adjusted regression coefficients for pain scores were calculated by multiple regression analysis after adjustment for age, BMI, gender, grip strength, and pain score at baseline in the overall population and after adjustment for age, BMI, grip strength, and pain score at baseline in men and women

[b] Adjusted regression coefficients for pain scores were calculated by multiple regression analysis with age, BMI, gender, grip strength, pain score at baseline, OPA, and mJSW as explanatory variables in the overall population and with age, BMI, grip strength, pain score at baseline, OPA, and mJSW as explanatory variables in men and women

examined the association of OPA, mJSW, and BMI with physical function disability according to muscle strength (Supplementary Table 2). Results were similar to findings for pain.

To examine whether OPA and mJSW predicted worsening of pain during the 3-year follow-up, we calculated differences of the WOMAC pain scores between baseline and follow-up (Table 5). In the overall population, mJSW was a significant predictor for worsening of pain after adjustment for age, BMI, gender, and pain score at baseline, whereas OPA was not. When analyzed in men and women separately, OPA was a significant predictor for worsening of pain in men, whereas mJSW was a significant predictor for worsening of pain in women.

We also examined whether OPA and mJSW predicted worsening of physical functional disability during the 3-year follow-up (Table 6). In the overall population, OPA and mJSW were significant predictors for worsening of physical functional disability after adjustment for age, BMI, gender, grip strength, and physical function score at baseline. To

Table 4 Effect of OPA and mJSW at baseline on WOMAC physical function scores after 3 years

	Crude regression coefficient[b] (95 % CI)	*P* value	Adjusted regression coefficient[a] (95 % CI)	*P* value	Adjusted regression coefficient[b] (95 % CI)	*P* value	Standardized beta
Overall							
OPA (mm^2)	0.34 (0.29 to 0.40)	<0.0001	0.09 (0.04 to 0.14)	0.0002	0.05 (0.0004 to 0.10)	0.0480	0.04
mJSW (mm)	−3.24 (−3.73 to −2.75)	<0.0001	−1.36 (−1.80 to −0.92)	<0.0001	−1.22 (−1.68 to −0.76)	<0.0001	−0.12
Men							
OPA (mm^2)	0.35 (0.23 to 0.48)	<0.0001	0.19 (0.08 to 0.30)	0.0008	0.14 (0.02 to 0.26)	0.0204	0.09
mJSW (mm)	−2.21 (−2.99 to −1.44)	<0.0001	−1.07 (−1.77 to −0.37)	0.0027	−0.69 (−1.46 to 0.07)	0.0758	−0.07
Women							
OPA (mm^2)	0.34 (0.27 to 0.40)	<0.0001	0.06 (0.009 to 0.12)	0.0225	0.03 (−0.03 to 0.08)	0.3305	0.03
mJSW (mm)	−3.86 (−4.51 to −3.20)	<0.0001	−1.49 (−2.05 to −0.92)	<0.0001	−1.41 (−2.00 to −0.82)	<0.0001	−0.13

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, *CI* confidence interval, *OPA* osteophyte area, *mJSW* minimum joint space width

[a] Adjusted regression coefficients for physical function score were calculated by multiple regression analysis after adjustment for age, BMI, gender, grip strength, and physical function score at baseline in the overall population and after adjustment for age, BMI, grip strength, and physical function score at baseline in men and women

[b] Adjusted regression coefficients for physical function score were calculated by multiple regression analysis with age, BMI, gender, grip strength, physical function score at baseline, OPA, and mJSW as explanatory variables in the overall population and with age, BMI, grip strength, physical function score at baseline, OPA, and mJSW as explanatory variables in men and women

examine whether OPA and mJSW independently predicted worsening of physical functional disability, we used multiple regression analysis with age, BMI, gender, grip strength, physical function score at baseline, OPA, and mJSW as explanatory variables and found that mJSW was an independent predictor for worsening of physical functional disability, but the significant association of OPA disappeared. When analyzed in men and women separately, after adjustment for age, BMI, grip strength, and physical function scores at baseline, OPA and mJSW were significant predictors for worsening of physical functional disability in men; in women, mJSW was a significant predictor for worsening of physical functional disability, but OPA was not. To examine whether OPA and mJSW independently predicted worsening of physical functional disability in men, we used multiple regression analysis with age, BMI, grip strength, physical function score at baseline, OPA, and mJSW as explanatory variables and found that the significant association of OPA and mJSW with worsening in physical function disappeared.

Discussion

This is the first large-scale study to examine whether osteophytosis and joint space narrowing independently predict QOL decline measured by WOMAC pain and physical function score in a longitudinal model. In addition, osteophytosis and joint space narrowing were estimated not by categorical grade but by continuous values such as OPA and mJSW at the knee. In the present study, OPA, rather than mJSW, was an independent predictor for pain and physical functional disability after 3 years of follow-up in men. OPA, rather than mJSW, also predicted worsening of pain in men during the 3-year follow-up, whereas mJSW, rather than OPA, predicted worsening of pain in women.

Previous studies have shown that knee OA has a strong effect on QOL [13, 20–22]; however, the knee OA was defined by KL grade or other categorical methods. KL grade is the most conventional system to grade radiographic severity of knee OA, but in this categorical system, osteophyte formation and joint space narrowing are not assessed separately. Thus, we cannot clarify whether osteophytosis and joint space narrowing have distinct effects on QOL. In addition, our previous cross-sectional study showed that osteophytosis was not strongly related to joint space narrowing on plain radiographs [31]. Furthermore, our experimental mouse model for OA identified a cartilage-specific molecule, carminerin, that regulates osteophytosis without affecting joint cartilage destruction during OA progression [32, 33]. This accumulating evidence indicates that osteophytosis and joint space narrowing may have distinct etiologic mechanisms and their progression may be neither constant nor proportional. Thus, to examine factors associated with knee OA, these two OA features should be assessed separately. Furthermore, because categorical methods are statistically less powerful than continuous methods, the association between knee OA and QOL might have been underestimated in previous studies. In addition, most studies regarding the association of knee OA with QOL were cross-sectional designs; thus, a causal relationship could not be clarified. So far, the role of the osteophytes in OA is controversial, with several researchers believing that osteophytes are merely a reflection of age and not associated with any of the clinical symptoms of OA, though few reported data support or refute this argument. This study was the first longitudinal model to report that osteophytosis, rather than mJSW, predicted QOL decline in men.

Table 5 Effect of OPA and mJSW at baseline on worsening of WOMAC pain scores after 3 years

	Crude regression coefficient[a] (95 % CI)	*P* value	Adjusted regression coefficient (95 % CI)	*P* value
Overall				
OPA (mm^2)	0.01 (−0.004 to 0.03)	0.1443	–	–
mJSW (mm)	−0.16 (−0.29 to −0.03)	0.0132	−0.30 (−0.44 to −0.16)	<0.0001
Men				
OPA (mm^2)	0.04 (0.006 to 0.08)	0.0209	0.05 (0.01 to 0.08)	0.0078
mJSW (mm)	−0.06 (−0.28 to 0.15)	0.5684	–	–
Women				
OPA (mm^2)	0.006 (−0.01 to 0.02)	0.4880	–	–
mJSW (mm)	−0.24 (−0.41 to −0.07)	0.006	−0.41 (−0.58 to −0.23)	<0.0001

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, *CI* confidence interval, *OPA* osteophyte area, *mJSW* minimum joint space width

[a] Adjusted regression coefficients for change of scores were calculated by multiple regression analysis after adjustment for age, BMI, gender, grip strength, and pain score at baseline in the overall population and after adjustment for age, BMI, grip strength, and pain score at baseline in men and women

Table 6 Effect of OPA and mJSW at baseline on worsening of WOMAC physical function scores after 3 years

	Crude regression coefficient[b] (95 % CI)	*P* value	Adjusted regression coefficient[a] (95 % CI)	*P* value	Adjusted regression coefficient[b] (95 % CI)	*P* value	Standardized beta
Overall							
OPA (mm^2)	0.10 (0.05 to 0.14)	<0.0001	0.05 (0.002 to 0.10)	0.0393	0.01 (−0.04 to 0.06)	0.6078	0.01
mJSW (mm)	−1.44 (−1.84 to −1.04)	<0.0001	−1.14 (−1.58 to −0.69)	<0.0001	−1.10 (−1.57 to −0.63)	<0.0001	−0.14
Men							
OPA (mm^2)	0.18 (0.07 to 0.29)	0.0012	0.14 (0.03 to 0.26)	0.012	0.10 (−0.02 to 0.23)	0.1095	0.08
mJSW (mm)	−1.27 (−1.95 to 0.59)	0.0003	−0.93 (−1.65 to −0.21)	0.0113	−0.66 (−1.45 to 0.13)	0.1021	−0.08
Women							
OPA (mm^2)	0.08 (0.03 to 0.13)	0.0024	0.03 (−0.02 to 0.09)	0.2521	–	–	–
mJSW (mm)	−1.58 (−2.10 to −1.05)	<0.0001	−1.25 (−1.82 to −0.68)	<0.0001	–	–	

WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, *CI* confidence interval, *OPA* osteophyte area, *mJSW* minimum joint space width

[a] Adjusted regression coefficients for changes in physical function scores were calculated by multiple regression analysis after adjustment for age, BMI, gender, grip strength, and physical function score at baseline overall and after adjustment for age, BMI, grip strength, and physical function score at baseline in men and women

[b] Adjusted regression coefficients for changes in physical function scores were calculated by multiple regression analysis with age, BMI, gender, grip strength, OPA, mJSW, and physical function score at baseline as overall explanatory variables and with age, BMI, grip strength, OPA, mJSW, and physical function score at baseline as explanatory variables in men

The association of osteophytosis with QOL may be complex. Osteophytes may not have any primary effect themselves but rather serve as markers for factors that strongly affect QOL decline. First, osteophytosis appears to start from activation of periosteal layers, with initial generation of chondrophytes and subsequent calcification to real osteophytes. The process is probably an adaptive reaction of the joint to cope with joint instability, and thus, OPA may indicate the severity of joint instability [34], which might lead to pain and physical functional disability, particularly in men. In addition, it is possible that osteophytosis is strongly associated with patellofemoral disease [35], which is associated with knee pain [36]. This is an area where further research would be useful. Nevertheless, our results indicate that the presence or absence of osteophytosis, rather than joint space narrowing, is an appropriate method to predict QOL decline in men.

The present study revealed gender differences in the associations of osteophytosis and joint space narrowing with pain and physical functional disability. Joint space narrowing was an independent predictor for QOL decline measured by WOMAC pain and physical function scores in women, but not in men. Our previous cross-sectional study also showed that the odds ratio of knee pain for KL grade 3 or 4 knee OA was approximately twice as high in women as in men [4]. Considering the definition of KL grade [19], this finding may indicate that joint space narrowing is more strongly associated with pain in women than men. At the same time, osteophytosis is an independent predictor for QOL decline measured by the WOMAC pain and physical function scores in men, but not in women. As mentioned above, osteophytosis may represent joint instability or patellofemoral disease, which may be more strongly associated with pain and physical function than joint space narrowing due to cartilage loss in men. These findings may be partly explained by the lower muscle mass in women compared with men. Previous reports have shown that muscle mass is also associated with QOL [37, 38]. BMI also has different effect on QOL between men and women. To clarify the effect of muscle strength on the association of OPA, mJSW, and BMI with QOL, we classified subjects according to grip strength and examined the association of OPA, mJSW, and BMI with WOMAC pain score. In both men and women with strong muscle strength, OPA was associated with pain rather than mJSW or BMI, whereas in those with weaker muscle strength, mJSW and BMI were associated with pain rather than OPA. We also examined the association of OPA, mJSW, and BMI with WOMAC physical function score according to grip strength, and results were similar to those for pain. This means that muscle strength, rather than gender itself, may affect differences between men and women in the association of mJSW and OPA with QOL.

There is a limitation in the present study. We did not include other weight-bearing joints that can have OA, such as hip OA, in the analysis, although such disorders may also affect QOL decline. However, the prevalence of KL grade 3 or 4 hip OA is 1.4 and 3.5 % in Japanese men and women [39], respectively, which is much less than the prevalence of KL grade 3 or 4 knee OA (13.5 and 24.6 % in Japanese men and women, respectively) [4]. Thus, it is possible that hip OA would not strongly affect the results of the present study.

In conclusion, the present longitudinal study using a large-scale population from the ROAD study revealed

that osteophytosis is a predictor for QOL decline in men. We also revealed gender differences in the association of osteophytosis and joint space narrowing with QOL decline. Future studies, along with longitudinal surveys in the ROAD study, will help further the understanding of osteophytosis and joint space narrowing mechanisms at the knee and their relationship with QOL.

Acknowledgments The authors wish to thank Dr. Seizo Yamamoto, Hideaki Ishibashi, Yorito Anamizu, and the other members of the Department of Orthopedics at Tokyo Metropolitan Geriatric Medical Center; Mr. Kutsuma and the other members of the Department of Radiology at Tokyo Metropolitan Geriatric Medical Center; Mrs. Tomoko Takijiri and the other members of the Public Office in Hidakagawa Town; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, and the other members of the Public Office in Taiji Town, for their assistance in locating and scheduling participants for examinations.

Funding This work was supported by Grants-in-Aid for Scientific Research (S19109007, B20390182, C20591737, C20591774), for Young Scientists (A18689031), and for Exploratory Research (19659305) from the Japanese Ministry of Education, Culture, Sports, Science and Technology; grants (H17-Men-eki-009, H18-Choujyu-037, H20-Choujyu-009, H21-Chouju-Wakate-011, and H22-Chouju-Wakate-007) from the Ministry of Health, Labor and Welfare, a Research Aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006–1) and Grant No. 166 from the Japan Orthopaedics and Traumatology Foundation are also acknowledged.

Disclosures All authors have no conflicts of interest to state.

References

1. Sharma L, Kapoor D (2007) Epidemiology of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM (eds) Osteoarthritis diagnosis and medical/surgical management, 4th edn. Lippincott Williams & Wilkins, Philadelphia, pp 3–26
2. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, Kelly-Hayes M, Wolf PA, Kreger BE, Kannel WB (1994) The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health 84:351–358
3. Felson DT, Zhang Y (1998) An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 41: 1343–1355
4. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, Saika A, Suzuki T, Yoshida H, Ishibashi H, Yamamoto S, Nakamura K, Kawaguchi H, Yoshimura N (2009) Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthr Cartil 17:1137–1143
5. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Yamamoto S, Ishibashi H, Kawaguchi H, Nakamura K, Akune T (2009) Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: the Research on Osteoarthritis/osteoporosis Against Disability (ROAD). J Bone Miner Metab 27:620–628
6. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey (2007) Available at http://www.mhlw.go.jp/toukei/list/20-19-1.html
7. Linaker CH, Walker-Bone K, Palmer K, Cooper C (1999) Frequency and impact of regional musculoskeletal disorders. Baillieres Clin Rheumatol 13:197–215
8. Duncan R, Peat G, Thomas E, Hay E, McCall I, Croft P (2007) Symptoms and radiographic osteoarthritis: not as discordant as they are made out to be? Ann Rheum Dis 66:86–91
9. Hannan MT, Felson DT, Pincus T (2000) Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. J Rheumatol 27:1513–1517
10. Neogi T, Felson D, Niu J, Nevitt M, Lewis CE, Aliabadi P, Sack B, Torner J, Bradley L, Zhang Y (2009) Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. BMJ 339:b2844
11. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW (1988) Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 15:1833–1840
12. Thumboo J, Chew LH, Soh CH (2001) Validation of the Western Ontario and McMaster University osteoarthritis index in Asians with osteoarthritis in Singapore. Osteoarthr Cartil 9:440–446
13. Woo J, Lau E, Lee P, Kwok T, Lau WC, Chan C, Chiu P, Li E, Sham A, Lam D (2004) Impact of osteoarthritis on quality of life in a Hong Kong Chinese population. J Rheumatol 31:2433–2438
14. Dillon CF, Rasch EK, Gu Q, Hirsch R (2006) Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991–94. J Rheumatol 33:2271–2279
15. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, Fang F, Schwartz TA, Abbate LM, Callahan LF, Kalsbeek WD, Hochberg MC (2007) Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol 34:172–180
16. Yoshida S, Aoyagi K, Felson DT, Aliabadi P, Shindo H, Takemoto T (2002) Comparison of the prevalence of radiographic osteoarthritis of the knee and hand between Japan and the United States. J Rheumatol 29:1454–1458
17. Croft P, Cooper C, Wickham C, Coggon D (1990) Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 132:514–522
18. Spector TD, Hart DJ, Byrne J, Harris PA, Dacre JE, Doyle DV (1993) Definition of osteoarthritis of the knee for epidemiologic studies. Ann Rheum Dis 52:790–794
19. Kellgren JH, Lawrence JS (eds) (1963) The epidemiology of chronic rheumatism: atlas of standard radiographs of arthritis. Blackwell Scientific, Oxford
20. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML (1999) Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford) 38:870–877
21. Hill CL, Parsons J, Taylor A, Leach G (1999) Health related quality of life in a population sample with arthritis. J Rheumatol 26: 2029–2035
22. Muraki S, Akune T, Oka H, En-yo Y, Yoshida M, Saika A, Suzuki T, Yoshida H, Ishibashi H, Tokimura F, Yamamoto S, Nakamura K, Kawaguchi H, Yoshimura N (2010) Association of radiographic and symptomatic knee osteoarthritis with health-related quality of life in a population-based cohort study in Japan: the ROAD study. Osteoarthr Cartil 18:1227–1234
23. Altman RD, Hochberg M, Murphy WA Jr, Wolfe F, Lequesne M (1995) Atlas of individual radiographic features in osteoarthritis. Osteoarthr Cartil 3(Suppl A):3–70
24. Altman RD, Gold GE (2007) Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthr Cartil 15(Suppl A):A1–A56
25. Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H, Yamamoto S, Nakamura K, Yoshimura N, Kawaguchi H (2008)

Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthr Cartil 16:1300–1306
26. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T (2010) Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol 39:988–995
27. Muraki S, Oka H, Akune T, En-yo Y, Yoshida M, Nakamura K, Kawaguchi H, Yoshimura N (2011) Association of occupational activity with joint space narrowing and osteophytosis in the medial compartment of the knee: the ROAD study. Osteoarthr Cartil 19:840–846
28. Oka H, Muraki S, Akune T, Nakamura K, Kawaguchi H, Yoshimura N (2010) Normal and threshold values of joint space width, joint space area, osteophyte area and fibro-tibial angle using a computer-assisted measuring system (KOACAD) to evaluate knee osteoarthritis: The ROAD study. J Orthop Sci 15:781–789
29. Barr S, Bellamy N, Buchanan WW, Chalmers A, Ford PM, Kean WF, Kraag GR, Gerecz-Simon E, Campbell J (1994) A comparative study of signal versus aggregate methods of outcome measurement based on the WOMAC Osteoarthritis Index. J Rheumatol 21:2106–2112
30. Hashimoto H, Hanyu T, Sledge CB, Lingard EA (2003) Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities osteoarthritis index (WOMAC). J Orthop Sci 8:288–293
31. Muraki S, Oka H, Akune T, En-yo Y, Yoshida M, Suzuki T, Yoshida H, Ishibashi H, Tokimura F, Yamamoto S, Nakamura K, Kawaguchi H, Yoshimura N (2011) Independent association of joint space narrowing and osteophyte formation at the knee with health-related quality of life in Japan: a cross-sectional study. Arthritis Rheum 63: 3859–3864
32. Yamada T, Kawano H, Koshizuka Y, Fukuda T, Yoshimura K, Kamekura S, Saito T, Ikeda T, Kawasaki Y, Azuma Y, Ikegawa S, Hoshi K, Chung UI, Nakamura K, Kato S, Kawaguchi H (2006) Carminerin contributes to chondrocyte calcification during endochondral ossification. Nat Med 12:665–670
33. Kamekura S, Kawasaki Y, Hoshi K, Shimoaka T, Chikuda H, Maruyama Z, Komori T, Sato S, Takeda S, Karsenty G, Nakamura K, Chung UI, Kawaguchi H (2006) Contribution of runt-related transcription factor 2 to the pathogenesis of osteoarthritis in mice after induction of knee joint instability. Arthritis Rheum 54:2462–2470
34. van denBerg WB (1999) Osteophyte formation in osteoarthritis. Osteoarthr Cartil 7:333
35. Kijowski R, Blankenbaker D, Stanton P, Fine J, De Smet A (2006) Correlation between radiographic findings of osteoarthritis and arthroscopic findings of articular cartilage degeneration within the patellofemoral joint. Skelet Radiol 35:895–902
36. Chang CB, Han I, Kim SJ, Seong SC, Kim TK (2007) Association between radiological findings and symptoms at the patellofemoral joint in advanced knee osteoarthritis. J Bone Joint Surg (Br) 89:1324–1328
37. Iannuzzi-Sucich M, Prestwood KM, Kenny AM (2002) Prevalence of sarcopenia and predictors of skeletal muscle mass in healthy, older men and women. J Gerontol A Biol Sci Med Sci 57:M772–M777
38. Sayer AA, Syddall HE, Martin HJ, Dennison EM, Roberts HC, Cooper C (2006) Is grip strength associated with health-related quality of life? Findings from the Hertfordshire Cohort Study. Age Ageing 35:409–415
39. Inoue K, Wicart P, Kawasaki T, Huang J, Ushiyama T, Hukuda S, Courpied J (2000) Prevalence of hip osteoarthritis and acetabular dysplasia in French and Japanese adults. Rheumatology (Oxford) 39:745–748

J Orthop Sci (2015) 20:64–70
DOI 10.1007/s00776-014-0666-4

ORIGINAL ARTICLE

Patient satisfaction with double-door laminoplasty for cervical compression myelopathy

Junichi Ohya · Yasushi Oshima · Katsushi Takeshita · Hiroyuki Oka · Hirotaka Chikuda · Yuki Taniguchi · Yoshitaka Matsubayashi · Sakae Tanaka

Received: 27 June 2014 / Accepted: 12 October 2014 / Published online: 30 October 2014

Abstract
Background Patient satisfaction with posterior laminoplasty for cervical compression myelopathy is not yet established. Moreover, postoperative patient-reported outcomes (PROs) associated with patient satisfaction remain unclear. This study aimed to investigate patient satisfaction after double-door laminoplasty for cervical compression myelopathy, and to identify the postoperative patient-reported outcomes associated with patient satisfaction.
Methods This retrospective study included 97 patients with cervical compression myelopathy who underwent double-door laminoplasty between 2002 and 2010 in our institution [mean follow-up: 58 months (range 12–123 months)]. We assessed postoperative PROs from questionnaires administered before surgery and at the latest follow-up. These questionnaires included the Neck Disability Index, physical and mental component summary of Short Form-36, EuroQol-5 dimension, Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ), and a numerical rating scale of pain or numbness in the neck, arms, and scapular lesion. Satisfaction was evaluated on the basis of a seven-point scale. Patients were divided into two groups: satisfied (very satisfied, satisfied, slightly satisfied) and dissatisfied (neither satisfied nor dissatisfied, slightly dissatisfied, dissatisfied, very dissatisfied). All PROs and the effectiveness of surgical treatment assessed by JOACMEQ were compared between both groups.
Results The satisfied group comprised 69 patients (71 %). Univariate analysis revealed a significant difference in scapular pain, Neck Disability Index, physical component summary of Short Form-36, postoperative mental component summary of Short Form-36, and improvement of lower extremity function postoperatively between both groups. Multivariate analysis revealed that there was a significantly higher proportion of patients with improved lower extremity function in the satisfied group than in the dissatisfied group.
Conclusions In conclusion, 71 % of the patients who underwent double-door laminoplasty for cervical compression myelopathy were satisfied. The findings of this study, which examines the association between patient satisfaction and PROs, suggest that improvement in lower extremity function following surgical intervention affects patient satisfaction in those with cervical compression myelopathy.

Introduction

Posterior laminoplasty has been established as one of the primary interventions in patients with cervical compression myelopathy due to cervical spondylotic changes and ossification of the posterior longitudinal ligament (OPLL) [1, 2]. Although previous literature has reported satisfactory long-term results of laminoplasty, most of these evaluations conducted in the past relied upon the physicians' point of view using the Japanese Orthopaedic Association (JOA) scoring system and image findings such as range of motion [3–6]. Based on the concept that ultimately the evaluation

J. Ohya (✉) · Y. Oshima · K. Takeshita · H. Chikuda · Y. Taniguchi · Y. Matsubayashi · S. Tanaka
Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
e-mail: oyaj-ort@h.u-tokyo.ac.jp

H. Oka
Department of Joint Disease Research, 22nd Century Medical and Research Center, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Table 1 Characteristics of the study population

	$N = 97$
Age (year), mean (SD)	64.3 (10.6)
Gender	
Male	61
Female	36
Follow-up (m), mean (range)	58 (12–123)
Diagnosis (*N*)	
CSM	59
OPLL	36
OYL	1
CDH	1
Surgical levels (*N*)	
C3–7	62
C2–7	18
C3–6	6
Others	11

SD standard deviation, *CSM* cervical spondylotic myelopathy, *OPLL* ossification of the posterior longitudinal ligament, *OYL* ossification of the yellow ligament, *CDH* cervical disc hernia

of treatment results is best done by the patient, the use of several patient-reported outcomes (PROs), such as the visual analog scale, numerical rating scale (NRS), and Short-Form 36 (SF-36), after surgical intervention has become popular in the field of spine surgery [7–9].

A high level of patient satisfaction should be the most important goal following surgery. Although there have been many reports regarding the level of patient satisfaction following other surgeries [10–14], the level after cervical posterior laminoplasty remains to be established [15, 16]. In addition, PROs associated with patient satisfaction after surgical intervention for cervical compression myelopathy remain unclear.

This study aimed to investigate patient satisfaction with double-door laminoplasty for cervical compression myelopathy and to identify the PROs associated with postoperative patient satisfaction.

Materials and methods

Patient population

We reviewed 106 consecutive patients with cervical compression myelopathy who underwent double-door laminoplasty between 2002 and 2010 in our institution. Of these, four patients were lost to follow-up before evaluation of the postoperative outcome. Five patients developed complications, which included C5 motor palsy in four patients and concurrent reoperation for postoperative deterioration due to suboptimal decompression in one patient; these influenced patient satisfaction. The patients with complications were excluded from the study so as to evaluate the association between PROs and patient satisfaction in the population that had an uneventful postoperative course. Finally, 97 patients [mean follow-up 58 months (range 12–123 months)] were included. Of these, 59 had cervical spondylotic myelopathy, 36 had cervical ossification of the posterior longitudinal ligament, one had cervical ossification of the yellow ligament, and one had cervical disc hernia. The most common surgical levels were C3/C7 in 62 patients, followed by C2/C7 in 18 patients and C3/C6 in six patients (Table 1).

Informed consent was obtained from each patient, and the study was approved by the institutional review board of the University of Tokyo.

Double-door laminoplasty

We performed double-door laminoplasty as described in a previous report [6]. Cervical laminae were exposed laterally to the medial aspect of the facet joints, and the interspinous ligaments were removed. The spinous processes were split sagittally. After bilateral gutters for the hinges were carefully made at the transitional area between the facet joint and laminae, spinal canal enlargement was achieved by bilateral opening of the laminae. HA spacers (Boneceram; Olympus Terumo Biomaterials Corp., Tokyo, Japan) were placed between the opening laminae and fixed with nonabsorbable sutures. Patients wore a soft cervical orthosis for approximately 3 weeks.

Outcome measures and questionnaires

We assessed the preoperative outcome of patients from the questionnaires administered before surgery, during their hospital admission. Questionnaires included several PROs, as follows: the Neck Disability Index (NDI) [17], physical component summary (PCS) and mental component summary (MCS) of SF-36 [18], EuroQol-5 dimension, Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ) [19], and NRS of pain or numbness in the neck, arms, and scapular lesion. Postoperatively, questionnaires that included the above-mentioned PROs in addition to the original satisfaction scales that assessed the postoperative outcome were sent to the patients. Satisfaction was evaluated based on a seven-point scale as follows: very satisfied, satisfied, slightly satisfied, neither satisfied nor dissatisfied, slightly dissatisfied, dissatisfied, and very dissatisfied. Patients were divided into two groups: satisfied (very satisfied, satisfied, slightly satisfied) and dissatisfied (neither satisfied nor dissatisfied, slightly dissatisfied, dissatisfied, very dissatisfied). The

postoperative PROs were evaluated using questionnaires administered at the latest follow-up.

Assessment of the effectiveness of surgical intervention

Effectiveness of the treatment was assessed using the JOA scoring system administered by a physician and JOACMEQ.

The recovery rate of the JOA scoring system was calculated using Hirabayashi's method as follows [2]: Recovery rate (%) = (postoperative JOA score − preoperative JOA score) × 100/(17 − preoperative JOA score). Effectiveness of surgical treatment based on the JOA score was defined as a recovery rate ≥50 % [20].

According to a previous report [19], effectiveness of surgical treatment in each domain of the JOACMEQ was defined as follows: (1) the post-treatment score was higher than the pretreatment score by ≥20 points, and (2) the pretreatment score was <90 and the post-treatment score reached ≥90 points. Patients with preoperative and postoperative scores >90 were excluded from this analysis.

Statistical analysis

All PROs and the effectiveness of surgical treatment were compared between both groups. Continuous outcomes were compared using the one-factor analysis of variance, and categorical outcomes were compared using the Chi square test and Fisher's exact test. Multivariate logistic regression models were prepared to estimate patient satisfaction associated with potential predictors including demographic variables and postoperative PROs, which were significantly different between the two groups in univariate analysis. All statistical analyses were conducted using JMP Pro 10 (SAS Institute, Cary, NC). The threshold for significance was a p value <0.05.

Results

Table 2 demonstrates PROs compared between preoperative and postoperative assessments. NDI, PCS of SF-36, EQ-5D, arm pain, and arm numbness were improved between preoperative and postoperative assessment. The preoperative and postoperative differences in neck pain, scapular pain, neck numbness, and scapular numbness were not significant. The satisfied group comprised 69 patients (71 %) and the dissatisfied group comprised the remaining 28 patients (29 %). Table 3 demonstrates the pathology of cervical compression myelopathy (the presence or absence of OPLL), surgical level (including C2 or C7, or neither), and the preoperative PROs of the two groups. None of the preoperative PROs were significantly different between the two groups. There was no significant difference in the pathology of cervical compression myelopathy and surgical level between the two groups.

Table 2 Patient reported outcomes compared between preoperative and postoperative assessments [mean (SD)]

	Preoperative assessment	Postoperative assessment	p value
NDI	36.0 (20.7)	27.3 (16.1)	<0.0001
PCS	20.6 (18.6)	32.1 (18.6)	<0.0001
MCS	49.6 (10.6)	50.9 (9.5)	0.44
EQ-5D	0.55 (0.21)	0.70 (0.19)	<0.0001
NRS (pain)			
Neck	3.4 (3.2)	2.9 (2.6)	0.17
Arms	4.2 (3.2)	2.9 (2.7)	<0.01
Scapular lesion	2.0 (2.6)	2.2 (2.6)	0.19
NRS (numbness)			
Neck	2.2 (2.9)	1.9 (2.5)	0.20
Arms	5.2 (3.0)	4.0 (3.0)	<0.01
Scapular lesion	2.0 (2.4)	1.9 (2.5)	0.96

SD standard deviation, *NDI* Neck Disability Index, *PCS* physical component summary of Short-Form 36, *MCS* mental component summary of Short-Form 36, *EQ-5D* EuroQol 5 dimension, *NRS* numerical rating scale

Univariate analysis revealed that the postoperative scapular pain, NDI, PCS of SF-36, MCS of SF-36, and effectiveness of treatment in the lower extremity evaluated using the JOACMEQ were significantly different between the two groups (Table 4). Patients in the satisfied group showed a tendency toward a higher recovery rate of the JOA score, as evaluated by a physician, compared to the dissatisfied group; however, the difference was not significant. Table 5 demonstrates the results of multivariable logistic regression models for satisfaction with double-door laminoplasty. After adjusting for confounders, the effectiveness of treatment in the lower extremity, evaluated using the JOACMEQ, was significantly higher in the satisfied group than in the dissatisfied group (Odds ratio = 3.77; 95 % CI, 1.13–15.3; $p = 0.03$).

Discussion

This study had two main findings. First, 71 % of the patients who underwent double-door laminoplasty for cervical compression myelopathy were satisfied. Additionally, several PROs, including NDI, SF-36, NRS of pain in the scapular lesion, and effectiveness of treatment in the lower extremity, which was evaluated using the JOACMEQ, were associated with patient satisfaction. In particular,

Table 3 Pathology of cervical compression myelopathy, surgical level, and preoperative patient-reported outcomes compared between the two groups [mean (SD)]

	Satisfied group $N = 69$	Dissatisfied group $N = 28$	p value[a]
Pathology of myelopathy OPLL presence [*N* (%)]	27 (39)	9 (32)	0.52
Surgical level including C2 or C7 [*N* (%)]	63 (91)	28 (100)	0.18[b]
NDI	37.5 (20.8)	31.9 (20.1)	0.23
PCS	20.4 (19.1)	21.0 (17.7)	0.89
MCS	49.9 (10.9)	48.9 (9.9)	0.67
EQ-5D	0.54 (0.21)	0.58 (0.21)	0.45
NRS (pain)			
Neck	3.4 (3.3)	3.6 (3.1)	0.77
Arms	4.3 (3.1)	3.9 (3.6)	0.60
Scapular lesion	1.8 (2.5)	2.5 (2.7)	0.25
NRS (numbness)			
Neck	2.1 (2.9)	2.7 (2.9)	0.43
Arms	5.6 (2.9)	4.4 (3.0)	0.16
Scapular lesion	1.7 (2.3)	2.6 (2.6)	0.15

SD standard deviation, *OPLL* ossification of the posterior longitudinal ligament, *NDI* Neck Disability Index, *PCS* physical component summary of Short-Form 36, *MCS* mental component summary of Short-Form 36, *EQ-5D* EuroQol 5 dimension, *NRS* numerical rating scale

[a] For continuous outcomes, the comparisons were made by the one-factor analysis of variance. For categorical outcomes, the comparisons were made by the Chi square test

[b] Fisher's exact test

multivariate logistic regression analysis revealed that effectiveness of treatment in the lower extremity evaluated using the JOACMEQ was a significant independent factor associated with patient satisfaction.

Although several studies on postoperative satisfaction in patients with cervical compression myelopathy have been previously conducted, evidence of patient satisfaction with the posterior operative approach is limited. With regard to cervical myelopathy due to multilevel compression, Sampath et al. [21] reported a 75 % satisfaction rate after surgery in their prospective multicenter study. However, their study was limited by a small sample size of 20 patients who underwent surgical intervention, and also included several procedures, including decompression through a posterior approach, anterior cervical discectomy, and spinal fusion with or without internal fixation with instrumentation. Although Fujimori et al. [15] found that 80 % of patients with cervical myelopathy due to multilevel compression by OPLL were satisfied with the surgical results, this study also included both posterior and anterior procedures. The present study was superior compared to previous reports regarding patient satisfaction following posterior cervical surgery in that it used data from approximately 100 patients following double-door laminoplasty as a single procedure. Moreover, the satisfaction scale in this study was classified into seven categories, whereas previous classifications in other studies had only five categories. The detailed questionnaire in this study may reflect patient satisfaction with more accuracy.

This study found that several PROs, such as NDI, SF-36, scapular pain, and effectiveness of treatment in the lower extremity assessed using the JOACMEQ, were associated with patient satisfaction. In particular, the effect of surgical intervention on the lower extremity, which was evaluated using the JOACMEQ, was reflected in patient satisfaction with double-door laminoplasty independently. Fujimori et al. [15], in their study of 69 patients with cervical OPLL, found that lower extremity function correlated with patient satisfaction, which was similar to the findings of this study. With regard to the reason why lower extremity function was identified as an independent factor of satisfaction, we speculate the following possibilities. The patients who did not feel improvement in lower extremity function might be dissatisfied because lower extremity function reportedly correlates more directly with quality of life than upper extremity function [22]. Fujimori et al. [15] found that many patients who were dissatisfied reported inability to move around independently as the reason in their response to open-type questions. This study confirmed the postulate that lower extremity function correlated more strongly with patient satisfaction. Moreover, a previous study demonstrated that following surgery, neurological recovery in the lower extremity was less likely to achieve neurological improvement compared to the upper extremity [23, 24]; this may play a role in the dissatisfaction experienced by the patients. Furthermore, patients in the current study might have suffered from degenerative diseases that affect

Table 4 Postoperative outcomes compared between the two groups [mean (SD)]

	Satisfied group	Dissatisfied group	*p* value[a]
NDI	24.4 (15.8)	34.3 (14.8)	<0.01
PCS	34.8 (18.8)	25.6 (16.8)	0.03
MCS	52.2 (9.6)	47.7 (8.7)	0.04
EQ-5D	0.72 (0.21)	0.64 (0.14)	0.05
NRS (pain)			
Neck	2.6 (2.6)	3.6 (2.5)	0.12
Arms	2.6 (2.7)	3.5 (2.5)	0.14
Scapular lesion	1.8 (2.5)	3.2 (2.6)	0.01
NRS (numbness)			
Neck	1.7 (2.6)	2.3 (2.3)	0.29
Arms	3.9 (3.0)	4.4 (3.0)	0.46
Scapular lesion	1.6 (2.5)	2.5 (2.6)	0.12
JOA score recovery rate	46.8 % (41.5)	32.1 % (31.4)	0.10
Effectiveness of surgical treatment evaluated using JOACMEQ [*N* (%)]			
Cervical spine function	20 (34)	7 (29)	0.80
Upper extremity function	27 (44)	7 (28)	0.23
Lower extremity function	30 (49)	4 (16)	<0.01
Bladder function	15 (26)	2 (8)	0.08
Quality of life	14 (23)	4 (16)	0.57

SD standard deviation, *NDI* Neck Disability Index, *PCS* physical component summary of Short-Form 36, *MCS* mental component summary of Short-Form 36, *EQ-5D* EuroQol 5 dimension, *NRS* numerical rating scale, *JOA* Japanese Orthopaedic Association, *JOACMEQ* Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire

[a] For continuous outcomes, comparisons were made by the one-factor analysis of variance. For categorical outcomes, the comparisons were made by the Fisher's exact test

Table 5 Multivariable logistic regression models for patient satisfaction after double-door laminoplasty

	Patient satisfaction		
	OR	95 % CI	*p* value
Age	0.99	0.93–1.05	0.64
Female (ref. male)	0.99	0.30–3.47	0.99
Postoperative NDI	1.01	0.96–1.07	0.69
Postoperative PCS	0.99	0.94–1.03	0.55
Postoperative MCS	0.94	0.87–1.02	0.13
Postoperative scapular pain	1.05	0.83–1.34	0.65
Effective in LE (ref. non-effective)	3.77	1.13–15.3	0.03

OR odds ratio, *CI* confidence interval, *NDI* Neck Disability Index, *PCS* physical component summary of Short-Form 36, *MCS* mental component summary of Short-Form 36

lower extremity function (e.g., knee osteoarthritis or lumbar spinal stenosis). Prior study revealed a high prevalence of knee osteoarthritis and lumbar spinal stenosis in the Japanese elderly [25, 26], which may be one of the reasons for insufficient improvement in lower extremity function compared to that in upper extremity function.

Although several factors have been reported to affect the outcome following laminoplasty, so far, the predictor for optimal timing of laminoplasty remains unknown. Age, preoperative JOA, signal intensity change on MRI, and cervical lordotic angle were reportedly associated with postoperative outcome in a previous study [27–30]. However, the prognostic significance of these factors has not been established. Timing of surgical intervention should be decided according to the expected postoperative satisfaction in addition to neurological improvement. This study found that the improvement of lower extremity function following surgical intervention was as an independent factor associated with patient satisfaction, which suggests that the factors reflecting severity of myelopathy in the lower extremity is important. Two clinical tests were reported for evaluation of lower extremity function in patients with cervical myelopathy. Mihara et al. [31] demonstrated that the triangle step test was a very useful method for evaluation of lower extremity function in patients with cervical myelopathy. Nakashima et al. [20] reported the 10-s step test as a simple physical assessment for severity of cervical compression myelopathy, particularly for lower extremity dysfunction. These tests can be candidate predictors for patient satisfaction following laminoplasty. Further research regarding predictors reflecting satisfaction to decide optimal timing of surgical intervention is expected in the future.

According to previous studies, satisfaction in patients with cervical myelopathy following anterior approach procedures ranges from 80.6 to 94.7 % [11, 13]. In the current study, postoperative patient satisfaction following posterior approach procedures was 71 %, and including the five patients with complications (one patient in the satisfied group and four patients in the dissatisfied group) in the analysis, it was 69 %. Despite recognition of posterior laminoplasty as an established treatment for cervical compression myelopathy, our data revealed that patient satisfaction following double-door laminoplasty was relatively low compared to that following the anterior approach. One of the reasons for this might be that patients undergoing laminoplasty often complain of axial pain, which may play a role in decreasing the level of patient satisfaction. Indeed, the satisfied group had a significantly lower postoperative numerical scapular pain scale in this study. Several factors, including different surgical techniques, radiological assessment, and postoperative management, were reportedly

associated with axial pain following posterior cervical surgery [32]. Further surgical modification, such as less invasive surgery and postoperative management including medication and early removal of cervical orthosis, should be attempted for improvement of postoperative scapular pain.

This study had several limitations. First, the surgery was not performed by the same surgeon in all patients, nor was it performed to the same surgical level. Surgical invasion may vary slightly between surgeons and surgical level. The inconsistency of surgical techniques performed can affect postoperative outcome. Indeed, such surgical factors were reported to be associated with axial pain after posterior cervical spine surgery [32, 33]. However, the preservation of muscles attached at C2 or C7 was not associated with patient satisfaction in this study. Second, the questionnaire used in this study did not assess the outcomes immediately after the surgical procedure. In particular, perioperative pain immediately after surgical procedure may cause decreased satisfaction. Further study with perioperative evaluation that includes use of a numerical pain scale and medication and cervical immobilization is required to verify the influence of perioperative management on patient satisfaction. Third, although we examined the patients with cervical compression myelopathy, including those with and without OPLL, the difference in the pathomechanism of myelopathy between those with OPLL and others may have affected the surgical outcome. Although there was no significant difference in the pathology of cervical compression myelopathy between the satisfied and dissatisfied groups, further large-scale studies that take these differences in the pathology into consideration are warranted. Finally, the relationship between the physician and the patient was not evaluated in this study. There may be bias related to this factor. We used a questionnaire for evaluation of satisfaction instead of directing the questions to the physician to decrease this bias. Despite these limitations, we believe that this study has valuable information that is of clinical importance.

In conclusion, 71 % of the patients who underwent double-door laminoplasty for cervical compressive myelopathy were satisfied. The findings in this study suggest that improvement in lower extremity function following surgical intervention affects patient satisfaction.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Kurokawa T. Double-door laminoplasty by longitudinal splitting of the spinal processes. Bessatsu Seikeigeka. 1982;2:234–40 (in Japanese).
2. Hirabayashi K, Miyakawa J, Satomi K, Maruyama T, Wakano K. Operative results and postoperative progression of ossification among patients with ossification of cervical posterior longitudinal ligament. Spine (Phila Pa 1976). 1981;6(4):354–64.
3. Chiba K, Ogawa Y, Ishii K, Takaishi H, Nakamura M, Maruiwa H, Matsumoto M, Toyama Y. Long-term results of expansive open-door laminoplasty for cervical myelopathy average 14-year follow–up study. Spine (Phila Pa 1976). 2006;31(26):2998–3005.
4. Iwasaki M, Kawaguchi Y, Kimura T, Yonenobu K. Long-term results of expansive laminoplasty for ossification of the posterior longitudinal ligament of the cervical spine: more than 10 years follow up. J Neurosurg. 2002;96(2 Suppl):180–9.
5. Kimura A, Seichi A, Inoue H, Hoshino Y. Long-term results of double-door laminoplasty using hydroxyapatite spacers in patients with compressive cervical myelopathy. Eur Spine J. 2011;20(9):1560–6.
6. Seichi A, Takeshita K, Ohishi I, Kawaguchi H, Akune T, Anamizu Y, Kitagawa T, Nakamura K. Long-term results of double-door laminoplasty for cervical stenotic myelopathy. Spine (Phila Pa 1976). 2001;26(5):479–87.
7. Copay AG, Martin MM, Subach BR, Carreon LY, Glassman SD, Schuler TC, Berven S. Assessment of spine surgery outcomes: inconsistency of change amongst outcome measurements. Spine J. 2010;10(4):291–6.
8. DeVine J, Norvell DC, Ecker E, Fourney DR, Vaccaro A, Wang J, Andersson G. Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery. Spine (Phila Pa 1976). 2011;36(21 Suppl):S69–74.
9. McCormick JD, Werner BC, Shimer AL. Patient-reported outcome measures in spine surgery. J Am Acad Orthop Surg. 2013;21(2):99–107.
10. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. National joint registry for England and Wales. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National joint registry for England and Wales. J Bone Joint Surg Br. 2007;899(7):893–900.
11. Chagas H, Domingues F, Aversa A, Vidal Fonseca AL, de Souza JM. Cervical spondylotic myelopathy: 10 years of prospective outcome analysis of anterior decompression and fusion. Surg Neurol. 2005;64 Suppl 1(S1):30–5.
12. Gepstein R, Arinzon Z, Adunsky A, Folman Y. Decompression surgery for lumbar spinal stenosis in the elderly: preoperative expectations and postoperative satisfaction. Spinal Cord. 2006;44(7):427–31.
13. Riew KD, Buchowski JM, Sasso R, Zdeblick T, Metcalf NH, Anderson PA. Cervical disc arthroplasty compared with arthrodesis for the treatment of myelopathy. J Bone Joint Surg Am. 2008;90(11):2354–64.
14. Toyone T, Tanaka T, Kato D, Kaneyama R, Otsuka M. Patients' expectations and satisfaction in lumbar spine surgery. Spine (Phila Pa 1976). 2005;30(23):2689–94.
15. Fujimori T, Iwasaki M, Okuda S, Nagamoto Y, Sakaura H, Oda T, Yoshikawa H. Patient satisfaction with surgery for cervical myelopathy due to ossification of the posterior longitudinal ligament. J Neurosurg Spine. 2011;14(6):726–33.
16. Neo M, Fujibayashi S, Takemoto M, Nakamura T. Clinical results of and patient satisfaction with cervical laminoplasty for considerable cord compression with only slight myelopathy. Eur Spine J. 2012;21(2):340–6.
17. Vernon H, Mior S. The Neck disability index: a study of reliability and validity. J Manipulative Physiol Ther. 1991;14(7):409–15.
18. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30(6):473–83.

19. Fukui M, Chiba K, Kawakami M, Kikuchi S, Konno S, Miyamoto M, Seichi A, Shimamura T, Shirado O, Taguchi T, Takahashi K, Takeshita K, Tani T, Toyama Y, Yonenobu K, Wada E, Tanaka T, Hirota Y. Subcommittee of the clinical outcome committee of the Japanese Orthopaedic Association on low back pain and cervical myelopathy evaluation. JOA Back Pain Evaluation Questionnaire (JOABPEQ)/JOA Cervical Myelopathy Evaluation Questionnaire (JOACMEQ). The report on the development of revised versions. April 16, 2007. The subcommittee of the clinical outcome committee of the Japanese Orthopaedic Association on low back pain and cervical myelopathy evaluation. J Orthop Sci. 2009;14(3):348–65.
20. Nakashima H, Yukawa Y, Ito K, Machino M, Kanbara S, Morita D, Takahashi H, Imagama S, Ito Z, Ishiguro N, Kato F. Prediction of lower limb functional recovery after laminoplasty for cervical myelopathy: focusing on the 10-s step test. Eur Spine J. 2012;21:1389–95.
21. Sampath P, Bendebba M, Davis JD, Ducker TB. Outcome of patients treated for cervical myelopathy. A prospective, multicenter study with independent clinical review. Spine (Phila Pa 1976). 2000;25(6):670–6.
22. King JT Jr, McGinnis KA, Roberts MS. Quality of life assessment with the medical outcomes study short form-36 among patients with cervical spondylotic myelopathy. Neurosurgery. 2003;52(1):113–20.
23. Cheung WY, Arvinte D, Wong YW, Luk KD, Cheung KM. Neurological recovery after surgical decompression in patients with cervical spondylotic myelopathy—a prospective study. Int Orthop. 2008;32(2):273–8.
24. Chiles BW 3rd, Leonard MA, Choudhri HF, Cooper PR. Cervical spondylotic myelopathy: patterns of neurological deficit and recovery after anterior cervical decompression. Neurosurgery. 1999;44(4):762–9.
25. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, Saika A, Suzuki T, Yoshida H, Ishibashi H, Yamamoto S, Nakamura K, Kawaguchi H, Yoshimura N. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthritis Cartilage. 2009;17(9):1137–43.
26. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, Takiguchi N, Minamide A, Oka H, Kawaguchi H, Nakamura K, Akune T, Yoshida M. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama spine study. Osteoarthritis Cartilage. 2012;20(10):1103–8.
27. Kohno K, Kumon Y, Oka Y, Matsui S, Ohue S, Sakaki S. Evaluation of prognostic factors following expansive laminoplasty for cervical spinal stenotic myelopathy. Surg Neurol. 1997;48(3):237–45.
28. Tanaka J, Seki N, Tokimura F, Doi K, Inoue S. Operative results of canal-expansive laminoplasty for cervical spondylotic myelopathy in elderly patients. Spine (Phila Pa 1976). 1999;24(22):2308–12.
29. Suda K, Abumi K, Ito M, Shono Y, Kaneda K, Fujiya M. Local kyphosis reduces surgical outcomes of expansive open door laminoplasty for cervical spondylotic myelopathy. Spine (Phila Pa 1976). 2003;28(12):1258–62.
30. Yukawa Y, Kato F, Yoshihara H, Yanase M, Ito K. MR T2 image classification in cervical compression myelopathy: predictor of surgical outcomes. Spine (Phila Pa 1976). 2007;32(15):1675–8.
31. Mihara H, Kondo S, Murata A, Ishida K, Niimura T, Hachiya M. A new performance test for cervical myelopathy: the triangle step test. Spine (Phila Pa 1976). 2010;35(1):32–5.
32. Wang SJ, Jiang SD, Jiang LS, Dai LY. Axial pain after posterior cervical spine surgery: a systematic review. Eur Spine J. 2011;20(2):185–94.
33. Kato M, Nakamura H, Konishi S, Dohzono S, Toyoda H, ima W, Kondo K, Matsuda H. Effect of preserving paraspinal muscles on postoperative axial pain in the selective cervical laminoplasty. Spine (Phila Pa 1976). 2008;33(14):E455–9.

J Orthop Sci (2015) 20:71–77
DOI 10.1007/s00776-014-0669-1

ORIGINAL ARTICLE

Prospective multicenter surveillance and risk factor analysis of deep surgical site infection after posterior thoracic and/or lumbar spinal surgery in adults

Satoshi Ogihara · Takashi Yamazaki · Toru Maruyama · Hiroyuki Oka · Kota Miyoshi · Seiichi Azuma · Takashi Yamada · Motoaki Murakami · Naohiro Kawamura · Nobuhiro Hara · Sei Terayama · Jiro Morii · So Kato · Sakae Tanaka

Received: 24 July 2014 / Accepted: 16 October 2014 / Published online: 1 November 2014

Abstract
Background Surgical site infection is a serious and significant complication after spinal surgery and is associated with high morbidity rates, high healthcare costs and poor patient outcomes. Accurate identification of risk factors is essential for developing strategies to prevent devastating infections. The purpose of this study was to identify independent risk factors for surgical site infection among posterior thoracic and/or lumbar spinal surgery in adult patients using a prospective multicenter surveillance research method.
Methods From July 2010 to June 2012, we performed a prospective surveillance study in adult patients who had developed surgical site infection after undergoing thoracic and/or lumbar posterior spinal surgery at 11 participating hospitals. Detailed preoperative and operative patient characteristics were prospectively recorded using a standardized data collection format. Surgical site infection was based on the definition established by the Centers for Disease Control and Prevention.
Results A total of 2,736 consecutive adult patients were enrolled, of which 24 (0.9%) developed postoperative deep surgical site infection. Multivariate regression analysis indicated four independent risk factors. Preoperative steroid therapy ($P = 0.001$), spinal trauma ($P = 0.048$) and gender (male) ($P = 0.02$) were statistically significant independent

S. Ogihara (✉)
Department of Orthopaedic Surgery, Sagamihara National Hospital, 18-1 Sakuradai, Minami-ku, Sagamihara City, Kanagawa 252-0392, Japan
e-mail: s-ogihara@sagamihara-hosp.gr.jp

T. Yamazaki · S. Terayama
Department of Orthopaedic Surgery, Musashino Red Cross Hospital, Tokyo, Japan

T. Maruyama
Department of Orthopaedic Surgery, Saitama Medical Center, Saitama Medical University, Saitama, Japan

H. Oka
Department of Joint Disease Research, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

K. Miyoshi
Department of Orthopaedic Surgery, Yokohama Rosai Hospital, Kanagawa, Japan

S. Azuma
Department of Orthopaedic Surgery, Saitama Red Cross Hospital, Saitama, Japan

T. Yamada
Department of Orthopaedic Surgery, NTT Kanto Hospital, Tokyo, Japan

M. Murakami
Department of Orthopaedic Surgery, Toranomon Hospital, Tokyo, Japan

N. Kawamura
Department of Spine and Orthopaedic Surgery, Japanese Red Cross Medical Center, Tokyo, Japan

N. Hara · S. Tanaka
Department of Orthopaedic Surgery, Faculty of Medicine, University of Tokyo, Tokyo, Japan

J. Morii
Department of Orthopaedic Surgery, Sanraku Hospital, Tokyo, Japan

S. Kato
Department of Orthopaedic Surgery and Musculoskeletal Oncology, Tokyo Metropolitan Komagome Hospital, Tokyo, Japan

patient-related risk factors, whereas an operating time ≥3 h ($P < 0.001$) was a surgery-related independent risk factor.
Conclusion Preoperative steroid therapy, spinal trauma, male gender and an operating time ≥3 h were independent risk factors for deep surgical site infection after thoracic and/or lumbar spinal surgery in adult patients. Identification of these risk factors can be used to develop protocols aimed at decreasing the risk of surgical site infection.

Introduction

Surgical site infection (SSI) after spinal surgery is one of the most serious complications that occurs in 0.7–12% of patients and can lead to high morbidity, mortality and increased healthcare costs [1, 2]. In this regard, various risk factors for SSI have been investigated to prevent this devastating complication. Risk factors were separated into two main categories: patient-related risk factors and surgery-related risk factors. Patient-related risk factors include advanced age [3], male gender [4], obesity [5, 6], previous spinal surgery [5], diabetes [5–7], malnutrition [3], smoking [5], spinal trauma [8, 9] and corticosteroid use [5, 10]. Surgery-related risk factors include spinal instrumentation [11], posterior surgical approach [2], tumor resection [2], fusion extending to the sacrum [12], increased estimated blood loss [4, 7] and prolonged operating time [4, 11, 13]. However, many of these studies were performed retrospectively at individual institutions, and they are limited by their relatively small sample size that restricts the power to perform a multivariate analysis.

High quality studies based on a prospective design and a large sample size are required to identify precise independent risk factors for SSI following spinal surgery. Multivariate analysis should also be performed to adjust for the occurrence of multiple risk factors within individual patients. In addition, standardized, hospital-based, multicenter surveillance methods utilizing a standard definition of SSI have been recommended to help determine risk factors and are considered useful in reducing infection rates [14–16].

Therefore, the purpose of this study was to identify independent risk factors for adult patients who develop deep SSI after posterior thoracic and/or lumbar spinal surgery using a prospective multicenter surveillance research method.

Materials and methods

Study design and selection criteria

This surveillance study for SSI following posterior thoracic and/or lumbar spinal surgery in adult patients was conducted in a prospective manner from July 1, 2010 to June 30, 2012 at 11 participating Japanese hospitals. Patients included in the study had undergone surgery by orthopedic service only. Each patient had undergone follow-up for a minimum of one year. Detailed preoperative patient characteristics and operative characteristics were recorded prospectively using a standardized data collection format. The institutional review board at participating hospitals approved the present study and informed consent was obtained from each patient. Patients who underwent surgery for the treatment of spinal infection were excluded from the present analysis. For homogeneity of the study group, we also excluded patients aged <20 years, those who underwent posterior instrumentation removal, vertebroplasty (percutaneous or open surgery), endoscopic surgery or single-stage anterior-posterior surgery.

Identification of SSI

A patient was considered to have an infection on the basis of the SSI definition put forth by the Centers for Disease Control and Prevention [17]. Superficial SSI was defined as an infection occurring within 30 days after the operation and involving the skin or subcutaneous tissue only. Deep SSI was defined as an infection occurring within 30 days after the operation if no implant was left in place or within one year if the implant was left in place, if the infection appeared to be operation-related and involved deep soft tissues. A deep SSI was further characterized by the presence of one or more of the following [17]: (1) purulent drainage from the deep incision; (2) a deep incision spontaneously dehisces or is deliberately opened by a surgeon when the patient has at least one of the following signs or symptoms: fever (>38 °C), localized pain, or tenderness, unless the site is culture-negative; (3) an abscess or other evidence of infection involving the deep incision that is found on direct examination, during reoperation or by a histopathologic or radiologic examination; and (4) diagnosis of a deep incisional SSI by a surgeon or attending physician.

The incidence of SSI was confirmed after double-checking by the attending surgeons and colleagues involved in this study at the participating hospitals. Microbiologic culture results of each patient with deep SSI were recorded and assembled. In cases in which open debridement was performed, microbiologic cultures were taken to confirm the presence of SSI and to determine further treatment.

Data collection

At each study hospital, the medical records of eligible adult patients who had undergone posterior thoracic and/or lumbar spinal procedures were prospectively collected utilizing standardized patient charts.

The recorded preoperative patient characteristics included age at time of surgery, sex, height, weight and diagnosis (spinal trauma, disc herniation, spinal stenosis, tumor or cancer, inflammatory arthritis, osteoporosis or spinal deformity). Preoperative patient-related risk factors for SSI included smoking, diabetes mellitus, body mass index (BMI), the patient's American Society of Anesthesiologists (ASA) score [18], previous surgery and steroid use. In addition, surgery-related factors considered as possible risk factors for SSI were collected and analyzed. These included operating time, estimated blood loss, anatomic location (thoracic, lumbar and/or sacral), emergency surgery, use of instrumentation, iliac crest bone grafting, dural tear and use of intraoperative fluoroscopy.

Statistical analysis

Associations between deep SSI and potential risk factors were analyzed. Fisher's exact test was used for categorical variables and the Wilcoxon test was used for continuous variables. Multivariate analysis was performed to evaluate the risk factors for SSI. Significant variables and the variables that correlated ($P < 0.20$) with SSI in univariate analysis were entered into a stepwise multiple logistic regression model. Furthermore, to adjust for confounding factors, the BMI and anatomic location of the surgery were entered into this model, as in previous reports BMI [5, 6] and anatomic location of the surgery [12] were identified as risk factors of SSI after spinal surgeries. All analyses were performed using SPSS Statistics version 19 (IBM Corporation, Armonk, NY) with the significance threshold set at $P < 0.05$.

Results

From July 2010 to June 2012, a total of 2,736 consecutive patients (1,164 female, 1,572 male; mean age, 64.6 years; age range, 20–94 years) in 11 Japanese hospitals were enrolled. Overall, 24 patients (0.9%) developed postoperative deep SSI. The demographic characteristics of the patients included in the study are shown in Table 1.

The statistical relationships of all variables to deep SSI are reported in Table 2. Univariate analysis indicated several significant risk factors, including an ASA score $\geq$3, preoperative steroid use, spinal trauma, spinal instrumentation, use of intraoperative fluoroscopy and an operating time $\geq$3 h. The significant factors in the univariate analysis and the factors with a P value <0.20 in the univariate analysis (male sex, diabetes mellitus, previous surgery and emergency surgery) were included in a multivariate analysis to further examine the risk factors for deep SSI. Although the BMI and anatomic location of the surgery were not significantly associated with deep SSI in the univariate analysis, they were included in the multivariate models. The final multivariate model shown in Table 2 is the most parsimonious model, showing the independent risk factors for deep SSI after adjusting for other risk factors. According to logistic regression models, men had a higher risk of deep SSI than women [odds ratio (OR), 3.01; 95% confidence interval (CI), 1.15–8.94; $P = 0.02$]. The preoperative diagnosis was also found to be associated with an increased risk of SSI; patients with spinal trauma had a 4.04 times higher risk than patients with other diagnoses (95% CI, 1.01–14.49; $P = 0.048$). Patients with preoperative steroid therapy (oral intake) had an 8.53 times higher risk than patients without steroid therapy (95% CI, 2.49–25.82; $P = 0.001$). An operating time $\geq$3 h was also found to be significantly associated with an increased risk of SSI (OR, 10.28; 95% CI, 3.31–39.36; $P < 0.0001$).

Microbiologic cultures were routinely taken in all 24 patients who developed deep SSI and 87.5% (21/24) of the patients had a positive culture. Twenty-one of the 24 patients (87.5%) underwent open debridement. In three patients, no organisms were isolated; antibiotics were administered intravenously prior to open debridement in these three patients, after which open debridement was performed. Abscess formation in the deep soft tissues was observed in the thee patients. Twenty of 21 patients with positive cultures (95.2%) had a single organism isolated, while only one case demonstrated polymicrobacterial growth [methicillin-resistant *Staphylococcus aureus* (MRSA) + *Propionibacterium acnes*]. *Staphylococcus aureus* was present in 57.1% (12/21) of the positive cultures (including the case of polymicrobacterial growth), with 66.7% (8/12) of these isolates demonstrating MRSA. Coagulase-negative *Staphylococcus* was the next most common organism, with occurrence in 33.3% (7/21) of the positive cultures; methicillin resistance was noted in 71.4% (5/7) of the patients (Table 3).

Discussion

In this study, we identified independent risk factors for adult patients who develop a deep SSI after posterior thoracic and/or lumbar spinal surgery using a prospective multicenter surveillance research method. An operating time $\geq$3 h was the strongest independent risk factor for postoperative deep SSI after adjusting for all other variables. This result is consistent with those of previous studies that described a prolonged surgical procedure as a significant risk factor for SSI [4, 13]. Frequent release of the tension on self-retractors [13] can minimize tissue ischemia and necrosis caused by intraoperative wound retraction during long-duration operations. A longer operating time also

Table 1 Demographic characteristics of the patients included in the study

Characteristics	Deep SSI ($n = 24$)	Non-deep SSI ($n = 2,712$)	P value[a]
Age at surgery, mean (SD), year	67.5 (13.2)	64.0 (15.0)	0.34
Male, *n* (%)	18 (75.0)	1,554 (57.3)	0.08
BMI, mean (SD), kg/m^2	23.9 (3.3)	23.8 (3.6)	0.88
ASA score, *n* (%)			
1 and 2	18 (75.0)	2,462 (90.8)	0.008
≥3	6 (25.0)	250 (9.2)	
Diabetes mellitus, *n* (%)	5 (20.8)	308 (11.4)	0.14
Smoking, *n* (%)	4 (16.7)	332 (12.2)	0.51
Preoperative steroid therapy (oral intake), *n* (%)	5 (20.8)	89 (3.3)	<0.0001
Previous surgery, *n* (%)	7 (29.2)	420 (15.5)	0.07
Diagnosis, *n* (%)			
Spinal trauma	5 (20.8)	117 (4.3)	<0.0001
Other	19 (79.2)	2,595 (95.7)	
Anatomic location of the surgery, *n* (%)			
Sacrum included	2 (8.3)	212 (7.8)	0.72
Other	22 (91.7)	2,500 (92.2)	
Surgical variables, *n* (%)			
Instrumentation	18 (75.0)	1,388 (51.2)	0.02
Emergency surgery	3 (12.5)	107 (4.0)	0.03
Use of intraoperative fluoroscopy	4 (16.7)	135 (5.0)	0.009
Dural tear	3 (12.5)	274 (10.1)	0.70
Iliac crest bone graft	3 (12.5)	215 (8.0)	0.41
Operating time ≥3 h	20 (83.3)	952 (35.1)	<0.0001

ASA American Society of Anesthesiologists, *BMI* body mass index, *SSI* surgical site infection

[a] Fisher's exact test was used for categorical variables and the Wilcoxon test was used for continuous variables

increases the risk for bacterial contamination in the surgical wounds [19]; frequent saline irrigation of the surgical wound during the procedure can help prevent this complication [9].

Preoperative steroid use as a risk factor for SSI has been described in several previous studies on spinal surgery patients [5, 10]. On the other hand, there are other studies reporting that steroid use was not a risk factor for SSI [2, 6]. In our study, multivariate analysis showed a strong association between preoperative steroid therapy (oral intake) and postoperative deep SSI. Many of these previous studies were conducted at individual hospitals and, to our knowledge, this is the first study evaluating the association between preoperative steroid therapy and SSI following spinal surgery using a prospective multicenter design. There appears to be a paucity of literature on the relationship between steroid dosage and SSI following spinal surgeries and steroid dosage was not included in our study. For a more detailed evaluation of steroid use and SSI risk, additional high-quality research is needed in the future.

Several studies have reported on patients with spinal trauma and the incidence of SSI [8, 9, 20]. Watanabe et al. reported a strong association between trauma and SSI using multivariate analysis, as compared to patients who underwent elective surgery (OR, 9.42; 95% CI, 1.59–55.73) [9]. The results of our analysis are consistent with previous reports. Patients with a traumatized spine tend to include multisystem trauma, concomitant open wounds, head injury and/or cardiopulmonary instability. Since multisystem trauma can cause severe general conditions, the preoperative hospital stay for patients with spinal trauma tends to be long. Blam et al. [8] reported that surgical treatment of the spine >160 h after injury increased the incidence of infection by more than 8 times, compared with cases in which treatment began within 48 h after injury. Blam et al. [8] also reported that the duration of the postoperative intensive care unit stay was an independent significant risk factor for SSI. In patients with a traumatized spine, it is important to perform the surgery immediately after general conditions are stabilized and decrease the perioperative stay in the intensive care unit.

Male gender was also found to be significantly associated with an increased risk of deep SSI in our current multivariate analysis, even though it was not significant in the univariate analysis. To our knowledge, only one previous study found a statistically significant association between being male and SSI after spinal surgery [4], but this relationship has been reported in several studies on total knee arthroplasty [21, 22] and gastric surgery [23]. In order to evaluate the association of the male gender with SSI more precisely, additional prospective studies with large sample sizes are needed.

Table 2 Univariate and multivariate logistic regression analyses for the odds ratios (ORs) and 95 % confidence interval (CI) of risk factors for deep SSI

Demographic characteristics	Univariate		Multivariate	
	OR (95 % CI)	*P* value	OR (95 % CI)	*P* value
Sex				
Female	1.00	0.07	1.00	0.02
Male	2.24 (0.93–6.18)		3.01 (1.15–8.94)	
BMI				
+1 kg/m^2	0.99 (0.90–1.11)	0.88	0.98 (0.88–1.10)	0.69
ASA score				
1 and 2	1.00	0.02	1.00	0.42
≥ 3	3.28 (1.81–7.91)		1.55 (0.51–4.13)	
Diabetes mellitus				
Yes	2.05 (0.68–5.15)	0.19	1.16 (0.36–3.14)	0.78
No	1.00		1.00	
Smoking				
Yes	1.43 (0.41–3.82)	0.53		
No	1.00			
Preoperative steroid therapy (oral intake)				
Yes	7.76 (2.53–19.80)	0.001	8.53 (2.49–25.82)	0.001
No	1.00		1.00	
Previous surgery				
Yes	2.25 (0.86–5.24)	0.16	2.03 (0.73–5.15)	0.16
No	1.00		1.00	
Diagnosis				
Spinal trauma	5.84 (1.91–14.81)	0.037	4.04 (1.01–14.49)	0.048
Other	1.00		1.00	
Anatomic location of the surgery				
Sacrum included	1.07 (0.17–3.67)	0.49	1.00	0.49
Other	1.00		1.42 (0.24–1.87)	
Instrumentation				
Yes	2.86 (1.20–7.91)	0.017	1.70 (0.50–5.31)	0.38
No	1.00		1.00	
Emergency surgery				
Yes	3.48 (0.81–10.29)	0.085	2.93 (0.51–12.58)	0.21
No	1.00		1.00	
Use of intraoperative fluoroscopy				
Yes	3.81 (1.10–10.29)	0.037	3.34 (0.90–9.92)	0.07
No	1.00		1.00	
Dural tear				
Yes	1.27 (0.29–3.72)	0.71		
No	1.00			
Iliac crest bone graft				
Yes	1.66 (0.39–4.86)	0.44		
No	1.00			
Operating time				
≥3 h	9.24 (3.49–31.85)	<0.0001	10.28 (3.31–39.36)	<0.0001
<3 h	1.00		1.00	

ASA American Society of Anesthesiologists, *BMI* body mass index, *SSI* surgical site infection

Contrary to some reports [11, 24], the use of instrumentation was not an independent risk factor for SSI in the current analysis. According to the univariate analysis, the use of instrumentation showed a significant association with deep SSI; however, the instrumentation and operating time factors may be confounding in relation to the occurrence of

Table 3 Microbiologic characteristics of deep SSI

Organism(s)	No. of cases
MRSA	7
Methicillin-resistant CNS	5
Staphylococcus aureus	4
CNS	2
Pseudomonas aeruginosa	1
Corynebacterium sp.	1
MRSA + *Propionibacterium acnes*	1
Unknown	3

CNS coagulase-negative staphylococci, *MRSA* methicillin-resistant *Staphylococcus aureus*, *SSI* surgical site infection

deep SSI in our data. In fact, the operating time was ≥3 h in all patients who received spinal instrumentation and developed deep SSI in the current study. Reducing the surgical time may help prevent SSI following posterior thoracolumbar instrumentation fusion surgery. However, in instrumentation surgery, biofilm formation and treatment difficulty in cases of deep wound infections have been reported [24]. Therefore, care and attention should be paid particularly to patients undergoing instrumentation surgery for the prevention of SSI.

Several reports have described diabetes as a risk factor for SSI after spinal surgeries [5–7]; however, diabetes was not a significant risk factor for deep SSI in the present study according to univariate and multivariate analyses. There is a possibility that the diabetes patients in our study included well-controlled and poor-controlled cases. Hikata et al. [7] reported that poorly-controlled diabetes (HbA1c ≥ 7.0%) was statistically significantly associated with the development of SSI after posterior thoracolumbar spinal instrumentation surgeries compared to well-controlled diabetes (6.1 ≤ HbA1c < 7.0%), and SSI occurred in none of the patients (0%) with well-controlled diabetes (6.1 ≤ HbA1c < 7.0%) in their operated case series.

Several studies have demonstrated that obesity is a patient-related risk factor for SSI [5, 6]; however, BMI was not a significant factor according to univariate and multivariate analyses in our study. Yoshiike et al. [25] reported that the prevalence of obesity in Japanese adults, which is estimated using international criteria (BMI ≥ 30), is lower than that of Western populations. Flegal et al. [26] described that 33.8% of adults in the United States were obese (BMI ≥ 30, 2007 to 2008); whereas, in our case series, 167 of 2,736 patients (6.1%) were obese. This difference in the prevalence of obesity between Japanese and Western populations may affect our study findings. Mahta et al. [27, 28] demonstrated that the thickness of subcutaneous fat at the surgical site was an important risk factor for SSI in posterior spine surgeries, and they described that the thickness of subcutaneous fat at the surgical site is more significant for predicting SSI than BMI.

No patients in our study underwent intrawound application of vancomycin powder, although recent studies have reported its effectiveness in preventing SSI after spinal surgery [29, 30]. In patients considered at high risk for SSI, the use of this treatment may be effective for reducing the incidence of devastating wound infections following spinal surgery.

A limitation of this study is the relatively small sample size of infected patients ($n = 24$), as only patients with deep SSI following specific types of procedures (posterior thoracic and/or lumbar surgery) were included. This contrasts with previous research on SSI that generally focused on a wide variety of spinal procedures and all types of infection [2, 6]. Another limitation is the fact that malnutrition and the number of the operated levels were not included in the factors we assessed. The occurrence of selection bias in patient enrollment cannot be denied; however, we made an effort to minimize this bias by enrolling consecutive patients from multiple centers and not from a single center. The strengths of this study are the relatively large number of surgical procedures. In addition, the prospective multicenter surveillance design allowed for a detailed study of independent risk factors for SSI after spinal surgery by using multivariate logistic regression.

In conclusion, we identified that an operative time ≥3 h, preoperative steroid use, spinal trauma and male gender were independent risk factors for deep SSI following posterior thoracic and/or lumbar spinal surgery in adult patients. The SSI risk factors identified in this study may facilitate the design of protocols for reducing the incidence of SSI in the future.

Conflict of interest The authors declare no conflicts of interest.

References

1. Xing D, Ma JX, Ma XL, Song DH, Wang J, Chen Y, Yang Y, Zhu SW, Ma BY, Feng R. A methodological, systematic review of evidence-based independent risk factors for surgical site infections after spinal surgery. Eur Spine J. 2013;22(3):605–15.
2. Olsen MA, Mayfield J, Lauryssen C, Polish LB, Jones M, Vest J, Fraser VJ. Risk factors for surgical site infection in spinal surgery. J Neurosurg. 2003;98(2 Suppl):149–55.
3. Klein JD, Hey LA, Yu CS, Klein BB, Coufal FJ, Young EP, Marshall LF, Garfin SR. Perioperative nutrition and postoperative complications in patients undergoing spinal surgery. Spine. 1996;21(22):2676–82.
4. Rao SB, Vasquez G, Harrop J, Maltenfort M, Stein N, Kaliyadan G, Klibert F, Epstein R, Sharan A, Vaccaro A, Flomenberg P. Risk factors for surgical site infections following spinal fusion procedures: a case-control study. Clin Infect Dis. 2011;53(7):686–92.
5. Wimmer C, Gluch H, Franzreb M, Ogon M. Predisposing factors for infection in spine surgery: a survey of 850 spinal procedures. J Spinal Disord. 1998;11(2):124–8.
6. Olsen MA, Nepple JJ, Riew KD, Lenke LG, Bridwell KH, Fraser VJ. Risk factors for surgical site infection following orthopaedic spinal operations. J Bone Joint Surg Am. 2008;90(1):62–9.

7. Hikata T, Iwanami A, Hosogane N, Watanabe K, Ishii K, Nakamura M, Kamata M, Toyama Y, Matsumoto M. High preoperative hemoglobin A1c is a risk factor for surgical site infection after posterior thoracic and lumbar spinal instrumentation surgery. J Orthop Sci. 2014;19(2):223–8.
8. Blam OG, Vaccaro AR, Vanichkachorn JS, Albert TJ, Hilibrand AS, Minnich JM, Murphey SA. Risk factors for surgical site infection in the patient with spinal injury. Spine. 2003;28(13):1475–80.
9. Watanabe M, Sakai D, Matsuyama D, Yamamoto Y, Sato M, Mochida J. Risk factors for surgical site infection following spine surgery: efficacy of intraoperative saline irrigation. J Neurosurg Spine. 2010;12(5):540–6.
10. Klekamp J, Spengler DM, McNamara MJ, Haas DW. Risk factors associated with methicillin-resistant staphylococcal wound infection after spinal surgery. J Spinal Disord. 1991;12(3):187–91.
11. Maragakis LL, Cosgrove SE, Martinez EA, Tucker MG, Cohen DB, Perl TM. Intraoperative fraction of inspired oxygen is a modifiable risk factor for surgical site infection after spinal surgery. Anesthesiology. 2009;110(3):556–62.
12. Picada R, Winter RB, Lonstein JE, Denis F, Pinto MR, Smith MD, Perra JH. Postoperative deep wound infection in adults after posterior lumbosacral spine fusion with instrumentation: incidence and management. J Spinal Disord. 2000;13(1):42–5.
13. Sasso RC, Garrido BJ. Postoperative spinal wound infections. J Spinal Disord. 2000;13(1):42–5.
14. Emori TG, Culver DH, Horan TC, Jarvis WR, White JW, Olson DR, Banerjee S, Edwards JR, Martone WJ, Gaynes RP, Hughes JM. National nosocomial infections surveillance system (NNIS): description of surveillance methods. Am J Infect Control. 1991;19(1):19–35.
15. Surveillance National Nosocomial Infections. (NNIS) report, data summary from October 1986-April 1996, issued May 1996. A report from the National Nosocomial Infections Surveillance (NNIS) System. Am J Infect Control. 1996;24(5):380–8.
16. Roberts FJ, Walsh A, Wing P, Dvorak M, Schweigel J. The influence of surveillance methods on surgical wound infection rates in a tertiary care spinal surgery service. Spine. 1998;23(3):366–70.
17. Mangram AJ, Horan TC, Pearson ML, Silver LC, Jarvis WR. Guideline for prevention of surgical site infection, 1999. Hospital Infection Control Practices Advisory Committee. Infect Control Hosp Epidemiol. 1999;20(4):250–78.
18. Owens WD, Felts JA, Spitznagel EL Jr. ASA physical status classifications: a study of consistency of ratings. Anesthesiology. 1978;49(4):239–43.
19. Ahn DK, Park HS, Kim TW, Yang JH, Boo KH, Kim IJ, Lee HJ. The degree of bacterial contamination while performing spine surgery. Asian Spine J. 2013;7(1):8–13.
20. Lim MR, Lee JY, Vaccaro AR. Surgical infections in the traumatized spine. Clin Orthop Relat Res. 2006;444:114–9.
21. Namba RS, Inacio MC, Paxton EW. Risk factors associated with deep surgical site infections after primary total knee arthroplasty: an analysis of 56,216 knees. J Bone Joint Surg Am. 2013;95(9):775–82.
22. Song KH, Kim ES, Kim YK, Jin HY, Jeong SY, Kwak YG, Cho YK, Sung J, Lee YS, Oh HB, Kim TK, Koo KH, Kim EC, Kim JM, Choi TY, Kim HY, Choi HJ, Kim HB. Differences in the risk factors for surgical site infection between total hip arthroplasty and total knee arthroplasty in the Korean Nosocomial Infections Surveillance System (KONIS). Infect Control Hosp Epidemiol. 2012;33(11):1086–93.
23. Kim ES, Kim HB, Song KH, Kim YK, Kim HH, Jin HY, Jeong SY, Sung J, Cho YK, Lee YS, Oh HB, Kim EC, Kim JM, Choi TY, Choi HJ, Kim HY. Prospective nationwide surveillance of surgical site infections after gastric surgery and risk factor analysis in the Korean Nosocomial Infections Surveillance System (KONIS). Infect Control Hosp Epidemiol. 2012;33(6):572–80.
24. Kasliwal MK, Tan LA, Traynelis VC. Infection with spinal instrumentation: review of pathogenesis, diagnosis, prevention, and management. Surg Neurol Int. 2013;4(Suppl 5):S392–403.
25. Yoshiike N, Matsumura Y, Zaman MM, Yamaguchi M. Descriptive epidemiology of body mass index in Japanese adults in a representative sample from the National Nutrition Survey 1990–1994. Int J Obes Relat Metab Disord. 1998;22(7):684–7.
26. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. JAMA. 2010;303(3):235–41.
27. Mehta AI, Babu R, Karikari IO, Grunch B, Agarwal VJ, Owens TR, Friedman AH, Bagley CA, Gottfried ON. The distribution of body mass as a significant risk factor for lumbar spinal fusion postoperative infections. Spine. 2012;37(19):1652–6.
28. Mehta AI, Babu R, Sharma R, Karikari IO, Grunch BH, Owens TR, Agarwal VJ, Sampson JH, Lad SP, Friedman AH, Kuchibhatla M, Bagley CA, Gottfried ON. Thickness of subcutaneous fat as a risk factor for infection in cervical spine fusion surgery. J Bone Joint Surg Am. 2013;95(4):323–8.
29. Sweet FA, Roh M, Sliva C. Intrawound application of vancomycin for prophylaxis in instrumented thoracolumbar fusions: efficacy, drug levels, and patient outcomes. Spine. 2011;36(24):2084–8.
30. Caroom C, Tullar JM, Benton EG Jr, Jones JR, Chaput CD. Intrawound vancomycin powder reduces surgical site infections in posterior cervical fusion. Spine. 2013;38(14):1183–7.

Osteoarthritis and Cartilage 23 (2015) 874–881

Osteoarthritis and Cartilage

Joint space narrowing, body mass index, and knee pain: the ROAD study (OAC1839R1)

S. Muraki † *, T. Akune †, Y. En-yo ‡, M. Yoshida ‡, T. Suzuki §, H. Yoshida ‖, H. Ishibashi ¶, F. Tokimura #, S. Yamamoto ††, S. Tanaka ‡‡, K. Nakamura §§, H. Kawaguchi ‖‖, H. Oka ¶¶, N. Yoshimura ¶¶

† *Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Tokyo, Japan*
‡ *Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan*
§ *National Center for Geriatrics and Gerontology, Aichi, Japan*
‖ *Research Team for Promoting Independence of the Elderly, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan*
¶ *Department of Orthopaedic Surgery, Ina Hospital, Saitama, Japan*
# *Department of Orthopaedic Surgery, Tokyo Geriatric Medical Center, Tokyo, Japan*
†† *Department of Orthopaedic Surgery, Toranomon Hospital, Tokyo, Japan*
‡‡ *Department of Orthopaedic Surgery, Faculty of Medicine, the University of Tokyo, Tokyo, Japan*
§§ *National Rehabilitation Center for Persons with Disabilities, Saitama, Japan*
‖‖ *Department of Orthopaedic Surgery, Japan Community Health Care Organization Tokyo Shinjuku Medical Center, Tokyo, Japan*
¶¶ *Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*

ARTICLE INFO

Article history:
Received 16 July 2014
Accepted 20 January 2015

Keywords:
Cohort study
Epidemiology
Osteoarthritis
Pain

SUMMARY

Objective: The objective of the present study was to clarify the association of joint space narrowing with knee pain in Japanese men and women using a large-scale population-based cohort of the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.
Methods: This study examined the association between minimum joint space width (mJSW) in the medial compartment and pain at the knee. mJSW was measured in the medial and lateral compartments of the knee using a knee osteoarthritis (OA) computer-aided diagnosis system.
Results: From the 3040 participants in the ROAD study, the present study analyzed 2733 participants who completed the radiographic examinations and questionnaires regarding knee pain (975 men and 1758 women; mean age, 69.9 ± 11.2 years). Subjects with lateral knee OA were excluded. After adjustment for age and Body mass index (BMI), medial mJSW, as well as medial mJSW/lateral mJSW, was significantly associated with knee pain. Sex and BMI affected the association of medial mJSW with knee pain. The threshold of medial mJSW was approximately 3 mm in men and 2 mm in women, while that of medial mJSW/lateral mJSW was approximately 60% in both men and women. BMI was found to have a distinct effect on the association of mJSW with pain.
Conclusion: The present cross-sectional study using a large-scale population from the ROAD study showed that joint space narrowing had a significant association with knee pain. The thresholds of joint space narrowing for knee pain were also established.

Introduction

Knee osteoarthritis (OA) is a major public health issue that causes chronic pain and disability[1–3]. The prevalence of radiographic knee OA is high in Japan [4], with 25,300,000 persons aged 40 years and older estimated to have radiographic knee OA[5]. According to the recent National Livelihood Survey of the Ministry of Health, Labour and Welfare in Japan, OA is ranked fourth among diseases that cause disabilities that subsequently require support with activities of daily living[6].

* Address correspondence and reprint requests to: S. Muraki, Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. Tel: 81-3-5800-9178; Fax: 81-3-5800-9179.
E-mail address: murakis-ort@h.u-tokyo.ac.jp (S. Muraki).

http://dx.doi.org/10.1016/j.joca.2015.01.011
1063-4584/

Knee pain is the principal clinical symptom of knee OA[7]. Although much effort has been devoted toward a definition of knee pain, the correlation with radiographic severity of the knee OA was not as strong as one would expect[4,8–10]. One of the reasons for this apparent discrepancy may be the definition of knee OA. Kellgren Lawrence (KL) grading is the conventional system most used to grade the radiographic severity of knee OA[11], but in this categorical system, joint space narrowing is not assessed separately. A recent cross-sectional study showed that the association between joint space narrowing and osteophytosis was not as high as expected on plain radiographs[12]. In addition, joint space narrowing and osteophytosis had distinct effects on QOL[12]. These accumulating lines of evidence have indicated that joint space narrowing and osteophytosis may have distinct etiologic mechanisms, and their progression may be neither constant nor proportional. Although osteophytosis also has some effect on ADL and QOL[12], joint space narrowing is the primary outcome in studies of OA[13]. Thus, to examine the association between knee OA and pain, joint space narrowing should be assessed separately. Chan *et al.* examined the association of joint space narrowing and duration of pain in patients with knee OA and found a significant association[14], but joint space narrowing was defined by categorical methods. Because categorical methods are statistically less powerful than continuous methods, the association between pain and knee OA might have been underestimated. To overcome this, a fully automatic system that can quantify the joint space width of knee OA on standard radiographs and allows for objective, accurate, and simple assessment of the structural severity of knee OA was developed[15]. Thus far, Kinds *et al.* measured joint space width and found significant associations with clinical outcomes[16], but the threshold of joint space width for clinical outcomes remains unclear.

Sex differences have been observed in knee OA. The prevalence of knee OA is higher in women than men, and the association of knee pain with knee OA also differs by sex[4]. Thus, the impact of joint space narrowing and osteophytosis on QOL may also differ between the sexes. Obesity is also one of the few established risk factors for knee OA and pain[17–23]. This suggests that a distinct association of joint space narrowing with pain may be found in subjects with and without obesity. However, to the best of our knowledge, there are no population-based studies that assess the effect of obesity on the association of joint space narrowing with pain.

Therefore, the objective of this study was to clarify the association of joint space narrowing with pain at the knee among Japanese men and women using a fully automatic system to measure joint space width in a large-scale, population-based cohort from the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study. Furthermore, the threshold of minimum joint space width (mJSW) or medial mJSW/lateral mJSW for pain was determined using receiver operating characteristic (ROC) curve analysis.

Subjects and methods

Subjects

The ROAD study is a nationwide prospective study designed to establish epidemiologic indices for the evaluation of clinical evidence for the development of a disease-modifying treatment for bone and joint diseases (with OA and osteoporosis as the representative bone and joint diseases). It consists of population-based cohorts in several communities in Japan. A detailed profile of the ROAD study has been reported elsewhere[4,5,24], and, thus, only a brief summary is provided here. To date, we have completed the creation of a baseline database including clinical and genetic information for 3040 inhabitants (1061 men and 1,979 women) ranging in age from 23 to 95 years (mean, 70.3 years), who were recruited from resident registration listings in three communities: an urban region in Itabashi, Tokyo, a mountainous region in Hidakagawa, Wakayama, and a coastal region in Taiji, Wakayama. All participants provided their written, informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology. Anthropometric measurements, including height and weight, were taken, and body mass index (BMI; weight [kg]/height^2 [m^2]) was calculated. Furthermore, all participants were also interviewed by well-experienced orthopedists regarding pain in both knees, by asking: “Have you experienced right knee pain on most days in the past month, in addition to now?” and “Have you experienced left knee pain on most days in the past month, in addition to now?”. Subjects who answered “yes” were defined as having knee pain. Among the 3040 subjects who participated in the baseline study, 30 (1.0%) who underwent unilateral or total knee arthroplasty were excluded. In addition, 35 (1.1%) whose radiographic examinations were insufficient for measuring joint space width, and 195 (6.4%) with lateral knee OA were excluded. One reason for excluding lateral knee OA is that most knee OA in Japan is medial type[4]. The other reason is that medial and lateral knee OA have distinct characteristics, and joint space narrowing occurs in the medial compartment in medial knee OA, but medial joint space width may not change or be larger in lateral knee OA. Furthermore, 47 patients (1.5%) who provided incomplete questionnaires regarding pain and so on were excluded, leaving a total of 2733 (89.9%) subjects (975 men and 1758 women).

Radiographic assessment

All participants underwent radiographic examinations of both knees using an anterior-posterior view with weight-bearing and foot map positioning by experienced radiological technologists. The beam was positioned parallel to the floor with no angle and aimed at the joint space. To visualize the joint space properly and to centralize the patella over the lower end of the femur, fluoroscopic guidance with an anterior-posterior X-ray beam was used, and the images were downloaded into Digital Imaging and Communication in Medicine (DICOM) format files. Knee radiographs were read without knowledge of participant clinical status by a single experienced orthopedist (S.M.) using the KL radiographic atlas for overall knee radiographic grades[11], and knee OA was defined as KL grade 2 or greater. The KOACAD system was used to measure mJSW in the medial compartment and OPA at the medial tibia[15], and the knee with the lower mJSW was defined as the designated knee for a participant. The KOACAD system is a fully automatic system that can quantify the major features of knee OA on standard radiographs and allows for objective, accurate, and simple assessment of the structural severity of knee OA in general clinical practice. This system was programmed to measure mJSW in the medial and lateral compartments using digitized knee radiographs. The KOACAD system has been described in detail elsewhere[15,25]. The KOACAD system was applied to the DICOM data by the experienced orthopedist who developed this system (H.O.); measurement reliability has been shown to be good[15], and the intraclass coefficient of correlation for medial mJSW measured on radiographs for an individual with weight-bearing and foot map positioning was 0.96. Reference values for OPA and mJSW by sex and age strata in Japan using the KOACAD system have been published previously[25]. Lateral knee OA was defined as KL grade 2 or greater with lower lateral mJSW than medial mJSW.

Statistical analysis

Differences in age, height, weight, BMI, mJSW, and medial/lateral mJSW between men and women and between subjects with and

without pain were examined by the non-paired Student's *t*-test. The prevalence of knee OA was compared between men and women by the χ^2 test. Associations of age, BMI, mJSW, and medial/lateral mJSW with knee pain were determined using multiple logistic regression analysis after adjustment for age, sex, and BMI overall, and after adjustment for age and BMI in men and women. In addition, subjects were classified according to mJSW (<1 mm, ≥1–<2 mm, ≥2–<3 mm, ≥3–<4 mm, ≥4 mm), and the associations of mJSW<1 mm, ≥1–<2 mm, ≥2–<3 mm, and ≥3–<4 mm with pain were determined using multiple logistic regression analysis after adjustment for age and BMI, compared with mJSW≥4 mm. To clarify the effect of BMI on the association of mJSW with pain, subjects were further classified into 10 groups according to mJSW and BMI (BMI <23 kg/m^2: mJSW<1 mm, ≥1–<2 mm, ≥2–<3 mm, ≥3–<4 mm, ≥4 mm; BMI ≥ 23 kg/m^2: mJSW<1 mm, ≥1–<2 mm, ≥2–<3 mm, ≥3–<4 mm, ≥4 mm), and the association with pain was determined using multiple logistic regression analysis after adjustment for age, compared with BMI <23 kg/m^2 and mJSW ≥4 mm. Subjects were also classified according to medial/lateral mJSW (<30%, ≥30–<40%, ≥40–<50%, ≥50–<60%, ≥60–<70%, ≥70–<80%, ≥80%), and the associations of medial/lateral mJSW<30%, ≥30–<40%, ≥40–<50%, ≥50–<60%, ≥60–<70%, and ≥70–<80% with pain were determined using multiple logistic regression analysis after adjustment for age and BMI, compared with medial/lateral mJSW ≥80%. The thresholds of mJSW and medial/lateral mJSW for pain were determined using ROC curve analysis. Data analyses were performed using SAS version 9.0 (SAS Institute Inc., Cary, NC).

Results

The characteristics of the 2733 participants in the present study are shown in Table I. The prevalence of knee OA was significantly higher in women than in men. The mJSW and medial mJSW/lateral mJSW were significantly lower in women than in men. The participants in the present study were significantly younger than the non-participants ($P < 0.05$), while BMI was not significantly different between them (non-participants: age, 74.3 ± 7.9 years; BMI, 23.1 ± 3.2 kg/m^2).

Table II shows age, BMI, mJSW, and medial/lateral mJSW in subjects with and without pain. For the right knee, overall and in women, subjects with pain were older and had higher BMI, narrower mJSW, and smaller medial/lateral mJSW than those without pain. In men, subjects with pain had higher BMI, narrower mJSW, and smaller medial/lateral mJSW than subjects without pain, while age was not significantly different in men with and without pain. For the left knee, results were similar except for age in men. Associations of mJSW and medial/lateral mJSW with right and left knee pain were next examined using multiple logistic regression analysis after adjustment for age, sex, and BMI overall, and after adjustment for age and BMI in men and women (Table II). Odds ratios (ORs) of mJSW (1-mm decrease) for pain were higher than 2, and the ORs of medial/lateral mJSW (10% decrease) for pain were 1.2–1.3.

Subjects were then classified according to mJSW (<1 mm, ≥1–<2 mm, ≥2–<3 mm, ≥3–<4 mm, ≥4 mm), and the prevalence of knee pain was examined (Fig. 1, Supplementary Table I). The prevalence of knee pain was more than 60% in subjects with mJSW<1 mm, while it was less than 10% in those with mJSW ≥4 mm. The OR for pain was also calculated after adjustment for age and BMI. Men with mJSW <1 mm, ≥1–<2 mm, and ≥2–<3 mm had significantly higher rates of pain than those with mJSW ≥4 mm, but men with mJSW ≥3–<4 mm did not (Table III). The OR for pain in men with mJSW <1 mm was around 40. Women with mJSW <1 mm and ≥1–<2 mm had significantly higher rates of pain than those with mJSW ≥4 mm, but, women with mJSW ≥2–<3 mm and ≥3–<4 mm did not. The ORs for pain in women with mJSW <1 mm were 12–14. Subjects were further classified into 10 groups according to BMI and mJSW (BMI < 23 kg/m^2: mJSW < 1 mm, ≥1–<2 mm, ≥2–<3 mm, ≥3–<4 mm, ≥4 mm; BMI ≥ 23 kg/m^2: mJSW < 1 mm, ≥1–<2 mm, ≥2–<3 mm, ≥3–<4 mm, ≥4 mm), and the ORs for pain were calculated (Supplementary Table II). In men, mJSW<1 mm and ≥1–<2 mm with BMI <23 kg/m^2 and mJSW < 1 mm, ≥1–<2 mm, and ≥2–<3 mm with BMI ≥23 kg/m^2 were significantly associated with pain compared with mJSW ≥4 mm with BMI <23 kg/m^2. In women at the right knee, mJSW < 1 mm with BMI <23 kg/m^2 and mJSW ≥0–<1 mm and ≥1–<2 mm with BMI ≥23 kg/m^2 were significantly associated with pain compared with mJSW ≥4 mm with BMI <23 kg/m^2. In women at the left knee, mJSW < 1 mm and ≥1–<2 mm with BMI <23 kg/m^2 and mJSW < 1 mm, ≥1–<2 mm, and ≥2–<3 mm with BMI ≥23 kg/m^2 were significantly associated with pain compared with mJSW ≥4 mm with BMI <23 kg/m^2.

Subjects were also classified according to medial/lateral mJSW (<30%, ≥30–<40%, ≥40–<50%, ≥50–<60%, ≥60–<70%, ≥70–<80%, ≥80%), and the prevalence of knee pain was examined (Fig. 2, Supplementary Table III). The prevalence of knee pain was approximately 60% in subjects with medial mJSW/lateral mJSW < 30%, while it was approximately 10% in those with medial mJSW/lateral mJSW ≥80%. The ORs for pain were also calculated after adjustment for age and BMI (Table IV). Men and women with mJSW <30%, ≥30–<40%, ≥40–<50%, and ≥50–<60% had higher rates of pain compared with those with mJSW ≥80%, except for men with mJSW≥50–<60% at the right knee. The ORs for pain in men with mJSW <30% were 14–20. The OR for pain in women with mJSW <30% was around 10.

The threshold values of mJSW for knee pain were then determined using ROC curve analysis (Supplementary Fig. 1). In men, the threshold values of mJSW for pain at the right and left knees were 2.87 mm (sensitivity 0.67, specificity 0.65, AUC 0.70, 95% confidence interval (CI) 0.64–0.75) and 2.82 mm (sensitivity 0.62, specificity 0.67, AUC 0.72, 95% CI 0.66–0.77), respectively. In women, the threshold values of mJSW for pain at the right and left knees were 2.01 mm (sensitivity 0.43, specificity 0.689, AUC 0.69, 95% CI 0.66–0.73) and 2.44 mm (sensitivity 0.59, specificity 0.75, AUC 0.71, 95% CI 0.67–0.74), respectively. Threshold values of medial/lateral mJSW for knee pain were also determined using ROC curve analysis (Supplementary Fig. 2). In men, the threshold values of medial/lateral mJSW for pain at the right and left knees were 55.2% (sensitivity 0.45, specificity 0.68, AUC 0.66, 95% CI 0.60–0.72) and

Table I
Characteristics of the subjects in the present study

	Overall	Men	Women	*P* values
N	2733	975	1758	
Age, years	69.9 ± 11.2	70.8 ± 10.8	69.4 ± 11.4	0.0012
Height, cm	154.4 ± 8.9	162.5 ± 6.7	150.0 ± 6.4	<0.0001
Weight, kg	55.1 ± 10.3	61.4 ± 10.0	51.6 ± 8.7	<0.0001
BMI, kg/m^2	23.0 ± 3.3	23.2 ± 3.0	22.9 ± 3.5	0.0493
Right knee				
Knee OA, %	45.3	33.6	51.8	<0.0001
mJSW, mm	2.8 ± 1.0	3.2 ± 0.9	2.6 ± 0.9	<0.0001
medial mJSW/lateral mJSW, %	68.7 ± 30.1	71.1 ± 22.2	67.4 ± 33.6	0.0007
Left knee				
Knee OA, %	47.5	35.8	54	<0.0001
mJSW, mm	2.9 ± 1.0	3.3 ± 01.0	2.7 ± 0.9	<0.0001
medial mJSW/lateral mJSW, %	70.8 ± 26.3	73.9 ± 22.7	69.1 ± 28.0	<0.0001

Except where indicated otherwise, values are means ± SD.
Knee OA was defined as Kellgren Lawrence grade 2 or worse.
BMI, body mass index; OA, osteoarthritis; mJSW, minimum joint space width.

Table II
Associations of age, BMI, mJSW, and medial mJSW/lateral mJSW with knee pain

	Right knee					Left knee				
	Pain +	Pain −	Adjusted OR	95% CI	*P* values	Pain +	Pain −	Adjusted OR	95% CI	*P* values
Overall										
N										
Age, years	72.4 ± 8.6	69.3 ± 11.6*	1.01	1.00–1.03	0.0499	72.8 ± 8.4	69.3 ± 11.6*	1.01	0.99–1.02	0.3226
BMI, kg/m^2	24.4 ± 3.6	22.7 ± 3.2*	1.12	1.08–1.16	<0.0001	24.2 ± 3.5	22.8 ± 3.2*	1.10	1.06–1.14	<0.0001
mJSW, mm (1-mm decrease)	2.1 ± 1.1	2.9 ± 0.9*	2.17	1.92–2.50	<0.0001	2.2 ± 1.1	2.9 ± 0.9*	2.22	1.96–2.56	<0.0001
medial mJSW/lateral mJSW, % (10% decrease)	54.3 ± 30.7	71.9 ± 29.0*	1.30	1.24–1.37	<0.0001	56.6 ± 30.4	71.0 ± 29.5*	1.22	1.16–1.29	<0.0001
Men										
N										
Age, years	71.7 ± 9.4	70.7 ± 11.0	0.99	0.96–1.02	0.5095	72.8 ± 8.7	70.5 ± 11.0*	0.98	0.95–1.01	0.2278
BMI, kg/m^2	24.2 ± 3.0	23.0 ± 3.0*	1.09	1.01–1.18	0.0207	24.1 ± 3.1	23.0 ± 3.0*	1.08	0.995–1.17	0.0635
mJSW, mm (1-mm decrease)	2.4 ± 1.2	3.3 ± 0.9*	2.33	1.82–3.03	<0.0001	2.6 ± 1.1	3.2 ± 0.9*	2.50	1.92–3.23	<0.0001
medial mJSW/lateral mJSW, % (10% decrease)	57.2 ± 27.6	72.8 ± 20.8*	1.33	1.19–1.49	<0.0001	61.3 ± 29.6	72.3 ± 20.8*	1.28	1.15–1.43	<0.0001
Women										
N										
Age, years	72.6 ± 8.4	68.5 ± 12.0*	1.02	1.005–1.04	0.0138	72.7 ± 8.3	68.6 ± 11.9*	1.02	0.997–1.04	0.0964
BMI, kg/m^2	24.4 ± 3.7	22.5 ± 3.3*	1.12	1.08–1.17	<0.0001	24.3 ± 3.6	22.6 ± 3.4*	1.11	1.06–1.16	<0.0001
mJSW, mm (1-mm decrease)	2.0 ± 1.1	2.8 ± 0.8*	2.13	1.82–2.50	<0.0001	2.1 ± 1.1	2.7 ± 0.8*	2.17	1.85–2.56	<0.0001
medial mJSW/lateral mJSW, % (10% decrease)	52.2 ± 31.5	71.4 ± 33.0*	1.30	1.22–1.38	<0.0001	55.1 ± 30.6	70.3 ± 33.7*	1.21	1.14–1.28	<0.0001

*$P < 0.05$ by non-paired Student's *t*-test.
Adjusted ORs were calculated by multiple logistic regression analysis after adjustment for age, sex, and BMI overall and after adjustment for age and BMI in men and women.
BMI, body mass index; mJSW, minimum joint space width; OR, odds ratio; CI, confidence interval.

57.9% (sensitivity 0.49, specificity 0.84, AUC 0.70, 95% CI 0.64–0.76), respectively. In women, the threshold values of medial/lateral mJSW for pain at the right and left knees were 57.9% (sensitivity 0.57, specificity 0.75, AUC 0.69, 95% CI 0.66–0.73) and 57.7% (sensitivity 0.58, specificity 0.76, AUC 0.71, 95% CI 0.68–0.74), respectively.

Discussion

Joint space narrowing is the primary outcome in studies of knee OA[20], because cartilage damage, which is one of the main causes of knee symptoms, is seen as a smaller mJSW[16]. Previous studies have shown significant associations of joint space narrowing with pain[14,16], though the threshold of joint space width for pain remained unclear. This is the first study to clarify the effect of joint space narrowing on knee pain using a large-scale, population-based, cohort study. In addition, joint space narrowing was evaluated not by categorical grade but by continuous values, using mJSW at the knee. In the present study, mJSW < 3 mm in men and mJSW < 2 mm in women were significantly associated with knee pain, compared with mJSW ≥4 mm, and the OR of mJSW < 1 mm for knee pain was quite high, particularly in men. It was also found that the effect of mJSW on pain was affected by BMI. Medial mJSW/lateral/mJSW < 60% was also significantly associated with knee pain in men and women, compared with medial mJSW/lateral mJSW ≥ 80%. Using ROC curve analysis, the thresholds of mJSW in men and women were found to be approximately 3 mm in men and

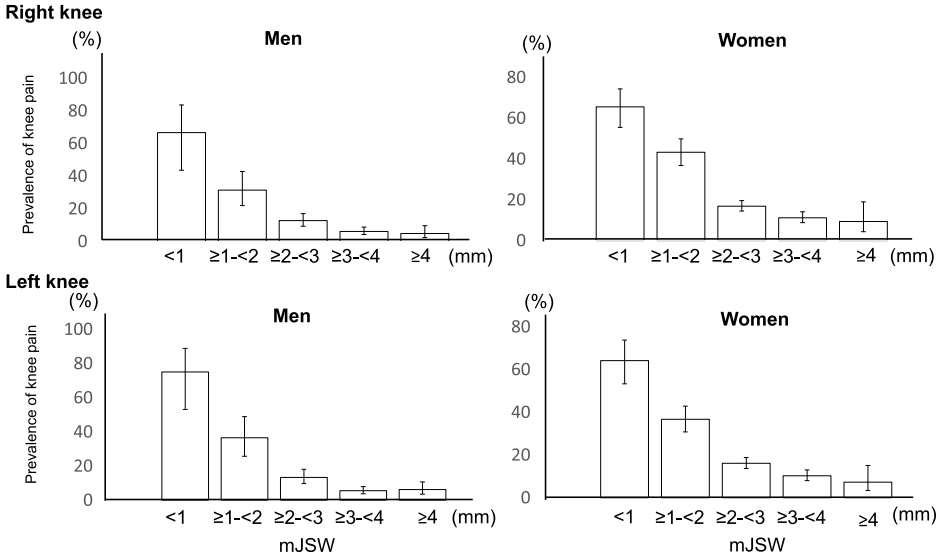


Fig. 1. Prevalence of knee pain by mJSW (mm) in men and women. mJSW, minimum joint space width.

Table III
OR of knee pain by medial mJSW

	Men				Women			
	Crude OR	95% CI	Adjusted OR	95% CI	Crude OR	95% CI	Adjusted OR	95% CI
Right knee								
<1 mm	39.0	11.9–146.4	39.4	11.6–151.8	18.4	7.7–51.9	12.3	5.0–35.3
≥1–<2 mm	9.0	3.8–23.9	8.5	3.5–23.0	7.4	3.3–19.7	5.9	2.6–15.9
≥2–<3 mm	2.9	1.3–7.2	3	1.4–7.6	2	0.9–5.2	1.8	0.8–4.9
≥3–<4 mm	1.3	0.6–3.3	0.8	0.2–3.2	1.2	0.5–3.3	1.3	0.6–3.5
≥4 mm	1		1		1		1	
Left knee								
<1 mm	45.5	14.9–163.3	38.1	11.9–142.1	22.7	9.4–64.3	14.0	5.7–40.2
≥1–<2 mm	8.7	4.0–20.2	8.5	3.8–20.1	7.3	3.3–19.5	5.3	2.3–14.2
≥2–<3 mm	2.3	1.2–5.0	2.3	1.2–5.1	2.4	1.1–6.4	2.0	0.9–5.3
≥3–<4 mm	0.9	0.4–1.9	0.9	0.4–1.9	1.4	0.6–3.8	1.4	0.6–3.7
≥4 mm	1		1		1		1	

Adjusted ORs were calculated by multiple logistic regression analysis after adjustment for age and BMI.
mJSW, minimum joint space width; CI, confidence interval.

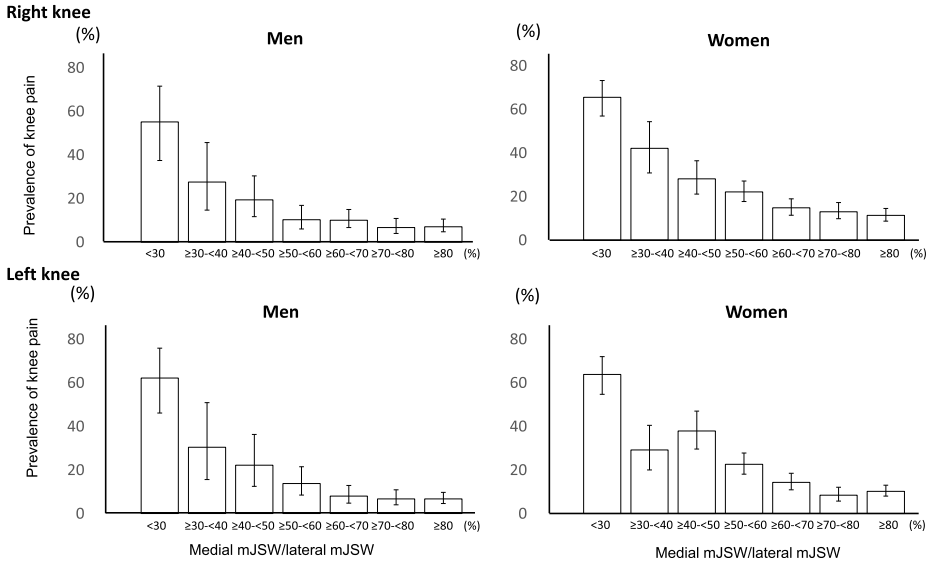


Fig. 2. Prevalence of knee pain by medial mJSW/lateral mJSW (%) in men and women. mJSW, minimum joint space width.

Table IV
OR of knee pain by the ratio of medial mJSW to lateral mJSW

	Men				Women			
	Crude OR	95% CI	Adjusted OR	95% CI	Crude OR	95% CI	Adjusted OR	95% CI
Right knee								
<30%	16.0	6.9–38.2	14.5	6.1–35.0	14.8	9.4–23.7	9.8	6.1–16.0
≥30–<40%	5.0	1.9–12.2	4.4	1.6–10.9	5.7	3.2–10.0	4.2	2.3–7.6
≥40–<50%	3.1	1.5–6.5	2.7	1.2–5.7	3.1	1.9–4.9	2.4	1.5–4.0
≥50–<60%	1.5	0.7–3.0	1.4	0.7–2.9	2.2	1.5–3.3	2.1	1.4–3.2
≥60–<70%	1.5	0.8–2.8	1.5	0.8–2.8	1.4	0.9–2.0	1.3	0.9–2.0
≥70–<80%	0.9	0.5–1.9	0.9	0.5–1.9	1.2	0.8–1.8	1.1	0.7–1.8
≥80%	1		1		1		1	
Left knee								
<30%	23.0	10.7–51.3	18.9	8.6–43.1	15.5	9.8–25.0	10.3	6.4–16.9
≥30–<40%	6.1	2.2–15.8	5.7	2.0–14.8	3.6	2.0–6.4	2.8	1.5–5.0
≥40–<50%	4.0	1.7–8.8	3.6	1.5–8.0	5.4	3.4–8.6	4.3	2.7–7.0
≥50–<60%	2.2	1.1–4.4	2.1	1.0–4.1	2.6	1.7–3.8	2.3	1.5–3.4
≥60–<70%	1.2	0.6–2.4	1.2	0.6–2.4	1.5	0.97–2.2	1.4	0.9–2.1
≥70–<80%	1.0	0.5–1.9	1.0	0.5–2.0	0.8	0.5–1.3	0.8	0.5–1.3
≥80%	1		1		1		1	

Adjusted ORs were calculated by multiple logistic regression analysis after adjustment for age and BMI.
mJSW, minimum joint space width; CI, confidence interval.

2 mm in women, while those of medial mJSW/lateral mJSW were approximately 60% in both men and women.

Although much effort has been devoted toward a definition of knee pain, the correlation with radiographic severity of knee OA was not as strong as one would expect[4,8–10]. In fact, our previous study showed that the OR of severe knee OA defined as KL grade 3 or 4 for knee pain was 8.6 in men and 4.4 in women[4], which was significant, but the OR was not as high as expected. One of the reasons for this is that knee pain may arise from a variety of structures other than joint cartilage, such as menisci, synovium, ligaments, bursae, bone, and bone marrow[26–30]. Another reason may be due to the definition of knee OA. Knee OA is characterized by the pathological features of joint space narrowing and osteophytosis. However, most conventional systems for grading radiographic severity have been categorical grades, such as KL grading[11], which cannot assess joint space narrowing individually. Several studies have shown that knee OA has a strong effect on QOL[31–34], but in these studies, knee OA was defined by categorical grades such as KL grade or American College of Rheumatology (ACR) grade, total knee arthroplasty, and self-questionnaire. In addition, joint space narrowing was separately evaluated using a radiographic atlas of individual features published by the Osteoarthritis Research Society International (OARSI) in 1995[35] and revised in 2007[36]. Chan *et al.* examined the association of joint space narrowing and duration of pain in patients with knee OA using categorical methods[14]. However, the grading is still limited in reproducibility and sensitivity due to the subjective judgment of individual observers and the categorical classification. Furthermore, because categorical methods are statistically less powerful than continuous methods, the association between pain and knee OA might have been underestimated in previous studies. Kinds *et al.* measured joint space width and found significant associations with clinical outcomes[16], but the threshold of joint space width for clinical outcomes remained unclear. In the present study, to overcome this problem, joint space width was evaluated using a fully automatic system, and the OR of mJSW <1 mm for knee pain was quite high, particularly in men, and it was possible to establish the threshold values of mJSW for knee pain, which may indicate that mJSW is better for diagnosing knee OA than KL grade. In the present study, 6% of men with mJSW <3 mm and 14% of women with mJSW <2 mm, which were the threshold values in the present study, had knee pain. In addition, our previous study showed that 10% of men without knee OA and 20% of women without knee OA had knee pain[4]. These subjects have knee pain, despite having no radiographical changes. This indicates that at least 10% and 20% of knee pain in men and women, respectively, may be explained by factors other than radiographical changes.

In the present study, sex differences were found in the association of mJSW with pain. These discrepancies between the sexes are explained by several factors. First, women are more susceptible to pain than men[4]. In fact, our previous study showed that the OR for knee pain in women without radiographic knee OA (KL = 0/1) was greater than that in men without radiographic knee OA[4]. In the present study, the prevalence of knee pain was 5–6% in men with mJSW ≥4 mm, while it was 7–9% in women with mJSW ≥4 mm. This high prevalence of knee pain in women with mJSW ≥4 mm, which are reference data, may partly explain the lower OR for knee pain in women than men. Second, men with normal knees had wider joint space widths than women with normal knees. Our previous study showed that mean mJSW in men with KL = 0 was approximately 4 mm, while that in women with KL = 0 was approximately 3 mm[25]. This means that subjects with mJSW = 3 mm have a normal knee in women, while they have joint space narrowing at the knee in men. In addition, mJSW = 1 mm means 75% cartilage loss in men, while it represent 67% cartilage loss in women. In fact, the associations of medial mJSW/lateral mJSW with pain were similar in men and women, which may also explain the lower OR for knee pain in women than men.

Obesity is one of the few established risk factors for knee OA and pain[17–23]. A clinical review article reported that 69% of knee replacements in middle-aged females can be attributed to obesity[22], and it has been estimated that, if overweight and obese individuals reduced their weight to reach normal BMIs, about 50% of knee OA cases would be eliminated[21]. However, to the best of our knowledge, there are no population-based studies that assess the effect of obesity on the association of joint space narrowing with pain. In the present study, a distinct effect of BMI was found on the association of mJSW with pain. For example, at the right knee in women, mJSW ≥1–<2 mm in women with BMI ≥23 kg/m^2 was significantly associated with pain, while mJSW ≥1–<2 mm in women with BMI <23 kg/m^2 was not, compared to mJSW≥4 mm with BMI <23 kg/m^2. In addition, the OR was similar between mJSW ≥1–<2 mm with BMI ≥23 kg/m^2 and mJSW <1 mm with BMI <23 kg/m^2. These indicate that weight loss may be an effective way to reduce knee pain even when joint space narrowing is present at the knee.

There are limitations in the present study. First, this was a large-scale, population-based, cross-sectional study of baseline data. Thus, causal relationships could not be determined. The ROAD study is a longitudinal survey, so further progress may help elucidate any causal relationships. Second, the threshold in the present study was calculated by a particular statistical method, but certain situations may favor sensitivity over specificity, e.g., screening. In addition, the sensitivity and specificity were modest in the present study. These may be partly explained by the fact that knee pain can arise from a variety of structures other than joint cartilage, such as menisci, synovium, ligaments, bursae, bone, and bone marrow[26–30], which are unable to be assessed radiologically. However, using the KOACAD system, it was possible to demonstrate strong associations of mJSW with knee pain and to establish the threshold of mJSW for knee pain, which may indicate that mJSW is more useful than categorical methods for diagnosing knee OA. Third, cases with lateral knee OA were excluded, leading to a selective sample. One reason for excluding lateral knee OA is that most knee OA in Japan is medial type. There are racial differences in the ratio of lateral to medial knee OA, and previous studies showed that the ratio of lateral to medial knee OA was 0.20 in Caucasian and 0.64 in Chinese populations[37]. In the Amsterdam OA Cohort, lateral knee OA is rather common, and it occurs in association with OA features in other knee compartments[38]. However, our previous study showed that the ratio of lateral to medial knee OA was 0.10 in Japan, which indicates that knee OA was medial type. The other reason for excluding lateral knee OA is that medial and lateral knee OA have distinct characteristics, because, in medial knee OA, there is narrowing of the medial mJSW, while in lateral knee OA, there is narrowing of the lateral mJSW. Thus, the effect of medial mJSW on pain may be obscured by lateral knee OA, because medial joint space width may not change or be larger in lateral knee OA. Thus, the aim of the present study was to clarify the effect of medial knee OA on pain, although excluding lateral OA leads to a selective sample. Lastly, it was not possible to clarify whether the threshold in the present study can apply to other races or populations, because the prevalence of knee OA and the ratio of medial knee OA/lateral knee OA are quite different among races[4,37,38], and the association of knee OA with pain among them may be quite different. To clarify this, international collaborative studies using the KOACAD system are needed.

In conclusion, the present cross-sectional study using a large population from the ROAD study showed that joint space narrowing was strongly associated with knee pain. The threshold of mJSW with knee pain was approximately 3 mm in men and 2 mm

in women, while the threshold of medial mJSW/lateral mJSW was approximately 60% in both men and women. BMI was found to have a distinct effect on the association of mJSW with pain. Further studies, along with continued longitudinal surveys in the ROAD study, will help improve our understanding of the mechanisms of joint space narrowing at the knee and their relationship with pain.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research (S19109007, B20390182, C20591737, C20591774), for Young Scientists (A18689031), and for Exploratory Research (19659305) from the Japanese Ministry of Education, Culture, Sports, Science, and Technology, grants (H17-Men-eki-009, H18-Choujyu-037, H20-Choujyu-009, H21-Chouju-Wakate-011 and H22-Chouju-Wakate-007) from the Ministry of Health, Labour and Welfare, a Research Aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1), and Grant No.166 from the Japan Orthopaedics and Traumatology Foundation.

The authors would like to thank Dr Anamizu and members of the Department of Orthopedics; Mr Kutsuma and other members of the Department of Radiology at Tokyo Metropolitan Geriatric Medical Center; Mrs Tomoko Takijiri and other members of the Public Office in Hidakagawa Town; and Mrs Tamako Tsutsumi and Mrs Kanami Maeda, and other members of the Public Office in Taiji Town, for their assistance in locating and scheduling participants for examinations.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.joca.2015.01.011.

Author contributions

All authors have made substantial contributions to all three of the following sections:

(1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data;
(2) drafting the article or revising it critically for important intellectual content; and
(3) final approval of the version to be submitted.

Conflicts of interest

There are no conflicts of interest.

References

1. Sharma L, Kapoor D. Epidemiology of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM, Eds. Osteoarthritis: Diagnosis and Medical/Surgical Management. 4th edn. Philadelphia: Lippincott Williams & Wilkins; 2007:3–26.
2. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, *et al*. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health 1994;84:351–8.
3. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41:1343–55.
4. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, *et al*. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthritis Cartilage 2009;17:1137–43.
5. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, *et al*. Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: the research on Osteoarthritis/Osteoporosis Against Disability (ROAD). J Bone Miner Metab 2009;27:620–8.
6. Ministry of Health, Labour and Welfare. The Outline of the Results of National Livelihood Survey 2010. Available at: http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/index.html.
7. Linaker CH, Walker-Bone K, Palmer K, Cooper C. Frequency and impact of regional musculoskeletal disorders. Baillieres Best Pract Res Clin Rheumatol 1999;13:197–215.
8. Summers MN, Haley WE, Reveille JD, Alarcon GS. Radiographic assessment and psychologic variables as predictors of pain and functional impairment in osteoarthritis of the knee or hip. Arthritis Rheum 1988;31:204–9.
9. Cicuttini FM, Baker J, Hart DJ, Spector TD. Association of pain with radiological changes in different compartments and views of the knee joint. Osteoarthritis Cartilage 1996;4:143–7.
10. Wluka AE, Wolfe R, Stuckey S, Cicuttini FM. How does tibial cartilage volume relate to symptoms in subjects with knee osteoarthritis? Ann Rheum Dis 2004;63:264–8.
11. Kellgren JH, Lawrence JS, Eds. The Epidemiology of Chronic Rheumatism: Atlas of Standard Radiographs of Arthritis. Oxford: Blackwell Scientific; 1963.
12. Muraki S, Oka H, Akune T, En-yo Y, Yoshida M, Suzuki T, *et al*. Independent association of joint space narrowing and osteophyte formation at the knee with health-related quality of life in Japan: a cross-sectional study. Arthritis Rheum 2011;63:3859–64.
13. Altman R, Brandt K, Hochberg M, Moskowitz R. Design and conduct of clinical trials in people with osteoarthritis. Osteoarthritis Cartilage 1996;4:217–43.
14. Chan WP, Huang GS, Hsu SM, Chang YC, Ho WP. Radiographic joint space narrowing in osteoarthritis of the knee: relationship to meniscal tears and duration of pain. Skeletal Radiol 2008;37:917–22.
15. Oka H, Muraki S, Akune T, Mabuchi A, Suzuki T, Yoshida H, *et al*. Fully automatic quantification of knee osteoarthritis severity on plain radiographs. Osteoarthritis Cartilage 2008;16:1300–6.
16. Kinds MB, Marijnissen AC, Bijlsma JW, Boers M, Lafeber FP, Welsing PM. Quantitative radiographic features of early knee osteoarthritis: development over 5 years and relationship with symptoms in the CHECK cohort. J Rheumatol 2014;40:58–65.
17. Schouten JSAG. A 12 Year Follow Up Study on Osteoarthritis of the Knee in the General Population [thesis]. Rotterdam: Erasmus University Medical School; 1991.
18. Hart DJ, Doyle DV, Spector TD. Incidence and risk factors for radiographic knee osteoarthritis in middle-aged women: the Chingford Study. Arthritis Rheum 1999;42:17–24.
19. Cooper C, Snow S, McAlindon TE, Kellingray S, Stuart B, Coggon D, *et al*. Risk factors for the incidence and progression of radiographic knee osteoarthritis. Arthritis Rheum 2000;43:995–1000.
20. Altman RD, Fries JF, Bloch DA, Carstens J, Cooke TD, Genant H, *et al*. Radiographic assessment of progression in osteoarthritis. Arthritis Rheum 1987;30:1214–25.
21. Coggon D, Reading I, Croft P, McLaren M, Barrett D, Cooper C. Knee osteoarthritis and obesity. Int J Obes Relat Metab Disord 2001;25:622.
22. D'Arcy Y. Pain and obesity. Nurs Manage 2012;43:21–6.

23. MacFarlane G, de Silva V, Jones G. The relationship between body mass index across the life course and knee pain in adulthood: results from the 1958 birth cohort study. Rheumatology 2011;50:2251–6.
24. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol 2010;39:988–95.
25. Oka H, Muraki S, Akune T, Nakamura K, Kawaguchi H, Yoshimura N. Normal and threshold values of joint space width, joint space area, osteophyte area and fibro-tibial angle using a computer-assisted measuring system (KOACAD) to evaluate knee osteoarthritis: the ROAD study. J Orthop Sci 2010;15:781–9.
26. Saito T, Koshino T. Distribution of neuropeptides in synovium of the knee with osteoarthritis. Clin Orthop Relat Res 2000;376:172–82.
27. Bollet AJ. Edema of the bone marrow can cause pain in osteoarthritis and other diseases of bone and joints. Ann Intern Med 2001;134:591–3.
28. Teichtahl AJ, Wluka AE, Morris ME, Davis SR, Cicuttini FM. The relationship between the knee adduction moment and knee pain in middle-aged women without radiographic osteoarthritis. J Rheumatol 2006;33:1845–8.
29. Thorp LE, Sumner DR, Wimmer MA, Block JA. Relationship between pain and medial knee joint loading in mild radiographic knee osteoarthritis. Arthritis Rheum 2007;57:1254–60.
30. Felson DT, Niu J, Guermazi A, Roemer F, Aliabadi P, Clancy M, *et al*. Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging. Arthritis Rheum 2007;56:2986–92.
31. Woo J, Lau E, Lee P, Kwok T, Lau WC, Chan C, *et al*. Impact of osteoarthritis on quality of life in a Hong Kong Chinese population. J Rheumatol 2004;31:2433–8.
32. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:870–7.
33. Hill CL, Parsons J, Taylor A, Leach G. Health related quality of life in a population sample with arthritis. J Rheumatol 1999;26:2029–35.
34. Muraki S, Akune T, Oka H, En-yo Y, Yoshida M, Saika A, *et al*. Association of radiographic and symptomatic knee osteoarthritis with health-related quality of life in a population-based cohort study in Japan: the ROAD study. Osteoarthritis Cartilage 2010;18:1227–34.
35. Altman RD, Hochberg M, Murphy Jr WA, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3(Suppl A):3–70.
36. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage 2007;15(Suppl A):A1–A56.
37. Felson DT, Nevitt MC, Zhang Y, Aliabadi P, Baumer B, Gale D, *et al*. High prevalence of lateral knee osteoarthritis in Beijing Chinese compared with Framingham caucasian asubjects. Arthritis Rheum 2002;46:1217–22.
38. van der Esch M, Knol DL, Schaffers IC, Reiding DJ, van Schaardenburg D, Knoop J, *et al*. Osteoarthritis of the knee: multicompartmental or compartmental disease? Rheumatology (Oxford) 2014;53:540–6.

PLOS | ONE

RESEARCH ARTICLE

Comparison of the Japanese Orthopaedic Association (JOA) Score and Modified JOA (mJOA) Score for the Assessment of Cervical Myelopathy: A Multicenter Observational Study

So Kato[1]*, Yasushi Oshima[1], Hiroyuki Oka[1], Hirotaka Chikuda[1], Yujiro Takeshita[2], Kota Miyoshi[2], Naohiro Kawamura[3], Kazuhiro Masuda[3], Junichi Kunogi[3], Rentaro Okazaki[4], Seiichi Azuma[4], Nobuhiro Hara[5], Sakae Tanaka[1], Katsushi Takeshita[1]

1 Department of Orthopaedic Surgery, the University of Tokyo, Tokyo, Japan, **2** Department of Orthopaedic Surgery, Yokohama Rosai Hospital, Yokohama, Japan, **3** Department of Spine and Orthopaedic Surgery, Japanese Red Cross Medical Center, Tokyo, Japan, **4** Department of Orthopaedic Surgery, Saitama Red Cross Hospital, Saitama, Japan, **5** Department of Orthopaedic Surgery, Musashino Red Cross Hospital, Musashino, Japan

* skatou-tky@umin.org

Citation: Kato S, Oshima Y, Oka H, Chikuda H, Takeshita Y, Miyoshi K, et al. (2015) Comparison of the Japanese Orthopaedic Association (JOA) Score and Modified JOA (mJOA) Score for the Assessment of Cervical Myelopathy: A Multicenter Observational Study. PLoS ONE 10(4): e0123022. doi:10.1371/journal.pone.0123022

Academic Editor: Michael Fehlings, University of Toronto, CANADA

Received: May 21, 2014

Accepted: February 25, 2015

Published: April 2, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Objectives

The Japanese Orthopaedic Association (JOA) score is widely used to assess the severity of clinical symptoms in patients with cervical compressive myelopathy, particularly in East Asian countries. In contrast, modified versions of the JOA score are currently accepted as the standard tool for assessment in Western countries. The objective of the present study is to compare these scales and clarify their differences and interchangeability and verify their validity by comparing them to other outcome measures.

Materials and Methods

Five institutions participated in this prospective multicenter observational study. The JOA and modified JOA (mJOA) proposed by Benzel were recorded preoperatively and at three months postoperatively in patients with cervical compressive myelopathy who underwent decompression surgery. Patient reported outcome (PRO) measures, including Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ), the Short Form-12 (SF-12) and the Neck Disability Index (NDI), were also recorded. The preoperative JOA score and mJOA score were compared to each other and the PRO values. A Bland-Altman analysis was performed to investigate their limits of agreement.

Results

A total of ninety-two patients were included. The correlation coefficient (Spearman's rho) between the JOA and mJOA was 0.87. In contrast, the correlations between JOA/mJOA

and the other PRO values were moderate (|rho| = 0.03 – 0.51). The correlation coefficient of the recovery rate between the JOA and mJOA was 0.75. The Bland-Altman analyses showed that limits of agreement were 3.6 to -1.2 for the total score, and 55.1% to -68.8% for the recovery rates.

Conclusions

In the present study, the JOA score and the mJOA score showed good correlation with each other in terms of their total scores and recovery rates. Previous studies using the JOA can be interpreted based on the mJOA; however it is not ideal to use them interchangeably. The validity of both scores was demonstrated by comparing these values to the PRO values.

Introduction

Cervical compressive myelopathy is a common disorder that frequently results in impairment of a patient's motor, sensory and bladder function. Several scales that measure severity of physical disability have been developed to assess a patient's pre- and post-treatment condition and the effectiveness of intervention. For example, the Japanese Orthopaedic Association (JOA) score was developed by the JOA in 1975. Since then, it has become one of the most frequently used outcome measures to evaluate functional status in patients with cervical myelopathy. Furthermore, and the concept of "recovery rate," advocated by Hirabayashi et al., has been widely accepted as an outcome measure [1]. Currently, the revised version of the JOA score (1994), which includes an assessment of the shoulder and elbow function, is the most frequently used [2, 3].

One of the drawbacks of the JOA score is that it evaluates the degree of motor dysfunction by assessing a patient's ability to use chopsticks. The use of chopsticks is limited to East Asian cultures including Japanese, Korean, Chinese and Vietnamese populations. The issues associated with using questionnaires related to cultural differences in eating methods have already been reported [4, 5]. Although chopsticks are now more widely used for eating, even in Western cultures, questionnaires using chopsticks cannot be readily applied to those who have not used them, or who do not use them regularly. Therefore, the adaptation of the JOA score to a Western population requires translation as well as modification [6]. Currently, there are three different kinds of so-called "modified JOA (mJOA) scores" [7–9]. However, the translation of these scores has not been validated and the scoring structure and content of evaluation items are substantially different. Despite their differences, the JOA score and the various modified scales are frequently confused with each other, and mistakenly discussed as being the same. Few comparisons of these scales have been made in the literature and few studies have assessed the validity of these scores. This causes confusion about which scale should be used in a certain population, and prevents us from comparing results of studies that used different modifications of the JOA score.

Therefore, it is very important to compare the properties of the JOA score and the mJOA score for the assessment of cervical myelopathy; the JOA score and the mJOA score. The objective of this study is to investigate the differences in and interchangeability of the JOA score and the mJOA score and to examine the validity of these scales by assessing correlations with other patient-reported outcome measures.

Materials and Methods

The study protocol was approved by the institutional review board of the Clinical Research Support Center of the University of Tokyo Hospital. In order to secure a sufficient number of participants, we called for volunteers from our research group, "The University of Tokyo Spine Group," and recruited five medical center institutions to participate in this prospective multicenter observational study. Ten surgeons in total were involved in this study. All eligible patients provided their written informed consent to participate in this study. All patients who underwent surgery for cervical compressive myelopathy between April 2013 and March 2014 were enrolled. Those with systemic diseases, including neurological disorders and rheumatoid arthritis, that could potentially affect motor function were excluded. Preoperatively, the surgeons recorded the following two scores.

JOA score (Table A in S1 File) [2, 3]

We used the latest version of the JOA score in Japanese. This scale consists of six domain scores (motor dysfunction in the upper extremities, motor dysfunction in the lower extremities, sensory function in the upper extremities, sensory function in the trunk, sensory function in the lower extremities, and bladder function), scaled from 0 to 4, 4, 2, 2, 2, and 3, respectively, with the minimum total score being 0 and the maximum total score being 17. Yonenobu et al. defined the myelopathy severity as mild if the JOA score is larger than 13, as moderate if the JOA score ranges from 9 to 13 and as severe if the JOA score is less than 9 [3]. Motor function in the fingers was assessed based on the ability to use chopsticks and button clothing. Keller et al. published the modified version in German [9, 10]. The authors did not mention the use of cutlery, but rather simply used the term "fine motor function" for the assessment of motor function in the upper extremities. The score proposed by Chiles et al. is similar, although the authors mentioned the use of a knife and fork [8]. The recovery rate was calculated according to the following formula (Hirabayashi method): Recovery rate (%) = (postoperative JOA −preoperative JOA) / (17 [full score]—preoperative JOA) × 100 [1].

Modified JOA score (Benzel et al.) (Table B in S1 File) [7]

This scale is the most commonly used among the so-called "mJOA" scores. Its scoring system differs from that of the original JOA in that it assesses only motor dysfunction in the upper and lower extremities, sensory function in the upper extremities, and bladder function, and does not include a scale for sensory function in the trunk and lower extremities. Each scale ranges from 0 to 7, 5, 3, and 3, respectively, with a total score of 0 to 18. Fehlings et al. defined the severity of myelopathy as mild if the mJOA score is 15 or larger, moderate if the mJOA score ranges from 12 to 14 or severe if the mJOA score is less than 12 [11]. This scale focused on the use of a spoon instead of chopsticks to evaluate motor function in the upper extremities. The recovery rate is calculated using the same formula as that applied for the original JOA, changing the full score from 17 to 18.

The differences between these scores are summarized in Table 1. The JOA score allocates 8 out of 17 (47%) points of the total score to motor function, while the mJOA score allocates 12 out of 18 (67%) points of the total score to motor function. These two scores were determined by the responsible surgeons at each institution. In addition to these scores, the Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ) [12], Short Form-12 (SF-12) [13] and the Neck Disability Index (NDI) [14] were recorded as patient reported outcome (PRO) measurements. These scales were completed by the patients in the form of questionnaires. The postoperative scores were recorded whenever possible at follow-up visits performed three months after surgery.

Table 1. A summary of the differences between the JOA score and modified JOA score.

	Structure							Assessment of MU
	MU	ML	SU+	ST	SL	BL	Total	
JOA [3]	4	4	2	2	2	3	17	Chopsticks
Modified JOA [7]	5	7	3	N/A	N/A	3	18	Spoon

JOA: Japanese Orthopaedic Association score, MU: motor function in the upper extremities, ML: motor function in the lower extremities, SU: sensory function in the upper extremities, ST: sensory function in the trunk, SL: sensory function in the lower extremities, BL: bladder function, N/A: not applicable

doi:10.1371/journal.pone.0123022.t001

The preoperative JOA and mJOA scores in each domain were compared with each other. The total scores were also compared with each other and to the PRO measurements. Furthermore, we compared the JOA and mJOA after dichotomizing the patients according to severity of motor function by the median of the JOA motor function scores. A prediction formula for the mJOA score was created using the JOA to enable direct comparisons between studies using these scores by converting the scores. We plotted the individual difference between the mJOA total score and the JOA total score (mJOA–JOA) against the average between the mJOA and JOA scores using a Bland–Altman plot. Bland-Altman analyses are now widely used for comparing two methods of measurement [15–19]. According to Bland and Altman, the limits of agreement can be estimated as the mean between duplicate measurements (the bias) ±1.96 SD, where the SD is the standard deviation of all of the paired differences [20]. This means that 95% of the differences will lie between these limits. Provided that differences within these lines are not clinically important, we could use the two measurement methods interchangeably. Although the minimally clinical important difference (MCID) of the JOA or mJOA has not been established, experts have argued that a difference of at least two points of mJOA is clinically important [21]. Therefore, the limits of agreement below 2 suggests the interchangeability of the two scores in the present study. Finally, among the patients whose postoperative scores at three months were available, the recovery rates for the JOA and mJOA scores were compared, and a Bland-Altman analysis was performed.

All analyses were carried out using the IBM SPSS Statistics Version 19 software package (SPSS, Inc., Somers, NY, USA). Correlations between the scores were analyzed by calculating the Spearman's rank correlation coefficient rho. P-values less than 0.05 were considered to be significant in all statistical tests. We defined the strength of the correlation according to the general guideline (rho $\geq$ 0.70: very strong, $\geq$ 0.50: strong, $\geq$ 0.30: moderate, $\geq$ 0.10: weak) [22].

Results

Ninety-two patients were included in the study. One patient whose bladder function could not be assessed due to anuria resulting from chronic renal failure was excluded. The mean age was 63.3 years (standard deviation: 12.9). The most common diagnosis indicated for surgery was cervical spondylotic myelopathy (58 patients), followed by ossification of the posterior longitudinal ligament (28 patients) and cervical disc herniation (six patients).

Comparisons of the scores in each domain

The correlations between the JOA and mJOA scores in each domain were strong to very strong, with correlation coefficients of 0.84 for motor function in the upper extremities ($p < 0.001$), 0.93 for motor function in the lower extremities ($p < 0.001$), 0.67 for sensory function in the upper extremities ($p < 0.001$) and 0.89 for bladder function ($p < 0.001$). The correlation

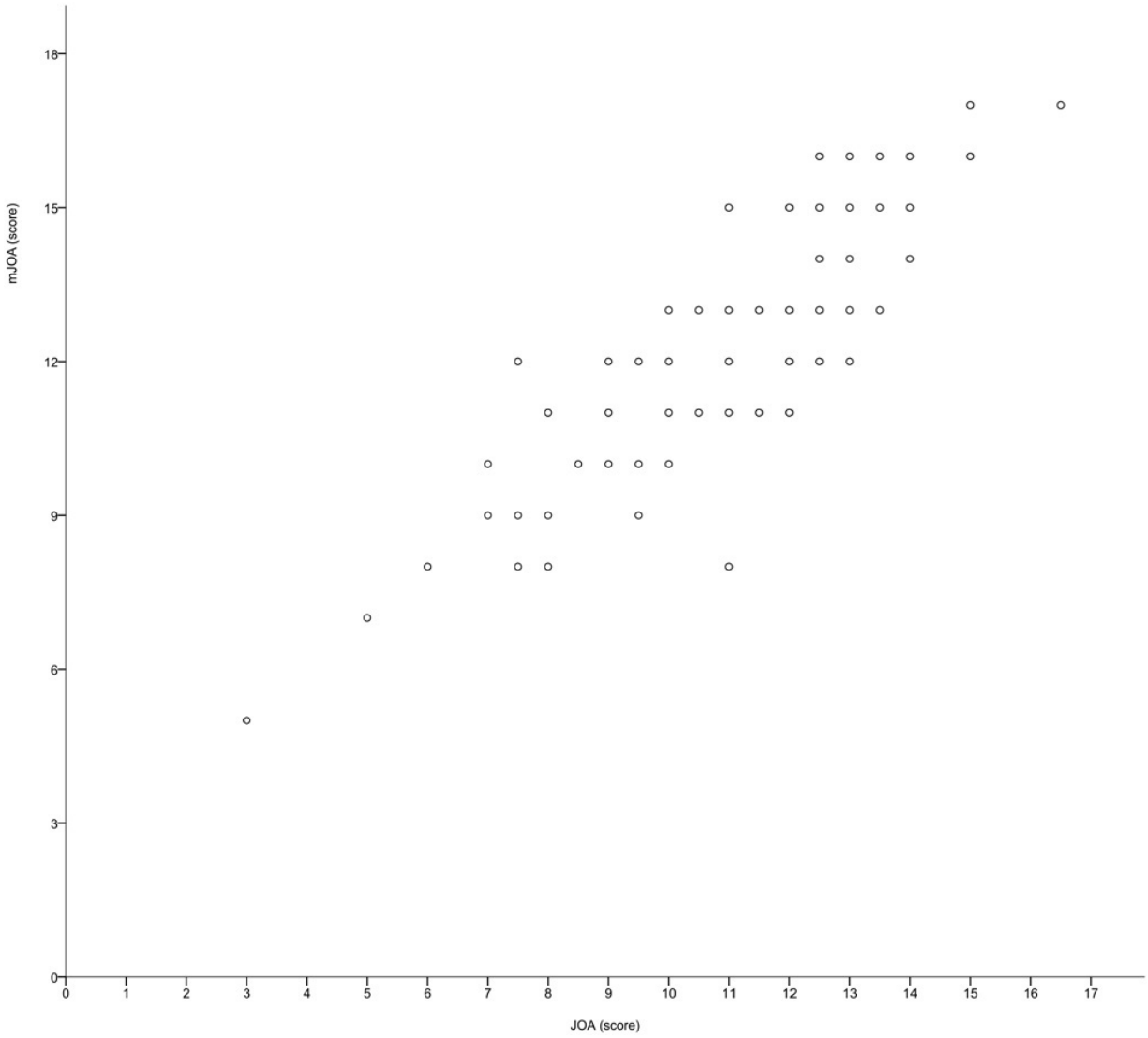


Fig 1. Scatterplot of the total scores for the JOA and mJOA scores (n = 92).

doi:10.1371/journal.pone.0123022.g001

between the total scores for motor function (the sum of the scores for the upper and lower extremities) was also very strong (rho = 0.90, p <0.001).

Total score

The mean preoperative JOA score was 11.2 (range: 3.0–16.5, standard deviation: 2.5), whereas the mean mJOA score was 12.4 (range: 5–17, standard deviation: 2.5). A scatterplot of the JOA and mJOA scores is shown in Fig 1, and the correlations between the preoperative scores are summarized in Table 2. The JOA and mJOA scores were very strongly correlated with each other (rho = 0.87, p <0.001). The median of the JOA motor function scores was 5. The correlation was found to be weaker in those with a motor function score less than 5 (n = 37,

Table 2. Correlations between the preoperative total scores among the JOA, modified JOA, JOACMEQ QOL score, SF-12 PCS, MCS and NDI (n = 92).

	JOA	Modified JOA	JOACMEQ QOL	SF-12 PCS	SF-12 MCS	NDI
JOA	1	0.87*	0.41*	0.50*	-0.05	-0.50*
Modified JOA		1	0.41*	0.47*	0.03	-0.51*
JOACMEQ			1	0.29*	0.40*	-0.66*
SF-12 PCS				1	-0.29*	-0.47*
SF-12 MCS					1	-0.17
NDI						1

* Statistical significance

doi:10.1371/journal.pone.0123022.t002

rho = 0.64) than in those with milder motor dysfunction (n = 55, rho = 0.77). On the other hand, the correlations between the JOA/mJOA scores and the other PRO values were not as strong. JOACMEQ QOL score, SF-12 PCS and NDI showed moderate correlations (|rho|: 0.41–0.51), whereas SF-12 MCS did not (|rho|: 0.03–0.05). While the very strong correlation between the JOA and mJOA scores demonstrates convergent validity, the moderate correlation with other PRO values suggests divergent validity. We created a prediction formula to calculate the total scores for the mJOA from the score of the JOA using linear regression analysis. The result is as follows:

$$\text{mJOA total} = 2.39 + 0.89 \times (\text{JOA total})$$

The R^2 of this equation was 0.78.

A Bland–Altman plot showing the differences between the two scores (mJOA–JOA) plotted against the mean of the two scores is shown in Fig 2. The mean difference between the two scores (the bias) was 1.2 (95% confidence interval: 0.9–1.5, standard deviation: 1.21). The upper and lower limits of agreement were 3.6 and -1.2, respectively. This range was well above the threshold we set based on an assumed MCID [21]; from this result, we were able to conclude that it is not ideal to interchange the JOA and mJOA.

Recovery rate (RR)

In 65 patients (71%) followed at three months postoperatively, the recovery rates were calculated using the Hirabayashi method and compared with each other. The mean JOA recovery rate was 45.1% (range: -33%– 100%, standard deviation: 30.8%), whereas the mean mJOA recovery rate was 38.2% (range: -200%– 100%, standard deviation: 43.0%). A scatterplot of the recovery rates for the JOA and mJOA is shown in Fig 3. In this figure, one outlier whose JOA RR was 0 and mJOA RR was -2.0 (deterioration), was omitted. Their correlations were very strong (rho: 0.75, $p < 0.001$). In two cases, one scale showed recovery while the other showed deterioration. Both of these patients had urinary symptoms. We created a prediction formula to calculate the mJOA RR from the JOA RR using linear regression analysis. The result is as follows:

$$\text{mJOA RR} = -0.05 + 0.95 \times (\text{JOA RR})$$

The R^2 value of this equation was 0.46.

A Bland–Altman plot showing the differences between the two recovery rates plotted against the mean of the two recovery rates is shown in Fig 4. The mean bias was -6.9% (95% confidence interval: -14.7%– 1.0%, standard deviation: 31.6%). The upper and lower limits of

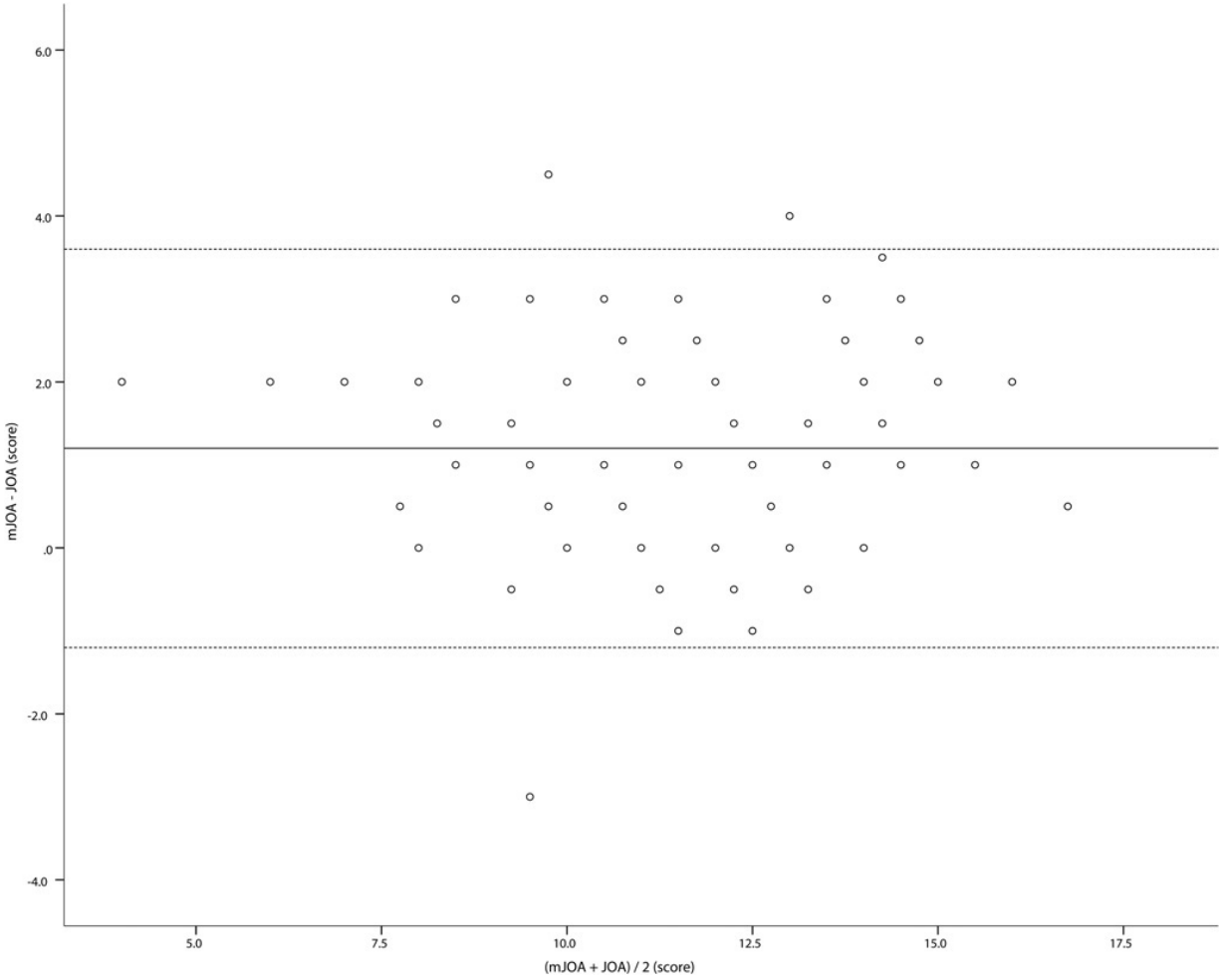


Fig 2. A Bland–Altman plot comparing the JOA and mJOA scores. The bias is shown as a solid line, and the upper and lower limits of agreement are shown as broken lines.

doi:10.1371/journal.pone.0123022.g002

agreement were 55.1% and -68.8%, respectively. This range is also substantial enough to consider that it is not ideal to interchange the recovery rates of the JOA and mJOA.

Discussion

There are two major findings in the present study. First, the domain and total scores of the JOA and mJOA were strongly correlated with each other. In addition, although the total scores and the recovery rates of the mJOA can be accurately predicted by the conversion formulas using the JOA score and its recovery rate, the Bland-Altman analyses showed they are not interchangeable. Second, the validity of the two types of JOA scores was demonstrated in comparisons with the PRO values.

No previous studies have directly compared the JOA and its modifications. The present study showed that the domain scores of the JOA and mJOA are strongly correlated, although the scoring structures of these scales differ in many domains, and the linearity of the scale is

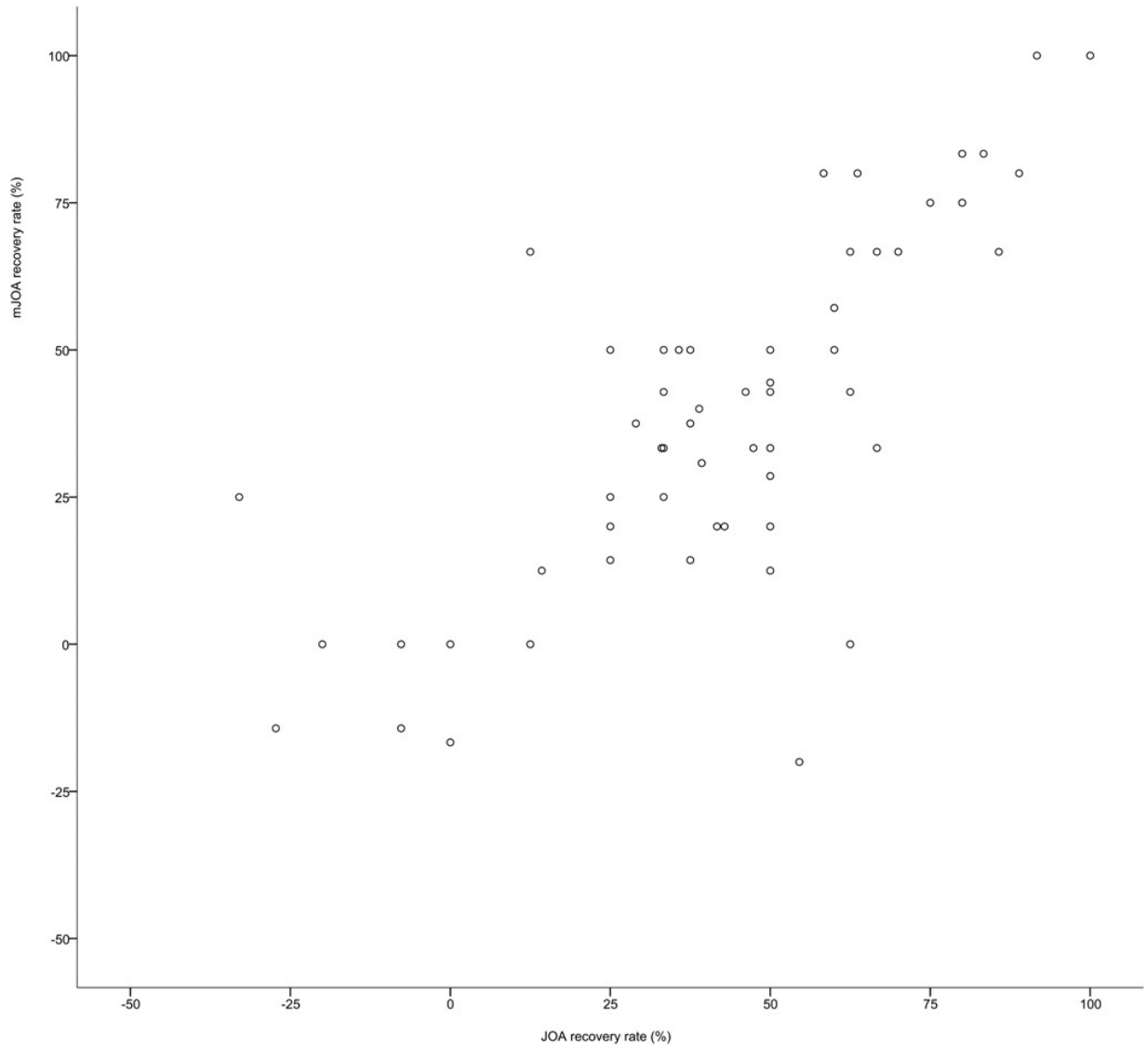


Fig 3. Scatterplot of the recovery rates for the JOA and mJOA scores. This figure includes only cases with a recovery rate from -1.0 to +1.0. Only two outliers were omitted (n = 63).

doi:10.1371/journal.pone.0123022.g003

not guaranteed. It is of note that the mJOA score exhibited a very strong correlation with the JOA score, despite that the mJOA lacks scores for sensory function in the trunk and lower extremities. This finding may be due to the fact that severe sensory disturbances in the trunk or lower extremities are relatively rare in operative candidates for cervical myelopathy. The correlation between the scores for the sensory function in the upper extremities was lower than that for the other domains. This result may be explained by the exaggerated construct differences in which the JOA has two points and the mJOA has three points. The correlation in the subjects with severer motor dysfunction was weaker. This finding is also understandable given that the mJOA score gives a higher proportion to the motor function score. In the two patients with

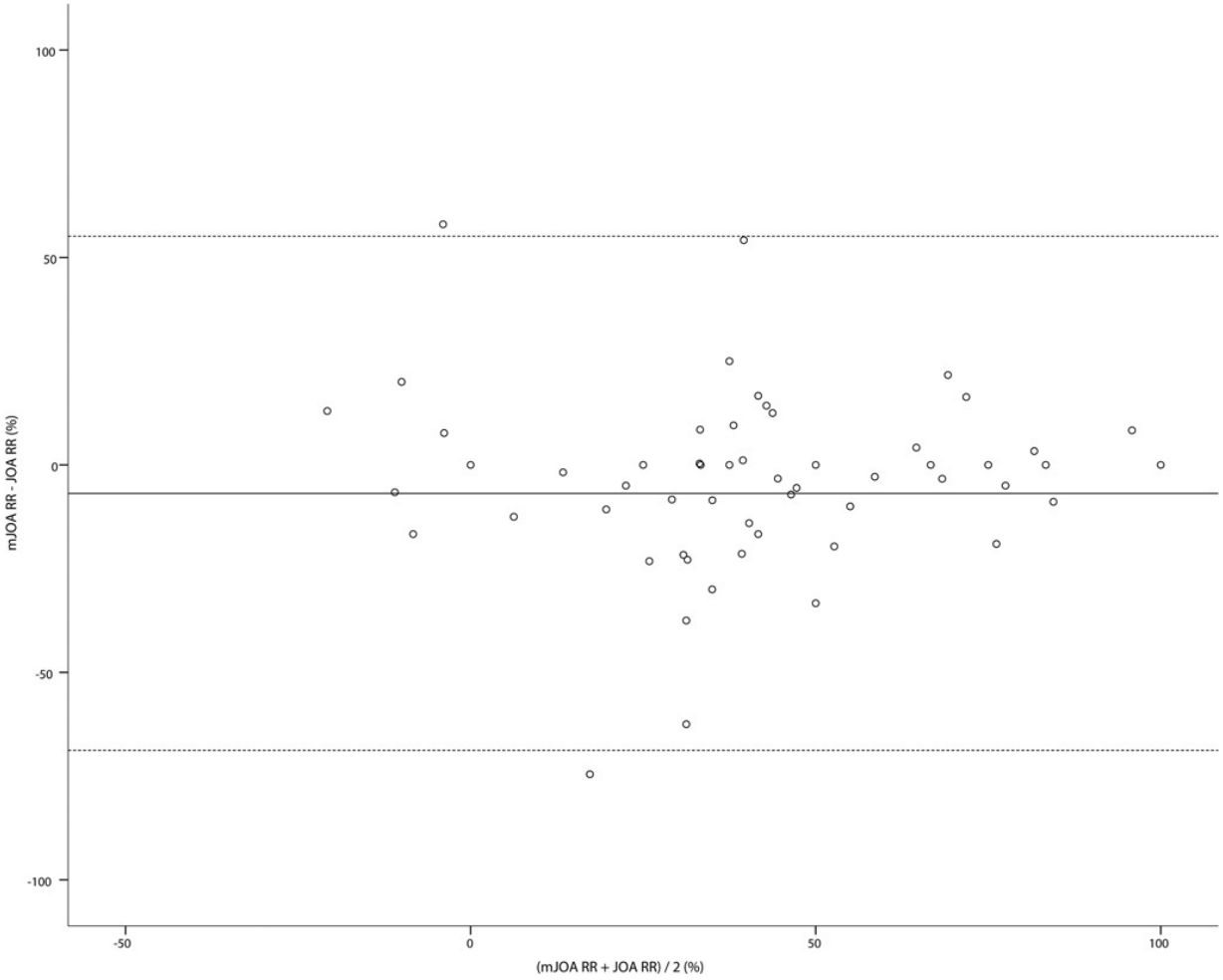


Fig 4. A Bland–Altman plot comparing the JOA and mJOA recovery rates. The bias is shown as a solid line, and the upper and lower limits of agreement are shown as broken lines.

doi:10.1371/journal.pone.0123022.g004

urinary symptoms, the recovery was not properly reflected in one scale. The bladder function score in the JOA tended to be exaggerated because the JOA criteria are more complicated than those of the mJOA. For example, the sense of urinary retention can lead to the patients receiving a score of 1, and this symptom is very common even in the elderly generation without myelopathy. These comparisons did not lead us to conclude that one scale had significant advantages over the other, and any of them can be used as desired based on the patients' cultural background. The mJOA would be more easily accepted for Asian populations, since many of them now use a spoon, than would the JOA for Western populations, although no validated translations in Asian languages exist, and this would be an obstacle for raters who do not understand English. Using our conversion formulas, it is possible to interpret the results of previous studies that used the mJOA according to the original JOA score. For example, if a study set a certain cut-off point to evaluate the effectiveness of a treatment using the recovery rate, we speculate that the evaluation might be slightly stricter when using the mJOA instead of the JOA.

The present study showed that the scores for the mJOA and JOA are strongly correlated; however, Anscombe suggested that data with nearly identical simple statistical properties may appear very different when graphed [23]. A further understanding of the relationship between the mJOA and JOA scores can be achieved by looking at the differences between the two methods plotted against the mean score for each subject. We therefore examined the two seemingly compatible scores using Bland-Altman plots. Although Bland-Altman analyses were originally developed to make comparisons of two methods using the same scale, many authors have since applied this technique to the comparisons of two different scales [16, 19]. The range of the JOA score and mJOA score differ slightly (0–17 vs. 0–18), but very few patients in the present study achieved a nearly full score, which theoretically maximizes the difference between the two scales. Since a Bland-Altman analysis is the best method for visualizing errors and because there are no alternatives, we believe that the application of this method to the present dataset is acceptable. In Fig 2, the error appears unbiased, as differences are spread evenly and randomly above and below zero points. We examined the agreement between these two methods by looking at the spread of differences. The variability between the two methods is reflected by the limits of agreement, which were substantial in the present study. Based on this difference, a patient can easily be categorized into different groups of severity by both the JOA and mJOA.

While the criterion validity of the JOA score has been discussed by comparing it to the results of multiple other scales, including the Cooper myelopathy scale (CMS) [10, 24], European myelopathy scale (EMS) [10, 25] and Short Form-36 (SF-36) PCS [26, 27], few studies have discussed the validation of mJOA based on comparisons of these scores with the PRO values [28]. The mJOA score has been compared with the Nurick grade [29–31], NDI and SF-36 [31]. We measured the concurrent validity by performing comparisons to the JOACMEQ, SF-12 and NDI. In the present study, we used the SF-12 instead of the SF-36 because the summary scores for the SF-12 have been shown to mirror those of the SF-36 [13]. All of these results suggest divergent validity. The PRO forms are completed by the patients, as opposed to the JOA and mJOA, and these scales are substantially affected by the patients quality of life. Meanwhile, the JOA and mJOA are more disease-specific for cervical myelopathy and likely measure a different construct. These results are in accordance with the findings of the study by Kopjar et al. that validated the mJOA score [31].

There are some limitations associated with the present study. First, the rate of follow-up was not as high as expected. Many patients dropped out after the surgery as they were satisfied with their postoperative results. Therefore, the analysis of the recovery rate may have been biased. Second, because the assessment for the JOA score and mJOA were produced in different languages, the translational validity was not verified. Finally, the inter-observer reliability and test-retest reliability were not investigated in the present study. However, the inter- and intra- observer reliability of the JOA is reported to be high [3]. The inter-observer reliability of the mJOA has also been reported to be high [32], although this finding should be interpreted with caution since a translated version of the scale was used in this study. Unfortunately, the test-retest reliability of the mJOA has not yet been established. Further studies may make it possible to compare the properties of these scores.

Conclusion

In conclusion, the mJOA score is very strongly correlated with the JOA, and previous studies using the JOA can be interpreted based on the mJOA based on this speculation, especially by using the conversion formulas advocated in this report. However, the Bland-Altman analysis revealed that it is not ideal to use these scoring systems interchangeably.

Supporting Information

S1 Dataset. Dataset of the outcomes in the participants.
(XLSX)

S1 File. Table A, Japanese Orthopaedic Association Score (English translation) [3].
Table B, Modified Japanese Orthopaedic Association Score [7].
(DOCX)

Author Contributions

Conceived and designed the experiments: SK YO KT. Performed the experiments: SK YO YT Kazuhiro. Masuda NK Kota. Miyoshi JK RO SA NH. Analyzed the data: SK HO. Contributed reagents/materials/analysis tools: SK. Wrote the paper: SK YO HO KT HC ST.

References

1. Hirabayashi K, Miyakawa J, Satomi K, Maruyama T, Wakano K. Operative results and postoperative progression of ossification among patients with ossification of cervical posterior longitudinal ligament. Spine (Phila Pa 1976). 1981; 6:354–64.
2. Japanese Orthopaedic Association. Scoring system for cervical myelopathy. J Jpn Orthop Assoc. 1994; 68:490–503.
3. Yonenobu K, Abumi K, Nagata K, Taketomi E, Ueyama K. Interobserver and intraobserver reliability of the japanese orthopaedic association scoring system for evaluation of cervical compression myelopathy. Spine (Phila Pa 1976). 2001; 26:1890–4; discussion 5.
4. Chen FF. What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. J Pers Soc Psychol. 2008; 95:1005–18. doi: 10.1037/a0013193 PMID: 18954190
5. Roh YH, Yang BK, Noh JH, Baek GH, Song CH, Gong HS. Cross-cultural adaptation and validation of the Korean version of the Michigan hand questionnaire. J Hand Surg Am. 2011; 36:1497–503. doi: 10.1016/j.jhsa.2011.06.006 PMID: 21783329
6. Leonardi M, Boos N. Degenerative Disorders of the Cervical Spine. In: Boos N, Aebi M, editors. Spinal Disorders. New York, NY: Springer-Verlag Berlin Heidelberg; 2008. p. 429–80.
7. Benzel EC, Lancon J, Kesterson L, Hadden T. Cervical laminectomy and dentate ligament section for cervical spondylotic myelopathy. J Spinal Disord. 1991; 4:286–95. PMID: 1802159
8. Chiles BW 3rd, Leonard MA, Choudhri HF, Cooper PR. Cervical spondylotic myelopathy: patterns of neurological deficit and recovery after anterior cervical decompression. Neurosurgery. 1999; 44:762–9; discussion 9–70. PMID: 10201301
9. Keller A, von Ammon K, Klaiber R, Waespe W. [Spondylogenic cervical myelopathy: conservative and surgical therapy]. Schweiz Med Wochenschr. 1993; 123:1682–91. PMID: 8211019
10. Vitzthum HE, Dalitz K. Analysis of five specific scores for cervical spondylogenic myelopathy. Eur Spine J. 2007; 16:2096–103. PMID: 17922150
11. Fehlings MG, Wilson JR, Kopjar B, Yoon ST, Arnold PM, Massicotte EM, et al. Efficacy and safety of surgical decompression in patients with cervical spondylotic myelopathy: results of the AOSpine North America prospective multi-center study. J Bone Joint Surg Am. 2013; 95:1651–8. doi: 10.2106/JBJS.L.00589 PMID: 24048552
12. Fukui M, Chiba K, Kawakami M, Kikuchi S, Konno S, Miyamoto M, et al. An outcome measure for patients with cervical myelopathy: Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ): Part 1. J Orthop Sci. 2007; 12:227–40. PMID: 17530374
13. Ware J Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996; 34:220–33. PMID: 8628042
14. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther. 1991; 14:409–15. PMID: 1834753
15. Ottenbacher KJ, Stull GA. The analysis and interpretation of method comparison studies in rehabilitation research. Am J Phys Med Rehabil. 1993; 72:266–71. PMID: 8398016
16. Madsen OR. Agreement between the DAS28-CRP assessed with 3 and 4 variables in patients with rheumatoid arthritis treated with biological agents in the daily clinic. J Rheumatol. 2013; 40:379–85. doi: 10.3899/jrheum.120594 PMID: 23457377

17. Gasparovic H, Gabelica R, Ostojic Z, Kopjar T, Petricevic M, Ivancan V, et al. Diagnostic accuracy of central venous saturation in estimating mixed venous saturation is proportional to cardiac performance among cardiac surgical patients. J Crit Care. 2014; 29:828–34. doi: 10.1016/j.jcrc.2014.04.012 PMID: 24857639

18. Randleman JB, Akhtar J, Lynn MJ, Ambrosio R Jr, Dupps WJ Jr, Krueger RR, et al. Comparison of objective and subjective refractive surgery screening parameters between regular and high-resolution Scheimpflug imaging devices. J Cataract Refract Surg. 2014.

19. Siemons L, Vonkeman HE, ten Klooster PM, van Riel PL, van de Laar MA. Interchangeability of 28-joint disease activity scores using the erythrocyte sedimentation rate or the C-reactive protein as inflammatory marker. Clin Rheumatol. 2014; 33:783–9. doi: 10.1007/s10067-014-2538-x PMID: 24562719

20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1:307–10. PMID: 2868172

21. Furlan JC, Kalsi-Ryan S, Kailaya-Vasan A, Massicotte EM, Fehlings MG. Functional and clinical outcomes following surgical treatment in patients with cervical spondylotic myelopathy: a prospective study of 81 cases. J Neurosurg Spine. 2011; 14:348–55. doi: 10.3171/2010.10.SPINE091029 PMID: 21235299

22. Kraemer HC, Morgan GA, Leech NL, Gliner JA, Vaske JJ, Harmon RJ. Measures of clinical significance. J Am Acad Child Adolesc Psychiatry. 2003; 42:1524–9. PMID: 14627890

23. Anscombe FJ. Graphs in Statistical Analysis. The American Statistician. 1973; 27:17–21.

24. Cooper PR, Epstein F. Radical resection of intramedullary spinal cord tumors in adults. Recent experience in 29 patients. J Neurosurg. 1985; 63:492–9. PMID: 4032012

25. Herdman J, Linzbach M, Krzan M. The European myelopathy score. In: Bauer B, Brock M, Klinger M, editors. Advances in Neurosurgery. Berlin: Springer; 1994. p. 266–8.

26. King JT Jr, Tsevat J, Moossy JJ, Roberts MS. Preference-based quality of life measurement in patients with cervical spondylotic myelopathy. Spine (Phila Pa 1976). 2004; 29:1271–80. PMID: 15167668

27. Singh A, Crockard HA. Comparison of seven different scales used to quantify severity of cervical spondylotic myelopathy and post-operative improvement. J Outcome Meas. 2001; 5:798–818. PMID: 16320550

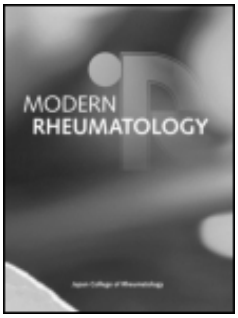
28. Singh A, Tetreault L, Casey A, Laing R, Statham P, Fehlings MG. A summary of assessment tools for patients suffering from cervical spondylotic myelopathy: a systematic review on validity, reliability and responsiveness. Eur Spine J. 2013.

29. Revanappa KK, Rajshekhar V. Comparison of Nurick grading system and modified Japanese Orthopaedic Association scoring system in evaluation of patients with cervical spondylotic myelopathy. Eur Spine J. 2011; 20:1545–51. doi: 10.1007/s00586-011-1773-y PMID: 21424591

30. Nurick S. The pathogenesis of the spinal cord disorder associated with cervical spondylosis. Brain. 1972; 95:87–100. PMID: 5023093

31. Kopjar B, Tetreault L, Kalsi-Ryan S, Fehlings M. Psychometric properties of the modified Japanese orthopaedic association scale in patients with cervical spondylotic myelopathy. Spine (Phila Pa 1976). 2015; 40:E23–8. doi: 10.1097/BRS.0000000000000648 PMID: 25341993

32. Bartels RH, Verbeek AL, Benzel EC, Fehlings MG, Guiot BH. Validation of a translated version of the modified Japanese orthopaedic association score to assess outcomes in cervical spondylotic myelopathy: an approach to globalize outcomes assessment tools. Neurosurgery. 2010; 66:1013–6. doi: 10.1227/01.NEU.0000368391.79314.6F PMID: 20404709



Taylor & Francis
Taylor & Francis Group

Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

Mutual associations among musculoskeletal diseases and metabolic syndrome components: A 3-year follow-up of the ROAD study

Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, Sakae Tanaka, Hiroshi Kawaguchi, Kozo Nakamura & Toru Akune

To cite this article: Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, Sakae Tanaka, Hiroshi Kawaguchi, Kozo Nakamura & Toru Akune (2015) Mutual associations among musculoskeletal diseases and metabolic syndrome components: A 3-year follow-up of the ROAD study, Modern Rheumatology, 25:3, 438-448, DOI: 10.3109/14397595.2014.972607

To link to this article: http://dx.doi.org/10.3109/14397595.2014.972607

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://informahealthcare.com/mor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2015; 25(3): 438–448

DOI: 10.3109/14397595.2014.972607

informa
healthcare

ORIGINAL ARTICLE

Mutual associations among musculoskeletal diseases and metabolic syndrome components: A 3-year follow-up of the ROAD study

Noriko Yoshimura[1], Shigeyuki Muraki[2], Hiroyuki Oka[1], Sakae Tanaka[3], Hiroshi Kawaguchi[4], Kozo Nakamura[5], and Toru Akune[2]

[1]*Department of Joint Disease Research, 22nd Century Medical and Research Center, University of Tokyo, Tokyo, Japan,* [2]*Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, University of Tokyo, Tokyo, Japan,* [3]*Department of Orthopaedic Surgery, Sensory and Motor System Medicine, Graduate School of Medicine, University of Tokyo, Tokyo, Japan,* [4]*JCHO Tokyo Shinjuku Medical Center, Shinjyuku-ku, Tokyo, Japan, and* [5]*National Rehabilitation Center for Persons with Disabilities, Saitama, Japan*

Abstract

Objective. This study aimed to assess the mutual associations between musculoskeletal diseases (knee osteoarthritis [KOA], lumbar spondylosis [LS], osteoporosis [OP]) and metabolic syndrome components (obesity [OB], hypertension [HT], dyslipidemia [DL], impaired glucose tolerance [IGT]).

Methods. Of the 1,690 participants (596 men, 1,094 women) at baseline, 1,384 individuals (81.9%; 466 men, 918 women) had complete data at the first follow-up in 2008. Logistic regression analysis included the occurrence or nonoccurrence of the musculoskeletal diseases or metabolic components as the outcome variable and the remaining musculoskeletal diseases and metabolic components at baseline as explanatory variables, adjusted for age, sex, residential region, smoking, and alcohol consumption.

Results. The risk of KOA occurring increased significantly with HT (odds ratio [OR], 2.57; 95% confidence interval [CI], 1.22–5.42; $p = 0.013$) and IGT (OR, 1.99; 95%CI, 1.07–3.70; $p = 0.029$). The risk of OP occurring at the lumbar spine increased with OP at the femoral neck (OR, 4.21; 95%CI 1.46–12.1; $p = 0.008$), and vice versa (OR, 2.19; 95%CI, 1.01–479; $p = 0.047$). KOA increased the risk of HT (Kellgren–Lawrence [KL] grade = 0, 1 vs. KL = 2: OR, 1.84; 95%CI, 1.09–3.12; $p = 0.024$) and DL (KL = 0, 1 vs. KL ≥ 3: OR, 1.66; 95%CI, 1.05–2.61; $p = 0.029$) occurring. Reciprocal relationships existed between the presence of metabolic components and the occurrence of the other metabolic components.

Conclusion. Mutual relationships existed between the occurrence and presence of musculoskeletal diseases, particularly KOA, and metabolic syndrome components.

Keywords

Incidence, Knee osteoarthritis, Lumbar spondylosis, Metabolic risk factor, Osteoporosis

History

Received 7 July 2014
Accepted 29 September 2014
Published online 7 November 2014

Introduction

The rapid aging of the Japanese society has increased the number of disabled elderly individuals requiring support or long-term care. The leading cause of disability is cardiovascular disease (21.5%), followed by dementia (15.3%), senility (13.7%), osteoarthritis (OA) (10.9%), and falls/osteoporotic fractures (10.2%) [1]. Owing to the similarity between the rate for cardiovascular disease and that for the combined contribution of musculoskeletal diseases (OA and osteoporosis [OP]), prevention for these diseases should be as important as that for cardiovascular disease, in order to reduce disability.

Most patients with cardiovascular disease have multiple risk factors for cardiovascular disease [2]; the presence of a specific combination of these risk factors (obesity [OB], hypertension [HT], dyslipidemia [DL], and impaired glucose tolerance [IGT] [3]) leads to metabolic syndrome—a condition that predisposes the affected individual to cardiovascular disease-associated morbidity and mortality.

It has been suggested that a significant relationship may exist between the components of metabolic syndrome and musculoskeletal diseases. Several studies have reported that OB or increased body mass index (BMI) increases the risk of the onset of knee osteoarthritis (KOA) [4–9]. We have previously confirmed that the presence of HT and IGT are risk factors for the occurrence of KOA, independent of OB [10]. In contrast, there have been few reports regarding a significant relationship between DL and musculoskeletal diseases.

Compared to the number of reports regarding the association between metabolic syndrome components and KOA, there have been relatively few reports regarding the relationship between metabolic syndrome components and lumbar spondylosis (LS), while we have previously reported a significant association between BMI and the incidence of LS [11].

Regarding the metabolic syndrome components and OP, a meta-analysis showed that high BMI is a protective factor for most osteoporotic fracture sites [12]. A review of cardiovascular medications used for HT reported that the medications were

Correspondence to: Noriko Yoshimura, MD, PhD, Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. Phone: +81-3-5800-9178. Fax: +81-3-5800-9179. E-mail: yoshimuran-ort@h.u-tokyo.ac.jp

associated with increased bone mineral density (BMD) and/or a reduction in osteoporotic fractures [13]. A systematic review indicated that type 2 diabetes mellitus increases the risk of osteoporotic fractures [14]. However, the influence of high-density lipoprotein cholesterol (HDL-cho) levels on OP remains controversial [15,16].

Regarding mutual relationships between the presence and occurrence of musculoskeletal diseases, we have previously reported that the presence of OP in women appears to reduce the occurrence of OA at the lumbar spine [17].

However, to the best of our knowledge, there have been no reports regarding mutual relationships between the presence and occurrence of the musculoskeletal diseases (KOA, LS, and OP) and metabolic components (OB, HT, DL, and IGT). The present study aimed to estimate the cumulative incidence of each of the musculoskeletal diseases and metabolic syndrome components and to clarify their mutual causalities and interactions, using a large-scale, population-based cohort entitled the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.

Patients and methods

Participants

The present study included the cohort established in 2005 for the ROAD study. Details of the cohort profile have been reported elsewhere [18,19]. In brief, from 2005 to 2007, we developed a baseline database that included clinical and genetic information on 3,040 Japanese residents (1,061 men, 1,979 women) who were recruited from resident registration listings in 3 communities with different characteristics: 1,350 subjects (465 men, 885 women) from an urban region in Itabashi, Tokyo; 864 (319 men, 545 women) from a mountainous region in Hidakagawa, Wakayama; and 826 (277 men, 549 women) from a coastal region in Taiji, Wakayama.

Of the total sample, the BMD measurement was performed only on subjects from the mountainous and coastal regions, who also underwent blood and urinary examinations. Therefore, the present study included data of those 1,690 subjects (596 men, 1,094 women).

All study participants provided written informed consent, and the study was approved by the ethics committees of the University of Wakayama Medical University (No. 373) and the University of Tokyo (Nos. 1264 and 1326).

Procedures in the baseline study

1) Questionnaire survey

All participants completed an interviewer-administered questionnaire modified from the Osteoporotic Fractures in Men Study [20] that contained 400 items. Lifestyle-related questions were asked regarding smoking habits, alcohol consumption, family medical history, physical activity, reproductive variables, health-related quality of life, and prescription medication use.

2) Radiographic assessment

Participants also underwent radiographic examination of both knees and the lumbar spine using anteroposterior and lateral views with weight-bearing and foot-map positioning. The radiographic severity of OA was determined according to the Kellgren–Lawrence (KL) grade: KL0, normal; KL1, slight osteophytes; KL2, definite osteophytes; KL3, joint or intervertebral space narrowing with large osteophytes; and KL4, bone sclerosis, joint or intervertebral space narrowing, and large osteophytes [21].

Radiographs of each site (vertebrae or knees) were examined by a single experienced orthopedic surgeon (SM) who was masked to the clinical status. In the present study, the maximum grade, diagnosed in at least one intervertebral level of the lumbar spine or at least one knee joint, was regarded as the subjects' KL grades. A participant with a KL grade ≥ 2 was defined as having radiographic OA.

3) Measurement of BMD

BMD was measured at the lumbar spine (L2–4) and proximal femur using dual-energy X-ray absorptiometry (DXA) (Hologic Discovery; Hologic, Waltham, MA, USA) on the same DXA equipment, and the same spine phantom was scanned daily to monitor the machine's performance in study populations from different regions. BMD of the phantom was adjusted to 1.032 (0.016) g/cm^2 ($\pm$ 1.5%) during all examinations.

OP was defined based on the World Health Organization (WHO) criteria, in which OP is diagnosed when BMD T-scores are lower than peak bone mass by $-$ 2.5 SDs [22]. The mean BMD of L2–4, measured using the Hologic DXA in Japan, was reported to be 1.011 (0.119) g/cm^2 for both young adult men and women. Therefore, we defined OP of the lumbar spine as L2–4 BMD < 0.714 g/cm^2. Further, the mean BMD of the femoral neck in young adult men and women was reported to be 0.863 (0.127) g/cm^2 and 0.787 (0.109) g/cm^2, respectively [23]. Therefore, we defined OP of the femoral neck as BMD < 0.546 g/cm^2 (men) and < 0.515 g/cm^2 (women).

4) Measurement of metabolic syndrome components

An experienced public health nurse measured systolic and diastolic blood pressures (BPs) using a mercury sphygmomanometer in addition to the anthropometric measurements (height, weight, and BMI). Hemoglobin A1c (HbA1c), blood glucose, HDL-cho, total cholesterol, and triglyceride levels were also measured between 09:00 and 15:00. All analyses were performed at the same laboratory within 24 h of extraction (Osaka Kessei Research Laboratories, Inc., Osaka, Japan).

In this study, definitions of metabolic syndrome components were based on criteria defined by the Examination Committee of Criteria for Metabolic Syndrome in Japan [24] and the Japan Society for the Study of Obesity [25]. However, because not all blood samples were obtained under fasting conditions, we also used indices from the National Health and Nutrition Survey in Japan, which were adopted as metabolic syndrome criteria in a previous national screening study due to the difficulty in collecting samples under fasting conditions [26]. The following definitions were used: OB, BMI > 27.5 kg/m [2,27]; HT, systolic BP ≥ 130 mmHg and/or diastolic BP ≥ 85 mmHg; DL, serum HDL-cho level < 40 mg/dL; and IGT, serum HbA1c level $\geq 5.5\%$ (measured according to the Japan Diabetes Society definition, and $\geq 5.9\%$ defined by the National Glycohemoglobin Standardization Program [28]). Furthermore, subjects taking medications for HT, DL, or diabetes mellitus were regarded as having HT, DL, or IGT, respectively.

Three-year follow-up

From 2008 to 2010, the 1,690 participants were invited to attend a 3-year follow-up of the second ROAD survey, involving repetition of the baseline examinations. The same experienced orthopedic surgeon (SM), while masked to the participants' clinical status, evaluated the knee and spine radiographs and categorized them using the KL grading scale.

A new case of OA of each joint was defined as a baseline KL grade < 2 for joints and a follow-up grade ≥ 2 for at least 1 joint. A new case of OP was diagnosed when an individual who did not show OP at baseline was determined to have OP at follow-up.

New cases of OB, HT, DL, and IGT were determined in a similar manner: individuals lacking the respective status at baseline but exhibiting the status at follow-up. Furthermore, subjects being treated with medication for HT, DL, or diabetes mellitus at baseline and follow-up were regarded as having HT, DL, or IGT, respectively.

Statistical analyses

Statistical analyses were performed using STATA statistical software (STATA Corp, College Station, TX). Differences in proportions were compared using the Chi-square test. Differences in continuous variables were tested for significance using analysis of variance for multiple groups or Scheffé's least significant difference test for pairs of groups. All p values and 95% confidence intervals (CIs) of two-sided analysis are presented.

To test the associations between musculoskeletal diseases and metabolic syndrome components, we performed multivariate logistic regression analyses. Occurrence or nonoccurrence of KOA over the 3 years (0, nonoccurrence; 1, occurrence) was used as the outcome variable, and the presence of LS classified using the KL grade (0, KL grades 0 and 1; 1, KL grade 2; 2, KL grades 3 and 4), OP at L2–4, OB, HT, DL, and IGT at baseline were used as explanatory variables after adjusting for age, sex, regional differences, smoking, and alcohol consumption. Regarding LS, occurrence over the 3 years (0, absence; 1, presence) was used as the outcome variable, and the presence of KOA classified by the KL grade, OP at L2–4, OB, HT, DL, and IGT at baseline were used as explanatory variables after adjusting for the variables mentioned above. OP at L2–4 was then replaced with OP of the femoral neck, and the analyses were run again. Regarding OP of the lumbar spine L2–4, occurrence was used as the outcome variable, and KOA, LS, OP of the femoral neck, OB, HT, DL, and IGT at baseline were used as explanatory variables after adjusting for the variables mentioned above. Regarding OP of the femoral neck, occurrence was analyzed separately as the outcome variable, and KOA, LS, OP at L2–4, OB, HT, DL, and IGT at baseline were used as explanatory variables after adjusting for the variables mentioned above.

Next, the occurrence of each metabolic component (OB, HT, DL, and IGT) was analyzed as the outcome variable, while KOA, LS, OP of L2–4 or the femoral neck, and the metabolic components not used as the outcome variable at baseline were used as explanatory variables after adjusting for age, sex, residing region, smoking, and alcohol consumption. Following the multivariate logistic regression analyses, the odds ratios (ORs) were evaluated.

Results

Eligible participants

Of the 1,690 participants in the baseline survey that was performed in mountainous and coastal regions, we analyzed the data of the 1,384 subjects (81.9%; 466 men, 918 women) who completed all of the examinations at both baseline and follow-up.

Table 1 shows the baseline characteristics of the subjects who participated in both the baseline and second surveys. Regarding the musculoskeletal diseases, the prevalence of each of radiographic

Table 1. Baseline characteristics of the subjects who participated in both baseline and second surveys in the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.

	Total $n = 1{,}384$	Men $n = 466$	Women $n = 918$	p value (men vs. women)
Age, n (%)				
≤ 39 y	39 (2.8)	10 (2.1)	29 (3.2)	0.23
40–49 y	135 (9.8)	40 (8.6)	95 (10.3)	
50–59 y	298 (21.5)	99 (21.2)	199 (21.7)	
60–69 y	413 (29.8)	131 (28.1)	282 (30.7)	
70–79 y	404 (29.2)	155 (33.3)	249 (27.1)	
≥ 80 y	95 (6.9)	31 (6.7)	64 (7.0)	
Total	1,384 (100.0)	466 (100.0)	918 (100.0)	
Selected characteristics, mean (SD)				
Age, y	63.9 (11.8)	64.9 (11.6)	63.4 (11.9)	0.0246*
Height, cm	155.6 (9.0)	164.0 (7.0)	151.3 (6.7)	< 0.001***
Weight, kg	56.0 (10.7)	62.1 (10.7)	52.5 (8.7)	< 0.001***
Prevalence of selected characteristics, %				
Residing in a coastal area	54.1	51.9	55.2	0.245
Current smoking habit (more than once a month)	12.3	29.4	3.5	< 0.001***
Current alcohol consumption (more than once a month)	40.6	68.2	26.6	< 0.001***
KL grade and BMD, mean (SD)				
KL grade of the knee (worst site)	1.48 (1.12)	1.24 (1.09)	1.61 (1.12)	< 0.0001***
KL grade of the knee (worst site)	2.12 (1.11)	2.29 (1.01)	2.03 (1.14)	0.002**
BMD of the lumbar spine (L2–4)	0.94 (0.20)	1.05 (0.19)	0.88 (0.18)	< 0.0001***
BMD of the femoral neck	0.68 (0.14)	0.75 (0.13)	0.64 (0.13)	< 0.0001***
Prevalence of musculoskeletal diseases, %				
Knee osteoarthritis (KL grade ≥ 2)	46.8	37.3	51.6	< 0.001***
Lumbar spondylosis (KL grade ≥ 2)	61.6	76.2	54.3	< 0.001***
Osteoporosis of the lumbar spine L2–4 (WHO criteria)	12.0	2.2	17.0	< 0.001***
Osteoporosis of the femoral neck (WHO criteria)	10.8	2.6	15.0	< 0.001***
Metabolic syndrome components, mean (SD)				
BMI, kg/m^2	23.1 (3.4)	23.4 (3.2)	22.9 (3.4)	0.0089**
Systolic blood pressure, mmHg	134.1 (20.4)	136.6 (18.3)	132.9 (21.4)	0.0015**
Diastolic blood pressure, mmHg	74.2 (11.4)	77.0 (11.5)	72.8 (11.0)	< 0.0001***
Serum HDL-cho level, mg/dL	61.2 (15.9)	55.8 (16.1)	64.0 (15.0)	< 0.0001***
Serum HbA1c level (Japan Diabetes Society), %	5.19 (0.73)	5.23 (0.85)	5.17 (0.67)	0.19
Prevalence of metabolic syndrome components, %				
Obesity	10.4	11.6	9.8	0.307
Hypertension	67.2	72.7	64.4	0.002**
Dyslipidemia	13.0	15.2	11.9	0.079
Impaired glucose tolerance	21.1	24.7	19.3	0.020*

BMD bone mineral density, *BMI* body mass index, *HbA1C* hemoglobin A1c, *HDL-cho* high-density lipoprotein cholesterol, *KL* Kellgren-Lawrence, *WHO* World Health Organization

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

KOA, OP of L2–4, and OP of the femoral neck was significantly higher in women than in men ($p < 0.001$). In contrast, the prevalence of radiographic LS was significantly higher in men than in women ($p < 0.001$). Regarding the metabolic syndrome components, the prevalence of HT and IGT was significantly higher in men than in women ($p < 0.05$).

Cumulative incidence of each musculoskeletal disease and comparison of the baseline metabolic syndrome component characteristics

Because we previously reported the age-sex distribution of the cumulative incidence of musculoskeletal diseases of the same subjects as in the present study [9,11], we report only the cumulative incidence for the total population at risk for KOA, LS, OP at L2–4, and OP at the femoral neck: 3.25%/year, 11.4%/year, 0.76%/year, and 1.83%/year, respectively.

Table 2 compares the baseline characteristics between the occurrence and nonoccurrence of the musculoskeletal diseases. The prevalence of each of LS, HT, and IGT at baseline was significantly higher with the occurrence of KOA than without (LS, $p = 0.009$; HT, $p < 0.001$; IGT, $p < 0.001$). The proportions of KOA and HT at baseline were significantly higher with the occurrence of LS than without (KOA, $p = 0.001$; HT, $p = 0.003$). The prevalence of OP of the femoral neck at baseline was significantly higher with the occurrence of OP of L2–4 than without ($p < 0.001$). Similarly, the prevalence of OP of L2–4 at baseline was significantly higher with the occurrence of OP of the femoral neck than without ($p < 0.001$). In addition, the prevalence of each of OB and HT at baseline was significantly lower with the occurrence of OP of the femoral neck than without (OB, $p = 0.031$; HT, $p = 0.009$).

Cumulative incidence of each metabolic syndrome component and comparison of baseline musculoskeletal disease characteristics

The cumulative incidence of each of OB, HT, DL, and IGT was 1.21%/year, 15.8%/year, 6.53%/year, and 4.56%/year, respectively. Table 3 compares the baseline characteristics between the occurrence and nonoccurrence of the above-mentioned metabolic components. The prevalence of DL was significantly higher with the occurrence of OB than without ($p = 0.013$), while the prevalence of OP at the femoral neck was significantly lower with the occurrence of OB than without ($p = 0.013$). The prevalence of each of KOA, OP of the lumbar spine L2–4, and OB was significantly higher with the occurrence of HT than without (KOA, $p < 0.001$; OP at L2–4, $p = 0.038$; OB, $p = 0.045$). Similarly, the prevalence of each of KOA, OP of the femoral neck, HT, and IGT were significantly higher with the occurrence of DL than without (KOA, $p = 0.005$; OP at femoral neck, $p = 0.003$; HT, $p = 0.008$; IGT, $p = 0.001$). The prevalence of OB was significantly higher with the occurrence of IGT than without ($p = 0.010$).

Mutual associations among musculoskeletal diseases and metabolic syndrome components during the 3-year follow-up

Regarding the musculoskeletal diseases, the risk of the occurrence of KOA increased significantly with the presence of HT ($p = 0.013$) and IGT ($p = 0.029$) (Table 4). Although the other metabolic syndrome components did not significantly influence the occurrence of musculoskeletal diseases, the presence of OB tended to increase the risk of the occurrence of LS ($p = 0.053$). The risk of the occurrence of OP at L2–4 was increased by the presence of OP at the femoral neck ($p = 0.008$), and the risk of the occurrence of OP at the femoral neck was increased by the presence of OP at L2–4 ($p = 0.047$) (Table 4). The presence of OP at lumbar L2–4 tended to decrease the risk of the occurrence of KOA, and the presence of LS tended to decrease the risk of the occurrence of OP at the femoral neck, although both of these associations were not statistically significant.

Table 5 shows the ORs associated with the presence of musculoskeletal diseases and metabolic syndrome components for the occurrence of the metabolic syndrome components. The presence of KOA increased the risk of the occurrence of HT and DL (HT, $p = 0.024$ and DL, $p = 0.029$); however, the other musculoskeletal diseases did not significantly influence the occurrence of the other metabolic syndrome components. In contrast, there was a mutual association between the occurrence of the metabolic syndrome components and the presence of the other metabolic syndrome components. The presence of HT significantly increased the risk of the occurrence of OB ($p = 0.007$), and, inversely, the presence of OB significantly increased the risk of the occurrence of HT ($p = 0.016$). The presence of OB also increased the risk of the occurrence of IGT ($p = 0.023$). Finally, the presence of IGT increased the risk of the occurrence of DL ($p = 0.004$).

Finally, Figure 1 summarizes the above-mentioned mutual associations between each factor for the musculoskeletal diseases and metabolic syndrome components (Figure 1).

Discussion

In the present study, using the data from a population-based cohort with a high participation rate (81.9%) over a period of 3 years, we observed that the risk of the occurrence of KOA was increased by the presence of HT and IGT, and, inversely, the presence of KOA increased the risk of the occurrence of HT and DL. The occurrence of LS tended to be influenced by the presence of OB. A reciprocal relationship existed between OP of the lumbar spine and OP of the femoral neck. Within the metabolic components, the presence of IGT increased the risk of the occurrence of DL, and there was a reciprocal relationship between OB and HT in addition to that between OB and IGT. However, the risk of OP occurring was not influenced by the presence of the metabolic syndrome components. Further, there was no significant association between HT and DL.

The influence of the metabolic components was the strongest with KOA, followed by LS; however, they did not appear to influence OP. Therefore, the results of the present study might indicate that the metabolic syndrome components influence the onset of osteoproliferative changes, such as those with OA, but not bone loss, such as that in OP.

The present findings regarding the relationship between the occurrence of KOA and the presence of HT or IGT support those that we previously reported [10]. At the same time, the risk of the occurrence of HT and DL was increased by the presence of KOA, independent of the presence of OB. Using the same individuals' data, confounding factors, such as bicycling, regular exercise, and a history of knee injuries, were all significantly associated with KOA in a previous study [29]; therefore, these factors could have influenced the relationships in the present study. With the addition of variables for regular cycling, regular exercise, and the history of injury to either knee at baseline to the logistic regression analysis in the present study, the adjusted ORs for HT and IGT, with the outcome as the occurrence of KOA, remained unchanged.

In addition, OB was not a significant risk factor for the occurrence of KOA in the present study. However, when the variable OB was replaced with the actual BMI values in the same model, the logistic regression analysis showed that greater BMIs significantly increased the risk of the occurrence of KOA (lumbar spine OP as an explanatory factor for OP, BMI + 1 kg/m^2; OR, 1.15; 95% CI, 1.05–4.73; $p = 0.038$; femoral neck OP as an explanatory factor for OP, BMI + 1 kg/m^2; OR, 1.15; 95% CI, 1.04–1.27; $p = 0.005$), indicating that the occurrence of KOA is associated with greater body composition, although the effect of OB, as defined by BMI,

Table 2. Comparison of baseline characteristics between the occurrence or nonoccurrence of each musculoskeletal disease during the 3-year follow-up period.

	Knee osteoarthritis (population at risk = 728)			Lumbar spondylosis (population at risk = 530)			Osteoporosis of lumbar spine (L2–4) (population at risk = 1,179)			Osteoporosis of femoral neck (population at risk = 1,187)		
	Occurrence ($n = 71$)	Nonoccurrence ($n = 657$)	p value (occurrence vs. nonoccurrence)	Occurrence ($n = 182$)	Nonoccurrence ($n = 348$)	p value (occurrence vs. nonoccurrence)	Occurrence ($n = 27$)	Nonoccurrence ($n = 1,152$)	p value (occurrence vs. nonoccurrence)	Occurrence ($n = 65$)	Nonoccurrence ($n = 1,122$)	p value (occurrence vs nonoccurrence)
Selected characteristics, mean (SD)												
Age, y	67.3 (8.2)	58.2 (11.8)	<0.0001***	63.3 (10.8)	56.8 (12.5)	<0.0001***	66.8 (8.9)	62.4 (11.8)	0.0584+	70.2 (9.0)	61.9 (11.5)	<0.0001***
Height, cm	153.9 (7.6)	158.8 (8.6)	<0.0001***	154.3 (9.2)	155.2 (7.9)	0.2573	151.9 (7.8)	157.0 (8.6)	0.0022**	151.4 (6.7)	157.2 (8.7)	<0.0001***
Weight, kg	56.0 (8.8)	56.8 (11.0)	0.556	54.9 (9.7)	53.6 (9.5)	0.1477	50.6 (7.4)	57.7 (10.3)	0.0004***	49.0 (6.4)	58.1 (10.3)	<0.0001***
BMI, kg/m^2	23.6 (2.9)	22.4 (3.2)	0.0035**	23.0 (3.1)	22.2 (3.3)	0.0107*	22.0 (3.0)	23.4 (3.3)	0.0312*	21.5 (3.2)	23.5 (3.3)	<0.0001***
Frequency of selected characteristics, %												
Female sex	74.7	58.6	0.009**	71.4	83.1	0.002*	85.2	61.0	0.011*	84.6	60.6	<0.001***
Residing in a coastal area	56.3	70.8	0.012*	42.3	61.5	<0.001***	48.2	56.4	0.392	52.3	56.4	0.516
Current smoking habit (more than once a month)	7.1	16.9	0.034*	14.4	9.8	0.118	3.9	13.7	0.147	5.0	13.8	0.051+
Current alcohol consumption (more than once a month)	35.2	47.9	0.041*	38.5	39.3	0.850	40.7	44.3	0.714	20.3	45.0	<0.001***
Prevalence of musculoskeletal diseases,%												
Knee osteoarthritis	–	–	–	44.8	30.5	0.001**	59.3	44.6	0.130	52.3	43.3	0.154
Lumbar spondylosis	67.6	51.3	0.009**	–	–	–	55.6	62.9	0.439	58.5	61.4	0.635
Osteoporosis of lumbar spine (L2–4)	14.1	10.2	0.312	18.7	14.9	0.268	–	–	–	26.2	4.9	<0.001***
Osteoporosis of femoral neck	11.3	7.2	0.213	14.3	8.9	0.058+	25.9	4.8	<0.001***	–	–	–
Prevalence metabolic syndrome components, %												
Obesity	9.9	7.8	0.535	11.0	6.6	0.080+	3.7	11.6	0.204	3.1	11.9	0.031*
Hypertension	85.7	55.6	<0.001***	63.2	49.4	0.003**	77.8	66.1	0.204	81.3	65.4	0.009**
Dyslipidemia	15.5	11.6	0.333	13.2	9.8	0.232	11.1	13.7	0.697	12.3	13.6	0.776
Impaired glucose tolerance	35.2	15.8	<0.001***	19.2	16.1	0.363	14.8	21.4	0.406	23.1	20.9	0.669

BMI body mass index.
"+$p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$".

DOI 10.3109/14397595.2014.972607

Table 3. Comparison of baseline characteristics between the occurrence or nonoccurrence of each metabolic syndrome component during the 3-year follow-up period.

	Obesity (population at risk = 1,239)			Hypertension (population at risk = 456)			Dyslipidemia (population at risk = 1,204)			Impaired glucose tolerance (population at risk = 1,092)		
	Occurrence ($n = 45$)	Nonoccurrence ($n = 1,194$)	p value (occurrence vs nonoccurrence)	Occurrence ($n = 216$)	Nonoccurrence ($n = 240$)	p value (occurrence vs nonoccurrence)	Occurrence ($n = 236$)	Nonoccurrence ($n = 968$)	p value (occurrence vs nonoccurrence)	Occurrence ($n = 146$)	Nonoccurrence ($n = 946$)	p value (occurrence vs nonoccurrence)
Selected characteristics, mean (SD)												
Age, y	58.8 (12.5)	64.3 (11.9)	0.0024**	60.1 (11.9)	55.9 (12.8)	0.0003***	66.2 (9.5)	63.0 (12.4)	0.0002***	65.4 (10.5)	62.5 (12.3)	0.0061**
Height, cm	155.5 (7.5)	155.5 (8.3)	0.9578	156.2 (9.1)	156.9 (7.9)	0.3902	155.7 (9.1)	154.0 (8.6)	0.0094**	154.3 (9.3)	156.0 (8.9)	0.0289*
Weight, kg	63.9 (6.8)	53.8 (8.8)	<0.0001***	54.7 (10.6)	53.1 (8.5)	0.0782+	56.5 (10.4)	55.2 (10.4)	0.1029	56.2 (12.3)	55.2 (10.2)	0.3097
BMI, kg/m^2	26.4 (1.0)	22.2 (2.5)	<0.0001***	22.3 (3.3)	21.5 (2.6)	0.0039**	23.7 (3.2)	22.7 (3.3)	<0.0001***	23.4 (3.7)	22.6 (3.1)	0.0041**
Frequency of selected characteristics, %												
Female sex	75.6	66.4	0.201	69.4	74.2	0.262	74.2	65.5	0.011*	72.6	67.1	0.187
Residing in a coastal area	82.2	52.3	<0.001***	61.1	63.3	0.625	60.6	50.7	0.006**	41.8	55.5	0.002**
Current smoking habit (more than once a month)	9.1	12.2	0.533	16.4	11.4	0.126	7.4	13.4	0.012*	10.7	12.7	0.509
Current alcohol consumption (more than once a month)	33.3	40.2	0.355	36.0	42.5	0.156	28.8	44.3	<0.001***	35.2	44.0	0.046*
Prevalence of musculoskeletal diseases, %												
Knee osteoarthritis	33.3	45.8	0.098+	41.7	25.8	<0.001***	54.5	44.2	0.005**	51.4	43.7	0.082+
Lumbar spondylosis	57.8	60.7	0.692	50.9	42.1	0.059+	64.8	59.7	0.149	61.0	59.5	0.740
Osteoporosis of lumbar spine (L2–4)	4.4	13.2	0.086+	14.0	7.9	0.038*	16.2	11.6	0.056+	13.1	12.3	0.778
Osteoporosis of femoral neck	0.0	12.3	0.013*	13.0	7.5	0.051+	16.7	9.9	0.003**	11.7	10.6	0.685
Prevalence of metabolic syndrome components, %												
Obesity	–	–	–	6.9	2.9	0.045*	12.3	9.3	0.169	13.7	7.4	0.010*
Hypertension	75.6	64.7	0.133	–	–	–	73.2	64.1	0.008**	63.7	63.1	0.888
Dyslipidemia	22.2	12.1	0.045*	11.1	7.9	0.244	–	–	–	11.6	10.4	0.638
Impaired glucose tolerance	28.9	18.8	0.093+	13.4	9.6	0.197	26.7	16.9	0.001**	–	–	–

BMI body mass index.
$+p < 0.1$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Table 4. Results of the logistic regression analyses for the occurrence of musculoskeletal diseases over a 3-year period.

		Occurrence of knee osteoarthritis			Occurrence of lumbar spondylosis			Occurrence of osteoporosis of the lumbar spine (L2–4)			Occurrence of osteoporosis of the femoral neck		
Explanatory variables	Reference (at baseline)	OR	95% CI	p value	OR	95% CI	p value	OR	95% CI	p value	OR	95% CI	p value
Musculoskeletal diseases													
Knee osteoarthritis	0: KL 0,1 vs 1: KL = 2	–	–	–	1.11	0.67–1.85	0.676	1.92	0.72–5.10	0.189	0.65	0.32–1.34	0.243
	0: KL 0,1 vs 2: KL ≥ 3	–	–	–	1.54	0.74–3.22	0.246	0.70	0.16–3.04	0.637	0.52	0.22–1.23	0.140
Lumbar spondylosis	0: KL 0,1 vs 1: KL = 2	1.25	0.61–2.58	0.546	–	–	–	0.73	0.24–2.15	0.563	0.82	0.38–1.79	0.616
	0: KL 0,1 vs 2: KL ≥ 3	1.35	0.69–2.64	0.382	–	–	–	0.52	0.19–1.44	0.209	0.54	0.27–1.08	0.086+
Osteoporosis of the lumbar spine (L2–4)	0: absent, 1: present	0.46	0.19–1.10	0.083+	0.78	0.43–1.39	0.395	–	–	–	2.19	1.01–4.79	0.047*
Osteoporosis of the femoral neck	0: absent, 1: present							4.21	1.46–12.1	0.008**	–	–	–
Metabolic syndrome components													
Obesity	0: BMI ≤ 27.5, 1: BMI > 27.5	1.59	0.59–4.30	0.357	1.97	0.99–3.92	0.053+	0.48	0.06–3.78	0.486	0.32	0.07–1.43	0.134
Hypertension	0: absent, 1: present	2.57	1.22–5.42	0.013*	1.20	0.79–1.82	0.395	1.81	0.67–4.87	0.243	1.36	0.66–2.79	0.400
Dyslipidemia	0: absent, 1: present	1.14	0.52–2.47	0.746	1.31	0.72–2.41	0.378	0.84	0.24–2.99	0.790	0.75	0.33–1.73	0.504
Impaired glucose tolerance	0: absent, 1: present	1.99	1.07–3.70	0.029*	0.73	0.43–1.24	0.243	0.65	0.21–2.03	0.456	0.90	0.45–1.80	0.773
Adjusted for:													
Age, years	+ 1 y	1.08	1.05–1.12	<0.001***	1.04	1.02–1.06	<0.001***	1.03	0.98–1.08	0.257	1.12	1.07–1.16	<0.001***
Sex	0: men, 1: women	4.73	2.22–10.1	<0.001***	0.48	0.28–0.85	0.011*	3.07	0.89–10.6	0.077+	4.75	1.91–11.9	0.001**
Residing region	0: mountainous area, 1: coastal area	0.63	0.36–1.11	0.112	0.60	0.38–0.94	0.024*	1.13	0.47–2.75	0.785	1.41	0.75–2.61	0.285
Smoking	0: ex or never smoker, 1: current smoker	1.05	0.36–3.01	0.931	1.31	0.67–2.57	0.431	0.56	0.07–4.69	0.593	1.16	0.31–4.13	0.819
Alcohol consumption	0: ex or never drinker, 1: current drinker	1.04	0.56–1.91	0.907	0.78	0.50–1.21	0.260	1.18	0.49–2.87	0.711	0.66	0.32–1.37	0.260

BMI body mass index, *CI* confidence interval, *KL* Kellgren–Lawrence grade, *OR* odds ratio.
$+p < 0.1$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.
Multivariate logistic regression analysis was performed for each of the outcome variables: occurrence of knee osteoarthritis (0, nonoccurrence; 1, occurrence), occurrence of lumbar spondylosis (0, absence; 1, presence), occurrence of osteoporosis of the lumbar spine (L2–4), and occurrence of osteoporosis of the femoral neck. All of the models included obesity, hypertension, dyslipidemia, and impaired glucose tolerance at baseline as explanatory variables in addition to the presence at baseline of each of the other musculoskeletal variables (knee osteoarthritis, lumbar spondylosis, osteoporosis at L2–4, osteoporosis of the femoral neck) after adjusting for age, sex, regional differences, smoking, and alcohol consumption.

DOI 10.3109/14397595.2014.972607

Table 5. Results of the logistic regression analyses for the occurrence of metabolic components over a 3-year period.

		Occurrence of obesity			Occurrence of hypertension			Occurrence of dyslipidemia			Occurrence of impaired glucose tolerance		
Explanatory variables	Reference	OR	95% CI	p value	OR	95% CI	p value	OR	95% CI	p value	OR	95% CI	p value
Musculoskeletal diseases													
Knee osteoarthritis	0: KL 0,1 vs 1: KL = 2	0.91	0.38–2.20	0.832	1.84	1.09–3.12	0.024*	1.17	0.79–1.73	0.440	1.06	0.67–1.67	0.797
	0: KL 0,1 vs 2: KL ≥ 3	1.09	0.39–3.08	0.866	1.96	0.89–4.34	0.096+	1.66	1.05–2.61	0.029*	0.70	0.39–1.29	0.256
Lumbar spondylosis	0: KL 0,1 vs 1: KL = 2	1.32	0.55–3.16	0.532	0.82	0.47–1.42	0.479	1.08	0.71–1.65	0.720	1.11	0.68–1.82	0.673
	0: KL 0,1 vs 2: KL ≥ 3	1.20	0.55–2.62	0.655	1.44	0.88–2.34	0.143	0.95	0.64–1.39	0.779	0.90	0.57–1.44	0.670
Osteoporosis of the lumbar spine (L2–4)	0: absent, 1: present	0.47	0.11–2.06	0.314	1.52	0.77–3.01	0.231	1.20	0.75–1.92	0.446	0.73	0.40–1.33	0.302
Components of metabolic syndrome													
Obesity	0: BMI ≤ 27.5, 1: BMI > 27.5	–	–	–	3.39	1.25–9.17	0.016*	1.04	0.63–1.72	0.866	2.42	1.34–4.39	0.004**
Hypertension	0: absent, 1: present	2.99	1.35–6.59	0.007**	–	–	–	1.29	0.89–1.85	0.174	0.82	0.54–1.22	0.326
Dyslipidemia	0: absent, 1: present	1.88	0.86–4.09	0.112	1.20	0.61–2.35	0.589	–	–	–	1.17	0.66–2.05	0.598
Impaired glucose tolerance	0: absent, 1: present	2.02	0.97–4.23	0.061+	1.17	0.62–2.21	0.620	1.53	1.06–2.20	0.023*	–	–	–
Adjusted factors													
Age, y	+ 1 y	0.95	0.91–0.98	0.002**	1.01	0.99–1.04	0.162	1.02	1.00–1.03	0.087+	1.02	1.00–1.05	0.025*
Sex	0: men, 1: women	1.46	0.61–3.47	0.396	0.77	0.45–1.33	0.352	1.28	0.85–1.93	0.245	1.35	0.81–2.25	0.245
Residing region	0: mountainous area, 1: coastal area	3.30	1.38–7.92	0.007**	1.25	0.80–1.97	0.327	1.84	1.31–2.59	<0.001***	0.65	0.43–0.98	0.038*
Smoking	0: ex or never smoker, 1: current smoker	0.66	0.20–2.25	0.510	1.82	0.98–3.40	0.060+	0.79	0.44–1.41	0.425	1.11	0.59–2.09	0.747
Alcohol consumption	0: ex or never drinker, 1: current drinker	0.80	0.39–1.64	0.549	0.78	0.50–1.20	0.259	0.71	0.50–1.01	0.060+	0.69	0.45–1.07	0.096+

BMI body mass index, *CI* confidence interval, *KL* Kellgren–Lawrence grade, *OR* odds ratio.
$+p < 0.1$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.
Multivariate logistic regression analysis was performed for each of the metabolic syndrome components: obesity, hypertension, dyslipidemia, and impaired glucose tolerance. All of the models included knee osteoarthritis (0, KL grades 0, 1; 1, KL grade 2; 2, KL grades, 3, 4), lumbar spondylosis (0, KL grades 0, 1; 1, KL grade 2; 2, KL grades 3, 4), osteoporosis (at L2–4) at baseline as explanatory variables in addition to the presence at baseline of each of the other metabolic syndrome components after adjusting for age, sex, regional differences, smoking, and alcohol consumption.

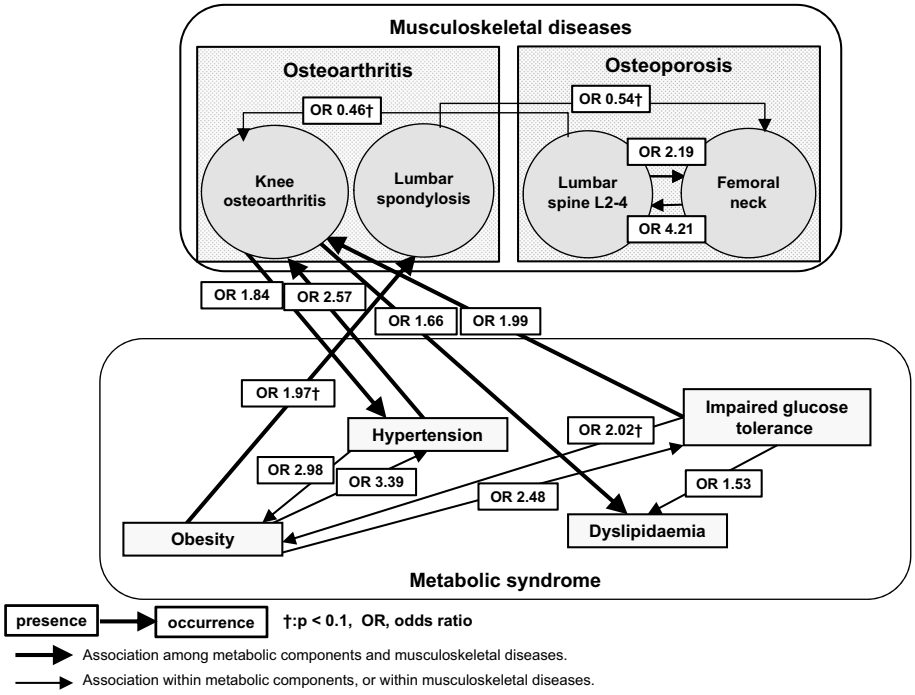


Figure 1. Summary of the mutual associations among musculoskeletal diseases and metabolic syndrome components.

was not significant. Therefore, to prevent new occurrence of KOA, the importance of body weight reduction remains unchanged.

With regard to the etiology of the association between OA and metabolic syndrome components, mechanical stress resulting from the weight-related load would be the first consideration. However, the results of the present study remained the same after the adjustment for OB. Zhuo et al. provided several explanations in their review in addition to summarizing the shared mechanisms involving inflammation, oxidative stress, common metabolites, and endothelial dysfunction that characterize the etiologies of OA and metabolic syndrome [30]. In the interest of determining whether inflammation played a role in the present study, we added baseline serum C-reactive protein (CRP) levels to the logistic regression analysis; the adjusted ORs for HT and IGT, with the outcome as the occurrence of KOA, remained unchanged. Because we used a standard method to measure CRP levels, further research is required to assess the effect of systemic inflammation on KOA using a more sensitive measurement method. With regard to oxidative stress, we plan to measure oxidative stress after we determine suitable indices and evaluate the potential role of oxidative stress as a shared risk factor for both musculoskeletal diseases and metabolic syndrome components.

The risk of the occurrence of LS was not influenced by any of the variables included in the present study; however, OB tended to increase the risk. Fewer environmental risk factors, including occupational activities, exist for LS than for KOA [31]. It is possible that lumbar osteoproliferative changes are not influenced by mechanical stress to the same extent as the knee. On the other hand, the definition in the present study of a new case of joint OA as a baseline KL grade <2 for joints and a follow-up grade ≥ 2 for at least one joint may have also influenced the findings. It is common to use a KL grade ≥ 2 for at least one joint as the criterion for the presence of OA, and, in the Japanese, a previously reported prevalence using this criterion was very high [18]. Therefore, the population at risk for the occurrence of LS was small with insufficient statistical power to detect significant relationships in the present study. However, the results were not affected when we changed the criterion to a more severe KL grade ≥ 3 for at least one joint in the logistic multivariate analysis, indicating that it is necessary to use another index to detect the risk factors for LS. We have started to measure osteophytes and minimum joint space using computer-assisted systems to enable the detection of risk factors for the development of OA of the lumbar spine.

Only the presence of OP at one site increased the risk of the occurrence of OP at the other site. Therefore, the treatment of OP at any site may prevent OP at other sites. In addition, OB tended to decrease the risk of the occurrence of OP. When we replaced OB with BMI in the logistic regression analysis, a greater BMI ($+1$ kg/m^2) significantly reduced the risk of the occurrence of OP at L2–4 and the femoral neck; it has previously been reported that emaciation increases the risk of OP [12].

We have previously reported that the presence of lumbar OP significantly reduces the risk of the occurrence of LS in a population-based cohort that consisted of 400 Japanese middle-aged and elderly people who were followed up for 10 years [17]. In the present study, LS tended to reduce the risk of the occurrence of OP at the femoral neck, and lumbar OP tended to reduce the risk of the occurrence of KOA. The low incidence of OP at both L2–4 (0.76%/year) and the femoral neck (1.83%/year) might not have resulted in sufficient statistical power to detect the risk of other musculoskeletal diseases. A longer observation period may be required to determine the significant risk factors, including musculoskeletal diseases and metabolic risk factors, for the occurrence of OP.

We also detected associations between the presence and occurrence of metabolic syndrome components, including OB and HT (reciprocal), OB and IGT (reciprocal), and IGT and DL. However, there was no significant relationship between HT and DL after adjustment for various confounders including smoking. When we replaced OB with the BMI values in the analysis, the significant associations that were present between OB and HT as well as IGT remained with BMI; furthermore, we found that BMI significantly increased the risk of the occurrence of DL (lumbar spine OP as an explanatory factor for OP, BMI + 1 kg/m^2; OR, 1.08; 95% CI, 1.03–1.14; $p = 0.001$; femoral neck OP as an explanatory factor for OP, BMI + 1 kg/m^2; OR, 1.09; 95% CI, 1.04–1.14; $p < 0.001$). Therefore, we emphasize that greater body composition increases the risk of other components of metabolic syndrome, and the

maintenance of a healthy body composition is important to reduce the occurrence of metabolic syndrome.

This study has certain limitations. First, the definitions for the metabolic syndrome components were not exactly the same as those used internationally by such associations as the Adult Treatment Panel (ATP) III, World Health Organization (WHO), and American Association of Clinical Endocrinologists (AACE) [32]. As there has been considerable debate regarding abdominal circumference ($\geq$ 85 cm in men, $\geq$ 90 cm in women) in the Japanese criteria [33], we decided to use BMI $\geq$ 27.5 kg/m^2 to indicate OB rather than abdominal circumference. Furthermore, because not all blood samples were obtained under fasting conditions, we did not use blood glucose or serum triglyceride levels as indicators. Therefore, our results may underestimate the presence of metabolic syndrome components, especially DL and IGT. However, we used an alternative index for each condition, as recommended by the National Health and Nutrition Survey, for cases in which samples could not be collected under fasting conditions [26]; thus, we believe our criteria could reflect dysfunction in lipid and glucose metabolisms. Finally, as previously mentioned, the incidence of some diseases, especially OP, may have been too low to detect risk factors in multivariate models. Therefore, longer observations should be conducted to determine the significant risk factors for OP. Similarly, the low incidence of OB might have biased the results, as indicated by the significant associations that were present with BMI but not OB. Because the main aim of the present study was to clarify the association between musculoskeletal diseases and metabolic components, we used the presence or occurrence of OB in the analysis. These results support the importance of body composition for the occurrence of musculoskeletal diseases or the components of metabolic syndrome, despite a lack of association with OB *per se*.

To conclude, we identified associations between the presence and occurrence of musculoskeletal diseases and metabolic syndrome components. Therefore, proactive prevention of metabolic syndrome or musculoskeletal diseases may also influence the occurrence of the other.

Acknowledgments

This work was supported by a Grant-in-Aid for H17-Men-eki-009 (Director, Kozo Nakamura), H20-Choujyu-009 (Director, Noriko Yoshimura), H23-Choujyu-002 (Director, Toru Akune), H-25-Choujyu-007 (Director, Noriko Yoshimura), and H25-Nanchitou(Men)-005 (Director, Sakae Tanaka) of the Ministry of Health, Labour and Welfare; and Scientific Research B23390172, B20390182, and Challenging Exploratory Research 24659317 to Noriko Yoshimura; B23390357 and C20591737 to Toru Akune; B23390356, C20591774, and Challenging Exploratory Research 23659580 to Shigeyuki Muraki; Challenging Exploratory Research 24659666 and 21659349 and Young Scientists A18689031 to Hiroyuki Oka; and Collaborating Research with NSF 08033011–00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. This study also was supported by grants from the Japan Osteoporosis Society (Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, and Toru Akune), and research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006–1 & 2010–2; Director, Hiroshi Kawaguchi).

The authors wish to thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka of the Gobo Public Health Centre; Dr. Naoki Hirabayashi of the Kawakami Clinic in Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Centre; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji, and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. We also thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, Mrs. Saeko Sahara, and Mr. Noriyuki Oe for their assistance in data reduction and administration.

Conflict of interest

All authors declare that (1) no author has received corporate support for the submitted work; (2) the authors have no relationships with companies that might have an interest in the submitted work in the previous 3 years; (3) the authors' spouses, partners, or children do not have financial relationships that may be relevant to the submitted work; and (4) the authors have no non-financial interests that may be relevant to the submitted work.

References

1. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2010. http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/(Accessed September 24, 2014) [in Japanese].
2. Dahlöf B. Cardiovascular disease risk factors: epidemiology and risk assessment. Am J Cardiol. 2010;105(1 Suppl):3A–9A.
3. Day C. Metabolic syndrome, or What you will: definitions and epidemiology. Diab Vasc Dis Res. 2007;4(1):32–8.
4. Lohmander LS, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Engström G. Incidence of severe knee and hip osteoarthritis in relation to different measures of body mass: a population-based prospective cohort study. Ann Rheum Dis. 2009;68(4):490–6.
5. Hart DJ, Doyle DV, Spector TD. Incidence and risk factors for radiographic knee osteoarthritis in middle-aged women: the Chingford Study. Arthritis Rheum. 1999;42(1):17–24.
6. Reijman M, Pols HA, Bergink AP, Hazes JM, Belo JN, Lievense AM, Bierma-Zeinstra SM. Body mass index associated with onset and progression of osteoarthritis of the knee but not of the hip: the Rotterdam Study. Ann Rheum Dis. 2007;66(2):158–62.
7. Cooper C, Snow S, McAlindon TE, Kellingray S, Stuart B, Coggon D, Dieppe PA. Risk factors for the incidence and progression of radiographic knee osteoarthritis. Arthritis Rheum. 2000;43(5):995–1000.
8. Engström G, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Lohmander LS. C-reactive protein, metabolic syndrome and incidence of severe hip and knee osteoarthritis. A population-based cohort study. Osteoarthritis Cartilage. 2009;17(2):168–73.
9. Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, et al. Incidence and risk factors for radiographic knee osteoarthritis and knee pain in Japanese men and women: a longitudinal population-based cohort study. Arthritis Rheum. 2012;64(5):1447–56.
10. Yoshimura N, Muraki S, Oka H, Tanaka S. Kawaguchi H, Nakamura K, Akune T. Accumulation of metabolic risk factors such as overweight, hypertension, dyslipidaemia, and impaired glucose tolerance raises the risk of occurrence and progression of knee osteoarthritis: a 3-year follow-up of the ROAD study. Osteoarthritis Cartilage. 2012;20(11):1217–26.
11. Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, et al. Incidence and risk factors for radiographic lumbar spondylosis and lower back pain in Japanese men and women: the ROAD study. Osteoarthritis Cartilage. 2012;20(7):712–8.
12. Johansson H, Kanis JA, Odén A, McCloskey E, Chapurlat RD, Christiansen C, et al. A meta-analysis of the association of fracture risk and body mass index in women. J Bone Miner Res. 2014;29(1):223–33.
13. Ghosh M, Majumdar SR. Antihypertensive medications, bone mineral density, and fractures: a review of old cardiac drugs that provides new insights into osteoporosis. Endocrine. 2014;46(3):397–405. [Epub ahead of print].
14. Janghorbani M, Van Dam RM, Willett WC, Hu FB. Systematic review of type 1 and type 2 diabetes mellitus and risk of fracture. Am J Epidemiol. 2007;166(5):495–505.

15. Ahmed LA, Schirmer H, Berntsen GK, Fønnebø V, Joakimsen RM. Features of the metabolic syndrome and the risk of non-vertebral fractures: the Tromsø study. Osteoporos Int. 2006;17(3):426–32.
16. Buizert PJ, van Schoor NM, Lips P, Deeg DJ, Eekhoff EM. Lipid levels: a link between cardiovascular disease and osteoporosis? J Bone Miner Res. 2009;24(6):1103–9.
17. Yoshimura N, Muraki S, Oka H, Mabuchi A, Kinoshita H, Yoshida M, et al. Epidemiology of lumbar osteoporosis and osteoarthritis and their causal relationship - is osteoarthritis a predictor for osteoporosis or vice versa?: the Miyama Study. Osteoporos Int. 2009;20(6): 999–1008.
18. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab. 2009;27(5):620–8.
19. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol. 2010;39(4):988–95.
20. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study–a large observational study of the determinants of fracture in older men. Contemp Clin Trials. 2005;26(5):569–85.
21. Kellgren JH, Lawrence JS. The Epidemiology of Chronic Rheumatism: Atlas of Standard Radiographs of Arthritis. Oxford: Blackwell Scientific; 1963.
22. World Health Organization. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. WHO Technical Report Series 843. Geneva, World Health Organization 1994.
23. Orimo H, Hayashi Y, Fukunaga M, Sone T, Fujiwara S, Shiraki M, et al.; Osteoporosis Diagnostic Criteria Review Committee: Japanese Society for Bone and Mineral Research. Diagnostic criteria for primary osteoporosis: year 2000 revision. J Bone Miner Metab. 2001;19(6):331–7.
24. Examination Committee of Criteria for Metabolic Syndrome. The definition and criteria of metabolic syndrome (in Japanese). J Jpn Soc Intern Med. 2005;94(4):794–809.
25. Examination Committee of Criteria for 'Obesity Disease' in Japan; Japan Society for the Study of Obesity. New criteria for 'obesity disease' in Japan. Circ J. 2002;66(11):987–92.
26. Ministry of Health, Labour and Welfare. The manual of the National Health and Nutrition Survey. http://www.mhlw.go.jp/bunya/shakaihosho/iryouseido01/pdf/tdfk13.pdf (Accessed May 11, 2014) [in Japanese].
27. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004;363(9403):157–63.
28. Seino Y, Nanjo J, Tajima N, Kadowaki T, Kashiwagi A, Araki E, et al. Report of the Committee on the classification and diagnostic criteria of diabetes mellitus: the Committee of the Japan Diabetes Society on the diagnostic criteria of diabetes mellitus. Diabetol Int. 2010; 1(5):2–20.
29. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Association of knee osteoarthritis with the accumulation of metabolic risk factors such as overweight, hypertension, dyslipidemia, and impaired glucose tolerance in Japanese men and women: the ROAD Study. J Rheumatol. 2011;38(5):921–30.
30. Zhuo Q, Yang W, Chen J, Wang Y. Metabolic syndrome meets osteoarthritis. Nat Rev Rheumatol. 2012;8(12):729–37.
31. Muraki S, Akune T, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Association of occupational activity with radiographic knee osteoarthritis and lumbar spondylosis in elderly patients of population-based cohorts: a large-scale population-based study. Arthritis Rheum. 2009;61(6):779–86.
32. Grundy SM, Brewer HB Jr, Cleeman JI, Smith SC Jr, Lenfant C; American Heart Association; National Heart, Lung, and Blood Institute. Definition of metabolic syndrome: report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. Circulation. 2004;109(3):433–8.
33. Shibata K, Suzuki S, Sato J, Ohsawa I, Goto S, Hashiguchi M, Tokudome S. Abdominal circumference should not be a required criterion for the diagnosis of metabolic syndrome. Environ Health Prev Med. 2010;15(4):229–35.

J Orthop Sci (2015) 20:287–294
DOI 10.1007/s00776-014-0677-1

ORIGINAL ARTICLE

Improved accuracy of diagnosis of lumbar intra and/or extra-foraminal stenosis by use of three-dimensional MR imaging: comparison with conventional MR imaging

Hiroshi Yamada · Masaki Terada · Hiroshi Iwasaki · Toru Endo · Motohiro Okada · Shinichi Nakao · Hiroshi Hashizume · Akihito Minamide · Yukihiro Nakagawa · Hideto Nishi · Syunji Tsutsui · Hiroyuki Oka · Munehito Yoshida

Received: 9 July 2014 / Accepted: 19 November 2014 / Published online: 11 December 2014

Abstract
Background The purposes of this study were to assess the reliability of 3-dimensional magnetic resonance (MR) imaging (3D MRI) and conventional MRI (CMRI) for detection of lumbar intra and/or extra-foraminal stenosis (LIEFS) and to compare the diagnostic accuracy of the 2 imaging modalities.
Methods A total of 60 sets of 3D MR and CMR images from 20 healthy volunteers and 40 LIEFS patients were qualitatively rated according to defined criteria by 3 independent, blinded readers. Kappa statistics were used to characterize intra and inter-reader reliability for qualitative rating of data. Multireader, multicase analysis was used to compare lumbar foraminal stenosis detection between the 2 modalities.
Results Intra-reader agreement for 3D MRI was excellent, with kappa = 0.90; that for CMRI was good, with kappa = 0.78. Average inter-reader agreement for 3D MRI was good, with kappa = 0.79, whereas that for CMRI was moderate, with kappa = 0.41. Average area under the ROC curve values (1st reading/2nd reading) for detection of lumbar foraminal stenosis using 3D MRI and CMRI were 0.99/0.99 and 0.94/0.92, respectively. Detection of LIEFS with 3D MRI was significantly better than with CMRI ($P = 0.0408/0.0294$).
Conclusions These results suggest that CMRI was of limited use for detection of the presence of LIEFS. Isolated imaging with CMRI may risk overlooking the presence of LIEFS. In contrast, reliability of 3D MRI for detection of LIEFS was good. Furthermore, readers' performance in the diagnosis of LIEFS can be improved by use of 3D MRI. Therefore, 3D MRI is recommended when using imaging for diagnosis of LIEFS.

H. Yamada (✉) · H. Iwasaki · T. Endo · M. Okada · S. Nakao · H. Hashizume · A. Minamide · Y. Nakagawa · H. Nishi · S. Tsutsui · M. Yoshida
Department of Orthopedic Surgery, Wakayama Medical University, 811-1 Kimiidera, Wakayama, Wakayama 641-8510, Japan
e-mail: yamacha@wakayama-med.ac.jp

M. Terada
Wamayama-Minami Radiology Clinic, Wakayama, Japan

H. Oka
Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Bunkyō, Japan

Introduction

Most surgery for lumbar spinal stenosis is considered successful for relief of symptoms, but a significant number of failures occur. The term "failed back surgery syndrome" (FBSS) is used to designate persistent complaints of low back pain and/or leg pain among patients who have undergone surgical procedures intended to relieve those complaints. Burton et al. reported the most common reason for FBSS as failure to diagnose lateral spinal stenosis, which includes intra and/or extra-foraminal nerve root entrapment [1]. This diagnostic limitation seems to be the result of difficulty identifying the lesion outside the spinal canal with conventional imaging modalities [2]. When the site of nerve compression is peripheral to the root sleeves, myelography provides no information [3]. Therefore, parasagittal magnetic resonance (MR) images have long been the recommended method for investigation of any abnormality in the region of the intervertebral foramen. Obliteration of the perineural fat surrounding the nerve root has been reported to

be the finding most suggestive of lumbar intra and/or extraforaminal stenosis (LIEFS). However, such MR images do not give complete information and sometimes result in false-positive or false-negative findings [4, 5].

The advent of 3-dimensional magnetic resonance imaging (3D MRI) has rapidly countered this difficulty in recent years [6, 7]. Direct visualization of the nerve root in the intervertebral foramen by use of 3D MRI has been a great benefit in the daily practice of spinal medicine. Nerve abnormalities outside the spinal canal can be easily diagnosed by using this modality. However, no reference standard for diagnosis of LIEFS on 3D MRI has been established, and the diagnostic performance of 3D MRI and conventional imaging modalities has not yet been directly compared.

The purposes of this study were to propose new diagnostic criteria for LIEFS on 3D MRI, to assess the reliability of 3D MRI and conventional MRI (CMRI) for the identification of LIEFS, and to compare LIEFS detection between the 2 imaging modalities.

Materials and methods

Patient population

The local ethics committee approved this retrospective study. Informed consent was obtained from all study participants. To collect CMRI and 3D MRI data samples for this reading study, positive images (radiographically abnormal cases) were obtained from patients who underwent surgery for LIEFS at our university hospital. To achieve uniform reading conditions, the responsible lesions were limited to the L5-S1 level, which is the most common site of involvement of LIEFS. To establish a clear diagnosis, only cases involving a single stenotic lesion were selected. That is, the patients had no concomitant intraspinal canal lesion at the time of surgery. Their leg symptoms were completely resolved after L5 nerve block, and the outcome of decompression surgery for LIEFS at L5–S1 was successful. LIEFS images that were of low diagnostic quality for the examinations, because of motion or other artifacts were excluded. After reviewing all the images and records, the final study group with positive images comprised 40 patients (22 male, 18 female; age 50–84 years, mean 70.0 years). The study group for negative images (radiographically normal cases) consisted of 20 normal healthy volunteers (10 male, 10 female; age 18–28 years; mean 21.3 years). These subjects had no current or previous history of low back pain and/or leg pain and no history of spinal disorders, and their spinal MRI revealed no disc degeneration over the entire lumbar spine. Thus, a total of 60 sets of 3D MRI and CMRI images from 20 normal healthy volunteers and 40 LIEFS patients were evaluated.

Imaging technique

MRI was performed with a 3.0 T MR scanner (Achieva; Philips Medical Systems, Best, Netherlands) using a 32-channel SENSE-Torso/Cardiac coil. The sequences for 3D MRI are called "Proset Myelo" in this system (3D FFE with, in principle, the selective excitation technique). The scanner settings were: TR = 20 ms, TE = 8 ms, flip angle = 15 degrees, slice thickness = 0.55 mm, field of view = 240 mm, and matrix = 256 × 512. Images were then subjected to postprocessing with multiplanar reformatting in a workstation to provide continuous longitudinal nerve images. For the conventional MRI procedure, the spin echo sequences of sagittal T1-weighted images were: TR = 650 ms, TE = 10 ms, flip angle = 90°, field of view = 280 mm, and matrix = 352 × 512.

Imaging analysis

Images were provided to the readers on CDs and viewed with commercial software (Virtual Place Liberty; AZE, Japan). Display monitors were not standardized across readers. Qualitative ratings were performed according to defined criteria by 3 independent readers (orthopedic surgeons and board-certified spinal experts of the Japanese Orthopedic Association) with no knowledge of the patients' clinical information. Each reader attended a lecture explaining the standardized definitions of imaging features from the first author (H.Y.), and consensus for interpretation of 3D MRI and CMRI findings was obtained among the readers before the start of the study. The images were scored as "definitely showing LIEFS", "probably showing LIEFS", "probably not showing LIEFS", or "definitely not showing LIEFS". All 3D MRI and CMRI images were independently evaluated twice by the 3 readers with a 1-month interval between readings.

The following reference standards were used for each imaging modality. On CMRI, obliteration of the perineural fat surrounding the nerve root in the intervertebral foramen was taken into consideration (parasagittal MRI reading technique). If the perineural fat was clearly visible, the image was considered normal. If the perineural fat was diminished because of disc height loss, osteoarthritic changes in the facet joints, buckling of the ligamentum flavum, or protrusion of the annulus fibrosis, the image was considered abnormal [8].

No widely used diagnostic criterion or grading system exists for LIEFS on 3D MRI. Hence, the authors reviewed the images of all the surgical cases of LIEFS and selected 4 representative imaging features as an index of highly suspicious findings of LIEFS:

1. transverse path of the nerve root and/or spinal nerve (Fig. 1a);

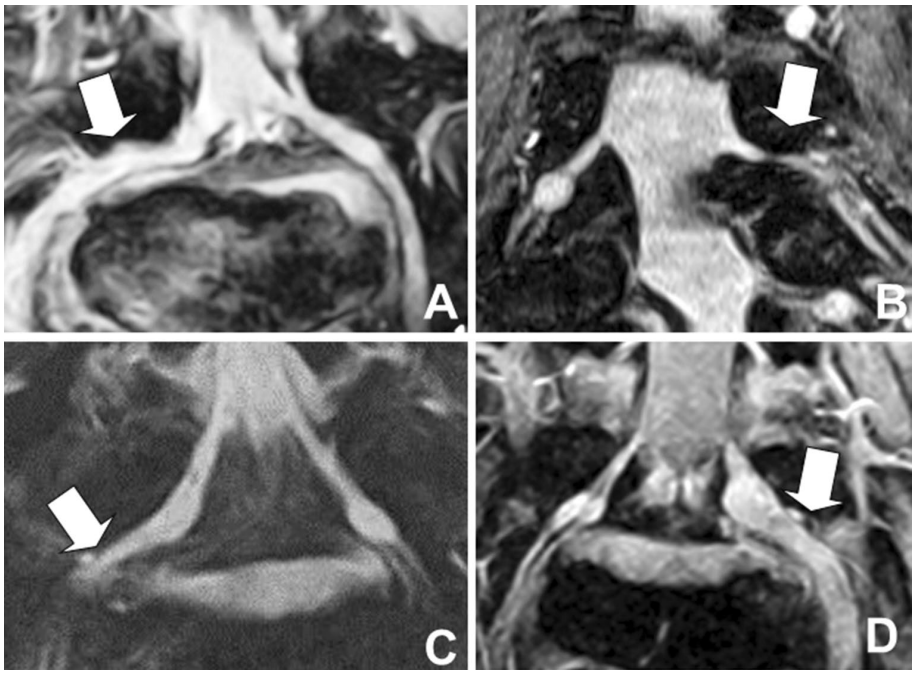


Fig. 1 Index of highly suspicious findings of LIEFS: **a** transverse path of the nerve root and spinal nerve, **b** obscurity of the spinal ganglion, **c** constriction of the spinal nerve, and **d** nerve swelling

2. obscurity of the dorsal root ganglion (DRG) (Fig. 1b);
3. spinal nerve indentation (Fig. 1c); and
4. nerve swelling (Fig. 1d).

These radiological findings on 3D MRI were highly consistent with the anatomical abnormalities in our surgical case series. Although all 4 radiological abnormalities were not always present on 3D MRI in cases with LIEFS, at least 1 imaging feature was always observed. Therefore, if any 1 of these 4 radiological findings was observed, it was regarded as abnormal. Images that did not have any of these radiological findings were considered normal.

As a basic step toward developing diagnostic criteria, we defined the area from the medial pedicle wall to the lateral pedicle wall as the intra-foraminal zone and the area beyond the lateral pedicle wall as the extra-foraminal zone. The spindle-shaped nerve root in the intra and/or extra-foraminal zone was defined as the DRG. The nerve root in the extra-foraminal zone beyond the DRG was defined as the spinal nerve. The nerve between the thecal sac and the DRG was defined as the nerve root [9]. Imaging features and their association with intraoperative findings were:

1. Transverse path of the nerve root and/or spinal nerve (Fig. 1a). The transverse path was defined as positive when either of the tilting angles of the nerve root and/or spinal nerve in the intra and/or extra-foraminal zone was larger than that of the normal contralateral side. The normal nerve root and spinal nerve basically run obliquely downward through the intervertebral foramina after branching from the dura mater. Hence, a transverse path is abnormal. The nerve root and/or spinal nerve were compressed and shifted upward by degenerating bulging discs. The transverse path of the nerve root and/or spinal nerve indicated the presence of up-down-type stenosis in the intra and/or extraforaminal zone [10, 11].
2. Obscurity of the DRG (Fig. 1b). Obscurity of the DRG was defined as positive when its configuration became unclear. DRGs are the largest neural structures around the intervertebral foramina and are ordinarily distinguishable from the nerve root and spinal nerve by their spindle shape on 3D MR images. Because the vast majority of DRGs are normally located in the intra-foraminal zone, morphological changes to the DRG imply significant diminishment of available space for the nerve root around the intervertebral foramen. Thus, when DRG configuration becomes unclear, the existence of LIEFS is strongly suspected. DRGs were compressed by degenerative bulging discs from the inferior and by superior articular facets from the posterior of the intra-foraminal zone. Obscurity of the DRG indicated circumferential-type stenosis in the intra-foraminal zone [4, 11].

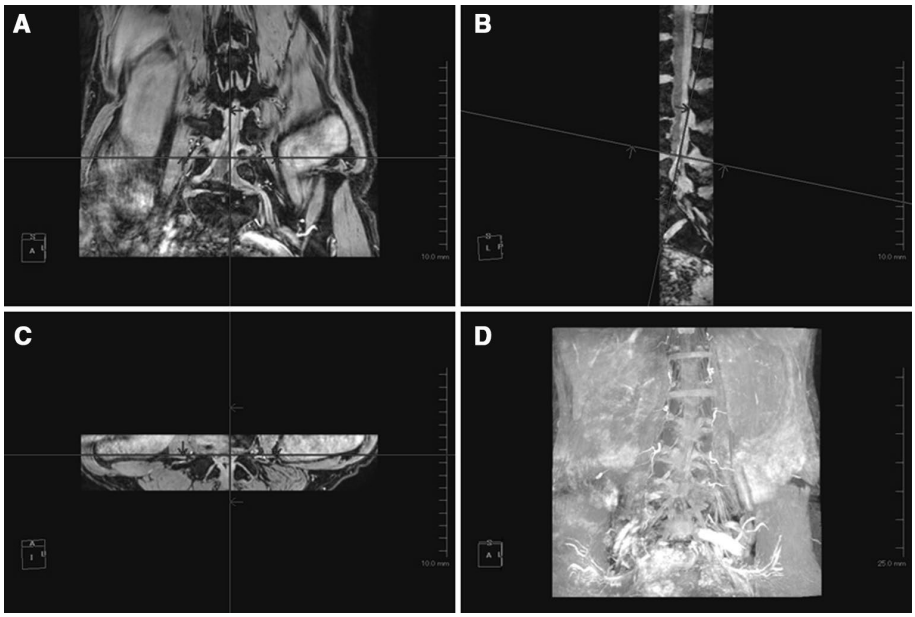


Fig. 2 An example of the 3D view of the 5th lumbar nerve roots and spinal nerves using MPR (multi-planar reconstruction) in **a** coronal view, **b** axial view, **c** sagittal view and the MIP (maximum intensity projection) technique (**d**)

3. Spinal nerve indentation (Fig. 1c). Spinal nerve indentation was regarded as positive when the nerve diameter narrowed locally. Because the nerve thickness is the same other than at the dorsal root ganglion, indentation is evidence suggesting the presence of stenosis. Spinal nerve indentation indicated spinal nerve entrapment on operative findings, which was mainly observed in cases of extra-foraminal stenosis at the lumbosacral junction. This lesion had several etiologies, for example impingement of the transverse process of L5 against the ala of the sacrum [12], entrapment by the lumbosacral ligament [13, 14], and entrapment by osteophytes of the L5 vertebral bodies and sacral ala [15]. All these lesions occurred in the lumbosacral tunnel [16] in the extra-foraminal zone. The lumbosacral tunnel is an osteofibrotic tunnel for the exiting nerve of L5, which is formed by developing osteophytes and the degenerating hypertrophied lumbosacral ligament. This unique anatomical structure contributed to spinal nerve entrapment.
4. Nerve swelling (Fig. 1d). Swelling of the nerve root, DRG, or spinal nerve was defined as positive when the size of any of these structures exceeded that of the normal contralateral side. Although morphological changes to the nerve do not always indicate symptomatic radiculopathy, nerve swelling has been reported as a good indicator of symptomatic LIEFS in previous studies [7]. Thus, the existence of nerve swelling at any level from the entrance zone to the extraforaminal zone and at any range from local to total was considered abnormal. This finding was often observed in cases with clinical symptoms of severe spontaneous pain. Thus, nerve swelling may indicate the existence of nerve inflammation.

An example of the 3D view of the 5th lumbar nerve roots and spinal nerves using multi-planar reconstruction and the maximum intensity projection technique is shown in Fig. 2. The readers could observe many different views of the nerves by adjusting the screen to detect the intra and/or extra-foraminal lesion.

Statistical methods

First, reliability was measured by use of kappa statistics. This statistical analysis was performed by using JMP version 10 (SAS Institute Japan, Tokyo, Japan). Intra and inter-reader reliability was assessed with the kappa coefficient, which was characterized as: <0.0 = "poor" agreement, 0.0–0.2 = "slight" agreement beyond chance, 0.21–0.4 = "fair" agreement, 0.41–0.60 = "moderate" agreement, 0.61–0.80 = "substantial" agreement, and 0.81–1.00 = "almost perfect" agreement [17]. To analyze the data obtained, readers' ratings were classified into 2 categories: "positive," consisting of definite or probable LIEFS, and "negative," consisting of cases defined as probably or definitely not showing LIEFS. Second, to compare the diagnostic performance of the 2 imaging modalities, receiver operating

characteristic (ROC) analysis based on a sequential test method was used [18, 19]. To analyze the data, readers' ratings were converted into a numerical rating scale as follows: "definitely showing LIEFS" = 4, "probably showing LIEFS" = 3, "probably not showing LIEFS" = 2, and "definitely not showing LIEFS" = 1. The area under the ROC curve (AUC) and 95 % confidence intervals (CI) were obtained with a quasi-maximum likelihood estimation of binomial distribution by using DBM MRMC software (version 2.2). Accuracy was determined by using the AUC. An area of 1 represents a perfect test; an area of 0.5 represents a worthless test. An approximately guide for classifying the accuracy of a diagnostic test is the traditional academic point system, with AUC values 0.90–1 = "excellent", 0.80–0.90 = "good", 0.70–0.80 = "fair", 0.60–0.70 = "poor", and 0.50–0.60 = "fail" [20]. The significance of the different AUC for CMRI and 3D MRI was tested by use of the Dorfman–Berbaum–Metz method, which included both reader variation and case sample variation, by means of an analysis of variance approach. *P* values <0.05 were regarded as indicative of a significant difference.

Results

The first and second reading results of each reader are summarized in Table 1. Average sensitivity and specificity for 3D MRI (1st reading/2nd reading) were 85.8/90.0 and 98.3/98.3 %, respectively. Those for CMRI were 59.2/63.3 and 100/100 %, respectively. Intra-reader reliability for CMRI and 3D MRI is summarized in Table 2. Intra-reader reliability for 3D MRI showed agreement was excellent, with kappa = 0.90, whereas agreement for CMRI was good, with kappa = 0.78. Inter-reader reliability for CMRI and 3D MRI are summarized in Tables 3 and 4, respectively. Inter-reader agreement for 3D MRI was good, with kappa = 0.79, whereas agreement for CMRI was moderate, with kappa = 0.41. The AUC values of the ROC curves for 3D MRI and CMRI for individual observers are summarized in Table 5. The mean AUC of the ROC curves (1st reading/2nd reading) for the 3 observers was 0.99/0.99 for 3D MRI and 0.94/0.92 for CMRI (Figs. 3, 4). AUC values for the 2 imaging modalities were statistically significantly different (*P* = 0.0408/0.0294, 95 % CI 0.10009–0.00178/0.00803–0.11867).

Table 1 Sensitivity and specificity of CMRI and 3D MRI for individual observers

Statistic and reader	3D MRI (1st reading/2nd reading)	CMRI (1st reading/2nd reading)
Sensitivity (TP/TP + FN) (%)		
A	82.5/82.5	70.0/67.5
B	82.5/92.5	70.0/57.5
C	92.5/95.0	37.5/65.0
Average	85.8/90.0	59.2/63.3
Specificity (TN/TN + FP) (%)		
A	100/95.0	100/100
B	95.0/100	100/100
C	100/100	100/100
Average	98.3/98.3	100/100

TP indicates true positive, *FN* false negative, *TN* true negative, *FN* false negative

Table 2 Intra-reader reliability of CMRI and 3D MRI for individual observers

Imaging modality and observer	κ Statistic	Standard error	Agreement
3D MRI			
A	0.8986	0.1291	58
B	0.8275	0.1291	55
C	0.9644	0.1291	59
Average	0.8968		
CMRI			
A	0.8661	0.1291	57
B	0.6089	0.1291	49
C	0.8636	0.1291	57
Average	0.7795		

Table 3 Inter-observer reliability of 3D MRI

Observers	κ Statistic	Standard error	Agreement
1st reading			
A and B	0.3114	0.1291	41
A and C	0.3407	0.1291	42
B and C	0.4139	0.1291	45
2nd reading			
A and B	0.5567	0.1291	47
A and C	0.3407	0.1291	42
B and C	0.4920	0.1291	47
Average	0.4092		

Discussion

Failure to diagnose LIEFS continues to be the most common reason for FBSS [1]. The difficulty in identifying LIEFS with conventional imaging modalities is well-recognized. LIEFS tends to be overlooked rather than overestimated. Therefore, new imaging techniques to detect LIEFS with certainty are required.

Table 4 Inter-observer reliability of CMRI

Observers	κ Statistic	Standard error	Agreement
1st reading			
A and B	0.7634	0.1291	53
A and C	0.7943	0.1291	54
B and C	0.7586	0.1291	53
2nd reading			
A and B	0.7586	0.1291	53
A and C	0.7917	0.1291	54
B and C	0.8933	0.1291	57
Average	0.7933		

Table 5 AUC values of the ROC curves for 3D MRI and CMRI for individual observers

Reader	3D MRI (1st reading/2nd reading)	CMRI (1st reading/2nd reading)
A	1.00/0.98	0.96/0.92
B	0.96/1.00	0.91/0.96
C	1.00/0.99	0.93/0.89
Average	0.99/0.99	0.94/0.92

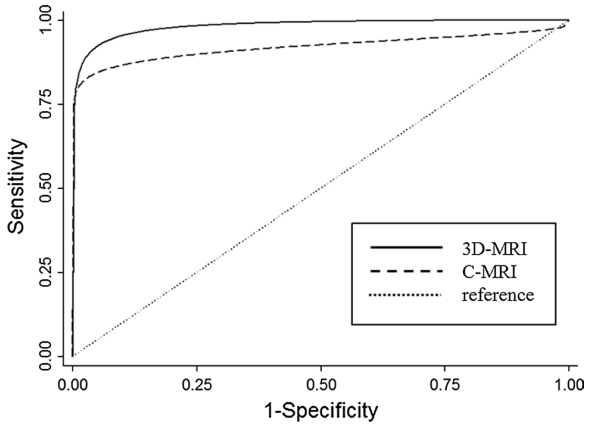


Fig. 3 Averaged ROC curves for 3 observers obtained from 1st reading results for 3D MRI and CMRI

This study confirms the low reliability of the parasagittal MRI reading technique for evaluating LIEFS reported in the past. Speciciale et al. [21] reported the lowest overall inter-observer reliability kappa value, 0.26, for ratings of stenosis severity, which included both foraminal and lateral recess and central stenosis. Lurie et al. [8] reported good agreement for intra-reader reliability for foraminal stenosis, with an overall kappa of 0.77, but moderate agreement for inter-reader reliability, with an overall kappa of 0.58. Our study also revealed good agreement for intra-reader

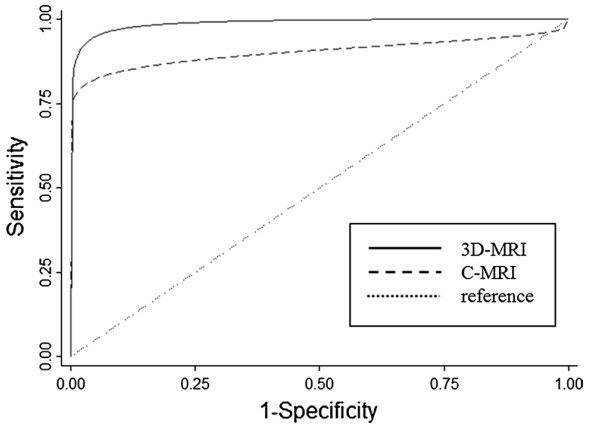


Fig. 4 Averaged ROC curves for 3 observers obtained from 2nd reading results for 3D MRI and CMRI

reliability and moderate agreement for inter-reader reliability, with overall kappa values of 0.78 and 0.41, respectively. Taking these findings into account, we conclude that parasagittal MRI reading of CMRI is not highly reliable for detection of LIEFS, especially with regard to inter-reader agreement. This situation is not ideal when considering surgical indications for LIEFS. Presurgical diagnosis may differ with different examiners, which may increase the number of FBSS patients.

The lower reliability of the parasagittal MRI reading technique for LIEFS identification has several possible causes. Evaluating a limited number of cross-sectional images is not sufficient to identify all abnormalities along the nerve pathways. This technique is limited to evaluation of foraminal stenosis and is inadequate for investigating extraforaminal pathology. Furthermore, accurate evaluation of nerve pathology in intra and/or extraforaminal zones where the nerve pathways progress in 3 dimensions is beyond the ability of 2-dimensional conventional imaging interpretation.

From this perspective, 3D MRI seems ideal for evaluating nerve lesions around the intervertebral foramen. Direct visualization of nerve morphology is very helpful for identifying nerve abnormalities with high reliability. As expected, this new imaging modality resulted in excellent agreement for intra-reader reliability, with overall kappa of 0.90, and good agreement for inter-reader reliability, with overall kappa of 0.79. In addition to improved imaging quality with high spatial resolution, we believe our diagnostic criteria for LIEFS on 3D MRI contributed to readers' understanding of the nerve pathology in the intra and/or extra-foraminal zone and served as a support tool for more reliable identification of nerve lesions.

No study comparing diagnostic performance between 3D MRI and CMRI using ROC analysis has been

conducted to date. ROC analysis is a crucial technique for evaluating diagnostic systems, and ROC curves have been regarded as imperative when comparing new imaging technology [22, 23]. Our study demonstrated that the AUC values of both imaging modalities were greater than 0.9, indicative of excellent diagnostic performance in clinical use, but the AUC of 3D MRI was significantly higher than that of CMRI. This result suggests that the detectability of LIEFS by 3D MRI is superior to that by CMRI. Although direct comparison of diagnostic performance between 3D MRI and conventional imaging studies has not been conducted previously, this study demonstrated that 3D MRI is superior to CMRI in reliability and detectability.

This study has some limitations. First, the number of imaging samples was small, but a significant difference was observed between 3D MRI and CMRI with regard to reliability and detectability. Second, this study only focused on comparing the detectability of abnormal morphology by the 2 imaging modalities. Therefore, we recruited control subjects with little or no degenerative changes for negative images (radiographically normal cases), whereas the study group for positive images was recruited from patients who had LIEFS surgery. LIEFS may be more difficult to distinguish radiologically from age-related changes, and future studies should examine the sensitivity and specificity of these changes for an age-matched control group.

Investigating the percentage of symptomatic nerve roots that are correctly identified as having these 3D MRI abnormalities is also very important. Several reports have indicated that abnormal CMRI findings of lumbar disc herniation or spinal stenosis are not always accompanied by symptoms [24, 25]. Radiographic abnormalities of the nerve are common, but symptomatic nerves are less prevalent. Therefore, the correlation between 3D MRI abnormalities and clinical symptoms must be investigated further. Otherwise, overdiagnosis may occur. To resolve this problem, functional diagnosis with use of selective nerve blocking, electromyographical study [26, 27], and diffusion MRI [28] may be help to improve the specificity of diagnosing symptomatic LIEFS. Nonetheless, improving the detectability of radiographic abnormalities indicative of LEIFS is important in daily spinal practice. Not all screening tests have been shown to benefit the person being screened; however, finding unrecognized LIEFS would be very beneficial for patients who are scheduled for nerve decompression surgery because misdiagnosis of LIEFS remains one of main reasons for FBSS.

In conclusion, CMRI demonstrated limitations in identifying the presence of LIEFS, and isolated imaging by use of CMRI may risk overlooking the presence of LIEFS. In contrast, use of 3D MRI for diagnosis of LIEFS resulted in good reliability and detectability. Our new diagnostic criteria for LIEFS on 3D MRI enable reliable discrimination between a normal root and LIEFS. This study demonstrated that reader performance in diagnosis of LIEFS can be improved by use of 3D MRI. Therefore, 3D MRI is recommenced when using imaging for diagnosis of LIEFS.

Acknowledgments The authors wish to thank Dr Hiroko Ihara, Dr Misato Okouchi, and Mr Yuji Nakao in Wakayama-minami radiology clinic for their technical assistance and for collecting data, and Dr Junji Shiraishi in School of Health Sciences, Kumamoto University for his comment on the study design and interpretation of the data.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Burton R, Kirkaldy-Willis W, Yong-Hing K, Heithoff K. Causes of failure of surgery on the lumbar spine. Clin Ortop Relat Res. 1981;157:191–7.
2. MacNab I. Negative disc exploration: an analysis of the causes of nerve root involvement in sixty-eight patients. J Bone Joint Surg Am. 1971;53(5):891–903.
3. Kornberg M. Extreme lateral lumbar disc herniations. Spine. 1987;12(6):586–9.
4. Kunogi J, Hasue M. Diagnosis and operative treatment of intraforaminal and extraforaminal nerve root compression. Spine. 1991;16(11):1312–30.
5. Cramer GD, Cantu JA, Dorsett RD, Greenstein JS, McGregor M, Howe JE, Glenn WV. Dimension of the lumbar intervertebral foramina as determined from the sagittal plane magnetic resonance imaging scans of 95 normal subjects. J Manipulative Physiol Ther. 2003;26(3):160–70.
6. Taira G, Endo K, Ito K. Diagnosis of lumbar disc herniation by three-dimensional MRI. J Orthop Sci. 1998;3(1):18–26.
7. Aota Y, Niwa T, Yoshikawa K, Fujiwara A, Asada T, Saito T. Magnetic resonance imaging and magnetic resonance myelography in the presurgical diagnosis of lumbar foraminal stenosis. Spine. 2007;32(8):896–903.
8. Lurie JD, Tosteson AN, Tosteson TD, Carragee E, Carrino J, Kaiser J, Blanco Sequeiros RT, Lecomte AR, Grove MR, Pearson LH, Weinstein JN, Herzog R. Reliability of readings of magnetic resonance imaging features of lumbar spinal stenosis. Spine. 2008;33(14):1605–10.
9. Byun WM, Jang HW, Kim SW. Three-dimensional magnetic resonance rendering imaging of lumbosacral radiculography in the diagnosis of symptomatic extraforaminal disc herniation with or without foraminal extension. Spine. 2012;37(10):840–4.
10. Jenis LG, An HS. Spine update: lumbar foraminal stenosis. Spine. 2000;25(3):389–94.
11. Baba H, Uchida K, Maezawa Y, Furusawa N, Okumura Y, Imura S. Microsurgical nerve root canal widening without fusion for lumbosacral intervertebral foraminal stenosis: technical notes and early results. Spinal Cord. 1996;34:644–50.
12. Wiltse LL, Guyer RD, Spencer CW. Alar transverse process impingement of the L5 spinal nerve: the far-out syndrome. Spine. 1984;9(1):31–41.
13. Olsewski JM, Simmons EH, Kallen FC, Mendel FC. Evidence from cadavers suggestive of entrapment of fifth lumbar spinal nerves by lumbosacral ligaments. Spine. 1991;16(3):336–47.
14. Transfeldt EE, Robertson D, Bradfold DS. Ligaments of the lumbosacral spine and their role in possible extraforaminal spinal nerve entrapment and tethering. J Spinal Disord. 1992;6(6):507–12.

15. Matsumoto M, Chiba K, Nojiri K, Ishikawa M, Toyama Y, Nishikawa Y. Extraforaminal entrapment of the fifth lumbar spinal nerve by osteophytes of the lumbosacral spine. Spine. 2002;27(6):E169–73.
16. Nathan H, Weizenbluth M, Halperin N. The lumbosacral ligament (LSL), with special emphasis on the "lumbosacral tunnel" and the entrapment of the 5th lumbar nerve. Int Orthop. 1982;6(3):197–202.
17. Raininko R, Manninen H, Battie MC, Gibbons LE, Gill K, Fisher LD. Observer variability in the assessment of disc degeneration on magnetic resonance images of the lumbar and thoracic spine. Spine. 1995;20(9):1029–35.
18. Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis: generalization to the population of readers and patients with the jackknife method. Invest Radiol. 1992;27(9):723–31.
19. Hills SL, Berbaum KS, Metz CE. Recent developments in the Dorfman–Berbaum–Metz procedure for multireader ROC study analysis. Acta Radiol. 2008;15(5):647–61.
20. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Hosmer DW, Lemeshow S, editors. Applied logistic regression. 2nd ed. New York: Wiley; 2000. p. 143–202.
21. Speciale AC, Pietrobon R, Urban CW, Richardson WJ, Helms CA, Major N, Enterline D, Hey L, Haglund M, Turner DA. Observer variability in assessing lumbar spinal stenosis severity on magnetic resonance imaging and its relation to cross-sectional spinal canal area. Spine. 2002;27(10):1082–6.
22. Obuchowski NA. ROC analysis. Am J Roentgenol. 2005;184(2):364–72.
23. Gur D. Technology and practice assessment: in search of a "desirable" statement. Radiology. 2005;234(3):659–60.
24. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am. 1990;72(3):403–8.
25. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, Takiguchi N, Minamide A, Oka H, Kawaguchi H, Nakamura K, Akune T, Yoshida M. Associations between radiographic lumbar spinal stenosis and clinical symptoms in the general population: the Wakayama Spine Study. Osteoarthr Cartil. 2013;21(6):783–8.
26. Ando M, Tamaki T, Kawakami M, Minamide A, Nakagawa Y, Maio K, Enyo Y, Yoshida M. Electrophysiological diagnosis using sensory nerve action potential for the intraforaminal and extraforaminal L5 nerve root entrapment. Eur Spine J. 2013;22(4):833–9.
27. Iwasaki H, Yoshida M, Yamada H, Hashizume H, Minamide A, Nakagawa Y, Kawai M, Tsutsui S. A new electrophysiological method for the diagnosis of extraforaminal stenosis at L5–S1. Asian Spine J. 2014;8(2):145–9.
28. Eguchi Y, Ohtori S, Yamashita M, Yamauchi K, Suzuki M, Orita S, Kamoda H, Arai G, Ishikawa T, Miyagi M, Ochiai N, Kishida S, Masuda Y, Ochi S, Kikawa T, Takaso M, Aoki Y, Toyone T, Suzuki T, Takahashi K. Clinical applications of diffusion magnetic resonance imaging of the lumbar foraminal nerve root entrapment. Eur Spine J. 2010;19(11):1874–82.

Research Report

Physical Performance Measures Associated With Locomotive Syndrome in Middle-Aged and Older Japanese Women

Misa Nakamura, PhD[1]; Hiroshi Hashizume, MD, PhD[2]; Hiroyuki Oka, MD, PhD[3]; Morihiro Okada, MD, PhD[1]; Rie Takakura, DPT[1]; Ayako Hisari, PT[1]; Munehito Yoshida, MD, PhD[2]; Hirotoshi Utsunomiya, DVM, PhD[4]

ABSTRACT

Background: The Japanese Orthopaedic Association proposed a concept called locomotive syndrome (LS) to identify middle-aged and older adults at high risk of requiring health care services because of problems with locomotion. It is important to identify factors associated with the development of LS. Physical performance measures such as walking speed and standing balance are highly predictive of subsequent disability and mortality in older adults. However, there is little evidence about the relationship between physical performance measures and LS.

Purpose: To determine the physical performance measures associated with LS, the threshold values for discriminating individuals with and without LS, and the odds ratio of LS according to performance greater than or less than these thresholds in middle-aged and older Japanese women.

Methods: Participants were 126 Japanese women (mean age = 61.8 years). Locomotive syndrome was defined as a score of 16 or more on the 25-question Geriatric Locomotive Function Scale. Physical performance was evaluated using grip strength, unipedal stance time with eyes open, seated toe-touch, and normal and fast 6-m walk time (6MWT). Variables were compared between LS and non-LS groups.

Results: Fourteen participants (11.1%) were classed as having LS. Unipedal stance time, normal 6MWT, and fast 6MWT were significantly different between the 2 groups. The LS group had a shorter unipedal stance time and a longer normal and fast 6MWT than the non-LS group. For these 3 variables, the area under the receiver operating characteristic curve was greater than 0.7, and the threshold for discriminating the non-LS and LS groups was 15 s for unipedal stance time, 4.8 s for normal 6MWT and 3.6 s for fast 6MWT. These variables were entered into a multiple logistic regression analysis, which indicated that unipedal stance time less than 15 s was significantly related to LS (odds ratio = 8.46; $P < .01$).

Conclusion: Unipedal stance time was the physical performance measure that was most strongly associated with LS. This measure may be useful for early detection of LS.

Key Words: GLFS-25, locomotive syndrome, physical performance, unipedal stance time

(*J Geriatr Phys Ther* 2015;38:202-207.)

INTRODUCTION

Japan is facing the advent of a super-aged society. The Japanese Statistics Bureau reported that, in 2013, people aged 65 years or older accounted for 24% of the Japanese population.[1] The number of individuals who require nursing care is increasing because of the increasing incidence of stroke, senility, falls or fractures, dementia, joint disorders, and other health issues.[2]

The Japanese Orthopaedic Association has proposed a concept called locomotive syndrome (LS) to identify middle-aged and older adults at high risk of requiring health care services because of problems with locomotion.[2] This syndrome is caused by reductions in muscular strength and balance that occur with aging and locomotive pathologies such as osteoporosis, osteoarthritis, and lumbar spinal stenosis.[2] In women, the syndrome may also be caused by reductions in physical activity and bone density that occur after menopause. To prevent LS, Nakamura[3] recommended that individuals avoid deterioration of the locomotive organs and the development of orthopedic problems and maintain or improve walking ability.

Recently, a quantitative and evidence-based screening tool called the 25-question Geriatric Locomotive Function Scale (GLFS-25) was developed to identify individuals with LS.[4] The GLFS-25 is a self-completed questionnaire that

[1]*Department of Rehabilitation, Osaka Kawasaki Rehabilitation University, Kaizuka, Osaka, Japan.*

[2]*Department of Orthopedic Surgery, Wakayama Medical University, Wakayama, Japan.*

[3]*Department of Joint Disease Research, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo, Japan.*

[4]*Department of Strategic Surveillance for Functional Food and Comprehensive Traditional Medicine, Wakayama Medical University, Wakayama, Japan.*

There are no potential conflicts of interest.

Address correspondence to: Misa Nakamura, PhD, Department of Rehabilitation, Osaka Kawasaki Rehabilitation University, 158 Mizuma, Kaizuka, Osaka 0104, Japan (nakamuram@kawasakigakuen.ac.jp).

Robert Wellman was the Decision Editor.

DOI: 10.1519/JPT.0000000000000033

consists of 25 items, including 4 questions on pain during the last month, 16 questions on activities of daily activity during the last month, 3 questions on social functions, and 2 questions on mental health status during the last month.[4] The 25 items are scored from 0 (no impairment) to 4 (severe impairment), and the total score ranges from 0 to 100 points. Higher scores indicate worse locomotive function, and the cutoff score for LS is 16 points, as determined by receiver operating characteristic (ROC) analysis. The validity of the GLFS-25 was confirmed by demonstrating significant correlation with the European Quality of Life Scale-5 Dimensions questionnaire.[4] Sasaki et al.[5] reported that the prevalence of LS was 21.2% in males and 35.6% in females and increased with age for both sexes. It is important to identify factors that are associated with the development of LS.

Physical function can predict future incidence of disability, dependence in activities of daily living, institutionalization, and death in initially nondisabled older adults.[6-13] A relatively strong relationship between general health and physical function in middle-aged and older adults is observed.[14] Several reports have shown that muscle strength, standing balance, and walking ability are key components of physical performance in older adults.[15,16] The Ministry of Education, Culture, Sports, Science, and Technology in Japan has proposed grip strength, coming up to a sitting position from a supine position, unipedal stance time with eyes open, seated toe-touch, 6-minute walk distance, and a 10-m obstacle walk time as variables that can be used to test the physical performance of older adults.[17] Physical performance measures such as walking speed, standing balance, and ability to repeatedly rise from a chair are highly predictive of subsequent disability.[9,12] In addition, hand grip strength is an important predictor of disability and mortality in older adults.[6,8,13]

Muramoto et al.[18,19] reported that the GLFS-25 score strongly correlated with several measures of physical performance. However, there is little evidence about the relationship between measures of physical performance and LS and little information on the threshold values of physical performance measures relevant to LS.[14,15] Furthermore, there are no reports on physical performance measures that are associated with the development of LS. In this study, we evaluated physical performance using grip strength, unipedal stance time with eyes open, seated toe-touch, and normal and fast 6-m walk time (6MWT). We hypothesized that these variables would be predictive of GLFS-25 score. The purpose of this study was to determine the physical performance measures associated with LS, the threshold values of these physical performance measures for discriminating individuals with and without LS, and the odds ratio of LS according to performance greater than or less than these thresholds in healthy middle-aged and older Japanese women.

METHODS

Participants

One hundred and thirty-four healthy adult Japanese women, who belonged to the Japan Agricultural Cooperative Kinan Women's Club and who attended a "Lecture meeting and checkup for health" supported by the local government in Tanabe, Wakayama, in 2013 volunteered to participate in this study. Eight women were excluded because they were unable to complete the physical performance test due to disabilities. The remaining 126 participants completed the physical performance tests and the GLFS-25 and were included in this study. Participants were aged from 34 to 84 years (mean age = 61.8 ± 10.2 years). All participants provided informed consent. The Ethics Committee for Human Research at Osaka Kawasaki Rehabilitation University approved the study protocol.

Outcome Measures

Status (presence and degree) of LS was evaluated using the GLFS-25.[4] Several measures of physical performance were obtained. Body weight was measured using a body weight meter (KaradaScan362, OMRON Co, Kyoto, Japan). Height was measured and used to calculate body mass index (BMI). Grip strength was measured bilaterally when the participant was in a standing position using a grip strength dynamometer (Appendix 1; T. K. K. 5401, Takei Scientific Instruments Co, Ltd, Niigata, Japan). Both hands were tested 2 times, and the maximum value was taken as grip strength. Unipedal stance time with eyes open was measured from the time when the foot was raised from the floor to the time when the foot was placed back on the floor, for a maximum of 60 s.[20] The time was measured for each lower limb, and the maximum value was taken as the unipedal stance time. Seated toe-touch was measured with the participant sitting on the floor with the knees extended. The participant was instructed to push their fingers forward and lean forward from the hip joint, and the distance reached by the fingertip was measured once using a digital long seat body anteflexion meter (Appendix 1; T.K.K. 5112, Takei Scientific Instruments Co, Ltd, Niigata, Japan). The 6MWT was measured as the time required to complete a 6-m walk in a straight line from a static start. The 6MWT test was performed once at normal walking speed and once as fast as possible.

Statistical Analysis

Participants were classified as LS (GLFS-25 score ≥16) or non-LS (GLFS-25 score <16), and independent variables were compared between the 2 groups. For numerical variables, normality of distribution and homogeneity of variance were tested prior to the comparison across groups. The Student t test was used when the assumptions of normal distribution and homogeneity of variance were met

in both groups, the Welch *t* test was used when the assumption of normal distribution was met but homogeneity of variance was not, and the Wilcoxon signed-rank test was used when the data were nonnormally distributed.

The threshold of each physical performance measure for discriminating the LS group and the non-LS group was evaluated using an ROC analysis. An area under an ROC curve (AUC) of 1.00 indicates perfect discrimination, whereas an AUC of 0.50 indicates complete absence of discrimination. Variables with AUC greater than 0.70[21] were included in a multiple logistic regression model to calculate the odds ratio of LS. Statistical analysis was conducted using JMP Pro 10.0 (SAS Institute Inc, Cary, NC). All statistical tests were 2-tailed, and a significance level of 0.05 was used.

Table 1. Characteristics of the Study Participants[a]

Variables	Mean Value (Standard Deviation)
Age, y	61.8 (10.2)
Height, cm	153.5 (5.8)
Weight, kg	54.9 (8.0)
BMI, kg/m^2	23.3 (3.1)
Grip strength, kg	26.2 (4.5)
Unipedal stance time, s	47.5 (19.6)
Seated toe-touch, cm	35.5 (9.7)
Normal 6-m walk time, s	4.5 (0.9)
Fast 6-m walk time, s	3.3 (0.6)
GLFS-25 score, points	7.6 (8.8)

[a]Data are mean ± standard deviation.

RESULTS

Age, BMI, each measure of physical performance, and the GLFS-25 score are shown in Table 1. Fourteen participants (11.1%) had a GLFS-25 score of 16 or more and were classed as LS. The LS group was older and had a higher BMI than the non-LS group. Significant differences between the 2 groups in unipedal stance time, normal 6MWT, and fast 6MWT were observed (Table 2). The LS group had a shorter unipedal stance time and a longer normal and fast 6MWT than the non-LS group. By contrast, there were no significant differences between the 2 groups in hand grip strength or seated toe-touch.

The ROC analysis was conducted for each physical performance measure, and the threshold corresponding to LS was identified. Unipedal stance time, normal 6MWT, and fast 6MWT had AUC greater than 0.7 in the ROC analysis. The threshold for discriminating the non-LS and LS groups was 15 s for unipedal stance time, 4.8 s for normal 6MWT, and 3.6 s for fast 6MWT. Figure 1 shows the ROC curves for grip strength, unipedal stance time, seated toe-touch, normal 6MWT, and fast 6MWT. The threshold values, AUC, sensitivity, and specificity are shown in Table 3. Unipedal stance time, normal 6MWT, and fast 6MWT were included in a multiple logistic regression model, which showed that unipedal stance time less than 15 s was significantly related to LS (odds ratio = 8.46; $P < .01$; Table 4).

DISCUSSION

The purpose of this study was to determine the physical performance measures associated with LS, the threshold values of these physical performance measures for discriminating individuals with and without LS, and the odds ratio of LS according to performance greater than or less than these thresholds in healthy middle-aged and older Japanese women. We found that the LS group had a shorter unipedal stance time and longer normal and fast 6MWT than the non-LS group. The threshold for discriminating the non-LS and LS groups was 15 s for unipedal stance time, 4.8 s for normal 6MWT, and 3.6 s for fast 6MWT. Moreover, we found that unipedal stance time less than 15 s was significantly related to LS, with an odds ratio of 8.46 identified by multiple logistic regression analysis.

Locomotive syndrome was proposed by the Japanese Orthopaedic Association in 2007 to identify individuals at high risk of requiring nursing care because of problems with locomotive organs.[2] The GLFS-25 has been developed to measure the presence and degree of LS in Japanese individuals.[4,18,19] However, the cutoff value to identify LS from the GLFS-25 score has been determined with reference to health-related quality of life, and limited information exists on the association of the GLFS-25 score with physical performance. Therefore, we examined the association between the presence of LS, as determined by the GLFS-25 score, and various measures of physical performance, including

Table 2. Comparison of Characteristics Between Nonlocomotive Syndrome and Locomotive Syndrome[a,b]

Variables	Non-LS (n = 112)	LS (n = 14)	*P*
Age, y	61.00 ± 9.96	68.21 ± 9.64	.0116[c]
BMI, kg/m^2	23.04 ± 2.92	25.49 ± 3.33	.0043[c]
Grip strength, kg	26.38 ± 4.63	24.43 ± 5.00	.1290[c]
Unipedal stance time, s	50.15 ± 16.88	26.42 ± 26.22	.0006[d]
Seated toe-touch, cm	35.67 ± 9.77	33.71 ± 9.36	.5141[d]
Normal 6-m walk time, s	4.45 ± 0.78	5.18 ± 1.31	.0122[d]
Fast 6-m walk time, s	3.27 ± 0.51	3.89 ± 0.66	.0004[d]

[a]Data are mean ± standard deviation.
[b]Locomotive syndrome: GLFS-25 score ≥16 points.
[c]The Student *t* test was applied for age, body mass index, and grip strength.
[d]The Wilcoxon signed-rank test was applied for unipedal stance time, seated toe-touch, normal 6-m walk time, and fast 6-m walk time.

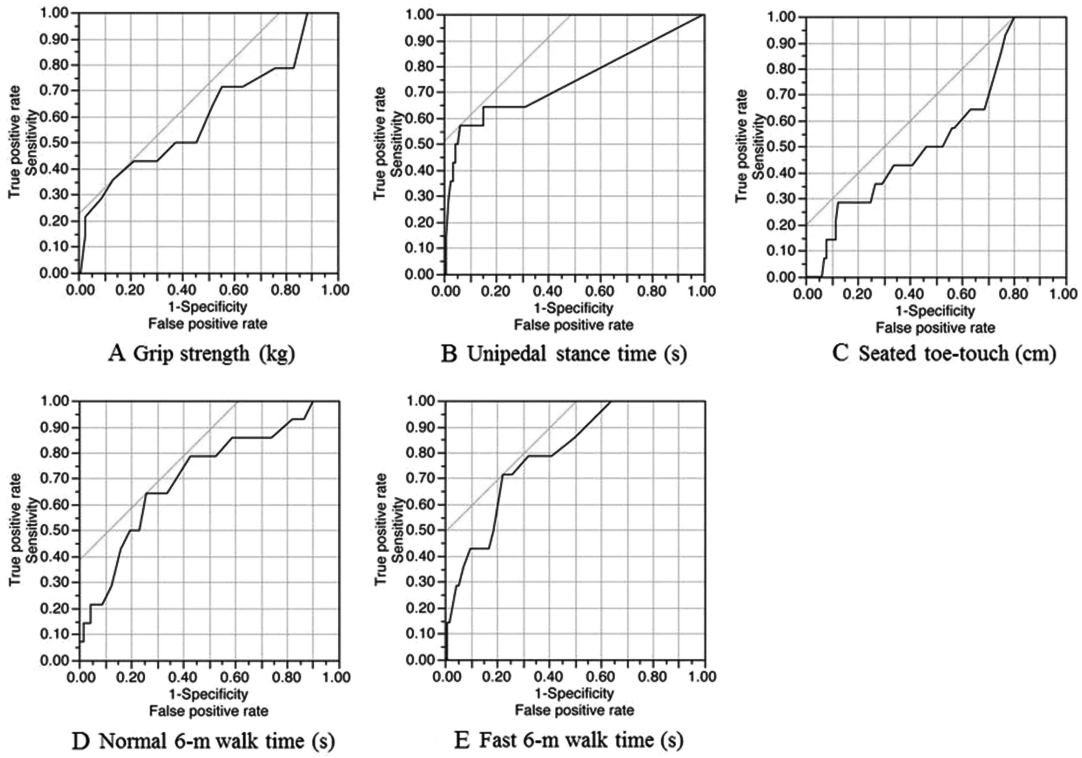


Figure 1. Receiver operating characteristic (ROC) curves showing the relationship between sensitivity and (1-specificity) for each measure of physical performance. In each plot, the light line represents a 45° tangent to the ROC curve and marks an optimal threshold where false negatives and false positives have similar costs. (A) ROC curve for grip strength. The area under the ROC curve was 0.606. The optimal threshold for detection of locomotive syndrome (LS) was 21 kg, with 35.7% sensitivity and 86.6% specificity. (B) ROC curve for unipedal stance time. The area under the ROC curve was 0.740. The optimal threshold for detection of LS was 15 s, with 57.1% sensitivity and 93.8% specificity. (C) ROC curve for seated toe-touch. The area under the ROC curve was 0.554. The optimal threshold for detection of LS was 43 cm, with 100% sensitivity and 19.6% specificity. (D) ROC curve for normal 6-m walk time. The area under the ROC curve was 0.706. The optimal threshold for detection of LS was 4.8 s, with 64.3% sensitivity and 74.1% specificity. (E) ROC curve for fast 6-m walk time. The area under the ROC curve was 0.792. The optimal threshold for detection of LS was 3.6 s, with 71.4% sensitivity and 77.7% specificity.

unipedal stance time with eyes open, grip strength, seated-toe touch, and normal and fast 6MWT. Our results showed that the LS group had a shorter unipedal stance time and longer normal and fast 6MWT than the non-LS group. The association between unipedal stance time and LS supports the findings of Muramoto et al.[18] The association between 6MWT and LS is consistent with previous reports that walking speed predicted the onset of functional decline in older adults living in rural Japan, and that there was a significant association between walking ability and disability in middle-aged and older adults.[4,22]

Previous reports have suggested that hand grip strength is a good indicator of overall muscle strength and an important predictor of disability and mortality in older adults.[23] The seated toe-touch test is a commonly used field measure of hamstring flexibility. Jones et al.[24] reported that a lack of hamstring flexibility was associated with low back pain, postural deviations, gait limitations, risk of falling, and straight leg raising range. However, in this study, we did not find a significant association between grip strength or seated toe-touch and LS. This may be due to a lack of statistical power, because the sample size of this study was small.

The threshold for discriminating between the non-LS and LS groups was 15 s for unipedal stance time, 4.8 s for normal 6MWT, and 3.6 s for fast 6MWT. Muramoto et al.[18] reported a threshold of 15 s for unipedal stance time, but this was the mean value from both lower limbs, in contrast with this study in which we adopted the higher value from the 2 lower limbs. Therefore, we cannot directly compare the threshold for unipedal stance time measured in this study with that reported by Muramoto et al.[18] Our calculation was more consistent with the standardized method recommended by the Japanese Orthopaedic Association.

Table 3. Threshold Values of the Physical Performance Tests for Locomotive Syndrome[a,b]

Physical Performance Tests	Threshold Values	Area Under the Curve	Sensitivity (%)	Specificity (%)
Grip strength (kg)	21	0.606	35.7	86.6
Unipedal stance time (s)	15	0.740	57.1	93.8
Seated toe-touch (cm)	43	0.554	100	19.6
Normal 6-m walk time (s)	4.8	0.706	64.3	74.1
Fast 6-m walk time (s)	3.6	0.792	71.4	77.7

[a]Data are mean ± standard deviation.
[b]Locomotive syndrome: GLFS-25 score ≥16 points.

The results of the multiple logistic regression analysis indicated that unipedal stance time less than 15 s was significantly associated with LS, with an odds ratio of 8.46. Unipedal stance time is used to evaluate both static and dynamic balance. Islam et al.[25,] reported that unipedal stance time was associated with activities of daily living.[26] Furthermore, Kikuchi et al.[27] reported that the fall-predicting score correlated with the duration of unipedal stance, as well as the timed up-and-go score, functional reach, grip strength, and tandem gait. In this study, LS was identified according to the results of a self-report questionnaire. The results of this study highlight the importance of measuring unipedal stance time in older adults, not only for predicting falls but also as a simple physical test to identify LS. Sakamoto et al.[28,29] reported that a unipedal stance exercise (the dynamic flamingo) reduced the incidence of falls and hip fracture. The Japanese Orthopaedic Association is now keen to identify exercises that are useful for preventing or improving LS.[3] It is possible that an interventional study using a unilateral stance exercise for preventing or improving LS will be conducted in the future.

This study has several limitations. First, the sample size of 126 was quite small. We did not find significant relationships between LS and handgrip strength or seated toe-touch, and this may be partly due to the lack of statistical power. Second, because the participants of this study were limited to women, the results cannot be generalized to men. Men should be included in future studies. We believe that this study contributes new information about the relationship between LS detected by self-report questionnaire and unipedal stance and has important implications for future studies.

Table 4. Multiple Logistic Regression Analysis for Odds Ratio and 95% Confidence Interval of the Physical Performance Tests for Locomotive Syndrome[a,b,c]

Physical Performance Tests	Odds Ratio	95% Confidence Interval	*P*
Unipedal stance time (s)	8.46	1.88-41.7	.0055
Normal 6-m walk time (s)	2.34	0.46-12.8	.3067
Fast 6-m walk time (s)	3.79	0.67-25.1	.1331

[a]Data are mean ± standard deviation.
[b]Data were adjusted by age and body mass index.
[c]Locomotive syndrome: GLFS-25 score ≥16 points.

CONCLUSION

Unipedal stance time was significantly associated with LS in middle-aged and older Japanese women. The threshold value for LS was 15 s. This measure may be useful for simple detection of LS.

ACKNOWLEDGMENTS

We thank Dr Hiroshi Kameda, Mr Shota Okumi, Mr Nobuhiro Koike, Mr Daiki Kanata, Mr Sho Tachibana, Mr Shodai Tanaka, Ms Sakiko Enomoto, Mr Takuma Nishimae, Mr Kenta Higashi, Mr Yusuke Saeki, Mr Ryosuke Hashikaku, Mr Yoshiki Kushi, Mr Taichi Takemoto, and Mr Yoshiki Kushi at Osaka Kawasaki Rehabilitation University for their cooperation.

REFERENCES

1. Statistics Bureau. Population estimates. http://www.stat.go.jp/english/data/jinsui/tsuki/index.Htm. Published July 2013. Accessed December 1, 2013.
2. Nakamura K. A "super-aged" society and the "locomotive syndrome". *J Orthop Sci.* 2008;3(1):1-2.
3. Nakamura K. The concept and treatment of locomotive syndrome: its acceptance and spread in Japan. *J Orthop Sci.* 2011;16(5):489-491.
4. Seichi A, Hoshino Y, Doi T, Akai M, Tobimatsu Y, Iwaya T. Development of a screening tool for risk of locomotive syndrome in the elderly: the 25-question Geriatric Locomotive Function Scale. *J Orthop Sci.* 2012;17(2):163-172.
5. Sasaki E, Ishibashi Y, Tsuda E, et al. Evaluation of locomotive disability using loco-check: a cross-sectional study in the Japanese general population. *J Orthop Sci.* 2013;18(1):121-129.
6. Shibata H, Haga H, Nagai H, et al. Predictors of all-cause mortality between ages 70 and 80: the Koganei study. *Arch Gerontol Geriatr.* 1992;14(3):283-297.
7. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol.* 1994;49(2):M85-M94.
8. Laukkanen P, Heikkinen E, Kauppinen M. Muscle strength and mobility as predictors of survival in 75-84-year-old people. *Age Ageing.* 1995;24(6):468-473.
9. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med.* 1995;332(9):556-561.
10. Gill TM, Williams CS, Tinetti ME. Assessing risk for the onset of functional dependence among older adults: the role of physical performance. *J Am Geriatr Soc.* 1995;43(6):603-609.
11. Sonn U. Longitudinal studies of dependence in daily life activities among elderly persons. *Scand J Rehabil Med Suppl.* 1996;34:1-35.
12. Ostir GV, Markides KS, Black SA, Goodwin JS. Lower body functioning as a predictor of subsequent disability among older Mexican Americans. *J Gerontol A Biol Sci Med Sci.* 1998;53(6):M491-M495.
13. Giampaoli S, Ferrucci L, Cecchi F, et al. Hand-grip strength predicts incident disability in non-disabled older men. *Age Ageing.* 1999;28(3):283-288.
14. Sato T, Demura S, Murase T, Kobayashi Y. Contribution of physical fitness component to health status in middle-aged and elderly males. *J Physiol Anthropol.* 2006;25(5):311-319.
15. Nagasaki H, Itoh H, Furuna T. The structure underlying physical performance measures for older adults in the community. *Aging (Milano).* 1995;7(6):451-458.
16. Ringsberg K, Gerdhem P, Johansson J, Obrant KJ. Is there a relationship between balance, gait performance and muscular strength in 75-year-old women? *Age Ageing.* 1999;28(3):289-293.

17. Ministry of Education, Culture, Sports, Science and Technology. New physical test implementation guidance. http://www.mext.go.jp/component/a_menu/sports/detail/__icsFiles/afieldfile/2010/07/30/ 1295079_04.pdf (Japanese). Accessed December 1, 2013.
18. Muramoto A, Imagama S, Ito Z, et al. Threshold values of physical performance tests for locomotive syndrome. *J Orthop Sci*. 2013;18(4):618-626.
19. Muramoto A, Imagama S, Ito Z, Hirano K, Ishiguro N, Hasegawa Y. Physical performance tests are useful for evaluating and monitoring the severity of locomotive syndrome. *J Orthop Sci*. 2012;17(6):782-788.
20. The Japanese Orthopaedic Association. Physical function test, single-leg stance time with eyes open 2288. http://www.joa.or.jp/jp/public/sick/condition/mads.html. Accessed February 1, 2013.
21. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Hosmer DW, Lemeshow S, eds. *Applied Logistic Regression*. 2nd ed. New York: Wiley; 2000:143-202.
22. Yoshimura N, Oka H, Muraki S, et al. Reference values for hand grip strength, muscle mass, walking time, and one-leg standing time as indices for locomotive syndrome and associated disability: the second survey of the ROAD study. *J Orthop Sci*. 2011;16(6):768-777.
23. Richards L, Palmiter-Thomas P. Grip measurement: a critical review of tools, methods, and clinical utility. *Crit Rev Phys Rehabil Med*. 1996;8: 87-109.
24. Jones CJ, Rikli RE, Max J, Noffal G. The reliability and validity of a chair sit-and-reach test as a measure of hamstring flexibility in older adults. *Res Q Exerc Sport*. 1998;69(4):338-343.
25. Islam MM, Nasu E, Rogers ME, Koizumi D, Rogers NL, Takeshima N. Effects of combined sensory and muscular training on balance in Japanese older adults. *Prev Med*. 2004;39(6):1148-1155.
26. Gawron W, Pospiech L, Orendorz-Fraczkowska K, Noczynska A. The influence of metabolic disturbances present in diabetes mellitus type I on vestiblo-spinal reflexes in children and young adults. *Otolaryngol Pol*. 2002;56(4):451-457.
27. Kikuchi R, Kozaki K, Iwata A, Hasegawa H, Toba K. Evaluation of risk of falls in patients at a memory impairment outpatient clinic. *Geriatr Gerontol Int*. 2009;9(3):298-303.
28. Sakamoto K, Nakamura T, Hagino H, et al.; Committee on Osteoporosis of The Japanese Orthopaedic Association. Effects of unipedal standing balance exercise on the prevention of falls and hip fracture among clinically defined high-risk elderly individuals: a randomized controlled trial. *J Orthop Sci*. 2006;11(5):467-472.
29. Sakamoto K, Endo N, Harada A, et al. Why not use your own body weight to prevent falls? A randomized, controlled trial of balance therapy to prevent falls and fractures for elderly people who can stand on one leg for ≤15 s. *J Orthop Sci*. 2013;18(1):110-120.

Appendix 1

Grip strength dynamometer (T.K.K. 5401, Takei Scientific Instruments Co, Ltd, Niigata, Japan); http://www.takei-si.co.jp/en/productinfo/detail/49.html

Digital long seat body anteflexion meter (T.K.K. 5112, Takei Scientific Instruments Co, Ltd, Niigata, Japan); http://store.shopping.yahoo.co.jp/hakaronet/training013.html

Original Article
EPIDEMIOLOGY/GENETICS

Obesity

Urinary 8-Iso-Prostaglandin F2α as a Marker of Metabolic Risks in the General Japanese Population: The ROAD Study

Kanae Mure[1], Noriko Yoshimura[2], Marowa Hashimoto[1], Shigeyuki Muraki[3], Hiroyuki Oka[2], Sakae Tanaka[4], Hiroshi Kawaguchi[5], Kozo Nakamura[6], Toru Akune[6], and Tatsuya Takeshita[1]

Objective: To determine whether 8-iso-prostaglandin F2α (8-iso-PGF2α) is a reliable biomarker of the accumulation of metabolic risks [e.g., overweight, hypertension, impaired glucose tolerance (IGT), and dyslipidemia].
Methods: This was a cross-sectional study of the baseline characteristics of a Japanese general population cohort study: Research on Osteoarthritis/Osteoporosis Against Disability (ROAD). Of 1,690 participants, 1,527 fulfilled all questionnaires and examinations. Free and conjugated urinary 8-iso-PGF2α levels and metabolic syndrome (MetS) components including blood pressure, HbA1c, total cholesterol, high-density lipoprotein cholesterol (HDL-C), and non-HDL-C were analyzed. The data were analyzed by ANCOVA, multiple regression analysis, and multinomial logistic analysis.
Results: 8-iso-PGF2α was significantly associated with HbA1c and significantly inversely associated with total cholesterol and non-HDL-C. Notably, IGT with an HbA1c cut-off of 5.5% was significantly associated with 8-iso-PGF2α level in participants aged ≤50 years. Multinomial logistic regression analysis revealed 8-iso-PGF2α level was significantly associated with a greater number of MetS risks present; this association was stronger in younger participants. In participants aged ≥71 years, 8-iso-PGF2α was significantly associated with a greater number of MetS risks with higher IGT cut-offs.
Conclusions: Urinary 8-iso-PGF2α can be a reliable marker of IGT and the accumulation of MetS risks, especially in younger people.

Obesity (2015) **23**, 1517–1524. doi:10.1002/oby.21130

Introduction

Metabolic syndrome (MetS) is a clustering of abdominal obesity, hypertension, dyslipidemia, and impaired glucose tolerance (IGT) and is strongly associated with the development of cardiovascular disease, diabetes mellitus, and atherosclerosis (1,2). The prevalence of MetS is increasing rapidly worldwide. Insulin resistance is the major underlying pathology, while oxidative stress is considered to be important as well (3,4).

Oxidative stress is broadly defined as an imbalance of free radicals and other reactive oxygen species or reactive nitrogen species and antioxidant defense systems (5). These reactive species can be

[1] Department of Public Health, Wakayama Medical University School of Medicine, Wakayama, Japan. Correspondence: Kanae Mure (kana@wakayama-med.ac.jp) [2] Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan [3] Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan [4] Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan [5] JCHO Tokyo Shinjuku Medical Center, Tokyo, Japan [6] National Rehabilitation Center for Persons with Disabilities, Saitama, Japan.

Funding agencies: This work was supported by Grants-in-Aid for Scientific Research to NY (B26293139, B23390172, B20390182, and Challenging Exploratory Research 24659317), TA (C20591737, B23390357), and SM (C20591774, B23390356, and Challenging Exploratory Research 23659580); for Young Scientists to HO (A18689031 and Challenging Exploratory Research 24659666); and for Collaborating Research with the National Science Foundation to NY (Director; 08033011-00262) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan as well as H17-Men-eki-009 (Director, KN), H18-Cho-ju-037 (Director, TN), H20-Cho-ju-009 (Director, NY), H23-Cho-ju-002 (Director, TA), H25-Nanchi-to (Men)-005 (Director, ST), and H25-Cho-ju-007 (Director, NY) from the Ministry of Health, Labour, and Welfare of Japan. This study was also supported by grants from the Japan Osteoporosis Society (NY, SM, HO, and TA) and by research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 and 2010-2; Director, HK).
Disclosure: The authors declared no conflict of interest.
Author contributions: KM and NY conceived the study. NY conceptualized the study, was primarily responsible for developing the protocol, and acts as the guarantor for this study. KM wrote the manuscript. KM and MH carried out experiments. SM, HO, ST, HK, KN, and TA collected the data. TT supported statistical analyses. All authors had final approval of the submitted and published versions.
Additional Supporting Information may be found in the online version of this article.
Received: 30 October 2014; **Accepted:** 3 April 2015; **Published online 7 June 2015. doi:10.1002/oby.21130**

generated by exogenous environmental sources such as smoking or endogenously during weight gain. They damage nucleic acids, proteins, and membrane lipids, leading to cell death (6). The consequences of these cellular modifications caused by oxidative stress are considered to be important in the pathogenesis of metabolic, inflammatory, and degenerative diseases including diabetes mellitus, atherosclerosis, and MetS (7). Thus, the assessment of oxidative stress levels in body fluids can be invaluable for predicting the progression of these diseases. However, as free radicals, reactive oxygen species, and reactive nitrogen species are highly reactive and short-lived, it is almost impossible to directly measure them.

F_2-isoprostanes are prostaglandin-like substances produced by the attack of free radicals on arachidonic acid (8). They are very stable, robust, and detectable molecules in body fluids such as plasma, bile, bronchial lavage fluid, cerebrospinal fluid, and urine. According to the Biomarkers of Oxidative Stress Study organized by the National Institutes of Health, isoprostanes are the most accurate and reliable markers of oxidative stress status *in vivo* (9). In particular, 8-iso-prostaglandin F2α (8-iso-PGF2α) is the best studied F_2-isoprostane. It appears to initially form on phospholipids and is subsequently hydrolyzed into free forms mainly by the platelet-activating factors acetylhydrolase enzymes in plasma before being rapidly excreted in urine (10). Urinary 8-iso-PGF2α is chemically very stable and is considered not to be formed artifactually in urine. Elevated levels of urinary 8-iso-PGF2α have been reported in metabolic, inflammatory, and degenerative diseases including type 2 diabetes, cardiovascular disease, and coronary heart diseases (11). Despite the rather larger number of studies assessing the association of MetS with oxidative stress, the direct association between 8-iso-PGF2α and MetS is seldom studied; there have been only 3 studies so far: 2 of them used univariate analyses and report conflicting results (12,13), while only 1, which involved women, showed a significant association in multivariate analysis (14). On the other hand, 8-iso-PGF2α level is reported to be significantly associated with insulin resistance in individuals with normal or prediabetic glucose levels independent of MetS (15).

Therefore, this study investigated whether urinary 8-iso-PGF2α is a reliable biomarker of the accumulation of MetS risks. Recent studies including our own have used elevated HbA1c as a criterion for IGT in different races because it reflects the average of blood glucose level for several months and its measurement does not require fasting (16-18). Meanwhile, adapting the same MetS criteria to different age groups is controversial (1,19). In addition, it has been suggested that the diagnostic cut-off of HbA1c should consider age (20). Thus, the present study evaluated different HbA1c levels as a criterion for IGT to clarify the optimal cut-off for identifying oxidative stress-related IGT in different age groups (21).

Methods

Participants

This study involved the cohorts established in 2005 for the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study. Details of the cohort profile have been reported elsewhere (22). Briefly, from 2005 to 2007, a baseline database including clinical information for 3,040 Japanese residents (men, 1,061; women, 1,979) was created. The participants were recruited from resident registration listings in two communities with different characteristics: 864 individuals (men, 319; women, 545) from a mountainous region in Hidakagawa, Wakayama, and 826 individuals (men, 277; women, 549) from a coastal region in Taiji, Wakayama. They were invited for the second survey of the ROAD study, which was a 3-year follow-up examination identical to the baseline examination. For this study, we enrolled all 1690 participants (men, 596; women, 1,094) who resided in the mountainous and coastal areas and underwent blood and urine examinations at baseline. Of the 1,690 participants, 1,527 who fulfilled all questionnaires and examinations were included. Written informed consent was provided by all participants, and the study was approved by the Ethics Committee of the University of Tokyo.

Baseline examination procedures

At the baseline examination, participants completed a 400-item interviewer-administered questionnaire that included lifestyle information such as alcohol consumption, smoking, and walking habits, and medical history. The participants were asked whether they took prescription medication daily or nearly every day (0: no, 1: yes). If they did not know what their medications were prescribed for, they were asked to bring their medications to the medical doctor (NY).

Anthropometric parameters and laboratory measurements

Age was classified as $\leq$50, 51–70, or $\geq$71 years. Anthropometric measurements including height and weight were measured on site to calculate body mass index [BMI: weight/height^2 (kg/m^2)]. Systolic (SBP) and diastolic (DBP) blood pressure were also measured on site by an experienced public health nurse. Blood samples were collected between 09:00 and 15:00. HbA1c (National Glycohemoglobin Standardization Program, NGSP), total cholesterol (T-chol), and high-density lipoprotein cholesterol (HDL-C) were measured at the same laboratory within 24 h of extraction (Osaka Kessei Research Laboratories, Osaka, Japan). Non-HDL-C was calculated by subtracting HDL-C from T-chol.

Definition of MetS risks

In the present study, hypertension (SBP $\geq$ 130 mmHg and/or DBP $\geq$ 85 mmHg), IGT (HbA1c >5.5%, $\geq$5.7%, $\geq$6.0%, or $\geq$6.5%, hereafter referred to as IGT criteria 1–4, respectively) and dyslipidemia (HDL < 40 mg/dl for men, <50 mg/dl for women) were evaluated as MetS components. Furthermore, taking antihypertensive, antihypercholesterolemic, or antidiabetic medications was counted as a MetS risk. The cut-off of a high risk of obesity was BMI $\geq$ 27.5 kg/m^2 according to the WHO Expert Consultation's recommendation for determining obesity in Asian populations (23).

Urinary 8-iso-PGF2α measurements

Gas chromatograph mass spectrometry (GC/MS) is a very accurate method of analyzing 8-iso-PGF2α; however, it requires time and is expensive. In contrast, the immunoassays have less specificity and quantitative ability as drawbacks but have large sample capacity, greater cost-effectiveness, and are strongly correlated with GC/MS results [$r = 0.63$, $P < 0.02$; (24)]. Furthermore, until recently, 8-iso-PGF2α in urine was considered to be the free form; however, approximately 40% of 8-iso-PGF2α in urine was recently reported to be conjugated as glucuronides (25). Therefore, we performed the

TABLE 1 Baseline participant characteristics

Parameter	Total	≤50 years	51-70 years	≥71 years	*P* for trend
***n* (women %)**		219 (68.0)	760 (66.6)	548 (60.4)	
Age (years)	64.7 ± 12.1	43.8 ± 5.6	62.0 ± 5.6	76.8 ± 4.4	<0.001[a]
BMI (kg/m^2)	23.0 ± 3.4	22.5 ± 3.9	23.3 ± 3.1	22.8 ± 3.4	0.964
SBP (mmHg)	134.9 ± 20.4	120.3 ± 19.1	134.8 ± 19.4	140.8 ± 19.4	<0.001
DBP (mmHg)	74.3 ± 11.3	71.4 ± 12.3	76.1 ± 10.9	73.1 ± 11.0	0.898
HbA1c (%) NGSP	5.56 ± 0.75	5.25 ± 0.60	5.59 ± 0.79	5.64 ± 0.72	<0.001
T-chol (mg/dl)	209.2 ± 34.7	209.5 ± 35.0	214.6 ± 35.4	201.6 ± 32.2	<0.001
HDL-C (mg/dl)	61.2 ± 15.8	66.7 ±18.0	61.5 ± 15.6	58.7 ± 14.4	<0.001
Non-HDL-C (mg/dl)	147.9 ± 33.8	142.9 ± 35.9	153.1 ± 34.3	142.9 ± 31.1	0.060
8-iso-PGF2α (ng/ml/M creatinine)	19.6 ± 21.7	20.4 ± 16.2	19.2 ± 18.1	19.8 ± 27.4	0.925
Prevalence of selected characteristics (%)					
Alcohol consumption[b]	606 (39.7)	110 (50.2)	316 (41.6)	180 (32.9)	<0.001[c]
Smoking habit[d]	198 (13.0)	49 (22.4)	93 (12.2)	56 (10.2)	<0.001
Walking habit[e]	1228 (80.4)	166 (75.8)	608 (80.0)	454 (82.9)	0.078
Prevalence of MetS risks (%)					
MetS I[f] **(≥3)**	206 (13.5)	12 (5.5)	105 (13.8)	89 (16.2)	<0.001
MetS II[g] **(≥3)**	153 (10.2)	10 (4.6)	73 (9.6)	70 (12.8)	0.003
MetS III[h] **(≥3)**	112 (7.3)	9 (4.1)	47 (6.2)	56 (10.2)	0.003
MetS IV[i] **(≥3)**	82 (5.4)	8 (3.7)	38 (5.0)	36 (6.6)	0.220

Data are mean ± SD.
[a]*P* for trend.
[b]Current alcohol consumption (regularly, ≥1 month).
[c]χ^2 test.
[d]Current smoking habit (regularly, ≥1 month).
[e]Regularly walking outside (≥5 times/week, including job).
[f]MetS risk I: sum of overweight (BMI ≥27.5 kg/m^2), hypertension (SBP ≥130 mmHg and/or DBP ≥85 mmHg or on medication), dyslipidemia (HDL <40 mg/dl for men and <50 mg/dl for women or on medication), and IGT criteria 1 (HbA1c ≥5.5% or on medication).
[g]MetS risk II: sum of overweight, hypertension, dyslipidemia, and IGT criteria 2 (HbA1c ≥5.7% or on medication).
[h]MetS risk III: sum of overweight, hypertension, dyslipidemia, and IGT criteria 3 (HbA1c ≥6.0% or on medication).
[i]MetS risk IV: sum of overweight, hypertension, dyslipidemia, and IGT criteria 4 (HbA1c ≥6.5% or on medication).

immunoassay with β-glucuronidase to assess both the free and conjugated forms of 8-iso-PGF2α in urine.

Urine samples were collected at the same time as blood samples. After centrifugation at 1,500 × *g* at 4°C for 15 min, samples were immediately placed on dry ice and stored at −80°C until the assay. After thawing on ice, the samples were centrifuged briefly at 1,500 × *g* at 4°C for 15 min to remove debris. Urinary 8-iso-PGF2α level was measured using an enzyme-linked immunosorbent assay kit (#EA85, Oxford Biomedical Research, Rochester Hills, MI, USA). Prior to the assay, sample was treated with β-glucuronidase provided in the kit. According to the manufacturer, the correlation (R^2) of this immunoassay with GC/MS was >0.8. Urinary creatinine level was measured using a colorimetric assay kit purchased from Cayman Chemical (Ann Arbor, MI). Urinary 8-iso-PGF2α level was corrected by creatinine clearance and is expressed as ng ml^{-1} creatinine (M)$^{-1}$.

Statistical analysis

All statistical analyses were performed using STATA version 13 (STATA Corp, College Station, TX). Differences in proportions were compared using the chi-square test. The associations of lifestyle factors with 8-iso-PGF2α level were analyzed by ANCOVA with a Bonferroni *post hoc* test after adjusted for age, BMI, and sex. Multivariate regression analyses were performed to study the associations between 8-iso-PGF2α level and each component of MetS risk after adjusting for age, BMI, and sex as well as current alcohol consumption, smoking, and walking habits. The associations between each MetS risk factor and 8-iso-PGF2α levels were analyzed by ANCOVA with a Bonferroni *post hoc* test adjusting for the same confounders listed above. Multinomial logistic regression analysis was used to test the association between 8-iso-PGF2α level and the presence of MetS risks. The number of MetS risks was the dependent variable (0: no MetS risks, 1: 1 MetS risk, 2: 2 MetS risks, 3: ≥3 MetS risks) and age (+1 year), BMI (+1 kg/m^2), sex (0: men, 1: women), current alcohol consumption (0: ex- or nondrinker, 1: current drinker), smoking habit (0: ex- or nonsmoker, 1: current smoker), and walking habit (0: regularly walking ≥5 times/week, 1: <5 times/week including on job) were independent variables. After the analysis, the relative risk ratios (RRRs) were evaluated. The level of significance was set at $P < 0.05$.

Results

The baseline characteristics of subjects in the ROAD study are shown in Table 1. SBP, HbA1c, and lipid panels were significantly associated with age (*P* for trend <0.001), whereas 8-iso-PGF2α level showed no association with age (Supporting Information Table

TABLE 2 Associations of lifestyle factors with urinary 8-iso-PGF2α level

	Regularly	No	*P*
Total			
Current alcohol consumption[a]	19.7 (0.9)[b]	19.6 (0.7)	0.924
n	606	921	
Current smoking habit[c]	23.3 (1.7)	19.1 (0.6)	0.018
n	198	1329	
Regular walking habit[d]	19.1 (0.6)	21.8 (1.3)	0.055
n	1228	299	
≤50 years			
Current alcohol consumption	17.1 (4.2)	19.6 (4.4)	0.284
n	110	109	
Current smoking habit	20.5 (4.8)	17.2 (4.2)	0.277
n	49	170	
Regular walking habit	18.0 (4.2)	16.7 (4.7)	0.617
n	166	53	
51-70 years			
Current alcohol consumption	19.3 (1.2)	18.5 (0.9)	0.578
n	316	444	
Current smoking habit	24.9 (2.1)	18.0 (0.8)	0.002
n	93	667	
Regularly walking habit	18.1 (0.8)	22.5 (1.6)	0.007
n	608	152	
≥71 years			
Current alcohol consumption	24.6 (3.8)	24.3 (3.6)	0.891
n	180	368	
Current smoking habit	24.0 (4.8)	24.5 (3.5)	0.896
n	56	492	
Regularly walking habit	24.0 (3.5)	26.1 (4.3)	0.505
n	454	94	

[a]Current alcohol consumption (regularly, ≥1 month).
[b]Data are estimated mean (SE) adjusted for age, BMI, and sex in ANCOVA.
[c]Current smoking habit (regularly, ≥1 month).
[d]Regular outdoor walking (≥5 times/week, including job).

S1). The prevalence of MetS risks was also significantly associated with age except that with the highest IGT criteria (≥6.5%).

The associations of selected lifestyle factors and 8-iso-PGF2α level are shown in Table 2. Although current smoking habit was significantly associated with higher 8-iso-PGF2α level ($P = 0.018$) and regular walking habit showed a trend of being inversely associated with 8-isoprotane level ($P = 0.055$), these associations were only significant in participants aged 51–70 years ($P = 0.002$ and $P = 0.007$, respectively).

The associations of 8-iso-PGF2α level with components of MetS risks are shown in Table 3. HbA1c was significantly associated with higher 8-iso-PGF2α level ($P = 0.003$), whereas T-chol and non-HDL-C showed significant inverse associations ($P = 0.013$ and 0.004, respectively). In participants aged ≤50 years, systolic and DBP, and HbA1c were significantly associated with higher 8-iso-PGF2α level ($P = 0.002$, 0.031, and <0.001, respectively). In participants aged 51–70 years, non-HDL-C was significantly inversely associated with 8-iso-PGF2α level ($P = 0.047$), and DBP and T-chol showed trends of being inversely associated with 8-iso-PGF2α level ($P = 0.059$ and 0.079, respectively). In participants aged ≥71 years, HbA1c was significantly associated with higher 8-iso-PGF2α level ($P = 0.030$), whereas non-HDL-C was significantly inversely associated ($P = 0.026$).

The associations of 8-iso-PGF2α level with each MetS risk factor are shown in Table 4. Only IGT criterion 4 (≥6.5%, the cut-off of diabetes) was significantly associated with 8-iso-PGF2α level ($P = 0.001$). Notably, in participants aged ≤50 years, 8-iso-PGF2α level was significantly associated with almost all MetS risk factors except dyslipidemia. In particular, IGT criterion 1 (>5.5%) was significantly associated with 8-iso-PGF2α level ($P = 0.020$), while IGT criteria 3 and 4 (≥6.0% and ≥6.5%, respectively) were strongly associated with 8-iso-PGF2α (both $P < 0.001$). However, in participants aged 51–70 years, there was no significant association between 8-iso-PGF2α level and any MetS risk factor ($P > 0.05$). In participants aged ≥71 years, 8-iso-PGF2α level was significantly associated with IGT criterion 4 ($P = 0.006$).

Multinomial logistic regression analyses were performed to study the associations of 8-iso-PGF2α level and the accumulation of MetS risk factors according to different IGT criteria among the 3 age groups (Table 5). Participants with higher MetS risks had a significantly higher RRR than those with no MetS risks with respect to all MetS risk criteria and IGT criteria (RRR = 1.01 with ≥3 MetS risks in all MetS risk criteria, $P = 0.041$, 0.046, 0.046, and 0.005, respectively). These associations were also significant in participants aged ≤50 years (RRR = 1.04 with ≥3 MetS risks in all MetS risk criteria, $P = 0.017$, 0.023. 0.025, and 0.010, respectively), but not in participants aged 51–70 years. In participants aged ≥71 years, significant associations or near-significant trends were observed between 8-iso-PGF2α levels and the accumulation of MetS risks with IGT criterion 4 (≥6.5%; RRR = 1.02, $P = 0.038$).

Discussion

To our knowledge, this is the first study to clarify the associations of both free and conjugated 8-iso-PGF2α in urine and MetS risks in a general population. A novel finding is 8-iso-PGF2α level is significantly associated with the accumulation of MetS risk with respect to age in a general population. In particular, 8-iso-PGF2α level reflected IGT with great sensitivity even just above the normal HbA1c cut-off (>5.5%) in participants aged ≤50 years, suggesting it is a reliable marker of IGT before establishing further risks in younger people. This finding also suggests the most widely used HbA1c cut-off might screen out younger participants at risk.

The community-based Framingham Study revealed that 8-iso-PGF2α level is significantly associated with insulin resistance and prediabetes assessed by fasting plasma glucose and insulin level (15). In this study, 8-iso-PGF2α was significantly associated with IGT with HbA1c at just above the normal cut-off (>5.5%), which is lower than the prediabetic cut-off (5.7%). To our knowledge, no study has investigated the direct association between 8-iso-PGF2α level and nondiabetic HbA1c levels. Both type 1 and 2 diabetes patients have elevated 8-iso-PGF2α levels (26), corroborating the

TABLE 3 Associations of urinary 8-iso-PGF2α level with each component of MetS risk

	β	P
Total (n = 1527)		
SBP (mmHg)	0.021	0.375
DBP (mmHg)	−0.006	0.795
HbA1c (%)	0.074	0.003
T-chol (mg/dl)	−0.063	0.013
HDL-C (mg/dl)	0.019	0.430
non-HDL-C (mg/dl)	−0.073	0.004
≤50 years (n = 219)		
SBP (mmHg)	0.192	0.002
DBP (mmHg)	0.132	0.031
HbA1c (%)	0.261	<0.001
T-chol (mg/dl)	−0.065	0.346
HDL-C (mg/dl)	−0.033	0.578
Non-HDL-C (mg/dl)	−0.047	0.472
51-70 years (n = 760)		
SBP (mmHg)	−0.053	0.128
DBP (mmHg)	−0.066	0.059
HbA1c (%)	0.032	0.379
T-chol (mg/dl)	−0.062	0.079
HDL-C (mg/dl)	0.016	0.641
Non-HDL-C (mg/dl)	−0.071	0.047
≥71 years (n = 548)		
SBP (mmHg)	0.048	0.244
DBP (mmHg)	0.010	0.803
HbA1c (%)	0.090	0.030
T-chol (mg/dl)	−0.070	0.098
HDL-C (mg/dl)	0.044	0.291
Non-HDL-C (mg/dl)	−0.093	0.026

β: standardized coefficient adjusted for age, BMI, and sex as well as current alcohol consumption, smoking, and walking habits in multivariate regression analyses.

significant positive association between 8-iso-PGF2α level and IGT with HbA1c at the diagnostic cut-off (≥6.5%) in all participants in the present study. Such a significant association was also seen in participants aged ≤50 and ≥71 years but not in those aged 51–70 years. Furthermore, in the study with type 1 and 2 diabetes patients mentioned above, insulin treatment decreased 8-iso-PGF2α levels, whereas oral hypoglycemic agents did not (26). The proportion of participants aged 51–70 years with HbA1c at the diagnostic cut-off on antidiabetic medication (57.4%) was higher than that in participants aged ≤51 and ≥71 years (25 and 55.2%, respectively, data not shown). It is possible that the proportion of participants on insulin therapy was higher in the participants aged 51–70 years, resulting lower 8-iso-PGF2α levels. However, the differences were not statistically significant, and details about their medications are unavailable. Therefore, this is merely speculation until confirmation from further investigation.

The associations between oxidative stress and other components of MetS risk are well studied. Overweight assessed by BMI was significantly associated with urinary 8-iso-PGF2α level in several studies including the Framingham study (27). Visceral fat accumulation is also positively associated with 8-iso-PGF2α (13). In this study, 8-iso-PGF2α level was significantly correlated with BMI in the total participants and those aged ≤50 and 51–70 years ($R^2 = 0.06$, $P = 0.015$; $R^2 = 0.17$, $P = 0.012$; $R^2 = 0.07$, $P = 0.043$; respectively) but not in participants aged ≥71 years ($R^2 = 0.03$, $P = 0.456$; Supporting Information Table S1). These results are corroborated by a recent study with older participants aged 70–79 years, in which 8-isoprostane was not associated with BMI (28).

In this study, 8-iso-PGF2α level was not associated with HDL-C; instead, it was significantly inversely associated with T-chol and non-HDL-C; this is concordant with other studies involving

TABLE 4 Associations of urinary 8-iso-PGF2α and MetS risk factors

	Yes	No	P
Total			
Hypertension[a]	19.0 (1.4)[b]	17.9 (1.6)	0.392
Dyslipidemia[c]	17.6 (1.8)	18.9 (1.4)	0.390
IGT 1[d]	19.1 (1.5)	18.2 (1.5)	0.464
IGT 2[e]	19.2 (1.6)	18.4 (1.4)	0.566
IGT 3[f]	20.5 (1.9)	18.2 (1.4)	0.152
IGT 4[g]	24.1 (2.1)	17.7 (1.4)	0.001
≤50 years			
Hypertension	25.0 (5.2)	20.9 (5.2)	0.095
Dyslipidemia	28.1 (6.2)	22.4 (5.1)	0.152
IGT 1	26.2 (5.2)	19.4 (5.3)	0.020
IGT 2	27.4 (5.6)	20.9 (5.2)	0.068
IGT 3	38.7 (6.6)	20.8 (5.0)	<0.001
IGT 4	50.6 (8.5)	22.3 (4.9)	<0.001
51-70 years			
Hypertension	15.6 (1.7)	17.1 (2.0)	0.314
Dyslipidemia	15.4 (2.1)	16.2 (1.7)	0.617
IGT 1	17.0 (1.8)	14.9 (1.8)	0.110
IGT 2	16.6 (1.9)	15.7 (1.8)	0.570
IGT 3	15.6 (2.3)	16.2 (1.7)	0.773
IGT 4	17.1 (2.6)	15.8 (1.7)	0.600
≥71 years			
Hypertension	25.0 (4.3)	20.9 (4.8)	0.206
Dyslipidemia	20.9 (4.9)	24.3 (4.2)	0.246
IGT 1	22.4 (4.4)	24.7 (4.3)	0.357
IGT 2	23.6 (4.5)	23.8 (4.3)	0.917
IGT 3	26.2 (4.7)	22.8 (4.3)	0.266
IGT 4	31.6 (5.0)	21.9 (4.2)	0.006

[a]Hypertension (SBP ≥130 mmHg and/or DBP ≥85 mmHg or on medication).
[b]Estimated mean (SE) adjusted for age and BMI as well as current alcohol consumption, smoking, and walking habits in ANCOVA.
[c]Dyslipidemia (HDL <40 mg/dl for men and <50 mg/dl for women or on medication).
[d]IGT criteria 1 (HbA1c >5.5% or on medication).
[e]IGT criteria 2 (HbA1c ≥5.7% or on medication).
[f]IGT criteria 3 (HbA1c ≥6.0% or on medication).
[g]IGT criteria 4 (HbA1c ≥6.5% or on medication).

TABLE 5 Associations of urinary 8-iso-PGF2α with the accumulation of MetS risks

	MetS risk I[a]			MetS risk II[b]			MetS risk III[c]			MetS risk IV[d]		
	RRR	95% CI	*P*	RRR	95% CI	*P*	RRR	95% CI	*P*	RRR	95% CI	*P*
Total												
0 reference												
1	1.01	1.00-1.02	0.026	1.01	1.00-1.02	0.025	1.01	1.00-1.02	0.056	1.01	1.00-1.02	0.035
2	1.01	1.00-1.02	0.052	1.01	1.00-1.02	0.150	1.01	1.00-1.02	0.164	1.01	1.00-1.02	0.180
≥3	1.01	1.00-1.02	0.041	1.01	1.00-1.02	0.046	1.01	1.00-1.02	0.046	1.01	1.00-1.02	0.005
≤50 years												
0 reference												
1	0.99	0.97-1.02	0.619	1.01	0.99-1.03	0.498	1.01	0.99-1.03	0.317	1.02	1.00-1.04	0.108
2	1.03	1.01-1.06	0.019	1.02	0.99-1.05	0.109	1.03	1.00-1.06	0.035	1.02	0.98-1.05	0.288
≥3	1.04	1.01-1.07	0.017	1.04	1.01-1.07	0.023	1.04	1.00-1.07	0.025	1.04	1.01-1.08	0.010
51-70 years												
0 reference												
1	1.01	0.99-1.02	0.372	1.00	0.99-1.01	0.538	1.00	0.99-1.01	0.956	1.00	0.99-1.01	0.829
2	1.01	1.00-1.02	0.178	1.00	0.99-1.01	0.917	1.00	0.98-1.01	0.499	1.00	0.98-1.01	0.518
≥3	1.00	0.99-1.02	0.654	1.00	0.99-1.02	0.712	1.00	0.99-1.02	0.796	1.01	0.99-1.02	0.513
≥71 years												
0 reference												
1	1.02	0.99-1.04	0.143	1.02	1.00-1.04	0.101	1.02	1.00-1.04	0.075	1.02	1.00-1.04	0.071
2	1.01	0.99-1.03	0.433	1.01	0.99-1.04	0.229	1.02	1.00-1.04	0.113	1.02	1.00-1.04	0.109
≥3	1.01	0.99-1.04	0.223	1.02	0.99-1.04	0.161	1.02	1.00-1.04	0.092	1.02	1.00-1.04	0.038

RRR: relative risk ratios in multinomial logistic analyses were determined, adjusting for age and sex as well as current alcohol consumption, smoking and walking habits.
[a]MetS risks I: sum of overweight (BMI ≥27.5 kg/m^2), hypertension (SBP ≥130 mmHg and/or DBP ≥85 mmHg or on medication), dyslipidemia (HDL <40 mg/dl for men and <50 mg/dl for women or on medication), and IGT criteria 1 (HbA1c ≥5.5% or on medication).
[b]MetS risk II: sum of overweight, hypertension, dyslipidemia, and IGT criteria 2 (HbA1c ≥5.7% or on medication).
[c]MetS risk III: sum of overweight, hypertension, dyslipidemia, and IGT criteria 3 (HbA1c ≥6.0% or on medication).
[d]MetS risk IV: sum of overweight, hypertension, dyslipidemia, and IGT criteria 4 (HbA1c ≥6.5% or on medication).

nonhypercholesterolemic subjects (27,28). The relationship between dyslipidemia and oxidative stress is rather complicated, because the functions of HDL and low-density lipoprotein (LDL) are more important for cellular oxidation than lipid profiles. HDL is considered to have an important role in protecting LDL from oxidization; meanwhile, it is also a major carrier of lipid-peroxides and F_2-isoprostanes (29). A recent interesting report explains the underlying mechanisms of this seemingly contradictory association, wherein subfractions of LDL show different effects on 8-iso-PGF2α and atherogenic risk (30); the most negatively charged fraction of LDL is significantly associated with non-HDL-C and 8-iso-PGF2α, whereas the least-charged fraction of LDL shows a significant inverse association. Although the proportion of LDL subfractions was not detected in the present study, this might partially explain the inverse association of non-HDL-C with 8-iso-PGF2α level. Taken together with the facts that HDL-C is a carrier of F_2-isoprostane and that LDL subfractions have different roles in systemic oxidation, further studies are warranted to clarify the relationship between oxidative stress and dyslipidemia.

Oxidative stress is also related to hypertension involving endothelial dysfunction mainly caused by decreases in nitric oxide bioavailability (31). In this study, 8-iso-PGF2α level was associated with hypertension only in participants aged ≤50 years. In an early study conducted with men aged 35–60 years, 8-iso-PGF2α level was significantly higher in hypertensive patients than that in normotensive subjects (32). In another previous study with older participants, 8-iso-PGF2α was significantly associated with hypertension in men but not women (28). Differences in the hypertension with respect to sex and age have been explained by the involvement of estradiol in nitric oxide stimulation (31,33). However, this does not fully explain the results in older male participants in this study; therefore, further study is needed to clarify this matter.

Although this is a cross-sectional study, which precludes the determination of causal relationships, a possible mechanism underlying the findings of this study is as follows. Metabolic, inflammatory, and degenerative diseases such as diabetes mellitus, atherosclerosis, and MetS share a signaling pathway involving a major transcription factor, nuclear factor kappa B (NF-κB) in response to oxidative stress. NF-κB controls the transcription of many pro-inflammatory mediators including tumor necrosis factor-α (TNF-α), interleukin-1, and interleukin-6 as well as cellular adhesion molecules (34). These inflammatory cytokines are associated with arachidonic acid-derived lipid mediators and play important roles in the development of atherosclerosis and diabetes. In an animal study, 8-iso-PGF2α formed in arachidonic acid decreased as NF-κB decreased as a result of antioxidant hemeoxygenase treatment to increase insulin sensitivity

(35). A recent detailed review suggests insulin regulates oxidative stress via the nuclear factor erythroid 2-related factor 2 (Nrf-2) signaling pathway (36). The transcription of hemeoxygenase is actually controlled by the Nrf-2 pathway. Studies investigating the mechanisms of NF-κB or Nrf-2 in metabolic and inflammatory diseases *in vivo* suggest they play important roles in humans. In fact, interleukin-1 receptor-associated kinase (IRAK), an inhibitor of NF-κB signaling and Nrf-2, is downregulated in MetS patients (37,38). Furthermore, genes involved in Nrf-2 pathways are significantly associated with oxidative stress defense systems in obese nondiabetic subjects (39). Nrf-2 controls not only antioxidant response genes, but also regulates genes related to adipogenesis, suggesting its importance in MetS. Thus, the results of the present study might be partially explained by the balance between oxidative stress and the Nrf-2/NF-κB defense system. Nevertheless, further study is required to determine the molecules involved in these pathways in participants in the ROAD study in order to clarify the underlying mechanisms.

There are several limitations of the present study that warrant consideration. Because we analyzed single spot urine samples, there might be additional unknown confounding factors. However, the good correlation between 8-iso-PGF2α levels between spot collection and 24-h collection might eliminate this concern (40). Another limitation is the cross-sectional design; therefore, the causal relationship between 8-iso-PGF2α and MetS risk remains unclear. However, further follow-up studies with our ROAD cohort will help clarify the causal relationship. In addition, as participants were not required to fast, we measured HbA1c instead of fasting plasma glucose, which is considered a MetS component in many international definitions. However, as mentioned before, HbA1c is a predictor of glucose intolerance. Therefore, the indices used in the present study accurately reflected the participants' physical condition. The influence of anti-inflammatory medications should be considered; however, details of those medications were not available. Moreover, we did not analyze genetic background but plan to in the near future. It is also worth noting that using the β-glucuronidase immunoassay allowed us to measure both the free and conjugated forms of 8-iso-PGF2α in urine. The formation of glucuronide conjugates of 8-iso-PGF2α might be a homeostatic mechanism to reduce oxidative damage to the kidneys (25). It is reasonable to consider that factors that affect glucuronidation, such as diet, may influence free 8-iso-PGF2α level in urine. Therefore, it is important to assess 8-iso-PGF2α level with β-glucuronidase to eliminate this concern. In other words, comparison of assessments with or without β-glucuronidase treatment will provide more information. Furthermore, we would plan a confirmatory study to compare our findings to those using GC/MS in the future.

In conclusion, the present study of a large general Japanese population from the ROAD study revealed a significant association of urinary 8-iso-PGF2α level with the accumulation of MetS risk. Our results suggest 8-iso-PGF2α can be a reliable marker of IGT and the accumulation of MetS risk besides diagnosis in people aged $\leq$50 years.

Acknowledgments

The authors wish to thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka of the Gobo Public Health Centre; Dr. Naoki Hirabayashi of the Kawakami Clinic, Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Kumiko Shinou, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, Ms. Ikuyo Ueyama, Mrs. Michiko Mori, Mrs. Hisayo Sugimoto, and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Centre; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mr. Shoichi Shimoichi, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji, and other members of the public office in Taiji Town for their assistance for examinations. We also thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, and Mrs. Saeko Sahara for their assistance with data reduction and administration. O

References

1. Cornier MA, Dabelea D, Hernandez TL, Lindstrom RC, Steig AJ, Stob NR, et al. The metabolic syndrome. *Endocr Rev* 2008;29:777-822.
2. Holewijn S, den Heijer M, Swinkels DW, Stalenhoef AF, de Graaf J. The metabolic syndrome and its traits as risk factors for subclinical atherosclerosis. *J Clin Endocrinol Metab* 2009;94:2893-2899.
3. Valle Gottlieb MG, da Cruz IB, Duarte MM, Moresco RN, Wiehe M, Schwanke CH, et al. Associations among metabolic syndrome, ischemia, inflammatory, oxidatives, and lipids biomarkers. *J Clin Endocrinol Metab* 2010;95:586-591.
4. Yubero-Serrano EM, Delgado-Lista J, Pena-Orihuela P, Perez-Martinez P, Fuentes F, Marin C, et al. Oxidative stress is associated with the number of components of metabolic syndrome: LIPGENE study. *Exp Mol Med* 2013;45:e28
5. Halliwell B, Whiteman M. Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean? *Br J Pharmacol* 2004;142:231-255.
6. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. *Curr Biol* 2014;24:R453-62. 4
7. Matsuda M, Shimomura I. Increased oxidative stress in obesity: implications for metabolic syndrome, diabetes, hypertension, dyslipidemia, atherosclerosis, and cancer. *Obes Res Clin Pract* 2013;7:e330-341.
8. Morrow JD, Roberts LJ. The isoprostanes: unique bioactive products of lipid peroxidation. *Prog Lipid Res* 1997;36:1-21.
9. Kadiiska MB, Gladen BC, Baird DD, Germolec D, Graham LB, Parker CE, et al. Biomarkers of oxidative stress study II: are oxidation products of lipids, proteins, and DNA markers of CCl4 poisoning? *Free Radic Biol Med* 2005;38:698-710.
10. Halliwell B, Lee CY. Using isoprostanes as biomarkers of oxidative stress: some rarely considered issues. *Antioxid Redox Signal* 2010; 13:145-156.
11. Basu SF2-isoprostanes in human health and diseases: from molecular mechanisms to clinical implications. *Antioxid Redox Signal* 2008;10:1405-1434.
12. Sjogren P, Basu S, Rosell M, Silveira A, de Faire U, Vessby B, et al. Measures of oxidized low-density lipoprotein and oxidative stress are not related and not elevated in otherwise healthy men with the metabolic syndrome. *Arterioscler Thromb Vasc Biol* 2005;25:2580-2586.
13. Fujita K, Nishizawa H, Funahashi T, Shimomura I, Shimabukuro M. Systemic oxidative stress is associated with visceral fat accumulation and the metabolic syndrome. *Circ J* 2006;70:1437-1442.
14. Kim M, Paik JK, Kang R, Kim SY, Lee SH, Lee JH. Increased oxidative stress in normal-weight postmenopausal women with metabolic syndrome compared with metabolically healthy overweight/obese individuals. *Metabolism* 2013;62:554-560.
15. Meigs JB, Larson MG, Fox CS, Keaney JF Jr., Vasan RS, Benjamin EJ. Association of oxidative stress, insulin resistance, and diabetes risk phenotypes: the framingham offspring study. *Diabetes Care* 2007;30:2529-2535.
16. Ong KL, Tso AW, Lam KS, Cherny SS, Sham PC, Cheung BM. Using glycosylated hemoglobin to define the metabolic syndrome in united states adults. *Diabetes Care* 2010;33:1856-1858.
17. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, et al. Accumulation of metabolic risk factors such as overweight, hypertension, dyslipidaemia, and impaired glucose tolerance raises the risk of occurrence and progression of knee osteoarthritis: a 3-year follow-up of the ROAD study. *Osteoarthritis Cartilage* 2012;20:1217-1226.
18. Janghorbani M, Amini M. Comparison of glycated hemoglobin with fasting plasma glucose in definition of glycemic component of the metabolic syndrome in an Iranian population. *Diabetes Metab Syndr* 2012;6:136-139.
19. Wu TW, Chan HL, Hung CL, Lu IJ, Wang SD, Wang SW, et al. Differential patterns of effects of age and sex on metabolic syndrome in Taiwan: implication for the inadequate internal consistency of the current criteria. *Diabetes Res Clin Pract* 2014;105:239-244.

20. Pani LN, Korenda L, Meigs JB, Driver C, Chamany S, Fox CS, et al. Effect of aging on A1C levels in individuals without diabetes: evidence from the framingham offspring study and the national health and nutrition examination survey 2001-2004. *Diabetes Care* 2008;31:1991-1996.
21. Cohen RM, Haggerty S, Herman WH. HbA1c for the diagnosis of diabetes and prediabetes: is it time for a mid-course correction? *J Clin Endocrinol Metab* 2010; 95:5203-5206.
22. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability study. *Int J Epidemiol* 2010;39:988-995.
23. WHO EC. Appropriate body-mass index for asian populations and its implications for policy and intervention strategies. *Lancet* 2004;363:157-163.
24. Proudfoot J, Barden A, Mori TA, Burke V, Croft KD, Beilin LJ, et al. Measurement of urinary F(2)-isoprostanes as markers of in vivo lipid peroxidation-a comparison of enzyme immunoassay with gas chromatography/mass spectrometry. *Anal Biochem* 1999;272:209-215.
25. Yan Z, Mas E, Mori TA, Croft KD, Barden AE. A significant proportion of F2-isoprostanes in human urine are excreted as glucuronide conjugates. *Anal Biochem* 2010;403:126-128.
26. Monnier L, Colette C, Mas E, Michel F, Cristol JP, Boegner C, et al. Regulation of oxidative stress by glycaemic control: evidence for an independent inhibitory effect of insulin therapy. *Diabetologia* 2010;53:562-571.
27. Keaney JF Jr., Larson MG, Vasan RS, Wilson PW, Lipinska I, Corey D, et al. Obesity and systemic oxidative stress: clinical correlates of oxidative stress in the framingham study. *Arterioscler Thromb Vasc Biol* 2003;23:434-439.
28. Kanaya AM, Wassel CL, Stoddard PJ, Harris TB, Cummings SR, Kritchevsky SB, et al. F2-isoprostanes and adiposity in older adults. *Obesity (Silver Spring)* 2011;19: 861-867.
29. Proudfoot JM, Barden AE, Loke WM, Croft KD, Puddey IB, Mori TA. HDL is the major lipoprotein carrier of plasma F2-isoprostanes. *J Lipid Res* 2009;50:716-722.
30. Urata J, Ikeda S, Koga S, Nakata T, Yasunaga T, Sonoda K, et al. Negatively charged low-density lipoprotein is associated with atherogenic risk in hypertensive patients. *Heart Vessels* 2012;27:235-242.
31. Landmesser R, Drexler H. Endothelial function and hypertension. *Curr Opin Cardiol* 2007;22:316-320.
32. Rodrigo R, Prat H, Passalacqua W, Araya J, Guichard C, Bachler JP. Relationship between oxidative stress and essential hypertension. *Hypertens Res* 2007;30:1159-1167.
33. Michel T, Vanhoutte PM. Cellular signaling and NO production. *Pflugers Arch* 2010;459:807-816.
34. Lawrence T. The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb Perspect Biol* 2009;1:a001651
35. Jadhav A, Tiwari S, Lee P, Ndisang JF. The heme oxygenase system selectively enhances the anti-inflammatory macrophage-m2 phenotype, reduces pericardial adiposity, and ameliorated cardiac injury in diabetic cardiomyopathy in zucker diabetic fatty rats. *J Pharmacol Exp Ther* 2013;345:239-249.
36. Wang X, Tao L, Hai CX. Redox-regulating role of insulin: the essence of insulin effect. *Mol Cell Endocrinol* 2012;349:111-127.
37. Hulsmans M, Geeraert B, De Keyzer D, Mertens A, Lannoo M, Vanaudenaerde B, et al. Interleukin-1 receptor-associated kinase-3 is a key inhibitor of inflammation in obesity and metabolic syndrome. *PLoS One* 2012;7:e30414
38. Santillan LD, Moyano M, Frau M, Flores O, Siewert S, Zirulnick F, et al. Reduced blood nrf-2 mRNA in local overweight boys at risk of metabolic complications: a study in San Luis city, San Luis, Argentina. *Metab Syndr Relat Disord* 2013;11: 359-365.
39. Das SK, Sharma NK, Hasstedt SJ, Mondal AK, Ma L, Langberg KA, et al. An integrative genomics approach identifies activation of thioredoxin/thioredoxin reductase-1-mediated oxidative stress defense pathway and inhibition of angiogenesis in obese nondiabetic human subjects. *J Clin Endocrinol Metab* 2011;96:E1308-1313.
40. Helmersson J, Basu S. F2-isoprostane excretion rate and diurnal variation in human urine. *Prostaglandins Leukot Essent Fatty Acids* 1999;61:203-205.

Osteoporos Int (2015) 26:2597–2605
DOI 10.1007/s00198-015-3184-0

ORIGINAL ARTICLE

Factors affecting changes in the serum levels of 25-hydroxyvitamin D: a 3-year follow-up of the ROAD study

N. Yoshimura[1] · S. Muraki[2] · H. Oka[3] · S. Tanaka[4] · H. Kawaguchi[5] · K. Nakamura[6] · T. Akune[6]

Received: 2 September 2014 / Accepted: 19 May 2015 / Published online: 19 June 2015

Abstract
Summary In this 3-year population-based cohort study, among 1346 subjects, the mean annual change in the serum 25-hydroxyvitamin D levels was 7.6 %/year, which tended to increase during the 3-year period. Multivariate regression analysis indicated that the L2-4 bone mineral density and total daily energy intake were significant independent associated factors.
Introduction The aim of this study was to clarify the change rate of the serum levels of 25-hydroxyvitamin D (25D) and the associated factors in a general Japanese population during a 3-year period.
Methods The baseline survey of Research on Osteoarthritis/osteoporosis Against Disability study (ROAD), a large-scale population-based cohort study, was performed between 2005 and 2007, and a follow-up survey was repeated 3 years later. Among 1690 participants at baseline, the change rate of the serum 25D levels were assessed in 1346 individuals (79.6 %; 458 men and 888 women) who completed measurements of 25D at both the baseline and follow-up examinations. The change rate was calculated, and the factors associated with the changes in the 25D levels were determined using multivariate regression analysis after adjustment for age, gender, body mass index, participated month, and regional differences at baseline.
Results The mean (standard deviation) change rate of the 25D levels in all subjects was 7.6 (13.3) %/year (men, 8.2 [12.4] %/year; women, 7.3 [13.7] %/year). Multivariate regression analysis indicated that higher bone mineral density at lumbar spine L2-4 (p=0.05) and total daily energy intake (p=0.04) were significantly associated with the change rate of the 25D levels.
Conclusions The serum levels of 25D tended to increase over the 3-year period, and higher lumbar bone mineral density and daily energy intake were found to be associated with increases in the 25D levels over time.

Keywords 25-Hydroxyvitamin D · Bone mineral density · Nutrition · Population-based cohort study · Risk factors

✉ N. Yoshimura
YOSHIMURAN-ORT@h.u-tokyo.ac.jp

[1] Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

[2] Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

[3] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

[4] Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

[5] JCHO Tokyo Shinjuku Medical Center, 5-1, Tsukudo-cho, Shinjyuku-ku, Tokyo 162-8542, Japan

[6] National Rehabilitation Center for Persons with Disabilities, 1, Namiki 4-chome, Tokorozawa City, Saitama Prefecture 359-8555, Japan

Introduction

Vitamin D (VD) is known to influence bone quality and is important for maintaining bone density [1, 2]. A number of studies have reported an association between inadequate VD intake and osteoporosis [3–6], and VD deficiency has moreover been reported as detrimental to numerous other conditions, including falls, fractures, type 2 diabetes, cardiovascular disease, certain cancers, autoimmune diseases, infections, and mortality [6]. Accordingly, preventing VD inadequacy is

considered highly beneficial for the prevention of morbidity; however, there are currently few reports clarifying how the VD levels change over time and what factors may affect the changes in the VD levels.

We have previously performed a population-based cohort survey using the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study cohorts, in which we first clarified the characteristics of subjects with VD deficiency and insufficiency, as defined by serum 25-hydroxyvitamin D (25D) levels <10 ng/mL and ≥10<30 ng/mL, respectively, using the baseline data of ROAD [7]. In this previous study, we found that VD deficiency was significantly associated with gender, residing region, serum intact parathyroid hormone (iPTH) levels, urinary β-isomerized C-terminal cross-linking telopeptide of type I collagen (β-CTX) levels, smoking habits, regular walking, calcium intake, and hyperparathyroidism. Next, using both the baseline and 3-year follow-up information of the ROAD study, we reported that the serum 25D levels at baseline could predict the occurrence of osteoporosis at the femoral neck within 3 years, but not the occurrence of knee osteoarthritis or lumbar spondylosis [8]. However, we have not yet evaluated the change rate of 25D and factors associated with the changes in the 25D levels over time.

With this in mind, in the present study, the changes from the baseline 25D levels in subjects with measurement taken both at the baseline and 3-year follow-up examinations of the ROAD study were evaluated to clarify the change rate of 25D. In addition, the factors affecting the changes in the 25D levels were assessed.

Methods

Study participants

The present study was performed using the ROAD study cohort established in 2005. The ROAD study is a national, prospective study of osteoarthritis that consists of population-based cohorts from several communities in Japan. Details of the cohort profile have been reported elsewhere [9, 10]. In brief, between 2005 and 2007, a baseline database was created that included clinical and genetic information of 3040 individuals (1061 men, 1979 women; mean age [standard deviation, SD], 70.3 [11.0] years). The subjects were recruited from resident registration listings in three communities with different characteristics: 1350 subjects from an urban region in Itabashi, Tokyo; 864 subjects from a mountainous region in Hidakagawa, Wakayama; and 826 subjects from a coastal region in Taiji, Wakayama.

In the present study, 1690 subjects (596 men, 1094 women; mean age, 65.2 [12.0] years; men, 66.3 [11.7] years; women, 64.7 [12.1] years) from the mountainous and coastal regions who participated in the ROAD study were analyzed. Data from those in the urban regions were not included, as baseline 25D measurements were not performed in that cohort. On the other hand, bone mineral density (BMD) measurements and blood and urinary examinations were performed in the participants from the mountainous and coastal regions.

All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo (No. 1264 and No. 1326) and the University of Wakayama Medical University (No. 373).

Baseline assessments

i. Interviewer-administered questionnaire

The participants completed an interviewer-administered questionnaire that consisted of questions related to lifestyle, including occupation, smoking habits, alcohol consumption, family history, medical history, physical activity, reproductive history, and health-related quality of life.

ii. Dietary assessment

A brief diet history questionnaire (BDHQ) was administered to assess the diet of the participants. The BDHQ was modified from a comprehensive, 16-page validated self-administered diet history questionnaire [11]. The BDHQ is a four-page structured questionnaire that includes questions about the frequency of consumption of 80 principal foods. The food serving size was defined as a normal portion according to the standard weight and volume of servings commonly consumed by the general Japanese population. A total of 141 variables, including dietary energy and nutrient intake, were calculated using an ad hoc computer algorithm for the BDHQ. Detailed explanations accompanied each questionnaire. Well-trained interviewers clarified any unclear sections of the questionnaire, which was completed by the participants at their leisure.

iii. Anthropometric measurements and medical history

Anthropometric measurements such as height and weight were measured in all participants, and body mass index (BMI) was calculated as weight (kg)/height (m^2). In addition, the grip strength of both hands was measured. Experienced orthopedic surgeons collected medical information about pain, swelling, and the range of motion of the knee.

iv. Blood and urinary examinations

Samples were collected between the end of October and the middle of January. All blood and urine samples were collected between 9 a.m. and 3 p.m. Blood samples were centrifuged to obtain sera. Sera and urine samples were immediately placed on dry ice and transferred to a deep freezer within 24 h. Samples were stored at −80 °C until assayed.

The serum 25D levels were measured using a radioimmunoassay with a ^{125}I-labeled tracer (DiaSorin, Stillwater,

MN, USA) [12]. The iPTH levels were measured using an electrochemiluminescence immunoassay (Roche Diagnostics GmbH, Menheim, Germany). Serum procollagen type I N-terminal propeptide (PINP), a marker of bone formation, was measured using a radioimmunoassay (Orion Diagnostics, Espoo, Finland). Urinary levels of β-CTX, a marker of bone resorption, were determined using enzyme-linked immunosorbent assay (Fujirebio, Inc., Tokyo, Japan). Urinary β-CTX values were standardized to the urinary creatinine concentrations.

v. Bone mineral density examination

The lumbar spine and proximal femur BMD values were determined using dual-energy X-ray absorptiometry (Hologic Discovery; Hologic, Waltham, MA, USA).

Three-year follow-up

Between 2008 and 2010, the 1690 participants with baseline 25D measurements were invited again to participate in the 3-year follow-up of the ROAD survey, which repeated the baseline examinations.

The follow-up samples were collected between October and January, in the same way as for the baseline study. As a general rule in the ROAD study, those who attended the baseline visit in, for example, October, subsequently attended the follow-up in October as well. However, some participants who attended the baseline visit in October might have attended the follow-up in a later month, owing to personal reasons.

Similar to for the baseline examination, all blood and urine samples at the follow-up were collected between 9 a.m. and 3 p.m. The blood samples were centrifuged to obtain sera. Sera and urine samples were immediately placed on dry ice and transferred to a deep freezer within 24 h, and the samples were stored at −80 °C until assayed.

Changes in the vitamin D levels

The changes in the serum levels (ng/mL) of 25D, and annual change rate of the 25D levels (%/year) over the 3-year period were calculated as follows:

Change in 25D levels (ng/mL)=25D levels measured in the second survey − 25D levels measured in the baseline survey.

Annual change rate of the 25D levels (%/year)=([(25D levels measured in the second survey − 25D levels measured in baseline survey)/25D levels measured in baseline survey]/3)×100.

Statistical analyses

All statistical analyses were performed using STATA statistical software (STATA Corp., College Station, TX, USA). Differences in proportions were compared using the chi-square test. The significance of differences between continuous variables was evaluated using analysis of variance for comparisons among multiple groups or Scheffe's least significant difference test for pairs of groups. All p values and 95 % confidence intervals are two-sided. A value of $p<0.05$ was considered statistically significant.

Multivariate regression analysis was used to test the association of the related factors with the changes in the 25D levels. In the analysis, we used the values of the annual change rate of serum 25D levels (%/year) as the objective variable and selected potential associated factors as explanatory variables, after adjusting for age (+1 year), gender (0, men; 1, women), BMI (+1 kg/m^2), participated month (0, October, November, December; 1, January), and regional differences (0, mountainous area; 1, coastal area) at the baseline survey, as previously described [7, 8].

As for the potential associated factors used as explanatory factors, we assessed the following factors with a significant or marginal ($p<0.1$) association with VD status in the simple linear analysis, which were assessed for the association with VD deficiency herein and in the previous reports of the ROAD study [7, 8]: smoking (0, never, ever; 1, current), alcohol consumption (0, never, ever; 1, current), regular walking outside (0, <5 times/week; 1, ≥5 times/week), regular exercising outdoors (e.g., football, tennis, baseball, and golf) after the most recent graduation (0, no; 1, yes), serum levels of iPTH (+1 pg/mL), baseline BMD values at the lumbar spine (+1 g/cm^2) or femoral neck (+1 g/cm^2), serum levels of PINP (+1 standard deviation, μg/L), urinary levels of β-CTX (+1 standard deviation, μg/mmol creatinine), and daily amounts of total energy (+100 kcal/day), calcium (+100 mg/day), and VD (+10 μg/day) intake, as calculated based on the BDHQ questionnaire.

Results

Eligible participants

Of the 1690 participants with 25D measurements at baseline, 251 (14.9 %) did not participate in the second survey. The reasons for these 251 dropouts were as follows: 40 individuals died, 97 were ill, 16 moved away, 8 were absent, 51 declined to participate in a second visit, and 39 stated "other" reasons. Furthermore, 55 individuals (3.3 %) did not complete the second survey, and the values of serum 25D were not obtained in 11 individuals at the follow-up because of insufficient serum samples for the measurement. In addition, 14 subjects were diagnosed as suspicion of hyperparathyroidism based on elevated serum iPTH >100 pg/mL at baseline, and 13 subjects prescribed oral VD at baseline and during the 3 years between the baseline and second surveys were excluded. As a result, in

the present study, 1346 subjects were analyzed (79.6 %; 458 men and 888 women; mean age, 63.8 [11.8] years).

Table 1 shows the baseline characteristics of the 1346 subjects classified by gender. The mean age (SD) of the subjects in the present study was 63.8 (11.8) years and the mean BMI (SD) was 23.1 (3.3) kg/m^2. More than half of all individuals lived in the coastal area, and approximately 30 % of all men were smokers, while the corresponding proportion of smokers in women was very small (3.5 %). More than half of all subjects had a habit of walking outside, including to, or during, work, and 15 % exercised regularly outdoors. The daily intakes of total energy (kcal), calcium (mg), and VD (μg) are also shown in Table 1, and these were all significantly higher in men than in women. Moreover, the prevalence of

Table 1 Background characteristics of the study subjects at baseline

	Total (n=1346)	Men (n=458)	Women (N=888)	p (men vs. women)
Mean values (SD) of 25D (ng/mL) and changes in the 25D levels (ng/mL)				
Baseline 25D levels	23.4 (6.5)	25.9 (6.4)	22.2 (6.2)	<0.0001***
Measurement of 25D in the second survey	27.7 (8.5)	31.4 (9.8)	25.8 (7.1)	<0.0001***
Changes in the 25D levels between baseline and the second survey	4.3 (7.3)	5.5 (8.5)	3.6 (6.5)	<0.0001***
Annual change rate (%/year) of 25D between baseline and the second survey	7.6 (13.3)	8.2 (12.4)	7.3 (13.7)	0.26
Mean values (SD) of selected characteristics				
Age (years)	63.8 (11.8)	64.9 (11.6)	63.2 (11.8)	0.0120*
Height (cm)	155.7 (9.0)	164.0 (7.0)	151.4 (6.6)	<0.0001***
Weight (kg)	56.1 (10.7)	52.5 (8.6)	63.1 (10.7)	<0.0001***
BMI (kg/m^2)	23.1 (3.3)	23.4 (3.2)	22.9 (3.4)	0.0068**
Prevalence of selected characteristics (%)				
Residing in the coastal area	53.6	51.8	54.6	0.32
Current smoking habit (regularly, ≥1 time/month)	12.4	29.4	3.5	<0.001***
Current alcohol consumption (regularly, ≥1 time/month)	40.8	68.5	26.6	<0.001***
Regularly walking outside (≥5 times/week, including for work)	56.5	62.1	53.6	0.004**
Regularly exercising outdoors (e.g., football, tennis, baseball, and golf) after the most recent school graduation	15.3	36.0	4.6	<0.001***
Mean values (SD) of selected measurements				
Serum levels of iPTH (pg/mL)	38.9 (14.7)	37.5 (13.8)	39.6 (15.1)	0.0155*
BMD (L2-4) (g/cm^2)	0.94 (0.20)	1.06 (0.20)	0.88 (0.18)	<0.0001***
BMD (femoral neck) (g/cm^2)	0.66 (0.16)	0.62 (0.12)	0.74 (0.14)	<0.0001***
Serum levels of PINP (μg/L)	57.9 (26.8)	46.9 (20.4)	65.5 (28.0)	<0.0001***
Urinary levels of β-CTX (μg/mmol Cr)	185.7 (126.8)	124.9 (73.0)	217.0 (136.9)	<0.0001***
Mean values (SD) of the total amounts of selected nutritional intakes (/day)				
Total energy (kcal)	1947 (584)	2326 (656)	1751 (428)	<0.0001***
Calcium (mg)	558 (230)	582 (259)	546 (215)	0.0069**
Vitamin D (μg)	20.4 (12.3)	22.6 (13.9)	19.3 (11.3)	<0.0001***
Prevalence of medication of osteoporosis prescribed by doctors (%)				
Calcium	0.4	0.2	0.5	0.51
Bisphosphonate	1.8	0.2	2.6	0.002**
SERMs	1.0	0.0	1.6	0.007**
Calcitonin	0.5	0.0	0.7	0.08
Prevalence of hyperparathyroidism (%)	6.1	5.2	6.5	0.35
Prevalence of osteoporosis (%) according to the WHO criteria				
Osteoporosis (L2-4)	13.6	2.4	19.4	<0.001***
Osteoporosis (femoral neck)	13.0	3.1	18.1	<0.001***

SD standard deviation, *25D* 25-hydroxyvitamin D, *BMI* body mass index, *iPTH* intact parathyroid hormone, *BMD* bone mineral density, *L2-4* lumbar spine L2-L4, *PINP* procollagen type I N-terminal propeptide, *β-CTX* β-isomerized C-terminal telopeptide cross-links of type I collagen, *Cr* creatinine, *SERMs* selective estrogen receptor modulators, *WHO* World Health Organization

*p<0.05; **p<0.01; ***p<0.001

osteoporosis measured in both the lumbar spine of L2-4 and femoral neck was significantly higher in women than in men, as were the mean values of iPTH. However, the proportion of medication use for osteoporosis was very small, even in women (Table 1).

Changes in the VD levels in subjects who participated in both the baseline and second surveys

In the total 1346 subjects, the mean change in the serum 25D level (SD) was 4.3 (7.3) ng/mL. The corresponding changes in men and women were 5.5 (8.5) and 3.6 (6.5) ng/mL, respectively, and the mean annual change rates in the 25D levels in the total subjects, men, and women were 7.6 (13.3), 8.2 (12.4), and 7.3 (13.7) %/year, respectively. There was no significant gender difference in the annual change rate of serum 25D (Table 1).

The mean annual change rates in the 25D levels (SD) in subjects in their 30s or younger, 40s, 50s, 60s, 70s, and 80s or older were 1.3 (7.4), 5.1 (12.2), 7.4 (17.5), 8.8 (12.6), 7.7 (11.1), and 8.2 (11.0) %/year, respectively, indicating that the serum levels of 25D increased in all age groups over time. Further, the 25D levels were significantly increased in subjects in their 60s compared to those in their 30s or younger (p=0.048).

Figure 1 shows the distribution of the annual change rates in the 25D levels (SD) classified by gender. In men, the mean annual change rates in the 25D levels in participants in their 30s or younger, 40s, 50s, 60s, 70s, and 80s or older were 2.2 (6.9), 2.8 (7.8), 7.5 (16.8), 10.5 (12.0), 8.7 (10.7), and 6.4 (8.1) %/year, respectively, and there was a significant difference between men in their 30s and younger and those in their 60s (p=0.036). The corresponding values in women were 1.0 (7.6), 6.2 (13.5), 7.4 (17.8), 8.0 (12.9), 7.0 (11.4), and 9.1 (12.2) %/year, respectively, and there was no significant difference between any particular age groups (Fig. 1).

Factors associated with changes in the serum 25D levels

Table 2 shows the results of the multivariate regression analysis assessing the associations with the annual change rate in the 25D levels. The analysis was adjusted for age (+1 year), gender (0, men; 1, women), BMI (+1 kg/m^2), participated month (0, October, November, December; 1, January), and regional differences (0, mountainous area; 1, coastal area) at baseline, as previously described [8]. The results of the analysis indicated that the BMD values measured at lumbar spine L2-4 and the total daily energy intake were significantly associated with the change rate of the 25D levels (Table 2). In other words, the higher the BMD of the lumbar spine and the higher the total daily energy intake, the higher the increase in the serum 25D levels.

To test the independence of the associations between BMD at the lumbar spine and total energy of daily food with the change in the 25D levels, the multivariate model was adjusted for both these variables; as a result, both variables remained significantly associated with the changes in the 25D levels, revealing that these variables were mutually and independently associated with the changes in the 25D levels (Table 2).

Discussion

In this 3-year follow-up of a large cohort of subjects, the serum levels of 25D at baseline and at a second survey were measured, and we estimated the change rate of 25D levels and the factors contributing to this change rate. This study showed that the mean change rate of the levels of serum 25D tended to

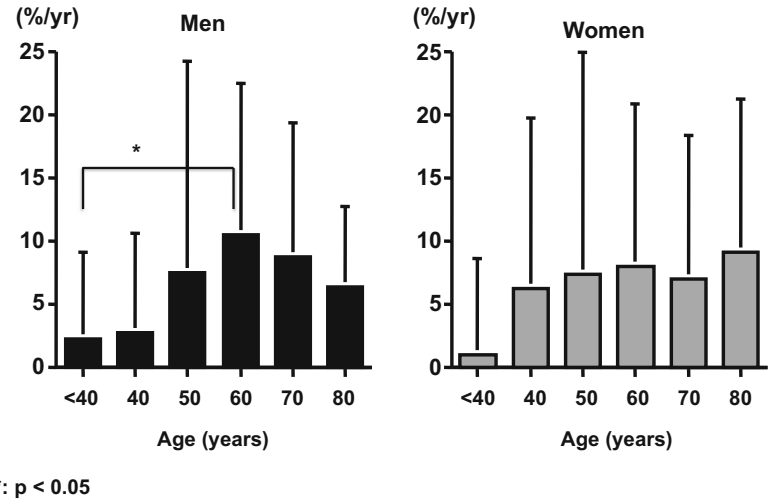


Fig. 1 Age-gender distribution (%) of the annual change rate of 25D

Table 2 Partial regression coefficients (beta) of potential factors associated with the annual change rate of serum 25D levels (%/year)

Explanatory variables collected at baseline	Reference	Multivariate regression analysis 1[a]			Multivariate regression analysis 2[b]		
		Beta	R^2	p	Beta	R^2	p
Lifestyle factors							
Smoking	0: ex or never smoker; 1: current smoker	0.04	0.143	0.18			
Alcohol consumption	0: ex or never drinker; 1: current drinker	0.12	0.142	0.46			
Regularly walking outside	0: no; 1: yes	0.02	0.135	0.46			
Regularly exercising outdoors (e.g. football, tennis, baseball, golf, etc.) after the most recent school graduation	0: no; 1: yes	−0.04	0.143	0.15			
Mean values of selected measurements							
Serum levels of iPTH (pg/mL)	+1 pg/mL	−0.01	0.141	0.62			
BMD (L2-4) (g/cm2)	+1 g/cm^2	0.06	0.144	0.05*	0.06	0.144	0.050*
BMD (femoral neck) (g/cm2)	+1 g/cm^2	−0.004	0.140	0.91			
Serum levels of PINP (μg/L)	+1 SD	0.009	0.140	0.73			
Urinary levels of β-CTX (μg/mmol Cr)	+1 SD	0.002	0.142	0.94			
Nutritional factors							
Total energy from daily food (kcal/day)	+100 kcal	0.06	0.145	0.039*	0.06	0.145	0.039*
Calcium from daily food (mg/day)	+100 mg	0.03	0.143	0.198			
Vitamin D from daily food (μg/day)	+1 μg	−0.01	0.142	0.578			

25D 25-hydroxyvitamin D, *iPTH* intact parathyroid hormone, *BMD* bone mineral density, *iPTH* intact parathyroid hormone, *PINP* procollagen type I N-terminal propeptide, *β*-CTX, β-isomerized C-terminal telopeptide cross-links of type I collagen, *Cr* creatinine

$^{+}p<0.1$;*$p<0.05$; **$p<0.01$; ***$p<0.001$

[a] Multivariate regression analysis was performed after adjustment for age (+1 year), gender (0: men; 1: women), body mass index (+1 kg/m2), participated month (0, October, November, December; 1, January), and regional differences (0: mountainous area; 1: coastal area) at the baseline survey

[b] Multivariate regression analysis was performed after adjustment for BMD (L2-4) and total daily energy intake mutually, in addition to age (+1 year), gender (0: men; 1: women), body mass index (+1 kg/m2), participated month (0, October, November, December; 1, January), and regional differences (0: mountainous area; 1: coastal area) at the baseline survey

increase during the 3-year period. There was no gender difference, and subjects in their 60s showed the largest increases in the change rate. Moreover, we found that the BMD values at the lumbar spine and the total daily energy intake at baseline were significantly associated with increases in the serum 25D levels.

The prevalence of VD inadequacy in postmenopausal women in Japan is well known to be very high [13, 14], and our previous report furthermore revealed a very high prevalence of VD insufficiency and a low prevalence of VD deficiency in Japanese men and women [7]. However, to our knowledge, this is the first report on the change rate of the serum levels of 25D measured in a general Japanese population.

As mentioned, in the present study, the mean changes in the serum 25D level tended to increase during the 3-year period in both men and women. It is possible that this result might be due to the intervention effect of the repeated surveys. We have performed individual consultations with all participants in order to explain the individual findings and inform them regarding the results of the survey. In such consultations, we provided the subjects with information on the association between osteoporosis and lifestyle factors, including nutrition, such as the importance of the total energy, calcium, and vitamin D intake from their daily food. Such consultations after each survey might have helped educate the participants, and as a result, it may have influenced the course of the 25D levels of the subjects. However, this procedure could not be avoided, and we believe that this is in fact a favorable effect, allowing for improved health outcomes of the study cohort, rather than a limitation of the present study.

Of note, at the translation of the changes of the measurements, there is one particular issue that should be considered, namely the regression to the mean (RTM) phenomenon. RTM is a statistical phenomenon that can cause natural variation. It happens when unusually large or small measurements tend to be followed by measurements that are closer to the mean [15]. To reduce the effects of this phenomenon, Barnett et al. recommended using a suitable study design, such as random allocation of subjects to the comparison groups and multiple measurements. Unfortunately, we could not change our study from a cohort study into a randomized controlled trial, or repeat the measurements of 25D at each visit, owing to budget limitations. Nonetheless, to reduce the RTM in the data

analysis, the authors also recommended adjusting each subject's follow-up measurements according to their baseline measurement, i.e., analysis of covariance [15]. Therefore, we here used the rate of the changes in the 25D levels as an objective variable and adjusted the baseline 25D levels in advance, which might have helped reducing the RTM phenomenon. By contrast, Sonderman et al. evaluated the reproducibility of 25D measurements using a prospective cohort study of 225 participants and indicated that the 25D serum measurements provided reasonably representative measures [16].

In this study, in the multivariate regression analysis, the BMD values at the lumbar spine, and the total daily energy intake measured at baseline were found to be associated with an increased change rate of the serum 25D levels, both mutually and independently from one another. As stated above, many studies have reported an association between inadequate VD intake and osteoporosis [3–6]. Cauley et al. reported that, in their observational study of a US cohort consisting of 1532 middle-aged women, the serum 25D levels were inversely associated with non-traumatic fractures, whereas menopause-related changes in the lumbar spine and femoral neck BMD values were not significantly associated with the serum 25D level [17]. In another cohort consisting of 1470 postmenopausal Japanese women, the incidence of proximal femur and long bone fractures tended to decrease as the serum 25D levels increased [18]. Regarding the relationship between 25D and the occurrence of osteoporosis, using the identical ROAD cohort as in the present study, we have previously shown that the serum 25D levels at baseline were significantly associated with subsequent osteoporosis occurrence, especially at the femoral neck [8]. These findings suggest that higher levels of 25D at baseline favorably impact future bone health. However, there is currently no report regarding the association between baseline BMD and changes in the 25D levels, as assessed in the present study.

Herein, we found that the baseline BMD values at lumbar spine L2-4 were significantly associated with increased 25D levels. Together with the results of our previous report in which we showed that the 25D levels at baseline influence the subsequent occurrence of osteoporosis [8], we conclude that the BMD values and 25D levels mutually influence one another over time. However, we did not examine the mechanisms of this mutual association in detail. Moreover, we also cannot explain why, in the present study, this association was observed only at the lumbar spine of L2-4 and not at the femoral neck. Further analysis is required to confirm this observed association.

Furthermore, among the nutritional factors examined, higher daily energy intake at baseline was found to significantly accelerate the increase in the serum levels of 25D. Adequate nutritional intake, especially adequate intake of calcium and VD, is recommended to maintain bone health [6, 19]. In addition, several other nutrients have been reported to improve bone health. Nieves reviewed several reports of nutrients and osteoporosis and concluded that the skeletal benefits of flavonoids, carotenoids, omega-3 fatty acids, and vitamins A, C, E, and K were limited to observational data or a few clinical trials [20]. Further, the author reported that potassium bicarbonate may improve calcium homeostasis, that high homocysteine levels may relate to the fracture risk, and that magnesium supplementation is likely only required in those with low magnesium levels [21]. Of note, regarding the associations between nutrients, BMD, and fractures, Nieves noted that the nutrients do not act in isolation [21], and it is hence important to evaluate the effects of different combinations of nutrients using analyses of dietary patterns. Accordingly, our results showed that the total daily energy intake, rather than daily VD intake, positively influenced the change rate of the 25D levels. Total energy can be considered a comprehensive factor containing most above mentioned nutrients, and we hence conclude that to prevent serum 25D decreases, it is important to improve energy-poor diets.

There are several limitations to this study. First, although the ROAD study includes a large number of participants, the participants in the present study (from the mountainous and coastal regions only) may not be completely representative of the general population. To address this issue, we compared the anthropometric measurements, smoking frequency, and alcohol consumption between the present study participants and the general Japanese population. The values for the general population were obtained from the report on the 2005 National Health and Nutrition Survey conducted by the Ministry of Health, Labour and Welfare, Japan [22], when our baseline ROAD started. The mean BMI values of men aged 30 or younger, 40, 50, 60, 70–74, 75–79, and 80 years or older, as reported in the National Health and Nutrition Survey, were 23.29 (3.68), 23.99 (3.27), 23.74 (3.07), 23.75(2.94), 23.68 (3.18), 23.31 (3.04), and 22.27 (2.64) kg/m^2, respectively, and those of women were 21.37 (3.81), 22.44 (3.49), 23.06 (3.37), 23.54 (3.66), 23.16 (3.42), 23.42 (3.53), and 22.50 (3.97) kg/m^2, respectively. In the present study, the corresponding mean BMI values were 23.51 (3.44), 24.70 (4.32), 23.60 (2.96), 23.86 (3.19), 23.11 (2.82), 22.24 (2.82), and 23.13 (2.74) kg/m^2, respectively, for men, and 21.14 (2.67), 21.82 (3.73), 23.03 (3.24), 23.25 (3.15), 23.65 (3.55), 22.44 (3.82), and 22.23 (3.21), respectively, for women. No significant differences were identified between our participants and the total Japanese population, except that male and female participants aged 75–79 years in the present study had significantly smaller BMI than the overall Japanese population ($p<0.05$). This difference should be taken into consideration when evaluating the potential risk factors in subjects aged 75–79 years. In addition, the proportion of current smokers and drinkers (those who regularly smoke or drink more than one drink/month) in the general Japanese population was compared with

that in the present study population. The proportion of current smokers was significantly higher in the general Japanese population than in our study population in both men and women (men, 34.8 vs. 29.4 %, $p<0.05$; women, 8.8 vs. 3.5 %, $p<0.05$). Moreover, the proportion of current drinkers was significantly higher in women in the general Japanese population than in women in the present study, while there was no significant difference in men (men, 69.8 vs. 68.5 %, $p=0.62$; women, 30.8 vs. 26.6 %, $p<0.05$), suggesting that the participants of the present study lead healthier lifestyles, at least in terms of their smoking habits, as compared to the general Japanese population. This selection bias should be taken into consideration when generalizing the results obtained from the present study. Second, the measurements of 25D were only conducted two times, once at baseline and once at the 3-year follow-up, and we could not exclude the effect of incidental life changes of the participants, such as holidays or dietary changes for special anniversaries around the examination date. Although this random effect on the subjects seems difficult to avoid, the large number of participants of the present study is considered to dilute any individual variance, and the high participation rate of the follow-up study is a major strength of the present study. Third, in the present study, limited information on the association between the changes in the exposure variables, including lifestyle factors, and the changes in the serum levels of 25D, was provided. As one example of a change in a potential associated factor, we assessed the association between new prescriptions of VD during the 3-year period between the baseline and the second survey and the changes in the serum levels of 25D. Over the 3-year follow-up, 32 individuals (3 men, 29 women) were newly prescribed oral VD supplementation. Multivariate regression analysis was performed to assess the association between the annual change rate in the 25D levels and new prescription of VD supplementation (1, yes; 0, no) after adjustment for age (+1 year), gender (0, men; 1, women), BMI (+1 kg/m^2), participated month (0, October, November, December; 1, January), and regional differences (0, mountainous area; 1, coastal area) at baseline. As a result, we found that there was no significant association between the presence of new VD supplementation and the changes in the serum levels of 25D (beta=−0.17, $p=0.51$). We plan on assessing the effects of changes in other potential associated factors, such as dietary-, physical-activity-, and other lifestyle-related factors on the changes in the serum levels of 25D over this 3-year period, in order to clarify the contributing factors to the increase in 25D, as a next step of our research.

In conclusion, we here found that the serum levels of 25D tended to increase over a 3-year period in a large cohort of general Japanese inhabitants and that higher lumbar BMD and total daily energy intake are significant independent factors associated with increases in the 25D levels over time.

Acknowledgments This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology for Scientific Research to NY (B26293139, B23390172, B20390182, and Challenging Exploratory Research 24659317), TA (C20591737, B23390357), and SM (C20591774, B23390356, Challenging Exploratory Research 23659580); for Young Scientists to HO (A18689031, Challenging Exploratory Research 24659666); and for Collaborating Research with NSF to NY (Director; 08033011-00262). Moreover, this work was supported by H17-Men-eki-009 (Director, KN), H18-Choujyu-037 (Director, TN), H20-Choujyu-009 (Director, NY), H23-Chojyu-002 (Director, TA), H25-Nanchi-to (Men)-005 (Director, ST), and H25-Chojyu-007 (Director, NY) from the Ministry of Health, Labour and Welfare in Japan, and by grants from the Japan Osteoporosis Society (NY, SM, HO, and TA), and research aids from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 and 2010-2; Director, HK). The authors wish to thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka of the Gobo Public Health Centre; Dr. Naoki Hirabayashi of the Kawakami Clinic, Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Kumiko Shinou, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, Ms. Ikuyo Ueyama, Mrs. Michiko Mori, Mrs. Hisayo Sugimoto, and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Centre; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mr. Shoichi Shimoichi, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji, and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. Finally, we also thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, and Mrs. Saeko Sahara for their assistance with data reduction and administration.

Conflicts of interest None.

References

1. Lips P (2001) Vitamin D deficiency and secondary hyperparathyroidism in the elderly: consequences for bone loss and fractures and therapeutic implications. Endocr Rev 22:477–501
2. Bischoff-Ferrari HA, Zhang Y, Kiel DP, Felson DT (2005) Positive association between serum 25-hydroxyvitamin D level and bone density in osteoarthritis. Arthritis Rheum 53:821–826
3. Bischoff-Ferrari HA, Kiel DP, Dawson-Hughes B, Orav JE, Li R, Spiegelman D, Dietrich T, Willett WC (2009) Dietary calcium and serum 25-hydroxyvitamin D status in relation to BMD among U.S. adults. J Bone Miner Res 24:935–942
4. Mithal A, Wahl DA, Bonjour JP, Burckhardt P, Dawson-Hughes B, Eisman JA, El-Hajj Fuleihan G, Josse RG, Lips P, Morales-Torres J, IOF Committee of Scientific Advisors (CSA) Nutrition Working Group (2009) Global vitamin D status and determinants of hypovitaminosis D. Osteoporos Int 20:1807–1820
5. Hanley DA, Cranney A, Jones G, Whiting SJ, Leslie WD, Cole DE, Atkinson SA, Josse RG, Feldman S, Kline GA, Rosen C, Guidelines Committee of the Scientific Advisory Council of Osteoporosis Canada (2010) Vitamin D in adult health and disease: a review and guideline statement from Osteoporosis Canada. CMAJ 182:E610–E618
6. Dawson-Hughes B, Mithal A, Boonen S, Bonjour JP, Burckhardt P, Ghada El-Hajj Fuleihan GEH, Josse R, Lips P, Morales-Torres J, Yoshimura N, for the IOF CSA Nutrition Working Group (2010) Vitamin D Recommendations for Older Adults. Osteoporos Int 21: 1151–1154
7. Yoshimura N, Muraki S, Oka H, Morita M, Yamada H, Tanaka S, Kawaguchi H, Nakamura K, Akune T (2013) Profiles of vitamin D insufficiency and deficiency in Japanese men and women:

association with biological, environmental, and nutritional factors and coexisting disorders: the ROAD study. Osteoporos Int 24: 2775–2787
8. Yoshimura N, Muraki S, Oka H, Nakamura K, Kawaguchi H, Tanaka S, Akune T Serum levels of 25-hydroxyvitamin D and occurrence of musculoskeletal diseases, such as osteoporosis, knee osteoarthritis and lumbar spondylosis: a three-year follow-up of the ROAD study. Osteoporos Int in press
9. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Ishibashi H, Kawaguchi H, Nakamura K, Akune T (2009) Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 27:620–628
10. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T (2010) Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol 39:988–995
11. Kobayashi S, Honda S, Murakami K, Sasaki S, Okubo H, Hirota N, Notsu A, Fukui M, Date C (2012) Both comprehensive and brief self-administered diet history questionnaires satisfactorily rank nutrient intakes in Japanese adults. J Epidemiol 22:151–159
12. Hollis BW, Kamerud JQ, Selvaag SR, Lorenz JD, Napoli JL (1993) Determination of vitamin D status by radioimmunoassay with an 125I-labeled tracer. Clin Chem 39:529–533
13. Lim SK, Kung AW, Sompongse S, Soontrapa S, Tsai KS (2008) Vitamin D inadequacy in postmenopausal women in Eastern Asia. Curr Med Res Opin 24:99–106
14. Lips P, Hosking D, Lippuner K, Norquist JM, Wehren L, Maalouf G, Ragi-Eis S, Chandler J (2006) The prevalence of vitamin D inadequacy amongst women with osteoporosis: an international epidemiological investigation. J Intern Med 260:245–254
15. Barnett AG, van der Pols JC, Dobson AJ (2005) Regression to the mean: what it is and how to deal with it. Int J Epidemiol 34:215–220
16. Sonderman JS, Munro HM, Blot WJ, Signorello LB (2012) Reproducibility of serum 25-hydroxyvitamin d and vitamin D-binding protein levels over time in a prospective cohort study of black and white adults. Am J Epidemiol 176:615–621
17. Cauley JA1, Greendale GA, Ruppert K, Lian Y, Randolph JF Jr, Lo JC, Burnett-Bowie SA, Finkelstein JS (2015) Serum 25 hydroxyvitamin d, bone mineral density and fracture risk across the menopause. J Clin Endocrinol Metab 100:2046–2054
18. Tanaka S, Kuroda T, Yamazaki Y, Shiraki Y, Yoshimura N, Shiraki M (2014) Serum 25-hydroxyvitamin D below 25 ng/mL is a risk factor for long bone fracture comparable to bone mineral density in Japanese postmenopausal women. J Bone Miner Metab 32:514–523
19. Nieves JW, Barrett-Connor E, Siris ES, Zion M, Barlas S, Chen YT (2008) Calcium and vitamin D intake influence bone mass, but not shortterm fracture risk, in Caucasian postmenopausal women from the National Osteoporosis Risk Assessment (NORA) study. Osteoporos Int 19:673–679
20. Nieves JW (2013) Skeletal effects of nutrients and nutraceuticals, beyond calcium and vitamin D. Osteoporos Int 24:771–786
21. Nieves JW (2014) Bone. Maximizing bone health–magnesium, BMD and fractures. Nat Rev Endocrinol 10:255–256
22. Ministry of Health, Labour and Welfare. The report of National Health and Nutrition Survey 2005. http://www.mhlw.go.jp/bunya/kenkou/eiyou07/01.htm

J Orthop Sci (2015) 20:896–905
DOI 10.1007/s00776-015-0741-5

ORIGINAL ARTICLE

Association between new indices in the locomotive syndrome risk test and decline in mobility: third survey of the ROAD study

Noriko Yoshimura[1] · Shigeyuki Muraki[2] · Hiroyuki Oka[3] · Sakae Tanaka[4] · Toru Ogata[5] · Hiroshi Kawaguchi[6] · Toru Akune[5] · Kozo Nakamura[5]

Received: 5 December 2014 / Accepted: 29 May 2015 / Published online: 25 June 2015

Abstract

Background We aimed to clarify the association between new indices in a locomotive syndrome risk test and decline in mobility.

Methods In the third survey of the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study, data on the indices were obtained from 1575 subjects (513 men, 1062 women) of the 1721 participants in mountainous and coastal areas. As outcome measures for decline in mobility, we used the five-times-sit-to-stand test (FTSST) and walking speed with cutoff values of 12 s and 0.8 m/s, respectively.

Results We first estimated the prevalence of the indices in locomotive syndrome risk test stage 1, including two-step test score <1.3, difficulty with one-leg standing from a 40-cm-high seat in the stand-up test, and 25-question GLFS score ≥7, which were found to be 57.4, 40.6, and 22.6 %, respectively. Next, we investigated the prevalence of the indices in locomotive syndrome risk test stage 2, including two-step test score <1.1, difficulty with standing from a 20-cm-high seat using both legs in the stand-up test, and 25-question GLFS score ≥16, which were found to be 21.1, 7.9, and 10.6 %, respectively. Logistic regression analysis using slow FTSST time or slow walking speed as the objective factor, and presence or absence of indices as the independent factor, after adjusting for confounders, showed all three indices in both stages 1 and 2 were significantly and independently associated with immobility. Finally, we clarified the risk of immobility according to an increasing number of indices in both stages 1 and 2 and found that the odds ratio for both slow FTSST time and slow walking speed increased exponentially.

Conclusion We found that the three indices independently predicted immobility and that accumulation of indices increased the risk of immobility exponentially.

Electronic supplementary material The online version of this article (doi:10.1007/s00776-015-0741-5) contains supplementary material, which is available to authorized users.

✉ Noriko Yoshimura
yoshimuran-ort@h.u-tokyo.ac.jp

[1] Department of Joint Disease Research, 22nd Century Medical and Research Center, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan

[2] Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo 113-8655, Japan

[3] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo 113-8655, Japan

[4] Department of Orthopaedic Surgery, Sensory and Motor System Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

[5] National Rehabilitation Center for Persons with Disabilities, Saitama 359-0042, Japan

[6] JCHO Tokyo Shinjuku Medical Center, Tokyo 162-8542, Japan

Introduction

According to the most recent National Livelihood Survey by the Ministry of Health, Labour, and Welfare in Japan, osteoporotic fracture and falls is ranked fourth and osteoarthritis is ranked fifth among conditions that cause disability and subsequently require support with regard to activities of daily living [1]. Given the increasing proportion of elderly individuals in the Japanese population, a comprehensive and evidence-based prevention strategy for musculoskeletal

diseases is urgently required. In 2007, the Japanese Orthopaedic Association (JOA) proposed that the term "locomotive syndrome" should be adopted to designate a condition requiring nursing care, or being at risk of developing such a condition, because of a decline in mobility resulting from a disorder of the locomotive system, which consists of bones, joints, muscles, and nerves [2]. Weakness of such locomotive components causes difficulty in mobility—defined as the ability to stand, walk, run, climb stairs, and perform other physical functions essential to daily life.

As candidate indices to assess the risk of locomotive syndrome, in 2013, the JOA proposed the following three tests: two-step test, stand-up test, and 25-question geriatric locomotive function scale (GLFS) [3]. With regard to the stand-up test, more than 50 % of subjects younger than 70 years old can stand up on one leg from a 40-cm-high seat [3]. The 25-question GLFS has already been assessed regarding its sensitivity and specificity for prediction of disability and was assigned a cutoff value of 16 by Seichi et al. [4]. However, there is little information regarding reference and/or cutoff values for the two-step test.

Recently, the JOA determined clinical decision limits of these three indices for assessing risk of locomotive syndrome [5]. In their proposal, clinical decision limits were established in two stages as follows:

Stage 1:

1. Two-step test score <1.3.
2. Difficulty with one-leg standing from a 40-cm-high seat in the stand-up test (either leg).
3. 25-question GLFS score ≥7.

When a subject meets any of the above-mentioned conditions, he/she is diagnosed as starting to decline in mobility.

Stage 2:

1. Two-step test score <1.1.
2. Difficulty with standing from a 20-cm-high seat using both legs in the stand-up test.
3. A 25-question GLFS score ≥16.

When a subject meets any of the above-mentioned conditions, he/she is diagnosed as progressing to a decline in mobility.

However, no report has evaluated such indices using data of the general population. From 2005 to 2007, we started a large-scale, population-based cohort investigation entitled the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study, consisting of 3040 participants in three communities located in urban, mountainous, and coastal areas. Following the baseline study, we performed a second survey in the same communities from October 2008 to January 2010, followed by a third survey from October 2012 to December 2013. In the third survey, participants completed the two-step test, stand-up test, and 25-question GLFS. In the present report, using data from the third survey of the ROAD study, we assessed the usefulness of these new indices for predicting immobility, which causes subsequent disability.

Participants and methods

Participants

Measurements were obtained from participants of the third survey of the ROAD study. The ROAD study, which began in 2005, is a nationwide prospective study comprising population-based cohorts established in several communities in Japan. Recruitment methods for this study have been described in detail elsewhere [6, 7]. To date, we have created a baseline database including clinical and genetic information of 3040 inhabitants (1061 men; 1979 women) aged 23–95 years who were recruited from listings of resident registrations in three communities. All participants provided written informed consent, and the study was conducted with approval from the ethics committees of the participating institutions.

The third survey of the ROAD study began in 2012 and was completed in 2013. All participants in the baseline study and second survey were invited to participate in the third survey. Besides former participants, inhabitants aged ≥40 years who were willing to attend the ROAD survey performed in 2012–2013 also were included as participants in the third survey. As a result, a total of 2566 (837 men, 1729 women; urban area, 845 individuals; mountainous area, 769 individuals; coastal area, 952 individuals) residents participated in the third survey.

In the present study, we used data from 1575 subjects (513 men; 1062 women) who completed the stand-up test, two-step test, and 25-question GLFS for disability among all 1721 participants in mountainous and coastal areas in the third survey.

At the third survey, participants completed an interviewer-administered questionnaire. Five interviewers, who had been trained by an expert (NY), were provided for this study. The questionnaire consisted of 200 items that included lifestyle information, such as primary occupation, smoking habits, alcohol consumption, physical activity, medical history, and prescription medication. Anthropometric measurements included height (cm), weight (kg), body mass index [BMI, weight (kg)/height $(m)^2$], and hand grip strength (kg).

Indices for risk of decline in mobility resulting from locomotive syndrome

In the present study, participants performed the following tests for assessment of decline in mobility.

Two-step test

This test measures the stride length to assess walking ability, including muscle strength, balance, and flexibility of the lower limbs. The two-step test was performed using the following procedure [3, 8, 9]: (1) subjects determined the starting line and stood with the toes of both feet behind it; (2) subjects were instructed to take two long steps (as long as possible) and then align both feet; (3) the length of the two steps from the starting line to the tips of the subject's toes where he/she stopped was measured. The two-step test score was calculated using the following formula: length of the two steps (cm) ÷ height (cm).

Stand-up test

This test assesses leg strength by having the subject stand up on one or both legs from a specified height. After preparation of four seats of different heights—40, 30, 20, and 10 cm—the subject stood up from each seat (in descending height order), first with both legs then with one leg. If the subject could stand up without leaning back to gain momentum and maintain the posture for 3 s, then he/she was diagnosed as having passed that height level [3, 9]. In the present study, if the subject was unable to stand up on one leg (right or left) from a height of 40 cm, then his/her stand-up test was considered failed.

The 25-Question GLFS

As mentioned above, the 25-question GLFS was developed by Seichi et al. [4]. It is a self-administered, comprehensive measure, consisting of 25 items that include four questions regarding pain during the last month, 16 questions regarding activities of daily living during the last month, three questions regarding social functions, and two questions regarding mental health status during the last month. These 25 items are graded with a five-point scale, from no impairment (0 points) to severe impairment (4 points), and then arithmetically added to produce a total score (minimum = 0, maximum = 100). Thus, a higher score is associated with worse locomotive function. Validity of the scale has been assessed, and a cutoff point of 16 was determined to have the highest sensitivity and specificity for indication of disability resulting from locomotive syndrome [3, 4].

Indices for decline in mobility resulting from locomotive syndrome

Because the present study utilized a cross-sectional design, not a longitudinal follow-up design, we could not evaluate the ability of the stand-up test or two-step test for prediction of disability resulting from locomotive syndrome. Therefore, in the present study, we used the following outcome measures as indices for decline in mobility.

Five-times-sit-to-stand test

There are several reports that inability to rise from a chair five times within a determined time is associated with increased disability and morbidity [10–12]. We have also reported that the longer the standing time is, the higher the incidence of disability [13]. The five-times-sit-to-stand test (FTSST) was performed according to the following procedure: (1) using a straight-back chair with a solid seat, participants were asked to sit on the chair with their arms folded across their chest; (2) participants were instructed to stand up and sit down as quickly as possible five times, keeping their arms folded across their chest; (3) the time when the participant stood for the fifth time was measured. In the present study, we used a cutoff value of 12 s to indicate a decline in mobility [14].

The 6-m walking time

As another outcome measure for decline in mobility, participants walked a 6-m course at their usual speed. The method of measurement of walking time was identical to that performed in the large-scale cohort study entitled Osteoporotic Fractures in Men (MrOS), which started prior to the ROAD study [15]. In the present study, we used a cutoff value of 0.8 m/s to indicate a decline in mobility [16].

In the present study, among the above-mentioned indices, mean scores and SDs for the two-step test were calculated according to participants' sex and age strata (<40, 40s, 50s, 60s, 70s, and ≥80 years). Then, we estimated the prevalence of each index in stages 1 and 2. Finally, we assessed the association between the cumulative number of indices and decline in mobility using multivariate analysis.

Statistical analysis

All statistical analyses were performed using STATA statistical software (STATA, College Station, TX, USA). Differences in proportions were compared using the chi-square test. The significance of differences in continuous variables was evaluated using analysis of variance for comparisons

among multiple groups or Scheffe's least significant difference test for pairs of groups. All p values and 95 % confidence intervals are two-sided. A p value of <0.05 was considered statistically significant.

Logistic regression analysis was used to test the association of each factor with the presence or absence of a decline in mobility. In the analysis, we used presence of immobility according to the FTSST time (>12 s = 1; ≤12 s = 0) and usual walking speed (<0.8 m/s = 1; ≥0.8 m/s = 0) as the objective variable, and presence or absence of new indices in stages 1 and 2 as explanatory variables, after adjusting for age (+1 year), sex (men = 0, women = 1), BMI (+1 kg/m^2), and regional difference (mountainous area = 0; coastal area = 1). Other factors were considered in the multivariate model after simple linear analysis; those used as explanatory factors are described in "Results."

Results

Summary characteristics, including age and BMI of the participants, are shown in Table 1. Two-thirds of the 1575 participants were women, and the mean age of women participants was 1 year less than that of men participants; however, this difference was not significant. By contrast, there was a significant difference in BMI between sexes ($p < 0.0001$). Table 1 also shows the age and sex distributions of mean FTSST time and walking speed. Both values tended to be significantly slower in participants aged in their 70s and 80s in both men and women, and there were no significant differences between sexes. Table 1 also shows the age and sex distributions of mean two-step test scores. Mean two-step test score was 1.25 (SD 0.20) in men and 1.23 (SD 0.21) in women; this difference was significant ($p < 0.0001$). Age differences indicated that the two-step test score was significantly lower according to age in both men and women ($p < 0.05$).

First, the prevalence of the indices in stage 1 and their association with decline in mobility described by slow FTSST time and slow walking speed were assessed (Table 2). Overall, the prevalence of two-step test score <1.3, difficulty with one-leg standing from a 40-cm-high seat in the stand-up test, and 25-question GLFS score ≥7 were 57.4, 40.6, and 22.6 %, respectively.

Prevalence of two-step test score <1.3 in subjects aged in their 30s and younger, 40s, 50s, 60s, 70s, and 80 years and older were 17.0, 28.5, 32.6, 51.5, 76.2, and 90.0 %, respectively, indicating that the prevalence increased according to age; even in subjects aged in their 40s and 50s, the prevalence was more than 30 %. Prevalence of subjects who could not stand with one leg from a 40-cm-high seat who were aged in their 30s and younger, 40s, 50s,

Table 1 Mean (SD) values for age, body mass index (BMI), five-times-sit-to-stand test (FTSST) time, walking speed, and two-step test score of participants classified by age and sex

Age strata (years)	n	Age (years)	BMI (kg/m^2)	FTSST time (s)	Walking speed (m/s)	Two-step test score
Men						
<40	23	32.8 (4.8)	24.5 (3.3)	6.96 (1.33)	1.26 (0.22)	1.49 (0.14)
40–49	38	44.7 (3.1)	25.4 (5.1)	6.79 (2.41)	1.25 (0.25)	1.41 (0.15)
50–59	82	55.2 (2.5)	24.2 (3.3)	7.11 (1.47)	1.25 (0.26)	1.36 (0.13)[a]
60–69	137	64.3 (2.7)	23.8 (3.4)	8.10 (2.51)	1.16 (0.26)	1.29 (0.15)[ab]
70–79	139	74.3 (2.8)	23.4 (2.9)	8.72 (2.18)[bc]	1.02 (0.24)[abcd]	1.20 (0.16)[abcd]
≥80	94	83.8 (3.1)	22.3 (3.0)	11.48 (4.72)[abcde]	0.81 (0.28)[abcde]	1.06 (0.22)[abcde]
Total	513	66.2 (13.7)	23.6 (3.4)	8.57 (3.17)	1.08 (0.30)	1.25 (0.20)
Women						
<40	36	34.4 (4.8)	20.7 (3.0)	7.11 (1.26)	1.28 (0.17)	1.40 (0.14)
40–49	85	44.9 (2.9)	21.9 (3.2)	7.19 (1.64)	1.25 (0.22)	1.35 (0.11)
50–59	195	54.7 (2.9)	23.0 (4.1)	7.10 (1.94)	1.26 (0.22)	1.35 (0.13)
60–69	309	64.7 (2.9)	22.8 (3.4)	7.90 (2.31)	1.18 (0.23)[c]	1.28 (0.18)[abc]
70–79	303	74.3 (2.9)	23.3 (3.3)	9.44 (3.57)[abcd]	1.02 (0.28)[abcd]	1.16 (0.18)[abcd]
≥80	134	83.1 (3.0)	22.0 (3.4)	11.89 (4.60)[abcde]	0.75 (0.28)[abcde]	0.97 (0.23)[abcde]
Total	1062	65.3 (12.6)	22.7 (3.5)	8.58 (3.31)	1.11 (0.30)	1.23 (0.21)

[a] Significantly different ($p < 0.05$) from values of those aged <40 years

[b] Significantly different ($p < 0.05$) from values of those aged in their 40s

[c] Significantly different ($p < 0.05$) from values of those aged in their 50s

[d] Significantly different ($p < 0.05$) from values of those aged in their 60s

[e] Significantly different ($p < 0.05$) from values of those aged in their 70s

Table 2 Prevalence of indices in the locomotive syndrome risk test (stage 1): two-step test score <1.3, difficulty with one-leg standing from 40-cm-high seat in the stand-up test (either leg), and 25-question geriatric locomotive function scale (GLFS) score ≥7 in participants classified by age and sex

Age strata (years)	Age (years) mean (SD)	Two-step test score <1.3 (%)	Difficulty with one-leg standing from 40-cm-high seat (either leg) (%)	25-question GLFS score ≥7 (%)
Men				
<40	32.8 (4.8)	13.0	4.4	4.4
40–49	44.7 (3.1)	21.1	15.8	10.8
50–59	55.2 (2.5)	34.6	15.9	7.4
60–69	64.3 (2.7)	49.3	30.7	12.0
70–79	74.3 (2.8)	71.7	47.8	19.9
≥80	83.8 (3.1)	84.6	78.0	44.0
Total	66.2 (13.7)	55.6	39.1	18.8
Women				
<40	34.4 (4.8)	19.4	11.1	0.0
40–49	44.9 (2.9)	31.8	12.9	8.3
50–59	54.7 (2.9)	31.8	23.6	13.0
60–69	64.7 (2.9)	52.4	33.9	19.7
70–79	74.3 (2.9)	78.3	56.2	31.6
≥80	83.1 (3.0)	93.8	78.1	54.3
Total	65.3 (12.6)	58.2	41.3	24.5*

* Significantly different ($p < 0.05$) from values of men

Table 3 Effect of presence of indices in the locomotive syndrome risk test (stage 1) for decline in mobility described by slow five-times-sit-to-stand test (FTSST) time and slow walking speed

Indices for decline in mobility	Reference	OR (95 % CI)	*p* value
FTSST time >12 s			
Two-step test score <1.3	Yes vs. no	3.28 (1.81–5.97)	<0.001
Difficulty with one-leg standing from 40-cm-high seat (either leg)	Yes vs. no	1.78 (1.17–2.69)	0.007
25-question GLFS score ≥7	Yes vs. no	2.51 (1.73–3.62)	<0.001
Walking speed <0.8 m/s			
Two-step test score <1.3	Yes vs. no	4.24 (2.18–8.22)	<0.001
Difficulty with one-leg standing from 40-cm-high seat (either leg)	Yes vs. no	2.01 (1.35–3.16)	0.001
25-question GLFS score ≥7	Yes vs. no	2.65 (1.82–3.86)	<0.001

After adjusting for age, sex, body mass index, and regional difference. Presence or absence of indices for stage 1 also was mutually adjusted

CI confidence interval, *GLFS* geriatric locomotive function scale, *OR* odds ratio

60s, 70s, and 80 years and older were 8.5, 13.8, 21.3, 32.9, 53.5, and 78.1 %, respectively, indicating that the prevalence increased according to age, similar to the two-step test; even in subjects aged in their 40s and 50s, the prevalence was around 20 %. Prevalence of 25-question GLFS score ≥7 in participants aged in their 30s and younger, 40s, 50s, 60s, 70s, and 80 years and older were 1.7, 9.1, 11.4, 17.4, 27.9, and 50.0 %, respectively, indicating the overall prevalence was lower than that of the other two indices, but it increased synergistically in those in their 80s and older. Regarding the sex difference of the indices in stage 1, although there were no significant differences between men and women with regard to two-step test score <1.3 and difficulty with one-leg standing from a 40-cm-high seat in the stand-up test, the prevalence of 25-question GLFS score ≥7 in women was significantly higher than that in men ($p < 0.05$).

Table 3 shows the results of logistic regression analysis using immobility described by slow FTSST time or slow walking speed as the objective factor and the presence or absence of new indices in stage 1 for a decline in mobility as explanatory factors, after adjusting for age (+1 year),

sex (men = 0; women = 1), BMI (+1 kg/m^2), and regional difference (mountainous area = 0; coastal area = 1). The analysis revealed that all three indices in stage 1 independently predicted immobility described by both slow FTSST time and slow walking speed.

Table 4 shows the association between accumulation of the indices in stage 1 and decline in mobility described by slow FTSST time or slow walking speed, after adjusting for age (+1 year), sex (men = 0; women = 1), BMI (+1 kg/m^2), and regional difference (mountainous area = 0; coastal area = 1). The analysis revealed that according to an increasing number of indices, the odds ratio of both slow FTSST time and slow walking speed increased exponentially.

Table 4 Effect of accumulation of indices (stage 1) for decline in mobility

No. of indices	OR (95 % CI)	*p* value
Five-times-sit-to-stand test time >12 s		
0 = reference	1.0	–
1	2.19 (0.98–4.87)	0.055
2	2.90 (1.30–6.47)	0.009
3	11.59 (5.18–25.93)	<0.001
Walking speed <0.8 m/s		
0 = reference	1.0	–
1	5.73 (1.71–19.16)	0.005
2	9.82 (2.96–32.52)	<0.001
3	32.21 (9.64–107.7)	<0.001

After adjusting for age, sex, body mass index, and regional difference

CI confidence interval, *OR* odds ratio

Next, the prevalence of the indices in stage 2 and their association with decline in mobility described by a slow FTSST time and slow walking speed were assessed (Table 5). Overall, the prevalence of two-step test score <1.1, difficulty with standing from a 20-cm-high seat using both legs in the stand-up test, and 25-question GLFS score ≥16 were 21.1, 7.9, and 10.6 %, respectively.

Prevalence of two-step test score <1.1 in subjects aged in their 30s and younger, 40s, 50s, 60s, 70s, and 80 years and older were 0.0, 1.6, 3.3, 11.3, 28.4, and 65.8 %, respectively, indicating that the prevalence was less than 5 % in those aged in their 50s and younger, around 10 % in those aged in their 60s, but more than 50 % in those aged 80 years and older. Prevalence of subjects who could not stand from a 20-cm-high seat using both legs who were aged in their 30 s and younger, 40s, 50s, 60s, 70s, and 80 years and older were 0.0, 0.8, 0.7, 5.0, 9.9, and 25.1 %, respectively, indicating that the prevalence was less than 5 % in those aged in their 60s and younger, around 10 % in those aged in their 70s, but dramatically increased in those aged 80 years and older. Prevalence of 25-question GLFS score ≥16 in participants aged in their 30s and younger, 40s, 50s, 60s, 70s, and 80 years and older were 0.0, 1.7, 3.3, 5.0, 12.9, and

Table 5 Prevalence of indices in the locomotive syndrome risk test (stage 2): two-step test score <1.1, difficulty with standing from 20-cm-high seat using both legs in the stand-up test, and 25-question geriatric locomotive function scale (GLFS) score ≥16 in participants classified by age and sex

Age strata (years)	Age, mean (SD) years	Two-step test score <1.1 (%)	Difficulty with standing from 20-cm-high seat (%)	25-question GLFS score ≥16 (%)
Men				
<40	32.8 (4.8)	0.0	0.0	0.0
40–49	44.7 (3.1)	2.6	2.6	0.0
50–59	55.2 (2.5)	3.7	0.0	1.2
60–69	64.3 (2.7)	8.8	3.7	6.0
70–79	74.3 (2.8)	23.9	2.9	8.1
≥80	83.8 (3.1)	58.2	16.5	27.5
Total	66.2 (13.7)	20.1	4.9	9.0
Women				
<40	34.4 (4.8)	0.0	0.0	0.0
40–49	44.9 (2.9)	1.2	0.0	2.4
50–59	54.7 (2.9)	3.1	1.0	4.2
60–69	64.7 (2.9)	12.4	5.5	4.6
70–79	74.3 (2.9)	30.4	13.1	15.1
≥80	83.1 (3.0)	71.1	31.3	39.4
Total	65.3 (12.6)	21.6	9.4*	11.4

* Significantly different ($p < 0.01$) from values of men

Table 6 Effect of presence of indices in the locomotive syndrome risk test (stage 2) for decline in mobility described by slow five-times-sit-to-stand test (FTSST) time and slow walking speed

Indices for decline in mobility	Reference	OR (95 % CI)	*p* value
FTSST time >12 s			
Two-step test score <1.1	Yes vs. no	3.03 (1.97–4.66)	<0.001
Difficulty with standing from 20-cm-high seat %	Yes vs. no	5.87 (3.48–9.89)	<0.001
25-question GLFS score ≥16	Yes vs. no	2.83 (1.77–4.54)	<0.001
Walking speed <0.8 m/s			
Two-step test score <1.1	Yes vs. no	4.19 (2.75–6.39)	<0.001
Difficulty with standing from 20-cm-high seat %	Yes vs. no	3.40 (1.99–5.82)	<0.001
25-question GLFS score ≥16	Yes vs. no	3.49 (2.15–5.65)	<0.001

After adjusting for age, sex, body mass index, and regional difference. Presence or absence of indices for stage 2 also was mutually adjusted

CI confidence interval, *GLFS* geriatric locomotive function scale, *OR* odds ratio

34.4 %, respectively, indicating that the overall prevalence increased according to age; the prevalence was less than 5 % in subjects aged in their 60s and younger, around 10 % in those aged in their 70s, but dramatically increased in those aged 80 years and older. Regarding the sex difference of the indices in stage 2, although there were no significant differences between men and women with regard to two-step test score <1.1 and 25-question GLFS score ≥16, the prevalence of difficulty with standing from a 20-cm-high seat using both legs in the stand-up test in women was significantly higher than that in men ($p < 0.01$).

Table 6 shows the results of logistic regression analysis using immobility described by a slow FTSST time or slow walking speed as the objective factor and the presence or absence of new indices in stage 2 for decline in mobility as explanatory factors, after adjusting for age (+1 year), sex (men = 0; women = 1), BMI (+1 kg/m^2), and regional difference (mountainous area = 0; coastal area = 1). The analysis revealed that all three indices in stage 2 independently predicted immobility described by both a slow FTSST time and slow walking speed. The odds ratios of all indices in stage 2 for slow FTSST time and slow walking speed were greater than those of all indices in stage 1.

Table 7 shows the association between accumulation of the indices in stage 2 and decline in mobility described by slow FTSST time or slow walking speed, after adjusting for age (+1 year), sex (men = 0; women = 1), BMI (+1 kg/m^2), and regional difference (mountainous area = 0; coastal area = 1). The analysis revealed that according to an increasing number of indices, the odds ratio of both slow FTSST time and slow walking speed increased exponentially. The odds ratios of accumulation of the indices in stage 2 for slow FTSST time and slow walking speed were greater than those of accumulation of the indices in stage 1.

Table 7 Effect of accumulation of new indices (stage 2) for decline in mobility

No. of new indices	OR (95 % CI)	*p* value
Five-times-sit-to-stand test time >12 s		
0 = reference	1.0	–
1	2.99 (1.85–4.84)	<0.001
2	12.35 (7.08–21.54)	<0.001
3	46.87 (19.37–113.45)	<0.001
Walking speed <0.8 m/s		
0 = reference	1.0	–
1	3.50 (2.19–5.59)	<0.001
2	12.52 (7.08–22.13)	<0.001
3	61.93 (24.92–153.87)	<0.001

After adjusting for age, sex, body mass index, and regional difference

CI confidence interval, *OR* odds ratio

Discussion

In the present study, we clarified the associations of three new indices for immobility, including the two-step test, stand-up test, and 25-question GLFS, represented by slow FTSST time and slow walking speed. We first tested the associations among the three indices by using the clinical decision limits for locomotive syndrome risk test stage 1. Next, we tested the three indices in stage 2. We clarified the age and sex distributions of the prevalence of these three indices and found that the three indices in both stages 1 and 2 significantly and independently predicted a decline in mobility and that the accumulation of such indices increased the risk of immobility exponentially.

First, we used both FTSST time and 6-m walking speed as outcome measures of immobility. As mentioned in "Methods", these two indices are both regarded as predictors for morbidity and disability [10–12, 16]. In the ROAD

study, we reported that the longer the standing time, the higher the incidence of disability [13] and that slow walking speed was also a predictor for the occurrence of disability [13]. Although we could not clarify the direct associations between these new indices and occurrence of disability because this study design was cross-sectional, these surrogate indices could help predict disability in the near future. Therefore, if we could clarify the significant associations between these new indices and FTSST time and 6-m walking speed, we might clarify the ability to predict disability indirectly. Based on this hypothesis, we used cutoff values of 12 s for FTSST time [13] and 0.8 m/s for 6-m walking speed [16].

The two-step test was developed to assess walking ability, including muscle strength, balance, and flexibility of the lower limbs. This test was first developed by Muranaga and Hirano in 2003 [3, 8]. They compared two-step test scores of 108 healthy volunteers (38 men, 70 women; mean age 59.0 years) with those of 108 hospital outpatients (56 men, 52 women; mean age 60.3 years) and found that the two-step test score was significantly associated with the risk of falling and degree of independence in daily life [7]. In the present study, we clarified mean two-step test scores in participants classified by age and sex and found mean scores of 2.07 in men and 1.87 in women. Scores of men were significantly higher than those of women, and age differences indicated that scores were significantly lower according to age in both men and women. These age and sex tendencies are consistent with those reported in a previous study [3].

Regarding the prevalence of the indices in locomotive syndrome risk test stages 1 and 2, we found that all prevalences increased with age. However, the distribution of each index seemed to differ. In stage 1, the prevalence was highest for a two-step test score <1.3, followed by difficulty with one-leg standing from a 40-cm-high seat in the stand-up test and 25-question GLFS score ≥7. By contrast, in stage 2, the prevalence also was highest for a two-step test score <1.1, but the prevalence of a 25-question GLFS score ≥16 was higher than that of difficulty with standing from a 20-cm-high seat using both legs in the stand-up test. The different age distributions of these three indices in both stages 1 and 2 might be conducive to their mutually independent associations with immobility. Our result that these three indices in both stages 1 and 2 independently predicted immobility shows that all three are important for predicting immobility. Furthermore, because these indices independently predicted immobility, risk severity may be classified by a positive number of indices present. In fact, in our study, accumulation of indices increased the risk of immobility exponentially, especially accumulation of indices in stage 2, which suggests the possibility of categorizing the severity of risk for immobility, such as 0, normal; 1, mild; 2, moderate; 3, severe.

With regard to the 25-question GLFS, there might be some concern that it is too cumbersome for older people to answer 25 questions. Seichi et al. also proposed a short version of the test using only five questions with a cutoff score of ≥6, on behalf of a 25-question GLFS score ≥16. They selected five questions from the 25-question GLFS using the Akaike information criterion (AIC) [17–19]. They determined that a cutoff score of 6 had the lowest AIC value, representing the model with the best fit, and finally concluded that the five-question GLFS can be applied as a rapid self-check tool for locomotive syndrome [4]. We compared the results of 1535 individuals who completed both the 25-question GLFS and five-question GLFS. Supplementary Table 1 shows that these two indices (five-question GLFS and 25-question GLFS) have a strong association (sensitivity 0.859; specificity 0.985). This result shows the possibility of using the five-question GLFS instead of the 25-question GLFS as a rapid-check tool for the prediction of immobility.

There are several limitations in the present study. First, as mentioned above, our study assessed the usefulness of three indices in locomotive syndrome risk test stages 1 and 2 for predicting immobility using FTSST time and walking speed as outcome variables. Although there has been significant evidence regarding these outcome measures, such as slow FTSST time and slow walking speed, and disability [8, 10–13, 16], including our report, the direct associations of these new indices and occurrence of disability are unclear. The proposal of these new indices was published in 2013 [3], so there was not enough time to observe future occurrence of disability in our cohort. Therefore, we should continue to observe our cohort and assess the ability of such indices to predict disability directly. Second, although the ROAD study includes a large number of participants, these participants do not truly represent the general population, since the subjects in the present study were recruited from only two areas. However, we have already confirmed the representativeness of the participants of the ROAD baseline study to the Japanese population by comparing anthropometric measurements and frequencies of smoking and alcohol consumption between participants and the general Japanese population [5]. Values for the general population were obtained from the 2005 National Health and Nutrition Survey conducted by the Ministry of Health, Labour and Welfare in Japan [20]. Regarding anthropometric measurements, we found no significant differences between our participants and the total Japanese population, except that men participants aged 70–74 years in the ROAD study were significantly smaller in terms of body structure than the overall population ($p < 0.05$). In addition, proportions of current

smokers and current drinkers were significantly higher in the general Japanese population than in the study population [5], suggesting that participants of the ROAD study have healthier lifestyles than the general population. This "healthy" selection bias should be taken into consideration when using reference values obtained in the ROAD study.

In conclusion, we have assessed whether the proposed clinical decision limits of the indices in locomotive syndrome risk test stages 1 and 2 could predict immobility represented by a slow FTSST time and slow walking speed using data from a population-based cohort of the third survey of the ROAD study. We found that all the indices in the locomotive syndrome risk test—the two-step test, stand-up test, and 25-question GLFS—could significantly and independently predict a decline in mobility and that the accumulation of such indices increased the risk of immobility exponentially.

Acknowledgments This work was supported by a Grant-in-Aid for H17-Men-eki-009 (Director, Kozo Nakamura), H20-Choujyu-009 (Director, Noriko Yoshimura), H23-Choujyu-002 (Director, Toru Akune), H-25-Choujyu-007 (Director, Noriko Yoshimura), and H25-Nanchitou(Men)-005 (Director, Sakae Tanaka) of the Ministry of Health, Labour, and Welfare; and Scientific Research B23390172, B20390182, and Challenging Exploratory Research 24659317 to Noriko Yoshimura; B23390356, C20591774, and Challenging Exploratory Research 23659580 to Shigeyuki Muraki; Challenging Exploratory Research 24659666 and 21659349 and Young Scientists A18689031 to Hiroyuki Oka; B23390357 and C20591737 to Toru Akune; Collaborating Research with NSF 08033011-00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. This study also was supported by grants from the Japan Osteoporosis Society (Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, and Toru Akune) and research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 & 2010-2; Director, Hiroshi Kawaguchi). The authors wish to thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka of the Gobo Public Health Centre; Dr. Naoki Hirabayashi of the Kawakami Clinic in Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Centre; Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji, and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. We also thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, Mrs. Saeko Sahara, and Mr. Noriyuki Oe for their assistance in data reduction and administration.

Conflict of interest No conflicts of interest have been declared by the authors.

References

1. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey (2013). http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa13/dl/06.pdf. Accessed 17 June 2015
2. Nakamura K. A "super-aged" society and the "locomotive syndrome". J Orthop Sci. 2008;13(1):1–2.
3. Locomotive syndrome. In: Locomotive Challenge! Council, editors. Locomotive syndrome pamphlet 2013. Japanese Orthopaedic Association: Tokyo; 2013.
4. Seichi A, Hoshino Y, Doi T, Akai M, Tobimatsu Y, Iwaya T. Development of a screening tool for risk of locomotive syndrome in the elderly: the 25-question Geriatric Locomotive Function Scale. J Orthop Sci. 2012;17(2):163–72.
5. Locomotive syndrome. In: Locomotive Challenge! Council, editors. Locomotive syndrome pamphlet 2015. Tokyo: Japanese Orthopaedic Association; 2015 (in press).
6. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, Saika A, Suzuki T, Yoshida H, Kawaguchi H, Nakamura K, Akune T. Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab. 2009;27(5):620–8.
7. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability (road) study. Int J Epidemiol. 2010;39(4):988–95.
8. Muranaga S, Hirano K. Development of a convenient way to predict ability to walk, using a two-step test. J Showa Med Assoc. 2003;63(3):301–8 **(in Japanese)**.
9. Muranaga S. Evaluation of the muscular strength of the lower extremities using the standing movement and clinical application. J Showa Med Assoc. 2001;61(3):362–7 **(in Japanese)**.
10. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir GV, Studenski S, Berkman LF. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci. 2000;55(4):M221–31.
11. Buatois S, Miljkovic D, Manckoundia P, Gueguen R, Miget P, Vançon G, Perrin P, Benetos A. Five times sit to stand test is a predictor of recurrent falls in healthy community-living subjects aged 65 and older. J Am Geriatr Soc. 2008;56(8):1575–7.
12. Bohannon RW. Reference values for the five-repetition sit-to-stand test: a descriptive meta-analysis of data from elders. Percept Mot Skills. 2006;103(1):215–22.
13. Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Tokimura F, Yoshida H, Suzuki T, Nakamura K, Yoshimura N. Incidence of certified need of care in the long-term care insurance system and its risk factors in the elderly of Japanese population-based cohorts: the ROAD study. Geriatr Gerontol Int. 2014;14(3):695–701.
14. Mong Y, Teo TW, Ng SS. 5-repetition sit-to-stand test in subjects with chronic stroke: reliability and validity. Arch Phys Med Rehabil. 2010;91(3):407–13.
15. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, Lewis C, Cawthon PM, Marcus R, Marshall LM, McGowan J, Phipps K, Sherman S, Stefanick ML, Stone K. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study—a large observational study of the determinants of fracture in older men. Contemp Clin Trials. 2005;26(5):569–85.
16. Chen LK, Liu LK, Woo J, Assantachai P, Auyeung TW, Bahyah KS, Chou MY, Chen LY, Hsu PS, Krairit O, Lee JS, Lee WJ,

Lee Y, Liang CK, Limpawattana P, Lin CS, Peng LN, Satake S, Suzuki T, Won CW, Wu CH, Wu SN, Zhang T, Zeng P, Akishita M, Arai H. Sarcopenia in Asia: consensus report of the Asian Working Group for Sarcopenia. J Am Med Dir Assoc. 2014;15(2):95–101.

17. Akaike H. Information theory and an extension of the maximum likelihood principles. In: Petrov B, Caski F, editors. The second international symposium on information theory. Budapest: Akademia Kiado; 1973. p. 267–81.
18. Sakamoto Y, Ishiguro M, Kitagawa G. Akaike information criterion statistics. Dordrecht: Reidel; 1986.
19. FORTRAN Program CATDAP-02. In: Sakamoto Y, ed. Categorical data analysis by AIC. Tokyo: KTK Scientific Publishers, Kluwer Academic Publishers; 1991: 163–206.
20. Ministry of Health, Labour and Welfare. The report of National Health and Nutrition Survey (2005). http://www.mhlw.go.jp/bunya/kenkou/eiyou07/dl/01-04.pdf. Accessed 17 June 2015

J Orthop Sci (2015) 20:811–817
DOI 10.1007/s00776-015-0743-3

ORIGINAL ARTICLE

Development of a support tool for the clinical diagnosis of symptomatic lumbar intra- and/or extra-foraminal stenosis

Hiroshi Yamada[1] · Hiroyuki Oka[2] · Hiroshi Iwasaki[1] · Toru Endo[1] · Masahiko Kioka[1] · Yuyu Ishimoto[1] · Keiji Nagata[1] · Noboru Takiguchi[1] · Hiroshi Hashizume[1] · Akihito Minamide[1] · Yukihiro Nakagawa[1] · Masaki Kawai[1] · Shunji Tsutsui[1] · Munehito Yoshida[1]

Received: 12 December 2014 / Accepted: 8 June 2015 / Published online: 25 June 2015

Abstract
Background Not all lumbar intra- and/or extra-foraminal stenosis (LIEFS) on MRI is symptomatic. Therefore, the establishment of clinical diagnostic tools that can identify patients with symptomatic LIEFS is crucial in the clinical setting. The aim of this study was to develop a support tool for clinical diagnosis of LIEFS.
Methods Patients with L5 radiculopathy alone were prospectively enrolled. Fifty-one patients with lumbar spinal canal stenosis only at the L4–5 level and 49 patients with LIEFS only at the L5–S1 level were extracted from this cohort. We compared the two groups with regard to 12 variables—three subjective and three objective items from the Japanese Orthopaedic Association (JOA) score; Kemp's sign; results of the lumbar flexion test, Bonnet test, and Freiberg test; pain on sitting; and pain when recumbent—to determine which factors were associated with a high index of clinical suspicion of LIEFS.
Results The significant predictors of a final diagnosis of LIEFS were identified as follows: pain when recumbent, Freiberg and Bonnet test results, and pain on sitting. To develop a diagnostic tool, a scoring system (0–20 points) was formulated on the basis of the contribution ratios of these risk factors. To determine the contribution ratio, an integer score was assigned to the identified risk factors as follows: pain when recumbent = 9 points, Freiberg = 5 points, Bonnet = 3 points, and pain on sitting = 3 points. The Hosmer–Lemeshow statistic for this scoring system was $p = 0.063$, and confirmed that it was a good model. Receiver operating characteristic (ROC) curve analysis demonstrated a cut-off value of 5 points, an area under the ROC curve of 0.87435, sensitivity of 75.5 %, and specificity of 82.3 %.
Conclusions We believe that the use of this tool in the clinical setting will improve the accuracy of diagnosing symptomatic LIEFS, which will lead to improved quality of patient care.

✉ Hiroshi Yamada
yamacha@wakayama-med.ac.jp; taigis54@yahoo.co.jp

[1] Department of Orthopedic Surgery, Wakayama Medical University, 811-1, Kimiidera, Wakayama, Wakayama 641-8509, Japan

[2] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

Introduction

Spinal stenosis can affect any part of the lumbar spine, but is most prevalent in the spinal canal, which consists of the central canal and lateral recess; this type of spinal stenosis has been termed lumbar spinal canal stenosis (LSCS). Moreover, spinal stenosis can also affect the intra-foraminal zone (defined as the area from the medial pedicle wall to the lateral pedicle wall) and the extra-foraminal zone (defined as the area beyond the lateral pedicle wall), which is referred to as lumbar intra- and/or extra-foraminal stenosis (LIEFS) [1]. In the area where the nerve root exits the spinal canal through the intervertebral foramen, pathologic changes (e.g., a bony spur or bulged annulus) that may have developed as a result of disc degeneration can impinge on or compress the surrounding area of the dorsal root ganglion (DRG) [2–5]. LIEFS is a leading cause of intense radicular symptoms, reported to occur in 8–11 % of surgical cases involving lumbar degenerative disease [6–8], and thus it should always be considered in the treatment of these diseases.

The significance of LIEFS in failed back surgery has long been debated [9, 10]. The condition is well known as a frequent cause of persistent radicular pain following decompression, as it is difficult to diagnose preoperatively [3]. Although parasagittal and coronal magnetic resonance (MR) images provide useful information regarding abnormalities around the intervertebral foramen [11, 12], not all cases of LIEFS detected through magnetic resonance imaging (MRI) are symptomatic [13]. Hence, LIEFS is often difficult to differentiate from LSCS, particularly when the nerve root is compressed at two adjacent sites—the spinal canal and intervertebral foramen—creating a so-called double-crush lesion [14]. To our knowledge, no useful method is currently available for identifying which level is responsible for the compression. Although securing decompression of both sites should be considered in a surgical strategy for patients with a double-crush lesion, this approach can lead to unnecessary treatment.

Establishing a diagnostic tool for identifying symptomatic LIEFS is crucial in the clinical setting, in order to prevent persistent radicular pain following decompression. In the present study, we aimed to clarify the unique clinical signs and symptoms of symptomatic LIEFS and to develop a support tool for its diagnosis.

Materials and methods

The objective of the present study was to discriminate between LIEFS and LSCS, which was diagnosed definitively using MRI and through surgical outcomes, by means of ordinary tests that can be performed easily in daily practice.

Because false-positive/negative results can occur with MRI, in cases where surgery on the preoperatively diagnosed affected level failed to improve symptoms, we considered the possibility of an inappropriate presurgical diagnosis. Conversely, when symptoms improved with surgery, we considered that an appropriate diagnosis had been made. Thus we examined only cases of successful surgery in the present study. Subjects in whom only LIEFS was detected on MRI and who had no lesions in the spinal canal were considered case subjects, whereas subjects in whom LIEFS was not detected on MRI and who had lesions only in the spinal canal were considered control subjects. The study was conducted with approval from our institution's ethics committees, and all participants provided written informed consent.

The clinical data of 1576 consecutive patients who underwent lumbar spine surgery between April 2007 and March 2011 at our institution were used in the present study. To clarify the unique clinical signs and symptoms associated with the nerve compression site, 217 consecutive surgical cases with only unilateral L5 radiculopathy were enrolled. In the absence of a universally accepted gold standard for diagnosis of lumbar spinal stenosis, including LSCS and LIEFS, the impressions of expert clinicians provide a reasonable method for establishing a clinical diagnosis. Hence, we employed this best-available gold standard to determine L5 radiculopathy based on patient history, routine physical examinations, and imaging studies, including lumbar radiography, selective radiculography, three-dimensional computed tomography (3D-CT), MRI, and three-dimensional MRI (3D-MRI).

In the imaging study, LSCS was confirmed through structural evidence of lateral recess stenosis at the L4–5 level using a combination of MRI and 3D-CT. We excluded cases of central–lateral spinal stenosis, wherein not only the L5 nerve root but also the cauda equina is compressed, as well as cases of combined stenosis with disc herniation. Confirmation of LIEFS was achieved using 3D-CT and 3D-MRI. A transverse path of the nerve root and/or spinal nerve, obscurity of the DRG, and spinal nerve indentation on 3D-MRI were considered a positive finding for LIEFS [15]. However, cases with combined lateral disc herniation were excluded. The inter- and intra-observer reliability of 3D-MRI readings were 0.4092 and 0.8968, respectively.

The final diagnosis in each patient was established using ipsilateral L5 nerve root blocking. A transforaminal injection was performed with 1 % lidocaine in a volume of 2 mL and 4 mg of betamethasone sodium phosphate. In order to eliminate the effect of the steroid, the pain score on the visual analog scale (VAS 100-mm method) was examined for the first 30 min after blocking. When the VAS of leg pain improved to less than 20 mm, it was considered a positive finding.

Operative findings and follow-up notes were also used to determine whether stenosis was confirmed intraoperatively and whether symptoms improved following surgery. The presence of a VAS pain score of >20 mm at the final follow-up was defined as an unsatisfactory result. Patients who did not achieve a satisfactory result were excluded from the study, given the possibility of an inappropriate presurgical diagnosis.

From this cohort, we selected patients with LSCS only at the L4–5 level and patients with LIEFS only at the L5–S1 level for the current analysis. Exclusion criteria included the following: LSCS at two or more levels, double-crush lesion, cervical or thoracic compressive myelopathy, previous lumbar spinal surgery, traumatic disorder, peripheral nerve neuropathy, and peripheral artery or inflammatory disease. All patients underwent posterior lumbar spinal decompression surgery with the use of a spinal microendoscope in the spinal canal at the L4–5 level or in the

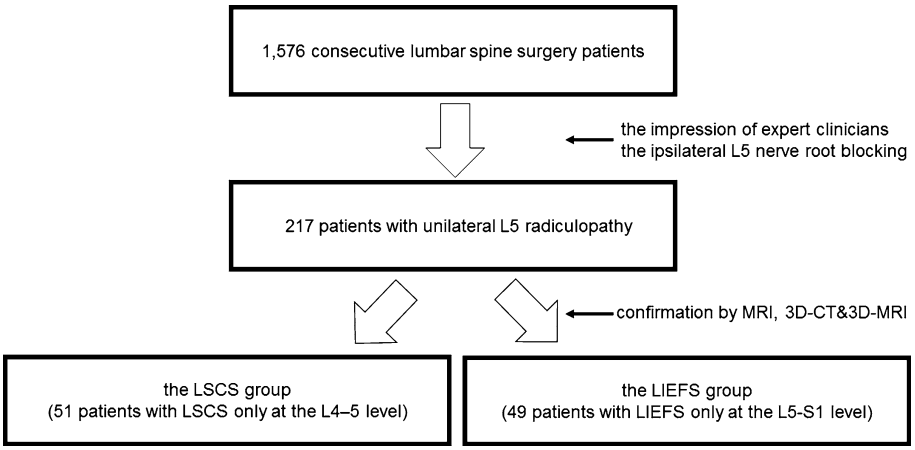


Fig. 1 Flow diagram depicting participants recruited to this study

intervertebral foramen at the L5–S1 level. Finally, a total of 100 patients were selected for the current analysis: 51 patients with LSCS only at the L4–5 level (LSCS group) and 49 patients with LIEFS only at the L5–S1 level (LIEFS group) (Fig. 1).

We compared the two groups with regard to 12 variables—three subjective and three objective items from the Japanese Orthopaedic Association (JOA) score; Kemp's sign; results of the lumbar flexion test, Bonnet test, and Freiberg test; pain on sitting; and pain when recumbent—to determine which factors were associated with a high index of clinical suspicion of LIEFS.

Outcome measures

We chose the 12 variables for the present study because they are simple tests that are used in daily practice. Subjective and objective items from the JOA score are shown in Table 1. We considered a score of 1 or 0 on the subjective JOA items and a score of 0 on the objective JOA items as positive findings for symptomatic LIEFS, and other scores were considered negative findings. When patients experienced buttock or leg pain with or without low back pain on sitting and experienced difficulty in maintaining this posture, pain on sitting was judged as a positive finding. When patients experienced buttock or leg pain in the supine position or in the lateral decubitus position on the affected side and experienced difficulty in maintaining this posture, pain when recumbent was judged as a positive finding. The pain provocation tests were performed in a standard manner. We modified the tests, which were originally described in previous publications, as follows: for Kemp's sign [16], we passively obliquely extended the patient's lumbar spine onto the affected side; in the lumbar flexion test, the patient was asked to lean forward with his/her feet together and knees

Table 1 Scoring system for low back pain proposed by the Japanese Orthopedic Association (maximum of 29 points)

Items	Score
Subjective symptoms (9 points)	
Low back pain	
None	3
Occasionally mild	2
Always present or sometimes severe	1
Always severe	0
Leg pain and/or numbness	
None	3
Occasionally mild	2
Always present or sometimes severe	1
Always severe	0
Walking ability	
Normal walking	3
Able to walk more than 500 m, pain/numbness/weakness present	2
Unable to walk 500 m due to pain/numbness/weakness	1
Unable to walk 100 m due to pain/numbness/weakness	0
Objective findings (6 points)	
Straight leg raising	
Normal	2
30–70°	1
<30°	0
Sensory function	
Normal	2
Mild sensory disturbance	1
Apparent sensory disturbance	0
Motor function	
Normal (normal MMT grades)	2
Slight decrease muscle strength (good MMT grades)	1
Markedly decrease muscle strength (MMT grades less than fair)	0

MMT manual muscle testing

straight as if he/she were to touch his/her toes; for the Bonnet test [16], the patient was placed in a supine position with his/her leg straight, raised to an angle of 70°, and the examiner forced the patient's hip to fully adduct and internally rotate; and in the Freiberg test [17, 18], the patient was placed in a supine position with his/her leg extended at an angle of 70°, and the examiner forced the patient's hip to fully adduct and internally rotate. Patients were asked to maintain these maneuvers for 30 s. If pain radiated to the buttock or lower extremity on the affected side during the maneuvers, the result was considered positive. Each examiner received an explanation of the standardized definitions of these provocation tests from the first author (HY), and consensus on the interpretation of test results was obtained among the examiners before the start of the study.

Statistical methods

We compared the LIEFS and LSCS groups using the Chi-square test for categorical variables and the one-factor analysis of variance test for numerical variables. Questionnaire items for the LIEFS diagnostic support tool were evaluated using simple logistic regression, and odds ratios were calculated. In the next step, variables with a P value of <0.2 in univariate analyses were included in a stepwise multivariate logistic regression model. We identified the significant ($P < 0.05$) predictors of a final diagnosis of LIEFS, and removed any variable with a P value of >0.05 in the final model. Discriminatory power, or the ability to identify the likelihood of patient outcomes, was determined using receiver operating characteristic (ROC) curve analysis, with an area of 1.00 under the ROC curve indicating perfect discrimination and an area of 0.50 indicating the complete absence of discrimination. Finally, to examine the performance of the support tool, we calculated its sensitivity and specificity. Statistical analyses were conducted using SAS version 9.2 statistical software (SAS Institute Inc., Cary, NC, USA). All statistical tests were two-tailed, and a significance level of 0.05 was used.

Results

As shown in Table 2, no differences were observed between the LIEFS and LSCS groups regarding the three subjective and three objective items from the JOA score. However, with the exception of the lumbar flexion test, differences in the pain provocation test results were significant (Kemp's sign, $P = 0.040$; Bonnet test, $P < 0.0001$; Freiberg test, $P < 0.0001$). The results for pain on sitting and pain when recumbent also differed significantly between the two groups (pain on sitting, $P < 0.0001$; pain when recumbent, $P < 0.0001$). Notably, pain when recumbent seemed to be more extreme in the LIEFS group than in the LSCS group.

Table 2 Participant characteristics

	LIEFS ($n = 49$)	LSCS ($n = 51$)	P value
Gender (male)	55.1 %	47.1 %	0.421
Age (years; mean ± SD)	68.7 ± 8.3	65.3 ± 10.4	0.072
BMI (mean ± SD)	23.6 ± 3.5	24.2 ± 3.1	0.385
Low back pain	49.0 %	43.1 %	0.558
Leg pain and/or numbness	87.8 %	86.3 %	0.826
Walking ability: IC (+)	69.4 %	76.5 %	0.425
Straight leg raising (+)	14.3 %	9.8 %	0.491
Sensory function: disturbance (+)	59.2 %	70.6 %	0.232
Motor function: disturbance (+)	24.5 %	13.7 %	0.170
Kemp signs (+)	79.6 %	60.8 %	0.040
Lumbar flexion test	2.0 %	0.0 %	0.305
Bonnet test (+)	65.3 %	19.6 %	<0.0001
Freiberg test (+)	44.9 %	7.8 %	<0.0001
Pain on sitting (+)	63.3 %	17.6 %	<0.0001
Pain when recumbent (+)	49.0 %	6.0 %	<0.0001

LIEFS lumbar intra- and/or extra-foraminal stenosis, *LSCS* lumbar spinal canal stenosis, *BMI* body mass index, *IC* intermittent claudication

We then conducted stepwise multivariate logistic regression analysis to select the most significant set of items predicting symptomatic LIEFS (Table 3). After variables with a P value of <0.2 in the univariate analyses were included in the model, the significant ($P < 0.05$) predictors of a final diagnosis of LIEFS were identified as follows: pain when recumbent ($P = 0.0009$), positive Freiberg test result ($P = 0.028$), positive Bonnet test result ($P = 0.0271$), and pain on sitting ($P = 0.0309$). The Hosmer–Lemeshow statistic for this analysis was $P = 0.063$, indicating that the model was good.

To develop a simple clinical diagnostic tool from the results of this analysis, an integer score derived from the β-coefficient was assigned to each identified risk factor, as follows: pain when recumbent, 9 points; positive Freiberg test result, 5 points; positive Bonnet test result, 3 points; and pain on sitting, 3 points (Table 4). For each patient, all applicable risk score values were summed to attain a total risk score, which ranged from 0 to 20. The results of ROC analysis were as follows: cut-off value, 5 points; area under the ROC curve, 0.87435; sensitivity, 75.5 %; and specificity, 82.3 % (Fig. 2).

Table 3 Univariate analyses for factors associated with the diagnosis of lumbar intra- and/or extra-foraminal stenosis

	Odds ratio	95 % CI	*P* value
Low back pain	1.26	0.58–2.80	0.558
Leg pain and/or numbness	1.14	0.35–3.80	0.826
Walking ability: IC (+)	0.70	0.28–1.69	0.425
Straight leg raising (+)	1.53	0.46–5.53	0.490
Sensory function: disturbance (+)	0.60	0.26–1.38	0.231
Motor function: disturbance (+)	2.04	0.74–5.98	0.168
Kemp signs (+)	2.52	1.05–6.35	0.039
Lumbar flexion test	1565190.8	0.18–N.D.	0.231
Bonnet test (+)	7.72	3.21–19.9	<0.0001
Freiberg test (+)	9.57	3.26–35.4	<0.0001
Pain on sitting (+)	8.04	3.30–21.2	<0.0001
Pain when recumbent (+)	15.4	4.79–69.2	<0.0001

LIEFS lumbar intra- and/or extra-foraminal stenosis, *IC* intermittent claudication

Table 4 Multivariable predictors for the diagnosis of lumbar intra- and/or extra-foraminal stenosis and the associated risk scoring system as a support tool

Characteristic	Regression β-coefficient	95 % CI	Risk score[a]
Bonnet test (+)	3.45	1.15–10.6	3
Freiberg test (+)	4.51	1.17–19.8	5
Pain on sitting (+)	3.28	1.08–10.2	3
Pain when recumbent (+)	9.03	2.38–45.2	9

The Hosmer–Lemeshow statistic was 14.82 ($P = 0.063$)

[a] The score was obtained by rounding the raw score to one decimal place if the coefficient was statistically significant

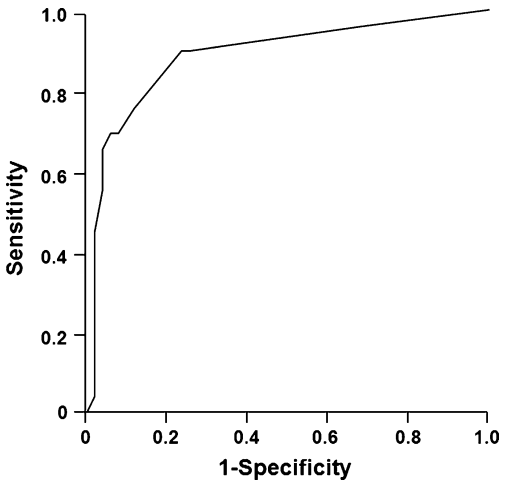


Fig. 2 Receiver operating characteristic curve for the lumbar intra- and/or extra-foraminal stenosis diagnostic support tool

Discussion

In the present study, we examined the unique clinical signs and symptoms in patients with symptomatic LIEFS. Due to the involvement of the DRG in these patients, they have generally been recognized as demonstrating more severe symptoms and signs than patients with LSCS [4, 8, 19–21]. As such, we defined a score of 1 or 0 on the subjective JOA items and a score of 0 on the objective JOA items as positive findings of symptomatic LIEFS. However, no differences were noted between the LIEFS and LSCS groups, thus indicating that the degree of severity of clinical signs and symptoms is not useful in differentiating symptomatic LIEFS from LSCS.

A high frequency of sciatic pain exacerbated by posterolateral bending of the lumbar spine (positive Kemp's sign) in LIEFS has been reported by several authors [8, 21]. In the present study, however, the frequency of a positive Kemp's sign in the LIEFS group was similar to that in the LSCS group. Hence, a positive Kemp's sign may not be a unique clinical feature of symptomatic LIEFS.

With regard to the North American Spine Society (NASS) clinical guidelines for degenerative LSCS, symptomatic LSCS has certain provocative and palliative features. Provocative features include exercise- or positionally induced neurogenic claudication, whereas the palliative features commonly include symptomatic relief with forward flexion, sitting, and/or recumbency [22]. These clinical characteristics contrast markedly with those of symptomatic herniated discs, in which forward flexion, sitting, and/or recumbency are well-known provocative features. Interestingly, LIEFS often has clinical features similar to those of herniated discs. Watanabe reported a high frequency of leg pain in a sitting position or at night in patients with LIEFS [21]. Kunogi, Porter, and Baba also reported the presence of these unique clinical findings in patients with LIEFS [6, 8, 19], where the frequency of leg pain under these two circumstances in their studies was found to be relatively high. In the present study, the frequency of pain on sitting and when recumbent demonstrated similar results, and significant differences were evident between the LSCS and LIEFS groups, which suggests that the clinical symptoms of pain on sitting and when recumbent are distinct characteristics of LIEFS. In contrast, results of the lumbar flexion test were negative in all patients in the LSCS group, and were positive in only two patients in the LIEFS group. Thus, a positive lumbar flexion test result was rarely observed in any type of stenosis, and is not a distinct characteristic of LIEFS, although it may be useful in distinguishing LSCS from herniated discs.

Positive Bonnet and Freiberg test results were also significant features of LIEFS in the present study. These tests are usually used to discriminate piriformis syndrome from

LSCS [23]. Piriformis syndrome is a neuromuscular disorder that occurs when the sciatic nerve is compressed or otherwise irritated by the piriformis muscle, causing pain, tingling, and numbness in the buttocks and along the path of the sciatic nerve descending through the lower thigh and into the leg. Although the exact mechanisms of the high frequency of positive results on our modified Bonnet and Freiberg tests are unknown, these procedures may enhance the compression or irritation of the DRG in the intervertebral foramen.

While these various characteristics may be useful in suggesting a diagnosis of symptomatic LIEFS, they are insufficient for use as a diagnostic tool for this condition, as their diagnostic performance when used in isolation was poor. The sensitivity of these clinical features is modest and not specific to LIEFS; therefore, we must conclude that single variables are inadequate for the diagnosis of symptomatic LIEFS. Hence, we developed a diagnostic tool using a scoring system (0–20 points) based on the contribution ratio of these risk factors. The Hosmer–Lemeshow statistic demonstrated that this scoring system was a good model, and ROC analysis also showed reasonable results.

This study has several limitations. First, this scoring system can be used only in patients with L5 lumbar radiculopathy. Future studies are needed to determine whether the tool can be used in patients with lumbar radiculopathy at other levels, as these may have different clinical features. However, a higher incidence of LIEFS has been noted in the lower lumbar segments. An and colleagues reported that the fifth lumbar root (75 %) is most commonly involved in LIEFS, followed by the fourth (15 %), third (5.3 %), and second (4 %) [20]. Based on these statistics, it is useful in general practice to understand the manner in which the level responsible for L5 lumbar radiculopathy can be analyzed. The second limitation of this scoring system is that it has yet to be validated in the clinical setting. However, as we encountered no problems in deciding where decompression should be performed in patients in the present study, the tool that we have designed and investigated is informative in a narrow population in which it is not required to guide clinical decision-making. Further research is needed to determine whether this tool is useful in patients with combined LSCS and LIEFS, and prospective randomized controlled studies should be conducted. Moreover, this scoring system will be validated if its use leads to more accurate diagnoses of symptomatic LIEFS and a reduction in the number of cases of persistent radicular pain following decompression.

In conclusion, we believe that use of this tool in the clinical setting will improve the accuracy of diagnosis of symptomatic LIEFS, thus leading to improved quality of patient care.

Compliance with ethical standard

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Arnoldi CC, Brodsky AE, Cauchoix J, Dommisse CGF, Edgar MA, Gargano FP, Jacobson RE, Kirkaldy-Willis WH, Kurihara A, Langenskiold A, Macnab I, McIvor WD, Paine KWE, Russin LA, Sheldon J, Tile M, Urist, Wilson WE, Wiltse LL. Lumbar spinal stenosis and nerve root entrapment syndromes: definition and classification. Clin Orthop Relat Res. 1976;115:4–5.
2. Briggs H, Krause J. The intervertebral foraminotomy for relief of sciatic pain. J Bone Joint Surg. 1945;27:475–8.
3. MacNab I. Negative disc exploration. An analysis of the causes of nerve root involvement in sixty-eight patients. J Bone Joint Surg Am. 1971;53(5):891–903.
4. Lee CK, Rauschning W, Glenn W. Lateral lumbar spinal canal stenosis: classification, pathologic anatomy, and surgical decompression. Spine. 1988;13(3):313–20.
5. Hasue M, Kunogi J, Konno S, Kikuchi S. Classification by position of dorsal root ganglia in the lumbosacral region. Spine. 1989;14(11):1261–4.
6. Porter R, Hibbert C, Evans C. The natural history of root entrapment syndrome. Spine. 1984;9(4):418–21.
7. Vanderlinden R. Subarticular entrapment of the dorsal root ganglion as a cause of sciatic pain. Spine. 1984;9(1):19–22.
8. Kunogi J, Hasue M. Diagnosis and operative treatment of intraforaminal and extraforaminal nerve root compression. Spine. 1991;16(11):1312–30.
9. Burton CV, Kirkaldy-Willis W, Yong-Hing K, Heithoff KB. Causes of failure of surgery on the lumbar spine. Clin Orthop Relat Res. 1981;157:191–9.
10. Schofferman J, Reynolds J, Herzog R, Covington E, Dreyfuss P, O'Beill C. Failed back surgery: etiology and diagnostic evaluation. Spine J. 2003;3:400–3.
11. Maher CO, Henderson FC. Lateral exit-zone stenosis and lumbar radiculopathy. J Neurosurg. 1999;90(1 Suppl):52–8.
12. Hashimoto M, Watanabe O, Hirano H. Extraforaminal stenosis in the lumbosacral spine. Efficacy of MR imaging in the coronal plane. Acta Radiol. 1996;37:610–3.
13. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am. 1990;72(3):403–8.
14. Yamada H, Yoshida M, Hashizume H, Minamide A, Nakagawa Y, Kawai M, Iwasaki H, Tsutsui S. Efficacy of novel minimally invasive surgery using spinal microendoscope for treating extraforaminal stenosis at the lumbosacral junction. J Spinal Disord Tech. 2012;25(5):268–76.
15. Yamada H, Terada M, Iwasaki H, Endo T, Okada M, Nakao S, Hashizume H, Minamide A, Nakagawa Y, Nishi H, Tsutsui S, Oka H, Yoshida M. Improved diagnostic accuracy of lumbar intra-and/or extra-foraminal stenosis by use of three dimensional MR imaging: comparison with conventional MR imaging. J Orthop Sci. 2015;20(2):287–94.
16. Miller KJ. Physical assessment of lower extremity radiculopathy and sciatica. J Chiropr Med. 2007;6:75–82.
17. Freiberg AH, Vinkle TH. Sciatica and the sacro-iliac joint. J Bone Joint Surg Am. 1934;16:126–36.
18. Nakamura H, Seki M, Konishi S, Yamano Y, Takaoka K. Pirifomis syndrome diagnosed by cauda equina action potentials: report of two cases. Spine. 2003;28(2):E37–40.

19. Baba H, Uchida K, Maezawa Y, Furusawa N, Okumura Y, Imura S. Microsurgical nerve root canal widening without fusion for lumbosacral intervertebral foraminal stenosis: technical notes and early results. Spinal Cord. 1996;34:644–50.
20. Jenis LG, An HS. Lumbar foraminal stenosis. Spine. 2000;25(3):389–94.
21. Watanabe K, Yamazaki A, Morita O, Sano A, Katsumi K, Ohashi M. Clinical outcomes of posterior lumbar interbody fusion for lumbar foraminal stenosis: preoperative diagnosis and surgical strategy. J Spinal Disord Tech. 2011;24(3):137–41.
22. Watters WC, Baisden J, Gilbert TJ, Kreiner S, Resnick DK, Bono CM, Ghiselli G, Heggeness MH, Mazanec DJ, O'Neill C, Reitman CA, Shaffer WO, Summers JT, Toton JF. Degenerative lumbar spinal stenosis: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis. Spine J. 2008;8:305–10.
23. Filler AG, Haynes J, Jordan SE, Prager J, Villablanca JP, Farahani K, McBride DQ, Tsuruda JS, Morisoli B, Batzdorf U, Johnson JP. Sciatica of nondisc origin and piriformis syndrome: diagnosis by magnetic resonance neurography and interventional magnetic resonance imaging with outcome study of resulting treatment. J Neurosurg Spine. 2005;2:99–115.

Muraki et al. BMC Musculoskeletal Disorders (2015) 16:305
DOI 10.1186/s12891-015-0737-5

BMC Musculoskeletal Disorders

RESEARCH ARTICLE Open Access

Quadriceps muscle strength, radiographic knee osteoarthritis and knee pain: the ROAD study

Shigeyuki Muraki[1*], Toru Akune[2], Masatoshi Teraguchi[3], Ryohei Kagotani[3], Yoshiki Asai[3], Munehito Yoshida[3], Fumiaki Tokimura[4], Sakae Tanaka[5], Hiroyuki Oka[6], Hiroshi Kawaguchi[7], Kozo Nakamura[2] and Noriko Yoshimura[8]

Abstract

Background: The objective of this study was to clarify the association of quadriceps muscle strength with knee pain using a large-scale, population-based cohort of the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.

Methods: From the 2566 subjects at the third visit of the ROAD study, the present study analyzed 2152 subjects who completed radiographic examinations and measurements of muscle strength and mass (690 men and 1462 women; mean age, 71.6 ± 12.2 years). Knee pain was assessed by an experienced orthopedist. Knee osteoarthritis (OA) was defined according to Kellgren-Lawrence (KL) grade. Quadriceps muscle strength and muscle mass at the lower limbs were measured by the Quadriceps Training Machine (QTM-05F, Alcare Co., Ltd. Tokyo, Japan) and the Body Composition Analyzer MC-190 (Tanita Corp., Tokyo, Japan), respectively.

Results: Quadriceps muscle strength and weight bearing index (WBI: quadriceps muscle strength by weight) were significantly associated with knee pain after adjustment for age and body mass index, whereas grip strength and muscle mass at the lower limbs were not. The significant association of quadriceps muscle strength with knee pain was independent of radiographic knee OA.

Conclusion: The present cross-sectional study showed an independent association of quadriceps muscle strength with knee pain.

Keywords: Cohort study, Epidemiology, Osteoarthritis, Pain, Muscle

Background

Knee osteoarthritis (OA) is a major public health issue that causes chronic pain and disability [1–3]. The prevalence of radiographic knee OA is high in Japan [4], with 25,300,000 persons aged 40 and older estimated to have radiographic knee OA [5]. According to the recent National Livelihood Survey of the Ministry of Health, Labour and Welfare in Japan, OA is ranked fourth among diseases that cause disabilities that subsequently require support with activities of daily living [6]. The principal clinical symptom of knee OA is knee pain [7]. Although much effort has been devoted toward a definition of knee pain, its correlation with radiographic severity of knee OA is not as strong as one would expect [4, 8–10]. In fact, our previous study showed that the odds ratio (OR) of severe knee OA defined as Kellgren-Lawrence (KL) grade 3 or 4 for knee pain was 8.6 in men and 4.4 in women [4], which was significant, but the OR was not as high as expected. In addition, 10 % of men without radiographic knee OA and 20 % of women without radiographic knee OA had knee pain [4]. This indicates that at least 10 % and 20 % of knee pain in men and women, respectively, may be explained by factors other than radiographic changes.

One of the factors contributing to knee pain other than radiographic knee OA may be quadriceps muscle weakness. Thus far, grip strength has been used as a useful clinical marker of sarcopenia [11], because measuring grip strength is easy. Although there is growing evidence that reduced grip strength is associated with adverse outcomes including morbidity [12], disability [13], falls [13], higher fracture rates [14], increased length of hospital stay [15], quality of life [16] and mortality [13], and grip strength is

* Correspondence: murakis-ort@h.u-tokyo.ac.jp
[1]Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan
Full list of author information is available at the end of the article

related to total muscle strength [17], quadriceps muscle strength may be more strongly associated with knee symptoms than grip strength. However, to the best of our knowledge, no population-based study has compared the association of knee pain with grip strength and quadriceps muscle strength because isokinetic devices such as Cybex, Biodex, and KIN-COM, which allow for the most detailed measurements regarding the quantitative evaluation of the quadriceps muscle strength, are expensive, large-scale, and impossible to move. Recently, the Quadriceps Training Machine (QTM) (QTM-05F, Alcare Co., Ltd. Tokyo, Japan) was developed to measure quadriceps muscle strength more easily [18]. The QTM has higher usability compared with other devices in terms of its small size, light weight, and good portability, as well as the fact that it has good correlation with Biodex and high credibility of measurements [18]. Although measurements of muscle mass are another method to evaluate muscle, the association between muscle strength and mass has been shown to be weak [19], indicating that a distinct association with knee symptoms between quadriceps muscle strength and muscle mass at the lower limb may be found. However, there are no population-based studies that compare the association of knee pain with quadriceps muscle strength and muscle mass at the lower limbs.

The objective of this study was to clarify the association of quadriceps muscle strength and muscle mass at the lower limbs with pain at the knee among Japanese men and women in a large-scale, population-based cohort from the Research on Osteoarthritis/osteoporosis Against Disability (ROAD) study.

Methods

Subjects

The ROAD study is a nationwide prospective study designed to establish epidemiologic indices for the evaluation of clinical evidence for the development of a disease-modifying treatment for bone and joint diseases (with OA and osteoporosis as the representative bone and joint diseases). It consists of population-based cohorts in several communities in Japan. A detailed profile of the ROAD study has been reported elsewhere [4, 5, 20], and thus, only a brief summary is provided here. To date, we have completed the creation of a baseline database including clinical and genetic information for 3040 inhabitants (1061 men and 1979 women) ranging in age from 23 to 95 years (mean, 70.3 years), who were recruited from resident registration listings in three communities: an urban region in Itabashi, Tokyo, a mountainous region in Hidakagawa, Wakayama, and a coastal region in Taiji, Wakayama. All participants provided written, informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology.

The third visit of the ROAD study began in 2011 and was completed in 2013. All participants in the baseline study were invited to participate in the third visit. In addition to the former participants, inhabitants aged ≥60 years in the urban area and those aged ≥40 years in the mountainous and coastal areas who were willing to participate in the ROAD survey performed in 2011–2013 were also included in the third visit.

Anthropometric measurements, including height and weight, were taken, and body mass index (BMI; weight [kg]/height^2 [m^2]) was calculated. Grip strength was measured on the right and left sides using a TOEI LIGHT handgrip dynamometer (TOEI LIGHTCO. LTD, Saitama, Japan). Isometric quadriceps muscle strength at the right and left knee was measured by the QTM one time each, and weight bearing index (WBI: quadriceps muscle strength/body weight) was calculated. Subjects carried out knee extension exercises by placing their knee joint on the QTM where specified; the load pressure applied to the QTM in the popliteal region was measured and displayed as the isometric knee extension muscle strength (quadriceps strength). The QTM has good correlation with Biodex and high credibility of measurement, and the method has been validated [18]. Lower limb muscle mass was measured by bioimpedance analysis [21–24] using the Body Composition Analyzer MC-190 (Tanita Corp., Tokyo, Japan), and muscle mass/height^2 (kg/m^2) was calculated. The protocol was described by Tanimoto and colleagues [25, 26], and the method has been validated [27].

All participants were also interviewed by well-experienced orthopedists regarding pain in both knees, by asking: "Have you experienced right knee pain on most days in the past month, in addition to now?" and "Have you experienced left knee pain on most days in the past month, in addition to now?". Subjects who answered "yes" were defined as having knee pain.

Radiographic assessment

All participants underwent radiographic examination of both knees using an anterior-posterior view with weight-bearing and foot map positioning by experienced radiological technologists. The beam was positioned parallel to the floor with no angle and aimed at the joint space. To visualize the joint space properly and to centralize the patella over the lower end of the femur, fluoroscopic guidance with an anterior-posterior X-ray beam was used, and the images were downloaded into Digital Imaging and Communication in Medicine (DICOM) format files. Knee radiographs were read without knowledge of participant clinical status by a single experienced orthopedist (S.M.) using the KL radiographic atlas for overall

knee radiographic grades [28], and knee OA was defined as KL grade 2 or greater. To evaluate the intraobserver variability of the KL grading, 100 randomly selected radiographs of the knee were scored by the same observer more than 1 month after the first reading. One hundred other radiographs were also scored by two experienced orthopedic surgeons (S.M. & H.O.) using the same atlas for interobserver variability. The intra- and inter-variabilities evaluated for KL grade (0-4) were confirmed by kappa analysis to be sufficient for assessment (0.86 and 0.80, respectively).

Statistical analysis

Differences in age, height, weight, BMI, muscle strength, WBI and muscle mass between men and women and between subjects with and without pain were examined using the non-paired student t-test. The prevalence of knee OA and pain was compared between men and women by the χ^2 test. Linear regression analysis was used to determine the association of age, muscle mass at the lower limb, and grip strength with quadriceps muscle strength. Associations of age, BMI, grip strength, quadriceps muscle strength, WBI and muscle mass at the lower limbs and KL grade with knee pain were determined using multiple logistic regression analysis after adjustment for age, sex, and BMI overall, and after adjustment for age and BMI in men and women. To determine the independent association of age, BMI, gender, muscle strength, and KL grade with knee pain, multiple logistic regression analysis was used with age, BMI, gender, muscle strength, and KL grade overall, and with age, BMI, muscle strength, and KL grade in men and women, as explanatory variables. To determine the independent association of WBI with knee pain, multiple logistic regression analysis was used with age, BMI, gender, WBI and KL grade, overall, and with age, BMI, WBI and KL grade in men and women as explanatory variables. In addition, subjects were classified according to muscle strength (<10 kgf, ≥10– < 20 kgf, ≥20– < 30 kgf, ≥30– < 40 kgf, ≥40 kgf), and the association of muscle strength <10 kgf, ≥10– < 20 kgf, ≥20– < 30 kgf, and ≥30– < 40 kgf with pain was determined using multiple logistic regression analysis after adjustment for age and BMI, compared with muscle strength ≥40 kgf). The thresholds of muscle strength or WBI for pain were determined using ROC curve analysis. Data analyses were performed using SAS version 9.0 (SAS Institute Inc., Cary, NC).

Results

Among the 2566 subjects who participated in the third visit of the ROAD study, 2303 (89.9 %) subjects underwent X-ray examinations at the knee. A total of 32 (1.3 %) subjects who underwent total knee arthroplasty before the third visit were excluded from the study. In addition, 12 (0.5 %) subjects who provided incomplete questionnaires regarding pain and 37 subjects (1.5 %) who did not undergo an examination of muscle strength or muscle mass were excluded. Further, 58 subjects (2.3 %) who were younger than 40 years were excluded, leaving a total of 2152 (85.1 %) subjects (690 men and 1462 women). The characteristics of the 2152 participants in the present study are shown in Table 1. Muscle strength and mass were significantly higher in men than women. WBI was not significantly different between men and women. The prevalence of knee OA and knee pain was significantly higher in women than in men. Quadriceps muscle strength was significantly associated with muscle mass at the lower limbs, but the association was weak (right: correlation coefficient =0.28 and 0.21 in men and women, respectively, $p < 00001$; left: correlation coefficient 0.34 and 0.37 in men and women, respectively, $p < 00001$). Quadriceps muscle strength was also significantly associated with grip strength, and the association was moderate (right: correlation coefficient =0.47 and 0.50 in men and women, respectively, $p < 00001$; left: correlation coefficient 0.50 and 0.52 in men and women, respectively, $p < 00001$). Quadriceps muscle strength was significantly associated with age in men and women ($p < 0.0001$) (Additional file 1: Figure S1).

Table 2 shows age, BMI, grip strength, quadriceps muscle strength, WBI and lower limb muscle mass/height^2 in subjects with and without pain. For the right knee, age, BMI, grip strength, quadriceps muscle strength and WBI were significantly different between subjects with and without pain, whereas muscle mass was not. Results were similar for the left knee. After adjustment for age and BMI, the significant association of grip strength with knee pain disappeared in men and women.

We next examined the prevalence of knee pain according to KL grade (Fig. 1). In the overall population, the prevalence of knee pain was 12.5 %, 19.1 % and 46.5 % in the right knee and 10.8 % 18.2 % and 45.3 % in the left knee in subjects with KL = 01, KL = 2 and KL = 3 or 4, respectively. After adjustment for age, gender and BMI, KL = 3 or 4 was significantly associated with knee pain compared with KL = 01 (right knee: odds ratio [OR] 4.16, 95 % confidence interval [CI] 3.10-5.61; left knee: OR 4.90, 95 CI 3.63-6.64). KL = 2 at the left knee was also significantly associated with pain (OR 1.52, 95 % CI 1.17-2.00), while KL = 2 at the right knee was not (OR 1.27, 95 % CI 0.94-1.71). The prevalence of knee pain was 9.9 %, 10.5 % and 48.9 % at the right knee and 9.1 %, 11.5 and 42.7 % at the left knee in men with KL = 01, KL = 2 and KL = 3 or 4, respectively, and 14.2 %, 21.7 % and 45.8 % at the right knee and 11.9 %, 20.8 % and 45.9 % at the left knee in women with KL = 01, KL = 2 and KL = 3or 4, respectively. In men and women, after adjustment for age and

Table 1 Subject characteristics

	Overall	Men	Women	*P* values
N	2152	690	1462	
Age, years	71.6 ± 12.2	72.5 ± 12.3	71.2 ± 12.1	0.0164
Height, cm	154.3 ± 9.2	163.1 ± 7.1	150.1 ± 6.8	<0.0001
Weight, kg	54.3 ± 10.7	61.6 ± 11.0	50.9 ± 8.6	<0.0001
BMI, kg/m^2	22.7 ± 3.4	23.1 ± 3.3	22.5 ± 3.5	0.0009
Right				
Grip strength	28.1 ± 9.6	37.2 ± 9.4	23.7 ± 5.8	<0.0001
Quadriceps muscle strength, kgf	28.1 ± 11.2	31.9 ± 12.7	26.2 ± 10.0	<0.0001
Weight bearing index	0.52 ± 0.20	0.52 ± 0.20	0.52 ± 0.20	0.8724
Lower limb muscle mass, kg	6.3 ± 1.6	7.9 ± 1.5	5.5 ± 0.8	<0.0001
Lower limb muscle mass/height^2, kg/m^2	2.6 ± 0.4	3.0 ± 0.4	2.4 ± 0.3	<0.0001
Knee OA (%)	44.1	31	50.3	<0.0001
Knee pain (%)	20.6	15.1	23.3	<0.0001
Left				
Grip strength	26.2 ± 9.4	35.2 ± 9.1	22.0 ± 5.9	<0.0001
Quadriceps muscle strength, kgf	26.9 ± 11.2	30.6 ± 12.6	25.1 ± 9.9	<0.0001
Weight bearing index	0.50 ± 0.20	0.50 ± 0.20	0.50 ± 0.20	0.9715
Lower limb muscle mass, kg	6.2 ± 1.6	7.8 ± 1.5	5.4 ± 0.8	<0.0001
Lower limb muscle mass/height^2, kg/m^2	2.6 ± 0.4	2.9 ± 0.4	2.4 ± 0.3	<0.0001
Knee OA (%)	45.2	33	51	<0.0001
Knee pain (%)	20	13.9	22.9	<0.0001

Except where indicated otherwise, values are means ± SD
Knee OA was defined as Kellgren-Lawrence grade 2 or worse
Weight bearing index was calculated as quadriceps muscle strength by weight
Differences between men and women were determined by non-paired student t test except for prevalence of knee OA and knee pain
Differences in prevalence of knee OA and knee pain between men and women were determined by chi-square test
BMI Body mass index, *OA* Osteoarthritis

BMI, KL = 3 or 4 was significantly associated with knee pain at the right knee (men: OR 6.82, 95 % CI 3.94-11.9, women: OR 3.52, 95 % CI 2.49-5.03) and the left knee (men: OR 5.64, 95 % CI 3.20-9.99, women: OR 4.83, 95 % CI 3.39-6.94). KL = 2 was not associated with knee pain except for the left knee in women (right knee, men: OR 0.91, 95 % CI 0.45-1.73, women: 1.32, 95 % CI 0.93-1.86; left knee, men: OR 1.08, 95 % CI 0.56-2.00, women: 1.68, 95 % CI 1.16-2.45).

To determine independent associations of age, BMI, gender, muscle strength and knee OA, we next used multiple logistic regression analysis with age, BMI, gender, quadriceps muscle strength and KL grade as explanatory variables in subjects overall, and with age, BMI, muscle strength and KL grade as explanatory variables in men and women (Table 3). Overall, BMI, gender, muscle strength and KL grade 3 or 4 were significantly associated with knee pain, whereas age and KL grade 2 were not. In men and women, BMI, muscle strength and KL grade 3 or 4 were significantly associated with knee pain. We also analyzed independent associations of age, BMI, gender, WBI and knee OA. Results for WBI were almost the same as those for quadriceps muscle strength (overall: OR 0.85, 95 % CI 0.79-0.91, $p = 0.0011$, men: OR 0.79, 95 % CI 0.69-0.90, $p = 0.0003$, women: OR 0.87, 95 % CI 0.80-0.94, $p = 0.0003$).

Next, to determine the prevalence of knee pain according to muscle strength, subjects were classified by muscle strength (<10 kgf, ≥10– < 20 kgf, ≥20– < 30 kgf, ≥30– < 40 kgf, ≥40 kgf). Prevalence of knee pain was 53.9 %, 27.0 %, 14.4 %, 11.6 and 9.8 % at the right knee and 33.3 %, 24.8 %, 12.2 %, 12.6 % and 6.5 % at the left knee in men with muscle strength <10 kgf, ≥10– < 20 kgf, ≥20– < 30 kgf, ≥30– < 40 kgf and ≥40 kgf, respectively, and 41.0 %, 31.0 %, 23.7 %, 16.3 % and 12.5 % at the right knee and 43.2 %, 31.0 %, 20.3 %, 16.1 % and 15.3 % at the left knee in women with muscle strength <10 kgf, ≥10– < 20 kgf, ≥20– < 30 kgf, ≥30– < 40 kgf and ≥40 kgf, respectively (Fig. 2). After adjustment for age, BMI and KL grade, subjects with muscle strength <10 kgf and ≥10– < 20 kgf had a significantly higher prevalence of knee pain compared with those with muscle strength ≥40kgf, except for left knee in women (Table 4). We also examined the

Table 2 Age, BMI, grip strength and lower limb muscle strength and muscle mass in subjects with and without knee pain

	Right knee					Left knee				
	Pain -	Pain +	Adjusted OR	95 % CI	*P* values	Pain -	Pain +	Adjusted OR	95 % CI	*P* values
Overall										
N	1708	444				1721	431			
Age, years	70.8 ± 12.5	74.8 ± 10.4*	1.04	1.03–1.035	<0.0001	70.8 ± 12.5	74.7 ± 10.4*	1.04	1.03–1.05	<0.0001
BMI, kg/m^2	22.5 ± 3.3	23.7 ± 3.7*	1.13	1.10–1.17	<0.0001	22.4 ± 3.3	23.8 ± 3.5*	1.14	1.11–1.18	<0.0001
Grip strength, kgf	28.7 ± 9.7	25.5 ± 8.5*	0.98	0.96–0.9996	0.0448	26.9 ± 9.5	23.8 ± 8.5*	0.99	0.97–1.01	0.3464
Quadriceps muscle strength, kgf (5kgf increase)	29.1 ± 11.2	23.9 ± 10.3*	0.83	0.78–0.88	<0.0001	27.9 ± 11.2	22.8 ± 10.2*	0.84	0.79–0.89	<0.0001
Weight bearing index, kgf/kg (0.1 kgf/kg increase)	0.54 ± 0.20	0.44 ± 0.18*	0.81	0.76–0.86	<0.0001	0.52 ± 0.19	0.42 ± 0.19*	0.83	0.77–0.88	<0.0001
Lower limb muscle mass/height^2, kg/m^2 (0.1kg/m^2 increase)	2.59 ± 0.43	2.58 ± 0.43	0.97	0.92–1.02	0.2421	2.56 ± 0.43	2.56 ± 0.42	0.98	0.94–1.03	0.4326
Men										
N	586	104				594	96			
Age, years	71.9 ± 12.5	76.1 ± 11.0*	1.04	1.02–1.07	<0.0001	72.0 ± 12.3	76.1 ± 11.7*	1.04	1.02–1.07	<0.0001
BMI, kg/m^2	22.9 ± 3.2	23.9 ± 3.8*	1.13	1.06–1.21	0.0002	22.9 ± 3.2	23.9 ± 4.0*	1.13	1.06–1.21	0.0003
Grip strength, kgf	37.5 ± 9.5	35.1 ± 9.0*	0.98	0.95–1.01	0.3070	35.5 ± 9.0	33.3 ± 9.7*	0.99	0.96–1.02	0.5031
Quadriceps muscle strength, kgf (5kgf increase)	32.9 ± 12.5	26.4 ± 12.0*	0.80	0.72–0.89	<0.0001	31.4 ± 12.6	25.4 ± 11.4*	0.82	0.91–1.23	0.0001
Weight bearing index, kgf/kg (0.1 kgf/kg increase)	0.54 ± 0.20	0.42 ± 0.19*	0.75	0.65–0.85	<0.0001	0.52 ± 0.20	0.41 ± 0.18*	0.78	0.68–0.89	0.0002
Lower limb muscle mass/height^2, kg/m^2 (0.1kg/m^2 increase)	2.95 ± 0.44	3.03 ± 0.47	1.01	0.92–1.10	0.8897	2.90 ± 0.44	2.97 ± 0.50	0.98	0.89–1.08	0.7281
Women										
N	1122	340				1127	335			
Age, years	70.2 ± 12.5	74.4 ± 10.2*	1.03	1.02–1.05	<0.0001	70.2 ± 12.6	74.3 ± 10.0*	1.03	1.02–1.05	<0.0001
BMI, kg/m^2	22.2 ± 3.3	23.7 ± 3.6*	1.13	1.09–1.18	<0.0001	22.2 ± 3.4	23.8 ± 3.4*	1.15	1.11–1.19	<0.0001
Grip strength, kgf	24.1 ± 5.8	22.5 ± 5.8*	0.98	0.95–1.004	0.1014	22.3 ± 6.0	21.1 ± 5.7*	0.99	0.97–1.02	0.6256
Quadriceps muscle strength, kgf (5kgf increase)	27.2 ± 9.9	23.2 ± 9.5*	0.84	0.78–0.91	<0.0001	26.1 ± 9.8	22.1 ± 9.7*	0.85	0.79–0.91	<0.0001
Weight bearing index, kgf/kg (0.1 kgf/kg increase)	0.55 ± 0.20	0.45 ± 0.18*	0.83	0.77–0.90	<0.0001	0.52 ± 0.19	0.43 ± 0.19*	0.84	0.78–0.91	<0.0001
Lower limb muscle mass/height^2, kg/m^2 (0.1kg/m^2 increase)	2.41 ± 0.27	2.45 ± 0.32*	0.95	0.89–1.02	0.1421	2.38 ± 0.29	2.44 ± 0.31*	0.99	0.93–1.04	0.6166

*$p < 0.05$ by non-paired student t test

Adjusted ORs were calculated by multiple logistic regression analysis after adjustment for age, sex, and BMI overall and after adjustment for age and BMI in men and women

BMI Body mass index, *mJSW*, Minimum joint space width

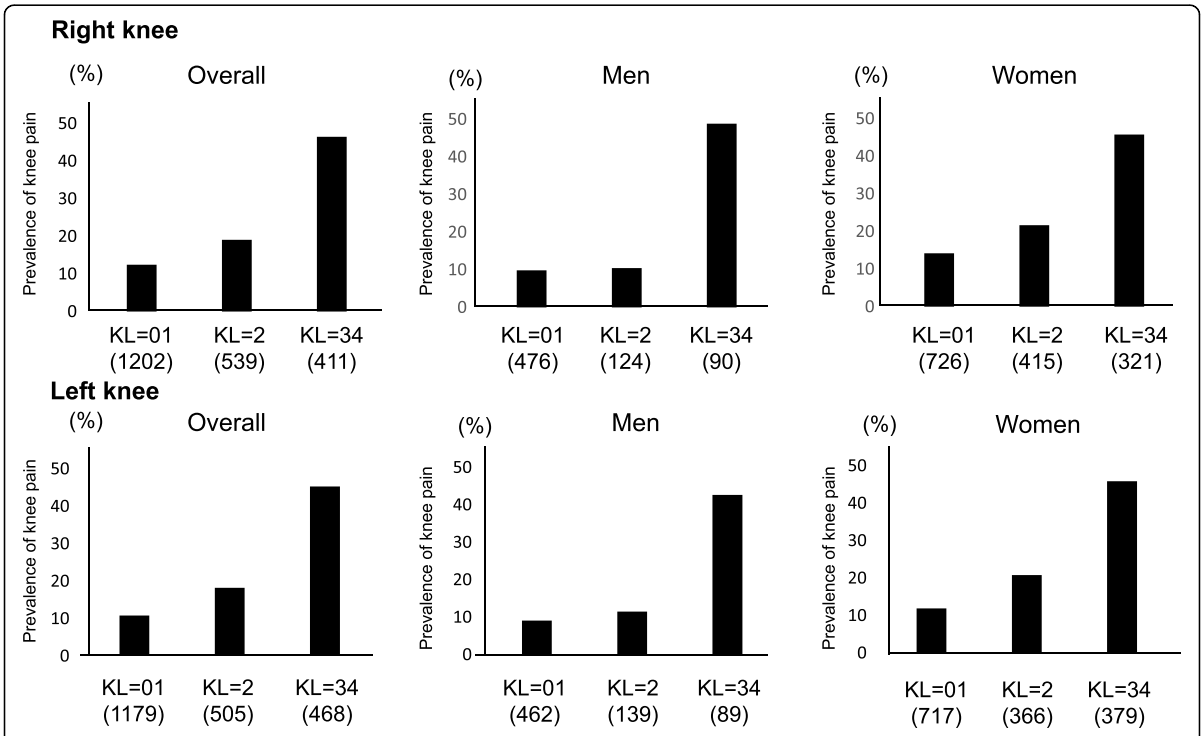


Fig. 1 Prevalence of knee pain by Kellgren-Lawrence (KL) grade. The number of subjects in each KL grade is shown in parentheses

prevalence of knee pain according to WBI and found similar results (Fig. 3).

The threshold values of muscle strength for knee pain were then determined using ROC curve analysis. At the right knee, the threshold values of muscle strength for pain were 27.5 kgf (sensitivity 0.58, specificity 0.64, AUC 0.64) and 27.0 kgf (sensitivity 0.72, specificity 0.48, AUC 0.62) in men and women, respectively. At the left knee, the threshold values of muscle strength for pain were 20 kgf (sensitivity 0.39, specificity 0.82, AUC 0.64) and 23.2 kgf (sensitivity 0.59, specificity 0.41, AUC 0.61) in men and women, respectively. Regarding WBI, the threshold values for pain were 0.43 kgf/kg (sensitivity 0.57, specificity 0.69, AUC 0.67) in men and 0.49 kgf/kg (sensitivity 0.64, specificity 0.59, AUC 0.64) in women at the right knee, and 0.37 kgf/kg (sensitivity 0.46, specificity 0.78, AUC 0.65) in men and 0.40 kgf/kg (sensitivity 0.49, specificity 0.74, AUC 0.64) in women at the left knee.

Discussion

This is the first study to clarify the effect of quadriceps muscle strength as well as muscle mass on knee pain using a large-scale, population-based, cohort study. In the present study, quadriceps muscle strength was significantly associated with knee pain, while grip strength and muscle mass of the lower limb were not. The significant association of quadriceps muscle strength with knee pain remained after adjustment for age, BMI, gender and knee OA.

The present study first clarified that quadriceps muscle strength and WBI were significantly associated with knee pain even after adjustment for radiographic knee OA, which means that the association of muscle strength with knee pain is independent of radiographic changes. In fact, our previous and other previous studies had already shown that the correlation with radiographic severity of the knee OA was not as strong as one would expect [4, 8–10], indicating that there may be some factors other than radiographical changes to explain knee pain. Our results in the present study indicate that not only radiographical changes but also quadriceps muscle strength has an important role in knee pain. The quadriceps muscle is the principal dynamic stabilizer of the knee joint; thus, quadriceps muscle weakness leads to instability of the knee, which may be one of the reasons for knee pain. This also means that knee pain may be prevented by muscle exercise. However, around 10 % of subjects with ≥40 kgf muscle strength had knee pain, indicating that several other factors such as synovitis, knee alignment, meniscal degeneration, thrust and so on may also affect knee pain.

In the present study, although the association of quadriceps muscle strength and grip strength was moderate,

Table 3 Association of age, BMI, gender, muscle strength and severity of knee OA with knee pain

	Right knee			Left knee		
	Adjusted OR	95 % CI	*P* values	Adjusted OR	95 % CI	*P* values
Overall						
Age	1.01	0.996–1.02	0.1698	1.00	0.99–1.02	0.6107
BMI	1.09	1.06–1.13	<0.0001	1.09	1.06–1.13	<0.0001
Women (vs Men)	1.34	1.03–1.76	0.0299	1.35	1.03–1.78	0.0321
Quadriceps muscle strength (5kgf increase)	0.87	0.82–0.92	<0.0001	0.88	0.82–0.93	<0.0001
KL 2	1.3	0.96–1.75	0.0929	1.54	1.12–2.12	0.0083
KL 3 or 4	3.77	2.79–5.10	<0.0001	4.49	3.31–6.10	<0.0001
Men						
Age	1.01	0.99–1.04	0.3309	1.01	0.99–1.04	0.3269
BMI	1.09	1.02–1.17	0.013	1.09	1.02–1.18	0.0152
Quadriceps muscle strength (5 kgf increase)	0.85	0.76–0.94	0.0019	0.86	0.77–0.96	0.0087
KL 2	1.06	0.54–2.16	0.8753	1.18	0.60–2.21	0.624
KL 3 or 4	5.98	3.42–10.54	<0.0001	4.99	2.79–8.93	<0.0001
Women						
Age	1.01	0.99–1.02	0.2416	1.00	9.99–1.02	0.945
BMI	1.09	1.05–1.14	<0.0001	1.09	1.05–1.14	<0.0001
Quadriceps muscle strength (5 kgf increase)	0.88	0.82–0.95	0.0007	0.89	0.82–0.96	0.003
KL 2	1.33	0.95–1.89	0.1007	1.67	1.15–2.44	0.0068
KL 3 or 4	3.24	2.28–4.64	<0.0001	3.37	3.12–6.44	<0.0001

Adjusted OR was calculated by multiple logistic regression analysis with age, BMI, gender, Quadriceps muscle strength and KL grade as explanatory variables
OR Odds ratio, *CI* Confidence interval, *BMI*, Body mass index

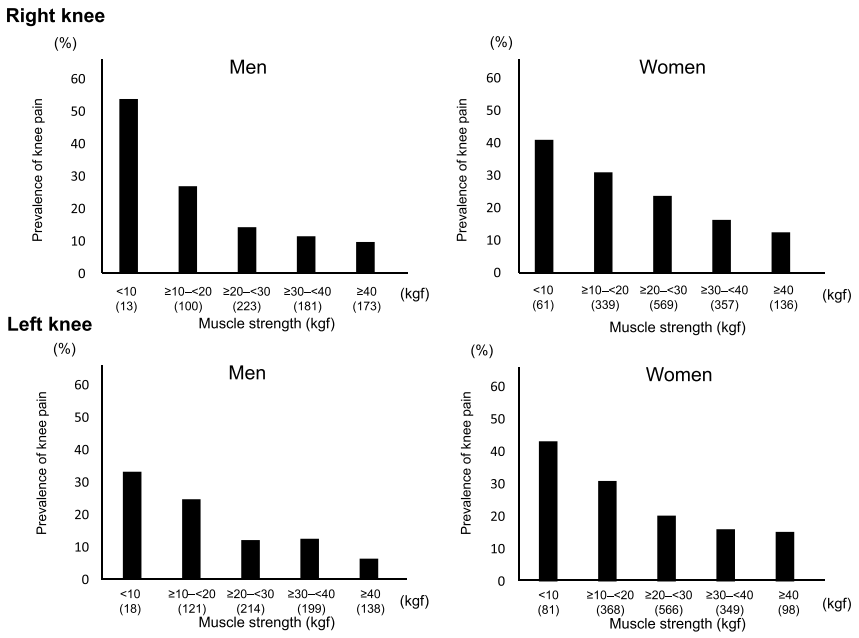


Fig. 2 Prevalence of knee pain by muscle strength at the lower limb in men and women. The number of subjects in each muscle strength is shown in parentheses

Table 4 Odds ratio for knee pain based on quadriceps muscle strength

	Right knee			Left knee		
	OR	95 % CI	*P* value	OR	95 % CI	*P* value
Men						
< 10 kgf	5.87	1.46–23.5	0.0131	4.00	1.002–15.4	0.0497
≥ 10- < 20 kgf	2.26	1.08–4.83	0.0312	3.03	1.30–7.59	0.0096
≥ 20- < 30 kgf	0.95	0.48–1.92	0.8909	1.39	0.61–3.44	0.4405
≥ 30- < 40 kgf	1.14	0.56–2.32	0.7230	1.74	0.79–4.16	0.1771
≥ 40 kgf	1			1		
Women						
< 10 kgf	2.78	1.28–6.13	0.0095	2.00	0.93–4.45	0.0783
≥ 10- < 20 kgf	1.82	1.01–3.42	0.0452	1.49	0.79–2.94	0.2253
≥ 20- < 30 kgf	1.7	0.98–3.10	0.0612	1.03	0.56–2.00	0.9227
≥ 30- < 40 kgf	1.11	0.62–2.08	0.7274	0.91	0.48–1.80	0.7879
≥ 40 kgf	1			1		

OR Odds ratio, *CI* Confidence interval

quadriceps muscle strength rather than grip strength was significantly associated with knee pain. The QTM used in the present study has higher usability compared with other devices. Thus, to use not only grip strength but also quadriceps muscle strength by the QTM may be recommended to estimate sarcopenia.

In the present study, we also examined muscle mass in the lower limbs and found that the association of muscle mass with quadriceps muscle strength was weak. This may be partly explained by impaired neuromuscular activation, which has an independent contribution to muscle strength after adjustment for muscle mass [29]. Furthermore, several studies reported that greater thigh adiposity is known to be associated with lower strength, worse mobility, and worse lipoprotein profiles in the elderly [30–32], which may obscure the association between muscle strength and mass at the lower limbs. This also may be partly explained by the fact that we examined muscle mass not at the quadriceps but at the whole limb on the right and left sides, because the Body Composition Analyzer MC-190 used in the present study cannot measure only quadriceps muscle mass. The present study also showed that muscle strength rather than muscle mass at the lower limbs was associated with knee pain. Previous studies found that lower limb muscle strength, but not muscle mass, was associated with quality of life [19]. Greater thigh adiposity and impaired neuromuscular function may also obscure the association of muscle mass with knee pain.

In the present study, sex differences were found in the association of quadriceps muscle strength with pain. The OR of muscle strength <10 kgf for pain was approximately 5 in men compared with muscle strength ≥40 kgf, while it was approximately 2 in women. These discrepancies between the sexes are partly explained by the

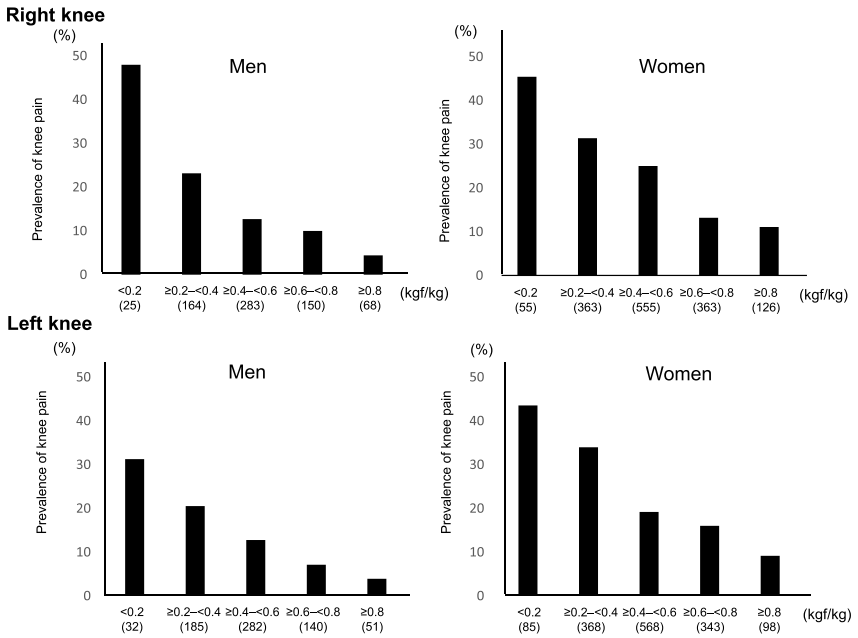


Fig. 3 Prevalence of knee pain by weight bearing index in men and women. The number of subjects in each muscle strength is shown in parentheses. Weight bearing index was calculated as quadiceps muscle strength/weight

fact that women are more susceptible to pain than men [4]. In fact, our previous study showed that the OR for knee pain in women without radiographic knee OA was greater than that in men without radiographic knee OA [4]. In the present study, the prevalence of knee pain was 6–10 % in men with muscle strength ≥40 kgf, and 15–16 % in women with muscle strength ≥40 kgf. This high prevalence of knee pain in women with muscle strength ≥40 kgf, which is the reference point, may partly explain the lower OR for knee pain in women than men. The threshold of muscle strength for knee pain was similar or higher in women than men, which may indicate that factors associated with knee pain include not only gender but also weaker muscle strength.

There are limitations to the present study. This was a large-scale, population-based, cross-sectional study of baseline data. Thus, causal relationships could not be determined. For example, subjects with knee pain may have less physical activity, thereby leading to muscle atrophy and decreases in strength. Or, those individuals with knee pain may be less likely to perform to maximum capacity on the quadriceps strength assessment. The ROAD study is a longitudinal survey, so further progress may help elucidate any causal relationships. In addition, knee pain due to knee OA may not be rest pain, but mainly motion pain, and we did not classify pain into motion pain and rest pain. Therefore, pain in the present study may include not only that from knee OA but also that from other knee pathology.

Conclusion

In conclusion, the present cross-sectional study using a large population from the ROAD study showed that quadriceps muscle strength rather than grip strength or muscle mass at the lower limbs was associated with knee pain. After adjustment for knee OA, muscle strength was independently associated with knee pain. The threshold of muscle strength for knee pain was similar in men and women. Further studies, along with continued longitudinal surveys in the ROAD study, will help improve our understanding of the relationship between muscle strength and pain.

Additional file

Additional file 1: Figure S1. Quadriceps muscle strength by age strata (PPT 188 kb)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SM, TA and NY conceptualized and designed the study, drafted the initial manuscript, and carried out the initial analyses. MY, FT, ST, HK and KN reviewed and revised the manuscript. MT, RK, YA and HO designed the data collection instruments, and coordinated and supervised data collection, critically reviewed the manuscript. All authors read and approved the final manuscript.

Authors' information

Not Applicable.

Acknowledgements

This study was supported by a Grant-in-Aid for H17-Men-eki-009 (Director, Kozo Nakamura), H20-Choujyu-009 (Director, Noriko Yoshimura), H23-Choujyu-002 (Director, Toru Akune), H-25-Choujyu-007 (Director, Noriko Yoshimura), and H25-Nanchitou (Men)-005 (Director, Sakae Tanaka) of the Ministry of Health, Labour and Welfare; and Scientific Research B23390172, B20390182, and Challenging Exploratory Research 24659317 to Noriko Yoshimura; H25-Choujyu-004 (Director, Atsushi Harada), H22-Choujyu-Wakate-007 (Director, Shigeyuki Muraki), B23390356, C20591774, and Challenging Exploratory Research 23659580 to Shigeyuki Muraki; Challenging Exploratory Research 24659666 and 21659349 and Young Scientists A18689031 to Hiroyuki Oka; B23390357 and C20591737 to Toru Akune; and Collaborating Research with NSF 08033011-00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. This study also was supported by grants from the Japan Osteoporosis Society (Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, and Toru Akune), grants from Mitsui Sumitomo Insurance Welfare Foundation (Shigeyuki Muraki), and research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 & 2010-2; Director, Hiroshi Kawaguchi).

The authors thank Dr. Anamizu and members of the Department of Orthopedics; Mr. Kutsuma and other members of the Department of Radiology at Tokyo Metropolitan Geriatric Medical Center. The authors also thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka of the Gobo Public Health Centre; Dr. Naoki Hirabayashi of the Kawakami Clinic in Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Centre; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji, and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. We also thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, Mrs. Saeko Sahara, and Mr. Noriyuki Oe for their assistance in data reduction and administration.

Author details

[1]Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. [2]National Rehabilitation Center for Persons with Disabilities, Saitama, Japan. [3]Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan. [4]Department of Orthopaedic Surgery, Tokyo Geriatric Medical Center, Tokyo, Japan. [5]Department of Orthopaedic Surgery, Faculty of Medicine, the University of Tokyo, Tokyo, Japan. [6]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Tokyo, Japan. [7]Department of Orthopaedic Surgery, Japan Community Health care Organization Tokyo Shinjuku Medical Center, Tokyo, Japan. [8]Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, the University of Tokyo, Tokyo, Japan.

Received: 23 June 2015 Accepted: 24 September 2015
Published online: 16 October 2015

References

1. Sharma L, Kapoor D. Epidemiology of osteoarthritis. In: Moskowitz RW, Altman RD, Hochberg MC, Buckwalter JA, Goldberg VM, editors. Osteoarthritis: diagnosis and medical/surgical management. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 3–26.
2. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health. 1994;84:351–8.
3. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum. 1998;41:1343–55.

4. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: The ROAD study. Osteoarthritis Cartilage. 2009;17:1137–43.
5. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: The Research on Osteoarthritis/osteoporosis Against Disability (ROAD). J Bone Miner Metab. 2009;27:620–8.
6. Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2010. Available at http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/index.html
7. Linaker CH, Walker-Bone K, Palmer K, Cooper C. Frequency and impact of regional musculoskeletal disorders. Baillieres Clin Rheumatol. 1999;13:197–215.
8. Summers MN, Haley WE, Reveille JD, Alarcon GS. Radiographic assessment and psychologic variables as predictors of pain and functional impairment in osteoarthritis of the knee or hip. Arthritis Rheum. 1988;31:204–9.
9. Cicuttini FM, Baker J, Hart DJ, Spector TD. Association of pain with radiological changes in different compartments and views of the knee joint. Osteoarthritis Cartilage. 1996;4:143–7.
10. Wluka AE, Wolfe R, Stuckey S, Cicuttini FM. How does tibial cartilage volume relate to symptoms in subjects with knee osteoarthritis? Ann Rheum Dis. 2004;63:264–8.
11. Roubenoff R. Sarcopenia: a major modifiable cause of frailty in the elderly. J Nutr Health Aging. 2000;4:4140–2.
12. Sayer AA, Syddall HE, Dennison EM, Martin HJ, Phillips DI, Cooper C, et al. Grip strength and the metabolic syndrome: findings from the Hertfordshire cohort study. QJM. 2007;100:707–13.
13. Bohannon RW. Hand-grip dynamometry predicts future outcomes in aging adults. J Geriatr Phys Ther. 2008;31:313–10.
14. Sirola J, Rikkonen T, Tuppurainen M, Jurvelin JS, Kroger H. Association of grip strength change with menopausal bone loss and related fractures: a population based follow-up study. Calcif Tissue Int. 2006;78:218–26.
15. Kerr A, Syddall HE, Cooper C, Turner GF, Briggs RS, Sayer AA. Does admission grip strength predict length of stay in hospitalised older patients? Age Ageing. 2006;35:82–4.
16. Muraki S, Akune T, Oka H, En-yo Y, Yoshida M, Saika A, et al. Association of radiographic and symptomatic knee osteoarthritis with health-related quality of life in a population-based cohort study in Japan. ROAD Stud Osteoarthritis Cartilage. 2010;18:1227–34.
17. Wind AE, Takken T, Helders PJ, Engelbert RH. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? Eur J Pediatr. 2010;169:281–7.
18. Omori G, Koga Y, Tanaka M, Nawata A, Watanabe H, Narumi K, et al. Quadriceps muscle strength and its relationship to radiographic knee osteoarthritis in Japanese elderly. J Orthop Sci. 2013;18:536–42.
19. Andrews JS, Trupin L, Schmajuk G, Barton J, Margaretten M, Yazdany J, et al. Muscle strength, muscle mass, and physical disability in women with systemic lupus erythematosus. Arthritis Care Res. 2015;67:120–7.
20. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol. 2010;39:988–95.
21. Janssen I, Heymsfield SB, Baumgartner RN, Ross R. Estimation of skeletal muscle mass by bioelectrical impedance analysis. J Appl Physiol. 2000;89:465–71.
22. Kyle UG, Genton L, Slosman DO, Pichard C. Fat-free and fat mass percentiles in 5225 healthy subjects aged 15 to 98 years. Nutrition. 2001;17:534–41.
23. Kyle UG, Genton L, Karsegard L, Slosman DO, Pichard C. Single prediction equation for bioelectrical impedance analysis in adults aged 20–94 years. Nutrition. 2001;17:248–53.
24. Roubenoff R, Baumgartner RN, Harris TB, Dallal GE, Hannan MT, Economos CD, et al. Application of bioelectrical impedance analysis to elderly populations. J Gerontol A Biol Sci Med Sci. 1997;52:M129–36.
25. Tanimoto Y, Watanabe M, Sun W, Sugiura Y, Tsuda Y, Kimura M, et al. Association between sarcopenia and higher-level functional capacity in daily living in community-dwelling elderly subjects in Japan. Arch Gerontol Geriatr. 2012;55:e9–13.
26. Tanimoto Y, Watanabe M, Sun W, Tanimoto K, Shishikura K, Sugiura Y, et al. Association of sarcopenia with functional decline in community-dwelling elderly subjects in Japan. Geriatr Gerontol Int. 2013;13:958–63.
27. Nemoto M, Yabushita N, Kim M, Matsuo T, Seino S, Songee J, et al. Validity of the bioelectrical impedance method for assessing body composition in non-frail and pre-frail older adults. Int J Body Comps Res. 2012;10:55.
28. Kellgren JH, Lawrence JS, editors. The epidemiology of chronic rheumatism: atlas of standard radiographs of arthritis. Oxford: Blackwell Scientific; 1963.
29. Metter EJ, Conwit R, Metter B, Pacheco T, Tobin J. The relationship of peripheral motor nerve conduction velocity to age-associated loss of grip strength. Aging (Milano). 1998;10:471–8.
30. Goodpaster BH, Carlson CL, Visser M, Kelley DE, Scherzinger A, Harris TB, et al. Attenuation of skeletal muscle and strength in the elderly: the health ABC study. J Appl Physiol. 2001;90:2157–65.
31. Visser M, Kritchevsky SB, Goodpaster BH, Newman AB, Nevitt M, Stamm E, et al. Leg muscle mass and composition in relation to lower extremity performance in men and women aged 70 to 79: the health, aging and body composition study. J Am Geriatr Soc. 2002;50:897–904.
32. Durheim MT, Slentz CA, Bateman LA, Mabe SK, Kraus WE. Relationships between exercise-induced reductions in thigh intermuscular adipose tissue, changes in lipoprotein particle size, and visceral adiposity. Am J Physiol Endocrinol Metab. 2008;295:e407–12.

Spine

SPINE Volume 40, Number 23, pp 1807–1815

CERVICAL SPINE

Efficacy of Posterior Segmental Decompression Surgery for Pincer Mechanism in Cervical Spondylotic Myelopathy: A Retrospective Case-controlled Study Using Propensity Score Matching

Akihito Minamide, MD, PhD,* Munehito Yoshida, MD, PhD,* Hiroshi Yamada, MD, PhD,* Hiroshi Hashizume, MD, PhD,* Yukihiro Nakagawa, MD, PhD,* Hideto Nishi, MD, PhD,* Hiroshi Iwasaki, MD, PhD,* Shunji Tsutsui, MD, PhD,* Otohiro Okada, MD, PhD,* Sae Okada, MD,* and Hiroyuki Oka, MD, PhD†

Study Design. Retrospective case-controlled study using propensity score matching.

Objective. We aimed to evaluate the efficacy of cervical microendoscopic laminoplasty (CMEL) of the articular segment in patients with cervical spondylotic myelopathy (CSM) by comparing the clinical results of CMEL with conventional expansive laminoplasty (ELAP) for CSM.

Summary of Background Data. A total of 259 patients undergoing CMEL or ELAP surgery for CSM at authors' institute were reviewed.

Methods. The patients were matched according to calculated propensity scores in a logistic regression model adjusted for age, sex, and preoperative severity of disorders and divided into the CMEL and ELAP groups. All patients were followed postoperatively for more than 2 years. The preoperative and 2-year follow-up evaluations included neurological assessment (Japanese Orthopaedic Association [JOA] score), recovery rates, the JOA Cervical Myelopathy Evaluation Questionnaire (JOACMEQ), axial pain (visual analog scale), and the Short Form 36 questionnaire (SF-36).

Results. There were 71 patients in each group (47 males and 24 females each). The mean ages of the CMEL and ELAP groups were 63.8 and 62.8 years, respectively. There was no significant difference in the preoperative JOA score between groups. The mean numbers of surgically affected levels in the ELAP and CMEL groups were 3.2 and 1.8 discs, respectively ($P \leq 0.05$). The groups exhibited similar recoveries of JOA, JOACMEQ, and SF-36 scores postoperatively. Sagittal alignment was maintained in both groups. However, postoperative neck axial complaints were significantly reduced in the CMEL group.

Conclusion. CMEL may be a useful and effective surgical procedure for CSM, providing similar results as ELAP. CMEL for CSM is indicated for posterior decompression of the articular segment along with a pincer mechanism. This minimally invasive technique may have potential advantages compared with conventional ELAP, and may provide an alternative surgical option.

Key words: articular segment, cervical laminoplasty, cervical spine, cervical spondylotic myelopathy, clinical outcome, endoscopic spinal surgery, minimum invasive surgery, pincer mechanism, propensity score matching, retrospective case-controlled study.

Level of Evidence: 4

Spine 2015;40:1807–1815

From the *Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan; and †Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo Hospital, Tokyo, Japan.

Acknowledgment date: January 21, 2015. First revision date: May 25, 2015. Acceptance date: June 29, 2015.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work.

Relevant financial activities outside the submitted work: grants.

Address correspondence and reprint requests to Akihito Minamide, MD, PhD, Department of Orthopaedic Surgery, Wakayama Medical University, 811–1 Kimiidera, Wakayama City, Wakayama 641–8510, Japan; E-mail: minamide@wakayama-med.ac.jp

DOI: 10.1097/BRS.0000000000001055

Cervical expansive laminoplasty (ELAP) for cervical myelopathy is a posterior decompression surgery that is reported to have favorable results.[1–9] However, some problems after conventional expansive laminectomy or laminoplasty have also been reported due to damage to the cervical posterior soft tissues including muscles and ligaments,[5,10–15] including persistent axial pain, restriction of neck motion, and loss of lordotic curvature. For the treatment of multisegment cervical myelopathy, the posterior arches of the cervical vertebrae and

attached deep extensor muscles are extensively compromised by consecutive laminectomy or laminoplasty throughout the affected levels.

Endoscopic surgery poses several challenges for endoscopic surgeons, particularly in terms of mastering hand-eye coordination. After training in live animal and cadaver surgery was introduced, technical progress has reduced some postoperative morbidities such as dural tear, neural deficit, *etc.*[16] Microendoscopic decompressive techniques were recently developed and applied to various spine pathologies including lumbar spinal stenosis and cervical radiculopathy and myelopathy.[1,4,7,9,13,16–21] Over 3000 patients with lumbar spinal canal stenosis have undergone microendoscopic decompression surgery at the authors' institution. The authors have performed cervical microendoscopic laminoplasty (CMEL) as a minimally invasive strategy for cervical posterior decompression surgery of the articular segment with a pincer mechanism.[21] This procedure is also a spinal cord decompression procedure that maintains the posterior structures.

Therefore, in the present study, we aimed to evaluate the efficacy of CMEL for the articular segment with pincer mechanism in patients with cervical spondylotic myelopathy (CSM) by comparing the clinical results of CMEL with conventional ELAP for patients with CSM.

MATERIALS AND METHODS

Study Design

This retrospective case-control study of the clinical outcomes of CMEL and ELAP for the treatment of CSM used the propensity score matching method.[22] A one-to-one matching analysis was performed between patients who underwent ELAP and CMEL on the basis of the estimated propensity scores of each patient.

CMEL Technique

First, the patient is secured in a Mayfield headholder and is turned to the prone position. The neck is fixed in a neutral position. The operator generally stands on the side of the approach, usually the left, with the video monitors opposite him/her. Under fluoroscopic guidance held lateral to the patient, the targeting level is marked on the side of the approach. A skin incision approximately 16-mm long is made at the spinal level to be decompressed. After splitting into the paravertebral muscles, a set of serial dilators from the METRx endoscopic system (Medtronic Sofamor Danek, Memphis, TN) is passed to gently dilate the cervical musculature. The tubular retractor lies on the lamina and facet joints, and is tilted to parallel to the intervertebral disc (Figure 1A). The endoscope is then attached to the tubular retractor. With the bony edges well visualized, the interlaminar space and medial edge of the fact complex are confirmed.

To begin the partial laminectomy, a high-speed drill with a long curved endoscopic bit (*e.g.*, Midas-Rex Legend; Medtronic, Fort Worth, TX) is used to thin the lamina to near the attachment of the ligamentum flavum. The endoscope is then swung medially to obtain a contralateral view (Figure 1B). After the basement of spinal process is drilled, the laminotomy is performed using a long curved high-speed drill with an endoscopic bit. Thus, the laminotomy is performed with the drilling and tunneling of the internal plate of the lamina through the spinal canal side. The scope is rotated to a lateral position to make use of its 25° viewing angle. As a result of these maneuvers, an excellent viewing angle of approximately 60° to 75° is usually obtained with good contralateral visualization.

The superior attachment of the ligamentum flavum is exposed, and the procedure is then continued to the superior portion of the inferior lamina. The inferior attachment of the ligamentum flavum is subsequently exposed. It is important to continue the contralateral procedure without removing the ligamentum flavum in order to protect the spinal cord.

When the spinal cord is completely decompressed, the floated ligamentum flavum is observed (Figure 1C). The ligamentum flavum is subsequently completely removed, revealing the dural pulsation (Figure 1D). When decompression surgery is required for an adjacent level, the tubular retractor is inclined cranially or caudally. Then, the same procedure is repeated at the adjacent level (Figures 1E, 2A–C). For cases requiring operation of more than 3 levels, another skin incision is added. For example, in a case of C3–C7 CMEL, the skin incisions are made at the C4 and C6 levels. A drain is placed at each level to prevent postoperative epidural hematoma. Finally, the tubular retractor and endoscope are removed, and the fascia and skin are closed using standard techniques.

Patient Population

This study was approved by the institutional review board of the authors' institution. Between 2004 and December 2011, consecutive patients diagnosed with CSM were enrolled. All patients presented with symptoms of cervical myelopathy, such as clumsiness, numbness of the upper and lower extremities, gait disturbance, and urinary disturbance. Cervical spinal cord compression was confirmed by magnetic resonance imaging (MRI), myelography and postmyelography computed tomography. The exclusion criteria were cervical myelopathy with tumor, trauma, ossification of the posterior longitudinal ligament, rheumatoid arthritis, pyogenic spondylitis, destructive spondyloarthropathies, and other combined spinal lesion.

A total of 259 patients underwent posterior decompression surgery using either CMEL[21] (Figure 1) or conventional cervical ELAP (French-door[23]or open-door type[24]) at the authors' institution. Postoperatively, the use of a neck brace was left to the patients' discretion. All patients were followed postoperatively for more than 2 years.

Assessment

Neurological evaluation and recovery rates were assessed postoperatively by using Hirabayashi's method with the

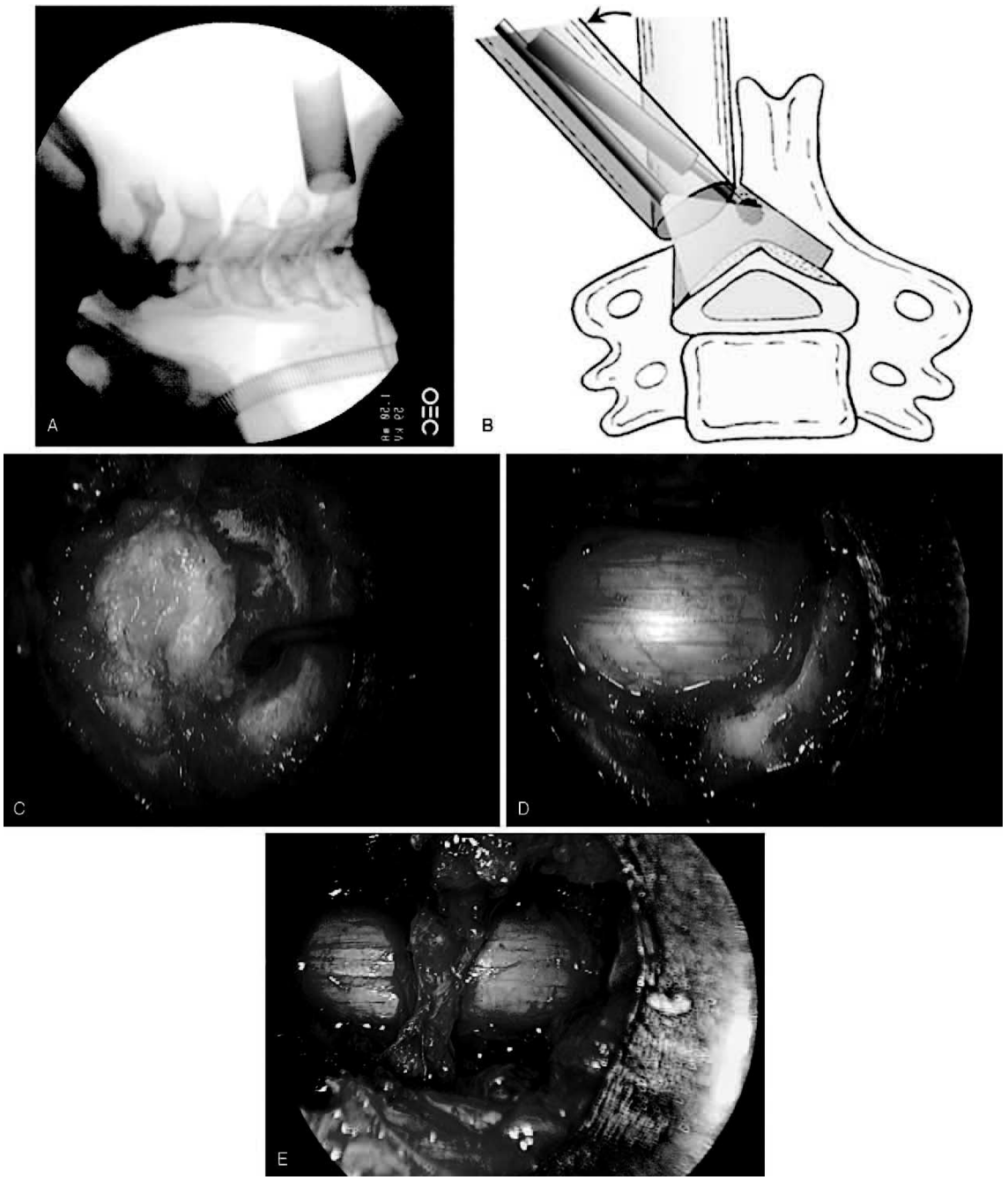


Figure 1. Cervical microendoscopic laminoplasty (CMEL) procedure. **(A)** The tubular retractor lies on the lamina and facet joints, and is tilted parallel to the intervertebral disc. The decompression surgery is performed using a high-speed air drill. **(B)** The hemilaminectomy is performed on the approaching side, followed by the laminotomy on the contralateral side done. Finally, the expansive laminotomy is completed to enlarge the spinal canal. **(C)** When the spinal cord is completely decompressed from all attachments of the ligamentum flavum, the floated ligamentum flavum is observed. **(D)** When the ligamentum flavum is completely removed, the ural pulsation is observed. The decompression is performed until the lateral edge of dural tube. **(E)** For an adjacent level, the tubular retractor is inclined cranially or caudally, and the above procedure is repeated.

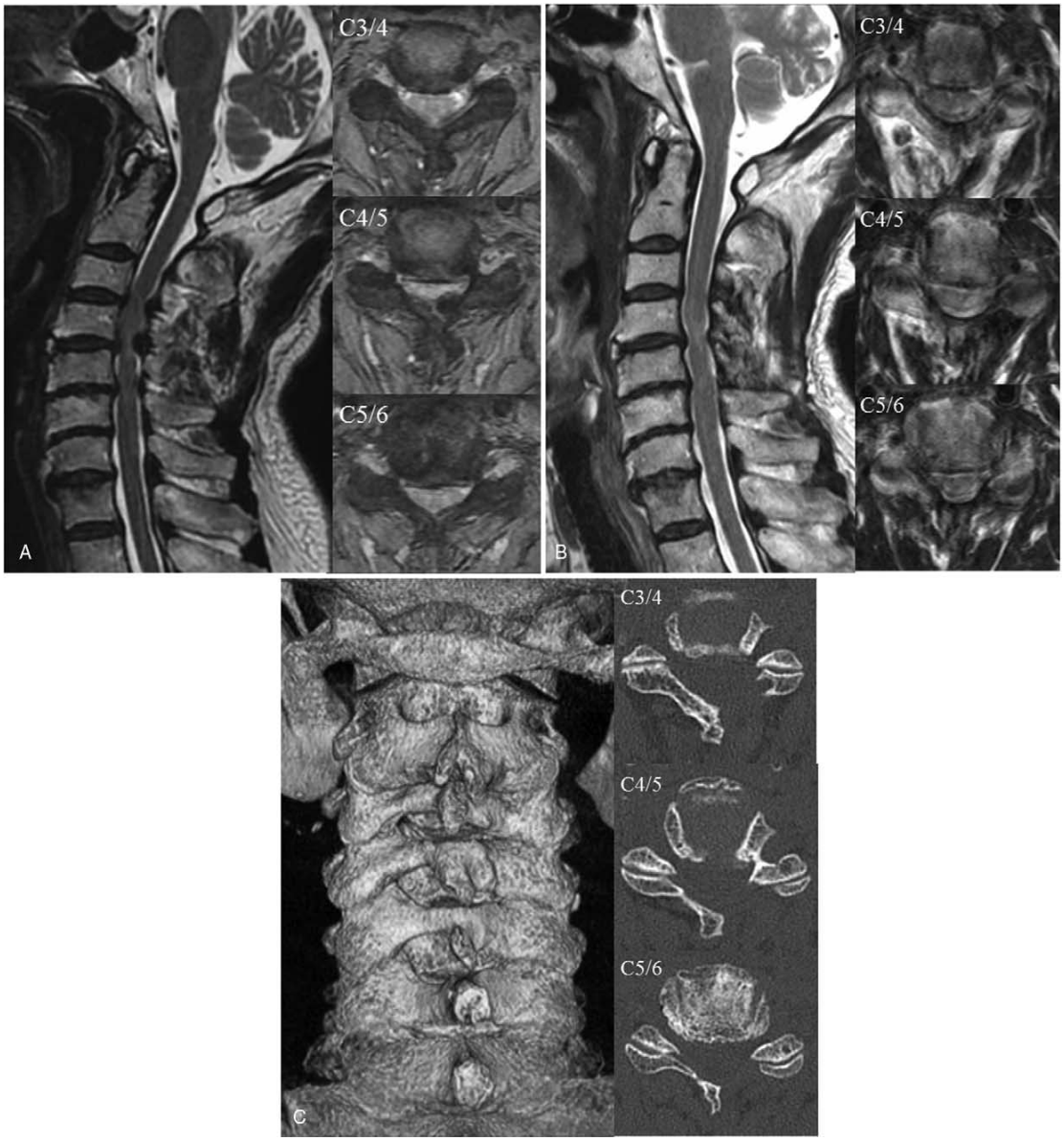


Figure 2. Illustrative case. A 74-yr-old female with cervical spondylotic myelopathy presented with clumsiness, numbness in the hands, crutch walking owing to spasticity, and urinary disturbance. **(A)** MRI showing spinal cord compression with spondylosis at the C3–C4, C4–C5, and C5–C6 levels, particularly severe spinal cord compression due to calcification of ligamentum flavum at C4–C5. The patient underwent CMEL from the C3–C6 levels. Her JOA score improved from 12 points preoperatively to 16 points 2 yr postoperatively. **(B, C)** MRI and CT yr postoperatively showing the success of spinal cord decompression. Axial images also show the successfully decompressed spinal cord at each level. By the unilateral approach, the partial hemilaminectomy and the laminotomy on the contralateral side are done at each level.

criteria proposed by the Japanese Orthopaedic Association scoring system (JOA score, maximum score: 17 points),[25] the JOA Cervical Myelopathy Evaluation Questionnaire (JOACMEQ),[26] the visual analog scale (VAS) for the assessment of axial pain, and the Short Form 36 (SF-36) survey. On the basis of JOACMEQ severity score points pre- and postoperatively, we investigated the effectiveness of treatment for cervical spine function, upper extremity function, lower extremity function, bladder function, and QOL. As for patient evaluations, we judged a treatment as effective when either of the following conditions was met: (1) postoperative score was increased by 20 or more compared with preoperative score or (2) preoperative score less than 90 reached 90 or more postoperatively. Furthermore, the effective rate of a group was calculated by dividing the number of subjects for whom treatment was judged as effective by the number obtained by subtracting the number of subjects with scores of 90 or more before surgery that remained 90 or

TABLE 1. Characteristics of Each Group on Matching by the Calculated Propensity Score

	ELAP*	CMEL†	P‡
Patients	71	71	
Sex	Male 47, female 24	Male 47, female 24	0.645
Age	63.8 ± 11.7	62.8 ± 13.7	0.96
Preoperative JOA§	10.1 ± 2.4	10.2 ± 2.6	<0.0001
Surgical levels¶	3.2 ± 0.6	1.8 ± 0.8	

*Conventional cervical expansive laminoplasty.
†Cervical microendoscopic laminoplasty.
‡$P < 0.05$ is statistically significant.
§Japanese Orthopaedic Association scoring system (JOA); full score 17 points.
¶The number of surgically affected levels.

more postoperatively from the total number of subjects constituting a group. Lateral radiographs were taken in the neutral position preoperatively and 2-years postoperatively. The lordotic angle was determined to be between the C2 and C7 angles at the neutral position using Cobb method.

Statistical Analysis

A one-to-one matching analysis was performed between the ELAP and CMEL groups on the basis of the estimated propensity scores of each patient. The propensity score approach addresses the selection bias inherent to retrospective observational studies, in which outcomes can reflect a lack of comparability in treatment groups rather than the actual effects of treatment.[22]

The matching procedure matched cases in the 2 groups according to the similarity of their propensity scores. A nearest-neighbor matching procedure was used, with the restriction that the propensities matched had to be within 0.05 units of each other. The application of propensity score matching involves the estimation of the propensity score followed by the matching of patients according to their estimated propensity score and comparison of outcomes in matched patients. To estimate the propensity score, we fitted a logistic regression model for the receipt of ELAP as a function of patient demographic factors including age, sex, and preoperative JOA score.

Demographic items were compared between surgical methods preoperatively and at the 2-year follow-up. Student *t* test was used to compare preoperative and postoperative recovery rates as well as JOA, JOACMEQ, VAS, and SF-36 scores between the CMEL and ELAP groups. All statistical analyses were performed using SPSS version 20.0 (SPSS Inc., Chicago, IL). The level of significance was set at $P < 0.05$.

RESULTS

There were 71 patients each in the CMEL and ELAP groups; each group comprised 47 males and 24 females. The mean ages at surgery in the ELAP and CMEL groups were 63.8 ± 11.7 and 62.8 ± 13.7 years, respectively ($P > 0.05$). The mean number of surgical levels involved in the ELAP and CMEL groups were 3.2 and 1.8, respectively ($P < 0.05$). The mean preoperative JOA scores in the ELAP and CMEL groups were 10.1 ± 2.4 and 10.2 ± 2.6 points, respectively ($P > 0.05$) (Table 1). The mean hospital stay was significantly shorter in the CMEL group than the ELAP group ($P < 0.01$) (Table 1). The mean recovery rates in the ELAP and CMEL groups were 56.3 ± 22.2% and 58.2 ± 23.7%, respectively; there was no significant difference in the JOA recovery rate between groups ($P = 0.35$) (Table 2).

Regarding perioperative complications, 1 patient each in the ELAP and CMEL groups had C5 nerve root palsy postoperatively and 2 patients developed postoperative epidural

TABLE 2. Clinical Outcomes on Each Surgical Procedure at the 2-yr Follow-up

	ELAP*	CMEL†	P‡
JOA§	13.9 ± 2.1	14.1 ± 1.9	0.485
JOA recovery rate	56.3 ± 22.2	62.8 ± 13.7	0.349
VAS¶	42.8 ± 32.4	24.5 ± 25.6	0.001
Satisfaction‖	7.8 ± 2.1	8.5 ± 1.8	0.036

*Conventional cervical expansive laminoplasty.
†Cervical microendoscopic laminoplasty.
‡$P < 0.05$ is statistically significant.
§Japanese Orthopaedic Association scoring system (JOA); full score 17 points.
¶Visual analogue scale for axial symptoms; full scale 100 mm.
‖Satisfaction for each surgery; full scale 10 points.

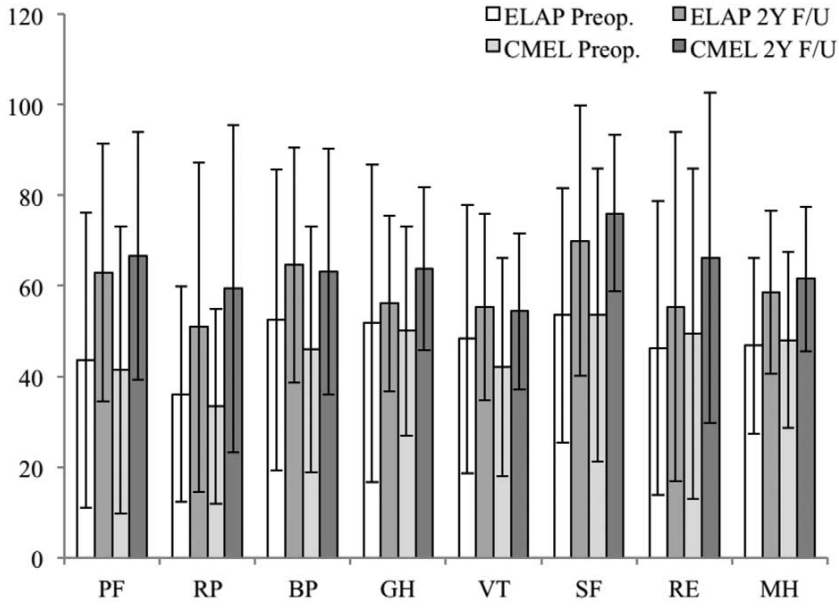


Figure 3. SF-36 scores. There were no significant differences in any subscale (physical functioning [PF], role physical [RP], bodily pain [BP], social functioning 1 [SF], general health perceptions [GH], vitality [VT], role emotional [RE], and mental health [MH]) between the CMEL and ELAP groups. (Preop: preoperatively, 2Y F/U: 2-yr follow-up).

hematoma in the CMEL group. All patients improved as a result of conservative treatment and had returned to daily life at the final follow-up. The VAS score for axial symptoms at the 2-year follow-up was significantly lower in the CMEL group than the ELAP group ($P = 0.001$) (Table 2). Regarding the JOACMEQ and SF-36 scores, there were no significant differences in any subscale between groups ($P > 0.05$) (Figure 3) (Table 3). The score on the patient satisfaction scale (scored on 10 points) was significantly higher in the CMEL group than the ELAP group ($P = 0.036$) (Table 2).

In the ELAP group, the mean lordotic angle was 8.9° preoperatively and 9.1° at the 2-year follow-up (Table 4); that in the CMEL group was 12.3° preoperatively and 13.6° at the 2-year follow-up. There were no significant differences in the lordotic angle pre- and postoperatively between groups ($P > 0.05$). 2 patients each in the ELAP and CMEL groups had local kyphosis that was greater than 13°[27]; 1 patient each in both groups exhibited improved local kyphosis after surgery, whereas the other patients still had kyphosis. The recovery rate was 36.4% in the ELAP group and 39.1% in the CMEL group. With regard to changes in the alignment of the lateral neutral position, 2 patients each in the ELAP and CMEL groups presented with preoperative lordosis that changed to kyphosis at the 2-year follow-up. Similarly, 8 (11.3%) and 4 (4.2%) of 71 patients in the ELAP and CMEL groups, respectively, exhibited an increase in the lordotic angle greater than10°. This suggests CMEL maintained lordosis better than ELAP.

DISCUSSION

Axial symptoms after cervical ELAP have recently been reported[10–14]; the frequency of such symptoms is reported to be 3 times that following cervical anterior interbody fusion.[10] Postoperative complications including persistent axial pain remain unresolved. Therefore, various modifications of surgical techniques as well as early neck mobilization have been developed for conventional cervical laminoplasty to prevent such morbidities.[23,28]

TABLE 3. Effective Rate of JOACMEQ* in Each Group

JOACMEQ*	ELAP† (%)	CMEL‡ (%)	P§
Cervical spine function	47.5	56.4	0.405
Upper extremity function	46.3	48.1	0.834
Lower extremity function	45.5	44.9	0.835
Bladder function	46.3	45.5	0.836
Quality of life	45.7	43.5	0.845

*Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire.
†Conventional cervical expansive laminoplasty.
‡Cervical microendoscopic laminoplasty.
§$P < 0.05$ is statistically significant.

TABLE 4. Between C2 and C7 Angle at the Neutral Position Using Cobb Method on Lateral Radiograph

	ELAP*	CMEL†	P‡
Preoperation	8.9 ± 10.7	12.3 ± 10.7	0.069
2-yr follow-up	9.1 ± 9.8	13.6 ± 10.6	0.011

*Conventional cervical expansive laminoplasty.
†Cervical microendoscopic laminoplasty.
‡$P < 0.05$ is statistically significant.

Intraoperative damage to the cervical posterior soft tissues including muscles and ligaments is reported to be one of the causes for these complications. Accordingly, the authors have been applying CMEL as a minimally invasive strategy for cervical posterior decompression surgery for cervical myelopathy.[21] This microendoscopic procedure involves a small skin incision that splits into the paravertebral muscles. CMEL also involves the combination of endoscopic hemilaminectomy and laminotomy. Compared with the conventional technique, endoscopic surgery substantially differs with respect to the influence on the soft tissues (Figure 4). In this study, the VAS scores for the assessment of axial pain indicate that CMEL resulted in less damage to the cervical soft tissues than ELAP. This difference is due to the smaller skin incision and less damage to the soft tissues, including muscle, incurred by the minimally invasive CMEL technique.

The main indication for CMEL is myelopathy with posterior factors such as calcification or ossification of the ligamentum flavum and degenerative spondylosis with a pincer mechanism.[29–31] The procedure extends the adaptation for CSM to multiple levels. In patients with CSM, this was possible with the decompression of the articular segment (Figure 5 A,B). This procedure can achieve the posterior decompression of the spinal cord associated with CSM by the decompression of the articular segment. The surgical procedures of this concept for posterior decompression include the segmental partial laminectomy and skip laminectomy.[27,32] However, the indication for articular segment decompression is limited to the cervical myelopathies excluding developmental spinal canal stenosis and spinal canal stenosis with severe anterior factors such as severe ossification of the posterior longitudinal ligament. Compared with the conventional laminoplasty technique, CMEL limits the enlarged area of the spinal canal because of the laminotomy technique. The posterior shift of the spinal cord with ELAP was thought to be necessary for patients with ossification of the posterior longitudinal ligament.[2,3] CMEL is indicated for the posterior decompression of the articular segment when the posterior indentation of the spinal cord is recognized with or without anterior compression of a bony spur and a degenerative bulging disc.

In this study, the surgical disc levels differed between the 2 groups, although there was no difference in the surgical selection criteria for the patients with CSM. Endoscopic surgery poses several challenges for endoscopic surgeons, particularly in terms of mastering hand-eye coordination. The CMEL surgery also has a learning curve. Therefore, the decision for the use of either ELAP or CMEL as the surgical method for patients with CSM was made as per the operator's discretion. Moreover, based on the general concept of the posterior shift of the spinal cord with ELAP, ELAP at C3–C5, C3–C6, or C3–C7 was performed even in patients with CSM in whom the main lesion responsible was located at 1 or 2 disc levels; for example, a patient with CSM

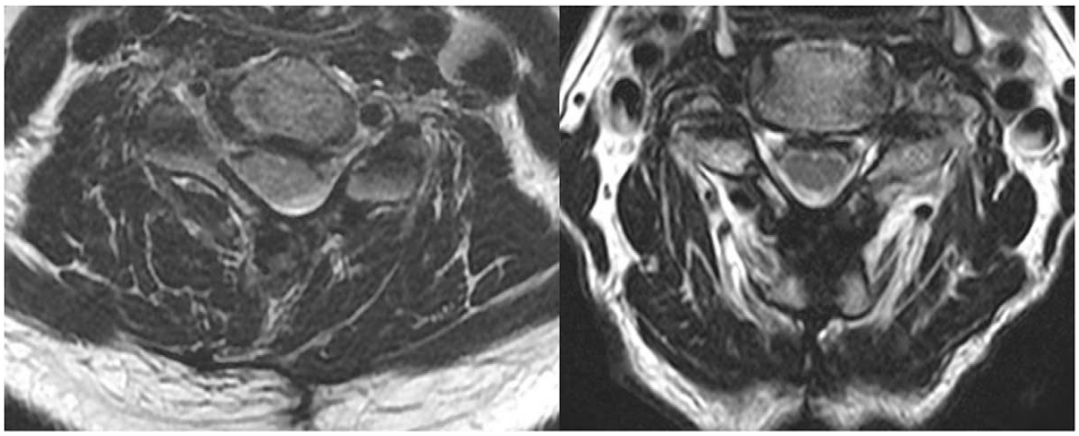


Figure 4. Influence of surgical techniques on soft tissues. The influence of endoscopic surgery (left) on the soft tissues differs substantially from that of conventional laminoplasty (right). MRI shows no changes of fatty degeneration in the paravertebral muscles after CMEL.

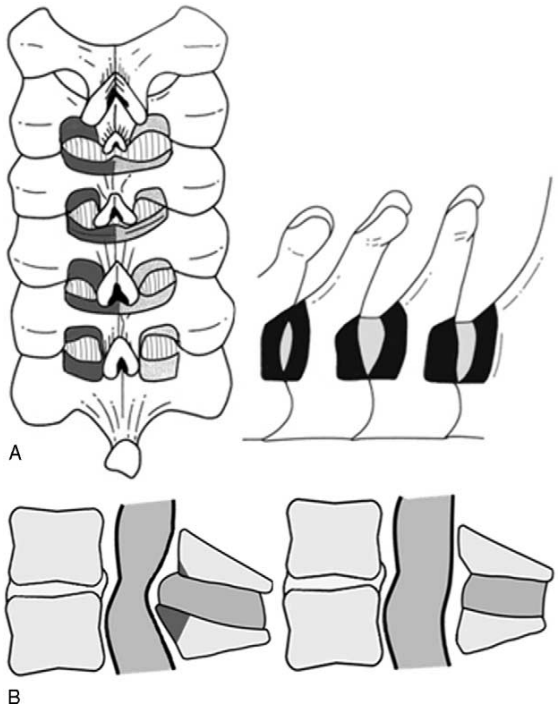


Figure 5. The microendoscopic procedure extends adaptation for patients of CSM. **(A)** Drilling and tunneling are performed from the inferior edge of the upper lamina to the superior edge of the lower lamina. The interlaminar space is enlarged until the attachment of the ligamentum flavum. **(B)** The ligamentum flavum is removed, and the spinal cord is decompressed, resulting in successful decompression of the articular segment.

with 1 lesion at the C4–C5 level underwent C3–C5 ELAP surgery, whereas a patient with CSM with 2 lesions at the C4–C5 and C5–C6 levels underwent C3–C6 ELAP surgery. However, CEL surgery was performed at the affected levels alone. Thus, it is possible that there may be another factor correlated with the decompression of the articular segment, besides the posterior shift of the spinal cord in CSM, unlike that in clinical conditions such as the ossification of the posterior longitudinal ligament.

The results of this study show that CMEL for CSM achieved a similar clinical outcome as ELAP with respect to the JOA, JOACMEQ, and SF-36 scores. Moreover, CMEL preserved the sagittal alignment of the cervical spine. This may indicate posterior decompression was sufficiently achieved and that the pincer mechanism was resolved. In addition to the recovery of spinal cord function, the invasion to the extensor musculature in the CMEL group may reduce postoperative axial neck pain.

The major limitation of this study is its retrospective case-control design with the use of propensity score matching, and the fact that it was not a randomized controlled trial, which would have been more desirable. The propensity score was estimated using preoperative clinical severity, which did not include imaging factors such as the stages of stenosis. Regardless, this study provides preliminary findings of this new surgical procedure, because the results indicate CMEL for CSM may have some advantages compared with laminoplasty. Nevertheless, further studies are required to clarify the efficacy and safety of this procedure.

In conclusion, this study indicates the clinical outcomes of CMEL for patients with CSM are similar to those of conventional laminoplasty. Posterior decompression of the articular segment with a pincer mechanism in CMEL can be indicated for patients with CSM. This minimally invasive technique may have potential advantages compared with conventional ELAP, and may provide an alternative surgical option.

➢ Key Points

- ❑ Compression of the cervical spinal cord in cervical spondylotic myelopathy (CSM) consists of a pincer mechanism due to bulging discs and a hypertrophied ligamentum flavum.
- ❑ Cervical microendoscopic laminoplasty (CMEL), in which the interlaminar space is enlarged until the attachment of the ligamentum flavum, which is then removed, exhibited comparable clinical outcomes as conventional expansive laminoplasty (ELAP) according to propensity score matching analysis.
- ❑ Posterior decompression of the cervical spinal cord in CSM is sufficient to remove the elements of the articular segment, such as the ligamentum flavum and the superior or inferior edge of the lamina.
- ❑ CMEL is promising for reducing postoperative neck and shoulder complaints caused by ELAP-induced soft-tissue damage.

References

1. Kimura I, Shingu H, Nasu Y. Long-term follow-up of cervical spondylotic myelopathy treated by canal-expansive laminoplasty. *J Bone Joint Surg Br* 1995;77:956–61.
2. Yoshida M, Tamaki T, Kawakami M, et al. Indication and clinical results of laminoplasty for cervical myelopathy caused by disc herniation with developmental canal stenosis. *Spine (Phila Pa 1976)* 1998;23:2391–7.
3. Hirabayashi K, Toyama Y, Chiba K. Expansive laminoplasty for myelopathy in ossification of the longitudinal ligament. *Clin Orthop Relat Res* 1999;359:35–48.
4. Saruhashi Y, Hukuda S, Katsuura A, et al. A long-term follow-up study of cervical spondylotic myelopathy treated by "French window" laminoplasty. *J Spinal Disord* 1999;12:99–101.
5. Heller JG,Edwards CC 2nd, Murakami H, et al. Laminoplasty versus laminectomy and fusion for multilevel cervical myelopathy: an independent matched cohort analysis. *Spine (Phila Pa 1976)* 2001; 26:1330–6.
6. Seichi A, Takeshita K, Ohishi I, et al. Long-term results of double-door laminoplasty for cervical stenotic myelopathy. *Spine (Phila Pa 1976)* 2001;26:479–87.
7. Kawaguchi Y, Kanamori M, Ishihara H, et al. Minimum 10-year followup after en block cervical laminoplasty. *Clin Orthop Relat Res* 2003;411:129–39.

8. Sakaura H, Hosono N, Mukai Y, et al. Long-term outcome of laminoplasty for cervical myelopathy due to disc herniation: a comparative study of laminoplasty and anterior spinal fusion. *Spine (Phila Pa 1976)* 2005;30:756–9.
9. Chiba K, Ogawa Y, Ishii K, et al. Long-term results of expansive open-door laminoplasty for cervical myelopathy-average 14-year follow-up study. *Spine (Phila Pa 1976)* 2006;31:2998–3005.
10. Hosono N, Yonenobu K, Ono K. Neck and shoulder pain after laminoplasty. A noticeable complication. *Spine (Phila Pa 1976)* 1996;21:1969–73.
11. Kawaguchi Y, Matsui H, Ishihara H, et al. Axial symptoms after en bloc cervical laminoplasty. *J Spinal Disord* 1999;12:392–5.
12. Satomi K, Ogawa J, Ishii Y, et al. Short-term complications and long-term results of expansive open-door laminoplasty for cervical stenotic myelopathy. *Spine J* 2001;1:26–30.
13. Yoshida M, Tamaki T, Kawakami M, et al. Does reconstruction of posterior ligamentous complex with extensor musculature decrease axial symptoms after cervical laminoplasty?. *Spine (Phila Pa 1976)* 2002;27:1414–8.
14. Hosono N, Sakaura H, Mukai Y, et al. C3-6 laminoplasty takes over C3-7 laminoplasty with significantly lower incidence of axial neck pain. *Eur Spine J* 2006;15:1375–9.
15. Liu J, Ebraheim NA, Sanford CG Jr, et al. Preservation of the spinous process-ligament-muscle complex to prevent kyphotic deformity following laminoplasty. *Spine J* 2007;7:159–64.
16. Guiot BH, Khoo LT, Fessler RG. A minimally invasive technique for decompression of the lumbar spine. *Spine (Phila Pa 1976)* 2002;27:432–8.
17. Adamason TE. Microendoscopic posterior cervical laminoforaminotomy for unilateral radiculopathy: results of a new technique in 100 cases. *J Neurosurg* 2001;95 (1 Suppl):51–7.
18. Khoo LT, Fessler RG. Microendoscopic decompressive laminotomy for the treatment of lumbar stenosis. *Neurosurgery* 2002;51 (5 Suppl):S146–54.
19. Ikuta K, Arima J, Tanaka T, et al. Short-term results of microendoscopic posterior decompression for lumbar spinal stenosis. Technical note. *J Neurosurg Spine* 2005;2 (5):624–33.
20. Yabuki S, Kikuchi S. Endoscopic partial laminectomy for cervical myelopathy. *J Neurosurg Spine* 2005;2:170–4.
21. Minamide A, Yoshida M, Yamada H, et al. Clinical outcomes of microendoscopic decompression surgery for cervical myelopathy. *Eur Spine J* 2010;19:487–93.
22. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat* 1985;39:33–8.
23. Yoshida M, Otani K, Shibasaki K, et al. Expansive laminoplasty with reattachment of spinous process and extensor musculature for cervical myelopathy. *Spine (Phila Pa 1976)* 1992;17:491–7.
24. Hirabayashi K, Watanabe K, Wakano K, et al. Expansive open-door laminoplasty for cervical spinal stenotic myelopathy. *Spine (Phila Pa 1976)* 1983;8:693–9.
25. Yonenobu K, Abumi K, Nagata K, et al. Interobserver and intraobserver reliability of the japanese orthopaedic association scoring system for evaluation of cervical compression myelopathy. *Spine (Phila Pa 1976)* 2001;26:1890–4; discussion 1895.
26. Nikaido T, Kikuchi S, Yabuki S, et al. Surgical treatment assessment using the Japanese orthopedic association cervical myelopathy evaluation questionnaire in patients with cervical myelopathy: a new outcome measure for cervical myelopathy. *Spine (Phila Pa 1976)* 2009;34:2568–72.
27. Suda K, Abumi K, Ito M, et al. Local kyphosis reduces surgical outcomes of expansive open-door laminoplasty for cervical spondylotic myelopathy. *Spine (Phila Pa 1976)* 2003;28:1258–62.
28. Shiraishi T, Fukuda K, Yato Y, et al. Results of skip laminectomy: minimum 2-year follow-up study compared with open-door laminoplasty. *Spine (Phila Pa 1976)* 2003;28:2667–72.
29. Miyasaka K, Kaneda K, Sato S, et al. Myelopathy due to ossification or calcification of the ligamentum flavum: radiologic and histologic evaluations. *AJNR Am J Neuroradiol* 1983;4:629–32.
30. Baba H, Maezawa Y, Kawahara N, et al. Calcium crystal deposition in the ligamentum flavum of the cervical spine. *Spine (Phila Pa 1976)* 1993;18:2174–81.
31. Kokubun S, Sato T, Ishii Y, et al. Cervical myelopathy in the Japanese. *Clin Orthop Relat Res* 1996;323:129–38.
32. Otani K, Sato K, Yabuki S, et al. A segmental partial laminectomy for cervical spondylotic myelopathy anatomical basis and clinical outcome in comparison with expansive open-door laminoplasty. *Spine (Phila Pa 1976)* 2009;34:268–73.

The Journal of Arthroplasty 30 (2015) 403–406

Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Total Hip Arthroplasty After Rotational Acetabular Osteotomy

Hideya Ito, MD [a,b], Yoshio Takatori, MD, PhD [c], Toru Moro, MD, PhD [d], Hirofumi Oshima, MD [d], Hiroyuki Oka, MD [e], Sakae Tanaka, MD, PhD [b]

[a] Bone and Joint Orthopaedic Surgery, Japanese Red Cross Medical Center, Shibuya-ku, Tokyo, Japan
[b] Sensory and Motor System Medicine, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan
[c] Japan Community Health Care Organization Yugawara Hospital, Ashigara-gun, Kanagawa, Japan
[d] Division of Science for Joint Reconstruction, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan
[e] Department of Joint Disease Research, 22nd Century medical and Research Center, Graduate, School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

ARTICLE INFO

Article history:
Received 21 August 2014
Accepted 1 October 2014

Keywords:
rotational acetabular osteotomy
periacetabular osteotomy
acetabular dysplasia
total hip arthroplasty
acetabular cups

ABSTRACT

In this study, we aimed to determine whether the outcomes of total hip arthroplasty (THA) after rotational acetabular osteotomy (RAO) are equal to those of primary THA, and to elucidate the characteristics of THA after RAO. The clinical and radiographic findings of THA after RAO (44 hips), with minimum 24 months of follow-up, were compared with a matched control group of 58 hips without prior RAO. We found that the outcomes in terms of functional scores and complication rates did not differ between THA after RAO and THA without previous pelvic osteotomy, indicating that the results of THA after RAO are equivalent to those of primary THA. Although THA after RAO requires technical considerations, similar clinical outcomes to primary THA can be expected.

Rotational acetabular osteotomy (RAO) is a type of periacetabular osteotomy used to treat symptomatic dysplasia of the acetabulum [1]. This procedure involves restoration of the femoral head coverage, resulting in pain relief and delays or prevention of the onset of arthritis. In Japan, there are reportedly a higher proportion of patients with dysplastic hips than in other countries [2], and many of these patients have undergone RAO. While some studies have reported good results of RAO [3–6], some patients require subsequent total hip arthroplasty (THA) because of pain secondary to progression of arthritis.

Several reports are available on THA after periacetabular osteotomy [7–11]. Most authors reported that THA after periacetabular osteotomy requires technical consideration and careful radiographic evaluation because the acetabulum may undergo morphologic changes. In terms of clinical results, one study reported that Bernese periacetabular osteotomy does not compromise the outcome of THA [11], whereas another study reported that the outcomes of THA after triple innominate osteotomy were not equivalent to those of primary THA [8]. However, it should be noted that these studies all had small sample sizes or were not comparative studies. To date, only one published case report of THA after RAO is available [12], and the effects of a previous RAO on subsequent THA are still unknown.

In this study, we aimed to determine whether the outcomes of THA after RAO are equal to those of primary THA, and to elucidate the characteristics of THA after RAO by comparing the clinical and radiographic findings of patients who underwent THA after RAO with matched controls who underwent THA without prior RAO.

Materials and Methods

This investigation was a retrospective chart and radiographic review comparing two groups of patients. We obtained institutional ethics board approval for the study, which was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. All patients provided informed consent to participate in the study. Between 1999 and 2011, we performed THA on 45 hips in 43 patients who had previously undergone RAO. One patient was lost to follow-up, resulting in the study group comprising of 44 hips in 42 patients. For comparative purposes, 58 age- and gender-matched hips in 58 patients who had undergone THA for osteoarthritis secondary to hip dysplasia during same period were identified and included as the control group. None of the patients in the control group had undergone any prior pelvic osteotomy.

The preoperative data analyzed included age at THA, gender, interval from RAO to THA (years), body mass index (BMI), the Crowe classification [13] of hip joints, pre-THA contralateral hip joint status, and previous femoral osteotomy. Post-operative data comprised the follow-up duration after THA.

Conflict of interest statement: One of the co-author receives finance as a consultant for Msd K. K., Asahi Kasei Pharma Corporation, Teijin Pharma Limited, Daiichi Sankyo Company, Limited, and has also received finance for presentations at Eisai Co., Ltd., Ono Pharmaceutical Co., Ltd., Taisho Toyama Pharmaceutical Co., Ltd., Chugai Pharmaceutical Co., Ltd., Eli Lilly Japan K. K., Pfizer Japan Inc., Msd K. K., Asahi Kasei Pharma Corporation, Teijin Pharma Limited, Daiichi Sankyo Company, Limited.

The Conflict of Interest statement associated with this article can be found at http://dx.doi.org/10.1016/j.arth.2014.10.002.

Reprint requests: Hideya Ito, MD, Bone and Joint Orthopaedic Surgery, Japanese Red Cross medical Center, 4-1-22 Hiroo Shibuya-ku, Tokyo 150-8935, Japan.

http://dx.doi.org/10.1016/j.arth.2014.10.002
0883-5403/

Surgical Procedure

All THAs were performed in the lateral decubitus position and through a posterior approach. The incision used differed from that used in the preceding RAO, which had been performed through a combined anterior and posterior approach with a single incision, as described by Ninomiya et al [1]. We used the combined approach described by Lusskin et al [14] for 35/44 hips in the study group and 42/58 hips in the control group. We did not perform trochanteric osteotomy in any joints. We attempted to place the acetabular cup with an abduction angle of between 30° and 50° [15]. After the acetabular preparation, the center of reaming was decided, and a gouge was used to remove the subchondral bone to measure the distance to the medial wall. Initial medialization of the acetabular reaming was performed using the smallest reamer, after which the diameter of the reamer was gradually increased. When there was uniform contact between the reamer and acetabular bone, a cup of that size was selected. All patients received a cementless acetabular component with 4 fins and additional screw fixation if required. After the final femoral reaming and rasping, trial reduction was performed. If a bony impingement occurred, any osteophytes of the acetabulum were removed using a chisel or bone rongeur forceps luer. Upon resolving the bony impingement, the final implantation of the femoral component was performed. All femoral components used were also cementless devices. The Mallory-Head acetabular and Bimetric stem systems (Biomet, Warsaw, IN, USA) were used on 32 hips in the study group and 38 hips in the control group, whereas the Q5LP acetabular and K-MAX stem systems (Kyocera Medical Corp, Osaka, Japan) were used in 12 and 20 hips in the study and control groups, respectively.

Computed tomography (CT) scans were obtained in all patients in the study group in order to determine the three-dimensional shape of the acetabulums.

Operative Data and Clinical Evaluation

Operative data, including the operative time, intraoperative estimated blood loss, removal of osteophytes, and the size of acetabular cups used, were obtained using clinical records.

Hip joint function was evaluated according to the Merle d'Aubigné-Postel score [16] preoperatively and at the final follow-up. Reoperation and complications, including infection, venous thromboembolism, dislocation, nerve palsy, and wound healing problems, were recorded.

Radiographic Evaluation

Radiographic evaluations were performed using anteroposterior radiographs taken before and immediately after THA, and at the final follow-up. The acetabular cup position was evaluated on the radiographs obtained immediately post-surgery. We measured the abduction angle of the acetabular cup and the hip joint center position. The hip joint center position was defined as the vertical and horizontal distances from the teardrop, as described by Fukui et al [17] (Fig. 1). The magnification of each radiograph was calibrated from the known and measured diameters of the prosthetic femoral head. Loosening of the acetabular cup and heterotopic bone formation were evaluated on the radiographs obtained immediately post-THA and at the final follow-up. The acetabular cup was considered to be loosening if there was more than 3 mm of migration or a change of at least 4° in the abduction angle [18]. We used the classification system developed by Brooker et al [19] to qualitatively evaluate heterotopic bone formation.

Statistical Analysis

Statistical analysis of the differences between the study and control groups was conducted using JMP Pro 10.0 (SAS Institute, Cary, NC, USA). The independent-sample *t* test was used for continuous variables, whereas the chi-square test or Fisher's exact test was used for dichotomous values according to the validity conditions. All statistical tests were two-tailed, and a significance level of 0.05 was used.

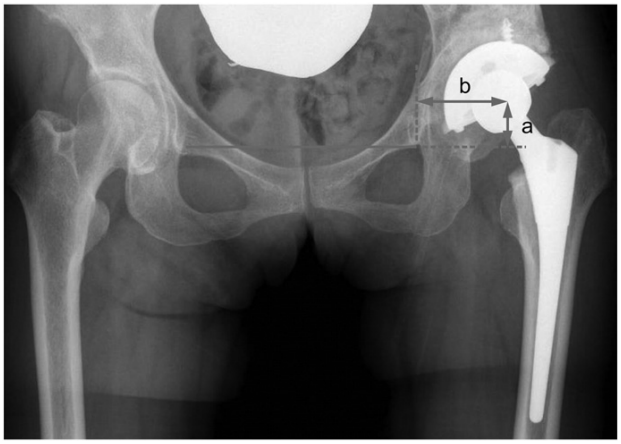


Fig. 1. Measurement of hip joint center position. The hip joint center position was defined as the vertical distance (a) and horizontal distance (b) from the teardrop.

Results

Demographics

The demographic and clinical baseline data of the patients are shown in Table 1. The mean age at the time of THA, gender, BMI, previous femoral osteotomy, and follow-up duration did not differ significantly between the two groups. Furthermore, the ratio of the Crowe classification of the preoperative hip joints and the contralateral hip joint status were also not significantly different between the groups. The average time interval between RAO and THA was 21 years (range, 7–37 years).

Operative Data and Clinical Evaluation

The operative and clinical data are shown in Table 2. The operative time in the study group was significantly longer than in the control

Table 1
Baseline Characteristics of the Study Patients (n = 100).

	Study Group	Control Group	*P*
Number of hips	44	58	
Number of patients	42	58	
Gender (M/F)	2/40	2/56	
Age at RAO (years)	34 ± 12.4 (11–53)	N/A	
Age at THA (years)	55.6 ± 7.8 (36–72)	56 2 ± 5.1 (46–67)	0.64
Interval from RAO to THA (years)	21 ± 7.3 (7–37)	N/A	
BMI (kg/m^2)	22.8 ± 3.4 (17.3–32.0)	22.3 ± 2.7 (17.5–27.6)	0.65
Crowe classification			0.87
I	29 (66.0%)	39 (67.3%)	
II	10 (22.7%)	14 (24.1%)	
III	3 (6.8%)	4 (6.9%)	
IV	2 (4.5%)	1 (1.7%)	
Contralateral joint			0.77
Normal	14 (31.8%)	21 (38.1%)	
OA	20 (45.5%)	27 (46.6%)	
THA	10 (22.7%)	10 (17.2%)	
Previous femoral osteotomy	2 (4.5%)	4 (6.9%)	1.00
Follow-up (months)	55.8 ± 36.2 (24–107)	62.9 ± 28.4 (24–95)	0.06

Data are presented as mean ± standard deviation (range) or number (%). Abbreviations: M, male; F, female; RAO, rotational acetabular osteotomy; THA, total hip arthroplasty; N/A, not applicable; BMI, body mass index; OA, osteoarthritis.

Table 2
Operative Data and Clinical Evaluations.

	Study Group (n = 44)	Control Group (n = 58)	P
Operative time (min)	177 ± 41 (115–227)	161 ± 36 (91–206)	0.03[a]
Blood loss (g)	567 ± 232 (140–1445)	524 ± 254 (50–1710)	0.15
Osteophyte removal	40 (90.9%)	31 (53.4%)	<0.001[a]
Combined approach	35 (79.5%)	42 (72.4%)	0.33
Cup size (mm)	50 ± 3.2 (46–58)	48 ± 2.1 (46–54)	<0.001[a]
MA score (preoperative)			
Total	8.3 ± 1.7 (5–12)	8.4 ± 2.1 (4–13)	0.74
Pain	2.3 ± 0.7 (1–4)	2.2 ± 0.8 (1–4)	0.26
Mobility	3.2 ± 1.2 (1–6)	3.6 ± 1.3 (1–6)	0.04[a]
Walking	2.8 ± 0.8 (1–5)	2.7 ± 0.8 (1–5)	0.31
MA score (last follow-up)			
Total	15.2 ± 1.7 (11–18)	15.7 ± 1.8 (11–18)	0.12
Pain	5.3 ± 0.6 (4–6)	5.5 ± 0.6 (4–6)	0.05
Mobility	5.1 ± 0.9 (3–6)	5.4 ± 0.8 (3–6)	0.12
Walking	4.8 ± 0.8 (3–6)	4.8 ± 1 (2–6)	0.78
Complications			
Infection	0	1	1.00
VTE	0	2	1.00
Dislocation	0	0	1.00
Nerve palsy	0	0	1.00
Wound healing problems	0	0	1.00
Reoperation	0	1	1.00

Data are presented as mean ± standard deviation (range) or number (%). Abbreviations: MA, Merle d'Aubigné-Postel; VTE, venous thromboembolism.
[a] $P < 0.05$.

group ($P = 0.029$); however, there were no differences in the estimated blood loss. In 40/44 hips (90.9%) in the study group, removal of osteophytes of the acetabular anterior wall was performed because of bony impingement, whereas this procedure was performed in only 31/58 hips (53.4%) in the control group.

The mean diameter of the acetabular cup used was 50 mm (range, 46–58 mm) in the study group, and 48 mm (range, 46–54 mm) in the control group ($P < 0.001$). There were no differences in the preoperative total Merle d'Aubigné-Postel score between the groups; however, the mobility score in the study group was significantly lower than in the control group ($P = 0.043$). At the last follow-up, the total Merle d'Aubigné-Postel, pain, mobility, and walk scores in the study and control groups were significantly improved compared with the preoperative scores. However, no significant differences were observed between the groups in terms of the improvements in the clinical results from before THA and the last follow-up. Reoperation was not needed for any patient. Postoperative complications included one case of infection and 2 cases of venous thromboembolism in the control group, whereas there were no cases of dislocation, wound healing problems, or nerve palsy.

Radiographic Evaluation

The radiographic data are shown in Table 3. The mean acetabular cup abduction angles were 40.7° (range, 30°–52°) and 43.5° (range, 22°–66°) in the study and control groups, respectively. Outliers of acetabular cup abduction angle were one hip >50° in the study group, and 10 hips >50° in the control group ($P = 0.021$).

The mean vertical distances of the hip joint center position after THA were 25.7 mm (range, 11–40 mm) and 23.7 mm (range, 13–41 mm) in the study and control groups, respectively. The mean horizontal distances were 31.2 mm (range, 21–42 mm) and 28.1 mm (range, 19–37 mm) in the study and control groups, respectively. While there was no significant difference in the vertical distance, the horizontal distance in the study group was found to be significantly larger than in the control group ($P = 0.002$), suggesting that the acetabular cup of THA after RAO was placed laterally. Moreover, there was no loosening of the acetabular and femoral component in either group. Heterotopic

Table 3
Radiographic Evaluations.

	Study Group (n = 44)	Control Group (n = 58)	P
Loosening	0	0	
Heterotopic ossification			0.008[a]
0	29 (65.9%)	52 (89.7%)	
1	11 (25%)	6 (10.3%)	
2	3 (7.0%)	0 (0%)	
3	1 (2.3%)	0 (0%)	
Cup abduction (°)	40.7 ± 5.2 (30–52)	43.5 ± 8.2 (22–66)	0.02[a]
Hip joint center			
Vertical distance (mm)	25.7 ± 6.5 (11–40)	23.7 ± 5.7 (13–41)	0.09
Horizontal distance (mm)	31.2 ± 5.3 (21–42)	28.1 ± 3.8 (19–37)	0.002[a]

Data are presented as mean ± standard deviation (range) or number (%).
[a] $P < 0.05$.

bone formations were seen in 15/44 hips (34.1%) (Grade I: 11 hips, Grade II: 3 hips, and grade III: 1 hip) in the study group and in 6/58 hips (10.3%) (Grade I: 6 hips) in the control group ($P = 0.008$).

Discussion

In this study, we demonstrated that the results of THA after RAO were comparable to those of primary THA, and reported on 7 specific characteristics of THA after RAO. We found that the outcomes in terms of the functional scores and complication rates did not differ between THA after RAO and THA without previous pelvic osteotomy, indicating that the results of THA after RAO are equivalent to those of primary THA. The characteristics of THA after RAO (study group) were as follows: the preoperative range of hip motion was poorer, the operative time was longer, the acetabular cups used were larger, removal of osteophytes was needed in more cases, heterotopic bone formations after THA were seen more frequently, the abduction angles of the acetabular cups were smaller, and their position tended to be lateral.

In most patients, removal of large osteophytes was needed after RAO, and we speculate that the presence of osteophytes might be associated with a poorer preoperative range of hip joint motion. In turn, removal of the osteophytes and the poor hip joint motion might be responsible for the prolonged operation time observed in the study group. Moreover, the acetabular cups used in the study group were larger than in the control group, indicating that the acetabulums after RAO may become wider than before RAO.

Interestingly, the abduction angles of the acetabular cups were lower in the study group than those in the control group. The abduction angles of 43/44 (97.7%) acetabular cups in the study group were within the target range, compared to only 48/58 (82.8%) acetabular cups in the control group. All outliers were >50°. These data indicated that the acetabular cups in the control group were occasionally placed too steep, likely because of the presence of acetabular dysplasia [20]. Correction of acetabular dysplasia by RAO may help surgeons place the acetabular cups in an adequate abduction angle. The acetabular cup position in the study group tended to be more lateralized than in the control group. Kaneuji et al [21] reported that the normal hip joint center was 31.5 ± 5 mm lateral from the teardrop. While the acetabular cup position in our study seemed to be largely acceptable, it has been recognized that acetabular cups in the upper and lateral position may lead to poor results during THA [22–24]. The appropriate cup position depends on the position and shape of the acetabulum, and RAO prior to THA may influence the cup position. Thus, this should be evaluated both during the preoperative planning and intraoperatively.

Although anteversion of the acetabular component was not measured in the present study, signs of retroversion of the acetabulum after periacetabular osteotomy have been previously reported [11], and preoperative CT is effective for three-dimensional evaluations of the acetabulum and osteophyte (Fig. 2).

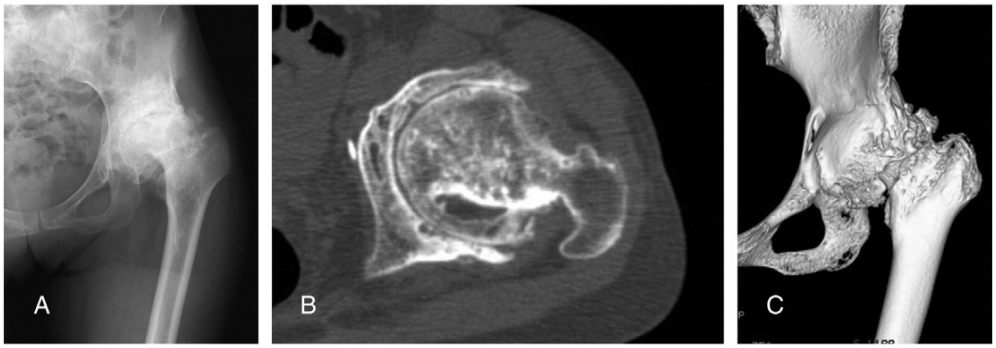


Fig. 2. The preoperative left hip of a 63-year-old woman with osteoarthritis after rotational acetabular osteotomy. (A) Anteroposterior radiograph showing progressive osteoarthritis changes. (B) Axial computed tomography image showing a large osteophyte of the acetabulum needing removal, and the depth of the acetabulum. (C) Three-dimensional computed tomography image showing the three-dimensional shape of the acetabulum.

In the present study, heterotopic bone formations after THA were seen more frequently in the study group (34.1%) than in the control group (10.3%). Similarly, previous studies reported that heterotopic bone formations were seen in 10%–42% of cases of THA after periacetabular osteotomy [10,11]. However, since most cases were classed as stage I or II in this study, heterotopic bone formations did likely not affect the clinical outcomes [25]. Moreover, we were concerned that wound healing problems would occur in the study group, since two separate incisions were performed, however, this did not occur in any case. We often used the combined approach in cases with poor range of hip motion or large osteophytes needing removal, and did not have to use trochanteric osteotomy in any case, suggesting that the approach is useful for exposing the anterior hip joint. Although we did not use any cemented devices or reinforced rings in this study, no component was loose. Thus, standard cementless devices appear to be useful in THA after RAO.

Previous studies have shown that RAO for dysplastic hips results in short-term hip pain relief and intermediate-term prevention of progression of arthritis, and the present study showed that the results of THA after RAO are equivalent to those of primary THA. Accordingly, we believe that RAO followed by THA could have long-term effects and result in long-term maintenance of the hip joint function for young patients with dysplastic hips, and longer follow-up studies are needed to confirm this.

In previous studies on the topic, the average time interval between a preceding periacetabular osteotomy and THA has been reported to range between 3.3 and 7.5 years [7,8,10,11]. In the present study, the average time interval between RAO and THA was 21 years. Thus, our results suggest that RAO may be a good procedure of joint preservation for dysplasia of the acetabulum. However, further large-scale studies are still needed in the future to investigate which periacetabular osteotomy procedure preserves the joint for the longest duration

This study has several limitations, including its retrospective nature and relatively short follow-up period. Moreover, three-dimensional radiographic evaluation was not performed.

In conclusion, although THA after RAO requires technical considerations, similar clinical outcomes to primary THA can be expected. However, further large-scale, long-term studies using three-dimensional radiographs are needed in the future to confirm our findings.

References

1. Ninomiya S, Tagawa H. Rotational acetabular osteotomy for the dysplastic hip. J Bone Joint Surg Am 1984;66:430.
2. Yoshimura N, Campbell L, Hashimoto T, et al. Acetabular dysplasia and hip osteoarthritis in Britain and Japan. Br J Rheumatol 1998;37:1193.
3. Ninomiya S. Rotational acetabular osteotomy for the severely dysplastic hip in the adolescent and adult. Clin Orthop Relat Res 1989;247:127.
4. Nakamura S, Ninomiya S, Takatori Y, et al. Long-term outcome of rotational acetabular osteotomy: 145 hips followed for 10–23 years. Acta Orthop Scand 1998;69:259.
5. Okano K, Enomoto H, Osaki M, et al. Outcome of rotational acetabular osteotomy for early hip osteoarthritis secondary to dysplasia related to femoral head shape: 49 hips followed for 10–17 years. Acta Orthop 2008;79:12.
6. Takatori Y, Ninomiya S, Nakamura S, et al. Long-term results of rotational acetabular osteotomy in patients with slight narrowing of the joint space on preoperative radiographic findings. J Orthop Sci 2001;6:137.
7. Baqué F, Brown A, Matta J. Total hip arthroplasty after periacetabular osteotomy. Orthopedics 2009;32:399.
8. Peters CL, Beck M, Dunn HK. Total hip arthroplasty in young adults after failed triple innominate osteotomy. J Arthroplasty 2001;16:188.
9. Wozniak W, Nikratowicz P, Owczarski T, et al. Total hip arthroplasty following Ganz periacetabular osteotomy. Cases study. Ortop Traumatol Rehabil 2010;12:561.
10. Hartig-Andreasen C, Stilling M, Søballe K, et al. Is cup positioning challenged in hips previously treated with periacetabular osteotomy? J Arthroplasty 2014;29:763.
11. Parvizi J, Burmeister H, Ganz R. Previous Bernese periacetabular osteotomy does not compromise the results of total hip arthroplasty. Clin Orthop Relat Res 2004;423:118.
12. Shinoda S, Hasegawa Y, Kawabe K, et al. Total hip arthroplasty for failed rotational acetabular osteotomy: a report of three cases. Nagoya J Med Sci 1998;61:53.
13. Crowe JF, Mani VJ, Ranawat CS. Total hip replacement in congenital dislocation and dysplasia of the hip. J Bone Joint Surg Am 1979;61:15.
14. Lusskin R, Goldman A, Absatz M. Combined anterior and posterior approach to the hip joint in reconstructive and complex arthroplasty. J Arthroplasty 1988;3:313.
15. Lewinnek GE, Lewis JL, Tarr R, et al. Dislocations after total hip-replacement arthroplasties. J Bone Joint Surg Am 1978;60:217.
16. D'Aunigné RM, Postel M. Functional results of hip arthroplasty with acrylic prosthesis. J Bone Joint Surg Am 1954;36:451.
17. Fukui K, Kaneuji A, Sugimori T, et al. A radiological study of the true anatomical position of the acetabulum in Japanese women. Hip Int 2011;21:311.
18. Massin P, Schmidt L, Engh CA. Evaluation of cementless acetabular component migration. An experimental study. J Arthroplasty 1989;4:245.
19. Brooker AF, Bowerman JW, Robinson RA, et al. Ectopic ossification following total hip replacement. Incidence and a method of classification. J Bone Joint Surg Am 1973;55:1629.
20. Rittmeister M, Callitsis C. Factors influencing cup orientation in 500 consecutive total hip replacements. Clin Orthop Relat Res 2006;445:192.
21. Kaneuji A, Sugimori T, Ichiseki T, et al. Minimum ten-year results of a porous acetabular component for Crowe I to III hip dysplasia using an elevated hip center. J Arthroplasty 2009;24:187.
22. Pagnano W, Hanssen AD, Lewallen DG, et al. The effect of superior placement of the acetabular component on the rate of loosening after total hip arthroplasty. J Bone Joint Surg Am 1996;78:1004.
23. Doehring TC, Rubash HE, Shelley FJ, et al. Effect of superior and superolateral relocations of the hip center on hip joint forces. An experimental and analytical analysis. J Arthroplasty 1996;11:693.
24. Morag G, Zalzal P, Liberman B, et al. Outcome of revision hip arthroplasty in patients with a previous total hip replacement for developmental dysplasia of the hip. J Bone Joint Surg (Br) 2005;87:1068.
25. Kocic M, Lazovic M, Mitkovic M, et al. Clinical significance of the heterotopic ossification after total hip arthroplasty. Orthopedics 2010;33:16.

The Spine Journal 15 (2015) 622–628

Clinical Study

The association of combination of disc degeneration, end plate signal change, and Schmorl node with low back pain in a large population study: the Wakayama Spine Study

Masatoshi Teraguchi, MD, PhD[a], Noriko Yoshimura, MD, PhD[b], Hiroshi Hashizume, MD, PhD[a,*], Shigeyuki Muraki, MD, PhD[c], Hiroshi Yamada, MD, PhD[a], Hiroyuki Oka, MD[b], Akihito Minamide, MD, PhD[a], Hiroyuki Nakagawa, MD, PhD[a], Yuyu Ishimoto, MD, PhD[a], Keiji Nagata, MD, PhD[a], Ryohei Kagotani, MD, PhD[a], Sakae Tanaka, MD, PhD[d], Hiroshi Kawaguchi, MD, PhD[e], Kozo Nakamura, MD, PhD[f], Toru Akune, MD, PhD[b], Munehito Yoshida, MD, PhD[a]

[a]*Department of Orthopaedic Surgery, Wakayama Medical University, 811-1 Kimiidera, Wakayama City, Wakayama 641-8510, Japan*
[b]*Department of Joint Disease Research, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*
[c]*Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*
[d]*Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*
[e]*Japan Community Health Care Organization Tokyo Shinjuku Medical Center, 6-1-1 Shinjuku, Shinjuku-ku, Tokyo 160-8402, Japan*
[f]*Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, 1 Namiki 4-chome, Tokorozawa City, Saitama 359-8555, Japan*

Received 17 July 2014; revised 21 October 2014; accepted 20 November 2014

Abstract

BACKGROUND CONTEXT: Disc degeneration (DD) reportedly causes low back pain (LBP) and is often observed concomitantly with end plate signal change (ESC) and/or Schmorl node (SN) on magnetic resonance imaging.

PURPOSE: The purpose of this study was to examine the association between DD and LBP, considering ESC and/or SN presence, in a large population study.

STUDY DESIGN/SETTING: Cross-sectional population-based study in two regions of Japan.

PATIENT SAMPLE: Of 1,011 possible participants, data from 975 participants (324 men, 651 women; mean age, 66.4 years; range, 21–97 years) were included.

OUTCOME MEASURES: Prevalence of DD, ESC, and SN alone and in combination in the lumbar region and the association of these prevalence levels with LBP.

METHODS: Sagittal T2-weighted images were used to assess the intervertebral spaces between L1–L2 and L5–S1. Disc degeneration was classified using the Pfirrmann classification system (grades 4 and 5 indicated degeneration); ESC was defined as a diffuse high signal change along

FDA device/drug status: Not applicable.

Author disclosures: ***MT:*** Grant: Grant-in-Aid for Young Scientists, B26861206 (D, Paid directly to institution). ***NY:*** Grant: Grant-in-Aid for Scientific Research, B26293139 (F, Paid directly to institution) and Grant-in-Aid for Challenging Exploratory Research, 24659317 (D, Paid directly to institution). ***HH:*** Grant: Grant-in-Aid for Scientific Research, C26462249 (D, Paid directly to institution). ***SM:*** Grant: Grant-in-Aid for Scientific Research, B23390356 (F, Paid directly to institution). ***HY:*** Grant: Grant-in-Aid for Scientific Research, C25462305 (D, Paid directly to institution). ***HO:*** Grant: Grant-in-Aid for Challenging Exploratory Research, 24659666 (D, Paid directly to institution). ***AM:*** Nothing to disclose. ***HN:*** Nothing to disclose. ***YI:*** Grant: Grant-in-Aid for Young Scientists, B25860448 (D, Paid directly to institution). ***KNag:*** Nothing to disclose. ***RK:*** Grant: Grant-in-Aid for Young Scientists, B26860419 (D, Paid directly to institution). ***ST:*** Nothing to disclose. ***HK:*** Nothing to disclose. ***KNak:*** Nothing to disclose. ***TA:*** Grant: Grant-in-Aid for Challenging Exploratory Research, 25670293 (D, Paid directly to institution). ***MY:*** Grant: Health Labour Sciences Research Grant, H20-Choju-Ippan-009 (D, Paid directly to institution).

The disclosure key can be found on the Table of Contents and at www.TheSpineJournalOnline.com.

This study was funded by the Ministry of Education, Culture, Sports, Science and Technology, and the Ministry of Health, Labour and Welfare in Japan. The sponsors did not contribute to the study design, data collection, data analysis, data interpretation, or the writing of the manuscript.

* Corresponding author. Department of Orthopedic Surgery, Wakayama Medical University, 811-1 Kimiidera, Wakayama City, Wakayama 641-8510, Japan. Tel.: +81-73-441-0645; Fax: +81-73-448-3008.

E-mail address: hashizum@wakayama-med.ac.jp (H. Hashizume)

http://dx.doi.org/10.1016/j.spinee.2014.11.012

either area of the end plate, and SN was defined as a small well-defined herniation pit with a surrounding wall of hypointense signal. Logistic regression analysis was used to determine the odds ratios (ORs) and confidence intervals (CIs) for LBP in the presence of radiographic changes in the lumbar region and at each lumbar intervertebral level, compared with patients without radiographic change, after adjusting for age, body mass index, and sex.

RESULTS: The prevalence of lumbar structural findings was as follows: DD alone, 30.4%; ESC alone, 0.8%; SN alone, 1.5%; DD and ESC, 26.6%; DD and SN, 12.3%; and DD, ESC, and SN, 19.1%. These lumbar structural findings were significantly associated with LBP in the lumbar region overall, as follows: DD, ESC, and SN, OR 2.17, 95% CI 1.2–3.9; L1–L2, OR 6.00, 95% CI 1.9–26.6; L4–L5, OR 2.56, 95% CI 1.4–4.9; and L5–S1, OR 2.81, 95% CI 1.1–2.3. The combination of DD and ESC was significantly associated with LBP as follows: L3–L4, OR 2.43, 95% CI 1.5–4.0; L4–L5, OR 1.82, 95% CI 1.2–2.8; and L5–S1, OR 1.60, 95% CI 1.1–2.3.

CONCLUSIONS: Our data suggest that DD alone is not associated with LBP. By contrast, the combination of DD and ESC was highly associated with LBP.

Keywords: Disc degeneration; End plate signal change; Schmorl node; Low back pain; Large population study; ROAD study

Introduction

Low back pain (LBP) causes functional impairment, diminished quality of life, loss of working ability, potential psychological distress, and increased health-care costs [1–3]. Magnetic resonance imaging (MRI) has become widely used in LBP diagnosis [4–15].

From the perspective of discogenic pain, the association between disc degeneration (DD) and symptoms remains controversial. Several reports have found that DD was a source of LBP [4–7], but others reported no association between DD and LBP [8,9]. This discrepancy is partly explained by the fact that DD often occurs concomitantly with various radiographic changes such as end plate signal change (ESC) and Schmorl node (SN). However, few studies have reported on the association of ESC and SN with LBP [10–14], and furthermore, to the best of our knowledge, no population-based study has examined the association of the combination of DD, ESC, and SN with LBP.

The purposes of this study were to examine the prevalence of combinations of DD, ESC, and SN in the lumbar region overall and to clarify the associations between LBP and combinations of DD, ESC, and SN in a large population.

Methods

Participants

The Wakayama Spine Study is a population-based study of degenerative spinal disease [15–17] performed in a subcohort of the large-scale population-based cohort study Research on Osteoarthritis/Osteoporosis against Disability (ROAD) [18,19]. Research on Osteoarthritis/Osteoporosis against Disability is a nationwide prospective study of bone and joint diseases consisting of population-based cohorts established in three communities in Japan. The participants were recruited from listings of resident registrations in three communities that have different characteristics: an urban region in Itabashi, Tokyo; a mountainous region in Hidakagawa, Wakayama; and a coastal region in Taiji, Wakayama. The inclusion criteria, apart from residence in the communities mentioned previously, were the ability to walk to the survey site, report data, and understand and sign an informed consent form. The age of the participants recruited from the urban region was 60 years or older and that of the participants from the other two regions was 40 years or older [18]. A second visit of the ROAD study to the mountainous region of Hidakagawa and the coastal region of Taiji was performed between 2008 and 2010. From the inhabitants participating in the second visit of the ROAD study, 1,063 volunteers were recruited to MRI examinations. Among these volunteers, 52 declined to attend the examination and 1,011 provided an additional written informed consent for the mobile MRI examination and were recruited for registration in the Wakayama Spine Study. Among the 1,011 participants, those who had an MRI-sensitive implanted device (e.g., pacemaker) or other disqualifier were excluded. Ultimately, 980 individuals underwent whole-spine MRI. One participant who had undergone a previous cervical operation and four participants who had undergone previous posterior lumbar fusion were excluded from the analysis. Thus, whole-spine MRI results were available for 975 participants (324 men and 651 women) with an age range of 21 to 97 years (mean, 67.2 years for men and 66.0 years for women). Experienced board-certified orthopedic surgeons also asked all participants the following question regarding LBP: ‘‘Have you experienced LBP on most days during the past month, in addition to now?’’ Those who answered ‘‘yes’’ were defined as having LBP based on the previous studies [20–24]. All study participants provided informed consent, and the

EVIDENCE & METHODS

Context
Controversy persists regarding the association of abnormal magnetic resonance imaging (MRI) findings with chronic axial and mechanical back pain. As part of a population-based study in Japan, the authors sought to correlate degenerative changes as observed on MRI with patient-reported back pain.

Contribution
A total of 975 patients were evaluated as part of this arm of a wider study. Degenerative changes, end plate signal changes and Schmorl's nodes were present in nearly 1 in 5 study participants. Degenerative changes alone were identified in approximately 30% of the population. The authors maintain that degenerative changes alone were not associated with low back pain, while the presence of degenerative and end plate signal changes significantly increased the odds of a patient reporting low back pain.

Implications
These findings reinforce data presented in prior work. While this investigation enjoys the advantage of a large population-based study, it suffers from the fact that it was cross-sectional and therefore cannot speak to trends over time. In addition, it seems that the determination of low back pain was based on patient responses to a single study question. This raises the possibility for bias to be introduced from misinterpretation, misunderstanding, or misinformation. Furthermore, sociodemographic and cultural factors unique to the population under study may impair the capacity for generalization of these findings to patients of other nationalities or those living in different social contexts.

—The Editors

study design was approved by the appropriate ethics review boards.

Magnetic resonance imaging

A mobile MRI unit (Excelart 1.5 T; Toshiba, Tokyo, Japan) was used, and whole-spine MRI was performed for all participants on the same day as the examination. The participants were supine during the MRI, and those with rounded backs used triangular pillows under their head and knees. The imaging protocol included sagittal T2-weighted fast spin echo (repetition time, 4,000 ms/echo; echo time, 120 ms; and field of view, 300×320 mm) and axial T2-weighted fast spin echo (repetition time, 4,000 ms/echo; echo time, 120 ms; and field of view, 180×180 mm). Sagittal T1-weighted images were omitted owing to cost and time limitations; only T2-weighted images were obtained.

Radiographic assessment

Sagittal T2-weighted images were used to assess DD, ESC, and SN at all intervertebral levels from C2–C3 to L5–S1. The present study assessed L1–L2 to L5–S1 in the lumbar region.

Disc degeneration

Disc degeneration grading was performed by a board-certified orthopedic surgeon who was blinded to the background of the participants. The degree of DD on MRI was classified into five grades based on the Pfirrmann classification system [25], with grades 4 and 5 indicating DD. To evaluate intraobserver variability, 100 randomly selected MR images of the entire spine were rescored by the same observer more than 1 month after the first reading. Furthermore, to evaluate interobserver variability, 100 other MR images were scored by 2 orthopedic surgeons using the same classification. The intraobserver and interobserver variabilities of DD, as evaluated by kappa analysis, were 0.94 and 0.94, respectively.

End plate signal change

End plate signal change was defined as diffuse areas of high signal change along the end plates, tending to be linear and always parallel to the vertebral end plates on sagittal T2-weighted images. However, discerning the type of Modic changes [26,27] was not possible because of cost and time limitations of this large-scale study. Because the T1 sequence was not obtained, we considered Modic Type I/II (T2 high signal intensity end plate change) to reflect the presence of ESC and T2 isosignal intensity and Modic Type III (T2 low signal intensity end plate change) to reflect the absence of ESC. To evaluate the intraobserver and interobserver variabilities, two orthopedic surgeons scored MR images in the same manner. The intraobserver and interobserver variabilities of ESC evaluated by kappa analysis were 0.86 and 0.82, respectively.

Schmorl node

Schmorl node was characterized by a localized defect at the rostral, caudal, or both end plates, with a well-defined herniation pit in the vertebral body with or without a surrounding sclerotic rim (low signal on T2-weighted image) [14,28]. Erosive defects in the end plate in degenerate segments were not considered as SN [14,28]. To evaluate intraobserver and interobserver variabilities, two orthopedic surgeons scored MR images in the same manner. The intraobserver and interobserver variabilities for SN evaluated by kappa analysis were 0.92 and 0.84, respectively.

Statistical analyses

Radiographic changes were compared between sexes using the chi-square test. Multivariate logistic regression analysis was used to estimate the radiographic changes of DD, ESC, and SN presence in the lumbar region as dependent variables, with LBP as the independent variable, after adjustment for age, body mass index (BMI), and sex. Multivariate logistic regression analysis was used to estimate the respective associations of eight combinations of radiographic changes (none; DD alone; SN alone; ESC alone; DD and SN; DD and ESC; SN and ESC; and DD, ESC, and SN) in the lumbar region as dependent variables, with LBP as the independent variable, after adjustment for age, BMI, and sex. Multivariate logistic regression analysis was also used to estimate the association of 8 combinations of radiographic changes at each intervertebral level (L1–L2 to L5–S1) in the lumbar region after adjustment for age, BMI, and sex. All statistical analyses were performed using JMP, version 8 (SAS Institute Japan, Tokyo, Japan).

Prevalence of DD, ESC, and/or SN, defined as the proportion of the number of participants who demonstrated the presence of DD, ESC, and/or SN in the lumbar region divided by the total number of participants, was used to describe the frequency of DD, ESC, and/or SN. In this analysis, to clarify the associated factors using multivariate logistic regression analysis, we entered a variable reflecting the observation of DD, ESC, and/or SN (1, presence; 0, absence) as a dependent variable.

Results

Table 1 shows the characteristics of the 975 participants in the present study including age and demographic measurements. Two-thirds of the participants were women.

The prevalence of DD, ESC, and SN in the lumbar region overall, without considering other radiographic changes, was 86.7%, 44.1%, and 29.6% in men and 89.6%, 48.7%, and 35.2% in women, respectively. Table 2 shows the prevalence of combinations of radiographic changes according to sex. DD alone demonstrated the highest prevalence, followed by DD and ESC and DD, ESC, and SN in both sexes. The prevalence of SN alone, ESC alone, or the combination of SN and ESC was small. The prevalence of combinations of radiographic changes in the lumbar region did not significantly differ between men and women.

When we evaluated DD, ESC, and SN in the lumbar region overall without considering other radiographic change, DD presence and ESC presence in the lumbar region were each significantly associated with LBP (DD: odds ratio [OR] 1.58, 95% confidence interval [CI] 1.02–2.49; ESC: OR 1.36, 95% CI 1.04–1.76). On the other hand, SN presence in the lumbar region was not significantly associated with LBP (OR 1.27, 95% CI 0.96–1.68). Next, to determine the effect of the combination of DD, ESC, and SN on LBP, we classified participants into eight groups: none; DD alone; ESC alone; SN alone; DD and ESC; DD and SN; SN and ESC; and DD, ESC, and SN. As shown in Table 3, the combination of DD, ESC, and SN in the lumbar region was significantly associated with LBP. Disc degeneration alone was not an associated factor for LBP.

Furthermore, as shown in Table 4, the effect of combinations of radiographic change at each intervertebral level from L1–L2 to L5–S1 on LBP was evaluated: the

Table 1
Characteristics of the 975 participants in the present study

	Overall	Men	Women
No. of participants	975	324	651
Demographic characteristics			
Age, y	66.4±13.5	67.2±13.9	66.0±13.4
Height, cm	156.4±9.4	164.6±7.2	151.5±7.2
Weight, kg	56.8±11.5	64.5±11.6	53.0±9.4
BMI, kg/m^2	23.3±3.6	23.6±3.4	23.1±3.7
Symptom			
LBP (%)	393 (40.3)	119 (36.7)	274 (42.1)

BMI, body mass index; LBP, low back pain.
Note: Values are the mean±standard deviation.

Table 2
Prevalence of combination of radiographic change in the lumbar region according to sex

	Overall (%)	Men (%)	Women (%)
Total	975	324	651
None	85 (8.7)	35 (10.8)	50 (7.7)
DD alone	296 (30.4)	104 (32.1)	192 (29.5)
ESC alone	8 (0.8)	3 (0.9)	5 (0.8)
SN alone	15 (1.5)	6 (1.9)	9 (1.4)
DD and ESC	259 (26.6)	85 (26.2)	174 (26.7)
DD and SN	120 (12.3)	37 (11.4)	83 (12.8)
SN and ESC	6 (0.6)	0 (0)	6 (0.9)
DD, ESC, and SN	186 (19.1)	54 (16.7)	132 (20.3)

DD, disc degeneration; ESC, end plate signal change; SN, Schmorl node.
Note: Chi-square test was used to determine differences in radiographic change between men and women.

Table 3
Association between LBP and radiographic changes in the lumbar region

	Proportion of participants with LBP (%)	OR (95% CI)
None	25/85 (29.4)	1
DD alone	112/296 (37.8)	1.35 (0.8–2.3)
ESC alone	0/8 (0)	—
SN alone	5/15 (33.3)	1.14 (0.3–3.6)
DD and ESC	107/259 (41.3)	1.51 (0.9–2.6)
DD and SN	45/120 (37.5)	1.26 (0.7–2.3)
SN and ESC	3/6 (50.0)	2.06 (0.4–11.9)
DD, ESC, and SN	96/186 (51.6)	2.17 (1.2–3.9)*

CI, confidence interval; DD, disc degeneration; ESC, end plate signal change; LBP, low back pain; OR, odds ratio; SN, Schmorl node.
Note: Proportion of participants with LBP means the number of participants with LBP/the number of participants with each radiographic change. ORs were calculated by multivariate logistic regression analysis after adjustment for age, body mass index, and sex.
* $p<.01$.

Table 4
Association between LBP and radiographic changes at each level in the lumbar region

	L1–L2		L2–L3		L3–L4		L4–L5		L5–S1	
	Proportion of participants with LBP (%)	OR (95% CI)	Proportion of participants with LBP (%)	OR (95% CI)	Proportion of participants with LBP (%)	OR (95% CI)	Proportion of participants with LBP (%)	OR (95% CI)	Proportion of participants with LBP (%)	OR (95% CI)
None	225/593 (37.9)	1	153/416 (36.8)	1	121/347 (34.9)	1	74/228 (32.5)	1	99/284 (34.9)	1
DD alone	76/193 (39.4)	1.0 (0.7–1.4)	138/322 (42.9)	1.19 (0.9–1.6)	161/400 (40.3)	1.17 (0.9–1.6)	184/449 (41.0)	1.36 (0.9–1.9)	176/440 (40.0)	1.14 (0.8–1.6)
ESC alone	11/27 (40.7)	1.06 (0.5–2.3)	9/23 (39.1)	0.96 (0.4–2.3)	3/10 (30.0)	0.74 (0.2–2.7)	3/16 (18.8)	0.50 (0.1–1.6)	6/15 (40.0)	1.24 (0.4–3.6)
SN alone	19/46 (41.3)	1.07 (0.6–2.0)	15/40 (37.5)	1.03 (0.5–2.0)	10/22 (45.5)	1.58 (0.6–3.8)	5/14 (35.7)	1.18 (0.3–3.5)	1/1 (100)	—
DD and ESC	16/36 (44.4)	1.20 (0.6–2.4)	33/69 (47.8)	1.37 (0.8–2.3)	55/93 (59.1)	2.43 (1.5–4.0)*	81/164 (49.4)	1.82 (1.2–2.8)†	96/201 (47.8)	1.60 (1.1–2.3)*
DD and SN	31/61 (50.8)	1.47 (0.8–2.6)	32/73 (43.8)	1.14 (0.7–1.9)	24/70 (34.3)	0.84 (0.5–1.5)	16/52 (30.8)	0.84 (0.4–1.6)	3/15 (20.0)	0.45 (0.1–1.5)
SN and ESC	2/3 (66.7)	2.48 (0.2–53.8)	0/6 (0)	—	2/2 (100)	—	1/2 (50.0)	2.04 (0.08–52.5)	0/0 (0)	—
DD, ESC, and SN	13/16 (81.3)	6.00 (1.9–26.6)‡	13/26 (50.0)	1.49 (0.7–3.4)	17/31 (54.8)	2.07 (0.9–4.5)	29/50 (58.0)	2.56 (1.4–4.9)‡	12/19 (63.2)	2.85 (1.1–2.3)*

CI, confidential interval; DD, disc degeneration; ESC, end plate signal change; LBP, low back pain; OR, odds ratio; SN, Schmorl node.
Note: Proportion of participants with LBP means the number of participants with LBP/the number of participants with each radiographic change. Multivariate logistic regression analysis of radiographic change was associated with LBP after adjustment for age, body mass index, and sex.
* $p<.05$.
† $p<.01$.
‡ $p<.005$.

combination of DD, ESC, and SN was significantly associated with LBP at L1–L2, L4–L5, and L5–S1. Furthermore, the combination of DD and ESC was significantly associated with LBP at L3–L4, L4–L5, and L5–S1.

Discussion

The prevalence of DD, ESC, or SN in the lumbar region has been examined in some previous studies [11–15,26–32], but, to the authors' knowledge, no population-based studies have assessed the prevalence of the combination of DD, ESC, and SN in a large population using MRI. First, we found that prevalence of combinations of DD, ESC, and SN in the lumbar region was approximately 20%. By contrast, the prevalence of ESC alone, SN alone, and combination of SN and ESC was quite small, which is partly explained by the fact that DD was reported to have a strong positive linear relationship with ESC and/or SN in the previous studies [14,28,33].

The association of DD with LBP remains controversial. An association between DD in the lumbar region and LBP was previously demonstrated in a twin study and other previous studies [15,30,31]. However, some reports have observed a high prevalence of DD among asymptomatic volunteers, with no association between DD and LBP [8,9]. These studies may have been limited in that they did not account for interactions between radiographic changes including DD, ESC, and SN. The present study found that the combination of DD, ESC, and SN was significantly associated with LBP, whereas DD alone was not.

The association of Modic changes, which are the gold standard to diagnose ESC, with clinical symptoms, has been controversial in the clinical studies based on patient series and a population-based cohort [10–12,29,32]. The present study found that the combination of DD and ESC at L3–L4, L4–L5, and L5–S1 was significantly associated with LBP. Degenerative change of end plates becomes a source of LBP and affects DD. Because the lumbar vertebral end plate contains immunoreactive nerves, as shown in the studies of sheep and humans [33,34], it has been reported that an increased number of tumor necrosis factor–immunoreactive spinal nerve cells and fibers are present in end plates demonstrating ESC [35]. Therefore, pain may originate from damaged end plates in patients with ESC.

Another possibility is that ESC is a proxy for discogenic pain, as ESC is most often seen in association with DD [12,36] and tumor necrosis factor–immunoreactive nerves have also been reported in DD [37]. Both the present results and a previous study indicate that the association between ESC and LBP appears to be stronger than that between DD and LBP [29]. Regarding the association between the level of ESC and LBP, Kuisma et al. [32] reported that both Modic type I and II lesions at L5–S1, but not at upper levels, are associated with LBP. However, the present study showed that the OR of LBP with Modic type I and II

lesions at L3–L4 was higher than that at L5–S1. We speculate that the association of LBP symptoms with the L3–L4 level might be because of mechanical factors and alignment of the whole spine, but the pathophysiology of this phenomenon needs further investigation.

In the present study, SN alone was not significantly associated with LBP. In addition, most SNs were combined with DD. In fact, SN was not itself a risk factor for back pain but was an indicator of DD in the previous reports [13,14,28]. Furthermore, SN occurs when the cartilaginous end plate of the vertebral body has been disrupted [38]. Such a disruption can be produced by an intrinsic abnormality of the end plate itself or by alterations in the subchondral bone of the vertebral body [39]. Therefore, LBP might be a multifactorial condition arising from a combination of DD, ESC, and SN in the lumbar region.

Study limitations

The present study has several limitations. First, it is a cross-sectional study, so the transition from DD, ESC, and SN cannot be clarified. Second, more than 1,000 participants included in the present study may not represent the general population, as they were recruited from only 2 areas. To confirm whether the participants of the Wakayama Spine Study are representatives of the Japanese population, we compared anthropometric measurements and frequencies of smoking and alcohol drinking between the study participants and the general Japanese population. No significant differences in BMI were observed (men: 23.7 and 24.0, $p=.33$ and women: 23.1 and 23.5, $p=.07$). Furthermore, the proportions of current smokers and drinkers (those who regularly smoked or drank more than one drink per month) among men and that of current drinkers among women were significantly higher in the general Japanese population than in the study population, and no significant difference in current smokers was observed among women (men smokers, 32.6% in the Japanese population and 25.2% among study participants, $p=.015$; women smokers, 4.9% in the Japanese population and 4.1% among study participants, $p=.50$; men drinkers, 73.9% in the Japanese population and 56.8% among study participants, $p<.0001$; and women drinkers, 28.1% in the Japanese population and 18.8% among study participants, $p<.0001$). These results suggest the likelihood that participants in this study had healthier lifestyles than the general Japanese population [40]. This ''healthy'' selection bias should be taken into consideration when generalizing the results obtained from the Wakayama Spine Study. Third, distinction of the type of Modic changes was not possible because owing to cost and time limitations, only T2-weighted images were obtained. However, Ohtori et al. [35] reported that both Modic type I and II changes were significantly associated with inflammation induced by tumor necrosis factor. Furthermore, the prevalence of Modic Type I was lower than that of Type II in a systematic review [41]. Therefore, we propose that high-intensity ESC on T2-weighted images is informative in the assessment of LBP. Fourth, the definition of LBP is different among many studies [24,32], and the result of association between LBP and radiographic change might be changed depending on the definition. We decided that the definition of LBP was ''LBP on most days during the past month, in addition to now'' from the previous reports [20–24].

Finally, the radiographic changes of DD, ESC, and SN in the lumbar region might not be strongly correlated with LBP. Low back pain can be caused by multiple factors including osteoporosis, back muscle strain, and psychosocial problems; thus, we can explain only a portion of the associated factors of LBP from MRI findings. Future investigations should include continued follow-up surveys of psychosocial and other factors.

Conclusions

We first investigated combinations of the radiographic changes DD, ESC, and SN in the lumbar region and at each intervertebral level and their association with LBP in a large population of individuals ranging in age from 21 to 97 years old. Our data suggest that DD alone is not an associated factor for LBP. By contrast, the combination of DD, ESC, and SN at L1–L2, L4–L5, and L5–S1 was significantly associated with LBP. Furthermore, the combination of DD and ESC at L3–L4, L4–L5, and L5–S1 was also significantly associated with LBP. Low back pain is caused by multiple factors beyond the scope of MRI findings. However, this study clarified that DD alone was not associated with LBP, whereas, by contrast, the combination of DD, ESC, and/or SN was associated with LBP. Although they may not be immediately applicable to clinical practice, these findings contribute to the progress of LBP research. Further investigations along with continued follow-up surveys will continue to elucidate the causes of LBP.

Acknowledgements

The authors wish to thank Mrs Tomoko Takijiri and other members of the Public Office in Hidakagawa Town and Mrs Tamako Tsutsumi, Mrs Kanami Maeda, and other members of the Public Office in Taiji Town for their assistance in the location and scheduling of participants for examinations. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

References

[1] Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J 2008;8:8–20.

[2] Deyo RA, Tsui-Wu YJ. Descriptive epidemiology of low-back pain and its related medical care in the United States. Spine 1987;12:264–8.

[3] Andersson GB. Epidemiological features of chronic low-back pain. Lancet 1999;354:581–5.

[4] Luoma K, Riihimaki H, Luukkonen R, Raininko R, Viikari-Juntura E, Lamminen A. Low back pain in relation to lumbar disc degeneration. Spine 2000;25:487–92.
[5] Chou D, Samartzis D, Bellabarba C, Patel A, Luk KD, Kisser JM, et al. Degenerative magnetic resonance imaging changes in patients with chronic low back pain: a systematic review. Spine 2011;36:S43–53.
[6] Takatalo J, Karppinen J, Niinimaki J, Taimela S, Nayha S, Mutanen P, et al. Does lumbar disc degeneration on magnetic resonance imaging associate with low back symptom severity in young Finnish adults? Spine 2011;36:2180–9.
[7] Carragee E, Alamin T, Cheng I, Franklin T, van den Haak E, Hurwitz E. Are first-time episodes of serious LBP associated with new MRI findings? Spine J 2006;6:624–35.
[8] Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects: a prospective investigation. J Bone Joint Surg Am 1990;72:403e8.
[9] Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994;331:69–73.
[10] Weishaupt D, Zanetti M, Hodler J, Min K, Fuchs B, Pfirrmann CW, et al. Painful lumbar disc derangement: relevance of endplate abnormalities at MR imaging. Radiology 2001;218:420–7.
[11] Mitra D, Cassar-Pullicino VN, McCall IW. Longitudinal study of vertebral type-1 end-plate changes on MR of the lumbar spine. Eur Radiol 2004;24:1574–81.
[12] Toyone T, Takahashi K, Kitahara H, Yamagata M, Murakami M, Moriya H. Vertebral bone-marrow changes in degenerative lumbar disc disease. An MRI study of 74 patients with low back pain. J. Bone Joint Surg Br 1994;76:757–64.
[13] Kjaer P, Leboeuf-Yde C, Korsholm L, Sorensen JS, Bendix T. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. Spine 2005;30:1173–80.
[14] Williams FMK, Manek NJ, Sambrook PN, Spector TD, Macgregor AJ. Schmorl's nodes: common, highly heritable, and related to lumbar disc disease. Arthritis Rheum 2007;57:855–60.
[15] Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, et al. Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study. Osteoarthritis Cartilage 2014;22:104–10.
[16] Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012;20:1103–8.
[17] Nagata K, Yoshimura N, Muraki S, Hashizume H, Ishimoto Y, Yamada H, et al. Prevalence of cervical cord compression and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Spine 2012;37:1892–8.
[18] Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability (ROAD) study. Int J Epidemiol 2010;39: 988–95.
[19] Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 2009;27:620–8.
[20] Muraki S, Oka H, Akune T, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of radiographic lumbar spondylosis and its association with low back pain in elderly subjects of population-based cohorts: the ROAD study. Ann Rheum Dis 2009;68:1401–6.
[21] Muraki S, Akune T, Oka H, En-Yo Y, Yoshida M, Saika A, et al. Impact of knee and low back pain on health-related quality of life in Japanese women: the Research on Osteoarthritis against Disability (ROAD). Mod Rheumatol 2010;20:444–51.
[22] Muraki S, Akune T, Oka H, En-Yo Y, Yoshida M, Saika A, et al. Health-related quality of life in subjects with low back pain and knee pain in a population-based cohort study of Japanese men: the ROAD study. Spine 2011;36:1312–9.
[23] Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, et al. Incidence and risk factors for radiographic lumbar spondylosis and lower back pain in Japanese men and women: the ROAD study. Osteoarthritis Cartilage 2012;20:712–8.
[24] Dionne CE, Dunn KM, Croft PR, Nachemson AL, Bunchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine 2008;33:95–103.
[25] Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine 2001;26:1873–8.
[26] Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disc disease: assessment of changes in vertebral body marrow with MR imaging. Radiology 1988;166:193–9.
[27] Modic MT, Masaryk TJ, Ross JS, Carter JR. Imaging of degenerative disc disease. Radiology 1988;168:177–86.
[28] Mok FP, Samartzis D, Karppinen J, Luk KD, Fong DY, Cheung KM. ISSLS prize winner: prevalence, determinants, and association of Schmorl's nodes of the lumbar spine with disc degeneration: a population-based study of 2449 individuals. Spine 2010;35:1944–52.
[29] Kjaer P, Korsholm L, Bendix T, Sorensen JS, Leboeuf-Yde C. Modic changes and their associations with clinical finding. Eur Spine J 2006;15:1312–9.
[30] Battié MC, Videman T, Kaprio J, Gibbons LE, Gill K, Manninen H, et al. The Twin Spine Study: contributions to a changing view of disc degeneration. Spine J 2009;9:47–59.
[31] Cheung KM, Karppinen J, Chan D, Ho DW, Song YQ, Sham P, et al. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine 2009;34:934–40.
[32] Kuisma M, Karppinen J, Niinimaki J, Ojala R, Haapea M, Heliovaara M, et al. Modic changes in endplates of lumbar vertebral bodies: prevalence and association with low back and sciatic pain among middle-aged male workers. Spine 2007;32:1116–22.
[33] Rajasekaran S, Babu JN, Arun R, Armstrong BR, Shetty AP, Murugan S. ISSLS prize winner: a study of diffusion in human lumbar discs: a serial magnetic resonance imaging study documenting the influence of the endplate on diffusion in normal and degenerate discs. Spine 2004;29:2654–67.
[34] Brown MF, Hukkanen MV, McCarthy ID, Redfern DR, Batten JJ, Crock HV, et al. Sensory and sympathetic innervation of the vertebral endplate in patients with degenerative disc disease. J Bone Joint Surg Br 1997;79:147–53.
[35] Ohtori S, Inoue G, Ito T, Koshi T, Ozawa T, Doya H, et al. Tumor necrosis factor-immunoreactive cells and PGP 9.5-immunoreactive nerve fibers in vertebral endplates of patients with discogenic low back pain and Modic type 1 or type 2 changes on MRI. Spine 2006;31:1026–31.
[36] Albert HB, Manniche C. Modic changes following lumbar disc herniation. Eur Spine J 2007;16:977–82.
[37] Freemont AJ, Peacock TE, Goupille P, Hoyland JA, O'Brien J, Jayson MI. Nerve ingrowth into diseased intervertebral disc in chronic back pain. Lancet 1997;350:178–81.
[38] Schmorl G. Uber Knorpelknotchen an den Wirbelbandscheiben (About cartilaginous node of the intervertebral disc). Fortschr Rontgenstr 1928;38:265–79.
[39] Wagner AL, Murtagh FR, Arrington JA, Stallworth D. Relationship of Schmorl's nodes to vertebral body endplate fractures and acute endplate disc extrusions. AJNR Am J Neuroradiol 2000;21:276–81.
[40] Ministry of Health, Labour and Welfare. The outline of the results of National Livelihood Survey 2007. Available at: http://www.mhlw.go.jp/toukei/list/20-19-1.html. Accessed December 24, 2014.
[41] Jensen TS, Karppinen J, Sorensen JS, Niinimaki J, Leboeuf-Yde C. Vertebral endplate signal changes (Modic change): a systematic literature review of prevalence and association with non-specific low back pain. Eur Spine J 2008;17:1407–22.

Clinical Interventions in Aging

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Efficacy of a trunk orthosis with joints providing resistive force on low back load during level walking in elderly persons

This article was published in the following Dove Press journal:
Clinical Interventions in Aging
10 November 2016
Number of times this article has been viewed

Junji Katsuhira[1,2]
Ko Matsudaira[2]
Hiroyuki Oka[2]
Shinno Iijima[3]
Akihiro Ito[3]
Tadashi Yasui[4]
Arito Yozu[5]

[1]Faculty of Medical Technology, Department of Prosthetics and Orthotics and Assistive Technology, Niigata University of Health and Welfare, Niigata, [2]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo, [3]International University of Health and Welfare Hospital, Tochigi, [4]Kawamura-Gishi Company, Ltd., Osaka, [5]Department of Rehabilitation Medicine, The University of Tokyo Hospital, Tokyo, Japan

Purpose: The effects of lumbosacral and spinal orthoses on low back pain and gait are not exactly clear. We previously developed a trunk orthosis with joints providing resistive force on low back load to decrease such load, and confirmed its positive effects during level walking in healthy young adults. Therefore, we aimed to determine the efficacy of this trunk orthosis during level walking in healthy elderly subjects.

Methods: Fifteen community-dwelling elderly subjects performed level walking at a self-selected speed without an orthosis, with our orthosis, and with a lumbosacral orthosis. Kinematic and kinetic data were recorded using a three-dimensional motion analysis system, and erector spinae activity was recorded by electromyography.

Results: When comparing the three conditions, our orthosis showed the following effects: it decreased the peak extension moment, increased the peak flexion moment, decreased the lateral bending angle, increased the peak thoracic extension angle, and had significantly lower erector spinae activity and significantly larger peak pelvic forward tilt angles.

Conclusion: Our orthosis with joints providing resistive force decreased low back load and modified trunk and pelvis alignments during level walking in healthy elderly people.

Keywords: biomechanics, orthosis, gait, low back pain, joint moment, motion analysis

Background

The lifetime prevalence of low back pain (LBP) is high; 70% of adults have had LBP at some time.[1] Moreover, the number of patients with LBP in developed countries is increasing in line with the proportion of elderly.[2,3]

Conservative and postoperative treatments for LBP include the use of class 1 medical devices such as a lumbosacral orthosis (LSO).[4] Cholewicki et al observed that one of the causes of LBP is excessive erector spinae muscle activity, which could be reduced with an LSO.[5] Any decrease in the compressive force exerted on the vertebral body by reducing such activity with an LSO would benefit those with osteoporosis and vertebral compression fracture, conditions to which elderly people are vulnerable.

However, a review of data held in the Cochrane Database found no evidence for the efficacy of lumbar supports alone in preventing and treating LBP.[6] Although Pfeifer et al reported that their newly designed spinal orthosis had several positive effects on muscle strength, body balance, kyphosis angle, and vital capacity in elderly patients with osteoporosis,[7] to our knowledge, no previous studies have reported any significant effects of wearing a trunk orthosis to specifically decrease erector spinae activity and low back load in elderly people. To address this issue, we previously designed a trunk orthosis to improve trunk and pelvic stability and alignment by means of resistive force

Correspondence: Junji Katsuhira
Faculty of Medical Technology, Department of Prosthetics and Orthotics and Assistive Technology, Niigata University of Health and Welfare, 1398 Shimamichou, Kita-ku, Niigata 950-3198, Japan
Tel +81 025 257 4621
Fax +81 025 257 4621
Email katsuhira@nuhw.ac.jp

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/CIA.S108033

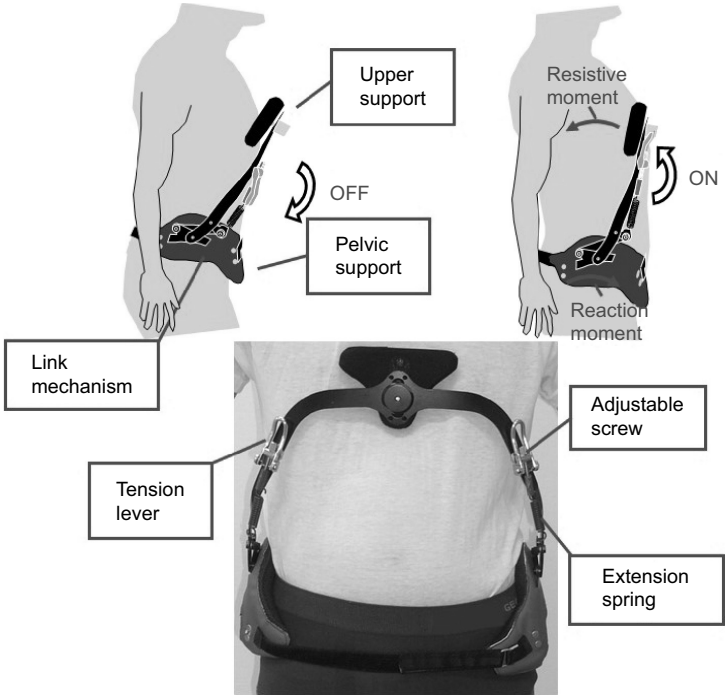


Figure 1 Our trunk orthosis with joints providing resistive force.

provided by joints with springs (Figure 1).[8] This orthosis with joints providing resistive force (ORF) creates resistive force to produce a resistive moment that rotates the trunk backward and pelvis forward. In our previous study, we reported the effect of this ORF in modifying trunk alignment and decreasing activity of the erector spinae during static standing in elderly subjects.[9]

A previous study reported that an LSO was effective for decreasing erector spinae activity in an unstable sitting position, where adjustment was needed to balance the upper body,[5] while another study showed no positive effects of an LSO on decreasing low back load by decreasing low back muscle activity and increasing intra-abdominal pressure.[10] To date, there have been no studies on the efficacy of a typical LSO or spinal orthosis for reducing low back load during level walking, where the demands for adjusting balance are high. We previously reported that our ORF prototype increased superficial abdominal muscle activity and decreased erector spinae activity during level walking in healthy young adults,[11] and here we sought to explore whether our findings can be extended to the trunk muscular activities and the low back joint moment (LBM) during level walking in healthy elderly subjects.

This biomechanics study examined the effects of the ORF on the gait of healthy elderly people during level walking and compared the effects with those obtained without an orthosis and with an LSO. We hypothesized that the ORF and LSO would both effectively decrease low back load measured by joint moment and trunk muscular activities during level walking, but that the ORF, with its biomechanical function of decreasing low back load, would show a superior effect. This follow-up to our previous study[9] was performed to confirm the hypothesis that ORF might decrease low back load not only in static standing but also in level walking, using a new biomechanical method.

Materials and methods

Subjects

From 31 community-dwelling elderly subjects who were candidates for this study, 15 were enrolled (all males; mean age, 67.7±6.1 years; mean height, 162.4±5.7 cm; mean weight, 62.3±7.8 kg) after excluding those with neurological disease, pain, history of an orthopedic surgical procedure, history of orthopedic treatment within the past 5 years, and history of LBP within the past 1 year. The study subjects were the same as those of our previous study.[9] The study was approved by the ethics committee of the International University of Health and Welfare (11–191). All the subjects provided written informed consent to participate.

Features of the ORF

The ORF is shown in Figure 1 and its features are described in our previous study.[8] Briefly, pelvic and upper supports

are positioned on the ileum and sternum, respectively. Stainless steel joints, connected to the upper support with a nylon pad and to the pelvic support, produce resistive force through the use of extension springs. A link mechanism translates the spring-generated tension into a resistive moment on the chest and a reaction moment on the posterior pelvis. The ORF weighs 0.99 kg and has a range of motion of 40 degrees. The upper support initially inclines backward to exert resistive force on the chest. The ORF has a release mechanism that releases the resistive force by pulling tension levers downward. Adjustment screws control the magnitude of the spring-generated resistive force. The ORF is currently an investigational product that has not been approved by the Food and Drug Administration or by a corresponding national agency for the indication described herein.

Experimental conditions

The subjects walked 10 m on a level surface at a self-selected speed in a laboratory setting under three conditions: without an orthosis, with the ORF, and with an LSO (Damen Corset, Pacific Supply, Osaka, Japan). The Damen corset was selected as it is frequently prescribed for patients with LBP. After completing three walking trials without any orthosis, they completed three trials in the two orthosis conditions in a randomized order. A minimum rest interval of 5 minutes was set between the conditions.

The subjects were allowed 5 minutes to accustom themselves to wearing the ORF and the LSO. They then practiced level walking in the laboratory before measurements were taken. Resistive force on the chest provided by the joints was measured in real time using a strain gauge (Kyowa, Tokyo, Japan) and the force data were transferred to a laptop computer by Bluetooth (Figure 2) and the force was set to a magnitude of 20–25 N during static standing. The pressure between the corset and abdomen was set to 10 mmHg in all measurement conditions.[12]

Experimental setup

Gait was recorded with a three-dimensional motion capture system (Vicon 612, Vicon, Oxford, UK) consisting of six force plates (four from AMTI, Watertown, MA, USA; and two from Kistler, Winterthur, Switzerland) and 12 infrared (IR) cameras with a sampling rate of 120 Hz. Referring to a study by Seay et al,[13] 41 IR-reflective markers (diameter, 14 mm) were attached to each subject's body. Additionally, three markers were attached over a strain gauge and on bilateral joints of the ORF. To measure muscle activity during level walking, electromyograms (EMGs) were obtained (Biometrics, Newport, UK) at a sampling rate of 1,080 Hz for bilateral erector spinae (2 cm to the side between L4–L5 vertebrae).[14] Maximum voluntary contraction was measured while one physical therapist manually applied resistant force to the midpoints of the bilateral scapulae, with the subject lying in the prone position on a bed.

Data analysis

During acquisition, we performed full-wave rectification feeding into a band pass filter (20–420 Hz) to decrease noise and used Visual 3D analytical software (C-motion, Germantown, MD, USA). The obtained EMGs were normalized using maximal voluntary contraction during isometric

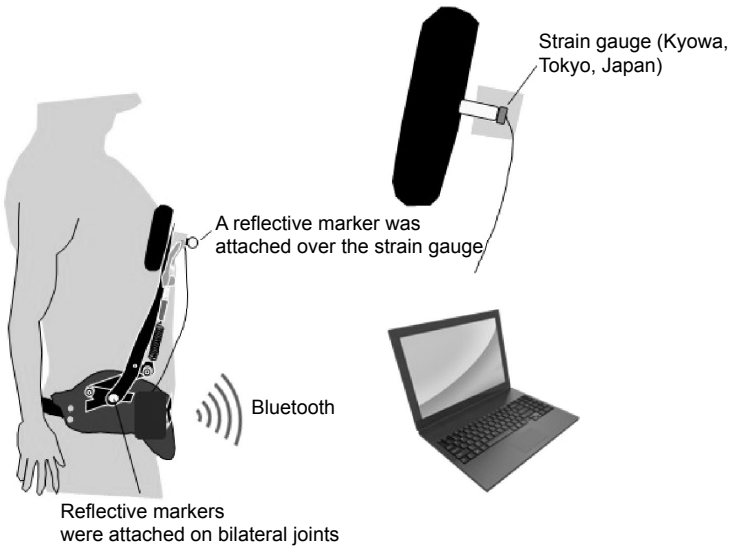


Figure 2 Trunk orthosis sensors and data transfer.

contraction (as a percent), and the root mean square was calculated for a 50 ms window. Subjects performed isometric contractions in prone against gravity with maximum resistance applied by the experimenter to obtain maximal voluntary contraction of the erector spinae.[15]

Visual 3D was used to perform kinetic and kinematic data analysis. The obtained physical coordinates and ground reaction force data were low-pass filtered with a second-order recursive Butterworth filter, with a cutoff frequency of 6 and 18 Hz, respectively, according to Winter's technique.[16] The link segment model consisted of 13 segments: head, trunk, pelvis, bilateral upper arms, forearms, thighs, shanks, and feet. Briefly, the low back extension and flexion moments were calculated using the ground reaction force data obtained from the force plates, the reaction force on the chest obtained from the strain gauge, and the coordinates of the IR-reflective markers on the bodies of the subjects and an ORF. Moment exerted by an ORF was calculated by multiplying the force measured by a strain gate and moment arm from joint of an ORF to the force. In our previous study analyzing the ORF effect during static standing,[9] we were not able to calculate LBM. The novelty of the present study lies in applying a new technique to calculate LBM during level walking while wearing an ORF. The moment was subtracted from the LBM calculated by using the ground reaction force data and the coordinates of the IR-reflective markers on the bodies of the subjects because the moment created by ORF joints equally gives forward rotation moment on the pelvis because of action–reaction law. In the analysis, segments were regarded as rigid and the joint moments were calculated using a link segment model in which segments were connected together at nodal points. To compute the joint moments, joints coordinate data were added to the ground reaction force data, in which the position of the center of mass, the weight portion, and the moment of inertia of each segment were used as parameters. The measurement data reported by Winter[16] were used as the body parameters necessary for calculating the LBM. Three-dimensional trunk and pelvic angles were calculated by the Eulerian method using coordinate systems as determined by markers on the trunk and pelvis, respectively. In this study, we defined LBM and bilateral erector spinae activities among these parameters as low back load because LBM and ES activities indicate the rotation force around the low back joint and the action of the low back muscles, respectively.

Statistical analysis

Peak values of kinetic and kinematic data acquired during level walking were extracted from the phase between mid-stance and terminal stance (MTS), and the pre-swing phase in one gait cycle of the right limb because it was not possible to calculate the LBM when a subject's posterior foot did not contact the force plates (Figure 3). Integral values of EMGs were calculated during stance. Mean

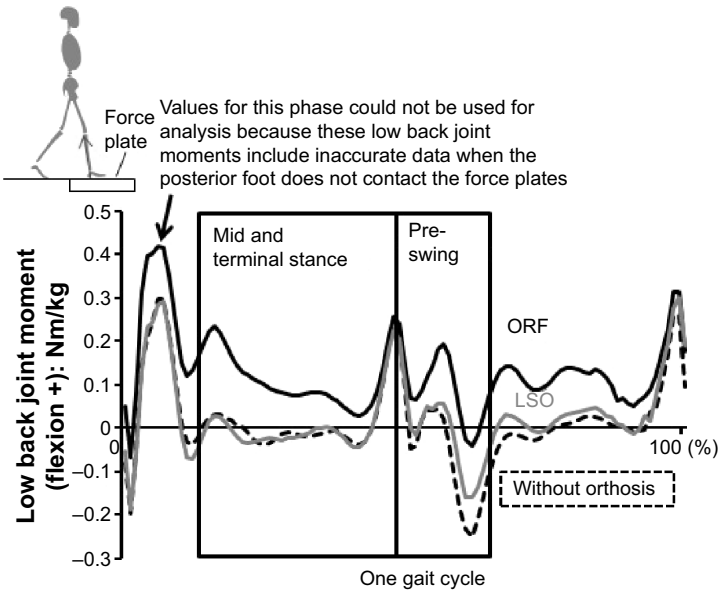


Figure 3 Average low back joint moments without orthosis (dotted line), with trunk orthosis with joints providing resistive force (ORF; solid black line), and with lumbosacral orthosis (LSO; solid gray line).

peak values of LBM, three-dimensional trunk and pelvic angles, and integral EMGs were calculated from the data obtained in three trials and were selected as representative values for analysis. Peak LBMs were normalized by subject weight (kg). Comparison was performed using repeated measures analysis of variance (ANOVA) after confirming non-deviation of the data and performing the Kolmogorov–Smirnov test, and variables showing a significant difference were subjected to multiple comparisons with Bonferroni correction. Significance was established at $P<0.05$. Statistical analysis was performed using SPSS 20 (SPSS Inc., Chicago, IL, USA).

Results

As shown in Tables 1 and 2, all kinematic and kinetic parameters differed significantly among the three conditions.

Low back load

Averaged LBM waveforms for all subjects in the three conditions are shown in Figure 3. The peak flexion moment was observed at the beginning of pre-swing and the peak extension moment at the end of pre-swing. Averaged waveforms of LBM and resistive force on the chest while wearing the ORF are shown in Figure 4. Force exerted on the chest was 34–37 N, which occurred during MTS and pre-swing. ANOVA indicated significant differences in the parameters showing low back load. Degrees of freedom for all data were 2 and 28. Significant main effects of an orthosis were observed for the peak flexion and extension moments not only in pre-swing but also in MTS (Table 1). Peak extension moments in MTS and pre-swing were significantly smaller with the ORF than in the other two conditions. The peak flexion moment in MTS was significantly larger with the ORF than with the LSO; and in pre-swing was significantly larger with the ORF than in the other two conditions.

Averaged waveforms of bilateral erector spinae activity for all subjects in the three conditions are shown in Figure 5. Peak activity was observed at the beginning of pre-swing. Significant main effects of an orthosis were observed in the integral of bilateral erector spinae activity during stance (Table 1). This integral was significantly smaller with the ORF than in the other two conditions.

Pelvic and thoracic angles

Significant main effects of an orthosis were observed in peak pelvic forward tilt angles in MTS and pre-swing, and in the peak pelvic leftward rotation angle in pre-swing (Table 2). Peak pelvic forward tilt angles in MTS and pre-swing were significantly larger with the ORF than in the other two conditions. The peak pelvic leftward rotation angle was significantly smaller with the ORF than with the LSO.

Significant main effects of an orthosis were observed in peak thoracic extension angles in MTS and pre-swing, and in peak right and left lateral flexion angles in pre-swing. Peak extension angles in MTS and pre-swing were significantly larger with the ORF than in the other two conditions. The peak right lateral bending angle in pre-swing was significantly larger with the ORF than with the LSO. The peak left

Table 1 Comparison of parameters indicating low back load in three conditions during level walking in 15 healthy elderly subjects

Parameter of low back load	Mean (95% confidence interval)			*F*-value	*P*-value from post-hoc test		
	W/O orthosis	With ORF	With LSO		W/O orthosis-with ORF	W/O orthosis-with LSO	With ORF-with LSO
Low back joint moment (Nm/kg)							
Peak extension moment in mid and terminal stance	0.19 (0.101–0.279)	0.07 (–0.019–0.159)	0.22 (0.126–0.314)	15.081***	0.002	0.777	0.002
Peak extension moment in pre-swing	0.29 (0.218–0.362)	0.11 (0.016–0.204)	0.27 (0.165–0.375)	17.658***	0.001	1.000	$P<0.001$
Peak flexion moment in mid and terminal stance	0.37 (0.254–0.486)	0.49 (0.363–0.617)	0.37 (0.259–0.481)	5.354*	0.118	1.000	0.013
Peak flexion moment in pre-swing	0.16 (0.044–0.276)	0.31 (0.188–0.432)	0.19 (0.085–0.295)	28.484***	$P<0.001$	0.337	$P<0.001$
Erector spinae activity (%IEMG)							
Integral of right side muscle activity during stance	8.22 (5.656–10.784)	7.22 (4.578–9.862)	8.21 (5.447–10.973)	7.459**	0.006	1.000	0.001
Integral of left side muscle activity during stance	10.57 (6.898–14.242)	7.32 (4.773–9.867)	9.4 (6.149–12.651)	14.917***	0.001	0.113	0.015

Notes: *$P<0.05$, **$P<0.01$, and ***$P<0.001$.
Abbreviations: LSO, lumbosacral orthosis; ORF, orthosis with joints providing restrictive force; W/O, without; IEMG, integral electromyogram.

Table 2 Pelvic and thoracic angles in three conditions during level walking in 15 healthy elderly subjects

Angle (degrees)	Mean (95% confidence interval)			*F*-value	*P*-value from post-hoc test		
	W/O orthosis	With ORF	With LSO		W/O orthosis-with ORF	W/O orthosis-with LSO	With ORF-with LSO
Pelvic angle							
Peak forward tilt angle mid and terminal stance	5.86 (3.673–8.767)	9.15 (6.431–12.789)	6.57 (4.771–8.969)	7.701**	0.011	0.841	0.071
Peak forward tilt angle in pre-swing	4.95 (2.032–7.868)	7.85 (4.195–11.505)	5.15 (2.652–7.648)	5.575**	0.040	1.000	0.064
Peak leftward rotation angle in pre-swing	3.19 (1.584–4.796)	1.75 (0.222–3.278)	3.01 (1.570–4.450)	4.473*	0.157	1.000	0.041
Thoracic angle							
Peak extension angle mid and terminal stance	1.07 (−1.007–3.147)	2.93 (0.781–5.079)	1.55 (−0.588–3.688)	16.373***	0.001	0.369	0.003
Peak extension angle in pre-swing	0.75 (−1.498–2.998)	2.80 (0.552–5.048)	1.45 (−0.826–3.726)	16.033***	0.001	0.116	0.006
Peak flexion angle in mid and terminal stance	1.15 (−0.738–3.038)	−0.71 (−2.726–1.306)	0.47 (−1.573–2.513)	12.833***	$P<0.001$	0.288	0.028
Peak flexion angle in pre-swing	0.22 (−1.918–2.358)	−1.82 (−3.93–0.29)	−0.49 (−2.639–1.659)	16.350***	0.001	0.117	0.005
Peak right lateral bending angle in pre-swing	0.64 (−0.174–1.454)	1.39 (0.465–2.315)	0.60 (−0.281–1.481)	5.548**	0.071	1.000	0.023
Peak left lateral bending angle in pre-swing	0.44 (−0.479–1.359)	−0.38 (−1.277–0.517)	0.29 (−0.591–1.171)	6.471**	0.024	1.000	0.156

Notes: *$P<0.05$, **$P<0.01$, and ***$P<0.001$.
Abbreviations: LSO, lumbosacral orthosis; ORF, orthosis with joints providing restrictive force; W/O, without.

lateral bending angle in pre-swing was significantly smaller with the ORF than in the other two conditions.

Gait performance

Walking velocity was 1.09±0.10 m/sec without an orthosis, 1.13±0.12 m/sec with the ORF, and 1.18±0.11 m/sec with the LSO. Repeated measures ANOVA showed a significant main effect ($P=0.002$). Walking velocity with the LSO was significantly faster than without an orthosis ($P=0.017$) but was not significantly different from with the ORF.

Discussion

We hypothesized that both the LSO and ORF would effectively decrease low back load during level walking in healthy

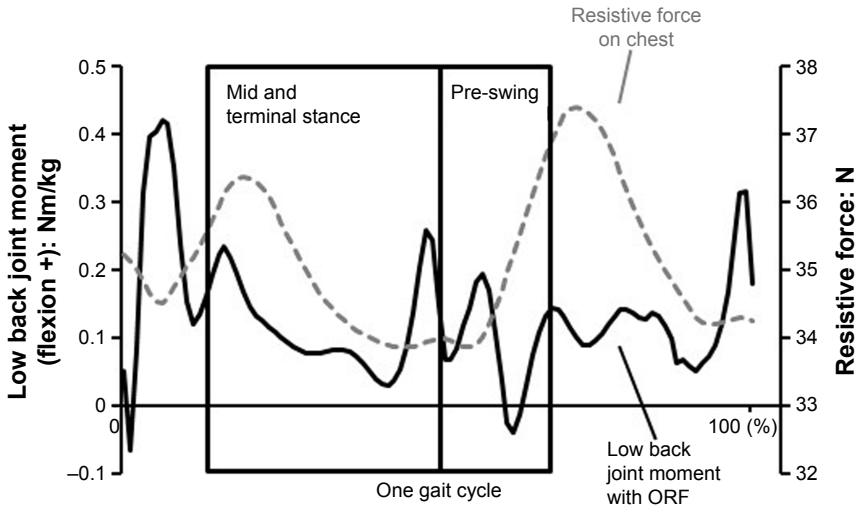


Figure 4 Average low back extension moment (solid line) and average resistive force on the chest (dotted line) with the trunk orthosis with joints providing resistive force (ORF).

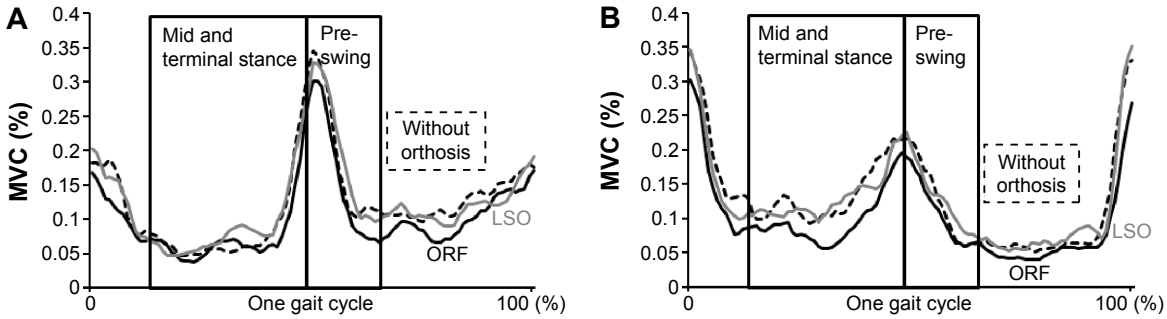


Figure 5 Right and left erector spinae activity (percent, MVC) without orthosis (dotted line), with trunk orthosis with joints providing resistive force (ORF; solid black line), and with lumbosacral orthosis (LSO; solid gray line).
Note: (**A**) Right erector spinae (**B**) Left erector spinae.
Abbreviation: MVC, maximum voluntary contraction.

elderly subjects but the ORF would have a superior effect. Our findings partially support this hypothesis. There were no significant differences between without an orthosis and with the LSO in the peak LBM and integral EMG of the erector spinae, or in the pelvic and thoracic angles. However, significant differences in all these parameters were observed with the ORF compared with no orthosis. Moreover, the low back extension moment and EMG of the erector spinae were significantly decreased with the ORF. Collectively, these results suggest that wearing the ORF during level walking should help to decrease the low back load in elderly people.

Several studies have suggested that using an LSO could stabilize the lumbosacral region but not decrease the low back load during static standing or lifting.[10] Our results indicate that the same applies during level walking also. Interestingly, the ORF not only decreased the activity of the erector spinae (which has higher fatigability in patients with LBP[17]) and the low back extension moment, but also increased the low back flexion moment created by the abdominal muscles.

The biomechanical function of the ORF can be explained by a simple model (Figure 6). The ORF can produce an extension moment for the upper trunk that decreases the low back extension moment. This extension moment also produces a resistive moment on the posterior pelvis; together, these moments could improve the malalignment commonly seen in elderly people (ie, lumbar kyphosis with pelvic backward tilt). Such malalignment increases both LBP and fall risks.[18,19] The extension moment served to extend the upper trunk, and the reaction moment acted as a forward rotation moment for the pelvis. In this way, the ORF increased the peak pelvic

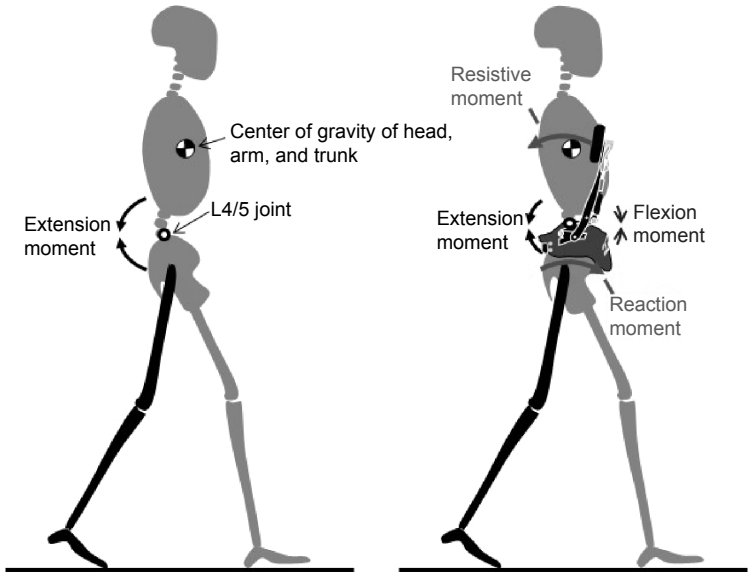


Figure 6 Biomechanical effects of the trunk orthosis with joints providing resistive force.

forward tilt angle and peak thoracic extension angle through MTS and pre-swing. This function, which can modify alignment in the elderly, would also be beneficial for decreasing the peak LBM, because modifications to the positioning of the pelvis and thorax are directly linked to a decrease in the lever arm from the low back joint to the head, arm, and trunk center of gravity (Figure 6).

A systematic review investigating the activation pattern of trunk muscles during walking in subjects with and without LBP indicated that those with LBP exhibit higher ES activity compared with asymptomatic subjects.[20] The magnitude of the decreases in erector spinae activity and the low back extension moment while wearing an ORF is relatively small. However, in daily life, people walk a large number of steps, so even though only a small decrease in low back load was evident with the ORF, the cumulative difference might have a distinct effect that can help treat and prevent LBP.[21] The spinal bones of elderly people, and particularly those of patients with osteoporosis, are more fragile than the middle-aged or young. Typical orthoses would not show a biomechanical effect of decreasing low back load during level walking because most correct only the abdominal region with compressive force or support the pelvis, thorax, and lower back with small resistive force. Only one eccentric type of orthosis, the rucksack-type orthosis, was found to decrease ES activity in elderly people during level walking in a previous study.[22] The rucksack-type orthosis controls the magnitude of force using weights to move the center of gravity of the upper body, thereby decreasing ES activity. This function is similar to that of the ORF, because both orthoses can control a relatively large magnitude of force applied to the upper trunk; however, the rucksack-type orthosis has the disadvantage of increasing low back compressive force in proportion to the amount of weight, which directly increases the gravitational force on the upper trunk. However, the ORF can apply resistive force horizontally to the chest, thereby avoiding an increase of low back compressive force.

An interesting feature of the ORF compared with other orthoses, including the rucksack-type, is its ability to increase the low back flexion moment produced by the abdominal muscles; the horizontally applied resistive force on the chest can not only decrease activity of the low back extension muscles using the support force, but also activate the abdominal muscles. The peak moment was larger with the ORF in MTS than with the LSO and in pre-swing than in the two other conditions. Rostami et al reported that using an LSO for 4 and 8 weeks decreased deep abdominal muscle thickness.[23] Although we did not examine deep abdominal muscle activity in this study, these muscles might contribute to increasing the low back flexion moment. This effect of the ORF would be beneficial for exercises aimed at improving abdominal muscle function during level walking in elderly subjects. As we previously found a positive training effect of the ORF on such function in hemiparetic patients,[8] a similar effect might be obtainable in elderly people.

A potentially negative feature of the ORF is that it limits pelvic rotation, which is one of the aspects of gait that increases step length and walking velocity.[24] Although walking velocity with the ORF was not faster than that without an orthosis, it was significantly faster with the LSO than without an orthosis. Thus, the stabilization of the pelvis and thorax achieved with an LSO might improve gait performance in elderly subjects, and the limited pelvic rotation with the ORF might decrease its overall positive effect for improving gait performance.

Limitations

This study has several limitations. Wearing the ORF during level walking served to decrease low back muscle activities and joint moment, and this might be effective in the prevention and treatment of LBP. However, we did not confirm the effects of long-term ORF use. Wearing LSOs and ORFs for long periods of time might adversely affect muscle control. Only low back extension and flexion moments in the three axial moments were calculated when wearing the ORF and therefore low back compressive force, which is a strong indicator of low back load, could not be calculated because the strain gauge measured the orthogonal resistive force on the chest produced by the ORF's joints. Finally, only healthy elderly male subjects participated in this study, and only a within-subject trial was conducted. Future studies should include healthy elderly female subjects and subjects who have LBP, and randomized controlled trials should be conducted.

Conclusion

In conclusion, we demonstrated that the ORF can decrease low back load during level walking in healthy elderly people by significantly decreasing LBM and increasing the abdominal moment. The ORF significantly modified malalignment commonly seen in elderly people. The ORF is a promising device for the prevention and treatment of LBP, and we plan to conduct randomized controlled trials with people who have LBP in the future.

Acknowledgments

This study was supported by the dissemination project of Clinical Research for Occupational Injuries and Illness from

the Ministry of Health, Labor and Welfare of Japan, and by a KAKENHI Grant (No 23700619) from the Japan Society for the Promotion of Science.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA*. 1992;268:760–765.
2. Buchbinder R, Blyth FM, March LM, Brooks P, Woolf AD, Hoy DG. Placing the global burden of low back pain in context. *Best Pract Res Clin Rheumatol*. 2013;27:575–589.
3. Dunn KM, Hestbaek L, Cassidy JD. Low back pain across the life course. *Best Pract Res Clin Rheumatol*. 2013;27:591–600.
4. Cholewicki J, Lee AS, Peter Reeves N, Morrisette DC. Comparison of trunk stiffness provided by different design characteristics of lumbosacral orthoses. *Clin Biomech* (*Bristol, Avon*). 2010;25:110–114.
5. Cholewicki J, Reeves NP, Everding VQ, Morrisette DC. Lumbosacral orthoses reduce trunk muscle activity in a postural control task. *J Biomech*. 2007;40:1731–1736.
6. van Duijvenbode I, Jellema P, van Poppel M, van Tulder MW. Lumbar supports for prevention and treatment of low back pain. *Cochrane Database Syst Rev*. 2008;(2):CD001823.
7. Pfeifer M, Kohlwey L, Begerow B, Minne HW. Effects of two newly developed spinal orthoses on trunk muscle strength, posture, and quality-of-life in women with postmenopausal osteoporosis: a randomized trial. *Am J Phys Med Rehabil*. 2011;90:805–815.
8. Katsuhira J, Miura N, Yasui T, Takane Mitomi, Yamamoto S. Efficacy of a newly designed trunk orthosis with joints providing resistive force in adults with post-stroke hemiparesis. *Prosthet Orthot Int*. 2016; 40:129–136.
9. Katsuhira J, Matsudaira K, Yasui T, Iijima S, Ito A. Efficacy of a trunk orthosis with joints providing resistive force on low back load in elderly persons during static standing. C*lin Interv Aging*. 2016;10: 1413–1420.
10. van Poppel MN, de Looze MP, Koes BW, Smid T, Bouter LM. Mechanisms of action of lumbar supports: a systematic review. *Spine*. 2000;25: 2103–2113.
11. Katsuhira J, Kikkawa K, Yasui T, et al. Invention and evaluation of trunk brace with joints providing resistance force: measurement of activities of trunk muscles during normal gait. *Bull Jpn Soc Prosthet Orthot Educ Res Dev*. 2011;27:112–119.
12. Lee YH, Chen CY. Lumbar vertebral angles and back muscle loading with belts. *Ind Health*. 1999;37:390–397.
13. Seay J, Selbie WS, Hamill J. In vivo lumbo-sacral forces and moments during constant speed running at different stride lengths. *J Sports Sci*. 2008;26:1519–1529.
14. De Foa JL, Forrest W, Biedermann H. Muscle fibre direction of longissimus, iliocostalis and multifidus: landmark-derived reference lines. *J Anat*. 1989;163:243–247.
15. Montgomery J, Hislop H, Connelly B. *Daniels and Worthingham's Muscle Testing: Techniques of Manual Examination*. 8th ed. Maryland Heights, MO: Saunders/Elsevier; 2007.
16. Winter DA. *Biomechanics and Motor Control of Human Movement*. 4th ed. Hoboken, NJ: John Wiley & Sons; 2009.
17. Sung PS, Lammers AR, Daniel P. Different parts of erector spinae muscle fatigability in subjects with and without low back pain. *Spine J*. 2009;9:115–120.
18. Glassman SDtsc, Bridwell K, Dimar JR, Horton W, Berven S, Schwab F. The impact of positive sagittal balance in adult spinal deformity. *Spine*. 2005;30:2024–2029.
19. Sinaki M, Brey RH, Hughes CA, Larson DR, Kaufman KR. Balance disorder and increased risk of falls in osteoporosis and kyphosis: significance of kyphotic posture and muscle strength. *Osteoporosis Int*. 2005;16:1004–1010.
20. Ghamkhar L, Kahlaee AH. Trunk muscles activation pattern during walking in subjects with and without chronic low back pain: a systematic review. *PM R*. 2015;7:519–526.
21. Adams MA, Dolan P. Spine biomechanics. *J Biomech*. 2005;38: 1972–1983.
22. Ishida H, Watanabe S, Yanagawa H, Kawasaki M, Kobayashi Y, Amano Y. Immediate effects of a rucksack type orthosis on the elderly with decreased lumbar lordosis during standing and walking. *Electromyogr Clin Neurophysiol*. 2007;48:53–61.
23. Rostami M, Noormohammadpour P, Sadeghian AH, Mansournia MA, Kordi R. The effect of lumbar support on the ultrasound measurements of trunk muscles: a single-blinded randomized controlled trial. *PM R*. 2014;6:302–308.
24. Saunders JB, Inman VT, Eberhart HD. The major determinants in normal and pathological gait. *J Bone Joint Surg Am*. 1953;35-A:543–558.

PLOS | ONE

RESEARCH ARTICLE

Diagnosing Discogenic Low Back Pain Associated with Degenerative Disc Disease Using a Medical Interview

Juichi Tonosu[1]*, Hirohiko Inanami[2,3], Hiroyuki Oka[4,5], Junji Katsuhira[6], Yuichi Takano[2,3], Hisashi Koga[2,3], Yohei Yuzawa[2,3], Ryutaro Shiboi[3], Yasushi Oshima[5], Satoshi Baba[3], Sakae Tanaka[5], Ko Matsudaira[4,5]

1 Department of Orthopedic Surgery, Kanto Rosai Hospital, Kanagawa, Japan, **2** Department of Orthopedic Surgery, Inanami Spine and Joint Hospital, Tokyo, Japan, **3** Department of Orthopedic Surgery, Iwai Orthopaedic Medical Hospital, Tokyo, Japan, **4** Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan, **5** Department of Orthopedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan, **6** Department of Prosthetics & Orthotics and Assistive Technology, Faculty of Medical Technology, Niigata University of Health and Welfare, Niigata, Japan

* juichitohnosu@yahoo.co.jp

Citation: Tonosu J, Inanami H, Oka H, Katsuhira J, Takano Y, Koga H, et al. (2016) Diagnosing Discogenic Low Back Pain Associated with Degenerative Disc Disease Using a Medical Interview. PLoS ONE 11(11): e0166031. doi:10.1371/journal.pone.0166031

Editor: Gayle E. Woloschak, Northwestern University Feinberg School of Medicine, UNITED STATES

Received: February 18, 2016

Accepted: October 21, 2016

Published: November 7, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Purposes

To evaluate the usefulness of our original five questions in a medical interview for diagnosing discogenic low back pain (LBP), and to establish a support tool for diagnosing discogenic LBP.

Materials and Methods

The degenerative disc disease (DDD) group (n = 42) comprised patients diagnosed with discogenic LBP associated with DDD, on the basis of magnetic resonance imaging findings and response to analgesic discography (discoblock). The control group (n = 30) comprised patients with LBP due to a reason other than DDD. We selected patients from those who had been diagnosed with lumbar spinal stenosis and had undergone decompression surgery without fusion. Of them, those whose postoperative LBP was significantly decreased were included in the control group. We asked patients in both groups whether they experienced LBP after sitting too long, while standing after sitting too long, squirming in a chair after sitting too long, while washing one's face, and in the standing position with flexion. We analyzed the usefulness of our five questions for diagnosing discogenic LBP, and performed receiver operating characteristic (ROC) curve analysis to develop a diagnostic support tool.

Results

There were no significant differences in baseline characteristics, except age, between the groups. There were significant differences between the groups for all five questions. In the

age-adjusted analyses, the odds ratios of LBP after sitting too long, while standing after sitting too long, squirming in a chair after sitting too long, while washing one's face, and in standing position with flexion were 10.5, 8.5, 4.0, 10.8, and 11.8, respectively. The integer scores were 11, 9, 4, 11, and 12, respectively, and the sum of the points of the five scores ranged from 0 to 47. Results of the ROC analysis were as follows: cut-off value, 31 points; area under the curve, 0.92302; sensitivity, 100%; and specificity, 71.4%.

Conclusions

All five questions were useful for diagnosing discogenic LBP. We established the scoring system as a support tool for diagnosing discogenic LBP.

Introduction

Low back pain (LBP) affects most adults at some point in their lives. In the last decade, LBP was continuously found to be the top leading cause of years lived with disability globally [1]. As in many industrialized countries, LBP is one of the most common health disabilities in Japan. In a population-based survey, the lifetime and 4-wk LBP prevalence were 83% and 36%, respectively [2].

It has been difficult to identify the cause of LBP. A specific source of pain can be identified in some cases of LBP; however, the source cannot be identified in other cases of LBP (i.e., nonspecific LBP) [3]. Magnetic resonance imaging (MRI) can identify underlying pathologies of LBP. However, the importance of MRI findings is unclear and controversial. Some reports have shown that disc degeneration was a source of LBP [4,5], whereas other reports have shown that there was no relationship between disc degeneration and LBP [6,7]. Reports have also shown that discogenic LBP associated with degenerative disc disease (DDD) is confirmed by the MRI findings and response to the injection of contrast media or local anesthesia into the disc [8–10]. Schwarzer et al. reported that 39% of cases of chronic LBP are discogenic, and the diagnosis is made by computed tomography after discography [11]. The technique of injecting local anesthesia into a disc is analgesic discography (discoblock). However, these procedures do not necessarily indicate high specificity findings of discogenic LBP [12, 13], and they are invasive and harmful to the disc [14, 15].

We hypothesized that discogenic LBP is one of the causes of LBP, and we sought to determine easier and less invasive means of diagnosing discogenic LBP. Few reports have specified that LBP in the sitting position can indicate discogenic LBP [16]. However, no report has found that LBP in standing position with flexion can indicate discogenic LBP. Based on our clinical experiences, we also hypothesized that discogenic LBP could be indicated in standing position with flexion and in sitting position. The purpose of the current study was to evaluate the usefulness of our original questions in a medical interview about LBP, which was intended to determine the characteristics of discogenic LBP, and establish a support tool for diagnosing discogenic LBP.

Materials and Methods

Subjects

In the current study, we defined the DDD group as those who suffered from discogenic LBP associated with DDD. The DDD group consisted of consecutive patients from November 2012

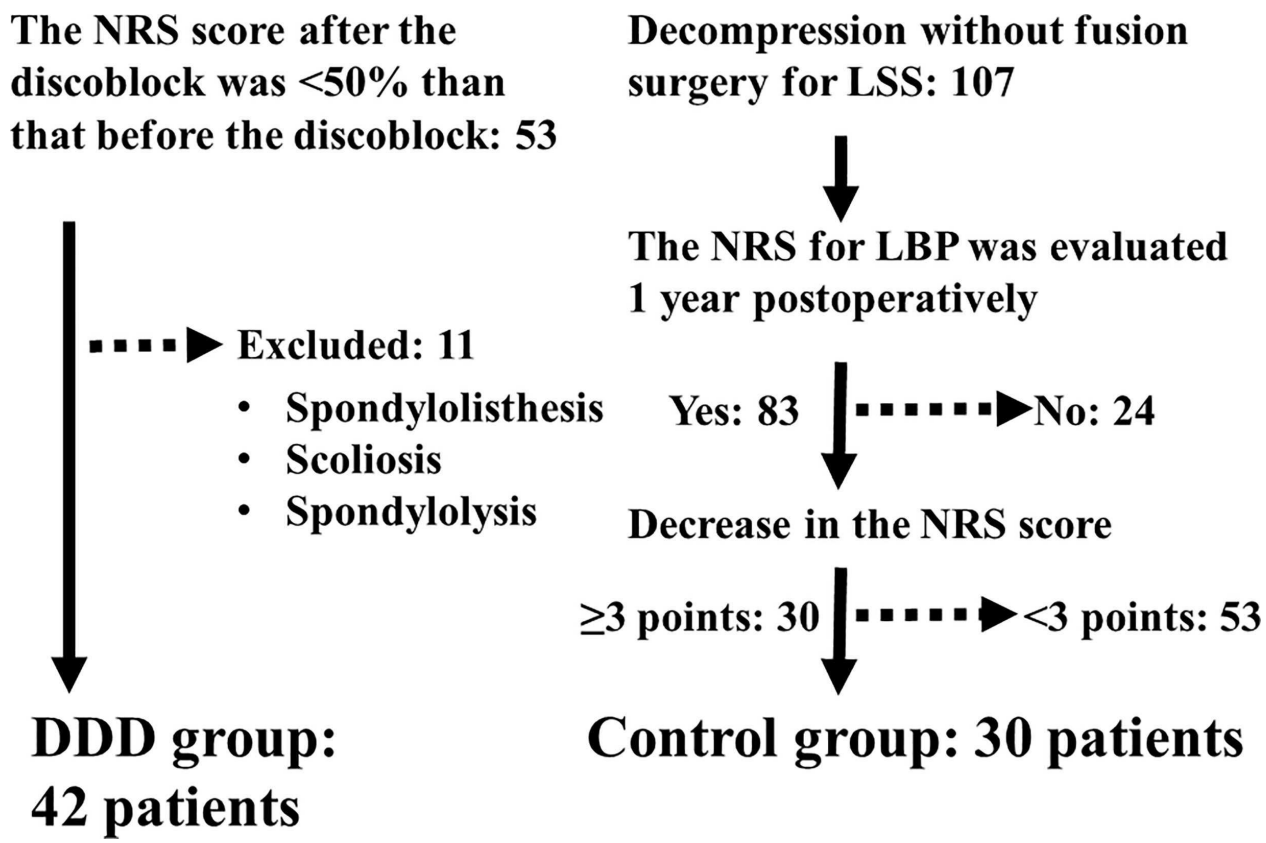


Fig 1. Study flow chart. Forty-two patients were included in the degenerative disc disease (DDD) group, and 30 patients were included in the control group. LBP, low back pain; NRS, numerical rating scale; LSS, lumbar spinal stenosis.

doi:10.1371/journal.pone.0166031.g001

to April 2014. Fifty-three patients had been diagnosed as having DDD by MRI and discoblock. Of 53 patients, we excluded 11 who had suffered from spondylolisthesis, scoliosis, and spondylolysis accompanied by DDD. Therefore, the DDD group consisted of 42 consecutive patients. We defined the control group as those who had suffered from LBP due to reasons other than DDD. We selected the control group from patients who had been diagnosed as having lumbar spinal stenosis (LSS) and had undergone posterior decompression surgery without fusion. The control group consisted of consecutive patients from April 2012 to December 2013. One hundred and seven patients had undergone decompression surgery for LSS. Of 107 patients, we could evaluate the numerical rating scale (NRS) score of 83 patients' LBP at 1 year postoperatively. Of 83 patients, 30 had a decrease in the postoperative NRS score of greater than or equal to 3 points compared with the preoperative NRS score. We included these 30 patients in the control group. In summary, 72 patients were included in this study, which consisted of 42 in the DDD group and 30 in the control group (Fig 1). We also collected patients' background information, including their age, sex, height, weight, and smoking habit, using a self-written questionnaire. We calculated the body mass index (BMI) from the data of height and weight. We also determined the NRS score of patients' LBP and assessed the Oswestry Disability Index (ODI) score [17] using a self-written questionnaire. We used a validated version of the Japanese ODI, which had been translated from the ODI version 2.0 [18]. The reliability and validity of this version was evaluated in their previous study, and was sufficient to use for outcome

studies in Japan. This study was approved by the medical/ethics review board of Iwai Orthopaedic Medical Hospital. Written informed consent was obtained from all the patients.

Definition of discogenic LBP

Although there is no consensus on how to diagnose discogenic LBP, we hypothesized and defined discogenic LBP as LBP that met the following criteria: MRI findings of a degenerated disc, and response to the discoblock into the disc suggestive of LBP. Although a discoblock may be harmful to the disc and it does not necessarily indicate high specificity findings of discogenic LBP, we hypothesized that a positive response to a discoblock indicates discogenic LBP.

At our hospital, well-trained medical clerks ask all patients about their medical history and symptoms during their first visit usually before they see the doctor. For patients who had suffered from lumbar diseases, medical clerks asked them about the following items in a medical interview: whether they had LBP from sitting too long, whether they had LBP while standing after sitting too long, whether they squirmed in a chair after sitting too long, whether they had LBP while washing their face, and whether they had LBP in standing position with flexion via a medical interview. Additionally, we precisely defined and evaluated the region of LBP and depicted it in a diagram for patients (i.e., pain localized between the costal margin and the inferior gluteal folds according to a previous report [19,20]). This was important for standardizing the study protocol for LBP.

We evaluated patients' physical findings, and radiography and MRI findings as needed. If the MRI showed only disc degeneration without disc herniation, spinal stenosis, or any other obvious findings, we suspected DDD. We evaluated disc degeneration on sagittal T2-weighted MRI based on Pfirrmann's grading system [21]. We considered grades $\geq$4 as disc degeneration. When we suspected discogenic LBP associated with DDD and the patient's disability was severe despite conservative therapies, we performed an additional examination (discoblock, a 1-mL injection of 1% lidocaine into the disc suggestive of LBP), and evaluated the degree of LBP both before and after the injection. We hypothesized discogenic LBP associated with DDD when the NRS score for LBP after the discoblock was $<$50% of that before the discoblock, although it was unclear whether the cutoff reduction rate of 50% was appropriate. When multilevel disc degeneration was shown on MRI, we performed the injections on a different day and evaluated the effectiveness of the injection for each disc.

Statistical methods

We compared the baseline characteristics of both groups, and analyzed the usefulness of the aforementioned five items of the medical interview for diagnosing discogenic LBP. For the age-adjusted analysis, we set the cut-off value at 65 years. Moreover, after identifying significant symptoms of discogenic LBP, we developed a support tool for diagnosing discogenic LBP.

Descriptive statistics were determined and presented as means and standard deviations or frequencies and percentages. Between-group differences in baseline characteristics were evaluated using the chi-square test for categorical variables, and Student's t-test for continuous variables. Age-adjusted odds ratios and 95% confidential intervals for each questionnaire were evaluated by logistic regression analyses. Moreover, we set the scores of each item as integral values from each age-adjusted odds ratio, and performed receiver operating characteristic (ROC) curve analysis to develop a support tool for diagnosing discogenic LBP. Finally, we calculated the area under the ROC curve (AUC), sensitivity, and specificity. An AUC of 1.0 indicated perfect discrimination, and in general, an AUC $\geq$0.7 was considered to indicate

acceptable discrimination. Statistical analysis was performed using the JMP 11.0 software program (SAS Institute, Cary, NC, USA). A p value <0.05 was considered significant.

Results

Patients' average age was 53.9 years in the DDD group and 71.1 years in the control group. The ratio of age $\geq$65 years was 16.7% in the DDD group and 30.1% in the control group. There was a significant difference in age between the DDD and control groups ($p < 0.0001$ and $p = 0.0002$, respectively). However, there were no significant differences in the other baseline characteristics such as sex, BMI, smoking habit, NRS score, and ODI score. There were significant differences between the groups for each item of the medical interview about LBP after sitting too long ($p < 0.0001$), LBP while standing after sitting too long ($p < 0.0001$), squirming in a chair after sitting too long ($p = 0.011$), LBP while washing one's face ($p < 0.0001$), and LBP in standing position with flexion ($p < 0.0001$) (Table 1).

In the age-adjusted analyses, the odds ratios of LBP after sitting too long, LBP while standing after sitting too long, squirming in a chair after sitting too long, LBP while washing one's face, and LBP in standing position with flexion were 10.5, 8.5, 4.0, 10.8, and 11.8, respectively. There were significant differences for all five items of the medical interview between the groups (Table 2). The integer scores were 11, 9, 4, 11, and 12, respectively, and the sum of the points of the five scores ranged from 0 to 47. Results of the ROC analysis were as follows: cut-off value, 31 points; AUC, 0.92302; sensitivity, 100%; and specificity, 71.4%.

Discussion

We examined five items of our medical interview regarding discogenic LBP. We hypothesized and defined discogenic LBP as a degenerated disc on MRI and response to a discoblock used for the disc suggestive of LBP.

Table 1. Baseline characteristics of patients in the DDD group and control group.

	DDD group (n = 42)	Control group (n = 30)	*P* value
• Age[1)]	53.4 ± 16.2	71.1 ± 9.4	<0.0001*
• Age ≥65 years[2)]	12 (16.7)	22 (30.1)	0.0002*
• Female sex[2)]	15 (35.7)	7 (23.3)	0.26
• BMI (kg/m^2)[1)]	24.2 ± 3.2	24.9 ± 2.7	0.36
• Smoking habit[2)]	28 (66.7)	16 (53.3)	0.25
• Current smoking habit[2)]	9 (21.4)	6 (20.0)	0.88
• NRS score[1)]	6.2 ± 2.3	6.2 ± 1.7	0.99
• ODI score[1)]	37.2 ± 13.3	37.8 ± 9.9	0.84
• LBP after sitting too long[2)]	35 (83.3)	9 (30.0)	<0.0001*
• LBP while standing after sitting too long[2)]	35 (83.3)	11 (36.7)	<0.0001*
• Squirming in a chair after sitting too long[2)]	33 (78.6)	15 (50.0)	0.011*
• LBP while washing one's face[2)]	31 (73.8)	6 (20.0)	<0.0001*
• LBP in standing position with flexion[2)]	22 (52.4)	2 (6.7)	<0.0001*

Data are shown as mean ± standard deviation or number of participants (%).

*: $P < 0.05$

[1)]: Student's t-test

[2)]: chi-square test.

DDD, degenerative disc disease; ODI, Oswestry Disability Index; NRS, numerical rating scale; LBP, low back pain.

doi:10.1371/journal.pone.0166031.t001

Table 2. Age-adjusted odds ratio, 95% confidential interval, and p-value for each item of the medical interview regarding LBP.

	Odds ratio	95% confidential interval	*P* value	Integer score
• LBP after sitting too long	10.5	3.3–39.4	<0.0001*	11
• LBP while standing after sitting too long	8.5	2.7–31.9	0.0002*	9
• Squirming in a chair after sitting too long	4.0	1.3–13.7	0.016*	4
• LBP while washing one's face	10.8	3.3–41.3	<0.0001*	11
• LBP in standing position with flexion	11.8	2.8–82.2	0.0004*	12

*: $P < 0.05$.

LBP, low back pain.

doi:10.1371/journal.pone.0166031.t002

No significant difference was observed in the baseline characteristics between the two groups, except with regard to age. The DDD group consisted of significantly younger patients and thus had a wider generation than the control group, although disc degeneration progresses with advancing age [5]. The reason for this may be caused by our definition of the control group.

We intended to define the control group as those who had LBP that was not mild for reasons other than discogenic LBP. The NRS score of 6.2 and ODI score of 37.8 in the control group indicate that the LBP was not mild, thus it was equivalent to that in the DDD group in terms of severity. However, it is difficult to confirm whether LBP was discogenic. Accordingly, we focused on patients who had been diagnosed as having LSS. Some patients with LSS have LBP, whereas other patients do not have LBP. The MRI scans of patients with LSS often show degenerated discs in addition to spinal stenosis, because disc degeneration progresses with advancing age [5] and LSS is often present in older people. However, the cause of LBP in patients with LSS is not necessarily derived from disc degeneration. It can often improve after decompression surgery without fusion [22], which indicates that compression of the dura itself can present as LBP in these patients. We hypothesized that improvement of LBP in the control group resulted from decompression of the dura itself and that it was not associated with disc degeneration, although the lack of a negative response to the discoblock in the control group was not evaluated in this study. We excluded patients who had undergone decompression combined with fusion surgery because the improvement of LBP by fusion surgery implies the co-existence of discogenic LBP. We defined a clinically meaningful improvement in LBP postoperatively as a reduction in the NRS score of $\geq$3 points, according to a study that reported that the cut-off value for the decrease in the NRS score is 2.5 for successful lumbar surgery [23]. Therefore, we considered the cause of LBP in the control group as LBP caused by LSS itself and that it was not associated with DDD. As LSS is often present in more aged people, there was a significant difference in age between the two groups. Therefore, we performed age-adjusted analyses. We defined the cut-off age as 65 years. Since the prevalence of LBP increases with advancing age [2], we did not underestimate the DDD group, which was younger than the control group.

There were significant differences between the groups in all the five items of our medical interview for analyses adjusted and not adjusted by age. The odds ratios of only the items of LBP after sitting too long, LBP while washing one's face, and LBP in standing position with flexion were >10. One reason for this may be because the results seemed to be associated with a higher intradiscal pressure in sitting and standing positions with flexion [24–28]. LBP while standing after sitting too long was significantly associated with discogenic LBP. The motion of standing after sitting too long includes changing the status of both the disc and facet, which is often degenerated in patients with LSS; however, our result indicated that LBP was discogenic. Therefore, the result may have been derived from the increasing intradiscal pressure. The items

LBP while washing one's face and LBP in standing position with flexion may be similar; however, we had intended to differentiate mild flexion of the trunk, such as the posture for washing one's face from full flexion of the trunk. In terms of the results, both items can significantly indicate discogenic LBP. The results may have also been derived from increasing intradiscal pressure.

Another reason why the results seemed to be similar to previous reports may be because being in one position too long advances disc degeneration [29, 30]. We assumed that being in one position too long caused discogenic LBP. LBP after sitting too long was also significantly associated with discogenic LBP. The high odds ratio of 10.5 for sitting too long in the current study may have been derived from being in the same position too long rather than from increasing intradiscal pressure itself. The symptom of squirming in a chair after sitting too long was also significantly associated with discogenic LBP. There has been no report about the relationship between LBP and squirming in a chair after sitting too long. This may have also been derived from being in the same position too long.

The AUC of 0.92302 was considered to indicate acceptable discrimination. ROC analysis indicated the cut-off value of 31 in our scoring system for diagnosing discogenic LBP, which meant that >31 points of the total 47 points are needed to diagnose discogenic LBP. Considering each item of the medical interview, we can diagnose discogenic LBP in all cases if four or five of the five items are positive, and in some cases if three of five items are positive.

There were some limitations to the current study. First, answers to the medical interview were not necessarily accurate. The subjective evaluation of the patients' own LBP can vary, i.e., positive or negative responses can differ depending on the medical interview. However, the results of the current study showed the high odds ratios of the five items, so we considered the results acceptable. Second, we did not evaluate the effectiveness of a discoblock in patients in the control group. The MRI scans of patients with LSS often show degenerated discs in addition to spinal stenosis. We omitted the evaluation of degenerated disc by discoblock in the control group. Patients had been diagnosed as having LSS based on clinical features such as lower limb symptoms and MRI findings; thus, an additional discoblock seemed unnecessary from clinical and ethical standpoints. However, if there were negative findings among the control group, our method for diagnosing discogenic LBP is more reliable. Third, we could not diagnose the degenerated disc responsible for LBP by the medical interview without performing the discoblock if there were multiple degenerated discs on the patient's MRI. Fourth, we did not evaluate any other diseases such as LBP associated with sacroiliac joint dysfunction, LBP of the zygapophyseal joint, and major disturbances of the central nervous system associated with chronic pain. Fifth, there was selection bias among our patients. All patients in the DDD group had undergone surgery, although the therapy for discogenic LBP associated with DDD was usually considered conservative. Although patients in the DDD group who were sent to our hospital had a tendency of having severe LBP, patients from one hospital cannot represent patients with discogenic LBP in general.

Conclusions

In accordance with our hypotheses that discogenic LBP exists and that a positive response to a discoblock indicates discogenic LBP, all five items of our medical interview about LBP (i.e., LBP after sitting too long, LBP while standing after sitting too long, squirming in a chair after sitting too long, LBP while washing one's face, and LBP in standing position with flexion) were useful for diagnosing discogenic LBP associated with DDD. We can diagnose discogenic LBP in all cases if four or five of the five items of the medical interview are positive, and in some cases, if three of five items are positive.

Supporting Information

S1 File. Supporting information. Dataset of this study.
(XLS)

Acknowledgments

We would like to thank Hiromi Yamanobe, Ai Ido, Kumiko Shimada, and Yuka Miura who are the medical clerks at the hospital for collecting the data.

Author Contributions

Conceptualization: JT KM HO ST.

Data curation: JT.

Formal analysis: JT HO.

Investigation: JT HI.

Methodology: JT HO.

Resources: JT HI YT YY HK RS YO SB.

Supervision: KM ST.

Writing – original draft: JT.

Writing – review & editing: JT HO JK.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380: 2163–2196. doi: 10.1016/S0140-6736(12)61729-2 PMID: 23245607
2. Fujii T, Matsudaira K. Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. Eur Spine J. 2013; 22: 432–438. doi: 10.1007/s00586-012-2439-0 PMID: 22868456
3. Deyo RA, Weinstein JN. Low back pain. N Engl J Med. 2001; 344: 363–370. doi: 10.1056/NEJM200102013440508 PMID: 11172169
4. Kjaer P, Leboeuf-Yde C, Korsholm L, Sorensen JS, Bendix T. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. Spine. 2005; 30: 1173–1180. PMID: 15897832
5. Cheung KM, Karppinen J, Chan D, Ho DW, Song YQ, Sham P, et al. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine. 2009; 34: 934–940. doi: 10.1097/BRS.0b013e3181a01b3f PMID: 19532001
6. Berg L, Hellum C, Gjertsen Ø, Neckelmann G, Johnsen LG, Storheim K, et al. Do more MRI findings imply worse disability or more intense low back pain? A cross-sectional study of candidates for lumbar disc prosthesis. Skeletal Radiol. 2013; 42: 1593–1602. doi: 10.1007/s00256-013-1700-x PMID: 23982421
7. Endean A, Palmer KT, Coggon D. Potential of magnetic resonance imaging findings to refine case definition for mechanical low back pain in epidemiological studies: a systematic review. Spine. 2011; 36: 160–169. doi: 10.1097/BRS.0b013e3181cd9adb PMID: 20739918
8. Eck JC, Sharan A, Resnick DK, Watters WC 3rd, Ghogawala Z, Dailey AT, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 6: discography for patient selection. J Neurosurg Spine. 2014; 21: 37–41. doi: 10.3171/2014.4.SPINE14269 PMID: 24980583
9. Manchikanti L, Benyamin RM, Singh V, Falco FJ, Hameed H, Derby R, et al. An update of the systematic appraisal of the accuracy and utility of lumbar discography in chronic low back pain. Pain Physician. 2013; 16: SE55–SE95. PMID: 23615887

10. Ohtori S, Kinoshita T, Yamashita M, Inoue G, Yamauchi K, Koshi T, et al. Results of surgery for discogenic low back pain: a randomized study using discography versus discoblock for diagnosis. Spine. 2009; 34: 1345–1348. doi: 10.1097/BRS.0b013e3181a401bf PMID: 19440168

11. Schwarzer AC, Aprill CN, Derby R, Fortin J, Kine G, Bogduk N. The prevalence and clinical features of internal disc disruption in patients with chronic low back pain. Spine. 1995; 20: 1878–1883. PMID: 8560335

12. Carragee EJ, Tanner CM, Khurana S, Hayward C, Welsh J, Date E, et al. The rates of false-positive lumbar discography in select patients without low back symptoms. Spine. 2000; 25: 1373–1380. PMID: 10828919

13. Carragee EJ, Lincoln T, Parmar VS, Alamin T. A gold standard evaluation of the "discogenic pain" diagnosis as determined by provocative discography. Spine. 2006; 31: 2115–2123. doi: 10.1097/01.brs.0000231436.30262.dd PMID: 16915099

14. Carragee EJ, Don AS, Hurwitz EL, Cuellar JM, Carrino JA, Herzog R. 2009 ISSLS Prize Winner: Does discography cause accelerated progression of degeneration changes in the lumbar disc: a ten-year matched cohort study. Spine. 2009; 34: 2338–2345. doi: 10.1097/BRS.0b013e3181ab5432 PMID: 19755936

15. Cuellar JM, Stauff MP, Herzog RJ, Carrino JA, Baker GA, Carragee EJ. Does provocative discography cause clinically important injury to the lumbar intervertebral disc? A 10-year matched cohort study. Spine J. 2016; 16: 273–280. doi: 10.1016/j.spinee.2015.06.051 PMID: 26133255

16. Young S, Aprill C, Laslett M. Correlation of clinical examination characteristics with three sources of chronic low back pain. Spine J. 2003; 3: 460–465. PMID: 14609690

17. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy. 1980; 66: 271–273. PMID: 6450426

18. Fujiwara A, Kobayashi N, Saiki K, Kitagawa T, Tamai K, Saotome K. Association of the Japanese Orthopaedic Association score with the Oswestry Disability Index, Roland-Morris Disability Questionnaire, and short-form 36. Spine. 2003; 28: 1601–1607. PMID: 12865852

19. Krismer M, van Tulder M; Low Back Pain Group of the Bone and Joint Health Strategies for Europe Project. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best Pract Res Clin Rheumatol. 2007; 21: 77–91. doi: 10.1016/j.berh.2006.08.004 PMID: 17350545

20. Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine. 2008; 33: 95–103. doi: 10.1097/BRS.0b013e31815e7f94 PMID: 18165754

21. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine. 2001; 26: 1873–1878. PMID: 11568697

22. Jones AD, Wafai AM, Easterbrook AL. Improvement in low back pain following spinal decompression: observational study of 119 patients. Eur Spine J. 2014; 23: 135–141. doi: 10.1007/s00586-013-2964-5 PMID: 23963487

23. Solberg T, Johnsen LG, Nygaard ØP, Grotle M. Can we define success criteria for lumbar disc surgery?: estimates for a substantial amount of improvement in core outcome measures. Acta Orthop. 2013; 84: 196–201. doi: 10.3109/17453674.2013.786634 PMID: 23506164

24. Nachemson AL. Disc pressure measurements. Spine. 1981; 6: 93–97. PMID: 7209680

25. Sato K, Kikuchi S, Yonezawa T. In vivo intradiscal pressure measurement in healthy individuals and in patients with ongoing back problems. Spine. 1999; 24: 2468–2474. PMID: 10626309

26. Nachemson A. The load on lumbar disks in different positions of the body. Clin Orthop Relat Res. 1966; 45: 107–122. PMID: 5937361

27. Nachemson A, Elfström G. Intravital dynamic pressure measurements in lumbar discs. A study of common movements, maneuvers and exercises. Scand J Rehabil Med Suppl. 1970; 1: 1–40. PMID: 4257209

28. Wilke HJ, Neef P, Caimi M, Hoogland T, Claes LE. New in vivo measurements of pressures in the intervertebral disc in daily life. Spine. 1999; 24: 755–762. PMID: 10222525

29. Adams MA, Hutton WC. The effect of posture on the fluid content of lumbar intervertebral discs. Spine. 1983; 8: 665–671. PMID: 6685921

30. McMillan DW, Garbutt G, Adams MA. Effect of sustained loading on the water content of intervertebral discs: implications for disc metabolism. Ann Rheum Dis. 1996; 55: 880–887. PMID: 9014581

Tsuji et al. BMC Musculoskeletal Disorders (2016) 17:447
DOI 10.1186/s12891-016-1304-4

BMC Musculoskeletal Disorders

RESEARCH ARTICLE Open Access

The impact of depression among chronic low back pain patients in Japan

Toshinaga Tsuji[1], Ko Matsudaira[2], Hiroki Sato[1] and Jeffrey Vietri[3*]

Abstract

Background: Chronic low back pain (CLBP) is associated with significant disability and reductions in health related quality of life (HRQoL), which can negatively impact overall function and productivity. Depression is also associated with painful physical symptoms, and is often present in patients with chronic pain. However, the incremental burden associated with depression or symptoms of depression among CLBP patients is not well understood. The objective of this study was to investigate the impact of depression on HRQoL in CLBP and to assess the relationship between depression and work impairment and healthcare use among CLBP patients in Japan.

Methods: Data were extracted from the 2014 Japan National Health and Wellness Survey ($N = 30{,}000$). CLBP was defined by report of diagnosed low back pain ≥3 months duration. Depression was assessed using the Patient Health Questionnaire (PHQ-9). Measurements assessed included pain, HRQoL, labor force participation, work productivity and healthcare utilization. Patients with depression (PHQ-9 ≥ 10) were compared to patients without depression (PHQ-9 < 10) using t-tests for continuous and count variables and chi-square for categorical variables, which were followed by generalized linear models adjusted for covariates. The association between presenteeism and other patient outcomes and characteristics was analysed using nonparametric correlations (Spearman's rho).

Results: Depressed CLBP patients had significantly more severe pain and higher levels of pain compared with patients without depression ($P < 0.001$). Depression was associated with worse HRQoL in CLBP patients. Presenteeism, overall work impairment and activity impairment were 1.8, 1.9 and 1.7 times as high, respectively, among those with depression relative to those without depression. CLBP patients with depression had almost twice as many healthcare provider visits in 6 months than those without depression. The pattern of results remained consistent after adjustment for sociodemographic and general health characteristics. Analysis also indicated presenteeism was closely related to overall work impairment (rho = 0.99).

Conclusions: Depression among CLBP patients in Japan was associated with higher pain scores and lower HRQoL scores, as well as lower labor productivity and increased healthcare use. Screening for depression in CLBP patients should be an essential part of CLBP patient care.

Keywords: Low back pain, Chronic low back pain, Depression, Quality of life

Background

Low back pain (LBP) is a common health issue affecting at least 80 % of individuals during their lifetime [1] and poses a severe economic burden on individuals and their communities [2–5]. The Global Burden of Disease Study 2013 found that globally, back pain was one of the leading cause of years lived with disability (YLDs) [6]. In Japan, back pain is the top cause of YLD and the 2nd and 4th most frequent reason for outpatient visits for women and men, respectively [6, 7].

One of the main characteristics of LBP is recurrence, and a number of patients develop chronic LBP (CLBP). In Japan CLBP is the most prevalent type of chronic pain [8], with a prevalence estimated at 23 %, and 11–12 % of the population is disabled by it [9]. Though considerable research has been directed at understanding back pain, most Japanese epidemiological studies examine LBP in general, with few focused on CLBP [10–12].

* Correspondence: jeffrey.vietri@kantarhealth.com
[3]Health Outcomes Practice, Kantar Health, 700 Dresher Road, Horsham, PA 19044, USA
Full list of author information is available at the end of the article

While burdensome in its own right, pain is also risk factor for depression, and many studies have examined the co-occurrence of pain and depression [13–16]. The comorbidity is clinically well established but the underlying mechanisms are not well understood, though a potential explanation is disruption of the mesolimbic dopamine system [17, 18]. Recent data from animal models indicate that regulation of dopamine activity in the ventral tegmental area (VTA) mediates depressive and anxiogenic responses [19] suggesting a neurological link between depression and chronic pain.

CLBP in particular is often co-morbid with depression [20], a main cause of disability worldwide [6]. Depression increases the risk of developing LBP [21], and CLBP is affected by the patient's mental state [22]. In spite of that, the mental state of most CLBP patients is not routinely assessed. Thus, in chronic pain, psychosocial risk factors become relevant, and are important to explain how individuals respond to back pain. Recent studies have demonstrated that psychosocial factors are important risk factors for LBP among Japanese workers [22, 23]; however, data examining the role of depression in CLBP patients in Japan is lacking.

The objective of this study was to investigate the impact of depression on health-related quality of life (HRQoL) in CLBP, as well as to assess the relationship between depression and work impairment and healthcare use among CLBP patients in Japan.

Methods

Sample

Data were extracted from the 2014 Japan National Health and Wellness Survey (NHWS) (Kantar Health, New York, USA), which is a general health survey designed to reflect the health of the population in Japan ($N = 30{,}000$). The survey is administered via the Internet, with potential respondents identified through opt-in survey panels. Participants were stratified by gender and age groups to ensure representative samples, with quotas set through the distribution of age and gender within the Japanese population aged ≥18 years.

Respondents were considered to have CLBP if they had been diagnosed with back pain by a doctor, reported experiencing back pain in the past month, and experienced back pain ≥ 3 months. Three months duration of LBP is considered chronic according to both Japanese and US treatment guidelines [24, 25]. Depression symptoms and severity of depression over the last two weeks was assessed using the Patient Health Questionnaire (PHQ-9), a validated scale used to screen for depression and assess its severity [26]. The scale evaluates depression by measuring the frequency of anhedonia, depressed mood, sleep disturbance, lack of energy, appetite disturbance, negative self-feelings, difficulty concentrating, psychomotor retardation or agitation, and thoughts of self-harm. A single-item measure of the interference of these symptoms was also included. Respondents who scored ≥ 10 (the cutoff associated with moderate depression) were considered to have depression regardless of whether they indicated a diagnosis of depression, and respondents scoring <10 (associated with minimal or mild depression) were considered not to have depression; this value has shown good sensitivity and specificity for major depression in previous research [27].

Measures

Using a 0–10 numeric rating scale (NRS) anchored by *No Pain* (0) and *Pain as Bad as You Can Imagine* (10), respondents rated the severity of their LBP, as well as the severity of their pain overall, as mild (0–3), moderate (4–6), or severe (7–10). The NRS was completed for both current and pain in the past week. Respondents indicated how frequently they experienced problems with pain on a 6-point scale ranging from *Daily* to *Once a month or less often*. HRQoL was measured using the revised Medical Outcomes Study 36-Item Short Form Survey Instrument (SF-36v2;[28]). This is a multipurpose, generic HRQoL instrument comprising 36 questions. The instrument is designed to report on eight health concepts (physical functioning (PF), role physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role emotional (RE), and mental health (MH)). The versions of the scores used in this study were based on the Japanese norms, which have a mean of 50 and standard deviation of 10 in the Japanese population [29]. Scores can be interpreted relative to this population average of 50 as well as with other comparison groups of interest. Higher scores indicate better quality of life.

Mental component summary (MCS), physical component summary (PCS), and short form 6D (SF-6D) health utility scores were also calculated according to the standard scoring algorithms. These scores are based on the US (MCS & PCS) and UK (SF-6D) general populations, but are commonly reported in studies outside those countries as the scores allow for comparison across international populations.

Labor force participation was defined as being employed or unemployed but looking for work. Work productivity was assessed using the Work Productivity and Activity Impairment (WPAI) questionnaire, a 6-item validated instrument which consists of four metrics: absenteeism (the percentage of work time missed because of one's health in the past seven days), presenteeism (the percentage of impairment experienced while at work in the past seven days because of one's health), overall work productivity loss (an overall impairment estimate that is a combination of absenteeism and presenteeism),

and activity impairment (the percentage of impairment in daily activities because of one's health in the past seven days) [30]. Only respondents who reported being employed full-time or part-time provided data for absenteeism, presenteeism, and overall work impairment. All respondents provided data for activity impairment.

Healthcare utilization was defined by the number of healthcare provider visits, the number of hospital emergency room (ER) visits, and the number of times hospitalized in the past six months. The reason for each visit was not included in the questionnaire.

Analysis

The analysis was primarily concerned with the association between the presence of depression, so patients with depression (PHQ-9 ≥ 10) were compared with those without depression (PHQ-9 < 10) using t-tests for continuous and count variables and chi-square for categorical variables. To ensure differences due to confounding variables were not attributed to depression, these tests were followed by regression modelling using generalized linear models adjusting for age, sex, length of LBP diagnosis, Charlson Comorbidity Index (CCI), household income, marital status, university education, body mass index (BMI), cigarette smoking, alcohol use, and exercise to account for sociodemographic characteristics and general health characteristics.

These comparisons according to were supplemented by correlational analysis, using the PHQ-9 score as a continuous measure. Because some outcomes were positively skewed rather than normally distributed, the association between presenteeism and other patient outcomes and characteristics was analysed using nonparametric correlations (Spearman's rho).

Results

Of the participants surveyed, 425 were identified as having CLBP. The average age of a respondent with CLBP was 54 years old, and 44 % were female (Table 1). When assessed according to depression status, CLBP patients with depression (PHQ-9 ≥ 10; $N = 70$) were younger than CLBP patients without depression (PHQ-9 < 10; $N = 355$) by approximately 9 years on average, but did not differ in terms of average CCI score, gender, or employment status. Patients with depression were less likely to be married or live with a partner (Table 1). Patients indicated their LBP was either mild (47 %) or moderate (44 %) rather than severe (9 %). Both overall severity of pain and current level of pain were near the midpoint of the NRS, and almost half reported daily problems with pain. Depression was significantly associated with more severe pain and higher levels of pain, current and in the prior week (Table 1).

CLBP patients with depression had worse HRQoL than CLBP patients without depression (Table 2). Depression was also associated with more impairment while at work (presenteeism). Overall work impairment, which is largely driven by presenteeism, was also significantly higher among CLBP patients with depression. There was no significant difference in absenteeism or rate of labor force participation between CLBP patients with and without depression. Depressed CLBP patients reported more activity impairment than those without depression. Depression was also associated with approximately two more healthcare provider visits among CLBP patients in the 6 month recall period (Table 2).

The pattern of results was consistent after covariates were incorporated into the regression analysis. Adjusted HRQoL scores were lower on all of the eight Japanese norm-based scores. Adjusted mean MCS and PCS using international norms were also lower (Fig. 1).

Regression-adjusted presenteeism and overall work impairment were 1.8 and 1.9 times as high, respectively, among those with depression relative to those without depression (Fig. 2). Activity impairment was 1.7 times as high in patients with depression compared with patients without depression after adjustment for covariates. HCP visits were almost twice as frequent in patients with depression compared with patients without depression. Likewise, work impairment was greater in patients with depression compared with patients without depression.

Analysis of depression based on PHQ-9 scores as a continuous variable also demonstrated the association between depression and pain among CLBP patients. Greater depression was significantly associated with more frequent problems with pain, greater current and past-week severity of pain (based on NRS scores), pain at more sites in addition to LBP, and more presenteeism and overall work impairment ($P < 0.001$, Table 3). Moreover, additional regression analysis conducted using PHQ-9 scores as a continuous variable corroborated the findings, indicating lower HRQoL scores with higher PHQ-9 scores, with the exception of the Japanese PCS score. Pain was likewise worse with greater depression as was presenteeism, overall work impairment, and activity impairment. Consistent with the results shown in Fig. 2, HCP visits were more frequent with greater depression scores, but there was no significant association with ER visits or hospitalizations (data not shown).

When assessing the relationship between work impairment and other characteristic and outcomes, presenteeism was very closely related to overall work impairment (rho = 0.99). Greater presenteeism was associated with more-severe LBP, more-severe pain in the prior week and currently based on the NRS. Although, there was a trend for greater presenteeism being associated with more frequent problems with pain,

Table 1 Characteristics of CLBP patients according to presence of depression

	Total (*N* = 425)	Depression (PHQ-9 ≥ 10) (*N* = 70)	No Depression (PHQ-9 < 10) (*N* = 355)	*P value*
Age, Mean ± SD	53.90 ± 14.16	45.91 ± 13.73	55.48 ± 13.73	<0.001
Female, n (%)	187 (44.00)	33 (47.14)	154 (43.38)	0.562
Employment status, n (%)				0.589
Not currently employed	164 (38.59)	25 (35.71)	139 (39.15)	
Employed	261 (61.41)	45 (64.29)	216 (60.85)	
Annual household income, n (%)				0.079
< ¥3million	83 (19.53)	22 (31.43)	61 (17.18)	
¥3million to < ¥5million	100 (23.53)	15 (21.43)	85 (23.94)	
¥5million to < ¥8million	113 (26.59)	15 (21.43)	98 (27.61)	
¥8million or more	97 (22.82)	12 (17.14)	85 (23.94)	
Decline to answer	32 (7.53)	6 (8.57)	26 (7.32)	
Marital status, n (%)				0.021
Single/Divorced/Separated/Widowed	138 (32.47)	31 (44.29)	107 (30.14)	
Married/living with partner	287 (67.53)	39 (55.71)	248 (69.86)	
Education level, n (%)				0.063
Less than university education	218 (51.29)	43 (61.43)	175 (49.30)	
University education or higher	207 (48.71)	27 (38.57)	180 (50.70)	
Body mass index category, n (%)				0.137
Underweight	52 (12.24)	9 (12.86)	43 (12.11)	
Normal weight	280 (65.88)	39 (55.71)	241 (67.89)	
Overweight	70 (16.47)	16 (22.86)	54 (15.21)	
Obese	19 (4.47 %)	4 (5.71)	15 (4.23)	
Decline to provide weight	4 (0.94 %)	2 (2.86)	2 (0.56)	
Smoking behavior, n (%)				0.088
Never smoked	182 (42.82)	34 (48.57)	148 (41.69)	
Former smoker	132 (31.06)	14 (20.00)	118 (33.24)	
Current smoker	111 (26.12)	22 (31.43)	89 (25.07)	
Alcohol use, n (%)				0.975
Do not drink	116 (27.29)	19 (27.14)	97 (27.32)	
Drink alcohol	309 (72.71)	51 (72.86)	258 (72.68)	
Vigorous exercise at least one day in the past month, n (%)				0.198
Do not exercise	213 (50.12)	40 (57.14)	173 (48.73)	
Exercise	212 (49.88)	30 (42.86)	182 (51.27)	
Charlson comorbidity index, Mean ± SD	0.51 ± 2.23	0.83 ± 3.64	0.44 ± 1.83	0.186
Sleep difficulties, n (%)	82 (19.29)	35 (50.00)	47 (13.24)	<0.001
Severity of LBP, n (%)				<0.001
Mild	186 (47.45)	19 (29.23)	167 (51.07)	
Moderate	172 (43.88)	34 (52.31)	138 (42.20)	
Severe	34 (8.67)	12 (18.46)	22 (6.73)	
Missing	33	5	28	
Severity of pain in the prior week (0–10), Mean ± SD	4.48 ± 2.31	5.80 ± 2.26	4.23 ± 2.23	<0.001
Current severity of pain (0–10), Mean ± SD	4.59 ± 2.28	5.86 ± 2.27	4.34 ± 2.20	<0.001

Table 1 Characteristics of CLBP patients according to presence of depression *(Continued)*

Frequency of problems with pain, n (%)				0.002
Daily	188 (44.24)	44 (62.86)	144 (40.56)	
4–6 times a week	63 (14.82)	12 (17.14)	51 (14.37)	
2–3 times a week	82 (19.29)	10 (14.29)	72 (20.28)	
Once a week	39 (9.18)	3 (4.29)	36 (10.14)	
2–3 times a month	35 (8.24)	0 (0.00)	35 (9.86)	
Once a month or less often	18 (4.24)	1 (1.43)	17 (4.79)	
Type of diagnosing doctor for LBP, n (%)				0.028
General internist	18 (4.24)	4 (5.71)	14 (3.94)	
Gynecologist	5 (1.18)	0 (0.00)	5 (1.41)	
Orthopedist	353 (83.06)	54 (77.14)	299 (84.23)	
Rheumatologist	4 (0.94)	3 (4.29)	1 (0.28)	
Pain management specialist	3 (0.71)	1 (1.43)	2 (0.56)	
Other	42 (9.88)	8 (11.43)	34 (9.58)	
Type of prescribing doctor, n (%)				0.150
General internist	28 (16.87)	6 (15.38)	22 (17.32)	
Gynecologist	2 (1.20)	0 (0.00)	2 (1.57)	
Orthopedist	116 (69.88)	24 (61.54)	92 (72.44)	
Rheumatologist	5 (3.01)	3 (7.69)	2 (1.57)	
Pain management specialist	1 (0.60)	0 (0.00)	1 (0.79)	
Other	14 (8.43)	6 (15.38)	8 (6.30)	
Missing	259	31	228	
Duration of LBP (months), Mean ± SD	112 ± 120	96 ± 99	115 ± 123	0.227
Current use of a prescription medication for pain, n (%)				0.002
No	259 (60.94)	31 (44.29)	228 (64.23)	
Yes	166 (39.06)	39 (55.71)	127 (35.77)	
Current use of NSAIDs prescription for pain, n (%)				0.049
No	40 (24.10)	14 (35.90)	26 (20.47)	
Yes	126 (75.90)	25 (64.10)	101 (79.53)	
Missing	259	31	228	
Use of an OTC product for pain, n (%)				0.861
No	306 (72.00)	51 (72.86)	255 (71.83)	
Yes	119 (28.00)	19 (27.14)	100 (28.17)	
Use of an herbal product for pain, n (%)				0.441
No	413 (97.18)	69 (98.57)	344 (96.90)	
Yes	12 (2.82)	1 (1.43)	11 (3.10)	

[a]NSAIDS are prescription drugs in Japan. *CLBP* chronic low back pain, *LBP* low back pain, *NSAIDs* non-steroidal anti-inflammatory drugs, *OTC* over-the-counter, *PHQ-9* Patient Health Questionaire-9

it did not reach statistical significance ($P = 0.08$). Presenteeism was moderately related to the severity of depression according to the PHQ-9 score (Table 4).

Discussion

Our results demonstrated that CLBP patients with depression had significantly more severe and higher levels of pain, as well as significantly worse HRQoL compared with CLBP patients without depression. These observations are consistent with those recently published by Hiyama et al., which showed that depressed patients and those with neuropathic LBP had a higher level of pain and poorer quality of life compared with non-depressed patients [16]. The majority of patients had mild (47 %) or moderate (44 %) LBP. Current and prior week pain severity scores were similar (4.6/10 vs 4.5/10)

Table 2 Outcomes among CLBP patients according to presence of depression

	Total (*N* = 425)	Depression (PHQ-9 ≥ 10) (*N* = 70)	No Depression (PHQ-9 < 10) (*N* = 355)	
	Mean ± SD	Mean ± SD	Mean ± SD	*P* value
Health status: Japanese norm-based scores				
Physical functioning	44.36 ± 15.43	37.73 ± 17.9	45.66 ± 14.57	<0.001
Role physical	42.26 ± 14.29	32.51 ± 16.28	44.19 ± 13.05	<0.001
Bodily pain	39.59 ± 8.89	34.61 ± 9.56	40.57 ± 8.43	<0.001
General health	42.59 ± 10.96	33.25 ± 9.24	44.43 ± 10.32	<0.001
Vitality	42.87 ± 10.92	30.98 ± 9.26	45.22 ± 9.63	<0.001
Social functioning	43.03 ± 13.36	30.51 ± 13.63	45.49 ± 11.85	<0.001
Role emotional	44.7 ± 13.14	32.34 ± 15.13	47.14 ± 11.22	<0.001
Mental health	45.17 ± 11.2	32 ± 9.54	47.77 ± 9.55	<0.001
Health status: International scores				
Mental component	45.01 ± 10.92	31.27 ± 10.04	47.72 ± 8.85	<0.001
Physical component	46.81 ± 7.65	44.08 ± 7.96	47.35 ± 7.48	0.001
Health utility (SF-6D)	0.67 ± 0.12	0.56 ± 0.09	0.69 ± 0.11	<0.001
Work impairment				
Absenteeism %	4.92 ± 17.87	7.33 ± 22.37	4.39 ± 16.75	0.335
Presenteeism %	31.59 ± 28.08	46.43 ± 26.12	28.43 ± 27.52	<0.001
Overall work impairment %	33.90 ± 30.08	49.81 ± 27.74	30.40 ± 29.50	<0.001
Activity impairment %	37.34 ± 29.90	56.00 ± 27.21	33.66 ± 29.05	<0.001
HCP visits (past 6 months)	12.64 ± 16.24	19.67 ± 21.07	11.25 ± 14.75	<0.001

HCP healthcare provider

and almost half of all patients reported daily pain problems. Overall sociodemographic patient characteristics were similar between the two groups of CLBP patients with the exception of age, marital status and sleeping difficulties. CLBP patients with depression were significantly younger, on average 9 years, compared with CLBP patients without depression. These observations tend to be consistent with observations for major depressive disorder where estimates in the general population are 15–17 %, while the 1-year prevalence rate in individuals ≥ 65 years is lower, at 1–4 % [31]. Significantly more CLBP patients with depression were single/divorced compared with CLBP patients without depression (44.3 % vs 30.1 %). However, differences in marital status and sleeping difficulties were consistent with differences observed in major depression disorder [27].

Epidemiological, cross-sectional, and prospective studies suggest that insomnia, chronic pain and depression are a cluster of symptoms that are mutually interactive. Studies using a variety of methods, including neuroimaging,

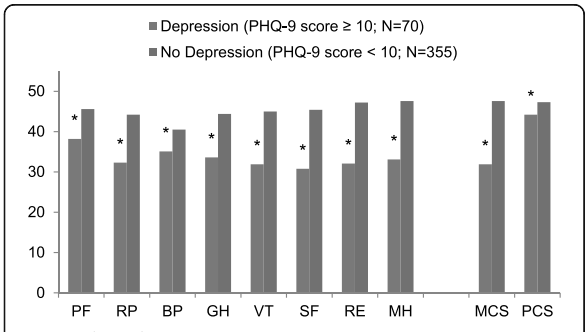


Fig. 1 Adjusted mean HRQoL scores among CLBP patients according to presence of depression. *$p < 0.05$

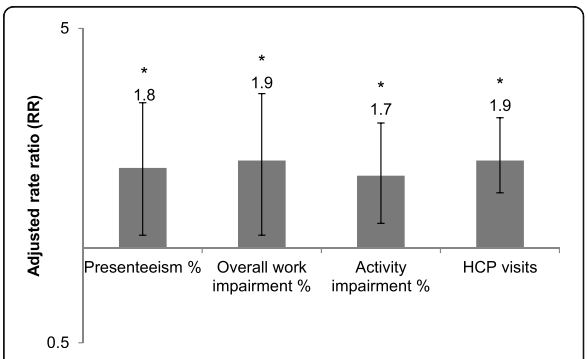


Fig. 2 Adjusted impairments and healthcare visit rates among depressed CLBP patients relative to those without depression. *$p < 0.05$. Results are presented on a logarithmic scale; values above 1 (x-axis) indicate increased impairment and resource use among CLBP patients with depression

Table 3 Correlations between depression, pain, and work impairment among CLBP patients

	Spearman's rho with PHQ-9 score[a]	*P* value
Frequency of problems with pain	0.236	<0.001
Current severity of pain (based on NRS score)	0.289	<0.001
Severity of pain in the prior week (based on NRS score)	0.332	<0.001
Additional pain sites (number, 0–6)	0.387	<0.001
Presenteeism %	0.340	<0.001
Overall work impairment %	0.342	<0.001

[a]Correlation is significant at the 0.01 level (2-tailed)
CLBP chronic low back pain, *NRS* numeric rating scale, *PHQ-9* Patient Health Questionnaire

suggest the mesolimbic dopamine system has been proposed as a key factor in promoting the comorbidity of this cluster of symptoms, [32] and our observations of both higher ratings of pain severity as well as greater prevalence of sleep difficulties among CLBP patients with depression are additional supportive evidence to this body of data.

The adjusted mean HRQoL scores in the CLBP depression group were lower than in the CLBP group without depression. The health status using both Japanese norm-based as well was international scores, indicated significantly poorer outcomes for CLBP patients with depression compared with CLBP patients without depression. Lower PCS scores in CLBP patients with depression are indicative that a decline of mental health could have an effect on physical health in CLBP patients. A similar relationship has been reported among CLBP patients in the United Kingdom, in whom depression as measured by the Hospital Anxiety and Depression Scale (HADS) was correlated with PCS scores [33]. Labor force and absenteeism did not differ by depression status, potentially because of Japanese working habits, where there is a tendency for less sick leave claims compared to other countries [34]. However, presenteeism, overall work and activity impairment were lower in CLBP patients with depression, demonstrating that, even though employees are present at work, they are less productive than those CLBP patients without depression. Additional analyses indicated that presenteeism was closely related to overall work impairment. The current study also demonstrates more frequent use of healthcare among CLBP patients who have depression, consistent with the relationship between depression and healthcare visits recently demonstrated in the US using National Health and Nutrition Examination Survey data [35].

Table 4 Correlations between presenteeism, pain, and depression among employed CLBP patients

	Spearman's rho with presenteeism	*P* value
Overall work impairment %	0.990[a]	<0.001
Severity of lower back pain	0.267[a]	<0.001
Severity of pain in the prior week (based on NRS)	0.297[a]	<0.001
Current severity of pain (based on NRS)	0.245[a]	<0.001
Frequency of problems with pain	0.115	0.077
Sites of pain in addition to LBP (number, 0–6)	0.239[a]	0.002
Depression severity based on PHQ-9	0.342[a]	<0.001

[a]Correlation is significant at the 0.01 level (2-tailed)

Treatment approaches, especially for Japanese workers, have focused on ergonomic approaches in the management of LBP. Consistent with a focus on musculoskeletal symptoms the majority of patients surveyed in our study were diagnosed with LBP by an orthopedist. However, recent studies highlight the importance of psychosocial risk factors in the development of CLBP [22, 23] and our data further highlights the need for mental health evaluation and treatment in addition to physical assessment and therapy.

One limitation of our study is that the analysis was cross-sectional. Therefore our results cannot indicate whether increased pain leads to depression, or whether depression leads to increased pain. Another limitation is selection bias that may not result in an all-encompassing representation of all patients with CLBP. The data were derived from opt-in surveys completed over the Internet. Compared to the general population our study population could be over-representative of individuals who live in urban environments and are technology literate.

Nakamura et al. has shown that chronic musculoskeletal pain does not necessarily improve with treatment and that patients have a high degree of dissatisfaction with it [11, 12]. Ineffective treatment may lead to "doctor shopping". In our study, a significantly higher number of CLBP patients with depression than those without depression were using prescription pain medication (55.7 % vs 35.8 %, $P = 0.002$) indicating that depressed CLBP patients not only suffer more but may also find treatment less effective. Moreover, increased mental and physical suffering often require assistance. All these factors pose undue strain and increase societal cost.

Conclusion

We have demonstrated that depression among CLBP patients is associated with higher pain scores and lower HRQoL scores, as well as lower labor productivity and increased healthcare use. Our results underscore the need to screen for depression in CLBP patients as an essential part of CLBP patient care.

Abbreviations

BMI: Body mass index; BP: Bodily pain; CCI: Charlson Comorbidity Index; CLBP: Chronic low back pain; ER: Emergency room; GH: General health;

HCP: Healthcare provider; HRQoL: Health-related quality of life; LBP: Low back pain; MCS: Mental component summary; MH: Mental health; NHWS: National Health and Wellness Survey; NRS: Numeric rating scale; NSAIDs: Non-steroidal anti-inflammatory drugs; OTC: Over-the-counter; PCS: Physical component summary; PF: Physical functioning; PHQ-9: Patient Health Questionnaire; RE: Role emotional; RP: Role physical; SF: Social functioning; SF-36v2: Medical Outcomes Study 36-Item Short Form Survey Instrument; SF-6D: Short form 6D; VT: Vitality; VTA: Ventral tegmental area; WPAI: Work Productivity and Activity Impairment; YLD: Years lived with disability

Acknowledgements
Writing assistance was provided by Ramona Pufan and funded by Kantar Health.

Funding
This study was funded by Shionogi & Co., LTD.

Availability of data and materials
The dataset supporting the conclusions of this article is proprietary to Kantar Health and will not be shared.

Authors' contributions
TT conceived the study idea. JV conducted the statistical analysis. TT, KM, HS, and JV participated in the interpretation of the results and revision of the manuscript for important intellectual content, and have read and approved the final version of the manuscript.

Competing interests
TT is a full-time employee and minor stock holder of Shionogi & Co., Ltd., & HS is a full-time employee of Shionogi & Co. Ltd. KM has received speaking fees from Shionogi & Co., Ltd., Ayumi Pharmaceutical Co., Eli Lilly Japan K.K., Ono Pharmaceutical Co., Ltd., Pfizer Japan Inc., Nippon Zoki Pharmaceutical Co., Ltd., Eisai Co., Ltd., and Teijin Pharma Ltd.; has received research grants from Pfizer Japan Inc., Eisai Co., Ltd., Ayumi Pharmaceutical Co., Nippon Zoki Pharmaceutical Co., Ltd., Ono Pharmaceutical Co., Ltd., Lilly Japan K.K., Sumitomo Dainippon Pharma Co., Ltd., Astellas Pharma Inc., TOTO Ltd., and Okamura Co.; and is a consultant to Shionogi & Co., Ltd. JV is an employee of Kantar Health, which received fees from Shionogi & Co. Ltd., for access to survey data, analysis, and reporting.

Consent for publication
Not applicable.

Ethics approval and consent to participate
The 2014 Japan NHWS was reviewed for exemption determination by Pearl IRB (Indianapolis, IN, USA; study number 14-KAN-106) prior to participant recruitment and found to meet the exemption requirements under 45CFR46.101(b)[2]. All respondents viewed an on-line informed consent form and indicated their consent to participate prior to responding to the survey. No ethical review was undertaken specific to the analysis of anonymous data presented in this report.

Author details
[1]Medical Affairs Department, Shionogi & Co., Ltd., Osaka, Japan. [2]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo, Japan. [3]Health Outcomes Practice, Kantar Health, 700 Dresher Road, Horsham, PA 19044, USA.

Received: 19 April 2016 Accepted: 17 October 2016
Published online: 27 October 2016

References

1. Hoy D, Brooks P, Blyth F, Buchbinder R. The Epidemiology of low back pain. Best Pract Res Clin Rheumatol. 2010;24(6):769–81.
2. Thelin A, Holmberg S, Thelin N. Functioning in neck and low back pain from a 12-year perspective: a prospective population-based study. J Rehabil Med. 2008;40(7):555–61.
3. Kent PM, Keating JL. The epidemiology of low back pain in primary care. Chiropr Osteopat. 2005;13:13.
4. Steenstra IA, Verbeek JH, Heymans MW, Bongers PM. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. Occup Environ Med. 2005;62(12):851–60.
5. Lidgren L. The bone and joint decade 2000–2010. Bull World Health Organ. 2003;81(9):629.
6. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I,Charlson F, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;386(9995):743–800.
7. Ministry of Health, Labor and Wellfare. Survey of Living coditions. 2010.
8. Takura T, Ushida T, Kanchiku T, Ebata N, Fujii K, daCosta DiBonaventura M, et al. The societal burden of chronic pain in Japan: an internet survey. J Orthop Sci. 2015;20(4):750–60.
9. Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. Lancet. 2012;379(9814):482–91.
10. Fujii T, Matsudaira K. Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. Eur Spine J. 2013;22(2):432–8.
11. Nakamura M, Nishiwaki Y, Ushida T, Toyama Y. Prevalence and characteristics of chronic musculoskeletal pain in Japan. J Orthop Sci. 2011;16(4):424–32.
12. Nakamura M, Nishiwaki Y, Ushida T, Toyama Y. Prevalence and characteristics of chronic musculoskeletal pain in Japan: a second survey of people with or without chronic pain. J Orthop Sci. 2014;19(2):339–50.
13. Vietri J, Otsubo T, Montgomery W, Tsuji T, Harada E. The incremental burden of pain in patients with depression: results of a Japanese survey. BMC Psychiatry [Internet]. 2015;15(1):104. Available from: http://www.scopus.com/inward/record.url?eid = 2-s2.0-84929148103&partnerID = tZOtx3y1.
14. Han C, Pae C-U. Pain and depression: a neurobiological perspective of their relationship. Psychiatry Investig. 2015;12(1):1–8.
15. Denkinger MD, Lukas A, Nikolaus T, Peter R, Franke S. Multisite pain, pain frequency and pain severity are associated with depression in older adults: results from the ActiFE Ulm study. Age Ageing. 2014;43(4):510–4.
16. Hiyama A, Watanabe M, Katoh H, Sato M, Sakai D, Mochida J. Effect of depression and neuropathic pain using questionnaires on quality of life in patients with low back pain; cross-sectional retrospective study. Eur Spine J. 2016;25:2750–60.
17. Taylor AMW, Castonguay A, Taylor AJ, Murphy NP, Ghogha A, Cook C, et al. Microglia disrupt mesolimbic reward circuitry in chronic pain. J Neurosci. 2015;35(22):8442–50.
18. Wood PB, Schweinhardt P, Jaeger E, Dagher A, Hakyemez H, Rabiner EA, et al. Fibromyalgia patients show an abnormal dopamine response to pain. Eur J Neurosci. 2007;25(12):3576–82.
19. Small KM, Nunes E, Hughley S, Addy NA. Ventral tegmental area muscarinic receptors modulate depression and anxiety-related behaviors in rats. Neurosci Lett. 2016;616:80–5.
20. Tetsunaga T, Misawa H, Tanaka M, Sugimoto Y, Tetsunaga T, Takigawa T, et al. The clinical manifestations of lumbar disease are correlated with self-rating depression scale scores. J Orthop Sci Japan. 2013;18(3):374–9.
21. Pinheiro MB, Ferreira ML, Refshauge K, Ordoñana JR, Machado GC, Prado LR, et al. Symptoms of depression and risk of new episodes of low back pain: a systematic review and meta-analysis. Arthritis Care Res (Hoboken). 2015;67(11):1591–603.
22. Matsudaira K, Kawaguchi M, Isomura T, Inuzuka K, Koga T, Miyoshi K, et al. Assessment of psychosocial risk factors for the development of non-specific chronic disabling low back pain in Japanese workers-findings from the Japan Epidemiological Research of Occupation-related Back Pain (JOB) study. Ind Health [Internet]. 2015;53(4):368–77. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4551067/pdf/indhealth-53-368.pdf.
23. Matsudaira K, Konishi H, Miyoshi K, Isomura T, Inuzuka K. Potential risk factors of persistent low back pain developing from mild low back pain in urban Japanese workers. PLoS One. 2014;9(4):5–10.
24. Japanese Orthopaedic Association. Clinical Practice Guideline for the Management of Low Back Pain. Tokyo: Nankodo Co., Ltd.; 2012.
25. Chou R, Qaseem A, Snow V, Casey D, Cross J, Thomas J, Shekelle P, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med [Internet]. 2007;147(7):478–91. Available from: http://dx.doi.org/10.7326/0003-4819-147-7-200710020-00006.
26. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.

27. Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, et al. Cross-national epidemiology of major depression and bipolar disorder. JAMA. 1996;276(4):293–9.
28. Maruish ME, editor. NEW SF36v2 User Guide. 3rd ed. 2011. p. 325.
29. Suzukamo Y, Fukuhara S, Green J, Kosinski M, Gandek B, Ware JE. Validation testing of a three-component model of Short Form-36 scores. J Clin Epidemiol [Internet]. 2011;64(3):301–8. Elsevier. Available from: http://dx.doi.org/10.1016/j.jclinepi.2010.04.017.
30. Spear J, Chawla S, O'Reilly M, Rock D. Does the HoNOS 65+ meet the criteria for a clinical outcome indicator for mental health services for older people? Int J Geriatr Psychiatry. 2002;17(3):226–30. England.
31. Glover J, Srinivasan S. Assessment of the person with late-life depression. Psychiatr Clin North Am. 2013;36(4):545–60.
32. Finan PH, Smith MT. The comorbidity of insomnia, chronic pain, and depression: dopamine as a putative mechanism. Sleep Med Rev. 2013;17(3):173–83.
33. Keeley P, Creed F, Tomenson B, Todd C, Borglin G, Dickens C. Psychosocial predictors of health-related quality of life and health service utilisation in people with chronic low back pain. Pain. 2008;135(1–2):142–50.
34. Matsudaira K, Palmer KT, Reading I, Hirai M, Yoshimura N, Coggon D. Prevalence and correlates of regional pain and associated disability in Japanese workers. Occup Environ Med. 2011;68(3):191–6.
35. Shmagel A, Foley R, Ibrahim H. Epidemiology of chronic low back pain in US adults: National Health and Nutrition Examination Survey 2009–2010. Arthritis Care Res (Hoboken). 2016. Epub ahead.

Skeletal Radiol (2016) 45:1577–1581
DOI 10.1007/s00256-016-2464-x

CASE REPORT

Potential use of ^{18}F-FDG-PET/CT to visualize hypermetabolism associated with muscle pain in patients with adult spinal deformity: a case report

Yuki Taniguchi[1] · Miwako Takahashi[2] · Ko Matsudaira[3] · Hiroyuki Oka[3] · Toshimitsu Momose[2]

Received: 17 June 2016 / Revised: 22 July 2016 / Accepted: 15 August 2016 / Published online: 26 August 2016

Abstract Patients with adult spinal deformity (ASD) are surgically treated for pain relief; however, visualization of the exact origin of the pain with imaging modalities is still challenging. We report the first case of a 60-year-old female patient who presented with painful degenerative kyphoscoliosis and was evaluated with flourine-18-fluoro-2-deoxy-D-glucose positron emission tomography/computed tomography (^{18}F-FDG-PET/CT) preoperatively. Because her low back pain was resistant to conservative treatment, she was treated with posterior spinal correction and fusion surgery from Th2 to the ilium. One year after the surgery, her low back pain had disappeared completely. In accordance with her clinical course, ^{18}F-FDG-PET imaging revealed the uptake of ^{18}F-FDG in the paravertebral muscles preoperatively and showed the complete absence of uptake at 1 year after surgery. The uptake site coincided with the convex part of each curve of the lumbar spine and was thought to be the result of the increased activity of paravertebral muscles due to their chronic stretched state in the kyphotic posture. This case report suggests the possibility of using ^{18}F-FDG-PET/CT to visualize increased activity in paravertebral muscles and the ensuing pain in ASD patients.

Keywords Adult spinal deformity · 18 F-FDG-PET/CT · Low back pain · Muscle pain · Kyphosis · Scoliosis

✉ Yuki Taniguchi
taniguchi-tky@umin.ac.jp

Miwako Takahashi
tmiwako-tky@umin.ac.jp

Ko Matsudaira
kmatsudaira-tky@umin.ac.jp

Hiroyuki Oka
okah-tky@umin.ac.jp

Toshimitsu Momose
tmomo-tky@umin.ac.jp

[1] Department of Orthopedic Surgery, The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

[2] Division of Nuclear Medicine, Department of Radiology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[3] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

Introduction

Adult spinal deformity (ASD) affects a large number of the elderly, and its prevalence is increasing [1–3]. ASD patients have greater functional limitations and worse quality of life than the normal population [4–8]. In particular, sagittal imbalance is correlated with pain; however, the precise etiology of this condition remains unknown [9]. The relationship between sagittal imbalance and low back pain (LBP) was first reported by Takemitsu et al., who defined a new condition called lumbar degenerative kyphosis [10]. In their study, the authors provided evidence that the paravertebral muscles in patients with lumbar degenerative kyphosis are weak and atrophic, with fatty infiltration, and speculated that LBP in these patients is probably caused by fatigue in these weak extensor muscles [10]. To date, some studies have reported increased activity in the paravertebral muscles in the kyphotic position; however, visualization of this increased activity or the ensuing pain in the muscles is still challenging [11, 12]. Here, we report the first case of a patient who presented with painful degenerative kyphoscoliosis and was evaluated with flourine-18-fluoro-2-deoxy-D-glucose positron-emission tomography/computed tomography (^{18}F-FDG-PET/CT) preoperatively. In

this case, the uptake of ^{18}F-FDG in the paravertebral muscles was detected preoperatively, and the increased accumulation was confirmed to disappear at 1 year after surgery, along with the disappearance of her pain.

This case report suggests for the first time the possibility of using ^{18}F-FDG-PET/CT to visualize the hypermetabolic painful paravertebral muscles in ASD patients. Because visualization of the exact origin of the pain in ASD patients has not been achieved by other imaging modalities to date, this case report may assist in opening a new frontier in the management of ASD patients.

Case report

A 60-year-old female patient presented with severe LBP. The radiographic examination of her spine showed degenerative kyphoscoliosis with rotation of the lumbar vertebrae (Fig. 1a, b). Although she was receiving treatment for rheumatoid arthritis (RA), her RA was well controlled and physical examination revealed that her LBP was caused by a spinal deformity. She had no diabetes mellitus. As her pain was resistant to conservative treatment, surgery was planned.

Incidentally, ^{18}F-FDG-PET/CT was performed immediately prior to the operation at the Department of Endocrine Surgery of our hospital for the follow-up of thyroid carcinoma that had been resected 1 year before her first visit to our department. The ^{18}F-FDG-PET/CT was acquired as follows: the patient fasted at least 5 h, and then FDG (296 MBq) was injected with the patient at rest in the supine position. Her plasma glucose level was 93 mg/dl. The PET scan was started 50 min after the injection with the PET-CT scanner (Aquiduo, Toshiba Medical Systems, Otawara, Japan). ^{18}F-FDG-PET/CT images revealed the uptake of ^{18}F-FDG in the bilateral paravertebral muscles of the lumbar spine (Fig. 1c-f). Notably, each uptake site was asymmetrical and coincided with the convex part of the respective curve of the lumbar spine (Fig. 1c–f). The maximum standardized uptake value (SUV-max) was 9.7 on the right side and 4.9 on the left side.

At first, soft tissue metastasis was suspected based on these results. However, magnetic resonance imaging (MRI) showed no evidence of soft tissue tumor (Fig. 1g, h); thus, the uptake was considered to be the result of increased activity in the paravertebral muscles due to the chronic stretching of the muscles in the kyphoscoliotic posture.

As the LBP was not improved by conservative treatment, we performed surgery. At first, the patient was treated with posterior spinal fusion from Th10 to the ilium with interbody fusion and decompression at the level of L3/4, L4/5, and L5/S. Although a spinal orthosis was applied postoperatively, proximal junctional failure with compression fracture of Th11, which caused severe paraplegia, occurred at 2 months after the primary operation. Therefore, we performed revision surgery, extending the fusion level to Th2 (Fig. 2a, b).

After the revision surgery, the postoperative course was uneventful, and the paraplegia gradually improved. One year after the revision surgery, the patient recovered her ability to walk alone with a single T-cane, and her LBP disappeared completely. The patient-oriented questionnaires that were completed preoperatively and at 1 year after the revision surgery revealed improvement of the patient's pain and quality of life. Besides the scores in the numeric rating scale for LBP, the disease-specific outcomes, including the results of the Roland-Morris Disability Questionnaire and Oswestry Disability Index test, improved remarkably (Table 1) [13–16]. To avoid the presence of overcorrection artifacts due to metallic implants on ^{18}F-FDG-PET/CT, conventional ^{18}F-FDG-PET was performed at 1 year after the revision surgery as follow-up of the patient's thyroid carcinoma. This imaging examination showed the complete disappearance of the increased uptake in the paravertebral muscles (Fig. 2c, d).

Discussion

With the development of surgical techniques and the improvement in implants, many ASD patients are treated surgically at present; however, the revision rates can reach 9 % in the long-term follow-up. Therefore, avoiding revision surgery is one of the aims in the care of ASD patients [17]. The main reasons for requiring surgery in ASD patients are persistent severe LBP, gastroesophageal reflux disease, and trunk imbalance, among which LBP is the most frequent reason. From this point of view, the preoperative accurate diagnosis of the origin of the pain is crucial to avoid multiple operations. Thus, visualization of pain in ASD patients is one of the ultimate goals; however, such a diagnostic imaging system had not been established to date.

^{18}F-FDG-PET/CT is a well-established hybrid imaging modality used in oncology, cardiology, and neurology; however, its application in the evaluation of skeletal muscles is still under investigation. ^{18}F-FDG is a glucose analog and is transported into the cells through glucose transporters; thus, uptake of ^{18}F-FDG reflects an increased turnover of glucose. On the basis of this mechanism, it has been reported that vigorous muscle exercise, stress-induced muscle tension, and activities such as talking or chewing can cause a physiological increase in the uptake of ^{18}F-FDG in the muscles involved [18–22]. Based on these observations, it is thought that, in this case, the chronically stretched extensor muscles due to the kyphoscoliotic posture showed a pathological uptake of ^{18}F-FDG due to the increased muscle activity and that the increased uptake completely disappeared after the appropriate posture was acquired [11, 12]. This idea is compatible with the fact that each site in the paravertebral muscles

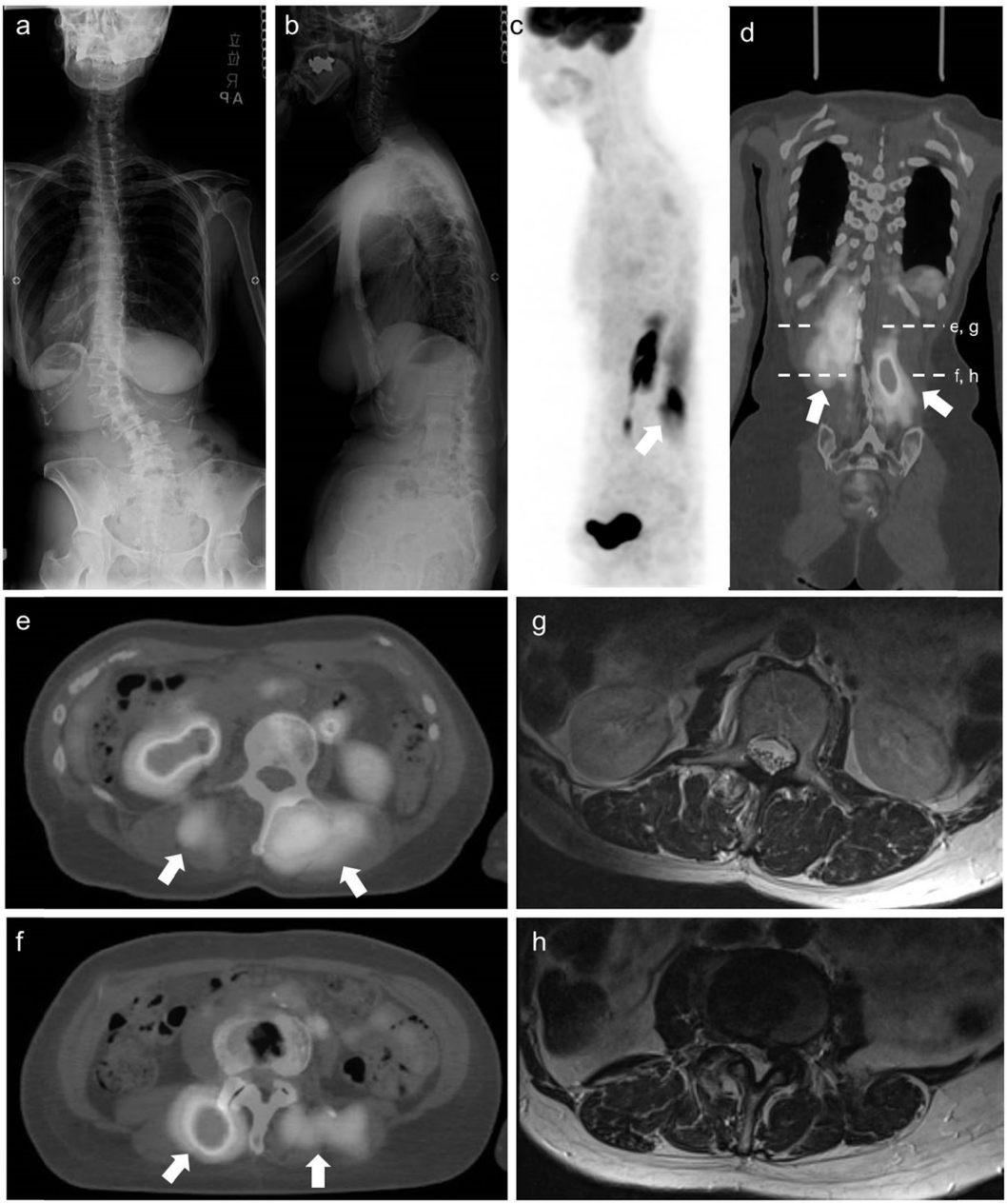


Fig. 1 A 60-year-old female with painful degenerative kyphoscoliosis. **a**, **b** Preoperative posterior-anterior and lateral plain radiographs of the whole spine showing degenerative kyphoscoliosis. **c** Lateral view of the preoperative maximum-intensity-projection image obtained by F-18-fluoro-2-deoxy-D-glucose positron-emission tomography/computed tomography (^{18}F-FDG-PET/CT) revealing abnormal uptake of FDG in the paravertebral muscles with normal distributions in the kidney and bladder. **d** Posterior-anterior view of the preoperative coronal ^{18}F-FDG-PET/CT image at the level of paravertebral and axial ^{18}F-FDG-PET/CT images at **e** L1/2 and **f** L3/4 levels, clearly showing that each uptake site is asymmetrical and coincides with the convex part of the respective curve of the lumbar spine (*white arrows*).**g**, **h** T2-weighted magnetic resonance imaging at the levels corresponding to **e** and **f**, respectively, showing no evidence of soft tissue metastasis. The white dotted lines in **d** indicate the level of each image of **e**, **g**, and **f**, **h**

showing increased uptake coincided with the convex part of each lumbar curve, because the paravertebral muscles are thought to be stretched most at this apex, where the rotation of the vertebra was also the most severe, causing three-dimensional elongation of the paravertebral muscles. As vertebral rotation is reported to be strongly correlated with pain in ASD patients, the pattern of the uptake of ^{18}F-FDG in this case was suggestive of an association with pain; however, further

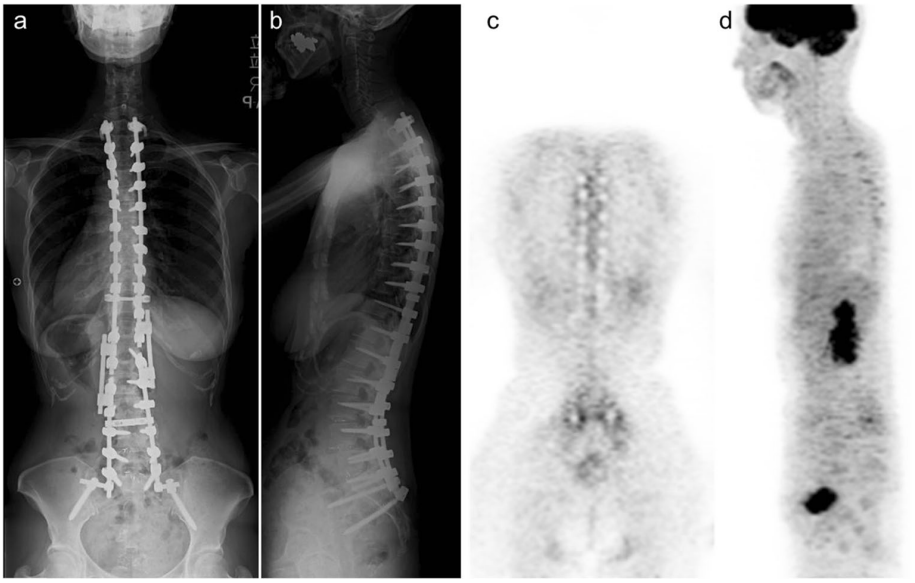


Fig. 2 **a**, **b** Posterior-anterior and lateral plain radiographs taken at 1 year after the revision surgery reveal that good spinal alignment has been achieved. **c**, **d** Posterior-anterior and lateral views of the maximum-intensity-projection images of ^{18}F-FDG-PET clearly show the disappearance of the abnormal uptake of FDG in the paravertebral muscles

investigation is required to determine whether this pathological uptake of ^{18}F-FDG in the paravertebral muscles reflects the pain experienced by ASD patients [23].

Some of the well-known factors causing LBP include intervertebral disc degeneration, facet joint arthritis, sacro-iliac joint dysfunction, and paravertebral muscle disorder. In this case, although the patient was receiving treatment for RA, arthritis of the facet joint or the sacro-iliac joint was not detected on ^{18}F-FDG-PET/CT. Furthermore, a high-intensity zone in the intervertebral discs, which is known to be correlated with LBP, was not seen on MRI. On the basis of these findings, the LBP in this case was considered to be caused by chronic fatigue in the paravertebral muscles, which was visualized with ^{18}F-FDG-PET/CT.

In conclusion, we present a case of painful degenerative kyphoscoliosis in which the paravertebral muscles at the convex region of the lumbar curves revealed an asymmetrically increased uptake of ^{18}F-FDG-PET/CT preoperatively. This study indicates the possibility of using ^{18}F-FDG-PET/CT to visualize the increased activity of painful paravertebral muscles in patients with ASD. Further investigation is mandatory to provide evidence of the usefulness of ^{18}F-FDG-PET/CT in the care and management of ASD patients.

Table 1 Results of the patient-oriented questionnaires

	Preoperative	Postoperative
NRS	8	0
RDQ	16 pts	0 pts
ODI	67 %	0 %

Each questionnaire was completed before the primary surgery and at 1 year after the revision surgery

NRS Numeric Rating Scale, *RDQ* Roland-Morris Disability Questionnaire, *ODI* Oswestry Disability Index

Compliance with ethical standards

Disclosure The authors declare that they have no conflict of interest concerning the materials or methods used in this study, or the findings specified in this paper.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Carter OD, Haynes SG. Prevalence rates for scoliosis in US adults: results from the first national health and nutrition examination survey. Int J Epidemiol. 1987;16(4):537–44.
2. Perennou D, Marcelli C, Herisson C, Simon L. Adult lumbar scoliosis: epidemiologic aspects in a low-back pain population. Spine (Phila Pa 1976). 1994;19(2):123–8.
3. Schwab F, Dubey A, Gamez L, et al. Adult scoliosis: prevalence, SF-36, and nutritional parameters in an elderly volunteer population. Spine (Phila Pa 1976). 2005;30(9):1082–5.
4. Baldus C, Bridwell KH, Harrast J, et al. Age-gender matched comparison of SRS instrument scores between adult deformity and normal adults: are all SRS domains disease specific? Spine (Phila Pa 1976). 2008;33(20):2214–8.
5. Baldus C, Bridwell K, Harrast J, et al. The scoliosis research society health-related quality of life (SRS-30) age-gender normative data:

an analysis of 1346 adult subjects unaffected by scoliosis. Spine (Phila Pa 1976). 2011;36(14):1154–62.
6. Berven S, Deviren V, Demir-Deviren S, Hu SS, Bradford DS. Studies in the modified scoliosis research society outcomes instrument in adults: validation, reliability, and discriminatory capacity. Spine (Phila Pa 1976). 2003;28(18):2164–9.
7. Weinstein SL, Dolan LA, Spratt KF, Peterson KK, Spoonamore MJ, Ponseti IV. Health and function of patients with untreated idiopathic scoliosis: a 50-year natural history study. JAMA. 2003;289(5):559–67.
8. Weinstein SL, Zavala DC, Ponseti IV. Idiopathic scoliosis: long-term follow-up and prognosis in untreated patients. J Bone Joint Surg Am. 1981;63(5):702–12.
9. Glassman SD, Bridwell K, Dimar JR, Horton W, Berven S, Schwab F. The impact of positive sagittal balance in adult spinal deformity. Spine (Phila Pa 1976). 2005;30(18):2024–9.
10. Takemitsu Y, Harada Y, Iwahara T, Miyamoto M, Miyatake Y. Lumbar degenerative kyphosis: clinical, radiological and epidemiological studies. Spine (Phila Pa 1976). 1988;13(11):1317–26.
11. Enomoto M, Ukegawa D, Sakaki K, et al. Increase in paravertebral muscle activity in lumbar kyphosis patients by surface electromyography compared with lumbar spinal canal stenosis patients and healthy volunteers. J Spinal Disord Tech. 2012;25(6):E167–73.
12. Williams M, Solomonow M, Zhou BH, Baratta RV, Harris M. Multifidus spasms elicited by prolonged lumbar flexion. Spine (Phila Pa 1976). 2000;25(22):2916–24.
13. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy. 1980;66(8):271–3.
14. Fujiwara A, Kobayashi N, Saiki K, Kitagawa T, Tamai K, Saotome K. Association of the Japanese Orthopaedic Sssociation score with the Oswestry disability index, Roland-Morris disability questionnaire, and short-form 36. Spine (Phila Pa 1976). 2003;28(14):1601–7.
15. Roland M, Morris R. A study of the natural history of back pain: part I development of a reliable and sensitive measure of disability in low-back pain. Spine (Phila Pa 1976). 1983;8(2):141–4.
16. Suzukamo Y, Fukuhara S, Kikuchi S, et al. Validation of the Japanese version of the roland-Morris disability questionnaire. J Orthop Sci. 2003;8(4):543–8.
17. Pichelmann MA, Lenke LG, Bridwell KH, Good CR, O'Leary PT, Sides BA. Revision rates following primary adult spinal deformity surgery: six hundred forty-three consecutive patients followed-up to twenty-two years postoperative. Spine (Phila Pa 1976). 2010;35(2):219–26.
18. Bar-Shalom R. Muscle uptake of 18-fluorine fluorodeoxyglucose. Semin Nucl Med. 2000;30(4):306–9.
19. Barrington SF, Maisey M. Skeletal muscle uptake of fluorine-18-FDG: effect of oral diazepam. J Nucl Med. 1996;37(7):1127–9.
20. Reinking MF, Osman MM. Prospective evaluation of physiologic uptake detected with true whole-body 18F-FDG PET/CT in healthy subjects. J Nucl Med Technol. 2009;37(1):31–7.
21. Shreve PD, Anai Y, Wahl RL. Pitfalls in oncologic diagnosis with FDG PET imaging: physiologic and benign variants. Radiographics. 1999;19(1):61–77.
22. Yeung HW, Grewal RK, Gonen M, Schoder H, Larson SM. Patterns of 18F-FDG uptake in adipose tissue and muscle: a potential source of false-positives for PET. J Nucl Med. 2003;44(11):1789–96.
23. Jackson RP, Simmons EH, Stripinis D. Coronal and sagittal plane spinal deformities correlating with back pain and pulmonary function in adult idiopathic scoliosis. Spine (Phila Pa 1976). 1989;14(12):1391–7.

PLOS ONE

RESEARCH ARTICLE

Psychometric Assessment of the Japanese Version of the Zurich Claudication Questionnaire (ZCQ): Reliability and Validity

Nobuhiro Hara[1], Ko Matsudaira[2]*, Kazuhiro Masuda[3], Juichi Tohnosu[4], Katsushi Takeshita[5], Atsuki Kobayashi[6], Motoaki Murakami[7], Naohiro Kawamura[8], Kiyohumi Yamakawa[9], Sei Terayama[10], Satoshi Ogihara[11], Hiroo Shiono[7], Jiro Morii[12], Keiji Hayakawa[13], So Kato[13], Kozo Nakamura[14], Hiroyuki Oka[2], Takayuki Sawada[15], Kyoko Inuzuka[15], Norimasa Kikuchi[15]

1 Department of Orthopaedic Surgery, Musashino Red Cross Hospital, Musashino, Tokyo, Japan, **2** Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, **3** Department of Orthopaedic Surgery, Tokyo Metropolitan Tama Medical Center, Fuchu, Tokyo, Japan, **4** Department of Orthopaedic Surgery, Kanto Rosai Hospital, Kawasaki, Kanagawa, Japan, **5** Department of Orthopaedic Surgery, Jichi Medical University, Shimotsuke, Tochigi, Japan, **6** Department of Orthopaedic Surgery, Tokyo Metropolitan Bokutoh Hospital, Sumida-ku, Tokyo, Japan, **7** Department of Orthopaedic Surgery, Toranomon Hospital, Minato-ku, Tokyo, Japan, **8** Department of Spine and Orthopaedic Surgery, Japanese Red Cross Medical Center, Shibuya-ku, Tokyo, Japan, **9** Department of Orthopaedic Surgery, Tokyo Metropolitan Komagome Hospital, Bunkyo-ku, Tokyo, Japan, **10** Department of Orthopaedic Surgery, Sangubashi Spine Surgery Hospital, Shibuya-ku, Tokyo, Japan, **11** Department of Orthopaedic Surgery, Sagamihara National Hospital, Sagamihara, Kanagawa, Japan, **12** Department of Orthopaedic Surgery, Sanraku Hospital, Chiyoda-ku, Tokyo, Japan, **13** Department of Orthopaedic Surgery, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, **14** National Rehabilitation Center for Persons with Disabilities, Tokorozawa, Saitama, Japan, **15** Clinical Study Support, Inc., Nagoya, Aichi, Japan

* kohart801@gmail.com

OPEN ACCESS

Citation: Hara N, Matsudaira K, Masuda K, Tohnosu J, Takeshita K, Kobayashi A, et al. (2016) Psychometric Assessment of the Japanese Version of the Zurich Claudication Questionnaire (ZCQ): Reliability and Validity. PLoS ONE 11(7): e0160183. doi:10.1371/journal.pone.0160183

Editor: Giovanni Grasso, Universita degli Studi di Palermo, ITALY

Received: March 1, 2016

Accepted: July 14, 2016

Published: July 28, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information file.

Funding: The authors received no specific funding for this work. N. Kikuchi is a board member of Clinical Study Support, Inc. Co-authors TS and KI are employed by Clinical Study Support, Inc. Clinical Study Support, Inc. provided support in the form of salaries for authors TS and KI, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Abstract

Purpose

The Zurich Claudication Questionnaire (ZCQ) is a self-administered measure to evaluate symptom severity, physical function, and surgery satisfaction in lumbar spinal stenosis (LSS). The purpose of this study is to assess the psychometric properties of the Japanese ZCQ in LSS patients.

Methods

LSS patients who are scheduled to undergo surgery were recruited from 12 facilities. Responses to several questionnaires, including the Japanese ZCQ; the visual analogue scale (VAS) to evaluate the degree of pain in the buttocks/legs, numbness in the buttocks/legs, and low back pain; the Oswestry Disability Index (ODI); and the SF-36v2, were collected before surgery and again 3 months after surgery (the post-surgery ZCQ was administered twice for test-retest reliability). For reliability, test-retest reliability was evaluated using the intra-class coefficient (ICC) and internal consistency was evaluated using Cronbach's alpha coefficient. Concurrent validity was assessed using Spearman's correlation

Competing Interests: The authors have read the journal's policy and the authors of this manuscript have the following competing interests: HO has received grants to his institution from Pfizer, Inc. N. Kikuchi is a board member of Clinical Study Support, Inc. TS and KI are employed by Clinical Study Support, Inc. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials. NH, K. Matsudaira, K. Masuda, JT, KT, AK, MM, N. Kawamura, KY, ST, SO, HS, JM, KH, SK, and KN have no conflict of interests to declare. There are no patents, products in development or marketed products to declare.

coefficients between the Japanese ZCQ and other questionnaires. Effect size (ES) and standard response mean were calculated for responsiveness. All analyses were performed individually for the Japanese ZCQ symptom, function, and satisfaction domains.

Results

Data from 180 LSS patients were used in this analysis. The ICCs were 0.81, 0.89, and 0.88 and Cronbach's alpha coefficients were 0.78, 0.84, and 0.92 for the Japanese ZCQ symptom, function, and satisfaction domains, respectively. Regarding the concurrent validity, strong correlations (±0.5) were demonstrated between the Japanese ZCQ domains and the VAS leg pain, ODI, and SF-36v2 physical functioning or bodily pain, whereas correlations were approximately 0.3 in scales measuring other symptoms that are less related to symptom, function, or satisfaction domains. ESs showed high values for the ZCQ symptom and function domains (-1.73 for both).

Conclusions

These psychometric assessments demonstrate that the Japanese ZCQ is a psychometrically reliable and valid measure in LSS. The Japanese ZCQ can evaluate both multi-dimensional aspects and the level of surgery satisfaction.

Introduction

Lumbar spinal stenosis (LSS) is a degenerative disorder that is characterized by a narrowing of the lumbar spinal canal, which entraps and compresses intraspinal vascular and nerve structures [1]. LSS results in neurological symptoms in the lower extremities, such as leg pain/numbness and gait disturbance, that dramatically deteriorate the patients' quality of life [2–4]. Conservative therapy is the primary treatment for LSS, and surgery is considered for LSS patients who do not improve [5]. Because pain or numbness is the primary complaint in LSS, the patient outcome measures have an important role in evaluating the treatment outcome.

Various outcome measures, such as the visual analogue scale (VAS) and Oswestry Disability Index (ODI) [6, 7], are used in research on LSS patients, but these measures are not disease specific. The Zurich Claudication Questionnaire (ZCQ), which is also known as both the Swiss Spinal Stenosis Measure and the Brigham Spinal Stenosis Questionnaire, was developed as a self-administered measure to assess symptom severity, physical function, and surgery satisfaction in LSS patients [8]. The questionnaire consists of three domains and uses a Likert-type scale. It includes 7 items for symptom severity with scores of 1 to 5, 5 items for functional disability with scores of 1 to 4, and 6 items for treatment satisfaction with score of 1 to 4. Higher scores indicate more severe LSS. The ZCQ demonstrates good validity and reliability in patients with LSS and is recommended as one of the appropriate methods for evaluating LSS treatment outcomes [9]. The ZCQ has been used worldwide in many studies on LSS [10–12].

To allow the use of the ZCQ in Japan, the English version was translated and linguistically validated as the Japanese ZCQ [13] following international guidelines [14, 15], but the psychometric validation has not yet been conducted.

The purpose of this study is to assess the psychometric properties of the Japanese ZCQ in LSS patients.

Materials and Methods

The study was approved by the Ethics Committee at the University of Tokyo. All patients who were enrolled in the study had provided written informed consent.

Participants

LSS patients between 20 and 85 years of age who were scheduled to undergo surgery were recruited. The inclusion criteria were as follows: 1) the presence of neurogenic intermittent claudication caused by numbness and/or pain in the lower limbs and 2) magnetic resonance imaging-confirmed symptomatic LSS that might explain the patient's symptoms. The exclusion criteria were as follows: 1) a positive straight leg raising test result (sciatic pain at < 70 degrees of leg elevation), which indicates that the pain is likely to be due to lumbar disc herniation; 2) presentation with lower-extremity peripheral arterial disease; 3) a history of spinal surgery; 4) complications causing disorders that interfere with gait, such as myelopathy; 5) peripheral neuropathy, such as diabetic neuropathy of the leg; or 6) disorders that potentially hinder gait other than LSS (e.g., rheumatoid arthritis). The study was conducted in 2010 at the University of Tokyo and 11 affiliated facilities, all of which are located in or near Tokyo.

Measures

Questionnaires were administered a maximum of 3 times: 1) before surgery, 2) 3 months after surgery, and 3) a few weeks after the second questionnaire if symptoms had not changed. The questionnaires administered before and 3 months after surgery included the Japanese versions of the ZCQ, ODI [6, 7], VAS for back/leg pain and leg numbness, and the SF-36 Ver.2 (SF-36v2). The ODI is a principal, condition-specific outcome measure used to assess disability from spinal disorders, particularly low back pain (LBP). The ODI consists of 10 items: pain intensity, personal care, lifting, walking, sitting, standing, sleeping, sex life, social life, and traveling. Each item is scored on a 6-point Likert-type scale with scores ranging from 0 to 5, and a higher score indicates a more severe disability. The reliability and validity of the Japanese version of the ODI were previously confirmed [16].

The degree of pain associated with buttock/leg pain or LBP was measured using a VAS covering the week prior to the relevant visit. We included three items: the degree of pain in the buttocks/legs, the degree of numbness in the buttocks/legs, and the degree of LBP. The scale ranged from 0 (no pain at all) to 10 (the worst pain/numbness imaginable).

The SF-36v2 is a questionnaire containing 36 items to assess the general health-related quality of life (QOL) [17]. The items are categorized into 8 domains: physical function, role limitations-physical, vitality, general health perception, bodily pain, social function, role limitations-emotional, and mental health. Each domain is scored from 0 to 100, with a higher score indicating a better QOL. A Japanese version of the SF-36v2, which has demonstrated good reliability and validity, was used in this study [18, 19].

Statistical analysis

Demographic and clinical characteristics of the participants were analysed descriptively. The Japanese ZCQ domain scores were summarized to examine missing data and the distribution. The scoring for each domain was carried out in the same way it would be for the English version of the ZCQ [8].

Psychometric properties of the Japanese ZCQ were assessed by evaluating reliability and validity. Responsiveness was also assessed. Reliability was evaluated by test-retest reliability and internal consistency. For test-retest reliability, the extent of agreement between two time

points was examined using the intra-class correlation coefficient (ICC) in patients with stable symptoms after spine surgery. The coefficient ranged from 0 to 1, with a higher value showing increased reliability. A coefficient greater than or equal to 0.7 was considered sufficient to determine test-retest reliability [20].

For internal consistency, the homogeneity of the items within the domain was evaluated using Cronbach's alpha coefficients of the pre-surgery responses for the symptom and function domains and of the post-surgery responses for the satisfaction domain. A Cronbach's alpha of 0.7 or higher was considered acceptable for internal consistency, while a score above 0.8 was good and above 0.9 was excellent [21].

For concurrent validity, the degree of correlation with the external criteria (ODI, VAS, and SF-36v2) was assessed using Spearman's correlation coefficient of the pre-surgery responses for the symptom and function domains and of the post-surgery responses for the satisfaction domain. Scales measuring similar concepts were expected to show a moderate to strong correlation, while those measuring different concepts were expected to show a weak correlation. For example, the VASs of pain in the buttocks/legs and numbness in the buttocks/legs were expected to correlate strongly with the Japanese ZCQ symptom domain, the ODI with the ZCQ function domain, and the SF-36v2 social functioning or vitality with the ZCQ satisfaction domain. The correlation coefficient was interpreted as follows: ±0.1 was considered weak, ±0.3 was considered moderate, and ±0.5 was considered to be a strong correlation [21].

Responsiveness was evaluated by the effect sizes (ESs) and standard response means (SRMs). The ES was obtained by calculating the mean change in scores from before to 3 months after surgery divided by the standard deviation of the pre-surgery score. An ES of 0.2 was considered small, 0.5 was moderate and 0.8 was large, following the guidelines proposed by Cohen [22]. The SRM was obtained by the mean change in scores from before to 3 months after surgery divided by the standard deviation of the mean change. The higher the ES or SRM, the greater was the level of sensitivity to detect change.

All statistical tests were two sided with a significance level of 5%. All analyses were performed using SAS release 9.3 (SAS Institute, Cary, NC, USA).

Results

Patient characteristics

A total of 195 participants were recruited for this study. Of those, 180 took part in the first questionnaire administration before surgery, and 135 provided answers after surgery. Demographic and clinical characteristics of the recruited patients at pre-surgery are summarized in Table 1. The mean (standard deviation, SD) age was 68.2 (9.9) years, and 57.8% of the patients were male. The mean (SD) duration of LSS was approximately 3.7 (4.6) years. Types of symptom and surgery were evenly distributed. Approximately half the patients had spondylosis, followed by degenerative spondylolisthesis as the second most common diagnosis.

The symptom, function, and satisfaction scores at pre- and post-surgery are summarized in Table 2. The mean and median were similar in all 3 domains at both time points. The mean or median scores of the symptom and function domains at 3 months after surgery are smaller than at pre-surgery.

Reliability

To analyse test-retest reliability, 30 participants who underwent surgery and answered the Japanese ZCQ twice were selected. The ICCs for the Japanese ZCQ symptom, function, and satisfaction domains were 0.81, 0.89, and 0.88, respectively.

Table 1. Demographic and clinical characteristics of the lumber spinal stenosis patients (n = 195).

Characteristics	n (%) or mean (SD)
Age (years)	68.2 (9.9)
Sex	
Male	111 (57.8)
Female	81 (42.2)
Disease duration (months)	43.9 (55.7)
Types of symptom	
Nerve root	56 (29.3)
Cauda equine	65 (34.0)
Mixed	70 (36.7)
Surgery type	
Decompression only	92 (47.2)
Decompression and fusion	103 (52.8)
Types of lumber spinal stenosis	
Spondylosis	98 (50.8)
Degenerative spondylolisthesis	73 (37.8)
Degenerative scoliosis	18 (9.3)
Isthmic spondylolisthesis	4 (2.1)
Working status	
At work	65 (33.7)
Out of work	128 (66.3)
Smoking	
Smokers	47 (24.2)
Non-smokers	147 (75.8)

Values are n (%) or mean (SD).

Not all groups above total 195 because some of the characteristics had missing values.

doi:10.1371/journal.pone.0160183.t001

The internal consistency was evaluated using the data collected from patients who replied to the Japanese ZCQ at pre-surgery for the symptom and function domains and at post-surgery for the satisfaction domain. Cronbach's alpha coefficients for the Japanese ZCQ symptom, function, and satisfaction domains were 0.78, 0.84, and 0.92, respectively.

Validity

To assess the concurrent validity, the correlation coefficients between the 3 domains of the Japanese ZCQ and the external criteria (VAS, ODI, and SF-36v2) were calculated (Table 3). All 3

Table 2. Distribution of the Zurich Claudication Questionnaire (ZCQ) subscales.

ZCQ Subscale	Pre-surgery (n = 180)*		3 months after surgery (n = 135)*	
	Mean (SD)	Median (minimum, maximum)	Mean (SD)	Median (minimum, maximum)
Symptom	3.41 ± 0.67	3.43 (1.57, 5.00)	2.25 ± 0.75	2.14 (1.00, 4.71)
Function	2.70 ± 0.57	2.80 (1.00, 3.75)	1.71 ± 0.66	1.50 (1.00, 3.60)
Satisfaction	–	–	1.97 ± 0.72	1.83 (1.00, 4.00)

Response of satisfaction was not obtained at pre-surgery. SD: Standard Deviation

*: Due to missing responses, domain scores for the symptom and function domains at 3 months after surgery could not be computed for one patient, and domain scores for the satisfaction domain at 3 months after surgery could not be computed for 2 patients.

doi:10.1371/journal.pone.0160183.t002

Table 3. Spearman's correlation coefficients (95% confidence interval) between the ZCQ and the ODI, VAS, and SF-36v2.

Measure	ZCQ Symptom (n = 180)	ZCQ Function (n = 180)	ZCQ Satisfaction (n = 135)
ZCQ symptom	–	0.63 (0.53, 0.71)	0.79 (0.72, 0.85)
ZCQ function	0.63 (0.53, 0.71)	–	0.73 (0.63, 0.80)
ZCQ satisfaction	–	–	–
VAS leg pain	0.50 (0.38, 0.60)	0.50 (0.38, 0.61)	0.73 (0.63, 0.80)
VAS leg numbness	0.58 (0.46, 0.67)	0.49 (0.36, 0.60)	0.67 (0.56, 0.76)
VAS low back pain	0.48 (0.35, 0.59)	0.42 (0.29, 0.54)	0.58 (0.44, 0.69)
ODI	0.63 (0.53, 0.71)	0.75 (0.67, 0.80)	0.74 (0.65, 0.80)
SF-36 physical functioning	-0.56 (-0.66, -0.45)	-0.62 (-0.70, -0.52)	-0.62 (-0.71, -0.50)
SF-36 role limitations-physical	-0.46 (-0.57, -0.34)	-0.51 (-0.61, -0.39)	-0.57 (-0.68, -0.44)
SF-36 bodily pain	-0.59 (-0.68, -0.48)	-0.63 (-0.71, -0.53)	-0.65 (-0.73, -0.53)
SF-36 social functioning	-0.44 (-0.55, -0.31)	-0.38 (-0.50, -0.24)	-0.61 (-0.71, -0.49)
SF-36 general health	-0.31 (-0.43, -0.17)	-0.28 (-0.41, -0.14)	-0.56 (-0.67, -0.43)
SF-36 vitality	-0.35 (-0.47, -0.21)	-0.41 (-0.52, -0.28)	-0.59 (-0.69, -0.46)
SF-36 role limitations-emotional	-0.46 (-0.57, -0.33)	-0.50 (-0.60, -0.38)	-0.50 (-0.61, -0.35)
SF-36 mental health	-0.31 (-0.44, -0.17)	-0.33 (-0.45, -0.19)	-0.54 (-0.65, -0.41)

Calculations of the ZCQ symptom and function domains were performed with the ZCQ responses at pre-surgery and those of the ZCQ satisfaction domain at 3 months after surgery.

95% Confidence interval is shown as lower and upper values. P < 0.0001 for all.

ZCQ, Zurich Claudication Questionnaire; ODI, Oswestry Disability Index; VAS, Visual Analogue Scale; SF-36, 36-Item Short-Form Health Survey Ver. 2.

doi:10.1371/journal.pone.0160183.t003

domains of the Japanese ZCQ correlated well with one another. Strong correlations with all 3 domains were observed in VAS leg pain by values of approximately 0.50–0.73, in ODI (0.63–0.75), and in physical functioning (-0.56–0.62) or bodily pain (-0.59–0.65) in the SF-36v2. Regarding the VAS, the coefficients of leg pain or leg numbness were generally higher than for the low back pain in all 3 Japanese ZCQ domains. With regard to each domain on the SF-36v2, strong correlations were observed between the Japanese ZCQ symptom domain and bodily pain or physical functioning; the ZCQ function domain and physical functioning, role limitations-physical, bodily pain, or role limitations-emotional; and the ZCQ satisfaction domain and all 8 SF-36v2 domains. Overall, all three Japanese ZCQ domains were correlated to all external criteria to a moderate to strong degree. However, all of the correlations between the Japanese symptom and function domains and the general health and mental health of the SF-36v2 were smaller: approximately 0.3.

Responsiveness

To assess the responsiveness, the ESs between the Japanese ZCQ and external criteria (ODI, VAS, and SF-36v2) were calculated (Table 4). The ES was highest in the Japanese ZCQ function domain and symptom domain, followed by VAS leg pain, SF-36v2 bodily pain, VAS leg numbness, ODI, and VAS LBP, which were all above 0.8. Japanese ZCQ symptom and function domains had the two highest SRMs among the measures in this study (1.54 and 1.38, respectively).

Discussion

The Japanese ZCQ was translated and linguistically validated prior to this study [13]. As a next step for developing a valid and reliable measure, the psychometric properties of the ZCQ were

Table 4. Responsiveness of outcome measures in the study.

Measure	Change at 3 months after surgery from before surgery		ES	SRM
	mean	SD		
ZCQ symptom	-1.16	0.75	-1.73	-1.54
ZCQ function	-1.01	0.71	-1.73	-1.38
ODI	-22.69	18.61	-1.22	-1.20
VAS leg pain	-3.86	3.41	-1.64	-1.20
VAS leg numbness	-3.37	2.98	-1.28	-1.20
VAS low back pain	-3.35	3.24	-1.15	-1.02
SF-36 physical functioning	12.98	16.89	0.74	0.83
SF-36 role limitations-physical	6.56	18.28	0.44	0.39
SF-36 bodily pain	12.52	13.34	1.44	0.91
SF-36 social functioning	6.14	18.61	0.40	0.36
SF-36 general health	4.40	8.08	0.45	0.52
SF-36 vitality	8.79	12.36	0.72	0.68
SF-36 role limitations-emotional	6.75	20.41	0.42	0.35
SF-36 mental health	7.38	14.03	0.54	0.52

Responsiveness of the ZCQ satisfaction domain was not calculated because a response on satisfaction was not obtained prior to surgery.

SD, Standard Deviation; ES, Effect Size; SRM, Standard Response Mean; ZCQ, Zurich Claudication Questionnaire; ODI, Oswestry Disability Index; VAS, Visual Analogue Scale; SF-36, 36-Item Short-Form Health Survey Ver. 2.

doi:10.1371/journal.pone.0160183.t004

assessed using the data collected from Japanese LSS patients. Based on the results of the current assessments, the Japanese ZCQ shows good validity and reliability. The responsiveness is also shown to be specific to LSS compared to other measures, such as the ODI or SF-36v2.

The ICCs for the 3 Japanese ZCQ domains (symptom, function, and surgery satisfaction) all satisfied the level of 0.7. The ICC was highest in the function domain and lowest in the symptom domain, but all were approximately 0.9. We set a satisfactory level of 0.7 for research use, but an ICC higher than 0.9 could be considered satisfactory for clinical practice in test-retest reliability [20, 21]. The ICC range was similar to other language versions, such as the original English (0.92), Norwegian (0.89–0.92), and simplified Chinese (0.91–0.95) versions [8, 23, 24].

Cronbach's alpha coefficients showed good to excellent levels of internal consistency for all 3 domains, and these were approximately 0.9 in the function and surgery satisfaction domains. Although a Cronbach's alpha of 0.7 is considered satisfactory for psychometric assessments, in clinical practice, values above 0.9 are considered suitable [20, 21]. Therefore, this measure is sufficient for application in clinical practice. Moreover, the range of Cronbach's alpha coefficients was similar to that in other language versions, including the original English (0.84–0.89), Norwegian (0.94–0.96), Iranian (0.88), and simplified Chinese (0.86–0.91) versions [8, 23–25]. On the basis of these reliability results, the Japanese ZCQ showed sufficient reliability.

The concurrent validity assessment showed moderate to strong agreement with the external criteria (ODI, VAS, and 8 domains of the SF-36v2). Scales measuring similar concepts with each Japanese ZCQ domain showed moderate to strong correlations, such as the correlation between the ZCQ symptom domain and the bodily pain scale of the SF-36v2. Similar findings were also observed in the original English and simplified Chinese versions [8, 24]. All 3 Japanese ZCQ domains correlated strongly with one another, with the highest correlation observed between the symptom and satisfaction domains. By contrast, scales that measure different

concepts showed smaller correlation coefficients, such as that between the Japanese ZCQ symptom and function domains and the general health scale of the SF-36v2. The correlation of the satisfaction domain was calculated using data obtained after surgery, when the mean changes in both symptom and function improved by a score of 1. This implies that patients in this study may have been satisfied when both their symptoms and function improved.

The Japanese ZCQ showed good responsiveness in both the symptom and function domains, and it showed better responsiveness than other measures, such as the ODI, which is commonly used to evaluate disability. The trend regarding responsiveness was similar to the results of other languages. For example, the ES and SRM of the symptom or function domain were between those of the original English (SRM = 1.48 and 1.6 in patients who were satisfied with surgery) and Norwegian (ES = 1.9 and 1.2) versions [23]. The current results showed that Japanese ZCQ symptom and function domains reflect changes in the post-operative condition of LSS in Japanese patients with a high degree of sensitivity. Because this measure can be used for multi-dimensional evaluations, is LSS specific, and has advantages in its simplicity and easy-to-answer format, the Japanese ZCQ may be useful for the elderly, who make up the majority of LSS patients. In addition, this questionnaire may enable better communication between the physician and the patient because sharing the responses of the patients evaluated in this study may enhance patient compliance with treatments.

There are several limitations of this study that should be mentioned. This study was carried out in Tokyo and its outlying areas. Because more than 10% of the Japanese population resides in this area, the study may over-represent the urban Japanese population. Therefore, we should consider the population of other suburban areas in Japan. Second, some physicians might have asked patients to complete the questionnaire in front of them and seen the responses the patients selected. This may have resulted in a bias in the responses being close to physicians' expectations. Third, we did not assess known-group validity; i.e., score changes between the groups in terms of severity. However, we were able to show good responsiveness of the symptom and function domains in this study. Lastly, the presence of dynamic instability was not assessed in this study, although it plays an important role in the decision of treatment/surgery in a clinical setting. Patients' responses to the questionnaire might have been influenced by the types of surgery/treatment; however, as this was a psychometric assessment study of the Japanese ZCQ, which will be widely used in patients with LSS, we consider that it is not a methodological problem to include patients with LSS regardless of the presence of dynamic instability.

Conclusion

The current psychometric assessments have demonstrated that the Japanese ZCQ is psychometrically reliable and valid measure in LSS. The Japanese ZCQ, which includes symptom and function domains, can evaluate multi-dimensional aspects along with the level of surgery satisfaction.

Supporting Information

S1 File. Supporting information. Dataset of this study.
(XLS)

Author Contributions

Conceived and designed the experiments: NH K. Matsudaira. Performed the experiments: NH K. Matsudaira K. Masuda JT KT AK MM N. Kawamura KY ST SO HS JM KH SK KN HO. Analyzed the data: NH TS. Wrote the paper: KI N. Kikuchi K. Matsudaira.

References

1. Amundsen T, Weber H, Nordal HJ, Magnaes B, Abdelnoor M, Lilleâs F. Lumbar spinal stenosis: conservative or surgical management?: A prospective 10-year study. Spine. 2000; 25: 1424–35; discussion 1435–6. PMID: 10828926
2. Berthelot JM, Bertrand VA, Rodet D, Maugars Y, Prost A. Lumbar spinal stenosis: a review. Revue du rhumatisme (English ed.). 1997; 64: 315–25.
3. Bolender NF, Schonstrom NSR, Spengler DM. Role of computed tomography and myelography in the diagnosis of central spinal stenosis. J Bone Joint Surg Am. 1985; 67: 240–6. PMID: 3968115
4. Takahashi K, Miyazaki T, Takino T, Matsui T, Tomita K. Epidural pressure measurements. Relationship between epidural pressure and posture in patients with lumbar spinal stenosis. Spine. 1995; 20: 650–653. PMID: 7604339
5. Mazanec DJ, Podichetty VK, Hsia A. Lumbar canal stenosis: start with nonsurgical therapy. Cleve Clin J Med. 2002; 69: 909–17. PMID: 12430977
6. Fairbank J. Use of Oswestry disability indez (ODI). Spine. 1995; 20: 1535–1537. PMID: 8623078
7. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy. 1980; 66: 271–273. PMID: 6450426
8. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. Spine. 1996; 21: 796–803. PMID: 8779009
9. North American Spine Society (NASS). Clinical Guidelines for Multidisciplinary Spine Care. Diagnosis and Treatment of Degenerative Lumbar Spinal Stenosis. Burr Ridge (IL): North American Spine Society (NASS); 2007.
10. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ, et al. A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. Eur Spine J. 2004; 13: 22–31. PMID: 14685830
11. Steurer J, Nydegger A, Held U, Brunner F, Hodler J, Porchet F. et al.; LumbSten Research Collaboration. LumbSten: the lumbar spinal stenosis outcome study. BMC Musculoskelet Disord. 2010; 11: 254. doi: 10.1186/1471-2474-11-254 PMID: 21044326
12. Moojen WA, Arts MP, Brand R, Koes BW, Peul WC. The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. BMC Musculoskelet Disord. 2010; 11: 100. doi: 10.1186/1471-2474-11-100 PMID: 20507568
13. Hara N, Matsudaira K, Terayama S, Takeshita K, Isomura T, Nakamura K. Development of the Japanese version of the Zurich claudication questionnaire (ZCQ): translation and linguistic validation. Seikei Geka. 2010; 61: 159–165. Japanese.
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000; 25: 3186–3191. PMID: 11124735
15. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality-of-life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993; 46: 1417–1432. PMID: 8263569
16. Fujiwara A, Kobayashi N, Saiki K, Kitagawa T, Tamai K, Saotome K. Association of the Japanese Orthopaedic Association score with the Oswestry Disability Index, Roland-Morris Disability Questionnaire, and short-form 36. Spine. 2003; 28: 1601–7. PMID: 12865852
17. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992; 30: 473–83. PMID: 1593914
18. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. J Clin Epidemiol. 1998; 51: 1037–44. PMID: 9817121
19. Fukuhara S, Ware JE, Kosinski M, Wada S, Gandek B. Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. J Clin Epidemiol. 1998; 51: 1045–53. PMID: 9817122
20. Fayers PM, Machin D. Quality of Life: The Assessment, Analysis, and Interpretation of Patient-reported Outcomes. 2nd ed. West Sussex: John Wiley & Sons Ltd; 2007.
21. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use. 4th ed. Oxford: Oxford University Press; 2008.
22. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
23. Thornes E, Grotle M. Cross-cultural adaptation of the Norwegian version of the spinal stenosis measure. Eur Spine J. 2008; 17: 456–62. doi: 10.1007/s00586-007-0576-7 PMID: 18193302

24. Yi H, Wei X, Zhang W, Chen Z, Wang X, Ji X, et al. Reliability and validity of simplified Chinese version of Swiss Spinal Stenosis Questionnaire for patients with degenerative lumbar spinal stenosis. Spine. 2014; 39: 820–5. doi: 10.1097/BRS.0000000000000273 PMID: 24525991

25. Azimi P, Ghandehari HS, Sadeghi S, Azhari S, Aghaei HN, Mohmmadi HR, et al. Severity of symptoms, physical functioning and satisfaction in patients with lumbar spinal stenosis: a validation study of the Iranian version of the Swiss Spinal Stenosis Score. J Neurosurg Sci. 2014; 58: 177–82. PMID: 25033977

PLOS ONE

RESEARCH ARTICLE

The Association between the Cross-Sectional Area of the Dural Sac and Low Back Pain in a Large Population: The Wakayama Spine Study

Hiroki Iwahashi[1], Noriko Yoshimura[2], Hiroshi Hashizume[1]*, Hiroshi Yamada[1], Hiroyuki Oka[3], Ko Matsudaira[3], Kazunori Shinto[1], Yuyu Ishimoto[1], Keiji Nagata[1], Masatoshi Teraguchi[1], Ryohei Kagotani[1], Shigeyuki Muraki[2], Toru Akune[4], Sakae Tanaka[5], Hiroshi Kawaguchi[6], Kozo Nakamura[4], Akihito Minamide[1], Yukihiro Nakagawa[1], Munehito Yoshida[1]

1 Department of Orthopaedic Surgery, Wakayama Medical University, 811–1 Kimiidera, Wakayama City, Wakayama 641–8510, Japan, **2** Department of Joint Disease Research, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113–8655, Japan, **3** Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113–8655, Japan, **4** Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, 1 Namiki 4-chome, Tokorozawa City, Saitama 359–8555, Japan, **5** Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113–8655, Japan, **6** Japan Community Health Care Organization Tokyo Shinjuku Medical Center, 6-1-1 Shinjuku, Shinjuku-ku, Tokyo 160–8402, Japan

* hashizum@wakayama-med.ac.jp

Citation: Iwahashi H, Yoshimura N, Hashizume H, Yamada H, Oka H, Matsudaira K, et al. (2016) The Association between the Cross-Sectional Area of the Dural Sac and Low Back Pain in a Large Population: The Wakayama Spine Study. PLoS ONE 11(8): e0160002. doi:10.1371/journal.pone.0160002

Editor: Masahiko Sumitani, The University of Tokyo Hospital, JAPAN

Received: February 21, 2016

Accepted: July 12, 2016

Published: August 3, 2016

Data Availability Statement: The present study used resident data from two communities in Wakayama prefecture. It is impossible for us to provide and upload these data in a public repository because we have confirmed with these municipalities and residents that data will remain confidential. We will provide anonymized data on request after discussing the contents with the municipalities, as long as researchers are qualified to request these data. Data requests can be made to the corresponding author at hashizum@wakayama-med.ac.jp

Abstract

Objective

The purpose of this study was to evaluate the relations between the degree of encroachment, measured as the cross-sectional area of the dural sac, and low back pain in a large population.

Methods

In this cross-sectional study, data from 802 participants (247 men, 555 women; mean age, 63.5 years) were analyzed. The measurement of the cross-sectional area of the dural sac from the level of L1/2 to L4/5 was taken using axial T2-weighted images. The minimum cross-sectional area was defined as the cross-sectional area of the dural sac at the most constricted level in the examined spine. Participants were divided into three groups according to minimum cross-sectional area measurement quartiles (less than the first quartile, between the first and third quartiles, and greater than the third quartile). A multivariate logistic regression analysis was used to estimate the association between the minimum cross-sectional area and the prevalence of low back pain.

Results

The mean minimum cross-sectional area was 117.3 mm^2 (men: 114.4 mm^2; women: 118.6 mm^2). A logistic regression analysis adjusted for age, sex, body mass index, and other

Funding: This study was supported by H23-Choujyu-002 (Director, TA), H-25-Choujyu-007 (Director, NY), H25-Nanchitou (Men)-005 (Director, ST), 201417014A (Director, NY), and H22-Choujyu-Wakate-007 (Director, SM) from the Ministry of Health, Labour and Welfare, [URL:http://www.mhlw.go.jp/]; a Grant-in-Aid for Scientific Research (B26293139, B23390172 to NY, B2629333, C20591774 to SM, C26462249 to HH, C25462305 to HY) and a Grant-in-Aid for Young Researcher (B25860448 to YI, B26861286 to MT, B26860419 to RK, B15K20013 to HI); and Grant-in-Aid for Challenging Exploratory Research (15K15219 to NY, 26670307 to SM, 24659666 to HO, 25670293 to TA) of JSPS KAKENHI grant, [URL:https://www.jsps.go.jp/j-grantsinaid/]; a Grant from the Japanese Orthopaedics and Traumatology Foundation, Inc. (No. 287) to MT, [URL:http://jotf.jp/invitation.html]; and Collaborating Research with NSF 08033011-00262 (Director, NY) from the Ministry of Education, Culture, Sports, Science and Technology in Japan, [URL: http://www.mext.go.jp/english/]. This study also was supported by grants from the Japan Osteoporosis Society (NY, SM, HO, and TA), [URL:http://www.josteo.com/ja/index.html]; a grant from JA Kyosai Research Institute (HO), [URL:http://www.jkri.or.jp/]; grants from Mitsui Sumitomo Insurance Welfare Foundation (SM), [URL:http://www.ms-ins.com/welfare/index.htm]; research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 & 2010-2; Director, HK), [URL:https://www.joa.or.jp/english/english_frame.html]; and Japan Society for the Promotion of Science a Grant-in-Aid for Scientific Research (C26462249) to Hiroshi Hashizume. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

confounding factors, including disc degeneration, showed that a narrow minimum cross-sectional area (smaller than the first quartile) was significantly associated with low back pain (odds ratio, 1.78; 95% confidence interval, 1.13–2.80 compared to the wide minimum cross-sectional area group: minimum cross-sectional area greater than the third quartile measured).

Conclusion

This study showed that a narrow dural sac cross-sectional area was significantly associated with the presence of low back pain after adjustment for age, sex, and body mass index. Further investigations that include additional radiographic findings and psychological factors will continue to elucidate the causes of low back pain.

Introduction

Low back pain (LBP) is a multifactorial symptom, a common cause of morbidity and disability, and was reported to have a prevalence of 28.5% in a recent study [1,2]. There are many causes of chronic LBP, one of which is lumbar spinal stenosis (LSS) [3]. According to the Evidence-Based Clinical Guidelines for Multidisciplinary Spine Care developed by The North American Spine Society [4], degenerative LSS describes a condition in which there is diminished space available for the neural and vascular elements in the lumbar spine secondary to degenerative changes in the spinal canal. When symptomatic, this condition causes a variable clinical syndrome of gluteal and/or lower-extremity pain and/or fatigue, which may occur with or without back pain. In reality, however, 67.5%–95% of patients with LSS experience LBP [5–7].

LBP in patients with LSS is also multifactorial. Patients with LSS often have facet arthrosis and degenerative discs. These pathologies may explain their back pain. Earlier findings from preoperative imaging studies of patients with central spinal stenosis have suggested that the cross-sectional area (CSA) of the dural sac was closely related to preoperative walking ability, health-related quality of life, leg pain, and LBP [8–10]. Recent studies have also reported the possibility of improving LBP following decompression surgery [11,12]. Thus, it is possible that constriction of the dural sac is also the cause of LBP in LSS patients.

However, no research to date has focused on the association between the prevalence of LBP and the cross-sectional area of the dural sac in the general population. Thus, the purpose of this study was to evaluate the relations between the degree of encroachment, measured as the CSA of the dural sac, and LBP in a large population.

Methods

Study design

We performed a cross-sectional, population-based study.

Participants

The present study design was approved by the Wakayama Medical University Ethics Committee. All participants provided their written informed consent. The Wakayama Spine Study is a population-based study of degenerative spinal disease [13–17] conducted with a sub-cohort of the large-scale, population-based cohort study Research on Osteoarthritis/Osteoporosis against Disability (ROAD) [18,19]. ROAD is a nationwide, prospective study of bone and joint diseases

consisting of population-based cohorts established in three communities in Japan. The participants were recruited from listings of resident registrations in three communities that have different characteristics: an urban region in I town, Tokyo; a mountainous region in H town, Wakayama; and a coastal region in T town Wakayama. The inclusion criteria, apart from residing in those communities, included the ability to walk to the survey site, to report data, and to understand and sign an informed consent form. No other exclusion criteria were used. A third visit of the ROAD study began in 2012 and was completed in 2013. From the volunteers participating in the third visit of the ROAD study, 1575 individuals (513 men, 1062 women), which included 718 individuals in the mountainous area and 857 individuals in the coastal area, were recruited to the second visit of the Wakayama Spine Study. Magnetic resonance imaging (MRI) was conducted only to the individuals in the coastal area because of the funding limitation. Thus, we evaluated data from 857 individuals in the coastal area for the present study. Among them, 42 participants with incomplete MRI records and one participant who had previously undergone posterior lumbar fusion were excluded from the analysis (Fig 1).

Experienced board-certified orthopedic surgeons also asked all participants the following question regarding LBP and buttock and leg pain: "Have you experienced LBP (or buttock and leg pain) on most days during the past month, in addition to now?" Those who answered "yes" were defined as having LBP (buttock and leg pain), based on previous studies [20–24]. Twelve participants who lacked information regarding LBP or buttock and leg pain were excluded. Thus, 802 participants (247 men and 555 women) ranging in age from 19 to 93 years (mean, 63.0 years for men and 63.8 years for women) were included for analysis (Fig 1). All study participants provided informed consent, and the study design was approved by the appropriate ethics review boards.

Magnetic resonance imaging

A mobile MRI unit (Achieva 1.5 T; Philips Medical Systems, Best, the Netherlands) was used, and whole-spine MRI was performed for all participants on the same day as the examination. The participants were supine during the MRI, and those with rounded backs used triangular pillows under their heads and knees. The imaging protocol included sagittal T2-weighted fast spin echo imaging (repetition time, 3,000 ms/echo; echo time, 120 ms; and field of view, 270 × 270 mm) and axial T2-weighted fast spin echo imaging (repetition time, 2,100 ms/echo; echo time, 100 ms; and field of view, 180 × 180 mm). Sagittal images were taken for the entire spine, but axial images were obtained for each lumbar intervertebral level (L1/2–L5/S1) parallel to the vertebral endplates.

The CSA of the dural sac

CSA measurement was performed with axial T2-weighted images using a radiological workstation specially designed for such purposes. The CSA of the lumbar dural sac, defined as the area occupied by the dural sac at the disc level, was measured from the level of L1/2 to L4/5 (Fig 2). The measurement was performed by an orthopedic surgeon who was blinded to the background of the participants. The CSA of the dural sac at the most constricted level in the examined spine was called the minimum CSA (mCSA) [8]. The participants were divided into three groups according to quartiles (the narrow group: mCSA less than the first quartile [Q1] measurement; the middle group: mCSA between the Q1 and the third quartile [Q3] measurements; and the wide group: mCSA greater than the measurement for Q3).

Disc degeneration

Disc degeneration (DD) grading was performed by a board-certified orthopedic surgeon who was blinded to the background of the participants, in accordance with previous studies. The

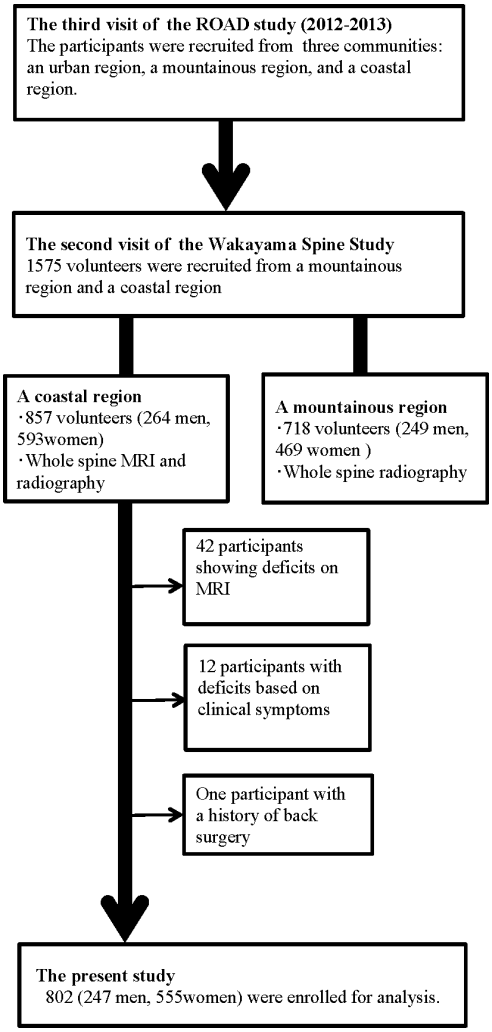


Fig 1. Flow diagram depicting participants recruited to the present study from the third visit of the ROAD study.

doi:10.1371/journal.pone.0160002.g001

degree of DD on MRI was classified into five grades based on the Pfirrmann classification system [25], with grades 4 and 5 indicating DD.

The signal intensity for grade 4 was intermediate to hypointense for the cerebrospinal fluid (dark gray), while the structure was inhomogeneous. Meanwhile, for grade 5, the signal intensity was hypointense for the cerebrospinal fluid (black), and the structure was likewise inhomogeneous. In addition, the disc space was collapsed

Statistical analysis

Radiographic changes were compared between sexes using the chi-squared test. The Jonckheere-Terpstra test was used to identify trends with regard to age or spinal levels for the CSA of dural sac. To test the association between the presence of LBP and mCSA, we used the Cochran-Armitage test and multiple logistic regression analysis. In the regression analysis, we used the presence or absence of LBP as the objective variable and mCSA (the wide group vs.

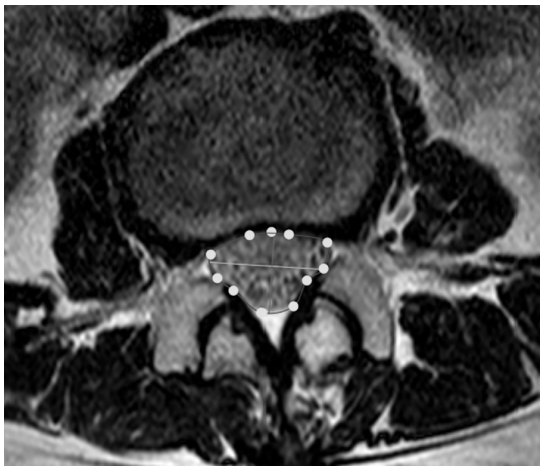


Fig 2. Illustration of the dural sac cross-sectional area measurement technique.

doi:10.1371/journal.pone.0160002.g002

the narrow group; the wide group vs. the middle group) and the presence or absence of DD and buttock and leg pain as explanatory variables, in addition to basic characteristics, such as age, sex, and body mass index (BMI). Receiver operator characteristic (ROC) curves and the corresponding areas under the curve (AUCs) were used to evaluate how the prediction model performed on the test data. The AUC was also calculated independently for the factors in the final model to demonstrate the additional value gained from the addition of each factor to the model. ROC curves plot the true-positive rate (sensitivity) vs the false-positive rate (1-specificity). All statistical analyses except Jonckheere-Terpstra test were performed using JMP, version 8 (SAS Institute Japan, Tokyo, Japan). The Jonckheere-Terpstra test was performed using SPSS statistics 23 (IBM Japan, Tokyo, Japan).

Results

Table 1 shows the characteristics of the 802 participants in the present study, including age, demographic measurements, and symptoms. The prevalence of LBP and buttock and leg pain was 38.6% and 23.3%, respectively. The mean mCSA was 117.3 mm^2 (men: 114.4 mm^2, women: 118.6 mm^2). Q1 and Q3 values for mCSA for the group overall were 85.8 mm^2 and

Table 1. Characteristics of participants.

	Overall	Men	Women
No. of participants	**802**	**247**	**555**
Demographic characteristics			
Age (years)	63.5±13.1	63.0±13.9	63.8±12.7
Height (cm)	157.4±8.9	166.8±6.8	153.4±6.4
Weight (kg)	57.3±11.4	66.7±10.8	53.0±8.9
Body mass index (kg/m^2)	23.0±3.6	24.0±3.5	22.6±3.6
Symptom			
Low back pain	309(38.6%)	94(38.2%)	215(38.7%)
Buttock and leg pain	186(23.3%)	47(19.1%)	139(25.0%)

Data are presented as means ± standard deviation or as n (%).

doi:10.1371/journal.pone.0160002.t001

147.2 mm^2, respectively. The proportion of each group divided into quartiles based on the mCSA values among the 802 participants was as follows: for men, 30.0% for the narrow group (mCSA $\leq$ Q1), 48.2% for the middle group (Q1 $<$ mCSA $\leq$ Q3), and 21.8% for the wide group (Q3 $<$ mCSA), and for women, 22.9% for the narrow group, 50.8% for the middle group, and 26.3% for the wide group. The proportion of men in the narrow group was significantly higher than the proportion of women in the narrow group ($p = 0.032$). The prevalence of disc degeneration in the 802 participants was 91.4% (men: 91.1%, women: 91.5%). There were no significant differences for the prevalence of DD between men and women.

Table 2 summarizes the distribution of the CSA of the dura of the 802 participants in the present study. A Jonckheere-Terpstra test for ordered alternatives showed that there was a statistically significant trend of smaller median CSAs with higher age strata at all intervertebral levels from L1/2 to L4/5 in both genders ($p < 0.0005$). The CSAs had a tendency to decrease with lower intervertebral levels in both genders ($p < 0.0005$).

On analyzing the relationship between the prevalence of LBP and mCSA, we found that the prevalence of LBP increased as mCSA decreased. The prevalence of LBP was 50.3% for the narrow group, 36.6% for the middle group, and 30.8% for the wide group. The participants who had narrower mCSA values were more likely to have LBP ($p < 0.0001$).

Logistic regression analyses were performed with LBP as the objective variable, mCSA as the explanatory variable, and patient characteristics, including age, sex, and BMI, as potential risk factors (model 1). Belonging to the middle group (Q1 $<$ mCSA $\leq$ Q3) was not significantly associated with LBP (odds ratio, [OR] 1.26; 95% confidence interval [CI], 0.87–1.82). On the other hand, belonging to the narrow group (mCSA $\leq$ Q1) was significantly associated with LBP (OR, 1.97; 95% CI, 1.27–3.04).

We then added the presence of DD as a dependent variable (model 2). Belonging to the middle group (Q1 $<$ mCSA $\leq$ Q3) was not significantly associated with LBP (OR, 1.19; 95% CI, 0.82–1.74). In contrast, belonging to the narrow group (mCSA $\leq$ Q1) was significantly associated with LBP (OR, 1.94; 95% CI, 1.25–3.02).

Finally, the presence of buttock and leg pain was added as a dependent variable (model 3). Belonging to the middle group (Q1 $<$ mCSA $\leq$ Q3) was not significantly associated with LBP (OR, 1.18; 95% CI, 0.80–1.73); however, belonging to the narrow group (mCSA $\leq$ Q1) was significantly associated with LBP (OR, 1.78; 95% CI, 1.13–2.80). The results of the logistic regression analysis for all models are summarized in Table 3.

Fig 3 shows the receiver operating characteristic (ROC) curves for the multiple logistic regression models for LBP. The AUC for model 1 was 0.59; for model 2, 0.60; and for model 3,

Table 2. Distribution of cross sectional area of dura (mm^2).

Men	Total	<50	50–59	60–69	70–79	≧80	Standardized Test Statistic	p-value
L1/2	172[149–192]	192[174–212]	177[154–193]	170[147–189]	154[132–181]	157[147–186]	-4.619	<0.0005
L2/3	146[120–172]	176[151–187]	157[136–174]	143[117–163]	126[107–158]	142[109–165]	-5.246	<0.0005
L3/4	132[102–165]	164 [146–181]	144[122–178]	122[102–149]	114[86–140]	119[88–137]	-5.652	<0.0005
L4/5	129[91–168]	166 [128–198]	140[111–186]	120 [88–157]	113[90–146]	99[73–154]	-5.538	<0.0005
Women	**Total**	**<50**	**50–59**	**60–69**	**70–79**	**≧80**	**Standardized Test Statistic**	**p-value**
L1/2	177[157–202]	204[176–220]	186[165–205]	176[156–199]	165[147–189]	168[145–191]	-7.506	<0.0005
L2/3	158[133–184]	188[168–210]	170[147–190]	153[130–178]	145[123–169]	148[112–161]	-8.915	<0.0005
L3/4	139[109–174]	179[149–197]	154[126–177]	134[107–163]	126[97–163]	113[73–153]	-8.775	<0.0005
L4/5	127[96–166]	149[111–189]	140[109–166]	123[96–157]	112[83–160]	107[77–141]	-6.003	<0.0005

Values are the median [first quartile- third quartile].

The CSAs had a tendency to decrease with age and lower intervertebral levels in both genders (Jonckheere-Terpestra test; $p < 0.0005$).

doi:10.1371/journal.pone.0160002.t002

Table 3. Association between low back pain and the minimum cross-sectional area in each logistic regression model.

	Explanatory variables	Category	OR	95%CI	AUC
model 1	mCSA	mCSA<Q1 vs. mCSA>Q3	2.02	1.30–3.12	0.59
		Q1≦mCSA<Q3 vs. mCSA>Q3	1.26	0.87–1.82	
model 2	mCSA	mCSA<Q1 vs. mCSA>Q3	1.94	1.25–3.02	0.6
		Q1≦mCSA<Q3 vs. mCSA>Q3	1.2	0.82–1.74	
	DD	1:presence 0:absence	2.41	1.23–4.73	
model 3	mCSA	mCSA<Q1 vs. mCSA>Q3	1.78	1.13–2.81	0.66
		Q1≦mCSA<Q3 vs. mCSA>Q3	1.18	0.80–1.73	
	DD	1:presence 0:absence	2.38	1.20–4.72	
	buttock and leg pain	1:presence 0:absence	3.31	2.33–4.69	

CI, confidence interval; DD, disc degeneration; mCSA, minimum cross-sectional area; OR, odds ratio; Q1, the first quartile; Q3, the third quartile; AUC, areas under the curve

Note: Multivariate logistic regression analysis of mCSA was associated with low back pain after adjustment for age, body mass index, and sex in each model. The minimum cross-sectional area of the dural sac is the cross-sectional area of the dural sac at the most constricted level in the examined spine from the level of L1/2 to L4/5. Q1, 85.8 mm2; median, 114.2 mm2; Q3, 147.2 mm

doi:10.1371/journal.pone.0160002.t003

0.66. The AUC for model 3 was significantly higher than those for the other two models ($p = 0.0008$).

Discussion

The purpose of this study was to evaluate the relations between the degree of encroachment, measured as the CSA of the dural sac, and LBP in a large population. In this study, narrowed dural sac CSA was significantly associated with the presence of LBP after adjustment for age, sex, and BMI. To the best of our knowledge, this is the first report of a positive association between LBP and the CSA of the dural sac in a large population of individuals ranging in age from 19 to 93 years old.

There are two possibilities to explain the narrow CSA of the dural sac. The narrow dural sac could be due to degenerative changes or to developmental stenosis of the dural sac. In this study, we considered that a narrow CSA of the dural sac was related to constriction of the dural sac due to degenerated tissue around the dural sac rather than to congenital stenosis, as 65.2%

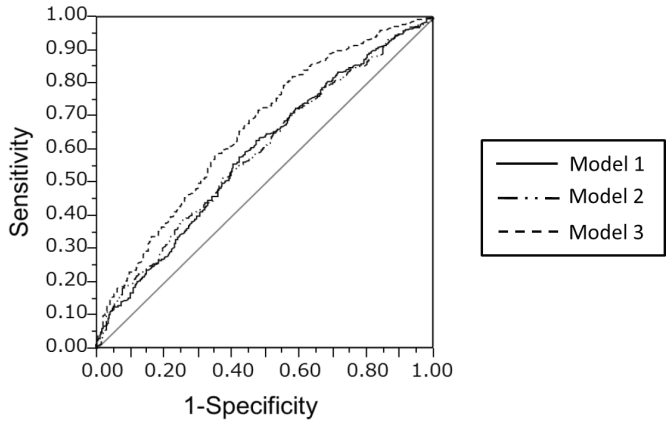


Fig 3. Receiver operating characteristic (ROC) curves for the multiple logistic regression models for low back pain. The area under the curve (AUC) for the ROC curves for models 1, 2, and 3 were 0.59, 0.60, and 0.66, respectively. The AUC for model 3 was significantly greater than the AUCs for models 1 and 2.

doi:10.1371/journal.pone.0160002.g003

of the participants were older than 60 years, and in fact, almost all patients with a small dural sac CSA had other degenerative findings, such as a bulging disc, thickened facet joint, and/or ligamentum flavum.

Findings from earlier preoperative imaging studies on patients with central spinal stenosis have suggested that the CSA of the dural sac was closely related to preoperative walking ability, health-related quality of life, leg pain, and LBP [8,10,26,27]. Moreover, it has been reported that LBP significantly improves following spinal decompression alone [12]. Jones et al. investigated Visual Analog Scale data for LBP in 119 patients with LSS, and they reported that there was a significant reduction in mean LBP from baseline to 6 weeks and 1 year postoperatively. Spinal decompression surgery has long been considered the gold standard surgical treatment for symptomatic LSS. The aim of decompression surgery is to improve radicular leg pain and walking distance. The authors concluded that there is a possibility for improvement in LBP after decompression surgery. We believe our current findings from an established, population-based cohort support the conclusions of the Jones et al. report. However, future studies are needed in order to identify patients who will show improvement in LBP after decompression surgery.

A potential explanation for LBP in LSS patients is the reduction in nutrient supply to ischemic nerves and hence the development of claudication pain originating from muscles supplied by the dorsal rami at the stenotic level [28]. Moreover, Konnai et al. reported that the lower lumbar dura mater is innervated by sensory nerves derived from upper lumbar dorsal root ganglia via the lumbar sympathetic trunk in rats. They concluded that these sensory nerves may mediate LBP and possibly interact with sympathetic nerves [29]. In LSS patients, the dural sac is encroached by degenerative tissue, such as a thickened ligamentum flavum, bulging disc, or osteophyte. Thus, sensory nerves innervating the dural sac can be pinched by degenerative tissues, which might cause LBP in LSS patients.

In the present study, disc degeneration was added as an objective variable to the multivariable regression model. As mentioned above, degenerative discs might be a potential source of LBP in LSS patients. Teraguchi et al. showed that there was a significant positive association between the presence of DD in the lumbar region and LBP [14]. According to that study, the presence of DD in the lumbar region was significantly associated with LBP. Thus, we added the presence of DD as an explanatory factor to adjust for the confounding effect of disc degeneration. Moreover, the presence of buttock and leg pain was added as a dependent variable to the multivariable regression model. LBP is defined as pain in the area bounded by the lowest palpable ribs superiorly and the gluteal folds inferiorly [30]. In this way, participants complaining of buttock pain due to radicular pain might be included in the group of patients with LBP. To adjust for the overlap, the presence of buttock and leg pain was also added as a dependent variable to the multivariable regression model. Finally, adjusting for buttock and leg pain would reinforce the hypothesis that LBP caused by constriction of the dural tube has a different pathology from that of radicular back pain.

After adding these two variables (DD, buttock and leg pain), a narrow CSA of the dural sac was still a significant variable associated with LBP. This result supports the hypothesis that a narrow CSA in the dural sac might be one of the reasons for LBP in LSS patients. Moreover, after adding DD and buttock and leg pain as dependent variables, the AUCs of the ROC curves for the multiple logistic regression analysis increased compared to the AUC before adding dependent variables. Adding the AUC for the ROC curve to the multiple logistic model after adding the presence of DD and buttock and leg pain was 0.66, which was not large. It is assumed that this small value indicates that the CSA of the dural sac might not be strongly correlated with LBP, because LBP can be caused by multiple factors, including osteoporosis, back muscle strain, poor alignment, and psychosocial difficulties. We could explain only a portion of the associated factors for LBP with one factor. However, adding some other factors to the models (MRI findings such as degenerative degeneration, or clinical findings such as buttock

and leg pain) yielded a better multivariate model for LBP. Future investigations should include continued follow-up surveys of other factors, such as facet arthropathy or end-plate change, which would enable us to explain more about nonspecific low back pain.

The present study has several limitations. First, although more than 800 participants were included in the present study, the study population may not be representative of the general population because participants were recruited from only one area of Japan. Anthropometric measurements were compared between the participants of the present study and those of the general Japanese population [31]. There was a significant difference in BMI for both men and women in our study and that of the general population (BMI [standard deviation] in men: 24.0 kg/m^2 [3.5 kg/m^2] vs. 23.4 kg/m^2 [3.36 kg/m^2], respectively, p = 0.00; BMI [standard deviation] in women: 22.57 kg/m^2 [3.62 kg/m^2] vs. 22.29 kg/m^2 [3.69 kg/m^2], respectively, p = 0.031). Therefore, the participants included in this study might have had a different prevalence of LBP or leg pain. However, we believe that the association between the CSA of the dural tube and LBP, which was shown in this study, could be generalized. Second, this is a cross-sectional study, so any causal relationship between symptomatic LSS and physical performance cannot be clarified. The Wakayama Spine Study is a longitudinal survey, so further progress will help elucidate any causal relationships. Third, the configuration of the dural sac was not taken into account. Generally, stenosis of the lumbar spinal canal is divided into the following categories: central stenosis, lateral recess stenosis, and foraminal stenosis. A comprehensive evaluation of spinal stenosis that includes the presence of lateral recess stenosis or foraminal stenosis, and not just the CSA of the dural tube, would be more appropriate for predicting LBP. Finally, the cut-off values for the CSA of the dural sac also posed a problem. In this study, the first and third quartiles of the mCSA were used as the cut-off values for all levels. However, it is inevitable that two nerve roots depart from the cauda at each vertebral level. Thus, it is reasonable to assume a gradually decreasing cut-off value in the distal direction. It might be more appropriate to use different cut-off values for each intervertebral level.

LBP is caused by multiple factors beyond the scope of MRI findings. However, this study clarified that a narrowed CSA of the dural sac was associated with LBP. Although a narrowed CSA might not be strongly correlated with LBP, these findings contribute to our understanding of LBP. Further investigations along with continued follow-up surveys, including additional radiographic findings and psychological and social factors, including occupation, will continue to elucidate the causes of LBP.

Author Contributions

Conceived and designed the experiments: HI NY HH HO HY AM YN MY KM TA ST HK K. Nakamura.

Performed the experiments: HI NY HH HO KS YI K. Nagata MT RK SM.

Analyzed the data: HI NY HH HO.

Contributed reagents/materials/analysis tools: HI NY HH HO.

Wrote the paper: HI.

References

1. Macfarlane GJ, Beasley M, Jones EA, Prescott GJ, Docking R, Keeley P, et al. The prevalence and management of low back pain across adulthood: results from a population-based cross-sectional study (the MUSICIAN study). Pain. 2012; 153: 27–32. doi: 10.1016/j.pain.2011.08.005 PMID: 21978663

2. Chou R, Qaseem A, Snow V, Casey D, Cross JT Jr, Shekelle P, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med. 2007; 147: 478–491. PMID: 17909209
3. Ganz JC. Lumbar spinal stenosis: postoperative results in terms of preoperative posture-related pain. J Neurosurg. 1990; 72: 71–74. PMID: 2136756
4. Kreiner DS, Shaffer WO, Baisden J, Gilbert T, Summers J, Toton J, et al. Evidence-based clinical guidelines for multidisciplinary spine care: Diagnosis and treatment of degenerative lumbar spinal stenosis. Burr Ridge, IL: North American Spine Society; 2011.
5. Amundsen T, Weber H, Nordal HJ, Magnaes B, Abdelnoor M, Lilleâs F. Lumbar spinal stenosis: conservative or surgical management?: A prospective 10-year study. Spine. 2000; 25: 1424–1436. PMID: 10828926
6. Amundsen T, Weber H, Lilleås F, Nordal HJ, Abdelnoor M, Magnaes B. Lumbar spinal stenosis: clinical and radiologic features. Spine. 1995; 20: 1178–1186. PMID: 7638662
7. Miyakoshi N, Hongo M, Kasukawa Y, Ishikawa Y, Shimada Y. Prevalence, spinal alignment, and mobility of lumbar spinal stenosis with or without chronic low back pain: a community-dwelling study. Pain Res Treat. 2011; 2011: 340629. doi: 10.1155/2011/340629 PMID: 22110922
8. Ogikubo O, Forsberg L, Hansson T. The relationship between the cross-sectional area of the cauda equina and preoperative symptoms in central lumbar spinal stenosis. Spine. 2007; 32: 1423–1429. PMID: 17545910
9. Arnoldi CC, Brodsky AE, Cauchoix J, Crock HV, Dommisse GF, Edgar MA, et al. Lumbar spinal stenosis and nerve root entrapment syndromes. Definition and classification. Clin Orthop Relat Res. 1976; 115: 4–5. PMID: 1253495
10. Coulier B. Evaluation of lumbar canal stenosis: decubitus imaging methods versus flexion-extension myelography and surface measurements versus the diameter of the dural sac. JBR-BTR. 2000; 83: 61–67. [Article in French] PMID: 10859898
11. Thomé C, Zevgaridis D, Leheta O, Bäzner H, Pöckler-Schöniger C, Wöhrle J, et al. Outcome after less-invasive decompression of lumbar spinal stenosis: a randomized comparison of unilateral laminotomy, bilateral laminotomy, and laminectomy. J Neurosurg Spine. 2005; 3: 129–141. PMID: 16370302
12. Jones AD, Wafai AM, Easterbrook AL. Improvement in low back pain following spinal decompression: observational study of 119 patients. Eur Spine J. 2014; 23: 135–141. doi: 10.1007/s00586-013-2964-5 PMID: 23963487
13. Nagata K, Yoshimura N, Muraki S, Hashizume H, Ishimoto Y, Yamada H, et al. Prevalence of cervical cord compression and its association with physical performance in a population-based cohort in Japan: The Wakayama Spine Study. Spine. 2012; 37: 1892–1898. doi: 10.1097/BRS.0b013e31825a2619 PMID: 22565382
14. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, et al. Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: The Wakayama Spine Study. Osteoarthritis Cartilage. 2014; 22: 104–110. doi: 10.1016/j.joca.2013.10.019 PMID: 24239943
15. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Associations between radiographic lumbar spinal stenosis and clinical symptoms in the general population: The Wakayama Spine Study. Osteoarthritis Cartilage 2013; 21: 783–788. doi: 10.1016/j.joca.2013.02.656 PMID: 23473979
16. Nagata K, Yoshimura N, Hashizume H, Muraki S, Ishimoto Y, Yamada H, et al. The prevalence of cervical myelopathy among subjects with narrow cervical spinal canal in a population-based magnetic resonance imaging study: The Wakayama Spine Study. Spine J. 2014; 14: 2811–2817. doi: 10.1016/j.spinee.2014.03.051 PMID: 24709229
17. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, et al. The association of combination of disc degeneration, end plate signal change, and Schmorl node with low back pain in a large population study: The Wakayama Spine Study. Spine J. 2015; 15: 622–628. doi: 10.1016/j.spinee.2014.11.012 PMID: 25433277
18. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability (ROAD) study. Int J Epidemiol. 2010; 39: 988–995. doi: 10.1093/ije/dyp276 PMID: 19749026
19. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab. 2009; 27: 620–628. doi: 10.1007/s00774-009-0080-8 PMID: 19568689

20. Muraki S, Oka H, Akune T, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of radiographic lumbar spondylosis and its association with low back pain in elderly subjects of population-based cohorts: the ROAD study. Ann Rheum Dis. 2009; 68: 1401–1406. doi: 10.1136/ard.2007.087296 PMID: 18718988

21. Muraki S, Akune T, Oka H, En-Yo Y, Yoshida M, Saika A, et al. Impact of knee and low back pain on health-related quality of life in Japanese women: the Research on Osteoarthritis against Disability (ROAD). Mod Rheumtol. 2010; 20: 444–451.

22. Muraki S, Akune T, Oka H, En-Yo Y, Yoshida M, Saika A, et al. Health-related quality of life in subjects with low back pain and knee pain in a population-based cohort study of Japanese men: the ROAD study. Spine. 2011; 36: 1312–1319. doi: 10.1097/BRS.0b013e3181fa60d1 PMID: 21730819

23. Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, et al. Incidence and risk factors for radiographic lumbar spondylosis and lower back pain in Japanese men and women: the ROAD study. Osteoarthritis Cartilage. 2012; 20: 712–718. doi: 10.1016/j.joca.2012.03.009 PMID: 22484574

24. Dionne CE, Dunn KM, Croft PR, Nachemson AL, Bunchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine. 2008; 33: 95–103. doi: 10.1097/BRS.0b013e31815e7f94 PMID: 18165754

25. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine. 2001; 26: 1873–1878. PMID: 11568697

26. Bolender NF, Schönström NS, Spengler DM. Role of computed tomography and myelography in the diagnosis of central spinal stenosis. J Bone Joint Surg Am. 1985; 67: 240–246. PMID: 3968115

27. Schönström NS, Bolender NF, Spengler DM. The pathomorphology of spinal stenosis as seen on CT scans of the lumbar spine. Spine (Phila Pa 1976). 1985; 10: 806–811.

28. Porter RW. Spinal stenosis and neurogenic claudication. Spine. 1996; 21: 2046–2052. PMID: 8883210

29. Konnai Y, Honda T, Sekiguchi Y, Kikuchi S, Sugiura Y. Sensory innervation of the lumbar dura mater passing through the sympathetic trunk in rats. Spine. 2000; 25: 776–782. PMID: 10751287

30. Hagen KB, Jamtvedt G, Hilde G, Winnem MF. The updated Cochrane review of bed rest for low back pain and sciatica. Spine. 2005; 30: 542–546. PMID: 15738787

31. Ministry of Health, Labour and Welfare. The Report of the National Health and Nutrition Survey 2013. Available: www.mhlw.go.jp/http://www.mhlw.go.jp/bunya/kenkou/eiyou/dl/h25-houkoku-05.pdf. Accessed: 24 July 2015.

BMJ Open

Influence of work-related psychosocial factors on the prevalence of chronic pain and quality of life in patients with chronic pain

Keiko Yamada,[1] Ko Matsudaira,[2,3] Hironori Imano,[1] Akihiko Kitamura,[1] Hiroyasu Iso[1]

To cite: Yamada K, Matsudaira K, Imano H, *et al*. Influence of work-related psychosocial factors on the prevalence of chronic pain and quality of life in patients with chronic pain. *BMJ Open* 2016;**6**:e010356. doi:10.1136/bmjopen-2015-010356

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2015-010356).

Received 23 October 2015
Accepted 23 March 2016

[1]Department of Public Health and Social Medicine, Osaka University Graduate School of Medicine, Suita-shi, Osaka, Japan
[2]Faculty of Medicine, Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo, Japan
[3]Japan Labour Health and Welfare Organization, Tokyo, Japan

Correspondence to
Dr Hiroyasu Iso;
iso@pbhel.med.osaka-u.ac.jp

ABSTRACT

Objectives: Working is a common cause of chronic pain for workers. However, most of them need to continue working despite the pain in order to make a living unless they get a sick leave or retirement. We hypothesised that the therapeutic effect of vocational rehabilitation may depend on psychosocial factors related to the workplace. To test this hypothesis, we examined the association of work-related psychosocial factors with the prevalence of chronic pain or health-related quality of life (HRQoL) among workers with chronic pain.

Methods: We examined 1764 workers aged 20–59 years in the pain-associated cross-sectional epidemiological survey in Japan. The outcomes were (1) chronic pain prevalence among all workers and (2) low Euro QoL (EQ-5D <0.76; mean value of the current study) prevalence among workers with chronic pain according to the degree of workplace social support and job satisfaction. Workplace social support and job satisfaction were measured using the Brief Job Stress Questionnaire. Multivariable-adjusted ORs were calculated using a logistic regression model including age, sex, smoking, exercise, sleep time, work hours, body mass index, personal consumption expenditure, intensity of pain and the presence of severe depressive symptoms.

Results: Chronic pain prevalence was higher among males reporting job dissatisfaction compared with those reporting job satisfaction. No difference was observed among women. Chronic pain prevalence did not differ between workers of either sex reporting poor workplace social support compared with those reporting sufficient support. Among workers with chronic pain, low HRQoL was more frequent in those reporting job dissatisfaction. Similarly, low HRQoL was more frequent in patients with chronic pain reporting poor social support from supervisors or co-workers compared with patients reporting sufficient support.

Conclusions: Work-related psychosocial factors are critical for HRQoL in patients with chronic pain.

Strengths and limitations of this study

- Our study included a large population and used the specialised questionnaire for pain medicine.
- This is the first study to investigate the association between work-related psychosocial factors and health-related quality of life of patients with chronic pain.
- Our questionnaire included only three psychosocial factors: social support from supervisors, support from co-workers and job satisfaction.
- Patients with severe chronic pain who took sick leave or had retired because of pain were not included in our study.

INTRODUCTION

The existence of chronic pain among workers is an economic burden and major public health problem.[1] Although most workers need to continue their work despite pain, in order to make a living, unless they get sick leave or retire, workers with chronic pain are likely to have lower productivity. Physical and mental overwork can cause chronic pain, which, according to the classical rehabilitation model, requires rest for relief and remission.[2] In contrast, the recent vocational rehabilitation model recommends continuing work or prompt return as soon as possible based on studies showing the benefits of remaining active despite pain.[2–4] However, patients with chronic pain may require added motivation or appropriate accommodation; thus, success of vocational rehabilitation may depend on psychosocial factors related to the workplace environment.

Poor work-related psychosocial factors were associated with a higher chronic pain prevalence among European and North American workers,[2 5 6] but these relationships have not been examined among Asian workers. Workplace environments vary among cultures; therefore, the influence of psychosocial factors on chronic pain may also differ in Asia.[7]

Health-related quality of life (HRQoL) measures are frequently used in epidemiology to quantify general health and functional status. Furthermore, HRQoL has been associated with worker productivity and is often used to calculate the cost-effectiveness of healthcare programmes.[8] Thus, HRQoL is an appropriate metric to evaluate the effects of work-related psychosocial factors on workers with health problems, but the relationship between work-related psychosocial factors and HRQoL of patients with chronic pain has not been explored.

Thus, we analysed the association between work-related psychosocial factors and chronic pain prevalence in the Japanese workplace. In addition, we examined the association between work-related psychosocial factors and HRQoL among patients with chronic pain.

MATERIALS AND METHODS

Study population

The pain-associated cross-sectional epidemiological study was an internet survey (conducted from 10 to 18 January 2009) designed to evaluate pain in a large Japanese population, using a self-reported questionnaire.[9] The sampling procedure ending in the sample being analysed in the current study is shown in figure 1. A total of 20 044 respondents (9746 men and 10 298 women) aged 20–79 years, matching the Japanese demographic composition in 2007,[10] were recruited by email from 1 477 585 candidates who registered with an internet survey company (Rakuten Research Inc, Tokyo, Japan).[11] Invitation emails containing a link to the first questionnaire were sent by computer system until the targeted sample number was achieved. Incomplete questionnaires were rejected automatically, so the response rate was not calculated. The first questionnaire included items on age, sex, job, HRQoL and pain. Subsequently, detailed questionnaires about lifestyle and psychosocial factors were sent to 5000 respondents aged 20–79 years who answered the first questionnaire, 2500 reporting pain and 2500 with no pain. The profile of these 5000 respondents was consistent with the Japanese demographic composition for sex and age in 2007.[10] A total of 2480 workers aged 20–59 years responded to the second questionnaire and 716 workers who had acute or subacute pain were excluded from our analyses. Thus, we included the data on 1764 workers aged 20–59 years, 532 with chronic pain and 1232 with no pain, in the analyses.

The proportions for the different job categories were 29.5% specialists, 8.6% managers, 28.2% white-collar workers, 8.4% sales workers, 3.3% service workers, 0.6% primary sector workers, 2.2% transportation or communication workers, 6.0% menial labourers, and 13.2% others. The majority, 86.2%, were full time while 13.8% were part time.

Ethics

All participants had given their informed consent before responding to the questionnaire. A credit point for internet shopping was given as an incentive to the respondents.

Measures

Job satisfaction and social support from supervisors and co-workers were measured using subscales of the Brief Job Stress Questionnaire.[12] The questionnaire section on social support from supervisors and co-workers consisted of three items ('How well do you get along with your supervisors/co-workers?, 'When you experience difficulties, how much do you rely on your supervisors/co-workers?' and 'How often do you consult your supervisors/co-workers about your private issues/problems?'). Each item was rated on a four-point scale ranging from 1 (sufficient) to 4 (poor), and the total score was calculated by summing the three items for a total score ranging from 3 to 12 points (with lower scores indicating greater levels of support). Subsequently, we calculated the quartiles of scores for social support from supervisors and from co-workers (higher quartile indicating greater level of support) and classified supervisor support as low (Q1, 12–10), intermediate (Q2, 9; Q3, 8–7) or high (Q4, 6–3 points), and co-worker support as low (Q1, 12–9), intermediate (Q2, 8; Q3, 7–6) or high (Q4, 5–3). Job satisfaction was classified into four categories: dissatisfied, somewhat dissatisfied, relatively satisfied or satisfied.

The primary outcome measure was chronic pain prevalence in the entire cohort. The participants also answered questions related to their own pain such as the pain sites, pain intensity at each site, the site of dominant pain, the duration of dominant pain and disability due to dominant pain. Pain intensities were scored on an 11-point Numerical Rating Scale (NRS) (0=no pain, 10=worst pain imaginable). A score ≥5 for the dominant pain site during the past 3 months was defined as chronic pain.

The secondary outcome was the prevalence of low Euro QoL (EQ-5D), defined as below the mean of 0.76 of the present study, in workers with chronic pain according to the NRS. We used the Japanese version of the EQ-5D instrument to measure HRQoL.[13] The EQ-5D includes five dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension is divided into three degrees of severity: 1 (no problem), 2 (moderate) and 3 (extreme problems). The five numbers expressing severity on the five dimensions (eg, 11 233, 22 112 or 11 333) are arranged in the order above, generating 3^5 (or 243) different health statuses. The 243 health statuses are then converted into a single index score called the 'utility value' from 0 (dead) to 1 (full health) according to the conversion table for the Japanese EQ-5D.[13]

The presence of severe depressive symptoms was treated as a confounding factor and used as an adjustment variable because depression is strongly associated with psychosocial factors, chronic pain and quality of

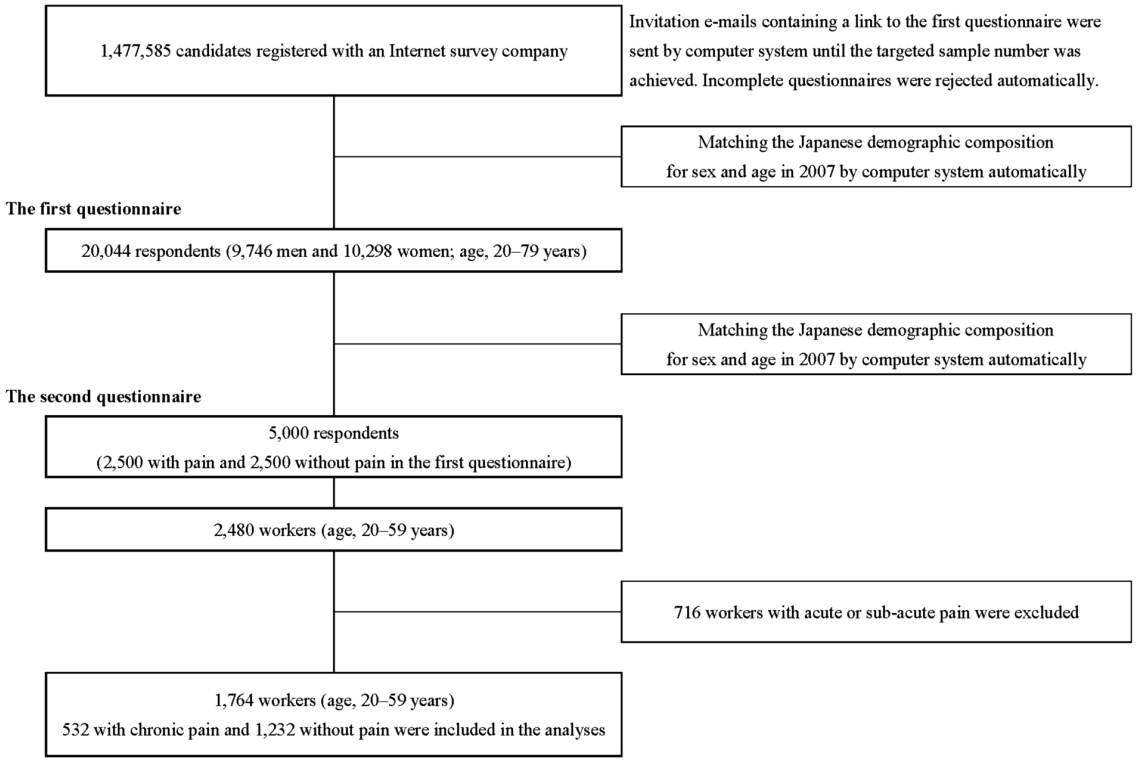


Figure 1 Flow chart of the sampling procedure ending in the sample being analysed in the current study.

life.[14] In the present study, the presence of severe depressive symptoms was defined as a Mental Health Inventory (MHI-5) score <52.[15] The MHI-5 is equal to the 36-item Short Form Health Survey (SF-36) 'mental health' domain.[16] The MHI-5 contains the following five questions: 'How much of the time during the last month have you: (1) been a very nervous person, (2) felt downhearted and blue, (3) felt calm and peaceful, (4) felt so down in the dumps that nothing could cheer you up and (5) been a happy person?' The respondents choose a number from 1 (all of the time) to 6 (none of the time). The score on the MHI-5, ranging from 5 to 30 points, is converted to a 100-point scale.[15] A previous Japanese study confirmed that the cut point of <52 on the MHI-5 (corresponding to ≥56 on the 20-item Zung Self-rating Depression Scale, ZSDS) was useful for screening severe depressive symptoms, with sensitivity of 91.8% and specificity of 84.6%.[15]

Statistical analysis

Analysis of covariance was used to test for differences in age-adjusted means and proportions of the various clinicodemographic characteristics recorded for the analysis. We analysed the association between work-related psychosocial factors and the prevalence of chronic pain among all workers, and the prevalence of low EQ-5D (<the mean of 0.76) among workers with chronic pain.

Multivariable-adjusted ORs and 95% CIs were calculated using the logistic regression model. The adjustment variables included age, sex, smoking status (never-smoker, ex-smoker or current smoker), exercise habit (exercise longer than 30 min more than twice a week; yes or no), sleeping time (hours/day), working hours (<40, 40 to 49, 50 to 59, 60 to 69 or >70 h/week), body mass index (kg/m^2, categorised in quintiles), personal consumption expenditure (JPY/month) and presence of severe depressive symptoms (MHI-5<53). When we analysed the association between work-related psychosocial factors and the EQ-5D of workers with chronic pain, we further adjusted for pain intensity as expressed on an 11-point NRS (0=no pain, 10=worst pain imaginable). We could not analyse the data according to job category because the numbers in each were small.

p Values of <0.05 for two-tailed tests were considered statistically significant. All statistical analyses were performed using SAS V.9.4 (SAS Institute Inc, Cary, North Carolina, USA).

RESULTS

We identified 532 workers aged 20–59 years (309 men and 223 women) who had chronic pain, and 1232 workers (783 men and 449 women) without pain. A total of 306 workers with chronic pain (57.5%) reported a severity of 5 points or more on the 11-point NRS for low back pain or neck pain. The prevalence of chronic pain in female workers was significantly higher than in

male workers (24.5% vs 19.7%, p<0.05). The prevalence of severe depressive symptoms in patients with chronic pain was 35.3% for men and 37.2% for women, about two-times higher than in pain-free workers (18.1% for men and 21.4% for women).

Table 1 shows the age-adjusted means and proportions of clinicodemographic characteristics according to work-related psychosocial factor category (support from supervisors, support from co-workers, job satisfaction). Workers who received poor support from supervisors and/or co-workers (Q1) and were dissatisfied with their jobs demonstrated a higher prevalence of severe depressive symptoms. A greater proportion of male workers who were dissatisfied with their job exhibited short sleep compared to those with good work-related psychosocial factors. Male workers receiving poor support from supervisors and/or co-workers, and who were dissatisfied with their jobs, exhibited a higher prevalence of chronic pain compared with those satisfied with their jobs and receiving sufficient support. Pain intensity of patients with chronic pain did not vary according to work-related psychosocial factor category.

Table 2 shows the age-adjusted and multivariable ORs of chronic pain patient characteristics according to work-related psychosocial factor category. Male workers with poor job satisfaction exhibited a higher prevalence of chronic pain, and the association remained statistically significant after adjustment for confounding factors and the presence of severe depressive symptoms. The degree of social support from supervisors/co-workers was associated with a progressive decrease in the prevalence of chronic pain among male workers. However, after adjustment for confounding variables and presence of severe depressive symptoms, most of these individual associations between chronic pain prevalence and support quartile were no longer statistically significant. No such associations were found for female workers.

Table 3 shows the age-adjusted and multivariable ORs for low EQ-5D (less than the mean of 0.76) in workers with chronic pain. The mean EQ-5D value of all workers was 0.90 (SD 0.14), 0.96 for those without chronic pain (SD 0.09) and 0.76 (SD 0.12) for those with chronic pain; therefore, 0.76 was used as the cut-off point. Work-related social factors were significantly associated with the low EQ-5D (less than the mean of 0.76) among workers of both sexes with chronic pain in the age-adjusted model and model 1. However, those associations were no longer statistically significant in the severe depressive syndrome-adjusted model (model 2), except for the category, support from supervisors/co-workers for female workers. In the entire cohort of patients with chronic pain (men and women), there was a significantly higher prevalence of low EQ-5D in those reporting poor support from supervisors and from those reporting poor support from co-workers. These associations were statistically significant in neither men nor women separately.

DISCUSSION

These results reveal a significant association between job dissatisfaction and the prevalence of chronic pain in Japanese male workers, in accord with previous findings of a link between job dissatisfaction and chronic musculoskeletal pain among European and North American workers.[2 5 6] However, no gender differences were found in those studies, while the link between job dissatisfaction and chronic pain was specific to men in the current study.

In general, there are gender differences in pain and analgaesia,[17 18] so we stratified participants by sex. Previous studies have found that women experience chronic pain more often than men.[17 18] Similarly, the prevalence of chronic pain was significantly higher in women than in men in the current study. However, job satisfaction was associated with the prevalence of chronic pain only among men. It was suggested that the incidence of chronic pain may be more influenced by job satisfaction among men than among women in Japan. Men generally work longer hours than do women in Japan, with an average daily working time of 416 min for men and 290 min for women in 2011.[19] On the other hand, many Japanese men do not share housework, averaging only 42 min daily compared with 215 min for women.[19] In our study, the proportion of male workers who had been working >60 h per week (15.5%) were more than that of female workers (4.2%).Thus, women may be more strongly affected by psychosocial factors at home than at work due to shorter working hours or cultural expectations. However, psychosocial factors in private life were not examined in this study, so we could not investigate the reasons for this gender difference in the impact of psychosocial factors on chronic pain.

Several recent studies have recommended that patients with chronic pain should continue to work and not take prolonged leave;[2–4] however, the success of this vocational rehabilitation model could depend on a favourable work environment. Indeed, support from supervisors and co-workers did have a positive effect on workers with chronic pain according to self-reported HRQoL. Similarly, supportive relationships at work led to better HRQoL of employees with severe mental illness.[4] This need for supportive supervisors and co-workers may result from a lower resiliency and capacity to cope with stress compared with healthy workers.[20]

We investigated the prevalence of low EQ-5D (defined as below the mean of 0.76) among workers with chronic pain as the secondary outcome of our research. In a Finnish study, the mean EQ-5D of the general population was 0.84 and that of the subpopulation with back pain was 0.74.[21] This is similar to the mean EQ-5D of Japanese workers in the current study, most of whom suffered from back or neck pain. Thus, EQ-5D appears to reflect impaired QoL resulting from chronic musculoskeletal pain. The EQ-5D can also detect meaningful changes in other clinical conditions.[22] For instance,

Table 1 Age-adjusted mean values and proportions of characteristics according to the categories of work-related psychosocial factors among workers aged 20–59 years

	Support from supervisors				Support from co-workers				Job satisfaction			
The quartiles of scores for social support Score; with lower scores indicating greater levels of support	**Q4 6–3 points**	**Q3 8–7 points**	**Q2 9 points**	**Q1 12–10 points**	**Q4 5–3 points**	**Q3 7–6 points**	**Q2 8 points**	**Q1 12–9 points**	**Satisfied**	**Relatively satisfied**	**Somewhat dissatisfied**	**Dissatisfied**
Men, n=1092												
n	285	332	232	243	200	351	226	315	136	638	234	84
Age (years)	40.0 (0.6)	42.7 (0.5)	40.3 (0.7)	41.2 (0.6)	40.6 (0.6)	43.0 (0.7)	41.4 (0.5)	39.6 (0.7)	39.6 (0.9)	41.6 (0.4)	41.1 (0.7)	40.4 (1.1)
Current smoker (%)	32.2	37.3	28.4	32.5	37.0	31.6	32.3	32.7	31.6	31.3	36.8	38.1
Have an exercise habit (%)	32.3	26.2	27.2	27.6	29	29.3	27	27.6	37.5	28.2	23.9	26.2
Short sleep (%)	6.0	2.7	3.0	8.2	3.5	5.1	3.5	6.3	3.7	4.4	3.8	13.1†
Overwork (%)	14.7	12.3	14.7	20.6	13.0	13.7	19.0	15.9	11.8	14.4	19.7*	15.5
Body mass index ≥25 (%)	25.6	26.2	26.7	25.9	23.5	24.2	29.2	27.6	22.8	27.0	26.5	23.8
Personal consumption expenditure (×10 000 JPY/ month)	29.1	32.2	28.7	27.6	26.3	34.7*	28.6	26.7	32.8	28.8	28.9	32.6
Severe depressive symptoms (%)	13.7	17.2	23.3†	41.6‡	11.5	14.8	19.5*	41.9‡	8.8	13.5	38.5‡	75.0‡
Chronic pain (%; number of participants with chronic pain, n=309)	23.5 (67)	25.9 (86)	30.6 (71)	35.0† (85)	25.0 (50)	21.4 (75)	30.1 (69)	36.5† (115)	19.9 (27)	25.7 (164)	32.1* (75)	51.2‡ (43)
Intensity of pain of participants with chronic pain (Numerical Rating Scale)	7.0	7.0	6.9	6.9	7.2	7.0	7.0	6.8	6.4	7.1*	6.9	7.0
Women, n=672												
n	160	210	110	192	149	215	95	213	74	411	134	53
Age (years)	38.9 (0.7)	38.4 (0.9)	40.1 (0.7)	40.2 (0.8)	38.7 (0.7)	40.3 (1.0)	39.6 (0.7)	40.0 (0.8)	40.3 (1.2)	39.8 (0.5)	38.9 (0.9)	37.2 (1.4)
Current smoker (%)	19.4	21.9	21.8	17.7	22.1	22.8	18.9	16.4	20.3	19.2	20.9	24.5
Have an exercise habit (%)	26.9	19.5	18.2	18.8	19.5	22.3	25.3	18.3	27.0	23.6	11.9	13.2
Short sleep (%)	4.4	4.3	1.8	6.3	3.3	4.2	7.4	4.2	1.4	4.1	7.5*	3.8

Continued

Table 1 Continued

The quartiles of scores for social support Score; with lower scores indicating greater levels of support	Support from supervisors				Support from co-workers				Job satisfaction			
	Q4	Q3	Q2	Q1	Q4	Q3	Q2	Q1				
	6–3 points	8–7 points	9 points	12–10 points	5–3 points	7–6 points	8 points	12–9 points	Satisfied	Relatively satisfied	Somewhat dissatisfied	Dissatisfied
Overwork (%)	5.6	5.2	1.8	3.1	5.4	4.7	2.1	3.8	6.8	4.4	3.7	0
Body mass index ≥25 (%)	13.1	13.3	15.5	14.1	12.8	15.8	10.5	14.1	9.5	14.8	17.2	3.8
Personal consumption expenditure (×10 000 JPY/ month)	25.6	27.3	29.5	24.6	24.3	27.3	26.2	27.3	24.2	25.8	30.1	26.3
Severe depressive symptoms (%)	10.6	23.3†	27.3†	43.2‡	12.1	23.3*	25.3*	40.8‡	10.8	16.1	46.3‡	81.1‡
Chronic pain (%; number of participants with chronic pain, n=223)	30.0 (48)	31.4 (66)	37.2 (41)	35.4 (68)	33.6 (50)	31.6 (68)	26.3 (25)	37.6 (80)	29.7 (22)	29.2 (120)	44.0* (59)	41.5 (22)
Intensity of pain of participants with chronic pain (Numerical Rating Scale)	7.5	6.9	7.4	7.8	7.5	7.2	6.8*	7.6	7.4	7.2	7.7	7.5

Test for significance from the category of Q4, or satisfied.
In parentheses: SEs.
The quartiles of scores for social support from supervisors and from co-workers were calculated, and classified as low (Q1), intermediate (Q2; Q3) or high (Q4).
*p<0.05, †<0.01, ‡<0.001.

Table 2 ORs (ORs, 95% CI) of chronic pain according to the categories of work-related psychosocial factors among workers aged 20–59 years

	Support from supervisors				Support from co-workers				Job satisfaction			
The quartiles of scores for social support Score; with lower scores indicating greater levels of support	**Q4** **6–3 points**	**Q3** **8–7 points**	**Q2** **9 points**	**Q1** **12–10 points**	**Q4** **5–3 points**	**Q3** **7–6 points**	**Q2** **8 points**	**Q1** **12–9 points**	**Satisfied**	**Relatively satisfied**	**Somewhat dissatisfied**	**Dissatisfied**
Men, n=1092												
Number of participants	285	332	232	243	200	351	226	315	136	638	234	84
Number of participants with chronic pain	67	86	71	85	50	75	69	115	27	164	75	43
Age-adjusted OR (95% CI)	1.00	1.14 (0.79 to 1.65)	1.44 (0.97 to 2.13)	1.75 (1.20 to 2.56)†	1.00	0.82 (0.54 to 1.23)	1.32 (0.86 to 2.03)	1.73 (1.16 to 2.56)†	1.00	1.40 (0.88 to 2.21)	1.91 (1.15 to 3.15)*	4.23 (2.32 to 7.72)‡
Model 1 OR (95% CI)	1.00	1.22 (0.84 to 1.78)	1.47 (0.98 to 2.19)	1.81 (1.23 to 2.68)†	1.00	0.88 (0.58 to 1.34)	1.35 (0.87 to 2.10)	1.81 (1.21 to 2.72)†	1.00	1.39 (0.87 to 2.22)	1.87 (1.12 to 3.14)*	4.35 (2.35 to 8.04)‡
Model 2 OR (95% CI)	1.00	1.18 (0.81 to 1.73	1.35 (0.90 to 2.03)	1.42 (0.95 to 2.13)	1.00	0.85 (0.56 to 1.30)	1.25 (0.80 to 1.95)	1.40 (0.92 to 2.13)	1.00	1.35 (0.84 to 2.16)	1.51 (0.89 to 2.57)	2.76 (1.43 to 5.31)‡
Women, n=672												
Number of participants	160	210	110	192	149	215	95	213	74	41	134	53
Number of participants with chronic pain	48	66	41	68	50	68	25	80	22	120	59	22
Age-adjusted OR (95% CI)	1.00	1.07 (0.68 to 1.67)	1.36 (0.81 to 2.27)	1.26 (0.80 to 1.98)	1.00	0.91 (0.58 to 1.42)	0.71 (0.40 to 1.25)	1.17 (0.76 to 1.82)	1.00	0.97 (0.56 to 1.67)	1.84 (1.00 to 3.36)*	1.62 (0.77 to 3.41)
Model 1 OR (95% CI)	1.00	1.03 (0.65 to 1.64)	1.45 (0.85 to 2.48)	1.31 (0.81 to 2.10)	1.00	0.95 (0.60 to 1.51)	0.78 (0.43 to 1.40)	1.28 (0.81 to 2.03)	1.00	1.03 (0.58 to 1.82)	1.66 (0.87 to 3.17)	1.98 (0.90 to 4.36)
Model 2 OR (95% CI)	1.00	0.95 (0.59 to 1.52)	1.29 (0.75 to 2.24)	1.05 (0.64 to 1.71)	1.00	0.86 (0.54 to 1.38)	0.69 (0.38 to 1.26)	1.04 (0.65 to 1.68)	1.00	0.98 (0.55 to 1.74)	1.31 (0.67 to 2.56)	1.23 (0.53 to 2.86)
Total, n=1764												
Number of participants	445	542	342	435	349	566	321	528	210	1049	368	137
Number of participants with chronic pain	115	152	112	153	100	143	94	195	49	284	134	65
Age-adjusted OR (95% CI)	1.00	1.12 (0.84 to 1.49)	1.41 (1.03 to 1.92)*	1.53 (1.15 to 2.05)†	1.00	0.85 (0.63 to 1.15)	1.07 (0.77 to 1.50)	1.47 (1.10 to 1.97)*	1.00	1.21 (0.86 to 1.72)	1.88 (1.28 to 2.77)†	2.94 (1.85 to 4.68)‡
Model 1 OR (95% CI)	1.00	1.15 (0.86 to 1.53)	1.44 (1.05 to 1.98)*	1.58 (1.17 to 2.12)†	1.00	0.90 (0.66 to 1.22)	1.01 (0.78 to 1.56)	1.55 (1.15 to 2.09)†	1.00	1.20 (0.84 to 1.72)	1.76 (1.18 to 2.62)†	3.13 (1.94 to 5.06)‡
Model 2 OR (95% CI)	1.00	1.08 (0.81 to 1.45)	1.31 (0.95 to 1.81)	1.25 (0.91 to 1.70)	1.00	0.84 (0.62 to 1.15)	1.01 (0.71 to 1.43)	1.22 (0.89 to 1.67)	1.00	1.15 (0.80 to 1.65)	1.42 (0.94 to 2.13)†	1.99 (1.19 to 3.32)†

Model 1: adjusted for age, sex, smoking, exercise habit, sleep time, working time, body mass index and personal consumption expenditure.
Model 2: adjusted for the aforementioned variables plus existence of severe depressive symptoms.
The quartiles of scores for social support from supervisors and from co-workers were calculated, and classified as low (Q1), intermediate (Q2; Q3) or high (Q4).
Test for significance from the category of Q4 or satisfied.
*p<0.05, †p<0.01, ‡p<0.001.

Table 3 ORs (ORs, 95%CI) of below Euro QoL (EQ-5D) mean value according to the categories of work-related psychosocial factors among workers aged 20–59 years

	Support from supervisors					Support from co-workers					Job satisfaction			
The quartiles of scores for social support Score; with lower scores indicating greater levels of support	**Q4**	**Q3**	**Q2**	**Q1**	**Q4**	**Q3**	**Q2**	**Q1**						
	6–3 points	**8–7 points**	**9 points**	**12–10 points**	**5–3 points**	**7–6 points**	**8 points**	**12–9 points**	**Satisfied**	**Relatively satisfied**	**Somewhat dissatisfied**	**Dissatisfied**
Men, n=309												
Number of participants	67	86	71	85	50	75	69	115	27	164	75	43
Number of participants with<EQ-5D mean value	20	24	21	47	12	20	24	56	8	45	29	30
Age-adjusted OR	1.00	0.87	0.99	2.87	1.00	1.08	1.57	2.91	1.00	0.87	1.44	5.56
		(0.43 to 1.78)	(0.48 to 2.06)	(1.46 to 5.65)†		(0.47 to 2.50)	(0.68 to 3.61)	(1.37 to 6.14)‡		(0.35 to 2.13)	(0.56 to 3.73)	(1.94 to 15.98)‡
Model 1 OR	1.00	0.81	0.84	2.60	1.00	1.03	1.45	2.70	1.00	0.78	1.25	5.09
		(0.39 to 1.70)	(0.39 to 1.82)	(1.28 to 5.29)†		(0.43 to 2.45)	(0.62 to 3.42)	(1.24 to 5.90)*		(0.31 to 1.96)	(0.47 to 3.34)	(1.70 to 15.25)†
Model 2 OR	1.00	0.61	0.66	1.43	1.00	0.87	1.10	1.48	1.00	0.74	0.71	2.09
		(0.27 to 1.36)	(0.29 to 1.51)	(0.65 to 3.12)		(0.34 to 2.24)	(0.44 to 2.75)	(0.63 to 3.47)		(0.28 to 1.99)	(0.24 to 2.10)	(0.62 to 7.09)
Women, n=223												
Number of participants	48	66	41	68	50	68	25	80	22	120	59	22
Number of participants with <EQ-5D mean value	13	25	17	41	14	22	13	47	6	42	32	16
Age-adjusted OR	1.00	1.64	1.90	4.09	1.00	1.23	2.78	3.66	1.00	1.44	3.15	7.24
		(0.73 to 3.69)	(0.78 to 4.64)	(1.84 to 9.10)‡		(0.55 to 2.74)	(1.02 to 7.58)*	(1.71 to 7.84)‡		(0.52 to 3.95)	(1.08 to 9.19)*	(1.90 to 27.55)†
Model 1 OR	1.00	1.63	1.71	3.95	1.00	1.11	2.49	3.45	1.00	1.42	3.43	7.55
		(0.68 to 3.92)	(0.67 to 4.35)	(1.69 to 9.22)†		(0.47 to 2.60)	(0.85 to 7.30)*	(1.55 to 7.66)†		(0.46 to 4.38)	(1.05 to 11.26)*	(1.77 to 32.20)†
Model 2 OR	1.00	1.19	1.22	2.54	1.00	0.86	1.60	2.59	1.00	1.36	2.06	3.49
		(0.47 to 3.01)	(0.45 to 3.32)	(1.04 to 6.21)*		(0.35 to 2.12)	(0.51 to 5.05)	(1.12 to 5.97)*		(0.42 to 4.38)	(0.58 to 7.28)	(0.74 to 16.36)
Total, n=532												
Number of participants	115	152	112	153	100	143	94	195	49	284	134	65
Number of participants with <EQ-5D mean value	33	49	38	88	26	42	37	103	14	87	61	46

Continued

Table 3 Continued

The quartiles of scores for social support Score; with lower scores indicating greater levels of support	Support from supervisors					Support from co-workers					Job satisfaction			
	6–3 points	8–7 points	9 points	12–10 points	Q1	5–3 points	7–6 points	8 points	12–9 points	Q1	Satisfied	Relatively satisfied	Somewhat dissatisfied	Dissatisfied
	Q4	Q3	Q2			Q4	Q3	Q2						
Age-adjusted OR	1.00	1.16	1.30	3.35		1.00	1.19	2.01	3.32		1.00	1.11	2.07	6.57
		(0.68 to 1.98)	(0.74 to 2.29)	(2.00 to 5.61)‡			(0.67 to 2.11)	(1.08 to 3.74)*	(1.95 to 5.67)‡			(0.56 to 2.17)	(1.02 to 4.22)*	(2.88 to 15.03)‡
Model 1 OR	1.00	1.07	1.15	3.03		1.00	1.12	1.82	3.18		1.00	1.06	1.94	6.42
		(0.62 to 1.85)	(0.65 to 2.06)	(1.78 to 5.15)‡			(0.62 to 2.02)	(0.97 to 3.44)	(1.84 to 5.50)‡			(0.53 to 2.15)	(0.92 to 4.09)	(2.72 to 15.19)‡
Model 2 OR	1.00	0.82	0.89	1.85		1.00	0.92	1.38	2.07		1.00	1.02	1.14	2.77
		(0.46 to 1.48)	(0.48 to 1.66)	(1.04 to 3.28)*			(0.49 to 1.74)	(0.70 to 2.73)	(1.15 to 3.70)*			(0.49 to 2.11)	(0.52 to 2.52)	(1.09 to 7.01)*

Model 1: adjusted for age, sex, smoking, exercise habit, sleep time, working time, body mass index, personal consumption expenditure, intensity of pain.
Model 2: adjusted for aforementioned variables plus existence of severe depressive symptoms.
The quartiles of scores for social support from supervisors and from co-workers were calculated, and classified as low (Q1), intermediate (Q2; Q3) or high (Q4).
Test for significance from the category of Q4 or satisfied.
*p<0.05, †p<0.01, ‡p<0.001.

intensity of migraine was correlated with EQ-5D.[23] Patients with chronic pain reporting poor support in the workplace showed a higher prevalence of low mean EQ-5D, indicating that these work-related psychosocial factors are important for maintenance of general health status and functional well-being.

Absence from work because of sickness for regional pain symptoms is much less common in Japan compared with that in the UK.[24] Compared with the UK, the reported rates of sick leave for regional pain symptoms in Japan are less than one-third (only 5%).[24 25] According to the population-based survey,[26] the prevalence of chronic pain in Japan (22.9%) is similar to that in Europe;[2] therefore, the number of people who are working with chronic pain without absence from work may be larger in Japan than in the UK. This cultural difference may reflect the result of a wide range HRQoL of workers with chronic pain being observed in the present study; thus, the association between work-related psychosocial factors and EQ-5D among workers with chronic pain may be detected sensitively in the current Japanese study.

Depression is strongly associated with psychosocial factors, chronic pain and QoL in the clinical setting.[14] The lifetime prevalence of major depression in primary care settings is 5–10%,[27] and the reported prevalence of pain in patients with depression averages about 65% (range 15–100%).[28] The estimated coexistence of major depression with chronic pain in the general population is 18% (4.7–22%).[28] In fact, the presence of severe depressive symptoms was the most powerful confounder in our study.

Limitations

There are several limitations to our study. First, the participants may not be truly representative of the general population. Although the demographic profile of respondents was consistent with the Japanese demographic composition for sex and age in 2007, the 1764 participants who answered the detailed questionnaire were selected purposefully (by eliminating those with acute and subacute pain). According to the Annual Report on the Labor Force Survey in 2009,[29] the percentage of the Japanese labour force aged 20–59 years was 92.5% for men and 69.9% for women, while 88.5% of male respondents and 52.2% of female respondents were currently in the labour force. The proportions of job categories were biased. Specialists and white-collar workers have a majority (57.7%), and the proportion of primary sector workers was very low (0.6%). Moreover, factors influencing the decision to respond to the web-based survey may have biased the distribution. For example, it may have selected against extremely busy workers or older workers less familiar with the internet. In addition, the respondents may have been particularly interested in pain research, possible due to personal affliction. The sampling issues of the web-based survey were noted before.[30] This difference, particularly the

low portion of female respondents in the workforce, could have influenced the results.

Second, our questionnaire included only three psychosocial factors, social support from supervisors, support from co-workers and job satisfaction. 'Job demand' and 'job control'[31] have also been included as work-related psychosocial factors, but these were not examined in this study. Our questionnaire did not include social support in private life. Furthermore, individual psychosocial factors such as 'fear avoidance', 'pain catastrophising' and 'resilience' were not investigated in the present study.

Third, patients with severe chronic pain who took sick leave or had retired due to pain were not included in our study. This limitation could reduce the statistical power to examine the association between work-related psychosocial factors and chronic pain.

CONCLUSION

Male workers reporting job dissatisfaction had a higher prevalence of chronic pain than those reporting job satisfaction. Among workers with chronic pain, those reporting poor social support and job dissatisfaction had a greater frequency of low HRQoL. Thus, work-related psychosocial factors are critical influences on the HRQoL of workers with chronic pain.

Acknowledgements The authors express their appreciation to Yasuo Takagi, professor of Keio University, Japan, for his valuable help in conducting the survey. The authors are grateful to all of the participants of this study. The authors thank Nagisa Mori for her linguistic advice. The authors would like to thank Enago (http://www.enago.jp) for the English language review.

Contributors KY conceived the original idea for the study, performed the analyses of the data and drafted the manuscript. KM and HI revised the manuscript, contributed to the interpretation of the data and added critical comments. HI and AK revised the manuscript and contributed with comments. All the authors approved the final version.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Obtained.

Ethics approval The Institutional Review Boards of Keio University and of the Japan Labour Health and Welfare Organization.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Goldberg DS, McGee SJ. Pain as a global public health priority. *BMC Public Health* 2011;11:770.
2. Croft P, Blyth FM, van der Windt D. *Chronic pain epidemiology: from aetiology to public health*. 1st edn. USA: Oxford University Press.
3. Waddell G, O'Connor M, Boorman S, *et al*. Working Backs Scotland: a public and professional health education campaign for back pain. *Spine* 2007;32:2139–43.
4. Rüesch P, Graf J, Meyer PC, *et al*. Occupation, social support and quality of life in persons with schizophrenic or affective disorders. *Soc Psychiatry Psychiatr Epidemiol* 2004;39:686–94.
5. Linton SJ. Occupational psychological factors increase the risk for back pain: a systematic review. *J Occup Rehabil* 2001;11:53–66.
6. Hoogendoorn WE, van Poppel MN, Bongers PM, *et al*. Systematic review of psychosocial factors at work and private life as risk factors for back pain. *Spine* 2000;25:2114–25.
7. Palmer KT, Calnan M, Wainwright D, *et al*. Disabling musculoskeletal pain and its relation to somatization: a community-based postal survey. *Occup Med (Chic Ill)* 2005;55:612–17.
8. Brouwer WBF, Meerding WJ, Lamers LM, *et al*. The relationship between productivity and health-related QOL: an exploration. *Pharmacoeconomics* 2005;23:209–18.
9. Yamada K, Matsudaira K, Takeshita K, *et al*. Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey. *Mod Rheumatol* 2013.
10. Japanese Ministry of Internal Affairs and Communications. http://www.stat.go.jp/data/jinsui/2007np/index.htm (accessed 14 Mar 2015).
11. Rakuten Research: English. http://research.rakuten.co.jp/en/ (accessed 14 Mar 2015).
12. Shimomitsu T, Haratani T, Nakamura K, *et al*. *The final development of the Brief Job Stress Questionnaire mainly used for assessment of the individuals*. Tokyo, 1990.
13. Williams A. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy (New York)* 1990;16:199–208.
14. Dworkin RH, Gitlin MJ. Clinical aspects of depression in chronic pain patients. *Clin J Pain* 1991;7:79–94.
15. Yamazaki S, Fukuhara S, Green J. Usefulness of five-item and three-item Mental Health Inventories to screen for depressive symptoms in the general population of Japan. *Health Qual Life Outcomes* 2005;3:48.
16. Fukuhara S, Bito S, Green J, *et al*. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. *J Clin Epidemiol* 1998;51:1037–44.
17. Woodrow KM, Friedman GD, Siegelaub AB, *et al*. Pain tolerance: differences according to age, sex and race. *Psychosom Med* 1972;34:548–56.
18. Craft RM, Mogil JS, Aloisi AM. Sex differences in pain and analgesia: the role of gonadal hormones. *Eur J Pain* 2004;8:397–411.
19. Ministry of Internal Affairs and Communications. Survey on time use and leisure activities (in Japanese). 2011. http://www.stat.go.jp/data/shakai/2011/pdf/houdou2.pdf (accessed 2 Jun 2015).
20. Friborg O, Hjemdal O, Rosenvinge JH, *et al*. Resilience as a moderator of pain and stress. *J Psychosom Res* 2006;61:213–19.
21. Saarni SI, Härkänen T, Sintonen H, *et al*. The impact of 29 chronic conditions on health-related quality of life: a general population survey in Finland using 15D and EQ-5D. *Qual Life Res* 2006;15:1403–14.
22. Payakachat N, Ali MM, Tilford JM. Can the EQ-5D detect meaningful change? A systematic review. *Pharmacoeconomics* 2015;33:1137–54.
23. Xu R, Insinga RP, Golden W, *et al*. EuroQol (EQ-5D) health utility scores for patients with migraine. *Qual Life Res* 2011;20:601–8.
24. Matsudaira K, Palmer KT, Reading I, *et al*. Prevalence and correlates of regional pain and associated disability in Japanese workers. *Occup Environ Med* 2011;68:191–6.
25. Madan I, Reading I, Palmer KT, *et al*. Cultural differences in musculoskeletal symptoms and disability. *Int J Epidemiol* 2008;37:1181–9.
26. Matsudaira K, Tatsuya I. Epidemiological findings from our studies of musculoskeletal pain. *Pain Clin* 2013;34:S53–61.
27. Katon W, Schulberg H. Epidemiology of depression in primary care. *Gen Hosp Psychiatry* 1992;14:237–47.
28. Bair MJ, Robinson RL, Katon W, *et al*. Depression and pain comorbidity: a literature review. *Arch Intern Med* 2003;163:2433–45.
29. Annual Report on the Labor Force Survey (in Japanese). http://www.stat.go.jp/data/roudou/report/2009/ (accessed 14 Mar 2015).
30. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. *J Epidemiol Community Health* 2003;57:68–73.
31. De Jonge J, Van Vegchel N, Shimazu A, *et al*. A longitudinal test of the demand-control model using specific job demands and specific job control. *Int J Behav Med* 2010;17:125–33.

J. Phys. Ther. Sci. 28: 802–806, 2016

The Journal of Physical Therapy Science

Original Article

Effect of pelvic forward tilt on low back compressive and shear forces during a manual lifting task

Shota Hayashi, RPT[1, 2)*], Junji Katsuhira, PhD[2, 3)], Ko Matsudaira, MD, PhD[3)], Hitoshi Maruyama, RPT, PhD[2)]

[1)] *Rehabilitation Center, Saiseikai Kanagawaken Hospital: 6-6 Tomiyacho, Kanagawa-ku, Yokohama-shi, Kanagawa 221-0821, Japan*
[2)] *Graduate School of International University of Health and Welfare, Japan*
[3)] *Department of Medical Research and Management for Musculoskeletal Pain 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo Hospital, Japan*

Abstract. [Purpose] To examine the effect of an instruction to increase pelvic forward tilt on low back load during a manual lifting task in the squat and stoop postures. [Subjects] Ten healthy males who provided informed consent were the subjects. [Methods] Kinetic and kinematic data were captured using a 3-dimensional motion analysis system and force plates. Low back compressive and shear forces were chosen as indicators of low back load. The subjects lifted an object that weighed 11.3 kg, under the following 4 conditions: squat posture, stoop posture, and these lifting postures along with an instruction to increase pelvic forward tilt. [Results] In the squat posture, the instruction to increase pelvic forward tilt reduced the low back compression and shear forces. [Conclusion] The present results suggest that a manual lifting task in the squat posture in combination with an instruction to increase pelvic forward tilt can decrease low back compression and shear forces, and therefore, might be an effective preventive method for low back pain in work settings.

Key words: Manual lifting task, Low back load, Motion analysis

(This article was submitted Sep. 9, 2015, and was accepted Nov. 26, 2015)

INTRODUCTION

A large number of people in developed countries have low back pain. The prevalence rate of lifetime low back pain in Japan was reported to be 83.5%[1)], and low back pain accompanies most occupational diseases. Manual lifting tasks are reported to confer the highest risk of low back pain in occupational work[2, 3)]. Lifting tasks are often conducted with 2 types of posture, namely the squat posture, with the knees and hips flexed and the back extended, and the stoop posture, with the hips and back flexed and the knees extended. Previous studies have compared low back load between these 2 conditions. The squat technique is widely recommended to prevent low back pain while conducting lifting tasks. However, Van Dieën et al.[4)] reported in a systematic review that no difference in low back load was observed between manual lifting tasks with the squat posture and those with the stoop posture. A large trunk forward bending angle is needed in combination with increase of pelvic forward tilt in the stoop posture. Less trunk forward bending angle in the squat posture than in the stoop posture, but the pelvic forward tilt angle decreases. Therefore, the appropriate posture for minimizing low back load in lifting tasks is still unclear. Low back load during a lifting task is biomechanically and directly affected by the lever arm, which is the distance from the center of the rotation of the low back joint to the center of gravity of the object. Accordingly, increasing pelvic forward tilt while executing the lifting task would decrease the lever arm, and thus, it might decrease low back load

*Corresponding author. Shota Hayashi (E-mail: 12S1090@g.iuhw.ac.jp)

during the lifting task. However, no previous study has compared low back load between with and without increase of pelvic forward tilt during the lifting task. Hence, the purpose of this study was to examine the effect of an instruction to increase pelvic forward while lifting an object with the squat and stoop postures on low back load.

SUBJECTS AND METHODS

The participants were 10 healthy male university students (mean ± SD: age, 20.9 ± 0.5 years; height, 174.9 ± 4.3 cm; weight, 64.1 ± 4.8 kg). The ethics committee of International University of Health and Welfare approved all study procedures (No. 12-210), which were consistent with the Declaration of Helsinki. The authors obtained informed consent from all the subjects prior to their participation in the study.

The experimental tasks were 3 trials of lifting an object from force plates under the following 4 conditions: 1) squat, hips and knees flexed and the back extended; 2) stoop, hips and back flexed and the knees extended; 3) squat and 4) stoop, respectively with an instruction to increase pelvic forward tilt. To increase pelvic forward tilt, participants were instructed to move their navel closer. After practicing each lifting posture with the instruction to increase pelvic forward tilt several times, subjects conducted the lifting tasks. In a pilot study, the subjects who performed squat and stoop lifting without an instruction to increase pelvic forward tilt changed their lifting maneuvers after conducting the lifting with the instruction. Therefore, the subjects first lifted the object with squat and stoop without any instruction in a random order, and then they lifted the object with squat and stoop with the instruction to increase pelvic forward tilt in a random order. Subjects lifted a box (37.5 × 50 × 21 cm; 11.3 kg) to waist height. A 10-kg weight was placed in the middle on the bottom of the box that weighed 1.3 kg. Previous studies have reported that the distance from the feet of subjects to an object affects the lifting posture and low back load[5, 6]. Hence, in this study, the object was placed on the centerline of 2 force plates, at one half the length of the foot from the toe, as in a previous study[7]. In addition, experiments were conducted after repeatedly practicing each task, and the subjects had enough time to rest, at least 5 minutes, between tasks (Fig. 1).

A 3-dimensional motion analysis system consisting of 10 infrared cameras (Vicon MX, Vicon, Oxford, UK) and 4 force plates (AMTI, Watertown, MA, USA) was used to record 3-dimensional marker displacements and ground reaction force data at a sampling frequency of 100 Hz. Forty-five reflective markers were attached to each subject according to the procedure described in the study of Katsuhira et al[7]. In addition, 4 markers were also attached to the upper frame of the box.

Several studies have used electromyography (EMG) to evaluate low back load during lifting tasks[8, 9]. Several studies have also used 3D motion analysis systems to measure low back compression force and low back shear force as parameters of low back load. The analysis of low back compression force has the advantage, that it can be compared with the safe limit proposed by National Institute of Occupational Safety and Health (NIOSH)[10]. Low back compressive and shear forces were chosen as indicators of low back load in the present study. The computation methods reported by Yamazaki et al.[11] and Katsuhira et al.[7] were used to obtain these forces in our study. Katsuhira et al. reported that low back compressive and shear forces almost simultaneously show peak values[7]. Therefore, they extracted the shear force at the time of the peak of the low back compressive force. As the same tendency was confirmed in our pilot study, the shear force was calculated at the time of the peak of the low back compressive force, and the pelvic angle and lever arm from the L4/5 joint to the center of the gravity of were calculated at the same time. Low back compressive and shear forces were normalized using the subjects'

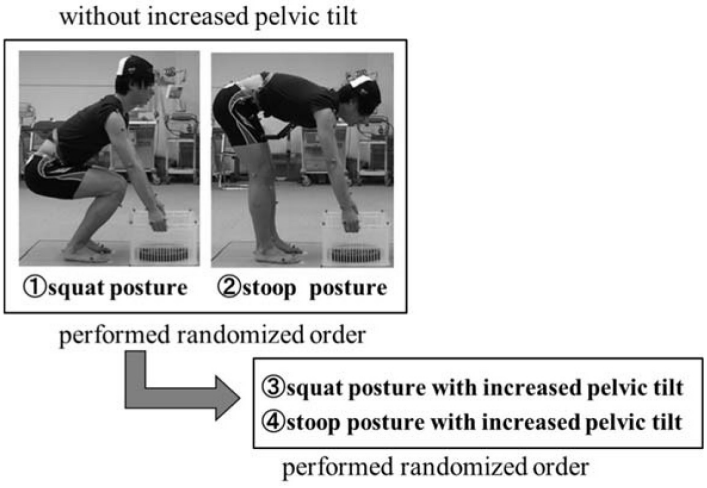


Fig. 1. Experimental condition

body masses to offset the difference in physical attributes between the subjects, in accordance with the method described in a previous study[7]. Moreover, the actual values of the low back compressive and shear forces before the normalization using body mass were compared with the safe limits of the compressive force reported by the NIOSH[10], and the shear force reported by Gallagher et al[12].

The paired t-test was used to assess individual differences between with and without the instruction to increase pelvic forward tilt in each posture. In addition, repeated-measures analysis of variance (ANOVA) was used to compare the differences among the 4 experimental conditions, and the Bonferroni post hoc test was conducted to identify which lifting condition showed the minimum value of low back load. P values < 0.05 were considered statistically significant. Statistical analysis was conducted by using the software package SPSS version 20 (IBM Inc., Armonk, NY, USA).

RESULTS

The mean values of the low back compression and shear forces are shown in Table 1. In the comparison between conditions with and without the instruction to increase pelvic forward tilt, the paired t-test showed there was a significant decrease in low back compression force in the squat posture with the instruction to increase pelvic forward tilt, but not in the stoop posture. In addition, one-way repeated-measures ANOVA and the post hoc test showed there was a significant increase in the low back compressive force in the squat posture without the instruction to pelvic forward tilt, compared the other 3 conditions. The mean increase peak values of the low back compression force in the present study were compared with the safety limit recommended by NIOSH, which is 3400 N. Low back compression force exceeded the safe limit under all 4 conditions.

The paired t-test showed there were no significant differences in standardized low back shear force in both the squat and stoop postures between with and without the instruction to pelvic forward tilt. Moreover, the one-way repeated-measures ANOVA and post hoc test showed there was a significantly smaller value of the low back shear force in the squat posture with the instruction to increase pelvic forward tilt than in the other 3 conditions.

In the comparison of the present results of low back shear force to the safe limit of the shear force reported by Gallagher et al., low back shear forces under all 4 conditions were lower than the safe limit.

The mean pelvic forward tilt angle and distance from the low back joint to the center of gravity of the object are shown in Table 2. The paired t-test showed there was a significant increase in pelvic forward tilt in the squat posture when subjects were instructed to increase pelvic forward tilt, but not in the stoop posture. Also, there was a significant decrease in the lever arm from the low back joint to the center of gravity of the object in the squat posture with the instruction to increase pelvic forward tilt but not in the stoop posture.

Table 1. Mean values of low back compression and shear forces

	Squat posture		Stoop posture		
	without increased pelvic tilt	with increased pelvic tilt	without increased pelvic tilt	with increased pelvic tilt	ANOVA
Normalized low back compression force (N/kg)	66.0±4.5*	59.5±5.5	59.60±5.1	58.2±4.9	a*, b*,c*
Low back compression force (N)	4,219.3±4.5	3,819.2±485.6	3,820.2±441.1	3,725.5±363.9	a*, b*,c*
Normalized low back shear force (N/kg)	1.5±0.5	1.15±0.6	1.8±0.3	1.8±0.4	a*,c*,d*,e*
Low back shear force (N)	92.7±31.1	75.4±38.7	113.4±20.4	117.1±25.1	a*,b*,c*,d*,e*

Mean ± SD, *p<0.05
a: Squat posture without increased pelvic tilt vs. squat posture with increased pelvic tilt
b: Squat posture without increased pelvic tilt vs. stoop posture without increased pelvic tilt
c: Squat posture without increased pelvic tilt vs. stoop posture with increased pelvic tilt
d: Squat posture with increased pelvic tilt vs. stoop posture without increased pelvic tilt
e: Squat posture with increased pelvic tilt vs. stoop posture with increased pelvic tilt
f: Stoop posture without increased pelvic tilt vs. stoop posture with increased pelvic tilt

Table 2. Mean values of pelvic forward tilt angle and distance from the low back joint to the center of gravity of the object

	Squat posture		Stoop posture	
	without increased pelvic tilt	with increased pelvic tilt	without increased pelvic tilt	with increased pelvic tilt
Pelvis forward tilt angle (°)	17.8±17.2	31.1±23.6*	41.3±12.5	44.2±12.7
Distance from low back joint to Center of gravity of an object (mm)	611.4±50.2*	569.2±45.4	505.3±38.9	487.6±35.7

Mean ± SD, *p<0.05

DISCUSSION

Giving an instruction to increase pelvic forward tilt significantly decreased the low back compression force only during lifting in the squat posture. Accordingly, an instruction to increase pelvic forward tilt might be more beneficial to decrease low back load during lifting in the squat posture than in the stoop posture. Low back compression force is an indicator of low back load which is related to low back joint moment. Low back extension moment especially relates to low back compression force in lifting tasks[7]. The distance from the L4/5 joint to the center of gravity of the object or the center of gravity of the head, trunk and arms is defined as the lever arm of the low back extension moment.

When the instruction to increase pelvic forward tilt was given in the squat position, pelvic forward tilt significantly increased with a significant decrease in the lever arm from the low back joint to the center of gravity of the object. The increase in pelvic forward tilt moved the L4/5 joint forward resulting in a decrease in the lever arm, and thus decreased the low back compressive force during lifting in the squat posture. However, no significant differences in the pelvic forward tilt or lever arm were found between with and without the instruction to increase pelvic forward tilt in the stoop posture. Lifting in the stoop posture is requires increase of pelvic forward tilt. Therefore, further increase in pelvic forward tilt might be difficult to perform.

Normalized low back shear force was the smallest in the squat posture with pelvic tilt. The trunk bending angle in the squat posture was smaller than that in the stoop posture. Low back shear force was calculated as the anteroposterior direction force applied to the L4/5 joint. Thus, a small trunk bending angle could decrease the low back shear force. Moreover, increasing pelvic forward tilt increases lumbar lordosis, which would have contributed to the decrease in the low back shear force.

Normalized low back compressive force during lifting in the squat posture without pelvic tilt was the greatest. No significant differences were observed among the other 3 conditions. Normalized low back shear force was significantly smaller during lifting in the squat posture with pelvic tilt. The low back compressive force exceeded the safe limit of 3400 N proposed by NIOSH[10]. Thus, smaller low back shear force would be advantageous the prevention of the risk of low back pain. The values of the low back shear force under all 4 conditions were lower than the safe limit of 700 N proposed by Gallagher et al[12]. However, a previous study suggested that even a small low back shear force might cause damage, resulting in spondylolysis[13]. Hence, the squat posture with an instruction to increase pelvic forward tilt, which can decrease both low back compressive and shear forces, be the recommended lifting posture.

The present study had several limitations. First, low back load was calculated using inverse kinematics. Hence, smaller low back load values were obtained than the actual values of low back load during co-contraction of both the abdominal and back muscles. The authors intend to construct a hybrid model using electromyography and inverse kinematics to obtain the low back load, taking into account co-contraction, in a future study. Second, the subjects of our study were healthy university students. Accordingly, the authors intend to study workers who engage in lifting tasks to confirm the effects of increasing pelvic forward tilt. The authors also intend to investigate the effects of work environment and mental conditions to clarify factors influencing low back load in lifting tasks.

In this study, the effects of an instruction to promote pelvic tilt on low back load during lifting an object from the ground were examined. Making workers aware of pelvic forward tilt during lifting in the squat posture could decrease both low back compressive and shear forces and might lower the incidence of low back pain. Low back pain caused by lifting in work settings has been a problem in both developing and developed countries. The authors recommend the lifting posture identified in this study and suggest that providing education on lifting posture would benefit workers who engage in lifting.

ACKNOWLEDGEMENT

This work was supported by Industrial Disease Clinical Research Grants (14020301-01).

REFERENCES

1) Fujii T, Matsudaira K: Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. Eur Spine J, 2013, 22: 432–438. [Medline] [CrossRef]
2) Hoogendoorn WE, Bongers PM, de Vet HC, et al.: Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: results of a prospective cohort study. Spine, 2000, 25: 3087–3092. [Medline] [CrossRef]
3) Waddell G, Burton AK: Occupational health guidelines for the management of low back pain at work: evidence review. Occup Med (Lond), 2001, 51: 124–135. [Medline] [CrossRef]
4) van Dieën JH, Hoozemans MJ, Toussaint HM: Stoop or squat: a review of biomechanical studies on lifting technique. Clin Biomech (Bristol, Avon), 1999, 14: 685–696. [Medline] [CrossRef]

5) Sasaki M, Horio A, Wakasa M, et al.: Influence of quadriceps femoris fatigue on low back load during lifting of loads at different distances from the toes. J Phys Ther Sci, 2008, 20: 81–89. [CrossRef]
6) Kingma I, Faber GS, Bakker AJ, et al.: Can low back loading during lifting be reduced by placing one leg beside the object to be lifted? Phys Ther, 2006, 86: 1091–1105. [Medline]
7) Katsuhira J, Matsudaira K, Iwakiri K, et al.: Effect of mental processing on low back load while lifting an object. Spine, 2013, 38: E832–E839. [Medline] [CrossRef]
8) Yoon JG: The correlation between the muscle activity and joint angle of the lower extremity according to the changes in stance width during a lifting task. J Phys Ther Sci, 2013, 25: 1023–1025. [Medline] [CrossRef]
9) In-gyu Y, Won-gyu Y: Effects of different transfer direction of manual material handling on trunk and lower extremity Muscles. J Phys Ther Sci, 2012, 12: 1281–1282.
10) Waters TR, Putz-Anderson V, Garg A, et al.: Revised NIOSH equation for the design and evaluation of manual lifting tasks. Ergonomics, 1993, 36: 749–776. [Medline] [CrossRef]
11) Yamazaki N, Yamamoto S, Inoue T: Measurement of transferring motions and evaluation of caregiver's low back load. Baiomekanizumu, 2000, 15: 195–205 (in Japanese).
12) Gallagher S, Marras WS: Tolerance of the lumbar spine to shear: a review and recommended exposure limits. Clin Biomech (Bristol, Avon), 2012, 27: 973–978. [Medline] [CrossRef]
13) Cyron BM, Hutton WC: The fatigue strength of the lumber neural arch in spondylolysis. J Bone Joint Surg Br, 1981, 60B: 234–238.

PLOS | ONE

RESEARCH ARTICLE

Descriptive Epidemiology of Somatising Tendency: Findings from the CUPID Study

Sergio Vargas-Prada[1,2,3], David Coggon[4,5]*, Georgia Ntani[4,5], Karen Walker-Bone[4,5], Keith T. Palmer[4,5], Vanda E. Felli[6], Raul Harari[7], Lope H. Barrero[8], Sarah A. Felknor[9,10], David Gimeno[9], Anna Cattrell[11], Matteo Bonzini[12], Eleni Solidaki[13], Eda Merisalu[14], Rima R. Habib[15], Farideh Sadeghian[16], M. Masood Kadir[17], Sudath S. P. Warnakulasuriya[18], Ko Matsudaira[19], Busisiwe Nyantumbu[20,21], Malcolm R. Sim[22], Helen Harcombe[23], Ken Cox[4], Leila M. M. Sarquis[24], Maria H. Marziale[25], Florencia Harari[7], Rocio Freire[7], Natalia Harari[7], Magda V. Monroy[8], Leonardo A. Quintana[8], Marianela Rojas[26], E. Clare Harris[4,5], Consol Serra[1,2,3,27], J. Miguel Martinez[28], George Delclos[1,2,3,9], Fernando G. Benavides[1,2,3], Michele Carugno[29], Marco M. Ferrario[12], Angela C. Pesatori[29,30], Leda Chatzi[13], Panos Bitsios[31], Manolis Kogevinas[2,3,32], Kristel Oha[33], Tiina Freimann[34], Ali Sadeghian[35], Roshini J. Peiris-John[36,37], Nalini Sathiakumar[38], A. Rajitha Wickremasinghe[39], Noriko Yoshimura[40], Helen L. Kelsall[22], Victor C. W. Hoe[41], Donna M. Urquhart[22], Sarah Derrett[42], David McBride[23], Peter Herbison[23], Andrew Gray[23], Eduardo J. Salazar Vega[43]

1 Center for Research in Occupational Health (CiSAL), Universitat Pompeu Fabra, Barcelona, Spain, **2** CIBER of Epidemiology and Public Health, Barcelona, Spain, **3** IMIM (Hospital del Mar Research Institute), Barcelona, Spain, **4** Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, United Kingdom, **5** Arthritis Research UK/MRC Centre for Musculoskeletal Health and Work, University of Southampton, Southampton, United Kingdom, **6** School of Nursing, University of São Paulo, São Paulo, Brazil, **7** Corporación para el Desarrollo de la Producción y el Medio Ambiente Laboral–IFA (Institute for the Development of Production and the Work Environment), Quito, Ecuador, **8** Department of Industrial Engineering, School of Engineering, Pontificia Universidad Javeriana, Bogotá, Colombia, **9** Southwest Center for Occupational and Environmental Health, The University of Texas Health Science Center at Houston School of Public Health, Houston, Texas, United States of America, **10** Center for Disease Control and Prevention/National Institute for Occupational Safety and Health, Atlanta, Georgia, United States of America, **11** North East London NHS Foundation Trust, Goodmayes Hospital, Ilford, United Kingdom, **12** Epidemiology and Preventive Medicine Research Center, University of Insubria, Varese, Italy, **13** Department of Social Medicine, Medical School, University of Crete, Heraklion, Greece, **14** Institute of Technology, Estonian University of Life Sciences, Tartu, Estonia, **15** Department of Environmental Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon, **16** Department of Occupational Health, School of Public Health, Shahroud University of Medical Sciences, Shahroud, Iran, **17** Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan, **18** Department of Medical Education and Health Sciences, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka, **19** Department for Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo Hospital, Tokyo, Japan, **20** National Institute for Occupational Health, National Health Laboratory Service, Johannesburg, South Africa, **21** Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa, **22** Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia, **23** Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand, **24** Federal University of Paraná, Curitiba-PR, Brazil, **25** School of Nursing of Ribeirão Preto, University of São Paulo, São Paulo, Brazil, **26** Program Health, Work and Environment in Central America, Institute for Studies on Toxic Substances (IRET), National University of Costa Rica, Heredia, Costa Rica, **27** Occupational Health Service, Parc de Salut MAR, Barcelona, Spain, **28** Servicio de Investigación y Análisis IT/EP, Departamento de Investigación y Análisis de Prestaciones, MC Mutual, Barcelona, Spain, **29** Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy, **30** Fondazione Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy, **31** Department of Psychiatry, Medical School, University of Crete, Heraklion, Greece, **32** Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain, **33** North Estonia Medical Centre, Tallinn, Estonia, **34** Tartu University Hospital, Tartu, Estonia, **35** Klinikum Leverkusen, Leverkusen, Germany, **36** Department of Physiology, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka, **37** Section of Epidemiology and Biostatistics, School of Population Health, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand,

Citation: Vargas-Prada S, Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, et al. (2016) Descriptive Epidemiology of Somatising Tendency: Findings from the CUPID Study. PLoS ONE 11(4): e0153748. doi:10.1371/journal.pone.0153748

Editor: Carlos M. Isales, Georgia Regents University, UNITED STATES

Received: December 16, 2015

Accepted: April 4, 2016

Published: April 29, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Monash University funded data collection in Australia through its grant schemes; National Health and Medical Research Council (Australia) supported HLK and DMU through fellowships; the Ministry of Higher Education in Malaysia supported VCWH in Australia; and the Health Research Council of New Zealand funded data collection in New Zealand. Data collection in Central America and Colombia was supported by the Southwest Center for Occupational and Environmental Health at the University of Texas Health Science Center research

training grant from the NIH Fogarty International Center.

Competing Interests: EJSV is now employed by AkzoNobel, USA, but his contribution to this study was completed before his employment began. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

38 Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **39** Faculty of Medicine, University of Kalaniya, Kelaniya, Sri Lanka, **40** Department of Joint Disease Research, 22nd Century Medical and Research Center, University of Tokyo, Tokyo, Japan, **41** Centre for Occupational and Environmental Health, Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia, **42** Injury Prevention Research Unit, Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand, **43** Health Safety and Environment Department, AkzoNobel, Houston, Texas, United States of America

* dnc@mrc.soton.ac.uk

Abstract

Somatising tendency, defined as a predisposition to worry about common somatic symptoms, is importantly associated with various aspects of health and health-related behaviour, including musculoskeletal pain and associated disability. To explore its epidemiological characteristics, and how it can be specified most efficiently, we analysed data from an international longitudinal study. A baseline questionnaire, which included questions from the Brief Symptom Inventory about seven common symptoms, was completed by 12,072 participants aged 20–59 from 46 occupational groups in 18 countries (response rate 70%). The seven symptoms were all mutually associated (odds ratios for pairwise associations 3.4 to 9.3), and each contributed to a measure of somatising tendency that exhibited an exposure-response relationship both with multi-site pain (prevalence rate ratios up to six), and also with sickness absence for non-musculoskeletal reasons. In most participants, the level of somatising tendency was little changed when reassessed after a mean interval of 14 months (75% having a change of 0 or 1 in their symptom count), although the specific symptoms reported at follow-up often differed from those at baseline. Somatising tendency was more common in women than men, especially at older ages, and varied markedly across the 46 occupational groups studied, with higher rates in South and Central America. It was weakly associated with smoking, but not with level of education. Our study supports the use of questions from the Brief Symptom Inventory as a method for measuring somatising tendency, and suggests that in adults of working age, it is a fairly stable trait.

Introduction

Somatising tendency is a predisposition to be unusually aware of, and to worry about, common somatic symptoms [1]. It can be measured through instruments such as the Somatic Symptom Scale [2], the Modified Somatic Perception Questionnaire [3], and a scale derived from the Brief Symptom Inventory [4], and is associated with various aspects of health and health-related behaviour. These include musculoskeletal pain [5–8], especially at multiple sites [9–15], sickness absence from work [16,17], medical consultation [18] and dissatisfaction with medical care [18]. Moreover, the relationship to pain has been observed in longitudinal as well as cross-sectional studies, indicating that tendency to somatise predicts, and is not simply a consequence of, other aspects of health [4–8,14,19,20].

In view of its potential to explain differences in health and behaviour, it is important to understand better the nature of somatising tendency and its descriptive epidemiology. It would be helpful to establish: i) how it can be assessed most efficiently (avoiding redundant information); ii) whether it should be viewed as a long-term trait or a variable state; iii) how it relates to personal characteristics such as sex, age and level of education; and iv) whether it varies

importantly between countries and cultures. To explore these questions, we used data from the Cultural and Psychosocial Influences on Disability (CUPID) study, a large international longitudinal investigation of musculoskeletal pain and its determinants [21].

Methods

The design of the CUPID study and its methods of data collection have been reported in detail elsewhere [21]. In brief, the study sample comprised a total of 12,426 participants aged 20–59 years from 47 occupational groups in 18 countries. The occupational groups fell into three broad categories–nurses (including nursing assistants), office staff who regularly used computers, and other workers (mainly manual employees carrying out repetitive tasks with their hands or arms). Each of the 12,426 participants completed a baseline questionnaire (either by self-administration, or in some occupational groups at interview), representing an overall response rate of approximately 70% among those who were eligible for inclusion [21]. After a mean interval of 14 months (range 3–35 months, 84% within 11–19 months), participants in 45 of the 47 occupational groups (n = 11,992) were asked to complete a shorter follow-up questionnaire, and responses were obtained from 9,305 (78%).

The questionnaires were originally drafted in English, and were then translated into local languages where necessary, accuracy being checked by independent back-translation. Among other things, the baseline questionnaire covered sex; age; age of completing full-time education; smoking habits; experience of pain in the past month at each of ten anatomical sites (low back; neck; and right and left shoulder, elbow, wrist/hand and knee) illustrated by diagrams; duration of sickness absence in the past 12 months because of illness other than a problem with the back, neck, upper limb or knees; and somatising tendency.

Somatising tendency was assessed through questions taken from the Brief Symptom Inventory [4], which asked how distressed or bothered (on a five-point ordinal scale from "not at all" to "extremely") the participant had been during the past seven days by each of: faintness or dizziness, pains in the heart or chest, nausea or upset stomach, trouble getting breath, numbness or tingling in parts of the body, feeling weak in parts of the body, and hot or cold spells. A symptom was deemed to occur if it was at least moderately distressing (i.e. in the highest three of the five levels). The same questions were asked both at baseline and at follow-up.

Statistical analysis was carried out with Stata (StataCorp LP 2012, Stata Statistical Software: Release 12.1, College Station, Texas, USA). Pairwise associations between somatic symptoms at baseline were summarised by odds ratios adjusted for sex and age, as were those between symptoms at baseline and at follow-up.

To explore the clustering of symptoms within individuals, we compared the frequency with which a given number of symptoms was reported with the frequency that would have been expected given the overall prevalence of each symptom, and assuming that their occurrence was mutually independent (for example, that experience of chest pain did not make it more or less likely that an individual would suffer from numbness or tingling). Within each of eight strata defined by combinations of sex and 10-year age band, the expected frequency of each possible combination of symptoms was calculated. These expected frequencies were then summed for combinations representing the same total number of symptoms, and the totals further summed across the eight strata to give the overall number of participants who would be expected to have that number of symptoms.

The relationship of different counts of somatic symptoms to multi-site pain in the past month (defined as pain at $\geq$4 of 10 anatomical sites) was assessed by Poisson regression, with adjustment for sex and age. Possible clustering of the pain outcome by occupational group was taken into account by random intercept, multi-level modelling. Associations were summarised

by prevalence rate ratios (PRRs) with 95% confidence intervals (95%CIs) based on robust standard errors. To explore whether somatising tendency could be adequately characterised without asking about all seven symptoms, we repeated the analysis, excluding data on specific symptoms in turn, and compared population attributable fractions (PAFs–defined as the proportions of cases in a population that would be eliminated if all people had the same risk as those in the reference category). Confidence intervals for PAFs were derived by bootstrapping. To check that findings were not specific to associations with multi-site pain, we repeated the analyses with an alternative outcome–absence from work for >5 days in the past year for reasons other than a problem with the back, neck, upper limb or knees.

We used simple descriptive statistics to summarise changes in the occurrence of somatic symptoms from baseline to follow-up, and the prevalence of symptoms by occupational group. To test whether there was greater similarity in the occurrence of symptoms within as compared to between countries, we calculated the intra-class correlation coefficient (ICC) for the mean numbers of symptoms by occupational group.

We also investigated the possibility that some occupational groups might have a different profile of somatic symptoms from others. For each combination of occupational group and symptom, we compared the number of participants in the group who reported the symptom, with the number that would have been expected to report it if, after allowance for sex and age, the frequency of the symptom as a proportion of all symptom reports in the occupational group were the same as that in the full study sample. A ratio of observed to expected greater than one was an indication that the occupational group experienced the symptom more often than would have been expected, given their overall tendency to somatise.

Finally, we used Poisson regression to assess the (mutually adjusted) cross-sectional associations of somatising tendency at baseline (defined as report of ≥3 somatic symptoms) with possible risk factors (sex, age, smoking habits and age finished full-time education). Again random intercept modelling was used to allow for possible clustering by occupational group.

Ethical approval for the study was provided by the relevant research ethics committee or institutional review board in each participating country (S1 Appendix).

Results

In one occupational group (office workers in Colombia), one of the questions about somatic symptoms had been omitted. Complete data on somatic symptoms at baseline were available for 12,072 men and women from the remaining 46 occupational groups (98% of all participants from those groups). Table 1 shows the prevalence of each symptom by sex and age. Among men, the prevalence of all symptoms except numbness or tingling was highest in the youngest age group (20–29 years). Women reported each of the seven symptoms more frequently than men, and particularly nausea or upset stomach, hot or cold spells (especially at older ages), and numbness or tingling (again more at older ages). Moreover, in contrast to men, the only symptoms that were most common at age 20–29 years were faintness or dizziness and nausea or upset stomach. In view of these differences, all subsequent analyses were adjusted for sex and age.

Table 2 summarises the associations between pairs of somatic symptoms at baseline. The strongest associations were for pain in the heart or chest with trouble getting breath (OR 9.3), and feeling weak in parts of the body with numbness or tingling in parts of the body (OR 7.9). However, all symptoms were associated with each other, the lowest odds ratio being 3.4.

Table 3 compares the frequency with which specified numbers of symptoms were reported and the frequency that would have been expected had the occurrence of each symptom been statistically independent. More participants than expected reported no symptoms at all (6,016

Table 1. Baseline prevalence (%) of distressing somatic symptoms in past 7 days by sex and age.

Symptom	Men				Women			
	20–29 years	30–39 years	40–49 years	50–59 years	20–29 years	30–39 years	40–49 years	50–59 years
	(N = 1,056)	(N = 1,379)	(N = 1,170)	(N = 641)	(N = 1,954)	(N = 2,487)	(N = 2,172)	(N = 1,213)
Faintness or dizziness	8.0 (85)	7.2 (99)	6.2 (73)	4.8 (31)	17.3 (339)	15.9 (395)	15.1 (328)	12.0 (146)
Pains in heart or chest	10.1 (107)	7.0 (97)	5.8 (68)	5.8 (37)	9.6 (188)	10.0 (248)	12.4 (269)	10.8 (131)
Nausea or upset stomach	16.3 (172)	12.8 (177)	11.4 (133)	9.7 (62)	27.0 (528)	25.3 (630)	22.9 (497)	18.3 (222)
Trouble getting breath	7.1 (75)	5.7 (79)	5.6 (65)	5.5 (35)	10.1 (197)	10.0 (149)	10.6 (230)	10.2 (124)
Hot or cold spells	16.7 (176)	11.8 (163)	10.8 (126)	9.8 (63)	21.6 (423)	21.5 (535)	26.9 (584)	35.1 (426)
Feeling weak in parts of your body	21.3 (225)	17.3 (238)	18.9 (221)	18.6 (119)	26.7 (522)	30.7 (763)	30.9 (671)	28.3 (343)
Numbness or tingling in parts of your body	14.8 (156)	11.6 (160)	16.0 (187)	14.8 (95)	17.2 (336)	25.0 (621)	30.8 (670)	29.6 (359)

Figures in brackets are the numbers of participants with the relevant symptom

doi:10.1371/journal.pone.0153748.t001

vs. 3,433). However, there were fewer than expected with 1–3 symptoms. The ratio of observed to expected numbers then increased progressively for report of larger numbers of symptoms, rising from 2.75 for four symptoms to 2000 for all seven symptoms.

Table 3 also shows the associations between the number of somatic symptoms and report of pain at ≥4 of 10 anatomical sites. Relative to no somatic symptoms, PRRs for multisite pain increased progressively from 2.3 (95%CI 2.0–2.7) for one somatic symptom to 5.9 (95%CI 4.8–7.4) for five somatic symptoms, and then remained at a similar level for six and seven symptoms. The right-hand columns of the table give the corresponding population attributable fractions (PAFs) and their 95% CIs. Overall, report of at least one somatic symptom accounted for 59.0% of the cases of multi-site pain in the study sample.

To explore whether information about any of the somatic symptoms was effectively redundant, we repeated the analysis of associations with multi-site pain excluding each of the seven symptoms in turn (Table 4). In each case, the PAF for multi-site pain that was associated with report of at least one of the remaining somatic symptoms was lower than in the analysis that included all somatic symptoms (53.2% to 58.6% vs. 59.0%), indicating that each symptom added to the characterisation of somatising tendency, although an index based on only six of the seven symptoms would still work well.

Table 2. Pairwise associations between specific somatic symptoms at baseline.

Symptom at baseline	Faintness or dizziness	Pains in heart or chest	Nausea or upset stomach	Trouble getting breath	Hot or cold spells	Feeling weak in parts of body
Pains in heart or chest	6.6 (492)					
Nausea or upset stomach	5.8 (802)	4.5 (587)				
Trouble getting breath	5.6 (429)	9.3 (450)	4.4 (547)			
Hot or cold spells	4.0 (706)	3.4 (548)	3.8 (1,050)	4.1 (544)		
Feeling weak in parts of body	5.1 (874)	4.7 (681)	3.9 (1,249)	4.9 (653)	4.9 (1,365)	
Numbness or tingling in parts of body	3.9 (705)	4.4 (607)	3.4 (1,032)	4.7 (587)	3.6 (1,113)	7.9 (1,625)

Associations are summarised by odds ratios adjusted for sex and age, with the number of participants reporting both symptoms in brackets

doi:10.1371/journal.pone.0153748.t002

Table 3. Observed and expected frequency of multiple somatic symptoms and associations with multi-site pain.

Number of somatic symptoms	Observed number of subjects	Expected number of subjects[a]	Ratio of observed to expected	Association with pain at ≥4 vs. 0 anatomical sites					
				Number with no pain	Number with pain at ≥4 sites	PRR[b]	(95% CI)	PAF[c] (%)	(95% CI)
0	6,016	3,433	1.75	3,125	374	1			
1	2,312	4,546	0.51	779	291	2.3	(2.0–2.7)	8.9	(6.8–11.0)
2	1,551	2,817	0.55	336	342	4.0	(3.3–4.8)	13.7	(11.5–15.8)
3	944	1,020	0.93	149	295	5.0	(4.1–6.1)	12.6	(10.7–14.6)
4	618	224.6	2.75	86	226	5.1	(4.1–6.5)	9.7	(7.6–11.7)
5	326	29.26	11.1	30	164	5.9	(4.8–7.4)	7.3	(5.5–9.1)
6	185	2.058	89.9	14	104	6.0	(4.6–7.7)	4.6	(3.2–6.0)
7	120	0.060	2000	15	78	5.8	(4.6–7.2)	3.4	(2.2–4.7)
≥1	6,056	8,639.18	0.70	1,409	1,500	3.8	(3.2–4.5)	59.0	(53.5–64.5)
≥3	2,193	1,276.23	1.72	294	867	5.6	(4.4–7.0)	37.9	(30.9–44.9)

[a] Expected number given the overall prevalence of each symptom, and assuming no association between the occurrence of one symptom and another after allowance for sex and age (in four 10-year strata)

[b] Prevalence rate ratio adjusted for sex and age (in four 10-year strata)

[c] Population attributable fraction

doi:10.1371/journal.pone.0153748.t003

To check that these patterns of association were not specific to pain outcomes, we repeated the analyses for Tables 3 and 4, taking as an alternative outcome sickness absence in the past 12 months for non-musculoskeletal reasons. In the analysis that included all seven somatic symptoms, PRRs rose progressively from 1.4 (95%CI 1.1–1.7) for report of one symptom to 3.2 (95%CI 2.4–4.2) for report of seven symptoms, and the PAF for report of at least one somatic symptom was 30.4% (Table 5). The PAFs when single somatic symptoms were disregarded ranged from 27.1% to 30.4% (Table 6).

Complete information about somatic symptoms at follow-up was available for 8,856 (73%) of the participants who provided satisfactory information at baseline, the follow-up rate being similar in those who initially did and did not have symptoms. Table 7 shows the number of somatic symptoms that they reported at follow-up, according to the number that were present at baseline. In general, participants reported similar numbers of symptoms at follow-up as at baseline, 6,677 (75%) having a change of zero or one in their symptom count. There were, however, notable exceptions. In particular, seven participants went from zero symptoms at baseline to seven at follow-up, and 19 changed to the same extent in the reverse direction. More detailed examination of the questionnaire responses for these 26 individuals indicated that for the most part, the changes represented substantial differences in the levels of distress reported from individual symptoms, and not simply a shift from their being "a little bit" to "moderately" distressing.

The 8,856 participants who provided complete information at both time-points reported a total of 10,326 somatic symptoms at baseline. Of these specific symptoms, 3,733 (36%) were

Table 4. Associations of multiple somatic symptoms with multi-site pain when one of the seven somatic symptoms was ignored.

Somatic symptom disregarded	Number of somatic symptoms														
	1		2		3		4		5		6		≥1 somatic symptom		
	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PAF[b] (%)
Faintness or dizziness	2.4	(2.1–2.8)	3.9	(3.3–4.7)	5.1	(4.2–6.3)	5.2	(4.2–6.5)	5.8	(4.6–7.5)	5.7	(4.6–7.1)	3.7	(3.2–4.4)	57.6
Pains in heart or chest	2.5	(2.1–2.9)	4.1	(3.4–5.0)	5.0	(4.1–6.2)	5.4	(4.3–6.8)	5.8	(4.5–7.4)	5.9	(4.7–7.3)	3.8	(3.2, 4.5)	58.6
Nausea or upset stomach	2.5	(2.1–3.0)	4.2	(3.5–5.0)	4.8	(3.9–5.9)	5.5	(4.4–6.8)	5.6	(4.4–7.1)	5.4	(4.4–6.7)	3.8	(3.2–4.4)	57.0
Trouble getting breath	2.4	(2.1–2.7)	4.1	(3.4–5.0)	5.0	(4.1–6.0)	5.3	(4.3–6.6)	6.1	(4.8–7.7)	5.8	(4.6–7.2)	3.8	(3.2–4.4)	58.6
Hot or cold spells	2.5	(2.2–3.0)	4.3	(3.6–5.2)	4.8	(3.8–6.0)	5.3	(4.2–6.7)	5.6	(4.4–7.1)	5.1	(4.0–6.5)	3.8	(3.2–4.5)	56.7
Feeling weak in parts of your body	2.5	(2.1–2.9)	3.7	(3.1–4.3)	4.3	(3.6–5.2)	4.7	(3.8–5.8)	5.1	(4.1–6.5)	4.8	(3.9–5.9)	3.3	(2.9–3.9)	53.2
Numbness or tingling in parts of your body	2.5	(2.2–2.9)	3.8	(3.2–4.6)	4.2	(3.5–5.1)	4.8	(3.9–6.0)	4.8	(3.9–6.0)	5.1	(4.0–6.3)	3.4	(3.0–4.0)	54.4

[a] Prevalence rate ratio, adjusted for sex and age (in four 10-year strata), for pain at ≥4 vs. 0 anatomical sites in participants with the specified number of somatic symptoms compared with no somatic symptoms. The specified number of symptoms was from the total of six that remained when the symptom in the left-hand column was disregarded.

[b] Population attributable fraction

doi:10.1371/journal.pone.0153748.t004

again reported at follow-up, while 6,593 (64%) had resolved. On the other hand, 4,123 (52%) of a total of 7,856 symptoms at follow-up were new since baseline. Table 8 summarises the pairwise associations between specific somatic symptoms at baseline and at follow-up. The highest

Table 5. Associations between number of somatic symptoms and sickness absence for >5 days in past 12 months for non-musculoskeletal reasons.

Number of somatic symptoms	Duration of sickness absence in past 12 months for reasons other than musculoskeletal pain					
	None	>5 days				
	N	N	PRR[a]	(95%CI)	PAF[b] (%)	(95%CI)
0	3,982	408	1			
1	1,377	225	1.4	(1.1–1.7)	5.1	(2.3–7.9)
2	863	200	1.8	(1.6–2.2)	7.7	(5.8–9.6)
3	453	135	2.3	(1.8–2.8)	6.3	(4.4–8.2)
4	266	104	2.8	(2.2–3.4)	5.5	(3.5–7.6)
5	154	60	2.7	(2.1–3.5)	3.2	(1.7–4.7)
6	70	34	2.9	(2.0–4.1)	1.9	(0.8–2.9)
7	44	28	3.2	(2.4–4.2)	1.6	(0.7–2.5)
≥1	3,227	786	1.9	(1.6–2.2)	30.4	(23.9–36.9)
≥3	987	361	2.5	(2.1–3.1)	18.4	(10.9–25.8)

[a] Prevalence rate ratio relative to no sickness absence in past 12 months for non-musculoskeletal reasons, adjusted for sex and age (in four 10-year strata)

[b] Population attributable fraction

doi:10.1371/journal.pone.0153748.t005

Table 6. Associations of multiple somatic symptoms with sickness absence for >5 days in past 12 months for non-musculoskeletal reasons when one of the seven somatic symptoms was ignored.

Somatic symptom disregarded	Number of somatic symptoms														
	1		2		3		4		5		6		≥1 somatic symptom		
	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PRR[a]	(95% CI)	PAF[b] (%)
Faintness or dizziness	1.4	(1.2–1.6)	2.0	(1.7–2.3)	2.4	(1.9–3.0)	2.4	(1.9–3.1)	2.8	(2.1–3.8)	3.0	(2.3–3.9)	1.8	(1.6–2.1)	29.3
Pains in heart or chest	1.4	(1.2–1.7)	1.8	(1.6–2.1)	2.3	(1.9–2.8)	2.8	(2.2–3.5)	2.6	(2.0–3.3)	3.6	(2.7–4.9)	1.8	(1.6–2.2)	29.6
Nausea or upset stomach	1.4	(1.2–1.7)	1.9	(1.6–2.2)	2.4	(2.0–2.9)	2.7	(2.1–3.4)	2.6	(1.9–3.6)	2.9	(2.1–3.9)	1.8	(1.6–2.1)	28.0
Trouble getting breath	1.4	(1.2–1.6)	1.9	(1.6–2.2)	2.1	(1.7–2.7)	2.7	(2.2–3.3)	2.9	(2.1–3.8)	2.9	(2.1–3.9)	1.8	(1.6–2.1)	29.4
Hot or cold spells	1.6	(1.3–1.8)	2.0	(1.7–2.3)	2.5	(2.0–3.1)	2.7	(2.1–3.5)	2.9	(2.2–3.9)	2.9	(2.2–3.8)	2.0	(1.7–2.3)	30.4
Feeling weak in parts of your body	1.4	(1.2–1.6)	1.9	(1.6–2.3)	2.4	(2.0–3.0)	2.9	(2.3–3.6)	3.0	(2.1–4.1)	2.9	(2.2–3.8)	1.8	(1.6–2.1)	27.1
Numbness or tingling in parts of your body	1.5	(1.2–1.7)	2.0	(1.7–2.3)	2.7	(2.3–3.3)	2.6	(2.1–3.3)	2.5	(1.7–3.7)	3.1	(2.4–4.0)	1.9	(1.6–2.2)	29.5

[a] Prevalence rate ratio, adjusted for sex and age (in four 10-year strata), for sickness absence in the past 12 months for non-musculoskeletal reasons vs. 0 days of sickness absence in participants with the specified number of somatic symptoms compared with no somatic symptoms. The specified number of symptoms was from the total of six that remained when the symptom in the left-hand column was disregarded.

[b] Population attributable fraction

doi:10.1371/journal.pone.0153748.t006

odds ratios (3.6 to 6.6) were for continuing presence of the same symptom at follow-up as at baseline, but all odds ratios were ≥1.6, and most were ≥2.0.

Fig 1 shows the prevalence of different numbers of somatic symptoms at baseline by occupational group. There was major variation between the groups–for example, the prevalence of ≥3 somatic symptoms ranged from 1.3% among office workers in Pakistan and 4.2% in sugar cane cutters in Brazil to 38.1% in office workers in Costa Rica and 51.8% in manual workers in Costa Rica. Apart from the Brazilian sugar cane cutters, rates in South and Central America were all relatively high. The mean numbers of symptoms by occupational group showed greater similarity within than between countries (ICC = 15%). However, there was no consistent pattern by type of occupation (nurse, office worker or other).

Table 7. Number of somatic symptoms reported at follow-up according to number of somatic symptoms reported at baseline.

Number of symptoms at baseline	Number of symptoms at follow-up							
	0	1	2	3	4	5	6	7
0	3,329	622	235	107	49	18	10	7
1	885	479	216	108	47	15	2	0
2	452	293	217	101	52	21	12	2
3	230	151	137	98	54	16	14	4
4	141	71	81	75	47	29	19	5
5	51	31	40	41	26	17	7	5
6	20	12	18	16	13	16	9	5
7	19	13	5	10	5	11	8	7

Analysis was restricted to the 8,856 participants who provided complete information about somatic symptoms at both baseline and follow-up.

doi:10.1371/journal.pone.0153748.t007

Table 8. Pairwise associations between specific somatic symptoms at baseline and at follow-up.

Symptom at baseline	Symptom at follow-up						
	Faintness or dizziness (n = 741)	Pains in heart or chest (n = 542)	Nausea or upset stomach (n = 1,254)	Trouble getting breath (n = 558)	Hot or cold spells (n = 1,445)	Feeling weak in parts of body (n = 1,778)	Numbness or tingling in parts of body (n = 1,538)
Faintness or dizziness (n = 1,030)	5.0 (293)	2.6 (149)	2.4 (295)	2.1 (133)	2.1 (304)	2.3 (377)	2.5 (351)
Pains in heart or chest (n = 753)	2.6 (155)	5.6 (182)	2.0 (196)	2.6 (115)	1.8 (216)	2.0 (264)	2.1 (252)
Nausea or upset stomach (n = 1,719)	2.6 (287)	2.0 (190)	3.6 (545)	1.7 (179)	1.6 (421)	2.0 (556)	1.7 (462)
Trouble getting breath (n = 748)	2.6 (148)	3.2 (132)	2.1 (202)	6.6 (199)	1.8 (215)	2.3 (286)	2.2 (253)
Hot or cold spells (n = 1,833)	2.2 (285)	2.1 (219)	2.0 (438)	1.9 (208)	3.9 (709)	2.1 (620)	2.0 (557)
Feeling weak in parts of body (n = 2,317)	2.6 (361)	2.4 (263)	2.1 (538)	2.2 (265)	2.0 (611)	4.3 (983)	2.9 (760)
Numbness or tingling in parts of body (n = 1,926)	2.5 (306)	2.7 (249)	2.0 (452)	1.9 (220)	1.9 (534)	2.7 (729)	5.1 (822)

Associations are summarised by odds ratios adjusted for sex and age (in 10-year strata), with the number of participants reporting both symptoms in brackets. Analysis was restricted to the 8,856 participants who provided complete information about somatic symptoms at both baseline and follow-up.

doi:10.1371/journal.pone.0153748.t008

To explore whether any occupational groups displayed a distinct profile of somatic symptoms, we compared the proportionate frequency of specific symptoms after standardisation for sex and age. The standardised proportions ranged from 0 for hot or cold spells in Brazilian sugar cane cutters and 0.15 for hot or cold spells in Sri Lankan postal workers to 1.98 for nausea or upset stomach in Japanese sales personnel and 2.10 for hot or cold spells in Pakistani postal workers. However, the large majority were between 0.67 and 1.5. The most salient patterns by country were high ratios for trouble getting breath in Brazil (1.28–1.56); low ratios for hot or cold spells in Greece (0.37–0.73); high ratios for faintness or dizziness (1.73 and 1.75) and low ratios for feeling weak (0.31 and 0.49) in Estonia; low ratios for pains in the heart or chest in Lebanon (0.42–0.63); low ratios for each of faintness or dizziness (0.58–0.72), pains in the heart or chest (0.23–0.78) and trouble getting breath (0.28–0.80), and high ratios for hot or cold spells (1.44–2.10) in Pakistan; high ratios for nausea or upset stomach (1.17–1.98) and low ratios for trouble getting breath (0.24–0.53) in Japan; and high ratios for pains in the heart or chest in South Africa (1.57 and 1.77). Further details are given in Table 9.

In a mutually adjusted analysis of the cross-sectional association between personal characteristics and somatising tendency (pragmatically specified as report of $\geq$3 somatic symptoms), there was a significantly elevated risk with female sex (PRR 1.8, 95%CI 1.5–2.1), and a weak but significant relationship to smoking habits (PRRs of 1.3 and 1.2 for current and ex- as compared with non-smokers). However, there was no association with age of finishing full-time education (data not shown).

Discussion

Within our large study sample, the seven somatic complaints that we examined were all mutually associated, such that report of multiple symptoms was much more frequent than would have been expected had their occurrence been unrelated. However, no cut-point in the number

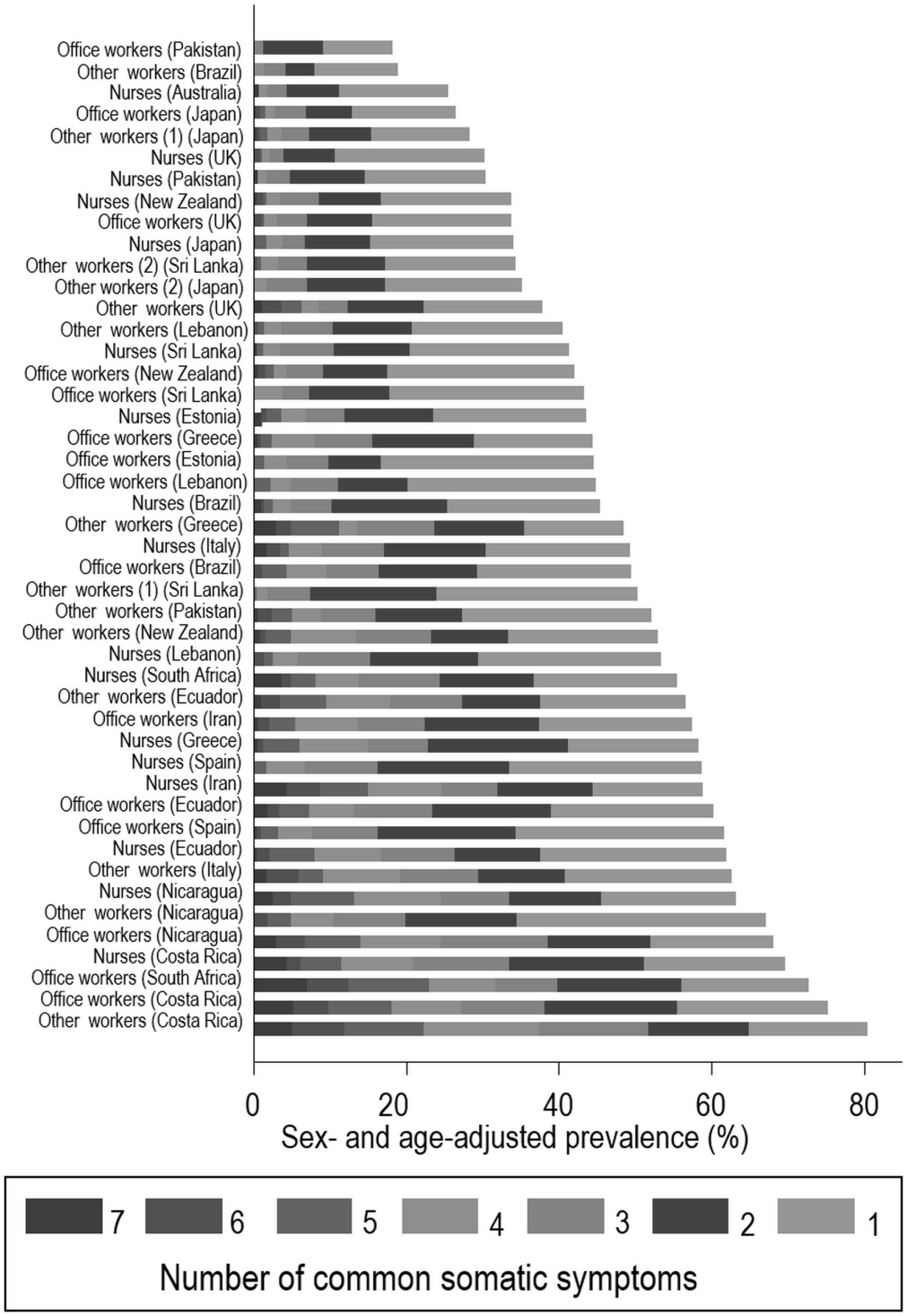


Fig 1. Frequency of somatic symptoms by occupational group.

doi:10.1371/journal.pone.0153748.g001

Table 9. Standardised proportions of specific symptoms by occupational group.

Occupational group	Symptom						
	Faintness or dizziness	Pains in heart or chest	Nausea or upset stomach	Trouble getting breath	Hot or cold spells	Feeling weak in parts of your body	Numbness or tingling in parts of your body
Brazil							
Nurses	0.58	1.03	0.77	1.28	0.80	1.06	1.43
Office workers	0.60	1.43	0.69	1.56	0.89	0.94	1.21
Other workers	1.32	1.27	0.72	1.54	0.00	0.83	1.89
Ecuador							
Nurses	1.13	0.94	1.00	0.63	1.29	0.96	0.84
Office workers	0.63	1.17	1.08	0.87	1.05	1.02	1.05
Other workers	1.06	1.37	0.91	1.12	1.11	0.98	0.75
Costa Rica							
Nurses	0.56	0.94	1.01	0.92	1.12	1.02	1.16
Office workers	0.71	1.02	1.12	1.16	1.15	0.83	1.04
Other workers	0.96	1.18	0.88	1.03	0.90	0.95	1.20
Nicaragua							
Nurses	0.78	0.80	1.03	0.89	1.12	0.91	1.19
Office workers	0.70	1.04	0.89	1.17	1.12	0.95	1.15
Other workers	0.53	0.54	0.83	1.52	1.23	1.07	1.16
UK							
Nurses	0.92	1.28	1.13	0.75	1.18	0.90	0.86
Office workers	1.05	1.06	1.05	1.08	1.05	0.96	0.88
Other workers	1.31	1.14	1.00	1.12	1.07	0.95	0.75
Spain							
Nurses	0.73	0.52	0.93	0.76	0.64	1.56	1.23
Office workers	0.66	0.56	0.89	0.88	1.17	1.24	1.09
Italy							
Nurses	1.16	1.06	1.24	1.18	0.87	0.88	0.86
Other workers	0.94	0.87	0.96	1.44	0.83	1.05	1.06
Greece							
Nurses	1.22	0.74	0.98	0.93	0.37	1.20	1.34
Office workers	1.08	1.18	0.96	0.63	0.73	1.11	1.19
Other workers	1.39	0.91	0.81	1.45	0.58	1.10	1.05
Estonia							
Nurses	1.75	1.77	0.84	1.15	1.03	0.49	0.90
Office workers	1.73	1.29	0.91	1.96	0.95	0.31	1.01
Lebanon							
Nurses	1.07	0.42	1.35	1.14	0.63	1.13	0.96
Office workers	0.71	0.48	1.25	1.68	0.57	1.15	1.11
Other workers	0.95	0.63	0.70	1.44	1.21	0.89	1.29
Iran							
Nurses	1.50	1.32	0.84	1.23	0.90	1.01	0.68
Office workers	1.86	1.05	0.38	0.74	1.40	1.00	0.79
Pakistan							
Nurses	0.58	0.36	0.50	0.36	1.44	1.47	1.54
Office workers	0.64	0.23	0.83	0.28	1.86	1.43	0.68
Other workers	0.72	0.78	0.72	0.80	2.10	1.09	0.57

(*Continued*)

Table 9. (*Continued*)

Occupational group	Symptom						
	Faintness or dizziness	Pains in heart or chest	Nausea or upset stomach	Trouble getting breath	Hot or cold spells	Feeling weak in parts of your body	Numbness or tingling in parts of your body
Sri Lanka							
Nurses	1.07	0.67	1.08	1.51	1.41	0.64	0.80
Office workers	0.76	1.11	1.01	0.87	1.33	0.72	1.18
Other workers (1)	0.83	1.26	0.80	0.99	0.15	1.40	1.20
Other workers (2)	0.82	1.59	1.05	0.79	1.15	0.95	0.79
Japan							
Nurses	1.38	0.87	1.56	0.53	0.90	0.85	0.66
Office workers	1.70	0.80	1.25	0.24	1.14	0.84	0.88
Other workers (1)	1.34	0.80	1.17	0.42	0.94	1.13	0.91
Other workers (2)	1.13	0.68	1.98	0.35	1.05	0.84	0.56
South Africa							
Nurses	1.20	1.57	0.83	1.16	1.06	0.83	0.88
Office workers	1.04	1.77	1.06	1.22	0.98	0.71	0.84
Australia							
Nurses	1.08	0.49	1.33	1.03	0.87	1.07	0.92
New Zealand							
Nurses	0.90	1.09	1.15	0.49	1.21	0.97	0.91
Office workers	0.62	0.58	1.36	0.83	1.12	1.05	0.97
Other workers	0.84	0.85	1.05	0.92	1.08	1.15	0.88

Standardised proportions were calculated as $O/\sum_i(n_i * S_i/N_i)$ where O was the observed frequency of the specified symptom in the occupational group, and within the i^{th} of 8 strata of sex and 10-year age band, n_i was the total number of symptom reports (any of the seven symptoms) in the occupational group, S_i was the number of reports of the specified symptom in all occupational groups combined, and N_i was the total number of symptom reports (any of the seven symptoms) in all occupational groups combined.

doi:10.1371/journal.pone.0153748.t009

of reported symptoms distinguished unequivocally between people with and without a somatising syndrome. Rather, there appeared to be a gradation in degrees of tendency to somatise. In most individuals, the level of somatising tendency (as assessed by the questionnaire) was little changed after a follow-up interval of approximately 14 months, although the specific symptoms reported at follow-up often differed from those at baseline. Tendency to somatise was more common in women than men, especially at older ages, and after allowance for sex and age, it varied markedly across the 46 occupational groups studied, with greater similarities within than between countries. It was weakly associated with smoking, but not with level of education.

As well as the size, geographical spread and cultural diversity of the study sample, our investigation benefitted from high response rates. However, it was limited to adults of working age, and the findings cannot necessarily be extrapolated to other age groups. It was also restricted to selected occupational groups, although apart perhaps from sugar cane cutters in Brazil, it seems unlikely that these will have been highly unrepresentative of the wider working populations in participating countries.

Somatising tendency was assessed through seven questions taken from the Brief Symptom Inventory, which has been established as a valid and reliable instrument [22] with the ability to predict future health outcomes in longitudinal investigations [6–8]. Moreover, where it was necessary to translate the questionnaire into local languages, care was taken to check accuracy through independent back-translation. Nevertheless, it is possible that symptoms were understood differently across varied cultural settings. Such variation may have contributed to differences in prevalence between countries, but would not explain associations with other variables measured at individual level in analyses that adjusted for possible clustering by job group.

We did not have information about personality traits or about other medical conditions such as cancer, which may have caused some of the symptoms that distressed participants. However, since our study sample comprised adults in active employment, the prevalence of serious co-morbidity will have been low, and should not have impacted importantly on our conclusions.

Understanding of terms for pain may have varied between participants speaking different languages, but the anatomical location of symptoms should have been unambiguous, since it was defined pictorially. Errors of interpretation are less likely to have occurred for other variables such as history of sickness absence, smoking habits and educational level, although they may have been liable to inaccurate recall. Provided inaccuracies were not differential in relation to somatising tendency, any resultant bias in associations with somatising tendency will have been towards the null.

Much of the literature on somatisation has focused on medically unexplained somatic symptoms as a reason for presentation to medical care, and a manifestation of hidden psychiatric morbidity. As defined in the tenth revision of the International Classification of Diseases (ICD10), somatisation disorder is generally infrequent, with prevalence rates among adults aged 18–65 years in a cross-cultural study of 14 countries mostly less than 2% [23]. However, our interest was in the wider spectrum of distress from common somatic symptoms, not necessarily leading to medical consultation of themselves, but collectively associated with other aspects of health and health-related behaviour. By limiting our enquiry to symptoms in the past week, we reduced the potential for errors in recall, which can be a problem when longer periods are considered [24].

Our results confirm that report of multiple distressing somatic symptoms constitutes a syndrome, the co-occurrence of symptoms being much more frequent than would be expected by chance. However, there was no clear dichotomy between people with and without somatising tendency. Thus, the strength of associations, both with multi-site pain and with sickness absence for non-musculoskeletal reasons, increased progressively with the number of symptoms reported, at least up to five. Because these associations were cross-sectional, they cannot necessarily be interpreted as causal, although longitudinal studies have indicated that people who complain of common somatic symptoms are more likely to develop multisite musculoskeletal pain subsequently [19,20]. We also found that all seven of the symptoms investigated contributed to the measurement of somatising tendency, with smaller attributable fractions for multi-site pain and non-musculoskeletal sickness absence when any one of the symptoms was disregarded. However, the differences in PAFs were generally small, and if resources were limited, it is likely that little would be lost if any one of the seven symptoms were omitted from the question set.

Follow-up of participants after approximately 14 months demonstrated that levels of somatising tendency were fairly stable within individuals over that timescale, and the observation that this occurred despite changes in the specific symptoms reported is evidence that the consistency reflects a continuing general predisposition to be aware of and report physical symptoms, rather than persistence of specific underlying disease. A similar pattern has been found

in earlier longitudinal studies [24]. It is notable, however, that a small minority of participants exhibited major changes in their degree of somatising tendency, suggesting that it is not entirely a fixed trait, and raising the possibility that it might in some cases be amenable to intervention. Another possibility is that these large changes reflected the development or resolution of co-morbidity.

The higher frequency of somatic symptoms among women than men accords with other studies [25–27]. It has been postulated that the imbalance may reflect innate differences in somatic and visceral perception; differences in symptom labelling, description and reporting; or a greater willingness of women to acknowledge and disclose discomfort [25]. It could also arise from a higher prevalence of depression in women.

Somatisation has also been reported to occur more commonly at older ages [23]. We too found a positive relationship to age in women, although in men, the prevalence of somatic symptoms was highest at younger ages. Because our analysis was cross-sectional, it was not possible to distinguish effects of age from trends across birth cohorts. However, the higher prevalence of hot or cold flushes among older women may have been a physiological effect of age.

The large differences between occupational groups and countries in the prevalence and degree of somatising tendency were apparent even after adjustment for differences in sex and age. As already discussed, the variation may have been, at least in part, a linguistic artefact. However, earlier research using different methods has also indicated unusually high rates of somatisation in South America [23]. In that study, there was no evidence that somatising patients from South America had a lower prevalence of co-occurring depression or generalised anxiety disorder, which suggests that their somatisation was not a manifestation of occult mental illness. Perhaps more likely is a culturally determined difference in the perception of bodily sensations and the importance that is attached to them, or in willingness to report them when they occur. There was also variation between countries in the relative frequency of specific somatic symptoms, but to a lesser extent.

Somatisation has previously been linked with an absence of formal education [23], but after allowance for sex, age and occupational group, we found no relationship to level of education. This may have been because within occupational groups there was too little heterogeneity for an effect to be discernible. We did, however, find a weak association with smoking, which is consistent with an earlier study in Finnish adolescents [28].

In summary, our study supports the use of questions from the Brief Symptom Inventory as a method for measuring tendency to somatise, each of the seven questions contributing to its assessment. The findings indicate that somatising tendency should be regarded as a quantifiable characteristic that exhibits an exposure-response relationship in its association with other health measures, and appears to be fairly stable over an interval of approximately one year, although the specific symptoms that individuals report frequently vary over time. It is more common in women than in men, especially at older ages, and its prevalence varies between countries with higher rates in South and Central America.

Given its potential to explain differences in disability and in economically important outcomes such as sickness absence from work, there is a need to understand further what drives somatising tendency, and whether and how it might be modified at a population level. There is evidence, for example, that it tracks across generations [29], and it may be a trait which is acquired early in life. Thus, there is a need for further research to establish how it evolves at younger ages, what influences its development, and also how constant it remains over longer follow-up periods.

Supporting Information

S1 Appendix. Committees which provided ethical approval for the CUPID study.
(DOCX)

S1 Dataset. Supporting Dataset.
(DTA)

S1 Metadata. Metadata.
(XLSX)

Acknowledgments

We thank: Pietro Muñoz, Patricio Oyos, Gonzalo Albuja, María Belduma and Francisco Lara for their assistance with data collection in Ecuador; Patrica Monge, Melania Chaverrri and Freddy Brenes, who helped with data collection in Costa Rica; Aurora Aragón, Alberto Berríos, Samaria Balladares and Martha Martínez who helped with data collection in Nicaragua; Alfredo José Jirón who assisted with data entry in Nicaragua; Catalina Torres for translation and piloting of the questionnaire in Spain; Ben and Marie Carmen Coggon for back translation of the Spanish questionnaire; Cynthia Alcantara, Xavier Orpella, Josep Anton Gonzalez, Joan Bas, Pilar Peña, Elena Brunat, Vicente San José, Anna Sala March, Anna Marquez, Josefina Lorente, Cristina Oliva, Montse Vergara and Eduard Gaynés for their assistance with data collection in Spain; Natale Battevi, Lorenzo Bordini, Marco Conti and Luciano Riboldi who carried out data collection in Italy; Paul Maurice Conway for back translation of the Italian questionnaire; Tuuli Sirk, who helped with data collection in Estonia; the Deputy for Training and Research, Shahroud University of Medical Sciences for financial support of data collection in Iran; Asad Ali Khan for supervision of data collection and checking in Pakistan; Khalil Qureshi for training of field workers and supervision of data collection and checking in Pakistan; and Masami Hirai, Tatsuya Isomura, Norimasa Kikuchi, Akiko Ishizuka and Takayuki Sawada for their help with data collection and management in Japan; Monash University which funded data collection in Australia through its grant schemes; NHMRC which supported Helen L Kelsall and Donna M Urquhart in Australia through fellowships; the Ministry of Higher Education in Malaysia which supported Victor C W Hoe in Australia; and the Health Research Council of New Zealand which funded data collection in New Zealand. Data collection in Central America and Colombia was supported by the Southwest Center for Occupational and Environmental Health at the University of Texas Health Science Center research training grant from the NIH Fogarty International Center

Sergio Vargas-Prada was supported by the program Rio-Hortega, Institute of Health Carlos III (ISCIII), Spain.

Eduardo J. Salazar Vega is now employed by AkzoNobel, USA but his contribution to this study was completed before moving to this company. AkzoNobel played no part in this study and have no financial interest in it.

We are particularly grateful to the Colt Foundation, which funded data collection in Brazil, Ecuador, Costa Rica, Nicaragua, UK, Greece, Estonia, Lebanon, Pakistan and South Africa; all of the organisations that allowed us to approach their employees; and all of the workers who kindly participated in the study.

Author Contributions

Conceived and designed the experiments: DC SVP. Performed the experiments: SVP DC GN KTP VEF RH LHB SAF DG AC MB ES EM RRH FS MMK SSPW KM BN MRS HH KC

LMMS MHM FH RF NH MVM LAQ MR ECH CS JMM GD FGB MC MMF ACP LC PB MK KO TF AS RJP-J NS ARW NY HLK VCWH DMU SD DM PH AG EJSV. Analyzed the data: GN DC KWB KTP. Contributed reagents/materials/analysis tools: N/A. Wrote the paper: SVP DC. Provided feedback on the initial draft manuscript and agreed the final changes: GN KTP KWB VEF RH LHB SAF DG AC MB ES EM RRH FS MMK SSPW KM BN MRS HH KC LMMS MHM FH RF NH MVM LAQ MR ECH CS JMM GD FGB MC MMF ACP LC PB MK KO TF AS RJP-J NS ARW NY HLK VCWH DMU SD DM PH AG EJSV.

References

1. Vargas-Prada S, Coggon D. Psychological and psychosocial determinants of musculoskeletal pain and associated disability. Best Pract Res Clin Rheumatol 2015; 29:374–390 doi: 10.1016/j.berh.2015.03.003 PMID: 26612236
2. Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, et al. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. JAMA Intern Med 2014; 174:399–407. doi: 10.1001/jamainternmed.2013.12179 PMID: 24276929
3. Main CJ. The Modified Somatic Perception Questionnaire (MSPQ). J Psychosom Res 1983; 27:503–514. PMID: 6229628
4. Derogatis LR, Melisoratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983; 13:595–605. PMID: 6622612
5. Macfarlane GJ. Hunt IM, Silman AJ. Role of mechanical and psychosocial factors in the onset of forearm pain: prospective population based study. Br Med J 2000; 321:676–679.
6. Palmer KT, Reading I, Calnan M, Linaker C, Coggon D. Does knee pain in the community behave like a regional pain syndrome? Prospective cohort study of incidence and persistence. Ann Rheum Dis 2007; 66:1190–119. PMID: 17114191
7. Palmer KT, Reading I, Linaker C, Calnan M, Coggon D. Population-based cohort study of incident and persistent arm pain: role of mental health, self-rated health and health beliefs. Pain 2008; 136:30–37. PMID: 17689865
8. Vargas-Prada S, Martínez JM, Coggon D, Delclos G, Benavides FG, Serra C. Health beliefs, low mood, and somatising tendency: contribution to incidence and persistence of musculoskeletal pain with and without reported disability. Scand J Work Environ Health 2013; 39:589–598. doi: 10.5271/sjweh.3377 PMID: 23955508
9. Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH, et al. Patterns of multi site pain and associations with risk factors. Pain 2013; 154:1769–1777. doi: 10.1016/j.pain.2013.05.039 PMID: 23727463
10. Croft P, Rigby AS, Boswell R, Schollum J, Silman A. The prevalence of chronic widespread pain in the general population. J Rheumatol 1993; 20:710–713. PMID: 8496870
11. Hunt IM, Silman A, Benjamin S, McBeth J, Macfarlane GJ. The prevalence and associated features of chronic widespread pain in the community using the 'Manchester' definition of chronic widespread pain. Rheumatology 1999; 38:275–279. PMID: 10325667
12. McBeth J, Macfarlane GJ, Hunt IM, Silman AJ. Risk factors for persistent chronic widespread pain: a community-based study. Rheumatology 2001; 40:95–101. PMID: 11157148
13. Palmer KT, Calnan M, Wainwright D, Poole J, O'Neill C, Winterbottom A, et al. Disabling musculoskeletal pain and its relation to somatization: a community-based postal survey. Occup Med 2005; 55:612–617.
14. Solidaki E, Chatzi L, Bitsios P, Coggon D, Palmer KT, Kogevinas M. Risk factors for new onset and persistence of multi-site musculoskeletal pain in a longitudinal study of workers in Crete. Occup Environ Med 2013; 70:29–34. doi: 10.1136/oemed-2012-100689 PMID: 22864252
15. Solidaki E, Chatzi L, Bitsios P, Markatzi I, Plana E, Castro F, et al. Work related and psychological determinants of multi-site musculoskeletal pain. Scand J Work Environ Health 2010; 36:54–61. PMID: 20011982
16. Carugno M, Pesatori AC, Ferrario MM, Ferrari AL, da Silva FJ, Martins AC, et al. Physical and psychosocial risk factors for musculoskeletal disorders in Brazilian and Italian nurses. Cadernos de Saúde Pública. *Rio de Janeiro*. 2012; 28:1632–1642. PMID: 23033179
17. Coggon D, Ntani G, Vargas-Prada S, Martinez JM, Serra C, Benavides FG, et al. International variation in musculoskeletal sickness absence: Findings from the CUPID study. Occup Environ Med 2013; 70:575–584. doi: 10.1136/oemed-2012-101316 PMID: 23695413

18. Palmer KT, Calnan M, Wainwright D, O'Neill C, Winterbottom A, Watkins C, et al. Upper limb pain in primary care: health beliefs, somatic distress, consulting and patient satisfaction. Family Practice 2006; 23:609–617. PMID: 17035285

19. Gupta A, Silman A, Ray D, Morriss R, Dickens C, Macfarlane GJ, et al. The role of psychosocial factors in predicting the onset of chronic widespread pain: results from a prospective population-based study. Rheumatology 2007; 46:666–671. PMID: 17085772

20. McBeth J, Harkness EF, Silman AJ, Macfarlane GJ. The role of workplace low-level mechanical trauma, posture and environment in the onset of chronic widespread pain. Rheumatology 2003; 42:1486–1494. PMID: 12867586

21. Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH, et al. The CUPID (Cultural and Psychosocial Influences on Disability) Study: Methods of Data Collection and Characteristics of Study Sample. PLoS One 2012; 7:e39820. doi: 10.1371/journal.pone.0039820 PMID: 22792189

22. Boulet J, Boss M. Reliability and validity of the Brief Symptom Inventory. J Consult Clin Psychol 1991; 3:433–437.

23. Gureje O, Simon GE, Ustun TB, Goldberg DP. Somatization in cross-cultural perspective: a World Health Organization study in primary care. Am J Psychiatry 1997; 154:989–995. PMID: 9210751

24. Rief W, Rojas G. Stability of somatoform symptoms–implications for classification. Psychosom Med. 2007; 69:864–869. PMID: 18040096

25. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. J Gen Intern Med 2001; 16:266–275. PMID: 11318929

26. Kroenke K, Spitzer RL. Gender differences in the reporting of physical and somatoform symptoms. Psychosom Med 1998; 60:150–155. PMID: 9560862

27. Wool CA, Barsky AJ. Do women somatize more than men? Gender differences in somatization. Psychosomatics 1994; 35:445–452. PMID: 7972659

28. Poikolainen K, Aalto-Setälä T, Marttunen M, Tuulio-Henriksson A, Lönnqvist J. Predictors of somatic symptoms: a five year follow up of adolescents. Arch Dis Child 2000; 83:388–392. PMID: 11040143

29. Craig TK, Cox AD, Klein K. Intergenerational transmission of somatization behaviour: a study of chronic somatizers and their children. Psychol Med 2002; 32:805–816 PMID: 12171375

Journal of Orthopaedic Science 21 (2016) 414–418

Contents lists available at ScienceDirect

Journal of Orthopaedic Science

journal homepage: http://www.elsevier.com/locate/jos

Original article

A population approach to analyze the effectiveness of a back extension exercise "One Stretch" in patients with low back pain: A replication study

Juichi Tonosu [a, *], Ko Matsudaira [b], Hiroyuki Oka [b], Hiroshi Okazaki [a], Takuya Oshio [c], Izumi Hanaoka [d], Yutaka Muraoka [d], Masahiro Midorikawa [c], Kikuo Wakabayashi [e], Sakae Tanaka [f]

[a] Department of Orthopedics, Kanto Rosai Hospital, Kizukisumiyoshi-cho 1-1, Nakahara-ku, Kawasaki City, Kanagawa 211-8510, Japan
[b] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan
[c] Careport Mimaki Special Nursing Home for the Aged, Japan
[d] Yodakubo Special Nursing Home for the Aged, Japan
[e] Bellport Maruko Special Nursing Home for the Aged, Japan
[f] Department of Orthopedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

ARTICLE INFO

Article history:
Received 5 September 2015
Received in revised form
25 February 2016
Accepted 10 March 2016
Available online 2 April 2016

ABSTRACT

Background: We examined the effectiveness of an intervention using a standing back extension exercise called "One Stretch", based on the McKenzie method, in improving or preventing low back pain and disability in Japanese care workers.
Methods: We conducted a non-randomized controlled trial in Japan. Care workers in the intervention group received an exercise manual and a 30-minute seminar on low back pain and were encouraged to exercise in groups, while care workers in a control group were given only the manual. All care workers answered questionnaires at baseline and after one year on the subjective improvement in low back pain, whether they had had a medical consultation for low back pain, and the exercise compliance. Low back pain with disability was assessed by the Oswestry Disability Index.
Results: Participants included 89 workers in the intervention group and 78 in the control group. Background characteristics did not differ significantly between the two groups. Compared to the control group, a greater number of care workers in the intervention group showed improvements in low back pain or prevented it, did not have a medical consultation for low back pain, and exercised regularly. Furthermore, significantly fewer care workers in the intervention group suffered from low back pain with disability by the end of the study period than in the control group.
Conclusion: The population approach about the exercise "One Stretch" led to better compliance with the exercise, and was effective for improving or preventing low back pain and in decreasing the likelihood of having a medical consultation for low back pain.

1. Introduction

Low back pain (LBP) is a major health problem, particularly in industrialized countries, and has affected people's lives and livelihood in various ways. To reduce the socioeconomic impact of LBP, it is important to prevent people without LBP symptoms from developing them. Physical exercises have been recommended as one means of LBP prevention, although there is as yet insufficient evidence for advocating any specific type or intensity of exercise [1].

McKenzie, who introduced a method of classifying LBP by subgroups, recommends extension exercises because the posterior displacement of the nucleus via such exercises can relieve LBP [2]. The McKenzie method involves classifying patients into specific subgroups primarily based on their symptomatic and mechanical responses to mechanical loadings. Most individuals with LBP can benefit in a short period from the back extension loading strategy, the theoretical explanation of which is based on the disc mode. This model proposes that posterior displacement of the nucleus can be

* Corresponding author. Tel.: +81 44 411 3131; fax: +81 44 433 3150.
E-mail address: juichitohnosu@yahoo.co.jp (J. Tonosu).

http://dx.doi.org/10.1016/j.jos.2016.03.002
0949-2658/

reduced via deliberate extension loading. Reducing how much the nucleus is displaced may result in decreased LBP. A meta-analysis regarding the McKenzie method concluded that this method may be effective for acute LBP patients; however, the magnitude of benefit may not be considered clinically worthwhile, and insufficient evidence exists for chronic LBP patients. Furthermore, the effectiveness of the classification-based McKenzie therapy has not yet been estimated [2].

The severity of LBP is positively correlated with disability, as shown by Von Korff et al. [3], and the present goal of LBP management is to become or remain free from LBP with disability. In Japan, in addition to an inadequate number of care workers and poor working environments for these, an increasing number of care workers are suffering from LBP with disability, which is a serious problem.

Therefore, we conducted an intervention trial utilizing a simple daily back extension exercise for LBP based on the McKenzie method called "One Stretch" for relieving LBP; however, LBP in the current study was not necessarily assumed to indicate discogenic LBP. Furthermore, there is insufficient evidence for the McKenzie method in treating chronic LBP patients, as noted above. Previously, this exercise was evaluated in a trial with 166 care workers [4], and the results indicated that the intervention group showed a greater improvement in objectively assessed LBP than did the control group. However, this study had the limitation in terms of the generalizability of the effects. Specifically, the intervention was delivered by a single individual, who was also one of the researchers involved in developing the intervention. Thus, a replication of this intervention in a different study center and conducted by someone outside the original development team is necessary. If similar effects were found, it would establish that this intervention of the population approach and the exercise "One Stretch" itself were responsible for the improvements, rather than the person delivering it, and that the intervention is generalizable across settings and populations.

2. Subjects and methods

2.1. Study population

This study was conducted at three health care facilities for the elderly in Nagano Prefecture, Japan. Eligible participants were Japanese care workers who worked at those facilities and supported the elderly in need of care. We excluded workers who would have had difficulty participating due to medical (e.g. spinal stenosis, rheumatoid arthritis, and ankylosing spondylitis) or other personal reasons. Furthermore, we excluded participants who did not complete the questionnaires at end of the 1-year study period. Written informed consent was obtained from all participants.

This study was approved by the medical/ethics review board of Kanto Rosai Hospital. We registered our study (ID: UMIN000006688) in the University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR).

2.2. Study design

This was a non-randomized controlled trial. Participants who worked at one health care facility were assigned to the control group and those at the other two facilities were assigned to the intervention group. The health care facility of the control group and those associated with the intervention group did not exhibit any differences with respect to the number of care workers, the age and sex ratio of care workers, the number of beds, the care need score of patients, the ratio of disabled patients, and the ratio of patients with dementia. All participants received an exercise manual, but only the intervention group received a 30-min seminar. The exercise manual described how to do the standing back extension exercise One Stretch (shown in Fig. 1). This exercise is an active extension of the back that is commonly used in physical therapy and is based on the exercises used for treating derangement syndrome, one of the subgroups classified by the McKenzie method [2]. The manual also included evidence-based information for treatment and prevention of LBP, including self-management and risk factors (e.g., psychosocial factors and fear-avoidance). The 30-min seminar was delivered by an orthopedist, and merely comprised a detailed explanation of the exercise manual and the One Stretch exercise.

Participants were asked to do this exercise regularly. To promote regular exercise in the intervention group, we encouraged them to perform it in groups. Specifically, the care workers in the intervention group performed the exercise in a group at the daily meeting in the health care facility. This approach spanned the entire study period.

2.3. Data collection

At baseline and end of the 1-year study period, data were collected using a self-administrated questionnaire. The baseline questionnaire assessed the following: age, sex, body mass index (BMI), smoking habit, whether they had a medical consultation for LBP (yes or no) at baseline, and the severity of LBP in the previous month. The severity of LBP was evaluated using Von Korff's grading system: 1) no pain, 2) LBP that does not interfere with work, 3) LBP that interferes with work, and 4) LBP that interferes with work such that it leads to sick leave [3]. We defined pain localized between the costal margin and the inferior gluteal folds as LBP [5]; to ensure that participants understood what we meant by LBP, we provided a diagram of LBP in the questionnaire. We also determined if participants had LBP with disability using the Oswestry Disability Index (ODI) [6]. In this analysis, we set the ODI cut-off value as 12 as per previous findings on the topic [7].

The questionnaire at the end of the study period assessed subjective improvement in LBP from baseline (improved, no change, or worsened), whether they had a medical consultation for LBP (yes or no) at the end of the study period, and overall compliance with the exercise during the study period (good or poor). Compliance was evaluated using self-reported answers. Participants who performed the given exercise at least once during their working day were defined as having "good compliance". Participants were asked to record whether they performed the exercise each day to evaluate overall compliance with the exercise during the study period.

In order to evaluate the effectiveness of the population approach, we compared results of the intervention group and the control group. Further, in order to evaluate the effectiveness of the "One Stretch" exercise itself, we compared the post-intervention subjective improvement in LBP against exercise compliance in both groups.

2.4. Statistical analysis

Descriptive statistics were determined and presented as means and standard deviations (SDs) or frequencies and percentages. Between-group differences in baseline characteristics were evaluated by using the χ^2 test or the Cochran–Armitage test for categorical variables and Student's t-test for continuous variables. Whether they had a medical consultation for LBP and the compliance with the exercise were evaluated by using the χ^2 test, and subjective improvement in LBP was evaluated by using the Cochran–Armitage test. The association between the intervention and the subjective improvement in LBP from baseline was evaluated using the Cochran–Mantel–Haenszel test with stratification according to levels of exercise compliance. All statistical tests were two-tailed and conducted with a significance level of 0.05.

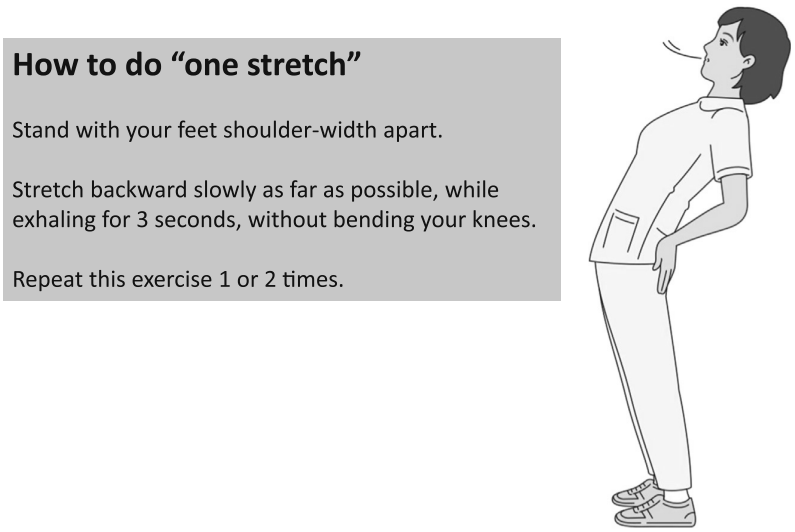


Fig. 1. Example of standing back extension “One Stretch”.

3. Results

A total of 167 care workers participated in this study and were assigned to either the intervention (n = 89) or the control group (n = 78). The mean age of the intervention group was 37.5 ± 12.4 years; 16.9% were men and 83.1% were women. The control group mean age was 37.6 ± 11.6 years, and 20.5% were men and 79.5% were women. The baseline characteristics of the groups are shown in Table 1. There were no significant differences in any characteristic, including the severity of LBP, and ODI between the two groups.

The post-intervention evaluations for the groups are shown in Table 2. Compared with the control group, the intervention group had a higher proportion of care workers who showed improvements in LBP, a lower proportion who had a medical consultation for LBP at the end of study period, and better compliance with the exercise. These differences were statistically significant ($P = 0.0001$, $P = 0.007$, and $P < 0.0001$, respectively). The proportion of participants in the intervention group with an ODI of greater than 12 at the end of the study period was less than that in the control group ($P = 0.04$); there were no significant differences between the groups at baseline.

The subjective improvement in LBP from baseline, compared with the pre-intervention severity of LBP, is shown in Table 3. Those from the intervention group having LBP that did not interfere with work showed a higher proportion of improvement in LBP, compared with those in the control group. The subjective improvement in LBP from baseline, according to exercise compliance, is shown in Table 4. Participants in the intervention group

Table 2
Post-intervention evaluations for intervention and control groups.

	Intervention (n = 89)	Control (n = 78)	*P*
Subjective improvement in LBP			*0.0001
Improved	27 (30.3)	6 (7.7)	
No change	25 (28.1)	44 (56.4)	
Worsened	5 (5.6)	9 (11.5)	
Medical consultation (yes)			
At baseline	9 (10.1)	16 (20.5)	0.06
At post-intervention	3 (3.4)	12 (15.4)	*0.007
Exercise compliance			*<0.0001
Good	47 (52.8)	13 (16.6)	
Poor	23 (25.8)	54 (69.2)	
ODI ≥ 12			
At baseline	30 (33.7)	32 (41.0)	0.33
At post-intervention	26 (29.2)	37 (47.4)	*0.04

Data are shown as number of participants (%). *: $P < 0.05$.
ODI, Oswestry Disability Index; LBP, low back pain.

Table 1
Baseline characteristics of the intervention and control groups.

	Intervention (n = 89)	Control (n = 78)	*P*
Age	37.5 ± 12.4	37.6 ± 11.6	0.97
Women	74 (83.1)	62 (79.5)	0.56
BMI	22.5 ± 3.8	22.3 ± 3.5	0.74
Smoking habit	39 (43.8)	37 (47.4)	0.28
Medical consultation (yes)	9 (10.1)	16 (20.5)	0.06
Severity of LBP in the previous 1 month			0.37
No pain	26 (29.2)	24 (30.8)	
LBP not interfering with work	54 (60.7)	40 (51.3)	
LBP interfering with work	4 (4.5)	11 (14.1)	
LBP interfering with work such that it leads to sick leave	1 (1.1)	1 (1.3)	
ODI	9.8 ± 1.0	11.5 ± 1.1	0.28
ODI ≥ 12	30 (33.7)	32 (41.0)	0.33

Data are shown as mean ± SD or number of participants (%).
ODI, Oswestry Disability Index; LBP, low back pain.

Table 3
Subjective improvement in LBP from baseline contrasted against the pre-intervention severity of LBP in the intervention and control groups.

Pre-intervention severity of LBP		n	Subjective improvement		
			Improved	No change	Worsened
No pain	Intervention	13	6 (46.2)	6 (46.2)	1 (7.7)
	Control	15	0 (0)	11 (73.3)	4 (26.7)
LBP not interfering with work	Intervention	40	20 (50.0)	17 (42.5)	3 (7.5)
	Control	33	3 (9.1)	27 (81.8)	3 (9.1)
LBP interfering with work such that it may or may not require sick leave	Intervention	3	0 (0)	2 (66.7)	1 (33.3)
	Control	9	3 (33.3)	4 (44.4)	2 (22.2)

Data are shown as number of participants (%).

Table 4
Subjective improvement in LBP from baseline contrasted against exercise compliance in the intervention and control groups.

Compliance		n	Improved	No change	Worsened
Good	Intervention	38	22 (57.9)	12 (31.6)	4 (10.5)
	Control	10	3 (30.0)	7 (70.0)	0 (0)
Poor	Intervention	17	4 (23.5)	12 (70.6)	1 (7.5)
	Control	48	3 (6.3)	37 (77.1)	8 (16.7)

Data are shown as number of participants (%).

that had good compliance showed a tendency towards greater subjective improvement in LBP, although this tendency did not reach statistical significance ($P = 0.054$).

4. Discussions

There were no significant differences in baseline characteristics between the intervention and control groups. The number of participants was almost equivalent to the previous study investigating this exercise [4], although the proportion of women was higher in the current study than in the previous study. The proportion of participants who had had a medical consultation for LBP at baseline, in both the intervention and control groups, was about ten times or more that of the Japanese general population, which was about 0.5% [8]. Thus, we might infer that care workers are disproportionately likely to have a medical consultation for LBP. The ODI at baseline in both groups was higher than the normative score of the ODI (8.73) [7], indicating that care workers were more likely to have LBP with disability. Specifically, the proportions of participants with an ODI $\geq$ 12 at baseline in intervention and control groups were 33.7% and 41.0%, respectively. However, for both groups, the proportion of participants who reported a subjective severity of LBP as "interferes with work" was lower than the proportion of participants with an ODI $\geq$ 12. Thus, it is possible that some participants who answered that their LBP severity was "does not interfere with work" might actually have LBP with disability.

In the subjective evaluations of the improvement, the intervention group had a higher proportion of care workers who had "improved" LBP, especially in those having LBP that did not interfere with work, compared to the control group. Furthermore, we noticed that subjective evaluations of "no change" included care workers without LBP. Six care workers who did not report LBP prior to intervention expressed an improvement in LBP post-intervention. In this case, it is possible that these 6 individuals forgot their initial response of "no pain." Conversely, it is also possible that they were not suffering from LBP at the time of the post-intervention questionnaire. Therefore, we regarded the proportion of participants responding with "improved" or "no pain" following an original response of "no pain" as individuals in whom the exercises prevented the development of LBP. These findings suggest that the population approach of the standing back extension exercise "One Stretch" is effective for improving and preventing LBP. This result is similar to what the previously trial on this exercise [4]. The intervention group had better compliance with the exercise than did the control group, which suggests that a population approach can encourage better compliance with the exercise. Generally, a population health approach is considered a powerful preventive strategy affecting causal behavior in health care activities [9]. This matches the present result regarding prevented LBP. It has been reported that some individuals need an individual approach in order to adopt preventive behaviors and that both population and individual approaches must complement each other [10]; however, individual approaches were not used in the current study.

A lower proportion of the intervention group had a medical consultation for LBP at the end of the study period. This finding is perhaps due not only to the subjective improvements in LBP but also to the fact that the intervention group participants had acquired a technique and conception for self-management of their LBP. Therefore, the intervention can be said to decrease the treatment costs for LBP, which is beneficial for both individuals and society.

Good exercise compliance was associated with a greater degree of improvement in LBP. This suggests that the "One Stretch" exercise is effective for improving and preventing LBP. That is also similar to the previous study on the One Stretch exercise [4]. Several other studies have supported the effectiveness of extension exercises for LBP. Long et al. [11] found that patients randomized to favorable directional preference exercises, consisting mostly of extension exercises, showed significant improvements in LBP compared to those randomized to opposite or mid-range movements. In a randomized controlled trial in which military conscripts performed either extension in lying exercises or a control group, the intervention group showed a significantly lower prevalence of LBP and care seeking for LBP compared to the control group [12]. Additionally, the extension approach inhibited developing back problems in the young male conscripts. This is similar to our study, even if there were differences in age, sex, and an exact posture of extensions. One of a possible mechanism for the clinical improvements seen in the present study was highlighted in a previous study using kinematic magnetic resonance imaging. Specifically, slightly degenerated intervertebral discs moved in a posterior direction during flexion and in an anterior direction during extension [13]. However, results of this study have shown that the effectiveness of the "One Stretch" exercise for LBP was not limited to discogenic LBP.

The proportion of participants with an ODI $\geq$ 12 in the intervention group at the end of the study period was lower than that in the control group, whereas there was no significant difference between the groups at baseline. This result indicates that LBP with disability can be reduced by the intervention. At present, LBP is prevalent in all adult populations, but only a few subjects become disabled. However, these patients are responsible for most of the

treatment costs for LBP [14]. As noted before, the severity of LBP correlates with disability [3] and the present goal of LBP management is to become or remain free from LBP with disability; as such, efforts to prevent LBP with disability instead of LBP in general are likely to be more effective and efficient. The intervention in the current study could be effective in terms of the aforementioned purpose. As such, the objective evaluation by the ODI is perhaps underestimated compared to the subjective improvement.

There were several limitations to this study. First, the questionnaire contained retrospective questions wherein participants assessed their LBP condition after one year; thus, the possibility for recall bias must be considered. Second, the sample was comparatively small. Due to the nature of the study, clustered randomized trials with adequate sample sizes are needed for evaluating the intervention. Thus, the generalizability of findings is limited and the findings should be interpreted with caution. We will perform further investigations of this topic in large-scale randomized controlled trials.

5. Conclusion

The population approach about the exercise “One Stretch” led to better compliance with the exercise, and was effective for improving or preventing LBP and in decreasing the likelihood of having a medical consultation for LBP. It is likely that the population approach about daily practice of this simple exercise and the exercise itself can benefit our society, especially in industrial health.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This study was supported by the dissemination project on the 13 fields of occupational injuries and illness of the Japan Labor Health and Welfare Organization. We would like to thank Yoshihiro Iwasada (PT, MS, Dip.MDT, the McKenzie Institute Japan) for his valuable advice.

References

[1] Burton AK, Balagué F, Cardon G, Eriksen HR, Henrotin Y, Lahad A, Leclerc A, Müller G, van der Beek AJ. Chapter 2. European guidelines for prevention in low back pain: November 2004. Eur Spine J 2006 Mar;15(Suppl. 2):S136–68.
[2] Machado LA, de Souza Mv, Ferreira PH, Ferreira ML. The McKenzie method for low back pain: a systematic review of the literature with a meta-analysis approach. Spine (Phila Pa 1976) 2006 Apr 20;31(9):E254–62.
[3] Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992 Aug;50(2):133–49.
[4] Matsudaira K, Hiroe M, Kikkawa M, Sawada T, Suzuki M, Isomura T, Oka H, Hiroe K, Hiroe K. Can standing back extension exercise improve or prevent low back pain in Japanese care workers? J Man Manip Ther 2015 Sep;23(4):205–9.
[5] Krismer M, van Tulder M, Low Back Pain Group of the B, Joint Health Strategies for Europe Project. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best Pract Res Clin Rheumatol 2007 Feb;21(1):77–91.
[6] Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980 Aug;66(8):271–3.
[7] Tonosu J, Takeshita K, Hara N, Matsudaira K, Kato S, Masuda K, Chikuda H. The normative score and the cut-off value of the Oswestry Disability Index (ODI). Eur Spine J 2012 Aug;21(8):1596–602.
[8] Ministry of Health. Labor and Welfare [Internet]. Comprehensive survey of living conditions. 2010 [cited 2015 Dec. 12]. Available from: http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/3-2.html.
[9] Rose G. Sick individuals and sick populations. Int J Epidemiol 1985 Mar;14(1):32–8.
[10] Doyle YG, Furey A, Flowers J. Sick individuals and sick populations: 20 years later. J Epidemiol Community Health 2006 May;60(5):396–8.
[11] Long A, Donelson R, Fung T. Does it matter which exercise? A randomized control trial of exercise for low back pain. Spine (Phila Pa 1976) 2004 Dec 1;29(23):2593–602.
[12] Larsen K, Weidick F, Leboeuf-Yde C. Can passive prone extensions of the back prevent back problems? A randomized, controlled intervention trial of 314 military conscripts. Spine (Phila Pa 1976) 2002 Dec 15;27(24):2747–52.
[13] Zou J, Yang H, Miyazaki M, Morishita Y, Wei F, McGovern S, Wang JC. Dynamic bulging of intervertebral discs in the degenerative lumbar spine. Spine (Phila Pa 1976) 2009 Nov 1;34(23):2545–50.
[14] Snook SH. Work-related low back pain: secondary intervention. J Electromyogr Kinesiol 2004 Feb;14(1):153–60.

PLOS | ONE

RESEARCH ARTICLE

Psychometric Properties of the Japanese Version of the STarT Back Tool in Patients with Low Back Pain

Ko Matsudaira[1]*, Hiroyuki Oka[1], Norimasa Kikuchi[2,3], Yuri Haga[2], Takayuki Sawada[2,3], Sakae Tanaka[4]

1 Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan, **2** Clinical Study Support, Inc., Nagoya, Aichi, Japan, **3** Department of Public Health, Aichi Medical University School of Medicine, Nagakute, Aichi, Japan, **4** Department of Orthopaedic Surgery, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

* kohart801@gmail.com

Citation: Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S (2016) Psychometric Properties of the Japanese Version of the STarT Back Tool in Patients with Low Back Pain. PLoS ONE 11(3): e0152019. doi:10.1371/journal.pone.0152019

Editor: Masahiko Sumitani, The University of Tokyo Hospital, JAPAN

Received: October 23, 2015

Accepted: March 8, 2016

Published: March 22, 2016

Data Availability Statement: All relevant data are within the paper.

Funding: The authors received no specific funding for this work. Clinical Study Support, Inc. provided support in the form of salaries for authors NK, YH and TS, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing Interests: The authors have the following interests. NK is a board member of Clinical

Abstract

Background and Objective

The STarT Back Tool uses prognostic indicators to classify patients with low back pain into three risk groups to guide early secondary prevention in primary care. The present study aimed to evaluate the psychometric properties of the Japanese version of the tool (STarT-J).

Methods

An online survey was conducted among Japanese patients with low back pain aged 20–64 years. Reliability was assessed by examining the internal consistency of the overall and psychosocial subscales using Cronbach's alpha coefficients. Spearman's correlation coefficients were used to evaluate the concurrent validity between the STarT-J total score/psychosocial subscore and standard reference questionnaires. Discriminant validity was evaluated by calculating the area under the curves (AUCs) for the total and psychosocial subscale scores against standard reference cases. Known-groups validity was assessed by examining the relationship between low back pain-related disability and STarT-J scores.

Results

The analysis included data for 2000 Japanese patients with low back pain; the mean (standard deviation [SD]) age was 47.7 (9.3) years, and 54.1% were male. The mean (SD) STarT-J score was 2.2 (2.1). The Cronbach's alpha coefficient was 0.75 for the overall scale and 0.66 for the psychosocial subscale. Spearman's correlation coefficients ranged from 0.30 to 0.59, demonstrating moderate to strong concurrent validity. The AUCs for the total score ranged from 0.65 to 0.83, mostly demonstrating acceptable discriminative ability. For known-groups validity, participants with more somatic symptoms had higher total scores. Those in higher STarT-J risk groups had experienced more low back pain-related absences.

Study Support, Inc. and co-authors YH and TS are employed by Clinical Study Support, Inc. HO has received grants to his institution from Pfizer, Inc. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

Conclusions

The overall STarT-J scale was internally consistent and had acceptable concurrent, discriminant, and known-groups validity. The STarT-J can be used with Japanese patients with low back pain.

Introduction

Low back pain (LBP) is a major musculoskeletal problem in the general population from childhood to older adulthood, affecting more than 632 million people worldwide [1]. The 2010 Global Burden of Diseases, Injuries, and Risk Factors Study reported that LBP was the leading cause of disability among 291 diseases and injuries globally, and LBP ranked as the highest global cause of years lived with disability [2]. This highlights the high prevalence of LBP worldwide, and may also reflect the difficulty of successful LBP management. In primary care, approximately 85% of patients with LBP have no specific underlying causes or pathology [3]. Patients with non-specific LBP often experience recurrent pain, and the majority of these patients suffer from chronic pain [4–5]. Recurrent and chronic LBP may result in a serious social and economic burden.

Psychological factors have been widely acknowledged as contributors to the chronicity of LBP [6–7]. These factors include pain catastrophizing, fear-avoidance beliefs, and psychological distress. A number of previous reports suggested an association between psychological factors and poor long-term outcomes [5, 8–9]. In primary care, cognitive behavioral therapy focused on psychological factors is a dominant treatment approach for people with LBP [5]. To provide efficient, targeted care, it is becoming common to stratify patients with LBP according to their risk for poor long-term outcomes [10]. Significant clinical benefits and cost-effectiveness of stratified care compared with non-stratified physiotherapy practice have been demonstrated in a randomized clinical trial [11].

The STarT Back Tool (STarT) has been widely used to stratify patients with LBP according to risk for chronicity (Fig 1). The STarT was originally developed as a screening tool for prognostic indicators of back pain to help primary care clinical decision-making in the UK [12]. The STarT consists of 9 items. Items 1–4 evaluate physical factors, and items 5–9 assess psychosocial factors. The STarT classifies patients into three risk groups: patients with a total score of 0–3 are classified as low-risk; patients with a total score of $\geq$ 4 but a psychosocial subscore of $\leq$ 3 as medium-risk; and patients with a psychosocial subscore of $\geq$ 4 are classified as high-risk [12] (Fig 2). Targeted treatments have been developed for patients in each risk group: a minimal intervention by general practitioners or physiotherapists for the low-risk group, physiotherapy to address pain and disability for the medium-risk group, and psychologically-informed physiotherapy to address pain and disability as well as psychosocial obstacles to recovery for the high-risk group [11, 13, 14].

Although the STarT has been translated into various languages, no validated Japanese version was available. In our previous study, we translated the original English version of the STarT into Japanese (STarT-J) and linguistically validated it [15]. As a next step, we conducted online surveys with Japanese people with LBP to evaluate the psychometric properties of the STarT-J. The present analysis aimed to evaluate the reliability and validity of the STarT-J in a large number of Japanese people with LBP, using cross-sectional data from these surveys.

The Keele STarT Back Screening Tool

Patient name: _____ Date: _____

Thinking about the **last 2 weeks** tick your response to the following questions:

		Disagree 0	Agree 1
1	My back pain has **spread down my leg(s)** at some time in the last 2 weeks	☐	☐
2	I have had pain in the **shoulder** or **neck** at some time in the last 2 weeks	☐	☐
3	I have only **walked short distances** because of my back pain	☐	☐
4	In the last 2 weeks, I have **dressed more slowly** than usual because of back pain	☐	☐
5	It's not really safe for a person with a condition like mine to be physically active	☐	☐
6	**Worrying thoughts** have been going through my mind a lot of the time	☐	☐
7	I feel that **my back pain is terrible** and **it's never going to get any better**	☐	☐
8	In general I have **not enjoyed** all the things I used to enjoy	☐	☐

9. Overall, how **bothersome** has your back pain been in the **last 2 weeks**?

Not at all	Slightly	Moderately	Very much	Extremely
☐	☐	☐	☐	☐
0	0	0	1	1

Total score (all 9): _____ Sub Score (Q5-9):_____

Funded by Arthritis Research UK

Fig 1. STarT Back Tool. Response options for items 1–8 are "disagree" (0 points) or "agree" (1 point). Responses to item 9 are on a scale of 1–5: "not at all," "slightly," "moderately," "very much," or "extremely." The first three options ("not at all," "slightly," and "moderately") are scored as 0, and the remaining two options ("very much" and "extremely") are scored as 1. Items 1–4 constitute the physical subscale. Items 5–9 constitute the psychosocial subscale.

doi:10.1371/journal.pone.0152019.g001

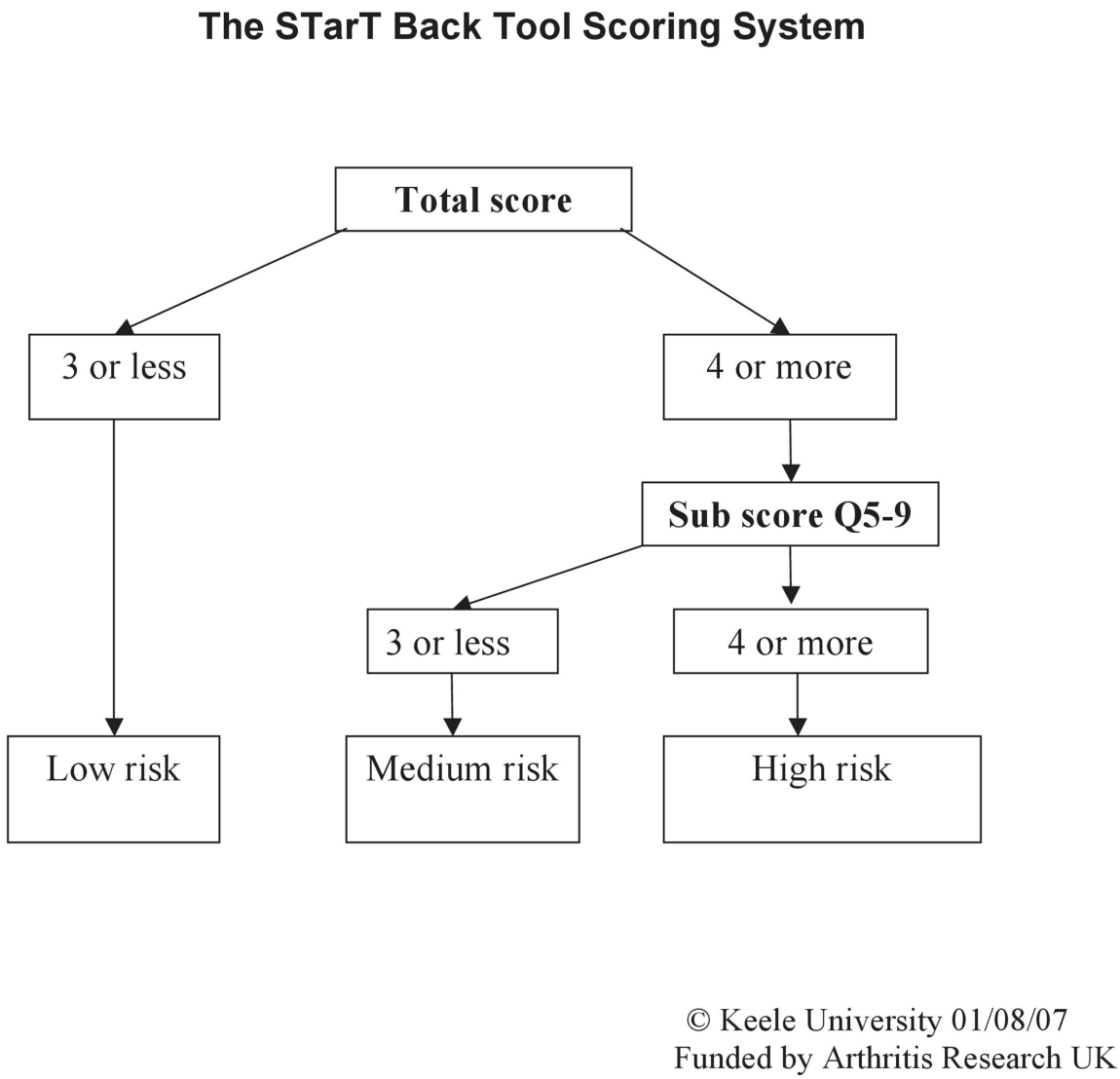


Fig 2. STarT Back Tool risk stratification. Sub score Q5–9: psychosocial subscale.

doi:10.1371/journal.pone.0152019.g002

Materials and Methods

Study population

To assess the psychometric properties of the STarT-J, we conducted online surveys collecting information on LBP in the Japanese population in January and February, 2014. Participants were recruited from an online panel provided by an Internet research company, UNITED, Inc. (Tokyo, Japan), which included approximately 1.25 million individuals aged 20–64 years registered as research volunteers. From these volunteers, 965,919 individuals were randomly selected and invited by e-mail to complete an online questionnaire on health problems associated with pain (first survey). We obtained 52,842 responses by the end of January 2014. Of these initial respondents, those who had LBP in the last 4 weeks were invited to complete

another online questionnaire (secondary survey). LBP was defined as pain in the lower back experienced in the last 4 weeks that lasted for more than 1 day, according to the standard definition of LBP proposed by Dionne et al. [16]. Pain associated with menstruation or pregnancy and pain during a feverish illness were excluded. A diagram showing the lower back area (between the inferior costal margin and gluteal folds) was provided in the questionnaire. The secondary survey closed on 7 February 2014, when the total number of responses reached 2000. The mean (standard deviation [SD]) age of respondents in the secondary survey was 47.7 (9.3) years and 54.1% were male. We conducted two subsequent surveys, 6 and 24 weeks after the secondary survey, to follow up respondents and investigate their LBP condition. In the present analysis, we analyzed secondary survey data to evaluate the psychometric properties of the STarT-J.

We obtained approval from the Medical/Ethics Review Board of the Japan Labour Health and Welfare Organization, Kanto Rosai Hospital (Approval number: 2012–22). Participation was voluntary, and no personal information was collected. Although no written informed consent was obtained, submitting a completed questionnaire was considered as evidence of consent. Potential participants first read an explanation of the aim of the survey and only those who agreed to participate could proceed to the questionnaire. As an incentive, participants received reward points for online shopping from the Internet research company.

Development of the linguistically-validated STarT-J

In our previous study [15], the STarT was translated into Japanese and linguistically validated in a general cross-cultural adaptation process [17–19]. This process occurred in three steps: (1) forward-translation (English to Japanese), (2) back-translation (Japanese to English), and (3) cognitive debriefing. In the third step, we conducted a pilot study to assess if the questions and response scales were understandable and correctly interpreted by Japanese patients. After considering their feedback, and consultation with a specialist as necessary, we published the STarT-J [15].

Measures

We included a number of measures in the online questionnaires.

Pain. The degree of pain associated with LBP during the last 4 weeks was assessed by a numerical rating scale (NRS), ranging from 0 (no pain at all) to 10 (the worst pain imaginable).

Disability caused by LBP. We used the Roland—Morris Disability Questionnaire (RDQ) to assess the LBP-related disability participants experienced in their daily lives. The RDQ comprises 24 Yes/No questions. The total score ranges from 0 to 24, with a higher score indicating greater disability. In this study, we used the Japanese version of the RDQ, for which the reliability and validity have been previously confirmed [20].

Fear-avoidance beliefs. Fear of pain can lead to avoidance of physical activity, an important indicator of a poor long-term LBP prognosis. The Fear-Avoidance Belief Questionnaire (FABQ), consisting of physical activity and work subscales, is widely used to assess fear-avoidance beliefs [21]. We used the FABQ physical activity subscale (FABQ-PA). The FABQ-PA score ranges from 0 to 30; a higher score indicates a stronger fear-avoidance belief. We also used the Tampa Scale of Kinesiophobia (TSK) [22–23], originally developed to measure the fear of movement or injury. The total TSK score sums the scores of 17 items (each rated on a scale of 1–4), and ranges from 17 to 68. A higher score indicates a higher level of kinesiophobia.

Catastrophizing. Pain catastrophizing is also an important indicator of poor LBP prognosis. Catastrophizing was assessed using the Pain Catastrophizing Scale (PCS), originally

developed to measure negative attitudes toward pain involving rumination, helplessness, and magnification. The PCS consists of 13 items. The total score ranges from 0 (no catastrophizing) to 52 (greater catastrophizing). We used the Japanese version of the PCS, for which the reliability and validity have been previously confirmed [24].

Depression and anxiety. A 14-item self-assessment scale, the Hospital Anxiety and Depression Scale (HADS), was used to measure anxiety and depression. The HADS comprises anxiety and depression subscales, each with seven items. The total score ranges from 0 to 21, with a higher score indicating more mental distress. The validity and reliability of the Japanese version of the HADS have been previously confirmed [25].

General health status. The EuroQol 5 Dimension (EQ-5D) [26] is an instrument that provides a simple, descriptive profile and single index value for general health status. The index score is derived from conversion of all responses, and ranges from −0.11 to 1.00. A score of 1 means "perfect health" and a score of 0 denotes "death."

Somatic symptoms. Somatization was assessed using the 7-item somatization subscale from the Brief Symptom Inventory (BSI) [27]. Seven symptoms (faintness or dizziness, pains in the heart or chest, nausea or upset stomach, trouble getting your breath, numbness or tingling in parts of the body, feeling weak in parts of the body, hot or cold spells) are rated on a 5-point scale: "not at all," "a little bit," "moderately," "quite a bit," and "extremely." We used the linguistically validated Japanese version of the BSI-somatization subscale [28].

Data analyses

Participants' demographic and clinical characteristics were summarized using descriptive statistics. To examine floor and ceiling effects, percentages of respondents with total scores of 0 and 9 were calculated. Floor and ceiling effects were considered to exist when more than 15% of respondents had the lowest or highest possible score [29]. To examine the reliability of the STarT-J, we evaluated internal consistency by calculating Cronbach's alpha coefficients for the overall scale and the psychosocial subscale. An alpha index more than 0.70 is considered to indicate satisfactory internal consistency [30].

Concurrent validity was evaluated by measuring correlations between the previously described reference instruments and the STarT-J total score and psychosocial subscore using Spearman's correlation coefficients. Correlation coefficients were evaluated according to the criteria for correlation strength in psychometric validation proposed by Cohen: 0.10 representing a weak, 0.30 a moderate, and 0.50 a strong correlation [31].

To assess discriminant validity, we calculated the area under the curves (AUCs) for the total scores and psychosocial subscores against the reference standards. We defined cases using the following cut-off values: a RDQ score of ≥ 7 for disability, a PCS score of ≥ 20 for catastrophizing, a TSK score of ≥ 41 for fear-avoidance beliefs, and a HADS score of ≥ 8 for depression and anxiety. In addition, a single question was used to determine the presence of referred leg pain within the last 4 weeks. Discriminative ability was interpreted according to the same criteria as used in the original STarT study: 0.70 to < 0.80 indicating acceptable discrimination, 0.80 to < 0.90 indicating excellent discrimination, and ≥ 0.90 indicating outstanding discrimination [12].

For known-groups validity, to test whether the STarT-J scores differentiated participants with known differences, we examined 1) total scores among the groups with a different number of somatic symptoms, and 2) the number of absences due to LBP among the three risk groups (low, medium, and high) using the Jonckheere—Terpstra test. If participants responded "moderately," "quite a bit," or "extremely" to a BSI item, they were considered to have that somatic symptom. Participants were then categorized into three groups according to the number of

somatic symptoms: no symptoms, one symptom, and two or more symptoms. With respect to the number of absences, days on which participants could not perform housework were counted, as well as absences from work. It was hypothesized that participants with more somatic symptoms would have higher total scores, and that participants in the high-risk group would have experienced more LBP-related absences.

All statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA). The level of significance was set at 0.05.

Results

Participant characteristics

The present analysis included data for 2000 Japanese patients with LBP. Table 1 presents a summary of participants' demographic and clinical characteristics. The mean (SD) age was 47.7 (9.3) years; 54.1% of participants were male. More than half (53.7%) of the participants had experienced LBP for more than 1 year. Most participants (92%) experienced recurrent LBP, and more than half (52.9%) reported having LBP 10 times or more.

Scores of the measures

The mean (SD) score for the STarT-J was 2.2 (2.1). No remarkable ceiling effect was observed as 0.9% of participants had the highest score of 9. However, a floor effect was observed as 23.4% of participants had the lowest score of 0. The score distribution for each item is shown in Table 2. Participants were classified into three risk groups according to their STarT-J score: 1557 (77.9%) into the low-risk group, 294 (14.7%) into the medium-risk group, and 149 (7.5%) into the high-risk group.

Reliability

The Cronbach's alpha coefficients were 0.75 for the overall scale and 0.66 for the psychosocial subscale.

Concurrent validity

To examine concurrent validity, Spearman's correlation coefficients were used to measure correlations between the STarT-J total score/psychosocial subscore and the pain NRS, RDQ, FABQ-PA, TSK, PCS, HADS, and the EQ-5D (Table 3). The correlation coefficients for the total score ranged from 0.30 (HADS depression) to 0.59 (RDQ), demonstrating a moderate to strong correlation with these reference standards. Similarly, correlation coefficients for the psychosocial subscore ranged from 0.33 (FABQ-PA) to 0.54 (RDQ), demonstrating a moderate to strong correlation. Both the total score and psychosocial subscore were strongly negatively correlated with the EQ-5D ($\gamma = -0.56$ and $\gamma = -0.53$, $p < 0.0001$). In terms of the correlation with psychosocial measures, the total score was moderately correlated with the TSK ($\gamma = 0.49$), whereas the psychosocial subscore was strongly correlated ($\gamma = 0.53$). Moderate correlation coefficients were observed for both the total score and psychosocial subscore with the PCS ($\gamma = 0.46$ and $\gamma = 0.49$) and the HADS ($\gamma = 0.40$ and $\gamma = 0.45$) ($p < 0.0001$ for all).

Discriminant validity

To assess discriminant validity, AUCs were calculated for the total score and psychosocial subscore against the cases defined by the reference standards (Table 4). The AUCs for the total score were all above 0.70, indicating acceptable to excellent discriminative ability, with the exception of depression and anxiety (0.65). For the psychosocial subscore, the AUCs ranged

Table 1. Participant characteristics: psychometric testing of the STarT-J (n = 2000).

Characteristics	n (%)	Mean (SD)
Sex		
Male	1081 (54.1)	
Female	919 (46.0)	
Age (years)		47.7 (9.3)
BMI $\geq$ 25 (kg/m^2)	506 (25.3)	
Duration of low back pain		
< 2 weeks	350 (17.5)	
$\geq$ 2 weeks, < 1 month	188 (9.4)	
$\geq$ 1, < 3 months	184 (9.2)	
$\geq$ 3, < 6 months	90 (4.5)	
$\geq$ 6 months, < 1 year	115 (5.8)	
$\geq$ 1, < 3 years	200 (10.0)	
$\geq$ 3 years	873 (43.7)	
Number of recurrence		
1	160 (8.0)	
2	135 (6.8)	
3–4	340 (17.0)	
5–9	308 (15.4)	
$\geq$10	1057 (52.9)	
STarT-J score		2.2 (2.1)
RDQ score		4.2 (4.7)
FABQ-PA score		12.9 (4.7)
TSK score		41.0 (6.5)
PCS total score		21.6 (10.0)
PCS rumination		10.6 (4.3)
PCS helplessness		6.2 (4.2)
PCS magnification		4.7 (2.7)
HADS total score		17.2 (6.7)
HADS anxiety		8.7 (3.4)
HADS depression		8.5 (4.1)
EQ-5D index score		0.78 (0.16)
NRS for low back pain		4.2 (1.8)

Values are n (%), or mean (SD).

STarT-J, the Japanese version of the STarT Back Tool; BMI, body mass index; RDQ, Roland—Morris Disability Questionnaire; FABQ-PA, Fear-Avoidance Belief Questionnaire Physical Activity Subscale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale; HADS, Hospital Anxiety and Depression Scale; EQ-5D, EuroQol 5 Dimension; NRS, numerical rating scale.

doi:10.1371/journal.pone.0152019.t001

from 0.67 (depression and anxiety) to 0.79 (disability), indicating poor to acceptable discriminative ability.

Known-groups validity

We examined the STarT-J total scores and risk groups among participants with known-differences. As hypothesized, participants with more somatic symptoms had higher total scores. The mean (SD) score of participants with no somatic symptoms was 1.71 (1.76), one somatic symptom was 2.73 (2.14), and two or more somatic symptoms was 3.76 (2.50) (Fig 3). A linear

Table 2. Score distribution of STarT-J items and risk group distribution (n = 2000).

Item	Number of participants who answered "agree" (1 point) n (%)
1	442 (22.1)
2	1069 (53.5)
3	317 (15.9)
4	264 (13.2)
5	574 (28.7)
6	652 (32.6)
7	425 (21.3)
8	351 (17.6)
9	239 (12.0)
Risk group distribution	
Low-risk	1557 (77.9)
Medium-risk	294 (14.7)
High-risk	149 (7.5)

Values are n (%).

STarT-J: The Japanese version of the STarT Back Tool. For item 9, answers of "very much" and "extremely" were scored as 1 point, and were counted as "agree"; the answers "not at all," "slightly," and "moderately" were scored as 0 points, and were not included.

doi:10.1371/journal.pone.0152019.t002

increasing trend in total score across groups with an increasing number of somatic symptoms was observed (Jonckheere-Terpstra test, $p < 0.0001$). With respect to the associations between risk groups and the number of absences, participants in the high-risk group reported a larger number of absences (Fig 4). The mean (SD) LBP-related absences in the low-risk group was 4.0 (5.4) days, 6.6 (8.3) days in the medium-risk group, and 12.6 (11.1) days in the high-risk group. A linear increasing trend in the number of absences across the risk groups was observed (Jonckheere-Terpstra test, $p < 0.0001$).

Table 3. Spearman's correlation coefficients for the STarT-J and related measures.

Measures	Total score Coefficients (95% CI)	Psychosocial subscore Coefficients (95% CI)
RDQ	0.59 (0.56–0.62)	0.54 (0.51–0.57)
FABQ-PA	0.34 (0.30–0.37)	0.33 (0.29–0.37)
TSK	0.49 (0.45–0.52)	0.53 (0.50–0.56)
PCS total	0.46 (0.42–0.49)	0.49 (0.46–0.52)
PCS rumination	0.43 (0.40–0.47)	0.44 (0.41–0.48)
PCS helplessness	0.39 (0.35–0.43)	0.43 (0.39–0.46)
PCS magnification	0.40 (0.36–0.44)	0.43 (0.39–0.47)
HADS total	0.40 (0.36–0.44)	0.45 (0.41–0.48)
HADS anxiety	0.42 (0.38–0.46)	0.46 (0.42–0.49)
HADS depression	0.30 (0.26–0.34)	0.35 (0.31–0.39)
EQ-5D	–0.56 (–0.59 –-0.52)	–0.53 (–0.56 –-0.50)
NRS for low back pain	0.42 (0.38–0.46)	0.39 (0.35–0.42)

Note: $p < 0.0001$ for all correlation coefficients. STarT-J, the Japanese version of the STarT Back Tool; CI, confidence interval; RDQ, Roland—Morris Disability Questionnaire; FABQ-PA, Fear-Avoidance Belief Questionnaire Physical Activity Subscale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale; HADS, Hospital Anxiety and Depression Scale; EQ-5D, EuroQol 5 Dimension; NRS, numerical rating scale.

doi:10.1371/journal.pone.0152019.t003

Table 4. AUCs for STarT-J total score and psychosocial subscore against reference standards.

Reference standards	Case definition	Total score AUC (95% CI)	Psychosocial subscore AUC (95% CI)
Disability	RDQ score $\geq$ 7	0.83 (0.81–0.85)	0.79 (0.77–0.82)
Referred leg pain	Yes	0.76 (0.73–0.79)	0.68 (0.65–0.72)
Fear-avoidance belief	PCS score $\geq$ 20	0.71 (0.69–0.73)	0.72 (0.70–0.74)
Catastrophizing	TSK score $\geq$ 41	0.74 (0.72–0.76)	0.75 (0.73–0.77)
Depression and anxiety	HADS score $\geq$ 8	0.65 (0.63–0.68)	0.67 (0.65–0.69)

AUC, area under the curve; STarT-J, the Japanese version of the STarT Back Tool; CI, confidence interval; RDQ, Roland—Morris Disability Questionnaire; PCS, Pain Catastrophizing Scale; TSK, Tampa Scale for Kinesiophobia; HADS, Hospital Anxiety and Depression Scale.

doi:10.1371/journal.pone.0152019.t004

Discussion

In this analysis, we evaluated the psychometric properties of the STarT-J. In summary, the overall scale of the STarT-J was internally consistent, and the STarT-J had acceptable concurrent validity, discriminant validity, and known-groups validity in Japanese patients with LBP.

The Cronbach's alpha coefficient for the overall scale (0.75) demonstrated sufficient internal consistency, and was similar to the original and other language versions: 0.79 for the original [12], 0.74 for the French [32], 0.74 for the Brazilian Portuguese [33], 0.82 for the Iranian [34], and 0.83 for the Persian [35] versions. Although these results could not be compared directly because the study methods varied, the similar values support that the overall scale of the STarT-J is internally consistent and no items are redundant. The Cronbach's alpha coefficient for the psychosocial subscale was 0.66, below the value of 0.70 considered necessary to claim

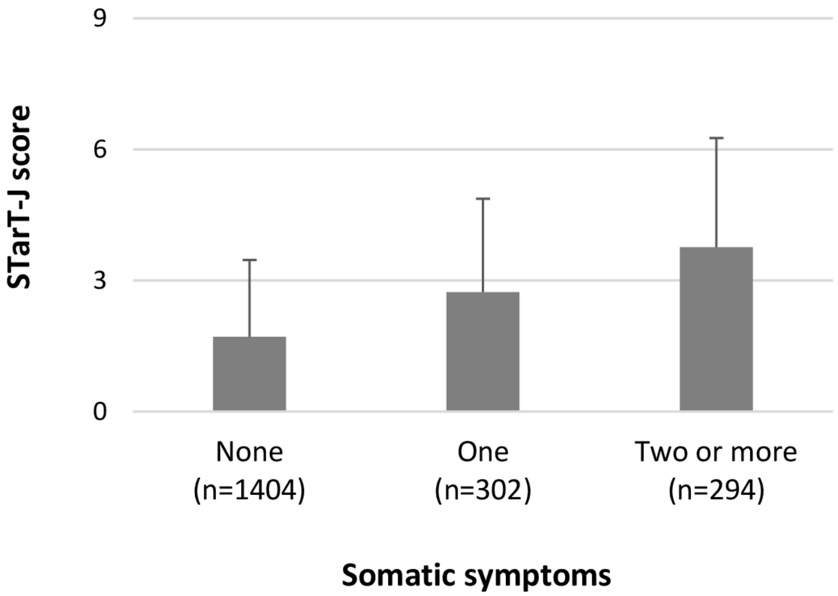


Fig 3. Mean STarT-J scores for participants with different numbers of somatic symptoms. The linear trend was tested using the Jonckheere-Terpstra test ($p < 0.0001$). STarT-J: The Japanese version of the STarT Back Tool. Number of somatic symptoms was assessed by the Brief Symptom Inventory somatization scale: a response of "moderately," "quite a bit," or "extremely" to an item was interpreted as the presence of that somatic symptom, and thus counted.

doi:10.1371/journal.pone.0152019.g003

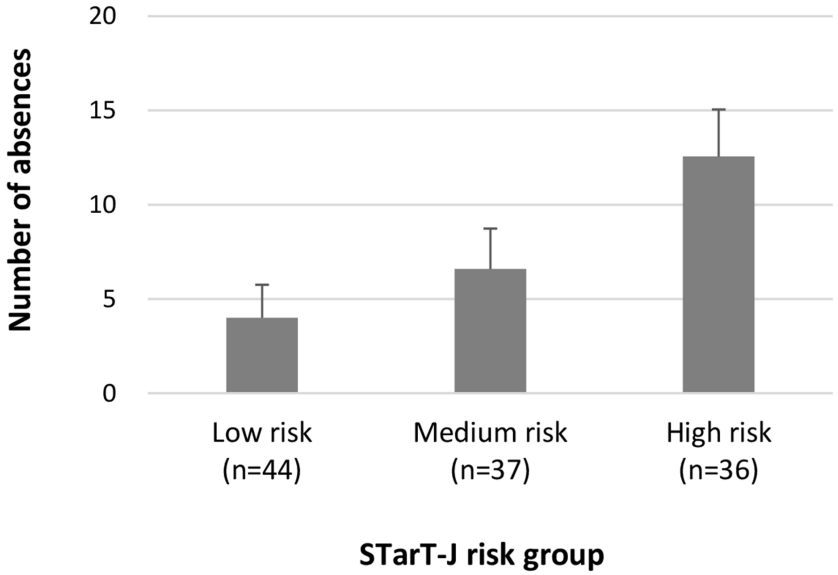


Fig 4. Mean number of absences for the three STarT-J risk groups. The linear trend was tested using the Jonckheere-Terpstra test ($p < 0.0001$). STarT-J: The Japanese version of the STarT Back Tool.

doi:10.1371/journal.pone.0152019.g004

the subscale is internally consistent. However, it should be taken into consideration that the coefficient for the subscale was also lower than for the overall scale in the original version, although it was still 0.74 [12].

To assess concurrent validity, we analyzed the correlations between the STarT-J and reference standards (the pain NRS, RDQ, FABQ-PA, TSK, PCS, HADS, and EQ-5D). Overall, the Spearman's correlation coefficients indicated that both the total score and the psychosocial subscore were moderately to strongly correlated with these existing scales. In particular, the STarT-J total score was strongly correlated with the RDQ ($\gamma = 0.59$). Similar results were observed in the German ($\gamma = 0.55$) [36], French ($\gamma = 0.74$) [32], and Persian ($\gamma = 0.811$) [35] versions. Although a direct comparison cannot easily be made, these similar results reinforce the concurrent validity of the STarT-J.

Discriminant validity was assessed by calculating the AUCs for the total score and the psychosocial subscore. For the total score, the AUCs for disability and referred leg pain were both higher than the AUCs for fear-avoidance beliefs, catastrophizing, and depression and anxiety. This demonstrated that the total score better discriminated cases defined by physical reference standards. However, for the psychosocial subscore, the AUCs for fear-avoidance beliefs, catastrophizing, and depression and anxiety were not remarkably higher than AUCs for the physical reference cases. These AUCs for the psychosocial reference cases were similar to those for the total score, indicating the psychosocial subscale might discriminate cases defined by the psychosocial reference standards at a similar level to the overall scale. A similar trend was observed in the original STarT [12], although overall, the AUCs were higher compared with the STarT-J.

To assess known-groups validity, we investigated relationships between total scores and the number of somatic symptoms, and between risk groups and the number of absences. Participants with more somatic symptoms had higher total scores, and those in the high-risk group had experienced greater LBP-related disability. This demonstrated that the STarT-J can differentiate patients with different levels of LBP-related problems.

The present study has some limitations. First, we did not examine the test—retest reliability. The intra-class, test—retest reliability over specific time intervals should therefore be evaluated in a future study. Second, the analysis might have included patients not targeted by the STarT, that is, patients who had specific causes of LBP. The diagnostic triage for LBP is to classify LBP into one of three categories: LBP with specific pathologic change ("red flag"), LBP with sciatica/radicular syndrome, or non-specific LBP [37]. According to this classification, six of the participants in the present analysis were probable "red flags," 308 had radicular syndrome, and the remaining 1686 participants were considered to have non-specific LBP. As the original study included patients with non-specific LBP who had referred leg pain [12], the STarT is considered applicable to patients with LBP potentially associated with sciatica/radicular syndrome. Therefore, assuming diagnoses were accurate, most participants probably fit into the STarT target group. However, it should be noted that these diagnoses might not necessarily be accurate as they were based on participants' self-report. Third, the study population might not be consistent with the primary care population. Our study included more low-risk participants and less high-risk participants compared with the original study [12]. This might be because we recruited from a general Japanese population registered with an online panel rather than from patients in hospitals. Our study population would therefore represent the general Japanese population with LBP. As the observed floor effect suggests, more patients might have LBP that was not sufficiently severe to require hospital care. Although our study population was broader than the primary care population, the percentage of patients with non-specific LBP was similar to that observed in primary care settings. In our study, 1686 participants (84.3%) probably had non-specific LBP. In primary care, approximately 85% of patients with LBP have non-specific LBP [3]. Therefore, our study population resembled the primary care population in terms of the distribution of non-specific LBP. Fourth, as this was a cross-sectional study, it did not assess the ability of the STarT-J to predict chronicity of LBP. To assess its predictive ability, longitudinal studies will be necessary to investigate associations between risk groups and long-term outcomes of patients with LBP.

In the present analysis, we evaluated the psychometric properties of the STarT-J to enable Japanese clinicians to use the scale in the early stages of LBP. The STarT is a simple and quick tool, and is suitable for use in primary care settings. Stratified care is a dominant approach in the management of LBP [10]. Stratified care based on the STarT risk groups has been shown to be clinically and economically beneficial for patients with LBP [11, 38]. Therefore, we expect that the STarT-J may facilitate early stratified care in primary care settings in Japan. This may alleviate the physical, social, and economical burden of LBP in the Japanese population.

In conclusion, acceptable internal consistency for the overall STarT-J scale demonstrated the reliability of the STarT-J in Japanese patients with LBP; acceptable concurrent validity, discriminant validity, and known-groups validity demonstrated the validity. In a subsequent analysis, the ability of the STarT-J to predict chronicity of LBP will be examined using longitudinal data, to validate its clinical use in Japanese patients.

Acknowledgments

We thank all those who participated in our online surveys.

Author Contributions

Conceived and designed the experiments: KM HO ST. Performed the experiments: KM HO. Analyzed the data: KM HO NK YH TS ST. Contributed reagents/materials/analysis tools: KM HO. Wrote the paper: KM HO NK YH TS ST.

References

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380: 2224–60. doi: 10.1016/S0140-6736(12)61766-8 PMID: 23245609
2. Buchbinder R, Blyth FM, March LM, Brooks P, Woolf AD, Hoy DG. Placing the global burden of low back pain in context. Best Pract Res Clin Rheumatol. 2013; 27: 575–89. doi: 10.1016/j.berh.2013.10.007 PMID: 24315140
3. Deyo RA, Weinstein JN. Low back pain. N Engl J Med. 2001; 344: 363–70. PMID: 11172169
4. Axen I, Leboeuf-Yde C. Trajectories of low back pain. Best Pract Res Clin Rheumatol. 2013; 27: 601–12. doi: 10.1016/j.berh.2013.10.004 PMID: 24315142
5. Dunn KM, Hestbaek L, Cassidy JD. Low back pain across the life course. Best Pract Res Clin Rheumatol. 2013; 27: 591–600. doi: 10.1016/j.berh.2013.09.007 PMID: 24315141
6. Hasenbring M, Hallner D, Klasen B. Psychological mechanisms in the transition from acute to chronic pain: over- or underrated? Schmerz. 2001; 15: 442–7. German. PMID: 11793149
7. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine (Phila Pa 1976). 2002; 27: E109–20.
8. Hill JC, Fritz JM. Psychosocial influences on low back pain, disability, and response to treatment. Phys Ther. 2011; 91: 712–21. doi: 10.2522/ptj.20100280 PMID: 21451093
9. Pincus T, McCracken LM. Psychological factors and treatment opportunities in low back pain. Best Pract Res Clin Rheumatol. 2013; 27: 625–35. doi: 10.1016/j.berh.2013.09.010 PMID: 24315144
10. Foster NE, Hill JC, O'Sullivan P, Hancock M. Stratified models of care. Best Pract Res Clin Rheumatol. 2013; 27: 649–61. doi: 10.1016/j.berh.2013.10.005 PMID: 24315146
11. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet. 2011; 378: 1560–71. doi: 10.1016/S0140-6736(11)60937-9 PMID: 21963002
12. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum. 2008; 59: 632–41. doi: 10.1002/art.23563 PMID: 18438893
13. Main CJ, Sowden G, Hill JC, Watson PJ, Hay EM. Integrating physical and psychological approaches to treatment in low back pain: the development and content of the STarT Back trial's 'high-risk' intervention (StarT Back; ISRCTN 37113406). Physiotherapy. 2012; 98: 110–6. doi: 10.1016/j.physio.2011.03.003 PMID: 22507360
14. Sowden G, Hill JC, Konstantinou K, Khanna M, Main CJ, Salmon P, et al.; IMPaCT Back study team. Targeted treatment in primary care for low back pain: the treatment system and clinical training programmes used in the IMPaCT Back study (ISRCTN 55174281). Fam Pract. 2012; 29: 50–62. doi: 10.1093/fampra/cmr037 PMID: 21708984
15. Matsudaira K, Kikuchi N, Kawaguchi M, Inuzuka K, Arisaka M, Hara N, et al. Development of a Japanese version of the STarT (Subgrouping for Targeted Treatment) Back screening tool: translation and linguistic validation. Journal of Musculoskeletal Pain Research. 2013; 5: 11–19. Japanese.
16. Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976). 2008; 33: 95–103.
17. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993; 46: 1417–32. PMID: 8263569
18. Suzukamo Y, Kumano H. Psychometrics. In: Ikegami N, Fukuhara S, Shimozuma K, Ikeda S, editors. QOL evaluation handbook for clinical diagnosis. Tokyo: Igaku Shoin; 2001. p. 8–13. Japanese.
19. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health. 2005; 8: 94–104. PMID: 15804318
20. Suzukamo Y, Fukuhara S, Kikuchi S, Konno S, Roland M, Iwamoto Y, et al. Validation of the Japanese version of the Roland-Morris Disability Questionnaire. J Orthop Sci. 2003; 8: 543–8. PMID: 12898308
21. Matsudaira K, Kikuchi N, Murakami A, Isomura T. Psychometric properties of the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ). J Orthop Sci. 2014; 19: 26–32. doi: 10.1007/s00776-013-0471-5 PMID: 24091984

22. Matsudaira K, Inuzuka K, Kikuchi N, Sakae C, Arisaka M, Isomura T. Development of a Japanese version of the Tampa Scale for Kinesiophobia (TSK-J): translation and linguistic validation. Seikei Geka (Orthopedic surgery). 2013; 48: 13–9. Japanese.

23. Kikuchi N, Matsudaira K, Sawada T, Oka H. Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury pain and/or low back pain. J Orthop Sci. Epub 2015 Jul 23.

24. Matsuoka H, Sakano Y. Assessment of cognitive aspect of pain: development, reliability, and validation of Japanese version of pain catastrophizing scale. Japanese Journal of Psychosomatic Medicine. 2007; 47: 95–102. Japanese.

25. Higashi A, Yashiro H, Kiyota K, Inokuchi H, Hatta H, Fujita K, et al. Validation of the hospital anxiety and depression scale in a gastro-intestinal clinic. The Japanese journal of gastro-enterology. 1996; 93: 884–92. Japanese. PMID: 8986079

26. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy. 1990; 16: 199–208. PMID: 10109801

27. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med. 1983; 13: 595–605. PMID: 6622612

28. Matsudaira K, Inuzuka K, Kikuchi N, Sakae C, Arisaka M, Isomura T. Development of the Japanese version of the brief symptom inventory-somatization scale: translation and linguistic validation. Seikei Geka (Orthopedic surgery). 2012; 63: 149–53. Japanese.

29. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007; 60: 34–42. PMID: 17161752

30. Nunnally JC. Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.

31. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

32. Bruyere O, Demoulin M, Beaudart C, Hill JC, Maquet D, Genevay S, et al. Validity and reliability of the French version of the STarT Back screening tool for patients with low back pain. Spine (Phila Pa 1976). 2014; 39: E123–8.

33. Pilz B, Vasconcelos RA, Marcondes FB, Lodovichi SS, Mello W, Grossi DB. The Brazilian version of STarT Back Screening Tool—translation, cross-cultural adaptation and reliability. Braz J Phys Ther. 2014; 18: 453–61. PMID: 25372008

34. Azimi P, Shahzadi S, Azhari S, Montazeri A. A validation study of the Iranian version of STarT Back Screening Tool (SBST) in lumbar central canal stenosis patients. J Orthop Sci. 2014; 19: 213–7. doi: 10.1007/s00776-013-0506-y PMID: 24343300

35. Abedi M, Manshadi FD, Khalkhali M, Mousavi SJ, Baghban AA, Montazeri A, et al. Translation and validation of the Persian version of the STarT Back Screening Tool in patients with nonspecific low back pain. Man Ther. Epub 2015 Apr 15.

36. Aebischer B, Hill JC, Hilfiker R, Karstens S. German Translation and Cross-Cultural Adaptation of the STarT Back Screening Tool. PLOS ONE. 2015; 10: e0132068. doi: 10.1371/journal.pone.0132068 PMID: 26161669

37. Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. Spine (Phila Pa 1976). 2001; 26: 2504–13; discussion 2513–4.

38. Foster NE, Mullis R, Hill JC, Lewis M, Whitehurst DG, Doyle C, et al. Effect of stratified care for low back pain in family practice (IMPaCT Back): a prospective population-based sequential comparison. Ann Fam Med. 2014; 12: 102–11. doi: 10.1370/afm.1625 PMID: 24615305

Original

The Current State along with Outstanding Issues Related to Email-Based Guidance by Physical Therapists Aiming to Prevent Low Back Pain among Workers

Takuo Nomura[1], Fuminari Asada[2], Kenichiro Takano[3] and Ko Matsudaira[4]

[1]Department of Rehabilitation Sciences, Faculty of Allied Health Sciences, Kansai University of Welfare Sciences

[2]Research Center for the Health Promotion and Employment Support, Osaka Rosai Hospital

[3]Research Center for the Health Promotion and Employment Support, Kansai Rosai Hospital

[4]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo Hospital

(Received: July 2, 2015)

Abstract

Low back pain (LBP) is more likely than any other symptom to prevent people from working, making the establishment of measures to prevent and reduce LBP vital in the workplace. In our previous study, we have conducted Physical Consultant research (PCo research) in order to verify the effect of email-based guidance provided by currently working physical therapists (the advisors) to workers (the clients) in preventing LBP. We found a significant improvement of Work Ability Index and Fear-Avoidance Beliefs Questionnaire among the clients in the PCo research. The purpose of this study was to consider means of improving the effectiveness of email-based guidance using computers and mobile terminals. The method involved the use of a PCo research database, with the results of the questionnaire survey carried out among advisors and clients analyzed once the study was completed. The results revealed that one advisor can receive and respond to questions from five clients at any one time. Questions from the clients included not only those related to LBP prevention, but also some related to other musculoskeletal symptoms, general lifestyle, and care for family members. We thought that the physical therapists are the profession most likely to be able to respond to such questions. The development of a dedicated system will be required in order to manage data from multiple clients and implement efficient and effective email-based guidance.

(JJOMT, 64: 113—118, 2016)

—Key words—
Physical therapist, Low back pain, Occupational health

Introduction

In Japan, from April 2008, medical insurers have been required to execute specific health examinations and offer specific health guidance (Specified Health Examinations and Guidance) focusing on visceral fat accumulation[1]. The subjects from 40 years old to 74 years old are categorized based on the results of examinations and questionnaires, according to their level of risk, into those requiring "motivational support" and those requiring "proactive support", with Specified Health Guidance provided depending on these categorizations. "Proactive support" is provided among individual in one format from interview, telephone, letters and electronic communication (fax and email, etc.), or as a combination of these.

In the field of research into physical activities, measures offering support for physical activity have also been provided by the use of mobile phone messaging[2]. Furthermore, trials using emails have begun in walking programs in occupational health[3]. We have already noted the short-term effects of guidance and support in im-

proving physical activity when provided by mobile telephone message to healthy young people, and provided by an electronic communication including mobile telephone message to healthy elderly people[4,5]. Yamatsu, et al. reported that the short-term success rate of intervention in physical activity using the internet and mobile phone terminals is at least 50%[6]. However, the long-term benefits of intervention in physical activity using the internet and mobile terminal devices are unclear, and the study of its long-term impact, including the development of a methodology for effective intervention, is an issue for the future. While the long-term effects of intervention are still an issue for future consideration, it is considered to be an "effective tool for promoting changes in peoples' behavior"-which is also part of the current Specified Health Guidance.

In Japan, it has been found that 60% of workplace-related illnesses are due to low back pain (LBP)[7], of which 85% is in the form of Non-Specific LBP (NSLBP), with no specific cause[8,9]. It has also recently been clarified that psychological and social factors also contribute to NSLBP and that the provision of correct information and an encouraging attitude among those around the subject can contribute to relief from LBP[10]. In terms of preventing LBP, the effect of guidance using electronic communication was not examined. In our previous study, we conducted a Physical Consultant research (PCo research) in order to verify the effects of email-based guidance provided by physical therapists (the advisors) to workers (the clients) in preventing LBP among workers who had in the past suffered from NSLBP[11].

In the present study, we used a database of PCo research and analyzed the results of a questionnaire conducted among advisors and clients once intervention had ended, in order to consider a more effective means of email-based guidance using computers and mobile terminal devices.

Methods

Definition of PCo research

This research project was conducted between June 2013 and February 2014, with the aim of verifying the effectiveness of email-based guidance aiming to prevent LBP among worker, and from this, propose a new business model for the utilization of physical therapists in the field of occupational health[11]. Fig. 1 shows a flow chart of the PCo research. The project allocated 20 currently working advisors with a minimum of three years' of clinical experience to 20 clients, for the purpose of the advisors providing individual guidance on how to prevent LBP by email. Two of the physical therapists who had experience providing advice regarding the prevention of LBP in worker, compiled a basic concept[12] for the prevention of LBP, with consultations and the content of guidance stored and shared via the cloud computing network at the research secretariat. The email-based guidance intervention was conducted for six months, during which time advisors sent a minimum of eight emails to clients. Moreover, in addition to the regular emails from advisors, the advisors also responded to queries from their clients.

The results of PCo research showed a significant increase in the Work Ability Index (WAI)[13] which is used to express how skillfully each client is engaging in his/her tasks, along with a tendency towards improvement within the Fear-Avoidance Beliefs Questionnaire (FABQ)[14] which expresses the subject's state of mind in regard to avoiding the fear of LBP, subsequent to the email-based guidance. In the present study, the Helsinki Declaration was fully adhered to in the use of the PCo research database and no individuals were identified.

Details of the questionnaire

The questionnaire was conducted after completion of the email-based guidance intervention. Questions asked to advisors included: "1) Did you need personal information about the client? (Y/N)"; "2) Number of days taken for the advisor to respond to a question from the client (No. of days)"; "3) Time spent on forming a response by the advisor to a question from a client (No. of minutes)"; "4) Number of responses by email by client to advisor (No. of responses)"; "5) Ideal amount of time spent by advisor in order to come up with a response to a question from a client (from the advisor's perspective) (No. of minutes)"; "6) Details of questions not related to LBP"; "7) No. of clients it would be possible to respond to in one month (No. of responses)"; and "8) Problems with the PCo research (from the advisor's perspective)". Questions 6) and 8) required a free response.

Questions asked to clients included: "1) Do you think that the advice provided by your advisor in emails was helpful in preventing LBP? (Y/N)"; and "2) Would you like to use email-based guidance again? (Y/N)".

The physical therapists dispatched by the research secretariat implemented a lecture for 90 minutes in regard to the prevention of low back pain for company employees. At this time, an explanation of the Physical Consultant research (PCo research) was provided orally, with those giving consent appointed as subjects (clients).

↓

Clients who gave their consent were allocated to a responsible physical therapist (advisor), in regard to whom, the clients' personal information was not disclosed. Personal information about the advisors was also not provided to the clients.

↓

Outline of Email-based guidance

Period	No. of times*	Advisor	Client
Start	1st time	Advisor sends email to client, providing guidance based on questions submitted by client.	Receives email from advisor, and sends questions to advisor.
Two weeks	2nd time		
First month	3rd time		
Second month	4th time		
Third month	5th time		
Fourth month	6th time		
Fifth month	7th time		
Sixth month	8th time		

↓ ↓

Survey on completion*1
Research secretariat implements questionnaire survey in regard to advisors in order to analyze issues related to the email-based guidance.

Survey on completion*2
Research secretariat implements questionnaire survey in regard to clients, in order to ascertain effectiveness of email-based guidance.

Fig. 1 Physical Consultant Research Flow Chart

Advisors: Physical therapists, Clients: Worker, Research secretariat: the Japanese Study Group of Physical Therapy in Occupational Health.

No. of times*: Basic number of times emails sent to clients by advisors. If the client asks questions, the number of times emails and responses are sent may increase based on the details of the advice sought and provided.

Survey on completion*: Data from both surveys on completion (1 and 2) implemented in regard to advisors and clients was utilized in this study.

Questions 3) and 5) for the advisors were compared using the Mann-Whitney U-test. The statistical software used was IBM SPSS Statistics ver. 19, with statistical significance defined at 5%.

Results

The results of the questionnaire survey conducted in regard to advisors are shown in Table 1. Since questions 6) and 8) required a free response, similar responses were compiled and the record shows responses in the order of popularity. No significant difference was between "Time spent on forming a response by the advisor to a question from a client" and "Ideal amount of time spent by advisor in order to come up with a response to a question from a client" (questions 3) and 5)).

The results of the questionnaire survey conducted in regard to clients revealed that 88% responded "Yes (it was helpful)" to the question "Do you think that the advice provided by your advisor in emails was helpful in preventing LBP?", and 76% responded "Yes (would like to use again)" in response to the question "Would you like to use email-based guidance again?"

Table 1 Results of Questionnaires given to Advisors

	Question	Response
1	Need for personal information about the client	All advisors required personal information from their clients
2	Time taken for the advisor to respond to a question from the client	2.0 ± 0.4 days
3	Time spent on creating a response by the advisor to a question from a client	22.0 ± 9.7 minutes
4	No. of responses by email by client to advisor	4.7 ± 4.2 times
5	Ideal amount of time spent by advisor in order to come up with a response to a question from a client	20.3 ± 13.7 min.
6	Questions related to topics other than low back pain	Stiff or painful shoulders: n = 2 Chills: n = 1 Stiffness upon waking in the morning: n = 1 Care for family members: n = 1 Obesity: n = 1 Lack of exercise: n = 1 Lower extremity pain: n = 1
7	No. of clients it would be possible to respond to in one month (from the advisor's perspective)	5.3 ± 7.4 people
8	Problems with the PCo research (from the advisor's perspective)	Need subject's information in advance: n = 4 Not clear whether or not client has seen email: n = 2 Need guidelines and examples for composing emails: n = 2 Need to decide how to follow-up if client does not respond: n = 2 Not clear if information being provided is what client requires: n = 1 Difficult to tell whether guidance is effective: n = 1 Differences in advice given by different advisors; this may cause dissatisfaction among clients: n = 1 Confused about what to do when clients do not respond: n = 1 Difficult to build trust with client due to anonymity: n = 1 Need for quantitative data and tools that allow for shared awareness between client and advisor: n = 1

Discussion

In the present study, we used a database of information from past PCo research and analyzed the results of a questionnaire conducted among advisors and clients once intervention had ended, in order to consider a more effective means of email-based guidance using computers and mobile terminal devices.

In the PCo research, we took care to ensure that the advisors could not identify the clients and that the clients could not identify the advisors even after the study ended. The results of the questionnaire implemented on advisors showed that all advisors stated they required personal information from clients in order to provide email-based guidance. In the field of physical activity research it has been reported that the greater the relevancy of messages sent to subjects in support of physical activities, the more likely it is that the message will be received and acted upon[15]. Obtaining personal information about the client is considered to facilitate the provision of guidance that is more relevant to factors in the client's background. As such, disclosure of personal information is an issue that requires more consideration. During the PCo research, a unified concept of guidance to prevent LBP[10] was provided; however, if intervention is to be considered in indeterminate multiple cases, even if personal information is not available. It will be necessary to compile a manual for email-based guidance methods to ensure a smooth response. Guidance for prevention of LBP needs to include guidance related to physical activity and exercise, as well as provisions for psychological and social factors such as FAB[16]. When giving individual advice, it is easy to respond to individual psychological and social factors. However, when giving guidance to a group, it is more difficult to focus on individual psychological and social factors, and as such, is an issue that will require consideration in the future.

The results of the questionnaire survey showed that the average ideal amount of time required by advisors in order to come up with a response to a question from a client was 20.3 ± 13.7 minutes, while the time taken to respond was in fact 22.0 ± 9.7 minutes. There was no significant difference noted between the two, in-

dicating that advisors were able to reply within their ideal time limit. When asked how many clients they could handle per month, the average response was diverse at 5.3 ± 7.4; however, it was believed that the group of currently active physical therapists in this study could manage email-based guidance with around 100 clients at any one time. Clients asked for advice on a diverse range of issues other than LBP, including other musculoskeletal symptoms, general lifestyle, and care for family members. It is thought that it would be difficult for people other than physical therapists, who have anatomical, physiological and other types of medical knowledge, as well as experience with rehabilitation, etc., to respond to medical and caregiving problems, to respond to these questions, and as such, there is a great need for physical therapists to act as advisors. The period of time taken for advisors to respond to clients was on average 2.0 ± 0.4 days, although confirmation of whether or not this was an acceptable period of time for clients was not carried out. Furthermore, advisors sent a minimum of eight emails to clients, with an average number of responses from clients of 4.7 ± 4.2, and it is not known whether the details of the emails sent were highly satisfactory or not. In a previous survey of 844 patients currently attending orthopedic clinics for treatment of chronic LBP, patients were asked to rank their "satisfaction with their current improvement in pain level" on a Rickert scale of 1 to 11; wherein, while the average response was 5.3, as many as 27% of respondents stated that they had "given up hope of any further relief" when asked "How much relief do you expect to gain from further treatment?"[17]. In this research, 88% of clients responded in the questionnaire that they felt the email-based guidance had been useful in preventing LBP, while 76% stated that they would like to receive further guidance. It should be noted in comparison that in prior research[17], although subjects were patients attending clinics, it is believed that this still represents a high level of satisfaction among clients with the email-based guidance.

In the future, in addition to considering the effectiveness of email-based guidance in research with regulated methodology, it is necessary to consider how to implement even more effective email-based guidance. The PCo research looked at 20 clients and the research secretariat was able to ascertain the content of emails between advisors and clients; however, when considering a response to a much larger number of clients, along with the development of advisors, it is believed that the development of a dedicated system will be required. Furthermore, cost/benefit analysis of email-based guidance must be carried out in order to acquire results that can demonstrate to corporations and clients the specific benefits of such a system.

Acknowledgement

This study was supported by the dissemination project of clinical research for occupational injuries and illness from Ministry of Health, Labour and Welfare of Japan. Furthermore, this study was published with the support of a research grant from the Japanese Physical Therapy Association. The authors would like to sincerely thank the Japanese Study Group of Physical Therapy in Occupational Health for their cooperation with this study.

References

1) Ministry of Health, Labour and Welfare of Japan: Special health checkups and special health-maintenance guidance. http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/kenkou/seikatsu/ (accessed 2015-5-30)
2) Kubota A: Specific approaches to regional health and welfare; How to increase the number of people with regular exercise habits: Development of program for improving lifestyles using IT (i-exer). JPFSM 54: 21—24, 2005 (in Japanese).
3) Sudo H, Hamasaki K, Harada K, et al: Development and evaluation of a workplace walking program. JJHEP 22: 133—145, 2014 (in Japanese).
4) Tomita Y, Nomura T, Sato A: About the exercise intervention effect that used a mobile telephone E-mail function: A trial of exercise intervention in youth. Bulletin of Kochi Women's University. Series of Faculty of Human Life and Environmental Science 56: 35—39, 2007 (in Japanese).
5) Nomura T, Akezaki Y, Tomita Y, et al: Promoting Exercise in Elderly Japanese People with Motivational Signs: Does Short-term Intervention Influence Exercise Habits and HQOL? JRHS 6: 5—10, 2008.
6) Yamatsu K, Kumagai S: Physical activity intervention based on information communication technology. Journal of Health Sciences 32: 31—38, 2010 (in Japanese).
7) Ministry of Health, Labour and Welfare of Japan: State of work related injuries & illnesses, 2013. http://www.mhlw.go.jp/bunya/roudoukijun/anzeneisei11/h25.html (accessed 2015-5-30)

8) Deyo RA, Rainville J, Kent DL: What can the history and physical examination tell us about low back pain? JAMA 268: 760—765, 1992.
9) Deyo RA, Weinstein JN: Low back pain. N Engl J Med 344: 363—370, 2001.
10) Matsudaira K: Psychological and social factors contribute to low back pain at work; Countermeasures to non-specific low back pain in the workplace. Occupational Health Journal 33: 60—66, 2010 (in Japanese).
11) Takano K, Asada F, Nomura T, et al: Approaches to preventing and improving low back pain through physical consulting. JJOMT 62(suppl.): 167: 2014 (in Japanese).
12) Matsudaira K: A New Concept of Back Pain for the 21st Century. The Occupational Health Promotion Foundation, 2012(in Japanese).
13) Ilmarinen J: The Work Ability Index (WAI). Occup Med 57: 160, 2007.
14) Matsudaira K, Inuzuka K, Kikuchi N, et al: Development of the Japanese version of the Fear-Avoidance Beliefs Questionnaire (FABQ-J), translation and linguistic validation. Orthopedic Surgery 62: 1301—1306, 2011 (in Japanese).
15) Kreuter MW, Wray RJ: Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 27: S227—232, 2003.
16) Mansell G, Kamper SJ, Kent P: Why and how back pain interventions work: what can we do to find out? Best Pract Res Clin Rheumatol 27: 685—697, 2013.
17) Takeuchi D, Taneichi H, Nohara Y: Satisfaction levels and adherence among patients being treated for chronic low back pain. Diagnosis and Treatment 99: 911—918, 2011 (in Japanese).

Reprint request:
Takuo Nomura
Department of Rehabilitation Sciences, Faculty of Allied Health Sciences, Kansai University of Welfare Sciences, 1-11, 3-chome Asahigaoka, Kashiwara-city, Osaka, 582-0026, Japan.

別刷請求先 〒582-0026 大阪府柏原市旭ヶ丘3丁目11番1号
関西福祉科学大学保健医療学部リハビリテーション学科
野村 卓生

腰痛予防を目的とした理学療法士によるメール指導の現状と課題

野村 卓生[1)]，浅田 史成[2)]，高野賢一郎[3)]，松平 浩[4)]
[1)]関西福祉科学大学保健医療学部リハビリテーション学科
[2)]大阪労災病院治療就労両立支援センター
[3)]関西労災病院治療就労両立支援センター
[4)]東京大学医学部附属病院22世紀医療センター運動器疼痛メディカルリサーチ＆マネジメント講座

—キーワード—
理学療法士，腰痛，産業保健

腰痛は，最も仕事に支障をきたしやすい疾患であり，いかに腰痛を予防し減らしていくかという職場での対策の確立が不可欠となっている．我々は以前に，労働者（以下，相談者）の腰痛予防を目的として理学療法士（以下，指導者）によるメール指導効果を検討するためにPhysical Consultant研究（以下，PCo研究）を実施した．PCo研究では，相談者の労働能力適応指標および腰痛恐怖回避思考の有意な改善を認めた．本研究の目的は，コンピューターや携帯端末機器を使用したメール指導に関して，より効果的な方法を検討することである．方法は，PCo研究のデータベースを用い，この研究で行われた指導者と相談者へのアンケート調査結果を分析することとした．結果，1人の指導者が同時期に5名の相談者の相談に対応可能なことが明らかとなった．相談者からは，腰痛予防に関連する内容だけではなく，他の筋骨格系の問題，生活習慣や家族の介護に関する相談があった．理学療法士は，これらの相談に対応できる最適な職種として考えられた．より多数の相談者および効果的なメール指導を行うにあたっては専用のシステム開発が必要と考えられた．

利益相反：利益相反基準に該当無し

（日職災医誌，64：113—118，2016）

PLOS | ONE

RESEARCH ARTICLE

Predictive Factors for Subjective Improvement in Lumbar Spinal Stenosis Patients with Nonsurgical Treatment: A 3-Year Prospective Cohort Study

Ko Matsudaira[1], Nobuhiro Hara[2], Hiroyuki Oka[1]*, Junichi Kunogi[3], Takashi Yamazaki[2], Katsushi Takeshita[4], Seichi Atsushi[5], Sakae Tanaka[6]

1 Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, University of Tokyo, Tokyo, Japan, **2** Department of Orthopaedic Surgery, Musashino Red Cross Hospital, Tokyo, Japan, **3** Department of Spine and Orthopaedic Surgery, Japanese Red Cross Medical Center, Tokyo, Japan, **4** Department of Orthopaedic Surgery, Jichi Medical University, Tochigi, Japan, **5** Department of Orthopaedic Surgery, Mitsui Memorial Hospital, Tokyo, Japan, **6** Department of Orthopaedic Surgery, University of Tokyo, Tokyo, Japan

* okah-tky@umin.ac.jp

Citation: Matsudaira K, Hara N, Oka H, Kunogi J, Yamazaki T, Takeshita K, et al. (2016) Predictive Factors for Subjective Improvement in Lumbar Spinal Stenosis Patients with Nonsurgical Treatment: A 3-Year Prospective Cohort Study. PLoS ONE 11(2): e0148584. doi:10.1371/journal.pone.0148584

Editor: Shervin Assassi, University of Texas Health Science Center at Houston, UNITED STATES

Received: June 24, 2015

Accepted: January 19, 2016

Published: February 10, 2016

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Objective

To assess the predictive factors for subjective improvement with nonsurgical treatment in consecutive patients with lumbar spinal stenosis (LSS).

Materials and Methods

Patients with LSS were enrolled from 17 medical centres in Japan. We followed up 274 patients (151 men; mean age, 71 ± 7.4 years) for 3 years. A multivariable logistic regression model was used to assess the predictive factors for subjective symptom improvement with nonsurgical treatment.

Results

In 30% of patients, conservative treatment led to a subjective improvement in the symptoms; in 70% of patients, the symptoms remained unchanged, worsened, or required surgical treatment. The multivariable analysis of predictive factors for subjective improvement with nonsurgical treatment showed that the absence of cauda equina symptoms (only radicular symptoms) had an odds ratio (OR) of 3.31 (95% confidence interval [CI]: 1.50–7.31); absence of degenerative spondylolisthesis/scoliosis had an OR of 2.53 (95% CI: 1.13–5.65); <1-year duration of illness had an OR of 3.81 (95% CI: 1.46–9.98); and hypertension had an OR of 2.09 (95% CI: 0.92–4.78).

Conclusions

The predictive factors for subjective symptom improvement with nonsurgical treatment in LSS patients were the presence of only radicular symptoms, absence of degenerative spondylolisthesis/scoliosis, and an illness duration of <1 year.

Introduction

Lumbar spinal stenosis (LSS) presents with neurological symptoms, such as numbness, pain, and intermittent claudication, in the lower extremities due to a narrowing of the intervertebral foramen and spinal canal, which serve as a passageway for nerves in the lumbar region.[1] Because of these symptoms, LSS is an important risk factor for decreased quality of life (QOL), particularly in the elderly. Previous epidemiological studies in Japan indicated a prevalence of LSS among people aged ≥70 years of approximately 10%.[2] With the aging society, the number of patients with LSS is predicted to rapidly increase. Thus, LSS is a disease that will be frequently encountered by primary care physicians.

With regard to LSS prognosis, several reports have demonstrated better outcomes with surgery compared with nonsurgical treatments.[3–5] Conversely, various other reports have revealed that, in some patient groups with relatively mild symptoms, the disease's natural course has a favourable prognosis.[6–10] However, patients with mild symptoms were excluded from some studies, and, in other cases, patients with severe symptoms requiring surgery were excluded. Therefore, it is not possible to draw conclusions regarding the natural history of LSS in all patients. To determine which patients have favourable prognoses, studies need to be conducted on a wide range of patients with LSS, regardless of the disease severity and therapeutic methods. However, to the best of our knowledge, no such study has been conducted.

Our hypothesis was that pre-treatment factors, such as duration of illness, types of symptoms, radiographic features, comorbidity, would predict patients' subjective improvement without surgical intervention. The aim of this study was to establish the evidence for favourable prognoses without surgical intervention.

Materials and Methods

Study design

This study was an investigator-initiated observational cohort study conducted at 17 medical centres in Japan, in which a wide variety of treatments, including surgical and conservative methods, were used in the treatment of spinal diseases. This study was approved by institutional review board of University of Tokyo, Tokyo Metropolitan Geriatric Hospital, Hitachi General Hospital, Asama General Hospital, MIshuku Hospital, Musashino Red Cross Hospital, Tokyo Metropolitan Tama Synthesis Medical Center, Japanese Red Cross Medical Center, Tokyo Yamate Medical Center, NTT Medical Center Tokyo, Sanraku Hospital, Kanto Central Hospital, Tokyo Metropolitan Hiroo Hospital, Tokyo Metropolitan Komagome Hospital, Kosei Hospital, Yokohama Rosai Hospital, Toranomon Hospital, and written informed consent was obtained from all participants.

Patient population

Patients with LSS were enrolled from the University of Tokyo Hospital and 17 related facilities between July 2002 and June 2003 based on the following eligibility criteria: aged 50–85 years

old and LSS based on the definition of Verbiest [11] (presence of paraesthesia or pain in the lower extremities, buttocks, perineum, or perianal region and magnetic resonance imaging showing the presence of spinal canal stenosis that may explain the patient's symptoms). Based on the pathogenesis, the patient's condition was required to be degenerative acquired stenosis (e.g., spondylosis, spondylolisthesis, or scoliosis), and patients with congenital, developmental, or post-traumatic LSS as well as those who underwent spinal surgery were excluded. The exclusion criteria were also as follows: presence of lumbar disc herniation (i.e., a positive straight leg raise test); arteriosclerosis obliterans (i.e., non-palpable foot arteries); complications causing disorders that interfere with gait, such as those after cerebral infarction or myelopathy; diagnosis of lower extremity symptoms because of peripheral nerve diseases; rheumatoid arthritis or Parkinson's disease; current administration of psychosomatic medicine or outpatient treatment at a psychiatric department; and compensation for damage.

Of the 314 patients that were screened, the study enrolled 274 patients (151 men, 123 women; mean age, 71 years) whose eligibility was guaranteed by a third-party evaluation.

In this study, a database was created by prospectively enrolling patients with LSS, regardless of the disease severity or treatment. Three years later, their prognosis was examined, and the factors that led to a subjective improvement in their symptoms without surgical intervention were assessed.

Study interventions

The treatment choice was made by the patients and physicians of each facility. The therapeutic methods included surgery (i.e., posterior lumbar decompression, posterior lumbar spinal fusion, or anterior lumbar interbody fusion) and nonsurgical methods (i.e., administration of non-steroidal anti-inflammatory drugs or prostaglandin E1 derivatives, exercise therapy, physical therapy, or nerve blocks). There was no limitation to the treatment selection.

Study measures

The following variables were examined at initial enrolment: degree of obesity (body mass index: $\geq$25 or $<$25 kg/m^2), educational background (at least a high school graduate, other), current comorbidities (hypertension, diabetes mellitus), duration of illness ($<$12 months, 12–59 months, or $\geq$60 months), types of symptoms (presence of cauda equina symptoms, at least the presence of bilateral numbness in the lower limbs), and presence of degenerative spondylolisthesis (% slip $\geq$5%) and scoliosis (Cobb angle $\geq$10 degrees) on radiographs. In addition, the Geriatric Depression Scale (GDS)-15, which is the abridged version of the GDS-30, was administered and assigned to tertiles defined by approximate thirds of the score distribution (0–2, 3–6, and $\geq$7) to assess depression.[12]

Three years after enrolment, a self-administered survey was delivered by mail to examine the patients' subjective improvement and determine whether surgery had been performed. In addition, the study centre also contacted survey non-respondents by telephone as an alternative form of contact to increase the response rates. The subjective degree of improvement was based on a 5-point scale, with 1 and 2 points indicating improvement without surgical intervention: 1) the condition has improved a lot; 2) the condition has improved; 3) nothing has changed; 4) the condition has become worse; and 5) the condition has become a lot worse.

Statistical analysis

A multivariable logistic regression model was used to assess the relationship between the candidate variables and patients' subjective improvement without surgical intervention. The following candidate variables were included in the final regression model when $P < 0.10$ in the

univariable analysis: age, sex, obesity, educational background, duration of illness, types of symptoms, and the presence of each of degenerative spondylolisthesis/degenerative scoliosis, hypertension, diabetes, and depression (GDS-15). Statistical analyses were performed using SPSS version 20.0 (IBM Corp., Armonk, NY, USA). A P value < 0.05 was considered to be statistically significant, and all reported P values are two sided.

Results

The 3-year follow-up rate was 67.5% (n = 185). There were no differences in the candidate variables between the 185 patients who completed the follow-up survey and the 89 patients who did not (Table 1).

Nonsurgical treatment resulted in subjective improvements in 56 (30.3%) of the 185 patients, and the condition worsened or did not change in 47 (25.4%) patients. In 82 patients (44.3%), surgery was performed within the 3-year follow-up (Fig 1). The proportion of patients with improvement was not significantly different between the groups (surgical treatment: 51/82, 62.2%; nonsurgical treatment: 57/103, 55.5%; P = 0.28).

The univariable analysis revealed that the duration of illness, types of symptoms, and the presence of each of degenerative spondylolisthesis/scoliosis, hypertension, and depression were significant explanatory variables ($P < 0.10$) (Table 2). The multivariable analysis with these explanatory factors showed that the absence of cauda equina symptoms (only radicular symptoms) had an odds ratio (OR) of 3.31 (95% confidence interval [CI]: 1.50–7.31); absence of degenerative spondylolisthesis/scoliosis had an OR of 2.53 (95% CI: 1.13–5.65); a <1-year duration of illness had an OR of 3.81 (95% CI: 1.46–9.98); and hypertension had an OR of 2.09 (95% CI: 0.92–4.78) (Table 3).

Table 1. Baseline characteristics, compared between the participants with lumbar spinal stenosis who did and did not complete the 3-year follow-up.

	Participants (n = 185)	Drop-outs (n = 89)	P-value
Age (years), mean (SD)	70.7 (7.4)	71.7 (7.6)	0.28
BMI (kg/m^2), mean (SD)	23.4 (3.1)	23.2 (3.1)	0.53
Gender (%)			
Female	77 (41.6)	46 (51.7)	0.12
Educational background			
At least a high school graduate	134 (72.4)	64 (71.9)	0.93
Cauda equina symptoms	78 (42.2)	44 (49.4)	0.26
Degenerative spondylolisthesis/scoliosis	99 (53.5)	47 (47.5)	0.91
Duration of illness (months)			
<12	48 (26.0)	23 (25.8)	0.99
12–59	80 (43.2)	38 (42.7)	
≥60	57 (30.8)	28 (31.5)	
Hypertension	120 (64.9)	57 (64.0)	0.89
GDS score (tertiles)			
0–2	73 (39.5)	27 (30.3)	0.13
3–6	64 (34.6)	29 (32.6)	
≥7	48 (25.9)	33 (37.1)	

BMI, body mass index; SD, standard deviation; GDS, Geriatric Depression Scale

The values are reported as n (%), unless indicated.

doi:10.1371/journal.pone.0148584.t001

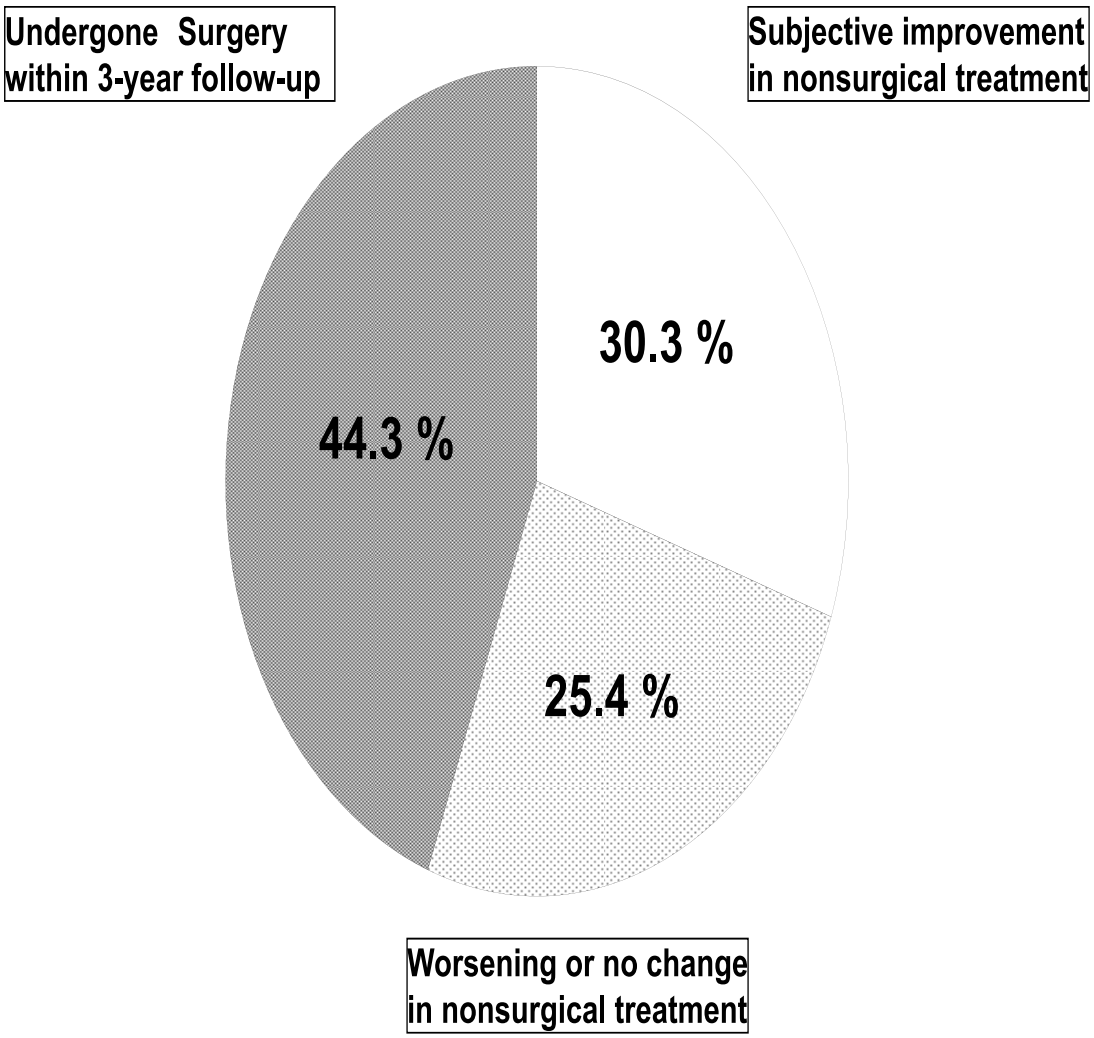


Fig 1. Response to the self-administered survey in 185 patients with lumbar spinal stenosis 3 years after treatment.

doi:10.1371/journal.pone.0148584.g001

Discussion

In patients with LSS from multiple medical centres and varying levels of disease severity and treatments, nonsurgical treatment resulted in subjective improvement of the symptoms at 3 years after enrolment in 30% of the patients; however, in 70% of the patients, the symptoms

Table 2. Univariable logistic regression analyses for 3-year subjective improvement in lumbar spinal stenosis symptoms through nonsurgical treatment.

Baseline factors	n	Odds ratio(95% CI)	P-value
Age (years)			
<65	40	1.69 (073–3.95)	0.22
65–74	80	1.21 (0.58–2.51)	0.61
≥75	65	1.00	
BMI (kg/m^2)			
<25	129	088 (0.45–1.73)	0.71
≥25	56	1.00	
Gender			
Female	77	1.19 (0.63–2.25)	0.85
Male	108	1.00	
Educational background (at least a high school graduate)			
Yes	134	1.00	
No	51	0.95 (0.47–1.91)	0.88
Cauda equina symptoms			
Yes	78	1.00	
No	107	4.42 (2.10–9.30)	< 0.001
Degenerative spondylolisthesis/degenerative scoliosis			
Yes	86	1.00	
No	99	2.11 (1.10–4.03)	0.03
Duration of illness (months)			
<12	47	3.68 (1.54–8.81)	0.003
12–59	79	1.72 (0.76–3.89)	0.2
≥60	59	1.00	
Hypertension			
Yes	65	1.00	
No	120	1.96 (0.97–3.95)	0.059
GDS score (tertiles)			
0–2	73	2.07 (0.88–4.14)	0.09
3–6	64	1.73 (0.88–4.83)	0.22
≥7	48	1.00	

BMI, body mass index; GDS, Geriatric Depression Scale; CI, confidence interval

doi:10.1371/journal.pone.0148584.t002

remained unchanged, worsened, or were treated surgically. Multivariable analysis showed that the factors associated with the improvement of subjective symptoms at 3 years after treatment were the presence of only radicular symptoms, the absence of degenerative spondylolisthesis and scoliosis, and an illness duration of <1 year.

The present study was conducted using a large-scale cohort of LSS patients from multiple medical centres, regardless of the disease severity, resulting in more representative data than previous studies. However, the present study did not include patients with very mild symptoms, who tend not to present at hospitals. Therefore, the prognosis may be slightly different from that in patients with more severe LSS. In addition, the degree of improvement of subjective symptoms was used as the measure of improvement; therefore, there may be differences in the actual improvement. However, the LSS severity is often defined on the basis of the intensity of lower extremity pain, and, because there are no well-defined classifications or criteria, the degree of subjective improvement may be closest to the actual degree of improvement. In a

Table 3. Multivariable logistic regression analyses for 3-year subjective improvement in lumbar spinal stenosis symptoms through nonsurgical treatment.

Baseline factors	Odds ratio (95% CI)	P-value
Cauda equina symptoms		
Yes	1.00	
No	3.31 (1.50–7.31)	0.003
Degenerative spondylolisthesis/degenerative scoliosis		
Yes	1.00	
No	2.53 (1.13–5.65)	0.024
Duration of illness (months)		
<12	3.81 (1.46–9.98)	0.007
12–59	1.87 (0.77–4.54)	0.17
≥60	1.00	
Hypertension		
Yes	1.00	
No	2.09 (0.92–4.78)	0.08
GDS score (tertiles)		
0–2	2.05 (0.80–5.25)	0.14
3–6	1.80 (0.70–4.68)	0.23
≥7	1.00	

GDS, Geriatric Depression Scale; CI, confidence interval

doi:10.1371/journal.pone.0148584.t003

study that compared surgically treated to conservatively treated patients and conducted follow-ups with 19 patients for an average of 31 months, [13] symptoms improved in 30% and remained unchanged in 60% of the conservatively treated patients who did not undergo any procedure. Despite the study's limitations, including its retrospective nature, unknown inclusion criteria for the conservatively treated patients, and small sample size, the rate of improvement was comparable to that of our cohort. Similarly, in a 5-year follow-up with 120 patients in whom conservative treatment was initially effective, an improvement was found in 43% of patients, the symptoms remained unchanged in 17%, and symptoms worsened in 40% at the final follow-up; however, the patients may have had relatively mild initial symptoms.[6] Moreover, in a prospective, randomised comparative study of surgical treatment for LSS, observations at 10 years after treatment in the 18 patients that received conservative treatment (control group) revealed mild pain in 2 patients (11%), moderate/severe pain in 6 patients (33%), and surgical treatment in 9 patients.[14] At the 2-year follow-up of a randomised cohort study with patients without spinal instability who were identified as surgical candidates and randomised to either surgical or conservative treatment, 43% of the patients with conservative treatment had to be re-assigned to the surgery group, while 28.7% reported an improvement in their symptoms.[15] However, because the patients with improved symptoms did not include those who were converted to the surgery group, it is possible that the percentage would be lower than those in the present study if the percentage was calculated in the same manner. Furthermore, the differences in results in these latter two studies, when compared with the present study, may be explained by the fact that the patients were indicated for surgery and may have had more severe conditions. However, in our study, if long-term follow-up was conducted, the percentage of patients with a favourable prognosis would likely decrease.

There are few reported studies regarding the predictive factors for the subjective improvement of LSS. However, Miyamoto et al. reported that the outcomes were favourable in patients with

radicular-type symptoms and in those who showed good improvement after the initial treatment, while outcomes were poor in patients with degenerative scoliosis.[6] Based on our experience, the prognosis of patients with radicular type LSS has been favourable; however, the underlying mechanism is not yet known. In the present study, the percentage of patients with cauda equina deficits whose treatment was converted to surgery was 3 times higher than that of patients with only radicular type LSS, which may support our experience. Degenerative scoliosis/spondylolisthesis was also predictive of poor prognosis in the present study; conservative treatment is reportedly less effective against degenerative scoliosis, [16] including at a 2-year follow-up.[17] It is possible that patients who repeatedly develop radiculopathy symptoms because of a susceptibility to physical compression have a poorer prognosis, and their treatment is likely to be converted to surgery. In addition, long illness duration has been associated with poor surgical outcomes in LSS; [18] likewise, our findings showed that, in the natural course of LSS, illness duration $\geq$1 year was also a factor for poor prognosis. A long illness duration likely leads to chronic nerve compression, which may cause oedema or Wallerian degeneration of the affected nerves.[19] Although hypertension was not a significant prognostic factor, it tended to be associated with a poor prognosis. Hypertension is more common in patients with LSS than in controls; [20,21] it causes arteriosclerosis and promotes degenerative changes in the spine and intervertebral discs.[22] Because it can also cause chronic obstructive arteriosclerosis, it may aggravate the prognosis; therefore, further studies are needed to determine if hypertension is related with prognosis in LSS.

This study has several limitations. First, because the follow-up rate was 67%, the presence of non-response bias is possible. Second, we intended to exclude lumbar disc herniation with the use of the straight leg raise test. However, the test was often negative in the elderly, even though they had undergone surgery for lumbar disc herniation. Furthermore, disc herniation is often prevalent in degenerative spine and is a concomitant cause of stenosis.[23] Thus, it was difficult to determine whether the cause of lumbar radiculopathy was lumbar disk herniation or LSS in our population, and it is possible that the influence of disk herniation was underestimated. Third, this study collected data at only a single time point, at 3 years from the date of enrolment. Therefore, the results failed to capture the time course of the disease, the rate of improvement, or requirement for surgical treatment. Additionally, we did not control for the nature, intensity, or duration of surgical or nonsurgical management.

The present study, with a wide range of patients with LSS, provided important findings that have not been reported previously and will aid decision-making regarding LSS treatment. In patients with radicular-type symptoms without degenerative scoliosis or spondylolisthesis and an illness duration of $<$1 year, the prognosis is likely to be favourable; however, in patients with cauda equina symptoms, degenerative scoliosis or spondylolisthesis, and a long disease duration, surgery may need to be proactively considered.

Future long-term follow-up of this cohort should be conducted, potentially with a questionnaire that more accurately measures disease severity and degree of satisfaction, such as the Zurich Claudication Questionnaire developed by Stucki et al., which is currently being used worldwide.[24] Determining the long-term prognosis of LSS may be useful for developing treatment guidelines.

Conclusion

In 30% of 274 patients with LSS, conservative treatment led to a subjective improvement in the symptoms at the 3-year follow-up; however, in 70% of the patients, the symptoms remained unchanged, worsened, or required surgical treatment. The predictive factors for improved subjective symptoms were the presence of only radicular symptoms, the absence of degenerative spondylolisthesis and scoliosis, and an illness duration of $<$1 year.

Supporting Information

S1 File. Supporting information. Dataset of this study.
(XLSX)

Author Contributions

Conceived and designed the experiments: KM HO. Performed the experiments: KM NH JK TY KT SA ST. Analyzed the data: HO. Contributed reagents/materials/analysis tools: KM HO. Wrote the paper: KM NH HO.

References

1. Katz JN, Harris MB. Clinical practice. Lumbar spinal stenosis. N Engl J Med 2008; 358:818–825. doi: 10.1056/NEJMcp0708097 PMID: 18287604
2. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012; 20:1103–1108. doi: 10.1016/j.joca.2012.06.018 PMID: 22796511
3. Athiviraham A, Yen D. Is spinal stenosis better treated surgically or nonsurgically? Clin Orthop Relat Res 2007; 458:90–93. PMID: 17308483
4. Kovacs FM, Urrútia G, Alarcón JD. Surgery versus conservative treatment for symptomatic lumbar spinal stenosis: a systematic review of randomized controlled trials. Spine 2011;15; 36:E1335–1351.
5. Munting E, Röder C, Sobottke R, Dietrich D, Aghayev E; Spine Tango Contributors. Patient outcomes after laminotomy, hemilaminectomy, laminectomy and laminectomy with instrumented fusion for spinal canal stenosis: a propensity score-based study from the Spine Tango registry. Eur Spine J 2014; 24:358–368. doi: 10.1007/s00586-014-3349-0 PMID: 24840246
6. Miyamoto H, Sumi M, Uno K, Tadokoro K, Mizuno K. Clinical outcome of nonoperative treatment for lumbar spinal stenosis, and predictive factors relating to prognosis, in a 5-year minimum follow-up. J Spinal Disord Tech 2008; 21:563–568. PMID: 19057249
7. Atlas SJ, Keller RB, Robson D, Deyo RA, Singer DE. Surgical and nonsurgical management of lumbar spinal stenosis: four-year outcomes from the Maine Lumbar Spine Study. Spine 2000; 25:556–562. PMID: 10749631
8. Atlas SJ, Keller RB, Wu YA, Deyo RA, Singer DE. Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the Maine Lumbar Spine Study. Spine 2005; 30:936–943. PMID: 15834339
9. Siebert E, Prüss H, Klingebiel R, Failli V, Einhäupl KM, Schwab JM. Lumbar spinal stenosis: syndrome, diagnostics and treatment. Nat Rev Neurol 2009; 5:392–403. doi: 10.1038/nrneurol.2009.90 PMID: 19578346
10. Benoist M. The natural history of lumbar degenerative spinal stenosis. Joint Bone Spine 2002; 69:450–457. PMID: 12477228
11. Verbiest H. A radicular syndrome from developmental narrowing of the lumbar vertebral canal. J Bone Joint Surg Br 1954; 36-B:230–237. PMID: 13163105
12. Almeida OP, Almeida SA. Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and SM-4. Int J Geriatr Psychiatry 1999: 14:858–865. PMID: 10521885
13. Johnsson KE, Udén A, Rosén I. The effect of decompression on the natural course of spinal stenosis. A comparison of surgically treated and untreated patients. Spine 1991; 16:615–619. PMID: 1862399
14. Amundsen T, Weber H, Nordal HJ, Magnaes B, Abdelnoor M, Lilleås F. Lumbar spinal stenosis: conservative or surgical management? A prospective 10-year study. Spine (Phila Pa 1976) 2000; 25:1424–1436.
15. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 2007; 356:2257–2270. PMID: 17538085
16. Everett CR, Patel RK. A systematic literature review of nonsurgical treatment in adult scoliosis. Spine (Phila Pa 1976) 2007; 32:S130–134.
17. Bridwell KH, Glassman S, Horton W, Shaffrey C, Schwab F, Zebala LP, et al. Does treatment (nonoperative and operative) improve the two-year quality of life in patients with adult symptomatic lumbar

scoliosis: a prospective multicenter evidence-based medicine study. Spine (Phila Pa 1976) 2009; 34:2171–2178.

18. Ng LC, Tafazal S, Sell P. The effect of duration of symptoms on standard outcome measures in the surgical treatment of spinal stenosis. Eur Spine J 2007; 16:199–206. PMID: 16496190
19. Kobayashi S, Yoshizawa H, Yamada S. Pathology of lumbar nerve root compression. Part 1: Intraradicular inflammatory changes induced by mechanical compression. J Orthop Res 2004; 22:170–179. PMID: 14656677
20. Lotan R, Oron A, Anekstein Y, Shalmon E, Mirovsky Y. Lumbar stenosis and systemic diseases: is there any relevance? J Spinal Disord Tech 2008; 21:247–251. PMID: 18525484
21. Uesugi K, Sekiguchi M, Kikuchi S, Konno S. Relationship between lumbar spinal stenosis and lifestyle-related disorders: a cross-sectional multicenter observational study. Spine (Phila Pa 1976) 2013; 38: E540–545.
22. Kauppila LI, Mikkonen R, Mankinen P, Pelto-Vasenius K, Mäenpää I. MR aortography and serum cholesterol levels in patients with long-term nonspecific low back pain. Spine (Phila Pa 1976) 2004; 29:2147–2152.
23. Rothoerl RD, Woertgen C, Holzschuh M, Schlaier J. Are there differences in the symptoms, signs and outcome after lumbar disc surgery in the elderly compared with younger patients? Br J Neurosurg 1998; 12:250–253. PMID: 11013689
24. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. Spine (Phila Pa 1976) 1996; 21:796–803.

Industrial Health 2016, **54**, 282–292

Original Article

Psychological detachment from work during non-work time: linear or curvilinear relations with mental health and work engagement?

Akihito SHIMAZU[1, 2]*, Ko MATSUDAIRA[3, 4], Jan DE JONGE[2, 5], Naoya TOSAKA[1], Kazuhiro WATANABE[1, 6] and Masaya TAKAHASHI[7]

[1]Department of Mental Health, The University of Tokyo, Graduate School of Medicine, Japan
[2]Asia Pacific Centre for Work Health and Safety, University of South Australia, Adelaide, Australia
[3]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Japan
[4]Clinical Research Center for Occupational Musculoskeletal Disorders, Kanto Rosai Hospital, Japan
[5]Human Performance Management Group, Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, The Netherlands
[6]Japan Society for the Promotion of Science, Japan
[7]National Institute of Occupational Safety and Health, Japan

Received May 19, 2015 and accepted January 12, 2016
Published online in J-STAGE January 30, 2016

Abstract: This study examined whether a higher level of psychological detachment during non-work time is associated with better employee mental health (Hypothesis 1), and examined whether psychological detachment has a curvilinear relation (inverted U-shaped pattern) with work engagement (Hypothesis 2). A large cross-sectional Internet survey was conducted among registered monitors of an Internet survey company in Japan. The questionnaire included scales for psychological detachment, employee mental health, and work engagement as well as for job characteristics and demographic variables as potential confounders. The hypothesized model was tested with moderated structural equation modeling techniques among 2,234 respondents working in the tertiary industries with regular employment. Results showed that psychological detachment had curvilinear relations with mental health as well as with work engagement. Mental health improved when psychological detachment increased from a low to higher levels but did not benefit any further from extremely high levels of psychological detachment. Work engagement showed the highest level at an intermediate level of detachment (inverted U-shaped pattern). Although high psychological detachment may enhance employee mental health, moderate levels of psychological detachment are most beneficial for his or her work engagement.

Key words: Psychological detachment, Mental health, Structural equation modeling, Work engagement, Curvilinearity

Introduction

In recent years, scholars have argued that not only on-job experiences (how employees spend their working time) but also off-job experiences (how they spend their private or leisure time) are crucial for understanding employee well-being[1)]. More specifically, better knowledge of off-job recovery from the demands experienced during working time is imperative[2)]. Recovery can be defined as a process during which individual functional systems that have been

*To whom correspondence should be addressed.
E-mail: ashimazu@m.u-tokyo.ac.jp

called upon during a stressful experience return to their initial, pre-stressor level[3]. Recovery can be regarded a process opposite to the strain process, during which the detrimental effects of stressful situations are alleviated or eliminated. Recovery is also regarded as an explanatory mechanism in the relation between acute stress reactions and chronic health impairment[4]. Certain experiences outside of work can help in alleviating reactions to work demands[5–7]. These so-called recovery experiences consist of psychological detachment, relaxation, mastery, and control[8]. Psychological detachment; i.e., the ability of individuals to mentally "switch off" from work by not doing work-related tasks and not thinking about work during non-work time, is considered the most crucial recovery experience for protecting one's well-being regarding job-related recovery[2, 9].

In the context of respites from work, detachment has been described as an "individual's sense of being away from the work situation"[10]. Psychological detachment has been further characterized as not being involved in work-related activities, such as phone calls, e-mails, or other work-related tasks, during off-work time[8]. Psychological detachment from work extends beyond the pure physical absence from the workplace during off-job time and abstaining from job-related tasks. It implies leaving the workplace behind oneself in psychological terms[11].

The relation between psychological detachment and well-being can be explained by COR theory[12] and the Effort-Recovery Model[3]. Conservation Of Resources (COR) theory asserts that an individual aspires to preserve, protect, and build resources. Resources are characterized as objects, conditions, personal characteristics, or energies that have specific importance for the individual. According to COR theory, stress occurs when individuals are threatened with resource loss, actually lose resources, or fail to gain resources following resource investment. The inability to replenish energy resources may lead to long-term fatigue, which hampers normal functioning in many aspects in daily life, including work. Thus, to recover from stress, individuals have to gain new resources and restore threatened or lost resources. Psychological detachment can contribute to gaining new resources and restore threatened or lost resources.

The Effort-Recovery Model[3] holds that effort expenditure at work leads to load reactions such as fatigue or physiological activation. Load reactions can accumulate and lead to impaired health and well-being, unless individuals can recover from work. By no longer being exposed to job-related demands, load reactions can return to pre-stressor levels, and recovery can occur before the next working period starts. This implies that recovery strategies such as psychological detachment during off-work time can be an opportunity to return to and stabilize at a baseline level. Thus, both the Effort-Recovery Model and COR theory suggest two complementary processes by which recovery occurs. First, it is important to refrain from work demands and to avoid activities that call upon the same functional systems or internal resources as those required at work. Second, gaining new internal resources such as energy, self-efficacy or positive mood will additionally help to restore threatened resources[8].

Previous studies that examined the relation between psychological detachment and well-being have revealed that psychological detachment is positively associated with mental health and negatively associated with job stress and burnout[6, 8, 11, 13, 14]. Therefore, we expect that a higher level of psychological detachment during non-work time will be associated with better mental health (Hypothesis 1).

Regarding positive aspects of employee well-being, the present study focuses on work engagement, which refers to a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption[15]. Previous studies have shown that psychological detachment is positively associated with work engagement[16–18], because detachment may contribute to the prevention of continued resource drain and restoration of resources[18]. If employees do not unwind from one's work, depleted resources can lead to low work engagement. Thus, we can assume that low levels of psychological detachment are associated with low work engagement.

However, the relation between psychological detachment and work engagement appears to be more complex. For instance, Shimazu *et al.*[19] showed a negative relation between these variables, suggesting that switching off mentally during off-job time did not improve work engagement, but rather decreased it. When individuals are highly detached from their jobs during off-job time, they may feel difficulty in "switching on" again in the next morning[14], and they may need more time to mobilize their energy for their job, which results in impaired work engagement.

These findings suggest that (very) low and (very) high levels of psychological detachment will be detrimental to work engagement. As a result, moderate levels of psychological detachment will be associated with the highest levels of work engagement. All these findings imply non-linear rather than linear relations between detachment and work engagement, which is in line with Warr's (1994) assumptions on work[20], mental health and well-being. Accordingly, we expect that psychological detachment will

have a curvilinear relation (inverted U-shaped pattern) with work engagement (Hypothesis 2).

Method

Study population

An Internet research company with 1.5 million registered research volunteers aged 20–69 years, was used to conduct a large Internet-based cross-sectional survey on occupation, health and well-being in 2011. We randomly selected 106,250 volunteers from 201,170 monitors, living in three greater metropolitan areas of Japan (23 wards of Tokyo, the City of Osaka, and the City of Nagoya). On March 25, 2011, the selected volunteers were invited to take part in the study via an e-mail containing a link to the survey. Participants received online shopping points as an incentive for participation. In order to prevent double registration, e-mail addresses were checked and a link to the questionnaire was disabled once the survey was completed. On March 31, 2011, the survey was closed when more than five thousand participants responded (a total of 5,860 surveys were collected). Therefore, a specific response rate could not be calculated for this survey.

Our respondents were very close to the people living in 23 wards of Tokyo, the City of Osaka, and the City of Nagoya in terms of mean age (45.2 years in our respondents, 43.9 in Tokyo, 44.8 years in Osaka, and 43.8 years in Nagoya), gender (50.8% in our respondents, 50.7% in Tokyo, 51.5% in Osaka, and 50.7% in Nagoya), and employment status (46.5% regular employment in our respondents, 46.1% in Tokyo, 46.2% in Osaka, and 50.1% in Nagoya). However, our respondents had higher educational level (40.9% undergraduate or higher) than those living in Tokyo (33.2%), in Osaka (20.8%), and in Nagoya (26.0%)[21, 22].

In our respondents, the proportion of respondents working within primary industries (e.g., agriculture, forestry, and fisheries) and secondary industries (e.g., mining, manufacturing, and constructions) was extremely low (0.1% and 7.6% respectively). Therefore, we analyzed responses only from those individuals working in tertiary industries (e.g., transport and postal activity, wholesale and retail trade, accommodations, eating and drinking services, finance and insurance, advertising, education and learning support, and medical, health care and welfare). Individuals with a reported age of either <20 years or ≥ 65 years, those with non-regular employment, or shift workers were excluded[23–25]. A total of 2,234 participants were retained and included in the analysis. The mean age of the participants was 41.7 years (SD = 11.3). Of the participants, 63.9% were male, 54.4% were married, 55.9% had a university degree or higher, and 12.2% worked more than 60 hours per week.

Measures

Psychological detachment

Psychological detachment was assessed using the corresponding subscale of the Japanese version of the Recovery Experience Questionnaire[8, 19], consisting of four items (i.e., "I forget about work," "I don't think about work at all," "I distance myself from my work," and "I get a break from the demands of work"). All items were scored on a five-point Likert scale, ranging from 1 (do not agree at all) to 5 (fully agree). Responses for the 4 items were summed to get a scale score. Cronbach's alpha coefficient was .86.

Mental health

Mental health was assessed using the corresponding subscale of the SF-36 version 1.2[26–28], consisting of five items (i.e., "Have you been a very nervous person?", "Have you felt so down in the dumps that nothing could cheer you up?", "Have you felt calm and peaceful? (reversed)", "Have you felt downhearted and blue?", and "Have you been a happy person? (reversed)"). All items were scored on a six-point Likert scale, ranging from 1 (all of the time) to 6 (none of the time). We used the SF-36 mental health summary score as a measure of mental health (Range: 0–100)[29]. Cronbach's alpha coefficient was .84.

Work engagement

Work engagement was assessed using the short form of the Utrecht Work Engagement Scale (UWES)[15], which has been validated in Japan[30]. The UWES includes three subscales that reflect the underlying dimensions of engagement: Vigor (3 items; e.g., "At my job, I feel strong and vigorous"), Dedication (3 items; e.g., "I am enthusiastic about my job"), and Absorption (3 items; e.g., "I am immersed in my work"). All items are scored on a seven-point Likert scale ranging from 0 (never) to 6 (always). Responses for the 3 items each were summed to get a scale score. Cronbach's alpha coefficients were .87 for vigor, .84 for dedication, and .86 for absorption.

Potential confounders

We controlled for two types of potential confounders; i.e., (1) job characteristics and (2) demographic characteristics. Their relation with detachment and our outcome measures is well-established in the literature[4, 9, 11].

Job characteristics were assessed using three scales of

the Brief Job Stress Questionnaire (BJSQ[31]): job demands, job control and workplace support. The first two scales consisted of 3 items each, for instance "My job requires working hard" and "I have influence over the pace of my work". Workplace support consisted of 6 items: 3 items for supervisor support and 3 items for coworker support. To receive a more parsimonious model and to avoid multi-collinearity, we combined the two subscales in overall workplace support due to a high bivariate correlation ($r=0.59$; $p<.001$). All items were scored on a four-point Likert scale, ranging from 1 (disagree) to 4 (agree). Cronbach's alpha coefficients were .81 for job demands, .85 for job control, and .86 for workplace support.

Demographic characteristics such as age, gender, marriage, education, and working hours per week were also included as potential confounders in the questionnaire.

Data analyses

To test the hypotheses, we conducted moderated structural equation modeling (MSEM) analyses, using the AMOS software package[32]. We preferred MSEM to hierarchical regression analyses, because MSEM allows multivariate testing of outcomes, allows assessing and correcting for measurement error, and provides measures of fit of the models under study. We followed the procedure proposed by Mathieu *et al.*[33] as described by Cortina *et al.*[34]. Linear psychological detachment and mental health had only one indicator that was the standardized (centered) scale score of the respective factor[33]. The indicator of the latent curvilinear psychological detachment was the squared term of the standardized (centered) scale score of psychological detachment. Work engagement had three indicators (i.e., vigor, dedication, and absorption). Correlation between linear psychological detachment and curvilinear one was constrained to be zero, whereas mental health and work engagement were allowed to correlate. The paths from the latent exogenous factors to their indicators were fixed using the square roots of the scale reliabilities, and the error variances of each indicator were set equal to the product of their variances and 1 minus their reliabilities. See Fig. 1 for our hypothesized model. For more details regarding the calculation of the reliability score of the curvilinear term,

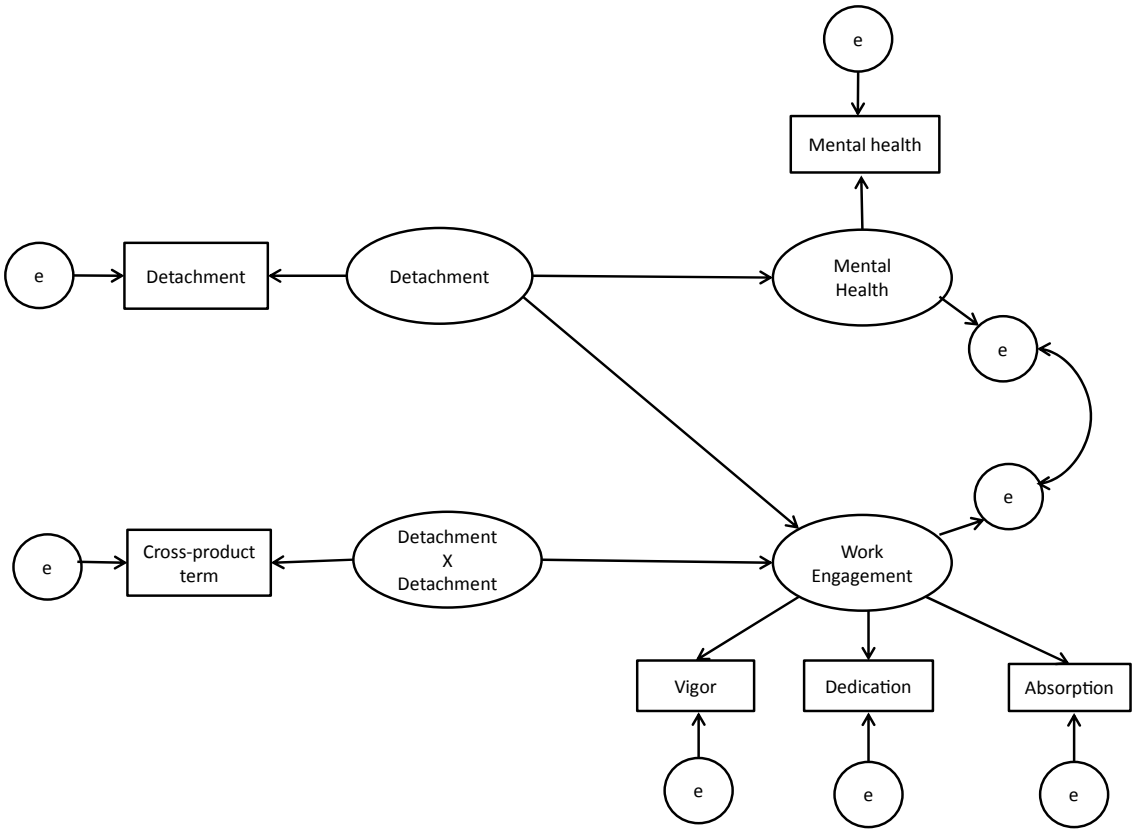


Fig. 1. Hypothesized model (Model 1).
Note: e=error.

we refer to Cortina *et al.*[34].

The fit of the models was assessed with the chi-square statistic, the goodness-of-fit index (GFI), the comparative fit index (CFI), the non-normed fit index (NNFI), and the root-mean-square error of approximation (RMSEA). It is suggested that GFI, CFI, and NNFI values that exceed .90 and RMSEA values as high as .08 are indicative of acceptable fit[35].

Ethics statement

This study was approved by the medical/ethics review board of the Japan Labour Health and Welfare Organization and The University of Tokyo medical department.

Results

Simple statistics

Zero-order correlation coefficients are shown in Table 1. Psychological detachment was positively correlated with mental health ($r=.22$, $p<.001$), and negatively correlated with vigor ($r=-.04$, $p<.05$), dedication ($r=-.06$, $p<.01$), and absorption ($r=-.14$, $p<.001$).

Results of MSES analyses

Results of the MSEM-analyses showed that the hypothesized model (Model 1) fits to the data ($\chi^2(8)=236.72$, $p<.001$, GFI=.97, NNFI=.93, CFI=.96) although RMSEA value exceeded .08 (RMSEA=.11). In line with Hypothesis 1, linear psychological detachment was positively related to mental health ($\beta=.24$, $p<.001$). As to Hypothesis 2, both linear and curvilinear psychological detachment were negatively related to work engagement ($\beta=-.10$, $p<.001$ and $\beta=-.06$, $p<.01$, respectively).

To ensure that no curvilinear relation existed between psychological detachment and mental health in addition to linear one, we examined the alternative model that adds the path from curvilinear psychological detachment to mental health. The model fit of the alternative model (Model 2: $\chi^2(7)=216.11$, $p<.001$, GFI=.97, NNFI=.92, CFI=.97, RMSEA=.12) was similar to one of the hypothesized model. However, the chi-square difference test, comparing the hypothesized model (Model 1) with the alternative model (Model 2), shows a significant improvement in model fit ($\Delta\chi^2(1)=20.61$, $p<.001$). This means that the alternative model (Model 2), including the path from curvilinear psychological detachment to mental health, offers a better account of the data than the hypothesized model (Model 1). Therefore, we decided to adopt the alternative model (Model 2) in further examination.

Table 1. Descriptive statistics for the key study variables (N=2,234)

	Variable	Range	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12
1	Age	20–64	41.74	11.31												
2	Gender[a]	0–1	.64	.48	−.10***											
3	Marriage[b]	0–1	.54	.50	−.41***	.31***										
4	Education[c]	0–1	.56	.50	−.08***	−.19***	−.03									
5	Working hours (per week)[d]	0–1	.12	.33	−.06**	−.15***	−.04*	.03								
6	Job demands	3–12	8.20	2.22	−.14***	−.07**	.02	.09***	.26***							
7	Job control	3–12	8.10	2.02	.18***	−.02	−.10***	.00	−.06**	−.16***						
8	Workplace support	6–24	15.20	3.89	−.03	.03	−.05*	−.01	−.01	−.02	.30***					
9	Psychological detachment	4–20	13.77	3.53	.01	.06**	.07**	−.05*	−.11***	−.25***	.07***	.05*				
10	Mental health	0–100	59.93	19.37	.14***	−.04*	−.12***	.03	−.09***	−.22***	.25***	.35***	.22***			
11	Vigor	0–18	6.93	3.67	.20***	.03	−.11***	.01	.01	.01	.31***	.30***	−.04*	.31***		
12	Dedication	0–18	8.25	3.77	.17***	.05*	−.09***	.01	.05*	.13***	.29***	.30***	−.06**	.25***	.82***	
13	Absorption	0–18	6.97	3.87	.14***	.00	−.08***	.03	.06**	.14***	.27***	.24***	−.14***	.18***	.79***	.83***

Note: *$p<.05$ **$p<.01$ ***$p<.001$. SD: Standard Deviation. [a] Gender was coded as 1 (men) and 0 (women). [b] Marriage was coded as 1 (yes) and 0 (no). [c] Education was coded as 1 (university or higher) and 0 (college or lower). [d] Working hours per week was coded as 1 (60=<) and 0 (<60).

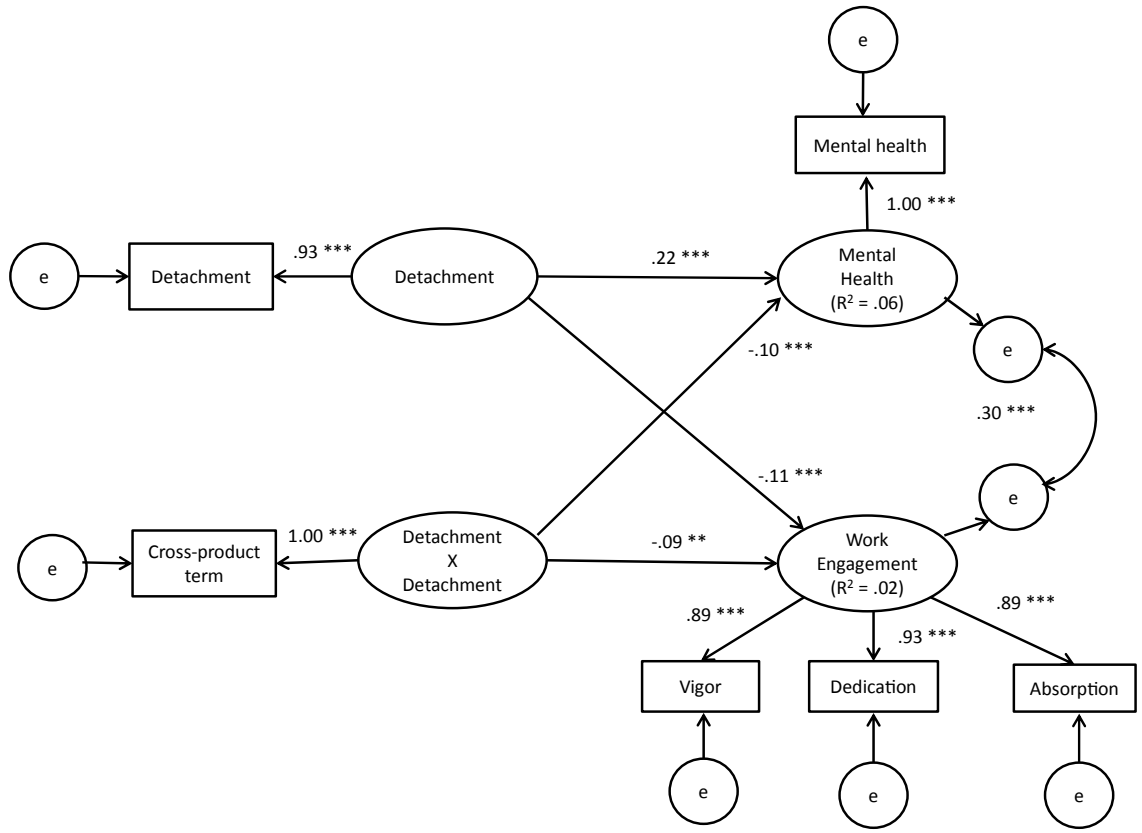


Fig. 2. Standardized solution (Maximum Likelihood estimates) of the final (alternative) model (Model 2: N=2,234).
Note: e=error. ***p<.001, **p< .01, *p<.05.

As can be seen in Fig. 2, linear psychological detachment was significantly and positively related to mental health (β=.22, p<.001) whereas curvilinear psychological detachment was also significantly but negatively related to it (β=−.10, p<.001). In addition, both linear and curvilinear psychological detachment were significantly and negatively related to work engagement (β=−.11, p<.001 and β=−.09, p<.01, respectively). Please note that the results regarding the curvilinear relationship between psychological detachment and work engagement were similar in all three sub dimensions of the construct (i.e., vigor, dedication, and absorption).

Regarding the curvilinear relation between psychological detachment and mental health, Fig. 3 shows that initially there is a positive relation: more detachment is associated with better mental health. However, at high levels of psychological detachment, the positive relation between psychological detachment and mental health became less prominent, and even seems to disappear. Mental health did not increase further and remained at a high level.

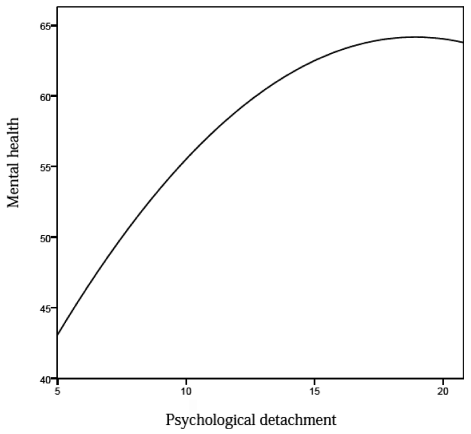


Fig. 3. Curve-fitting between psychological detachment and mental health.

With regard to the curvilinear relation between psychological detachment and work engagement, Fig. 4 shows that moderate levels of psychological detachment were associated with the highest levels of work engagement, whereas

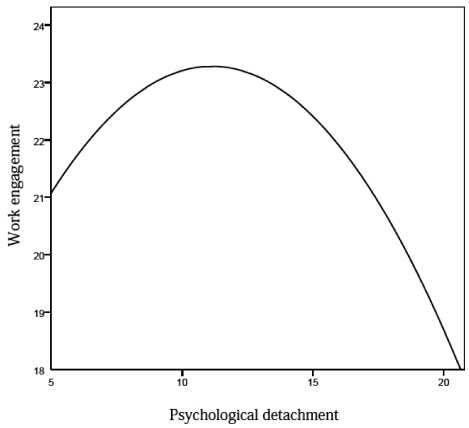


Fig. 4. Curve-fitting between psychological detachment and work engagement.

very low and very high detachment were associated with lower levels of work engagement (i.e., inverted U-shaped pattern).

In a final step, we conducted additional analysis to control for potential confounders (i.e., age, gender, marriage, education, working hours, job demands, job control, and workplace support). Specifically, each control variable was included in the alternative model (Model 2) as a manifest variable simultaneously and was allowed to relate to all variables in the model. After controlling for confounding variables, the path coefficients were virtually the same as those of the alternative model (Model 2), but the model fit decreased (χ^2 (35)=1538.06, $p<.001$, GFI=.91, NNFI=.53, CFI=.82, RMSEA=.14). These results indicate that the added relations of the control variables to the model variables were weak. Importantly, many control variables did not significantly affect the structural paths in the model (i.e., 18 out of 48 paths were not statistically significant). Therefore, the control variables were removed from the final model in Fig. 2.

Discussion

The aim of this large cross-sectional Internet survey study was to examine whether higher levels of psychological detachment during non-work time would be associated with improved employee mental health (Hypothesis 1). We also examined whether psychological detachment would have a curvilinear relation (i.e., inverted U-shaped pattern) with work engagement (Hypothesis 2). Examination of the curvilinear relation was novel, because prior research on the function of psychological detachment on work engagement is inconsistent in this respect[16–19].

As far as the relation between psychological detachment and mental health is concerned, MSEM revealed that not only linear psychological detachment ($\beta=.22$, $p<.001$) but also curvilinear detachment ($\beta=-.10$, $p<.001$) was significantly related to mental health. This result was contrary to our expectation. Examining Fig. 3, the positive relation between psychological detachment and mental health flattened after higher levels of psychological detachment. This pattern of findings suggests that mental health initially improves when people psychologically detach. However, employee mental health does not benefit any further from extremely high levels of psychological detachment. It is important to note that mental health does not suffer at such very high levels of psychological detachment. Although most previous studies showed that higher levels of psychological detachment during non-work time were associated with better employee mental health[6, 8, 11, 13], our result suggests that the favorable effect of psychological detachment may have an upper limit on mental health, at least among our participants. Future research needs to examine under which conditions and for whom psychological detachment has such a curvilinear relation with mental health.

As to the relation between psychological detachment and work engagement, we also found a curvilinear relation. Moderate levels of psychological detachment were associated with highest levels of work engagement, whereas very low and very high psychological detachment was associated with lower levels of work engagement (i.e., inverted U-shaped pattern). Very low levels of psychological detachment may drain one's resources and inhibit resource restoration, whereas very high levels of psychological detachment may require a longer time to get back into "working mode" in the next morning[9]. These may negatively impact work engagement, particularly at high levels of detachment. Finally, it is worth noting that the curvilinear relation between psychological detachment and work engagement resembles (albeit at a weaker level) a previously found relation between psychological detachment and job performance in earlier research[14]. Given that both of these are more strictly work-related variables, the current finding may have implications for future research on the topic.

Limitations and suggestions for future research

Next to several strengths such as a large sample size and sufficient study power, there are also several limitations of this study. First, we used self-report survey data. Self-report measures may be biased due to, for example, negative affect. Common method variance might have affected the results, suggesting that the true associations between

variables might be weaker than those observed in this study. Although several studies have shown that these influences are not as high as could be expected[36–38], our findings should be replicated using more objective measures (e.g., peer-ratings of mental health and work engagement) in the future.

Second, we used a cross-sectional study design, which precludes making causal inferences. For instance, our data showed that psychological detachment was related to better mental health. This might indicate that more psychological detachment leads to better mental health. It might also be that individuals enjoying better mental health are more likely to detach themselves from their work. Based on the cross-sectional analyses of the current study, it can only be concluded that psychological detachment is related to mental health and well-being. More longitudinal research is needed to uncover the causal sequence in the relation between psychological detachment and its consequences. However, it should be noted that there is a growing body of literature that demonstrates longitudinal effects of psychological detachment on health and well-being, particularly at day-level[39–42]. They support our causal inferences from both theoretical and empirical viewpoints.

Third, our data were collected from people living in three greater metropolitan areas of Japan (23 wards of Tokyo, the City of Osaka, and the City of Nagoya), which requires caution regarding the generalizability of our findings. Our sample may not represent other working populations quite well. Therefore, further studies are necessary to examine whether our results are applicable to workers in local areas.

Fourth, our data were collected via the Internet, which again requires caution regarding the generalizability of our findings. It has been claimed that the socioeconomic and educational status of the average Internet user is usually greater than that of the general population[43]. Indeed, our participants reported higher educational status than those completing nationwide paper-and-pencil surveys in Japan[44] and those living in Tokyo, in Osaka, and in Nagoya[21, 22]. Thus, similar to typical Internet studies, self-selection might be a limitation of the present study.

Finally, psychological detachment did not have much explanation for outcomes in our participants. Specifically, linear and curvilinear psychological detachment explained successively 6% and 2% of the variances of mental health and work engagement in Model 2. One possible explanation is that we did not examine the combined effects of psychological detachment and other types of recovery experiences. Until now, only bivariate associations of recovery experiences with outcome variables have mainly been investigated. However, in reality, it is less likely that people use either type of recovery experience exclusively. Rather, they may use various types of recovery experiences simultaneously given the positive correlations among them (e.g., $r = .16 - 63$ by Sonnentag[8], and $r = .26 - .70$ by Shimazu *et al.*[19]). Hence, it is important to examine the combined as well as independent associations of each type of recovery experience with well-being in employees. According to COR theory[12], employees using various type of recovery experiences simultaneously are assumed to experience better well-being because multiple recovery experiences may provide more opportunity for recovery from resource loss and for resource gain. Another possible explanation is that we did not consider conditions under which employees use psychological detachment. This suggests the possibility that psychological detachment may not be favorable for everybody and in all situations[45]. For instance, employees who experience their jobs as highly meaningful and enjoyable might find detachment difficult to achieve, but lack of detachment might be less of a problem for such people. Thus, job features might moderate the relation between psychological detachment and well-being. Future research needs to examine the conditions under which psychological detachment can have more favorable effects.

Implications for practice

Our findings have some implications for practice. A first implication is that psychological detachment during non-work time is associated with employee mental health and work engagement *in different ways*.

With regard to employee mental health, higher levels of detachment would facilitate better mental health (although the favorable effect of detachment had limitations). It is important that both organizations and supervisors should support employee detachment by advising that employees be as unavailable as possible (e.g., via e-mail, texting or phone) during their non-work time. It might be beneficial for workers to detach from work if they do not use their smartphones or tablets for work-related issues during free time[46–48]. However, it might also be possible that checking one's work e-mails helps to detach from work in particular circumstances. For example, if s/he is unsure whether s/he has forgotten to inform a colleague about an important work-related issue, to check the sent box of his/her e-mail account might help him/her thereafter to detach from work. Further research needs to examine whether the use of communication devices such as smartphones or tablets during non-work time can be beneficial or not for one's detachment from work. Organizations and supervisors can also support

employee detachment by not initiating work-related communication with their employees during non-work time, thereby allowing detachment to occur[14]. Supervisors can act as role models in this respect by not being available during non-work time. This is particularly important in a country like Japan, because those who are in charge of changing long working culture in Japan are often work addicts themselves[49]. Furthermore, improving working conditions to achieve adequate levels of job demands (e.g., reduce time pressure) can be a promising avenue to facilitate psychological detachment because high job demands can inhibit psychological detachment during off-work time[2].

It is also important for employees who are at risk for workaholism (i.e., working excessively with an obsessive manner[50]) to modify this tendency, since it inhibits psychological detachment[2]. Training programs that focus on time management and problem solving skills might be helpful, because workaholic employees take on more work than they can handle and accept new tasks before completing previous ones[51]. Rational emotive therapy[52] might be also helpful, since workaholic people suffer from the belief that they should be perfect[53].

With regard to work engagement, the relation with psychological detachment is more complex and suggest a different practical implication: Moderate levels of psychological detachment would be associated with the highest levels of work engagement. Although operationalizing the optimal level of psychological detachment seems to be not very easy, it should be noted that thinking about work may not be necessarily negative per se[9, 54]. Positively reflecting about one's work (e.g., thinking about a recent success or about an inspiring goal) might even improve work engagement, but this thinking should not be too much − there seems to be an upper limit for work reflection. Future research needs to clarify the preferable type and amount of work-related thoughts during off-job time to improve work engagement.

Conclusion

Although higher levels of psychological detachment may enhance employee mental health, it seems that moderate levels of psychological detachment are most beneficial for his or her work engagement. In future, more research is needed to address how, and under which conditions, to attain optimal levels of psychological detachment to achieve both better employee mental health and greater work engagement.

References

1) Shimazu A, Schaufeli WB (2009) Is workaholism good or bad for employee well-being? The distinctiveness of workaholism and work engagement among Japanese employees. Ind Health **47**, 495−502.
2) Shimazu A, de Jonge J, Kubota K, Kawakami N (2014) Psychological detachment from work during off-job time: predictive role of work and non-work factors in Japanese employees. Ind Health **52**, 141−6.
3) Meijman TF, Mulder G (1998) Psychological aspects of workload. In: Handbook of work and organizational psychology: Vol. 2. Work Psychology, Drenth PJD, Thierry H, de Wolff CJ (Eds.), 5−33, Psychology Press, Hove.
4) Geurts SAE, Sonnentag S (2006) Recovery as an explanatory mechanism in the relation between acute stress reactions and chronic health impairment. Scand J Work Environ Health **32**, 482−92.
5) Eden D (2001) Vacations and other respites: studying stress on and off the job. In: International review of industrial and organizational psychology, Cooper CL, Robertson IT (Eds.), 121−46, Wiley, Chichester.
6) Fritz C, Sonnentag S (2005) Recovery, health, and job performance: effects of weekend experiences. J Occup Health Psychol **10**, 187−99.
7) Sonnentag S (2001) Work, recovery activities, and individual well-being: a diary study. J Occup Health Psychol **6**, 196−210.
8) Sonnentag S, Fritz C (2007) The Recovery Experience Questionnaire: development and validation of a measure for assessing recuperation and unwinding from work. J Occup Health Psychol **12**, 204−21.
9) De Jonge J, Spoor E, Sonnentag S, Dormann C, Van den Tooren M (2012) "Take a break?!" Off-job recovery, job demands, and job resources as predictors of health, active learning, and creativity. Eur J Work Organ Psychol **21**, 321−48.
10) Etzion D, Eden D, Lapidot Y (1998) Relief from job stressors and burnout: reserve service as a respite. J Appl Psychol **83**, 577−85.
11) Sonnentag S, Bayer UV (2005) Switching off mentally: predictors and consequences of psychological detachment from work during off-job time. J Occup Health Psychol **10**, 393−414.
12) Hobfoll SE (1989) Conservation of resources. A new attempt at conceptualizing stress. Am Psychol **44**, 513−24.
13) Fritz C, Sonnentag S (2006) Recovery, well-being, and performance-related outcomes: the role of workload and vacation experiences. J Appl Psychol **91**, 936−45.
14) Fritz C, Yankelevich M, Zarubin A, Barger P (2010) Happy, healthy, and productive: the role of detachment from work during nonwork time. J Appl Psychol **95**, 977−83.
15) Schaufeli WB, Salanova M, Gonzalez-Romá V, Bakker AB (2002) The measurement of engagement and burnout: a two sample confirmatory factor analytic approach. J Happiness

Stud **3**, 71–92.
16) Kühnel J, Sonnentag S, Westman M (2009) Does work engagement increase after a short respite? The role of job involvement as a double-edged sword. J Occup Organ Psychol **82**, 575–94.
17) Siltaloppi M, Kinnunen U, Feldt T (2009) Recovery experiences as moderators between psychological work characteristics and occupational well-being. Work Stress **23**, 330–48.
18) Sonnentag S, Binnewies C, Mojza EJ (2010) Staying well and engaged when demands are high: the role of psychological detachment. J Appl Psychol **95**, 965–76.
19) Shimazu A, Sonnentag S, Kubota K, Kawakami N (2012) Validation of the Japanese version of the recovery experience questionnaire. J Occup Health **54**, 196–205.
20) Warr P (1994) A conceptual framework for the study of work and mental health. Work Stress **8**, 84–97.
21) Kawasaki city (2013) Kawasaki city judging from Big city comparison statistics chronological table. http://www.city.kawasaki.jp/shisei/category/51-4-9-6-0-0-0-0-0-0.html. Accessed October 14, 2015.
22) Yokohama city (2012) Big city comparison statistics chronological table (H22 National census). http://www.city.yokohama.lg.jp/ex/stat/daitoshi/index2.html#10. Accessed October 14, 2015.
23) Matsudaira K, Shimazu A, Fujii T, Kubota K, Sawada T, Kikuchi N, Takahashi M (2013) Workaholism as a risk factor for depressive mood, disabling back pain, and sickness absence. PLoS One **8**, e75140.
24) Nishikitani M, Tsurugano S, Inoue M, Yano E (2012) Effect of unequal employment status on workers' health: results from a Japanese national survey. Soc Sci Med **75**, 439–51.
25) Vyas MV, Garg AX, Iansavichus AV, Costella J, Donner A, Laugsand LE, Janszky I, Mrkobrada M, Parraga G, Hackam DG. (2012). Shift work and vascular events: systematic review and meta-analysis BMJ 345:e4800
26) Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K (1998) Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. J Clin Epidemiol **51**, 1037–44.
27) Fukuhara S, Ware JEJ, Kosinski M, Wada S, Gandek B (1998) Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. J Clin Epidemiol **51**, 1045–53.
28) Ware JEJ, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care **30**, 473–83.
29) Ware JEJ, Gandek B, Kosinski M, Aaronson NK, Apolone G, Brazier J, Bullinger M, Kaasa S, Leplège A, Prieto L, Sullivan M, Thunedborg K (1998) The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in ten countries: results from the IQOLA Project. J Clin Epidemiol **51**, 1167–70.
30) Shimazu A, Schaufeli WB, Kosugi S, Suzuki A, Nashiwa H, Kato A, Sakamoto M, Irimajiri H, Amano S, Hirohata K, Goto R, Kitaoka-Higashiguchi K (2008) Work engagement in Japan: validation of the Japanese version of Utrecht work engagement scale. Appl Psychol **57**, 510–23.
31) Shimomitsu T, Yokoyama K, Ono Y, Maruta T, Tanigawa T(1998) Development of a novel brief job stress questionnaire. In: Report of the research grant for the prevention of work-related diseases from the Ministry of Labour, Kato S (Ed.), 107–115, Ministry of Labour, Tokyo (in Japanese).
32) Arbuckle JL (2013) Amos (Version 22.0) [Computer Program]. SPSS, Chicago.
33) Mathieu JE, Tannenbaum SI, Salas E (1992) Influences of individual and situational characteristics on measures of training effectiveness. Acad Manage J **35**, 828–47.
34) Cortina JM, Chen G, Dunlap WP (2001) Testing interaction effects in LISREL: Examination and illustration of available procedures. Organ Res Methods **4**, 324–60.
35) Byrne BM (2001) Structural equation modeling with AMOS: basic concepts, applications, and programming. Lawrence Erlbaum Associates, Inc., New Jersey.
36) Edwards JR (2008) To prosper, organizational psychology should ... overcome methodological barriers to progress. J Organ Behav **29**, 469–91.
37) Spector PE (2006) Method variance in organizational research truth or urban legend? Organ Res Methods **9**, 221–32.
38) Spector PE, Zapf D, Chen PY, Frese M (2000) Why negative affectivity should not be controlled in job stress research: don't throw out the baby with the bath water. J Organ Behav **21**, 79–95.
39) Binnewies C, Sonnentag S, Mojza EJ (2010) Recovery during the weekend and fluctuations in weekly job performance: A week-level study examining intra-individual relationships. J Occup Organ Psychol **83**, 419–41.
40) Sonnentag S, Bayer UV (2005) Switching off mentally: predictors and consequences of psychological detachment from work during off-job time. J Occup Health Psychol **10**, 393–414.
41) Sonnentag S, Binnewies C, Mojza EJ (2008) "Did you have a nice evening?" A day-level study on recovery experiences, sleep, and affect. J Appl Psychol **93**, 674–84.
42) ten Brummelhuis LL, Bakker AB (2012) Staying engaged during the week: the effect of off-job activities on next day work engagement. J Occup Health Psychol **17**, 445–55.
43) Smith MA, Leigh B (1997) Virtual subjects: using the internet as an alternative source of subjects and research environment. Behav Res Methods Instrum Comput **29**, 496–505.
44) Oshio T, Kobayashi M (2010) Income inequality, perceived happiness, and self-rated health: evidence from nationwide surveys in Japan. Soc Sci Med **70**, 1358–66.
45) Sonnentag S, Fritz C (2015) Recovery from job stress: the stressor-detachment model as an integrative framework. J Organ Behav **36**, S72–103.
46) Derks D, Bakker A (2014) Smartphone use, work-home interference, and burnout: A diary study on the role of recovery. Appl Psychol **63**, 411–40.
47) Derks D, ten Brummelhuis LL, Zecic D, Bakker AB (2014) Switching on and off…: Does smartphone use obstruct the

possibility to engage in recovery activities? Eur J Work Organ Psychol **23**, 80–90.
48) Derks D, van Mierlo H, Schmitz EB (2014) A diary study on work-related smartphone use, psychological detachment and exhaustion: examining the role of the perceived segmentation norm. J Occup Health Psychol **19**, 74–84.
49) Japan Institute for Labour Policy and Training (2010) Research findings on work characteristics, personal characteristics, and working time. http://www.jil.go.jp/press/documents/20101207.pdf (in Japanese). Accessed April 2, 2015.
50) Schaufeli WB, Shimazu A, Taris TW (2009) Being driven to work excessively hard: the evaluation of a two-factor measure of workaholism in the Netherlands and Japan. Cross-Cultural Res **43**, 320–48.
51) van Wijhe C, Schaufeli WB, Peeters MCW (2010) Understanding and treating workaholism: setting the stage for successful interventions. In: Risky business: Psychological, physical and financial costs of high risk behavior in organizations, Burke RJ, Cooper CL (Eds.), 107–34, Farnham, England: Gower.
52) Ellis A (1995) Changing rational-emotive therapy (RET) to rational emotive behavior therapy (REBT). J Ration-Emot Cogn-Behav Ther **13**, 85–9.
53) Ng TWH, Sorensen KL, Feldman DC (2007) Dimensions, antecedents, and consequences of workaholism: a conceptual integration and extension. J Organ Behav **28**, 111–36.
54) Binnewies C, Sonnentag S, Mojza EJ (2009) Feeling recovered and thinking about the good sides of one's work. J Occup Health Psychol **14**, 243–56.

Industrial Health 2016, **54**, 230–236

Original Article

Potential risk factors for onset of severe neck and shoulder discomfort (Katakori) in urban Japanese workers

Takayuki SAWADA[1, 2]*, Ko MATSUDAIRA[3], Yumiko MUTO[1], Tadashi KOGA[1, 4] and Masaya TAKAHASHI[5]

[1]Clinical Study Support, Inc., Japan
[2]Department of Public Health, Aichi Medical University, Japan
[3]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, the University of Tokyo, Japan
[4]CPC Clinical Trial Hospital, Medipolis Medical Research Institute, Japan
[5]National Institute of Occupational Safety and Health, Japan

Received July 28, 2015 and accepted December 18, 2015
Published online in J-STAGE January 30, 2016

Abstract: Katakori is a Japanese word, and there is no clear English translation. Katakori consists of two terms, Kata means neck and shoulder, kori means stiffness. Consequently, Katakori is defined as neck and shoulder discomfort or dull pain. Katakori is a major somatic complaint and has a large impact on workers. To examine the association between onset of severe Katakori and potential risk factors in Japanese workers, a prospective cohort study, entitled "Cultural and Psychosocial Influence on Disability (CUPID)", was conducted. Self-administered questionnaires were distributed twice: at baseline and 1 year after baseline. Logistic regression was used to explore the risk factors of onset of severe Katakori. Of those 1,398, the incidence of severe Katakori onset after 1 year was 3.0% (42 workers). Being female (adjusted odds ratio: 2.39, 95% confidence interval: 1.18–4.86), short sleep duration (adjusted odds ratio: 2.86, 95% confidence interval: 1.20–6.82) and depressed mood with some issues at work (adjusted odds ratio: 3.11, 95% confidence interval: 1.38–7.03) were significantly associated with onset of severe Katakori. Psychosocial factors as well as gender difference were associated with onset of severe Katakori. We suggest that mental health support at the workplace is important to prevent severe Katakori.

Key words: Katakori, Prospective study, Risk factors, Japanese workers, Psychosocial factors

Introduction

Katakori is a Japanese word, and there is no clear English translation. Katakori consists of two terms, Kata means shoulder and kori means stiffness. Consequently, Katakori is defined as discomfort or dull pain caused by muscle stiffness around the back of the head and through the shoulders and/or shoulder blades[1)]. Katakori is usually classified as one of the cervico-omo-brachial syndrome. The symptoms of Katakori are considered to be close to "neck pain" or "chronic nonspecific neck pain" as expressed in the references[2–4)].

Katakori is classified into primary Katakori (essential Katakori) which does not identify any causable disease (organic disorder) and secondary Katakori (symptomatic Katakori) which is caused by disease. Examples of disease

*To whom correspondence should be addressed.
E-mail: takayuki_sawada@jp-css.com

which can be the cause of secondary Katakori include cervical spine disease, glenohumeral joint disease, cardiovascular disease, pulmonary disease, eye fatigue, temporomandibular arthrosis, and menopausal syndrome[5, 6].

The prevalence of Katakori is 6.1% among males and 13.1% among females in Japan[7], therefore Katakori is a major somatic complaint which is comparable to low back pain and has a large impact on people including workers with subjective symptoms, however, its pathogenesis is still unclear. Furthermore, the association between Katakori and potential risk factors has not been properly assessed in prospective epidemiological research.

There have been reports of several risk factors associated with Katakori: such being female[6–9], using a Visual Display Terminal (VDT)[6] and mental health[9, 10]. These factors have been identified based on the results of cross-sectional studies.

A prospective cohort study, entitled "Cultural and Psychosocial Influence on Disability (CUPID)", was conducted to explore further the impact of cultural and psychosocial influences on musculoskeletal symptoms and associated disability[11, 12]. A cross-sectional analysis of baseline data shows that being female and depressed mood have been associated with severe Katakori in urban Japanese workers[8]. In this study, using one year of follow-up data, we conducted a continued analysis to examine the association between onset of severe Katakori and potential risk factors in urban Japanese workers. To our knowledge, this was the first longitudinal study assessing the potential risk factors for onset of severe Katakori. In this study, we especially focused on severe Katakori since Katakori is a common symptom among Japanese workers.

Subjects and Methods

Data from a 1-year prospective cohort of the CUPID study were used for this analysis. The CUPID study is an international joint research project, which has involved 18 countries. In Japan, ethical approval for the study was obtained from the ethics committees of the University of Tokyo Hospital and review board of the Japan Labour Health and Welfare Organization. All participants provided written informed consent.

The workers around Tokyo including office workers, sales and marketing personnel, transportation workers, and nurses were recruited.

The board of each participating organization was asked to distribute a self-administrated questionnaire along with a cover letter from the study administration office to their workers. Responders were asked to return their completed questionnaires by mail and to provide their names and mailing addresses for direct correspondence from the study administration office for 1-year follow-up purposes.

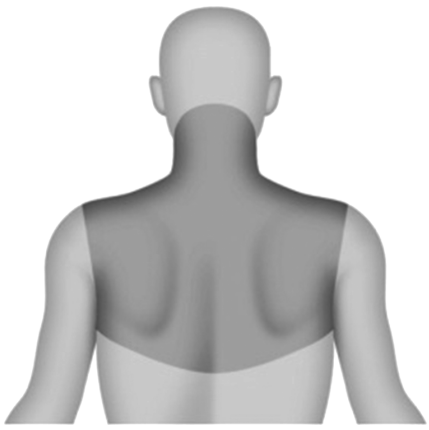


Fig. 1. Diagram showing pain area for Katakori.

The original questionnaire used in the CUPID study was translated into Japanese with some newly designed questions for Japanese workers regarding Katakori. The translation equivalence with the original questionnaire was checked through independent back-translation into English. For the participants, the pain area of Katakori was defined as the back of the head and through the shoulders and/or shoulder blades (Fig. 1). At baseline, respondents were asked about the frequency and severity of Katakori they had experienced during the previous month. The frequency of Katakori was assessed on a 6-point scale (1, always; 2, almost always; 3, often; 4, sometimes; 5, seldom; 6, never); the severity of Katakori was measured on an 11-point numerical rating scale (NRS) ranging from 0 (no Katakori) to 10 (severe Katakori). At follow-up, the frequency of Katakori was assessed using three duration periods (1–6 days, 1–2 weeks, or ≥2 weeks) and the severity of Katakori was measured by NRS.

In addition, the baseline questionnaire assessed individual characteristics (i.e., age, gender, age at the last educational status, body mass index (BMI), hours of sleep, marital status, regular exercise, smoking habits, visual fatigue, dental therapy, dental bite, and outpatient with articular and spine symptoms), ergonomic work demands (period of current service, working hours per week, VDT use, finger repetition, lifting, driving, standing, and work shift), and psychosocial factors (job satisfaction, job control, inadequate break time at work, worksite support, interpersonal stress at work, and experience of depressed mood with an issue at

work). Variables were categorized by the same methods previously used in the CUPID study for Katakori association[8]. Age was categorized as <30 years, 30–39 years, 40–49 years or ≥50 years. BMI was calculated by height and body weight recorded in a questionnaire; BMI ≥25 was defined as obesity. Age at the last educational status was categorized as ≤19 years or >19 years; low education was defined as ≤19 years. Regular exercise was defined as physical exercise performed more than twice a week for 20 minutes or longer during the previous 12 months. Short sleep duration was defined as an average of <5 hours. Low experience in current job was defined as <1 year of current service. Sixty hours of working hours per week was defined as high work demand. VDT was defined as work using the computer display for ≥4 hours per shift. Lifting was defined as a work to lift or move ≥25 kg (object or person) by hand. Driving was defined as ≥4 hours of car or truck driving per shift. Standing was defined as ≥4 hours standing per shift. Work shift was defined as irregular work shift such as night shift. To assess the level of job satisfaction, responders were asked, "Considering everything, how satisfied are you with your work?" Answers were the following four choices: "Very satisfied", "Satisfied", "Not well satisfied" and "Not satisfied at all". Low job satisfaction was defined as an answer of "Not well satisfied" or "Not satisfied at all". To assess the level of job control, responders were asked, "How much control do you have in your work?" These items had four response options: often, sometimes, seldom, and never/almost never. Low job control was defined as an answer of "seldom" or "never/almost never". To assess the level of worksite support, responders were asked, "When you have difficulties in your work, how often do you get help and support from your colleagues or supervisor/manager?" This item had five response options: often, sometimes, seldom, never, and not applicable. Low worksite support was defined as an answer of "seldom" or "never" for worksite support. Depressed mood with some issues at work was defined as experience of that in past 12 months.

The follow-up questionnaire was distributed 1 year after the baseline assessment, and the second questionnaire was sent only to the participants who returned the first one with their written consent of participating. Therefore, those who did not return a questionnaire did not participate in the study any longer.

The outcome of interest was onset of severe Katakori during the 1-year follow-up period. In this study, severe Katakori was defined as frequency more than 2 weeks in the previous month and as severity with NRS more than 7 points at the follow-up. Incidence was calculated for the participants who reported no severe Katakori at baseline, as we defined severe Katakori as frequency more than often and as severity with NRS more than 7 points during the previous month. Participants were excluded from the analysis if they had changed their job.

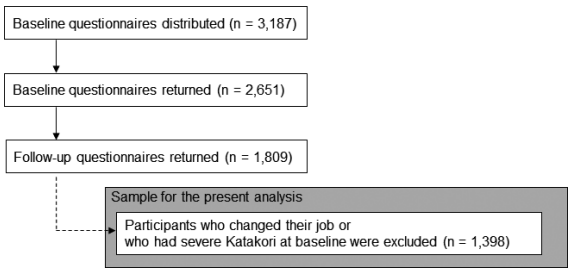


Fig. 2. Flowchart of the sample selection.

For statistical analysis, in addition to compiling descriptive statistics, logistic regression was used to explore the associations between risk factors and onset of severe Katakori. Results of logistic regression analyses were summarized by odds ratios (ORs) and the respective 95% confidence intervals (CIs). For the assessment of potential risk factors, crude ORs were initially estimated. Factors with p-values <0.1 were considered to be potential risk factors. We conducted a multivariate logistic regression analysis using potential risk factors in the model and then using a stepwise selection method in which terms were retained if they reached the 0.05 level of significance. All statistical tests were two-tailed, and conducted with a significance level of 0.05. The software package SAS Release 9.3 (SAS Institute Inc., Cary, NC) was used for statistical analyses.

Results

The baseline questionnaire was distributed to 3,187 participants and was completed by 2,651 participants. The following year, 1,809 participants successfully completed and returned the follow-up questionnaire, thereby yielding a follow-up rate of 68.2%.

Participants (n=411) were excluded from the analysis if they had severe Katakori at baseline (n=330) or those who changed their job (n=81). Thus, a total of 1,398 participants were included in the present analysis (Fig. 2).

Mean (SD: standard deviation) age was 37.3 (10.0) years, of which 1,398 of 73.8% of participants were male. Jobs were nurses (21%), office workers (15%), sales and marketing personnel (21%) and transportation operators (43%). [Table 1] The incidence of onset of severe Katakori

Table 1. Characteristics of responders

Characteristics	Severe Katakori	Non-Severe Katakori	Total
N (%)	42 (3.0%)	1,356 (97.0%)	1,398
Gender			
Male, n (%)	21 (2.0%)	1,011 (98.0%)	1,032 (73.8%)
Female, n (%)	21 (5.7%)	345 (94.2%)	366 (26.2%)
Age, mean (SD)	37.1 (9.0)	37.3 (10.0)	37.3 (10.0)
Job type			
Transportation operative	15 (2.5%)	585 (97.5%)	600 (43.0%)
Sales/marketing personnel	5 (1.7%)	289 (98.3%)	294 (21.0%)
Nurse	16 (5.4%)	278 (94.6%)	294 (21.0%)
Office workers	6 (2.8%)	204 (96.7%)	211 (15.1%)

Table 2. Crude odds ratios of the risk factors for onset of severe Katakori

Risk factors	%	Crude odds ratio (95%CI)	*p* value
Gender			
Male	73.8	1.00	
Female	26.2	2.92 (1.58–5.42)	0.001
Age (yr)			
<30	25.5	1.00	
30–39	37.3	1.79 (0.74–4.33)	0.197
40–49	22.6	1.64 (0.62–4.35)	0.324
≥50	14.6	1.51 (0.50–4.57)	0.462
Outpatient with articular and spine			
No	97.2	1.00	
Yes	2.8	0.82 (0.11–6.14)	0.850
Outpatient with dental therapy			
No	92.7	1.00	
Yes	7.3	1.35 (0.47–3.87)	0.537
Wrong dental bite			
No	83.8	1.00	
Yes	16.2	1.76 (0.85–3.65)	0.130
Visual fatigue			
No	56.3	1.00	
Yes	43.7	2.20 (1.15–4.21)	0.017
BMI			
<25 kg/m^2	84.0	1.00	
≥25 kg/m^2	16.0	1.50 (0.71–3.19)	0.291
Current smoking			
No	56.4	1.00	
Yew	43.6	1.44 (0.78–2.66)	0.245
Age at last educational status **(yr)**			
≥20	62.4	1.00	
<19	37.6	0.66 (0.33–1.29)	0.221
Regular exercise			
Yes	20.2	1.00	
No	79.8	1.50 (0.62–3.60)	0.367
Marital status			
Married	56.4	1.00	
Not married	43.3	1.20 (0.65–2.21)	0.568
Sleep duration			
≥5 h	56.4	1.00	
<5 h	43.3	2.75 (1.24–6.10)	0.013
Experience in current job			
≥1 yr	90.6	1.00	
<1 yr	9.4	1.32 (0.51–3.42)	0.569
Working hours per week			
Low	59.2	1.00	
High	40.8	0.89 (0.47–1.67)	0.715
Inadequate break time at work			
Not inadequate	45.6	1.00	
Inadequate	54.4	3.16 (1.50–6.66)	0.003
VDT			
Not VDT	75.3	1.00	
VDT	24.7	1.23 (0.62–2.42)	0.557
Finger repetition			
No	77.7	1.00	
Yes	22.3	1.09 (0.53–2.25)	0.811
Lifting			
No	47.4	1.00	
Yes	52.6	1.09 (0.59–2.03)	0.777
Driving			
No	64.5	1.00	
Yes	35.5	1.01 (0.53–1.91)	0.980
Standing			
No	43.1	1.00	
Yes	56.9	1.93 (0.98–3.80)	0.058
Work shift			
Regular shift	60.8	1.00	
Irregular shift	39.2	1.73 (0.94–3.21)	0.058
Job satisfaction			
Satisfied	43.4	1.00	
Not satisfied	56.6	1.38 (0.74–2.57)	0.310
Job control			
Controlled	46.4	1.00	
Not controlled	53.6	0.64 (0.35–1.19)	0.528
Worksite support			
Supported	91.3	1.00	
Not supported	8.7	1.15 (0.40–3.27)	0.800
Interpersonal stress at work			
Not stressed	51.2	1.00	
Stressed	48.8	1.93 (1.02–3.66)	0.045
Depressed mood with some issue at work			
Not feeling depressed	50.0	1.00	
Depressed	50.0	4.15 (1.89–9.07)	<0.001

CI: confidence interval.

Table 3. Adjusted odds ratios of risk factors which were significant for onset of severe Katakori

Risk factor	Adjusted odds ratio (95%CI)	*p* value
Gender		
Male	1.00	
Female	2.39 (1.18–4.86)	0.016
Sleep duration		
≥5 h	1.00	
<5 h	2.86 (1.20–6.82)	0.018
Depressed mood with some issue at work		
Not feeling depressed	1.00	
Depressed	3.11 (1.38–7.03)	0.006

CI: confidence interval.
Adjusted by gender, sleep duration and experience of depressed mood with some issue at work

in the follow-up period was 3.0% (42 workers), with mean (SD) age of 37.1 (9.0) years. Of those, 50% were males.

To assess the effect of the selected drop-out, the baseline characteristics of patients who were followed up (n=1,809) and those who dropped-out (n=842) are calculated. The mean (SD) age was 37.3 (10.0) years and 33.6 (8.5) years, respectively, and the majority were men in both groups (66.0% vs 57.7%). The prevalence of severe Katakori was 18.8% and 21.2%, respectively.

Crude odds ratios of baseline factors for onset of severe Katakori are shown in Table 2. The factors potentially relating to onset of severe Katakori were gender, visual fatigue, sleep duration, inadequate break time, standing, work shift, interpersonal stress and depressed mood with some issues at work. In psychosocial factors, depressed mood with some issues at work was only included, instead of interpersonal stress at work, because of its strong correlation (ρ= 0.4137, $p<0.0001$). The crude odds ratio of depressed mood with some issues at work was higher than the interpersonal stress at work, thus the higher factor was selected. Because 77% (281/366) of females were nurses, and 87% (255/294) of nurses were defined as irregular work shift, the correlation between female and irregular work shift was strong ($\rho=0.3422$, $p<0.0001$). Previous studies reported that Katakori was associated with females, so "female" was included in multivariate logistic regression analysis.

In the multivariate logistic regression analysis, these six factors were entered into the model. As a result, three potential risk factors were selected (Table 3).

A supplemental analysis was conducted to examine a combined impact of gender and nurses because 77% (281/366) were female nurses. We performed multivariate logistic regression analysis with the main three effects, nurse and interaction of gender and nurse. The adjusted odds ratios of main effects were similar to the main analysis, and the nurse effect as well as the interaction were not statistically significant. Based on these results, we propose three potential risk factors: gender, short sleep duration, and depressed mood with some issues at work which might associate with severe Katakori.

Discussion

To examine the association between onset of severe Katakori and potential risk factors, we conducted analyses using data from the CUPID study among urban workers in Japan. Although the incidence was small, severe Katakori occurred during the 1-year follow-up in some workers who had no or mild symptoms at baseline. A series of analyses showed gender, low sleep or depressed mood with some issues at work as important potential risk factors.

In our results, females showed higher odds (adjusted odds ratio=2.18) as a potential risk factor for onset of severe Katakori. According to the supplemental analysis, being female is potential risk factor of Katakori as it eliminates the possibility of nurses to affect the main result of this study. Based on these results, this study suggests the association of gender as a potential risk factor of severe Katakori. This finding is similar to those published previously[6, 8, 9]. We speculate this trend may be attributable to gender differences in muscle strength. Estrogen may also be involved in the pathogenesis of Katakori, although there is no scientific evidence for this assertion. Further studies will be required to explain the reason for gender differences in the manifestation of Katakori.

Being in a depressed frame of mind with some issues at work showed 3.1 times more increased risk of severe Katakori than those who are not. Previous cross-sectional studies suggest the association of Katakori and work stress, which was classified as a psychosocial factor[1, 6]. Krantz *et al.* have reported that emotional stress and psychologically stressful tasks are associated with increased electrographic activity in the trapezius muscle[13], and Hallman *et al.* have reported that autonomic imbalance is associated with neck shoulder pain, the Japanese definition of Katakori[14]. We suggest that psychosocial stress can progress to sympathetic and muscle stress, which may lead to the onset of Katakori.

In the present study, we found short sleep duration to be a potential risk factor. Mulligan *et al.* reported that nocturnal pain was associated with sleep quality, sleep duration, and habitual sleep efficiency in patients with shoulder dis-

orders[15]. Short sleep duration might delay a daily recovery of tissue damage and cause the onset of severe Katakori. In order to ensure an adequate sleep duration, individuals should be responsible in attaining the required sleep duration, and support can be provided by encouraging a non-stressful work environment. In the present study, we had assessed sleep duration only. Further studies are required to explore any association between Katakori and the quality of sleep, including insomnia and other sleep disorders.

Factors identified as potential risk factors in the present study can be explained by Eriksen's hypothesis that head-down and neck flexion positions and/or psychological stress increase the intracellular nitric oxide/oxygen ratio through sympathetic nerve activity, resulting in inhibition of cytochrome oxidase; and then, lactate production would follow activating nociceptive fibers[16].

There were some limitations in this study. First, the generalizability of the findings may be limited. The majority of participants were male, and therefore a broad range of Japanese occupations was not represented. The study cohort was not a representative sample of the entire spectrum of Japanese workers in urban areas. Being female was one of the potential risk factors of Katakori although no interaction effects of gender and nurse were found in our supplemental analysis. However, the majority of females in this study were nurses, and the sample size included in the supplemental analysis may not have been sufficient. Therefore, our results need to be interpreted with care. Second, misclassification, to some extent, is inevitable. Information might be subjective in the decision of Katakori or sicknesses and missing value cannot be avoided due to the nature of a self-assessment survey. Third, drop-out may introduce bias into the data analysis due to the low follow-up rate of this study, although we considered that the baseline characteristics of both the follow-up group and the drop-out group seemed to be similar. Fourth, this study may not cover some unquestioned items which were not involved in the questionnaire. For example, some peculiar characteristics of Japanese may not be addressed by the original CUPID questionnaire regarding stress at work. Also, there were some items which were not involved in the original CUPID questionnaire as follows: disabilities of the arm, shoulder and hand questionnaire scores correlated significantly with depressive symptoms, catastrophic thinking, kinesiophobia, and pain anxiety[17]. The aforementioned behavioral items may need to be included as additional potential risk factors of severe Katakori. At last, a more complicated analysis model might be suitable for further assessment to discover other potential risk factors, instead of the logistic regression models assessed for the present analysis.

In conclusion, being female, short sleep duration and depressed mood with some issues at work were associated with onset of severe Katakori. We suggest that mental health support including the lack of sleep is important to prevent severe Katakori, especially for females.

Acknowledgements

We thank Dr. David Coggon and Keith T. Palmer for organizing and leading the CUPID study; CUPID collaborators for all their dedications, and Dr. Noriko Yoshimura for data collection in Japan. The study was a part of clinical research projects conducted by the Japan Labour Health and Welfare Organization.

References

1) Yabuki S (2007) Pathogenesis of the Neck-shoulder Stiffness (Katakori). Rinsho Seikei Geka (Clinical Orthopedic Surgery) **42**, 413–7 (in Japanese).
2) Ijmker S, Huysmans MA, van der Beek AJ, Knol DL, van Mechelen W, Bongers PM, Blatter BM (2011) Software-recorded and self-reported duration of computer use in relation to the onset of severe arm-wrist-hand pain and neck-shoulder pain. Occup Environ Med **68**, 502–9.
3) McLean SM, May S, Klaber-Moffett J, Sharp DM, Gardiner E (2010) Risk factors for the onset of non-specific neck pain: a systematic review. J Epidemiol Community Health **64**, 565–72.
4) Kadi F, Waling K, Ahlgren C, Sundelin G, Holmner S, Butler-Browne GS, Thornell LE (1998) Pathological mechanisms implicated in localized female trapezius myalgia. Pain **78**, 191–6.
5) Takagishi K, Hoshino Y, Ide J, Sugihara T, Hata Y, Sano H, Hamada J, Yabuki S, Mochiduki Y, Suzuki K, Yanagawa T, Tamai K, Ogawa K, Atsuta Y, Shinozaki T (2008) A project study for Katakori, 2004–2006 (commentary). Nihon Seikeigeka Gakkai Zasshi (The Journal of the Japanese Orthopaedic Association) **82**, 901–11 (in Japanese).
6) Fujii T, Matsudaira K, Noma K, Isizuka A, Yamada K, Arisaka M, Higashikawa A (2012) Objective Measurement of Neck-shoulder Discomfort and Analysis of Associated Factors. Rinsho Seikei Geka (Clinical Orthopaedic Surgery) **47**, 619–24 (in Japanese).
7) Japan Ministry of Health, Labour and Welfare. Comprehensive Survey of Living Conditions 2013. http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa13/dl/04.pdf. (In Japanese) Accessed Jul 16, 2015.
8) Fujii T, Matsudaira K, Yoshimura N, Hirai M, Tanaka S (2013) Associations between neck and shoulder discomfort (Katakori) and job demand, job control, and worksite sup-

port. Mod Rheumatol **23**, 1198–204.
9) Kimura T, Tsuda Y, Uchida S, Eboshida A (2006) Association of perceived stress and stiff neck/shoulder with health status: multiple regression models by gender. Hiroshima J Med Sci **55**, 101–7.
10) Iizuka Y, Shinozaki T, Kobayashi T, Tsutsumi S, Osawa T, Ara T, Iizuka H, Takagishi K (2012) Characteristics of neck and shoulder pain (called katakori in Japanese) among members of the nursing staff. J Orthop Sci **17**, 46–50.
11) Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Serra C, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir M, Warnakulasuriya SS, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Marziale MH, Sarquis LM, Harari F, Freire R, Harari N, Monroy MV, Quintana LA, Rojas M, Salazar Vega EJ, Harris EC, Vargas-Prada S, Martinez JM, Delclos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Sirk T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kielkowski D, Kelsall HL, Hoe VC, Urquhart DM, Derrett S, McBride D, Gray A (2012) The CUPID (Cultural and Psychosocial Influences on Disability) study: methods of data collection and characteristics of study sample. PLoS One **7**, e39820.
12) Matsudaira K, Palmer KT, Reading I, Hirai M, Yoshimura N, Coggon D (2011) Prevalence and correlates of regional pain and associated disability in Japanese workers. Occup Environ Med **68**, 191–6.
13) Krantz G, Forsman M, Lundberg U (2004) Consistency in physiological stress responses and electromyographic activity during induced stress exposure in women and men. Integr Physiol Behav Sci **39**, 105–18.
14) Hallman DM, Ekman AH, Lyskov E (2014) Changes in physical activity and heart rate variability in chronic neck-shoulder pain: monitoring during work and leisure time. Int Arch Occup Environ Health **87**, 735–44.
15) Mulligan EP, Brunette M, Shirley Z, Khazzam M (2015) Sleep quality and nocturnal pain in patients with shoulder disorders. J Shoulder Elbow Surg, pii: S1058-2746(15)00086-5. doi:10.1016/j.jse.2015.02.013.
16) Eriksen W (2004) Linking work factors to neck myalgia: the nitric oxide/oxygen ratio hypothesis. Med Hypotheses **62**, 721–6.
17) Das De S, Vranceanu AM, Ring DC (2013) Contribution of kinesophobia and catastrophic thinking to upper-extremity-specific disability. J Bone Joint Surg Am **95**, 76–81.

Research Paper

PAIN®

Classification of neck/shoulder pain in epidemiological research: a comparison of personal and occupational characteristics, disability, and prognosis among 12,195 workers from 18 countries

Leila M.M. Sarquis[a,b,c], David Coggon[b,c,*], Georgia Ntani[b,c], Karen Walker-Bone[b,c], Keith T. Palmer[b,c], Vanda E. Felli[d], Raul Harari[e], Lope H. Barrero[f], Sarah A. Felknor[g,h], David Gimeno[g], Anna Cattrell[i], Sergio Vargas-Prada[j,k,l], Matteo Bonzini[m], Eleni Solidaki[n], Eda Merisalu[o], Rima R. Habib[p], Farideh Sadeghian[q], M. Masood Kadir[r], Sudath S.P. Warnakulasuriya[s], Ko Matsudaira[t], Busisiwe Nyantumbu[u,v], Malcolm R. Sim[w], Helen Harcombe[x], Ken Cox[b], Maria H. Marziale[y], Florencia Harari[e], Rocio Freire[e], Natalia Harari[e], Magda V. Monroy[f], Leonardo A. Quintana[f], Marianela Rojas[z], E. Clare Harris[b,c], Consol Serra[j,k,l,aa], J. Miguel Martinez[bb], George Delclos[g,j,k,l], Fernando G. Benavides[j,k,l], Michele Carugno[cc], Marco M. Ferrario[m], Angela C. Pesatori[cc,dd], Leda Chatzi[n], Panos Bitsios[ee], Manolis Kogevinas[k,l,ff], Kristel Oha[gg], Tiina Freimann[hh], Ali Sadeghian[ii], Roshini J. Peiris-John[jj,kk], Nalini Sathiakumar[ll], A. Rajitha Wickremasinghe[mm], Noriko Yoshimura[nn], Helen L. Kelsall[w], Victor C.W. Hoe[oo], Donna M. Urquhart[w], Sarah Derrett[pp], David McBride[x], Peter Herbison[x], Andrew Gray[x], Eduardo J. Salazar Vega[qq]

Abstract

To inform case definition for neck/shoulder pain in epidemiological research, we compared levels of disability, patterns of association, and prognosis for pain that was limited to the neck or shoulders (LNSP) and more generalised musculoskeletal pain that involved the neck or shoulder(s) (GPNS). Baseline data on musculoskeletal pain, disability, and potential correlates were collected by questionnaire from 12,195 workers in 47 occupational groups (mostly office workers, nurses, and manual workers) in 18 countries (response rate = 70%). Continuing pain after a mean interval of 14 months was ascertained through a follow-up questionnaire in 9150 workers from 45 occupational groups. Associations with personal and occupational factors were assessed by Poisson regression and summarised by prevalence rate ratios (PRRs). The 1-month prevalence of GPNS at baseline was much greater than that of LNSP (35.1% vs 5.6%), and it tended to be more troublesome and disabling. Unlike LNSP, the prevalence of GPNS increased with age. Moreover, it showed significantly stronger associations with somatising tendency (PRR 1.6 vs 1.3) and poor mental health (PRR 1.3 vs 1.1); greater variation between the occupational groups studied (prevalence ranging from 0% to 67.6%) that correlated poorly with the variation in LNSP; and was more persistent at follow-up (72.1% vs 61.7%). Our findings highlight important epidemiological distinctions between subcategories of neck/shoulder pain. In future epidemiological research that bases case definitions on symptoms, it would be useful to distinguish pain that is localised to the neck or shoulder from more generalised pain that happens to involve the neck/shoulder region.

Keywords: Neck pain, Shoulder pain, Diagnostic classification, Case definition, Disability, Associations, Prognosis

1. Introduction

Pain in the neck and/or shoulder(s) is a common problem in people of working age and an important cause of disability. Like other regional pain, it may arise from identifiable musculoskeletal pathology—for example, cervical spondylitis or subacromial bursitis. However, the relationship of such abnormalities to symptoms is imperfect,[17] and their occurrence in association with pain does not necessarily imply that they are responsible for it. Furthermore, some modes of investigation that might be used to detect relevant pathology, in particular, magnetic resonance imaging, are relatively expensive and not readily applicable in large surveys. Most epidemiological studies of neck and shoulder pain have therefore defined cases by the occurrence and characteristics of symptoms. Moreover, because neck/shoulder pain can be difficult for patients to localise precisely and commonly extends across both the neck and shoulders, it has often been treated as a single diagnostic entity. Using this approach, it has been linked with occupational physical activities such as manual materials handling,[20] awkward postures,[2,20,24] and use of computers[28]; psychological factors such as low mood[8] and tendency to somatise[15]; and various psychosocial aspects of work.[16,21,24]

The merits of aggregating all pain in the neck and shoulder region will depend on whether there are identifiable subsets of cases that differ importantly in their causes, prognosis, or response to treatment.[4] For example, pain in the neck/shoulder is often accompanied by pain at other anatomical sites.[1,13,14,25] Pain that is localised only to the neck or shoulder might be more reflective of local pathology, whereas some psychological factors might show stronger associations with pain, which, although involving the neck or shoulder, is more widespread. If such

distinctions occur, then associations with causes and effects of treatment may be diluted when all cases are aggregated.

The Cultural and Psychological Influences on Disability (CUPID) study is a large international investigation in which information about musculoskeletal pain, associated disability, and potential risk factors was collected from workers employed in 47 occupational groups distributed across 18 countries.[5] To explore whether there are differences between subcategories of neck/shoulder pain in associations and/or prognosis, we analysed data from the CUPID study, focusing in particular on pain that was limited to the neck or shoulder as compared with more generalised musculoskeletal pain that involved the neck or shoulder but also affected other anatomical sites. As well as comparing associations with demographic, physical, psychological, and psychosocial factors, we examined differences between the subcategories of pain in their relative prevalence by occupational group.

2. Methods

During 2006 to 2011, baseline information was collected by questionnaire from 47 occupational groups in 18 countries. Participants were aged 20 to 59 years and had been employed in their current job for at least 12 months. The occupational groups fell into 3 broad categories—nurses, office workers, and other workers (mostly performing manual tasks with their arms). In most groups, potentially eligible subjects were identified from employers' records, in some cases with random sampling to achieve the desired sample size (at least 200 per group if possible). The number of participants by group (mean response rate = 70%, response rate >80% in 33 of the 47 groups) varied from 92 to 1018, giving a total sample size of 12,426.

The questionnaire, which was completed either by self-administration or at interview, was originally drafted in English and then translated into local languages where necessary, with independent back-translation to ensure accuracy. Among other things, it covered demographic characteristics (sex and age), occupational activities (in an average working day), psychosocial aspects of work, somatising tendency, mental health, beliefs about arm pain, and experience of musculoskeletal pain and associated disability.

Somatising tendency was assessed through questions taken from the Brief Symptom Inventory[9] and was classified according to the number of symptoms from a total of five (faintness or dizziness, pains in the heart or chest, nausea or upset stomach, trouble getting breath, and hot or cold spells) that had been at least moderately distressing during the past week. Mental health was ascertained using elements from the Short Form-36 questionnaire[29] and was graded to 3 levels (good, intermediate, and poor) representing approximate thirds of the distribution of scores in the full study sample. Questions on beliefs about arm pain were adapted from the Fear Avoidance Beliefs Questionnaire.[27] Participants were classed as having adverse beliefs about the work-relatedness of arm pain if they completely agreed that such pain is commonly caused by work; about the impact of physical activity if they completely agreed that for someone with such pain, physical activity should be avoided as it might cause harm, and that rest is needed to get better; and about its prognosis if they completely agreed that neglecting such problems can cause serious harm and completely disagreed that such problems usually get better within 3 months.

The questions about musculoskeletal pain concerned 10 anatomical sites (low back; neck; and left and right shoulder, elbow, wrist/hand, and knee), which were illustrated with body mannequins. For each site, participants were asked whether they had experienced pain that had lasted for longer than a day (1) during the past 12 months and (2) during the past month. Pain in

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Federal University of Paraná, Curitiba-PR, Brazil,* [b] *Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, United Kingdom,* [c] *Arthritis Research UK/MRC Centre for Musculoskeletal Health and Work, University of Southampton, Southampton, United Kingdom,* [d] *School of Nursing, University of São Paulo, São Paulo, Brazil,* [e] *Corporación para el Desarrollo de la Producción y el Medio Ambiente Laboral—IFA (Institute for the Development of Production and the Work Environment), Quito, Ecuador,* [f] *Department of Industrial Engineering, School of Engineering, Pontificia Universidad Javeriana, Bogotá, Colombia,* [g] *Southwest Center for Occupational and Environmental Health, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA,* [h] *Center for Disease Control and Prevention/National Institute for Occupational Safety and Health, Atlanta, GA, USA,* [i] *Medical Research Council Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry Psychology and Neuroscience, Kings College London, London, United Kingdom,* [j] *Center for Research in Occupational Health (CiSAL), Universitat Pompeu Fabra, Barcelona, Spain,* [k] *CIBER of Epidemiology and Public Health, Barcelona, Spain,* [l] *IMIM (Hospital del Mar Research Institute), Barcelona, Spain,* [m] *Epidemiology and Preventive Medicine Research Center, University of Insubria, Varese, Italy,* [n] *Department of Social Medicine, Medical School, University of Crete, Heraklion, Greece,* [o] *Institute of Technology, Estonian University of Life Sciences, Tartu, Estonia,* [p] *Department of Environmental Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon,* [q] *Department of Occupational Health, School of Public Health, Shahroud University of Medical Sciences, Shahroud, Iran,* [r] *Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan,* [s] *Department of Medical Education and Health Sciences, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka,* [t] *Department for Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo Hospital, Tokyo, Japan,* [u] *National Institute for Occupational Health, National Health Laboratory Service, Johannesburg, South Africa,* [v] *Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa,* [w] *Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia,* [x] *Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand,* [y] *School of Nursing of Ribeirão Preto, University of São Paulo, São Paulo, Brazil,* [z] *Program Health, Work and Environment in Central America, Institute for Studies on Toxic Substances (IRET), National University of Costa Rica, Heredia, Costa Rica,* [aa] *Occupational Health Service, Parc de Salut MAR, Barcelona, Spain,* [bb] *Servicio de Investigación y Análisis IT/EP, Departamento de Investigación y Análisis de Prestaciones, MC Mutual, Barcelona, Spain,* [cc] *Department of Clinical Sciences and Community Health, Università degli Studi di Milano, Milan, Italy,* [dd] *Fondazione Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy,* [ee] *Department of Psychiatry, Medical School, University of Crete, Heraklion, Greece,* [ff] *Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain,* [gg] *North Estonia Medical Centre, Tallinn, Estonia,* [hh] *Tartu University Hospital, Tartu, Estonia,* [ii] *Klinikum Leverkusen, Leverkusen, Germany,* [jj] *Department of Physiology, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka,* [kk] *Section of Epidemiology and Biostatistics, School of Population Health, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand,* [ll] *Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA,* [mm] *Faculty of Medicine, University of Kelaniya, Kelaniya, Sri Lanka,* [nn] *Department of Joint Disease Research, 22nd Century Medical and Research Center, University of Tokyo, Tokyo, Japan,* [oo] *Centre for Occupational and Environmental Health, Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia,* [pp] *Injury Prevention Research Unit, Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand,* [qq] *AkzoNobel, Houston, TX, USA*

**Corresponding author. Address: MRC Lifecourse Epidemiology Unit, Southampton General Hospital, Southampton SO16 6YD, United Kingdom. Tel.: +44 2380 777624; fax: +44 2380 704021. E-mail address: dnc@mrc.soton.ac.uk (D. Coggon).*

PAIN 157 (2016) 1028–1036

http://dx.doi.org/10.1097/j.pain.0000000000000477

the neck or shoulder was classed as disabling if it had made it difficult or impossible to get dressed, do normal jobs around the house, or (for shoulder only) comb/brush hair or bath/shower.

Approximately 14 months after baseline, participants from 45 occupational groups who had given consent were asked to complete a shorter follow-up questionnaire, which used identical questions to ascertain experience of musculoskeletal pain in the past month. Follow-up was not possible for office workers in South Africa or for manual workers in Costa Rica. Further details of the methods of data collection in the CUPID study have been reported elsewhere.[5]

Statistical analysis was performed with Stata version 12.1 software (Stata Corp LP, College Station, TX) and focused on the 1-month prevalence at baseline of 2 main categories of pain: localised neck/shoulder pain (LNSP) and generalised pain involving the neck or shoulder(s) (GPNS). The former was defined as pain in the neck and/or shoulder(s) with no pain at any of the other 7 anatomical sites during the past 12 months. Neck or shoulder pain during the past month that occurred in a context of pain at one or more other anatomical sites during the past 12 months was classed as GPNS. Within the category of LNSP, we further distinguished 2 subsets of cases—those in whom all pain in the past 12 months was restricted to the neck (localised neck pain) and those in whom all pain in the past 12 months was limited to one or both shoulders (localised shoulder pain).

We first used simple descriptive statistics to describe the frequency and severity of different pain outcomes at baseline. Next, we applied Poisson regression with confidence intervals (CIs) based on robust SEs to assess the cross-sectional association of various personal and occupational factors with each of LNSP, GPNS, localised neck pain, and localised shoulder pain. In each analysis, the measure of pain was taken as the outcome variable, and the comparison group was people with no pain in either the neck or shoulders in the past 12 months. To account for possible clustering by occupational group, we used hierarchical, random intercept modelling, associations being summarised by prevalence rate ratios (PRRs). To assess the significance of differences in associations for LNSP and GPNS, we performed a further Poisson regression analysis with GPNS as the outcome and LNSP as the comparator. Similarly, the significance of differences in associations with localised neck as compared with localised shoulder pain was assessed through a model with localised neck pain as the outcome and localised shoulder pain as the comparator.

We then examined the prevalence of LNSP and GPNS by occupational group and their correlation after adjustment for other factors. To derive adjusted prevalence rates, we first took no neck/shoulder pain in the past 12 months as a comparator and estimated PRRs for LNSP and GPNS in each occupational group relative to a reference (office workers in the United Kingdom), using Poisson regression models that included the other factors. Next, we calculated the "adjusted numbers" of participants in each occupational group with the 2 pain outcomes that would give crude PRRs equal to those estimated from the regression model. We then used these adjusted numbers to calculate adjusted prevalence rates.

Finally, we explored the prevalence of continuing pain in the month preceding follow-up for different categories of neck/shoulder pain at baseline.

Ethical approval for the study was provided by the relevant research ethics committee in each participating country.[5]

3. Results

One hundred ninety of the 12,426 participants were excluded from the analysis because of missing information about pain in the neck or shoulders in the past month and/or 12 months, together with a further 41 participants who provided incomplete data on pain at other anatomical sites in the past 12 months. Among the remaining 12,195 subjects, 4241 (35%) were men.

Table 1 shows the 1-month prevalence and severity of different categories of neck/shoulder pain at baseline. Overall, neck/shoulder pain in the past month was common (40.7%) and occurred mostly in a context of more widespread pain in the past 12 months (prevalence = 35.1%). In contrast, the prevalence of pain that was localised to the neck and/or shoulders was lower (5.6%) and particularly that of pain which was limited entirely to the shoulders (1.2%). Most of the latter was restricted to a single

Table 1

One-month prevalence and severity of different categories of neck/shoulder pain.

Category of pain	Definition	Number of cases*	Prevalence percent (95% CI)	Proportion percent (95% CI) of cases in which pain			
				Was present for >14 d in the past month	Was disabling in the past month	Led to medical consultation in the past 12 mo	Caused absence from work in the past 12 mo
Localised neck pain	Pain in neck in the past month, no pain elsewhere in the past 12 mo	302	2.5 (2.2-2.8)	18.5 (14.3-23.4)	23.8 (19.1-29.1)	35.1 (29.7-40.8)	10.9 (7.6-15.0)
Localised shoulder pain	Pain in 1 or both shoulders in the past month, no pain elsewhere in the past 12 mo	152	1.2 (1.1-1.5)	22.4 (16.0-29.8)	48.0 (39.9-56.3)	28.9 (21.9-36.8)	12.5 (7.8-18.8)
Localised neck/shoulder pain	Pain in neck and/or shoulder(s) in the past month, no pain elsewhere in the past 12 mo	680	5.6 (5.2-6.0)	24.0 (20.8-27.4)	35.6 (32.0-39.3)	38.2 (34.6-42.0)	11.3 (9.0-13.9)
Generalised pain involving neck/shoulder	Pain in neck and/or shoulders in the past month, with pain at other sites in the past 12 mo	4282	35.1 (34.3-36.0)	30.5 (29.1-31.9)	56.9 (55.4-58.4)	49.1 (47.6-50.6)	20.3 (19.1-21.6)

* In addition, 5344 participants had no neck or shoulder pain in the past 12 months, and 1889 had neck and/or shoulder pain in the past 12 months, but not in the past month.

shoulder (122 of the 152 cases). Most participants with GPNS (93%) reported pain in the past 12 months in the low back and/or knees as well as any pain in the upper limb. Generalised pain involving the neck or shoulder(s) tended to be more troublesome and disabling than LNSP. For example, it had made everyday activities difficult or impossible during the past month in 56.9% of cases as compared with 35.6% of those with LNSP; and it had caused absence from work during the past year in 20.3% of cases as compared with 11.3% of those in whom the pain was localised. When the analysis was repeated separately for 8 strata of sex and age, the patterns were similar across each stratum.

Table 2 shows baseline associations with personal and occupational factors separately for LNSP and GPNS, in comparison with no pain in the neck or shoulders in the past 12 months. Both categories of pain were significantly more frequent in women than in men (PRRs of 1.4 and 1.3, respectively), and both were associated with a tendency to somatise. However, the relationship to somatising tendency was stronger for GPNS (PRR 1.6, 95% CI 1.5-1.8, for report of distress from 2 or more somatic symptoms compared to none, *P* for trend <0.001) than for LNSP (PRR 1.3, 95% CI 1.1-1.5, for report of distress from 2 or more somatic symptoms compared to none, *P* for trend <0.001). Generalised pain involving the neck or shoulder(s) also showed a stronger association with poor mental health, and unlike LNSP, its prevalence increased significantly with age. Direct comparison of the 2 pain outcomes (in a single Poisson regression model that effectively took those with GPNS as cases and those with LNSP as controls) indicated that the differences in associations with age, somatising tendency, and mental health were statistically significant ($P < 0.05$). In addition, both categories of neck/shoulder pain were more weakly linked with prolonged use of keyboards at work (PRR 1.3 for LNSP and 1.1 for GPNS) and adverse beliefs about the prognosis of arm pain (PRRs 1.3 and 1.1, respectively); and GPNS with occupational lifting, work with the hands above shoulder height and adverse beliefs about the work-relatedness of arm pain.

Table 3 presents corresponding risk estimates for localised neck pain and localised shoulder pain, defined as in **Table 1** (again the comparator was no pain in the neck or shoulders in the past 12 months). There were clear differences in the patterns of association, such that the prevalence of localised neck pain was markedly higher in women than in men (PRR 1.7, 95% CI 1.2-2.3), lower at older ages (*P* for linear trend across age categories = 0.04), and significantly associated with somatising tendency, lack of support at work, job insecurity, and particularly use of a keyboard for >4 hours in an average working day (PRR 1.9,

Table 2

Associations of neck/shoulder pain with personal and occupational factors.

Personal/occupational factor	Localised neck/shoulder pain			Generalised pain involving the neck/shoulder		
	Number of cases (%)	PRR*	95% CI	Number of cases (%)	PRR*	95% CI
Sex						
Male	211 (7.9)	1		1009 (29.1)	1	
Female	469 (14.0)	1.4	1.1-1.7	3273 (53.1)	1.3	1.2-1.4
Age, y						
20-29	180 (10.6)	1		856 (36.0)	1	
30-39	230 (11.8)	1.1	0.9-1.4	1329 (43.6)	1.2	1.1-1.3
40-49	193 (12.5)	1.1	0.9-1.4	1327 (49.6)	1.3	1.1-1.4
50-59	77 (9.3)	0.8	0.7-1.1	770 (50.7)	1.3	1.2-1.5
Activity in an average working day						
Lifting weights ≥25 kg	199 (9.4)	0.9	0.7-1.1	1602 (45.5)	1.1	1.0-1.1
Working with hands above shoulder height for >1 h	177 (10.3)	1.0	0.9-1.3	1508 (49.4)	1.1	1.1-1.2
Use of keyboard for >4 h	316 (15.6)	1.3	1.1-1.6	1725 (50.2)	1.1	1.0-1.2
Psychosocial aspects of work						
Work for >50 h per week	140 (8.2)	1.0	0.7-1.2	592 (27.5)	0.9	0.8-1.0
Time pressure at work	485 (10.9)	1.1	0.9-1.3	3350 (45.9)	1.2	1.1-1.2
Incentives at work	188 (10.5)	1.0	0.9-1.2	1140 (41.5)	0.9	0.9-1.0
Lack of support at work	174 (15.4)	1.1	0.9-1.4	1344 (58.5)	1.1	1.0-1.1
Job dissatisfaction	146 (11.7)	1.1	0.9-1.4	868 (43.9)	1.0	1.0-1.1
Lack of job control	130 (10.3)	1.0	0.8-1.2	963 (46.0)	1.0	0.9-1.1
Job insecurity	195 (10.0)	1.0	0.9-1.1	1331 (43.2)	1.0	1.0-1.1
Number of distressing somatic symptoms in the past week						
0	425 (9.9)	1		1859 (32.4)	1	
1	163 (15.1)	1.3	1.1-1.6	1059 (53.6)	1.4	1.3-1.5
2+	87 (14.8)	1.3	1.1-1.5	1320 (72.4)	1.6	1.5-1.8
Missing	5 (9.6)	1.0	0.3-3.0	44 (48.4)	1.3	1.1-1.6
Mental health						
Good	265 (10.4)	1		1311 (36.4)	1	
Intermediate	216 (12.3)	1.2	0.9-1.4	1333 (46.3)	1.2	1.1-1.3
Poor	195 (11.7)	1.1	0.9-1.4	1621 (52.5)	1.3	1.2-1.4
Missing	4 (9.3)	1.1	0.5-2.5	17 (30.4)	0.8	0.5,1.1
Adverse beliefs about arm pain						
Work-relatedness	159 (11.4)	1.1	0.9-1.3	1613 (56.6)	1.2	1.2-1.3
Physical activity	62 (8.7)	0.8	0.6-1.0	425 (39.5)	0.8	0.8-0.9
Prognosis	69 (15.2)	1.3	1.0-1.5	604 (61.0)	1.1	1.1-1.2

The denominators for percentages of cases are the total numbers of cases and referents with the relevant personal/occupational factor.

* Prevalence rate ratios (PRRs) derived from a single Poisson regression model, with random intercept modelling to allow for clustering by occupational group. The reference group was participants with no pain in the neck or shoulders in the past 12 months (n = 5344).

Table 3

Associations of localised neck pain and localised shoulder pain with personal and occupational factors.

Personal/occuptional factor	Localised neck pain			Localised shoulder pain		
	Number of cases (%)	PRR*	95% CI	Number of cases (%)	PRR*	95% CI
Sex						
Male	85 (3.3)	1		58 (2.3)	1	
Female	217 (7.0)	1.7	1.2-2.3	94 (3.2)	1.3	0.9-1.8
Age, y						
20-29	96 (5.9)	1		39 (2.5)	1	
30-39	104 (5.7)	1.0	0.8-1.3	44 (2.5)	1.0	0.7-1.5
40-49	76 (5.3)	0.9	0.7-1.2	46 (3.3)	1.3	0.8-2.0
50-59	26 (3.4)	0.6	0.3-1.0	23 (3.0)	1.2	0.7-1.8
Activity in an average working day						
Lifting weights ≥25 kg	96 (4.8)	0.9	0.6-1.3	42 (2.1)	0.7	0.5-1.2
Working with hands above shoulder height for >1 h	82 (5.0)	1.0	0.8-1.3	50 (3.1)	1.3	1.0-1.8
Use of keyboard for >4 h	155 (8.3)	1.9	1.5-2.4	53 (3.0)	1.0	0.7-1.4
Psychosocial aspects of work						
Work for >50 h per wk	71 (4.3)	1.2	0.8-1.8	35 (2.2)	0.9	0.5-1.5
Time pressure at work	215 (5.2)	1.0	0.7-1.2	117 (2.9)	1.3	0.9-2.0
Incentives at work	92 (5.4)	1.2	0.9-1.5	40 (2.4)	0.8	0.5-1.2
Lack of support at work	92 (8.8)	1.5	1.1-2.0	21 (2.2)	0.7	0.4-1.1
Job dissatisfaction	63 (5.4)	1.1	0.8-1.3	34 (3.0)	1.3	0.9-2.0
Lack of job control	49 (4.2)	0.7	0.6-0.9	33 (2.8)	1.1	0.7-1.7
Job insecurity	104 (5.6)	1.2	1.0-1.5	43 (2.4)	0.9	0.6-1.2
Number of distressing somatic symptoms in the past week						
0	179 (4.4)	1		102 (2.6)	1	
1	77 (7.7)	1.5	1.1-2.0	28 (3.0)	1.0	0.7-1.5
2+	45 (8.2)	1.5	1.1-2.1	18 (3.5)	1.2	0.7-2.0
Missing	1 (2.1)	1	0.1-7.7	4 (7.8)	2.2	0.7-6.8
Mental health						
Good	117 (4.9)	1		56 (2.4)	1	
Intermediate	98 (6.0)	1.2	0.9-1.5	50 (3.1)	1.3	0.8-2.1
Poor	87 (5.6)	1.1	0.9-1.5	43 (2.8)	1.3	0.8-2.1
Missing	0 (0)	0.0	0.0-0.0	3 (7.1)	1.8	0.6-5.3
Adverse beliefs about arm pain						
Work-relatedness	66 (5.1)	0.9	0.7-1.3	38 (3.0)	1.1	0.8-1.5
Physical activity	28 (4.1)	0.8	0.5-1.2	20 (3.0)	1.1	0.7-1.8
Prognosis	23 (5.6)	0.9	0.7-1.3	15 (3.7)	1.4	1.0-2.1

The denominators for percentages of cases are the total numbers of cases and referents with the relevant personal/occupational factor.

* Prevalence rate ratios (PRRs) derived from a single Poisson regression model, with random intercept modelling to allow for clustering by occupational group. The reference group was participants with no pain in the neck or shoulders in the past 12 months (n = 5344).

95% CI 1.5-2.4). In contrast, localised shoulder pain was associated with work with the hands above shoulder height (PRR 1.3, 95% CI 1.0-1.8) and belief that arm pain has a poor prognosis (PRR 1.4, 95% CI 1.0-2.1). Direct comparison of localised neck pain with localised shoulder pain indicated that the differences in associations with age, use of a keyboard, and lack of support at work were statistically significant ($P < 0.05$).

Figure 1 summarises the crude prevalence of LNSP and GPNS across the 47 occupational groups. In almost all groups, GPNS predominated. Rates for LNSP ranged from 0.5% among flower plantation workers in Ecuador and 1.1% in Colombian office workers to 11.2% and 11.8% in office workers from Spain and Sri Lanka, respectively. For GPNS, the variation between occupational groups was even greater with zero prevalence among sugar cane cutters in Brazil and a rate as high as 67.6% in Costa Rican telephone call centre workers. Office workers tended to have higher prevalence of LNSP than nurses, but there were no consistent differences in GPNS by type of occupation. The proportion of neck/shoulder pain that was localised did not relate consistently to its overall prevalence.

Figure 2 plots the prevalence of LNSP by occupational group against that for GPNS, after adjustment for the factors listed in **Table 2**. After such adjustment, the variation in GPNS was a little reduced, and there was little correlation between the 2 types of pain (Spearman correlation coefficient = 0.22 overall, 0.46 in nurses, -0.47 in office workers and 0.45 in other workers).

Follow-up data were sought for 45 of the 47 occupational groups, and among the 11,764 participants in these groups, 9150 (78%) provided satisfactory information about neck and shoulder pain after a mean interval of 14 months (range 3-35 months, 84% within 11-19 months) from baseline. **Table 4** shows the prevalence of continuing pain at follow-up for different categories of neck/shoulder pain at baseline. The persistence of neck/shoulder pain was significantly higher ($P = 0.003$) when it was associated with pain at other anatomical sites in the past year (72.1%) than when it was localised (61.7%). Persistence was lowest for pain that was localised to the shoulders, with a prevalence of 31.9% for pain in the shoulder(s) during the month before follow-up and 41.6% for pain in the neck or shoulders.

4. Discussion

Our findings indicate that most neck and shoulder pain occurs in a context of current or recent pain at other anatomical sites, and that in these circumstances, it tends to be more troublesome and disabling than pain that is localised entirely to the neck/shoulder

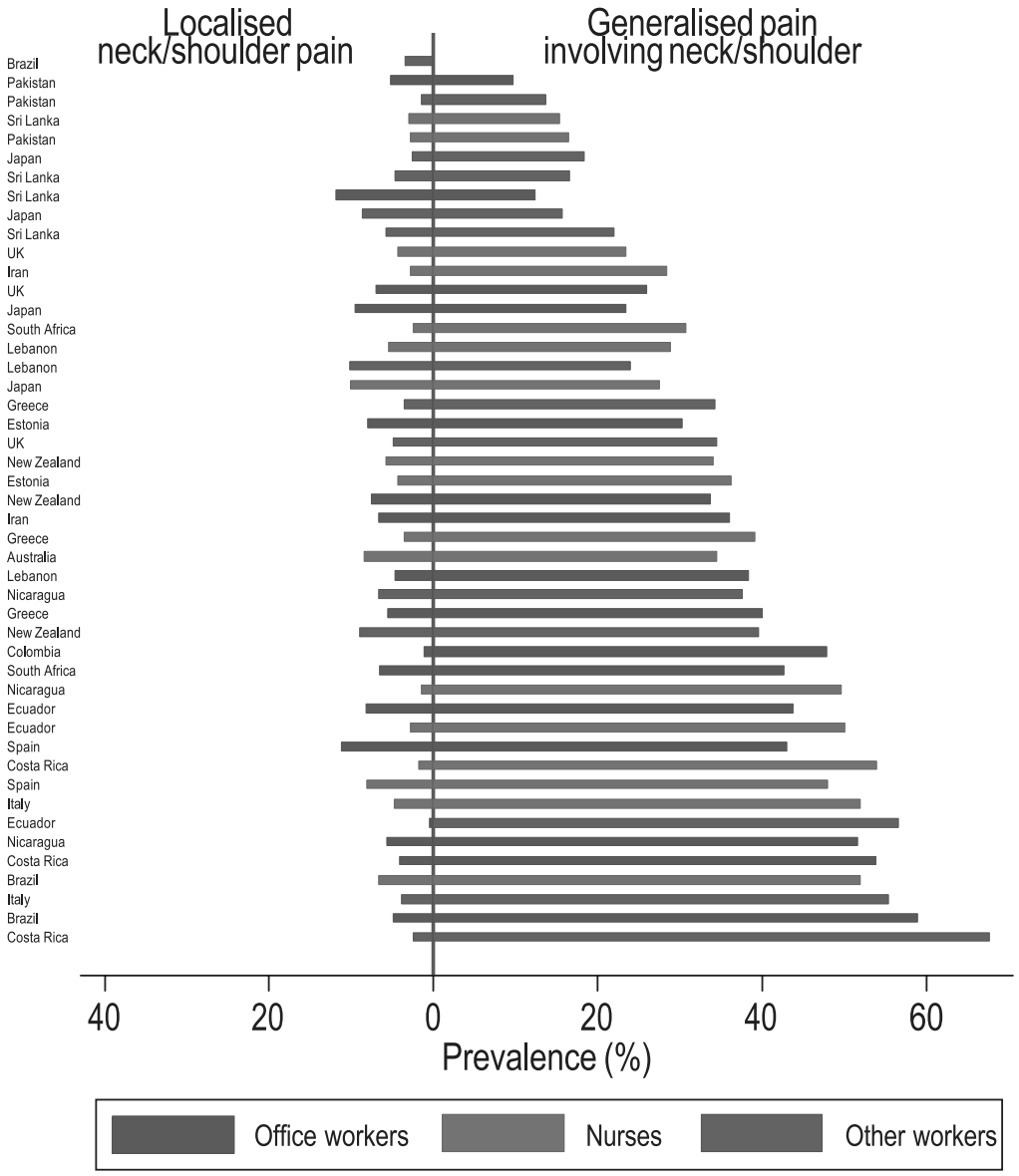


Figure 1. Crude 1-month prevalence of localised neck–shoulder pain and generalised pain involving neck/shoulder by occupational group.

area. Furthermore, it seems that these 2 subcategories of neck/shoulder pain differ importantly in their epidemiology. Thus, GPNS showed stronger associations with somatising tendency, poor mental health and older age, greater variation between the occupational groups studied (which correlated poorly with the variation in LNSP), and tended to be more persistent. There were also differences between pain that was localised to the neck and that which was localised to one or both the shoulders, the former being less frequent at older ages and strongly associated with prolonged use of keyboards at work.

Several earlier reports have described the occurrence and determinants of neck/shoulder pain in specific occupational groups from the CUPID study in Australia,[15] Brazil and Italy,[3] Estonia,[10,22] Iran,[26] Japan,[19] New Zealand,[11,12] and Sri Lanka.[30] However, the much larger size of the current analysis made it possible to examine diagnostic subgroups in a way that could not be done meaningfully with smaller data sets. It also enabled comparison of prevalence rates across a large number of occupational groups in a diversity of countries.

Information was collected through a standardised questionnaire, and while in some occupational groups, interviews were used as an alternative to self-administration, there is no reason to expect that this would have differentially affected the reporting of localised neck/shoulder pain as compared with pain that was more generalised. Translation of the questionnaire into local languages was checked by independent back-translation, and errors in reporting should have been further reduced by the use of pictures to define the anatomical areas of interest. In the cross-sectional analyses for **Tables 2 and 3**, there was a possibility of bias if experience of pain modified recall of occupational exposures, and also of reverse causation if, for example, neck/shoulder pain led to greater awareness of somatic symptoms. Moreover, the prevalence of pain may have been reduced through healthy worker selection. Again, however, there is no

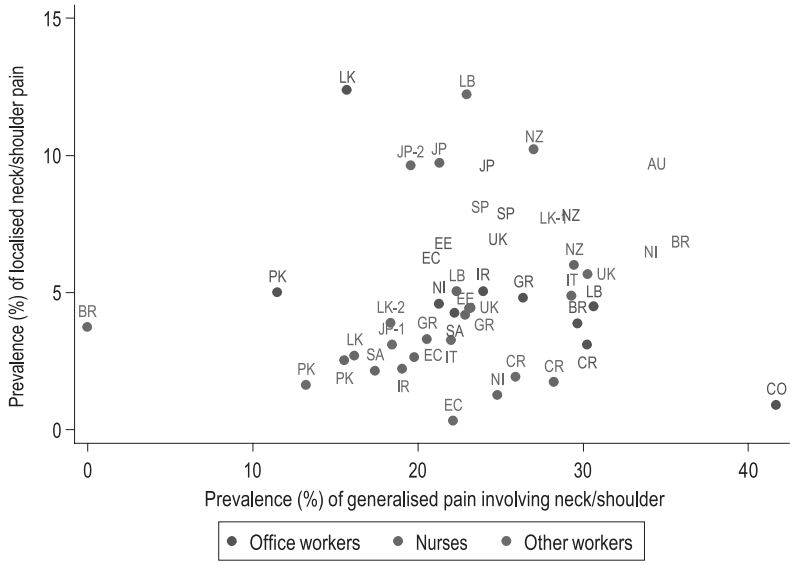


Figure 2. Adjusted 1-month prevalence of localised neck–shoulder pain and generalised pain involving neck/shoulder by occupational group. Prevalence rates are adjusted for all of the personal and occupational factors in Table 2 (See Section 2 on Methods). Key to countries: AU, Australia; BR, Brazil; CO, Colombia; CR, Costa Rica; EC, Ecuador; EE, Estonia; GR, Greece; IR, Iran; IT, Italy; JP, Japan; LB, Lebanon; LK, Sri Lanka; NI, Nicaragua; NZ, New Zealand; PK, Pakistan; SA, South Africa; SP, Spain; UK, United Kingdom.

reason to expect major differences in such effects according to whether pain was localised to the neck and shoulder or more generalised.

The criteria by which we distinguished between LNSP and GPNS were to some extent arbitrary. Subject to the limits of participants' recall, they ensured that those with LNSP had not suffered from pain at other anatomical sites in the 12 months before baseline. However, it remains possible that this group included some people with a predisposition to pain at multiple sites which would have become manifest had they been studied over a longer period. Nevertheless, the case definitions were adequate to reveal important differences in epidemiological features. Some GPNS may have reflected radiation to the distal arm of pain that arose from primary pathology in the neck, but most participants with GPNS (93%) reported pain in the past 12 months in the low back and/or knees as well as any pain in the upper limb.

Several previous studies have documented the frequent co-occurrence of neck and shoulder pain[24] and also their association with pain in the lower back and at other anatomical sites.[8,18,21] However, we have been unable to identify any investigations that focused on pain limited only to the neck or shoulders.

Studies that have examined neck pain overall have found higher rates in women than in men[21,23] and at older ages.[21] This is consistent with our findings for GPNS (which included most of the participants with neck pain in the past month at baseline). Moreover, pain that was localised entirely to the neck showed an even greater difference by sex (PRR 1.7, 95% CI 1.2-2.3). However, in contrast to GPNS, localised neck pain was less prevalent at older ages.

Many studies have examined the relationship of neck and/or shoulder pain to physical activities at work, associations being found most consistently with manual material handling[20] and awkward postures,[2,20,24] including work with the hands above shoulder height,[20] and to a lesser extent with computer work.[28] We found that localised neck pain was more prevalent among participants who reported prolonged use of computer keyboards (PRR 1.9, 95% CI 1.5-2.4) and a borderline association of localised shoulder pain with prolonged elevation of the arms at work, but otherwise there were no clear relationships with physical activities for any of the pain outcomes examined. This may have been because within each occupational group, exposures to physical activities were fairly homogeneous, making

Table 4

One-month prevalence of pain in neck and/or shoulders at follow-up by category of neck/shoulder pain at baseline.

Category of pain at baseline	Number of cases eligible for follow-up	Number of cases (%) who completed follow-up	Number (%)* of cases with pain at the same site in the past month at follow-up			Number (%)* of cases with pain in neck and/or shoulder(s) in the past month at follow-up		
			Number of cases	Percentage*	95% CI	Number of cases	Percentage*	95% CI
Localised neck pain	289	219 (75.8)	123	56.2	49.3-62.8	137	62.6	55.8-69.0
Localised shoulder pain	148	113 (76.4)	36	31.9	23.4-41.3	47	41.6	32.4-51.2
Localised neck/shoulder pain	660	501 (75.9)				309	61.7	57.3-66.0
Generalised pain involving neck/shoulder	4047	3253 (80.4)				2344	72.1	70.5-73.6

Analysis was restricted to the 9150 cases with satisfactory information about neck/shoulder pain at follow-up.
* Percentage of those who completed follow-up.

it difficult to detect effects in analyses that adjusted for occupational group.

Homogeneity of exposures within occupational groups may also have limited our ability to discriminate associations with psychosocial aspects of work, which again have been implicated previously in the occurrence of neck and shoulder pain,[16,21,24] although mostly with estimates of relative risk less than 2.[16] It was not possible to identify any clear differences between LNSP and GPNS, and although lack of support at work carried a higher risk of localised neck pain (PRR 1.5, 95% CI 1.1-2.0), this may have been a chance observation in the context of multiple testing.

In contrast, stronger associations were observed with somatising tendency and poor mental health. For poor mental health, which has been linked previously with neck pain,[28] the relationship was limited to GPNS, whereas for somatising tendency, it extended to localised pain but was stronger for GPNS (PRR for ≥2 vs 0 distressing somatic symptoms 1.6, 95% CI 1.5-1.8). The last finding is consistent with the earlier observation that within the CUPID study, somatising tendency is associated particularly with multisite musculoskeletal pain.[7] It might be expected that people who are prone to worry about other common somatic symptoms would also be more aware of musculoskeletal pain and more likely to report it. The weaker relationship to localised pain in the shoulder(s) suggests that the latter may be determined more by localised factors (eg, pathology in the shoulder or health beliefs relating specifically to the shoulder).

The association that we observed between localised shoulder pain and adverse beliefs about the prognosis of arm pain (PRR 1.4 95% CI 1.0-2.1) may reflect a relationship more to the persistence than the incidence of symptoms. Prevalence depends on both incidence and persistence, and our cross-sectional analysis could not distinguish between effects on one as compared with the other.

We also found marked differences in the prevalence of neck/shoulder pain by occupational group and in the proportion of such pain that was localised to the neck/shoulder region (evidenced by the absence of clear positive correlation between the 2 categories of pain in **Figure 1**). The larger variation was for GPNS and tended to parallel that reported previously for disabling pain in the low back and wrist/hand regions.[6] This may be because, like neck/shoulder pain, most low back and wrist/hand pain occurs in people with a high susceptibility to musculoskeletal pain in general. The differences were somewhat reduced after control for known and suspected risk factors, but remained large. They might in part reflect differences in understanding of pain, across different cultures and especially between populations speaking different languages. However, such differences could not explain the variation between occupational groups in the proportion of neck/shoulder pain that was localised, which again was marked. The lack of correlation between the prevalence of LNSP and GPNS suggests differences in general predisposition to complain of musculoskeletal pain, perhaps culturally determined, that are not explained by differences in somatising tendency or other known or suspected risk factors for such pain. Whatever the explanation, the differential variation adds to the case for treating LNSP and GPNS as separate entities.

That case is further supported by the observation that in comparison with LNSP, GPNS tended to be more persistent at follow-up. This pattern would be expected if the general predisposition to pain that seems to drive rates of GPNS were a fairly unchanging personal characteristic, whereas LNSP was more influenced by transient factors such as reversible injuries to local tissues in the neck and shoulders.

In conclusion, our findings point to important distinctions between subcategories of neck/shoulder pain. It is uncommon for people with neck/shoulder pain not to have experienced pain also at other anatomical sites during the past year and those whose pain is limited to the neck and/or shoulders tend to be younger, to somatise less, and to be less disabled by their pain. Localised neck/shoulder pain is also less persistent than that which is associated with pain elsewhere and shows stronger associations with occupational physical activities, perhaps reflecting specific effects on local tissues (eg, muscle fatigue from postures associated with prolonged use of keyboards). In future research on neck/shoulder pain that bases case definitions only on symptoms, it would be useful to distinguish pain that is localised to the neck or shoulder from more generalised pain that happens to involve the neck/shoulder region as well as other parts of the body.

Conflict of interest statement

The authors have no conflicts of interest to declare.

Acknowledgements

The authors thank Pietro Muñoz, Patricio Oyos, Gonzalo Albuja, María Belduma, and Francisco Lara for their assistance with data collection in Ecuador; Patrica Monge, Melania Chaverrri, and Freddy Brenes who helped with data collection in Costa Rica; Aurora Aragón, Alberto Berríos, Samaria Balladares, and Martha Martínez who helped with data collection in Nicaragua; Alfredo José Jirón who assisted with data entry in Nicaragua; Catalina Torres for translation and piloting of the questionnaire in Spain; Ben and Marie Carmen Coggon for back translation of the Spanish questionnaire; Cynthia Alcantara, Xavier Orpella, Josep Anton Gonzalez, Joan Bas, Pilar Peña, Elena Brunat, Vicente San José, Anna Sala March, Anna Marquez, Josefina Lorente, Cristina Oliva, Montse Vergara, and Eduard Gaynés for their assistance with data collection in Spain; Natale Battevi, Lorenzo Bordini, Marco Conti, and Luciano Riboldi who performed data collection in Italy; Paul Maurice Conway for back translation of the Italian questionnaire; Tuuli Sirk who helped with data collection in Estonia; the Deputy for Training and Research, Shahroud University of Medical Sciences for financial support of data collection in Iran; Asad Ali Khan for supervision of data collection and checking in Pakistan; Khalil Qureshi for training of field workers and supervision of data collection and checking in Pakistan; and Masami Hirai, Tatsuya Isomura, Norimasa Kikuchi, Akiko Ishizuka, and Takayuki Sawada for their help with data collection and management in Japan; Monash University which funded data collection in Australia through its grant schemes; NHMRC which supported Helen Kelsall and Donna Urquhart in Australia through fellowships; the Ministry of Higher Education in Malaysia which supported Victor Hoe in Australia; and the Health Research Council of New Zealand which funded data collection in New Zealand. Data collection in Central America and Colombia was supported by the Southwest Center for Occupational and Environmental Health at the University of Texas Health Science Center research training grant from the NIH Fogarty International Center.
Coordenação de Aperfeiçoamento de Pessoal de nível Superior (CAPES), Brasilia, DF, Brazil supported Leila Mansano Sarquis through a postdoctoral fellowship (BEX no 6841/14-7), enabling her to work on this article during an attachment at University of Southampton, United Kingdom.
Sergio Vargas-Prada was supported by the program Rio-Hortega, Institute of Health Carlos III (ISCIII), Spain.
Eduardo J. Salazar Vega is now employed by AkzoNobel, USA.

The authors are particularly grateful to the Colt Foundation, which funded data collection in Brazil, Ecuador, Costa Rica, Nicaragua, United Kingdom, Greece, Estonia, Lebanon, Pakistan, and South Africa; all organisations that allowed us to approach their employees; and all of the workers who kindly participated in the study.

Article history:
Received 1 October 2015
Received in revised form 17 December 2015
Accepted 22 December 2015
Available online 5 January 2016

References

[1] Andersen LL, Clausen T, Carneiro IG, Holterman A. Spreading of chronic pain between body regions: prospective cohort study among health care workers. Eur J Pain 2012;16:1437–43.
[2] Ariëns GAM, van Mechelen W, Bongers PM, Bouter LM, van der Wal G. Physical risk factors for neck pain. Scand J Work Environ Health 2000;26: 7–19.
[3] Carugno M, Pesatori AC, Ferrario MM, Ferrari AL, Silva FJ, Martins AC, Felli VE, Coggon D, Bonzini M. Physical and psychosocial risk factors for musculoskeletal disorders in Brazilian and Italian nurses. Cad Saude Publica 2012;28:1632–42.
[4] Coggon D, Martyn C, Palmer KT, Evanoff B. Assessing case definitions in the absence of a diagnostic gold standard. Int J Epidemiol 2005;34:949–52.
[5] Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Serra C, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SSP, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H; other members of the CUPID Collaboration. The CUPID (Cultural and Psychosocial Influences on Disability) Study: methods of data collection and characteristics of study sample. PLoS ONE 2012;7:1–12.
[6] Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Serra C, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SSP, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Marziale MH, Sarquis LM, Harari F, Feire R, Harari N, Monroy MV, Quintana LA, Rojas M, Salazar Vega EJ, Harris EC, Vargas-Prada S, Martinez JM, Declos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Sirk T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VCW, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A. Disabling musculoskeletal pain in working populations: is it the job, the person, or the culture?. PAIN 2013;154:856–63.
[7] Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH, Felknor SA, Gimeno D, Cattrell A, Serra C, Bonzini M, Solidaki E, Merisalu E, Habib RR, Sadeghian F, Kadir MM, Warnakulasuriya SSP, Matsudaira K, Nyantumbu B, Sim MR, Harcombe H, Cox K, Marziale MH, Sarquis LM, Harari F, Feire R, Harari N, Monroy MV, Quintana LA, Rojas M, Salazar Vega EJ, Harris EC, Vargas-Prada S, Martinez JM, Declos G, Benavides FG, Carugno M, Ferrario MM, Pesatori AC, Chatzi L, Bitsios P, Kogevinas M, Oha K, Sirk T, Sadeghian A, Peiris-John RJ, Sathiakumar N, Wickremasinghe AR, Yoshimura N, Kelsall HL, Hoe VCW, Urquhart DM, Derrett S, McBride D, Herbison P, Gray A. Patterns of multi-site pain and associations with risk factors. PAIN 2013;154:1769–77.
[8] Croft PR, Lewis M, Papageorgiou AC, Thomas E, Jayson MI, Macfarlane GJ, Silman AJ. Risk factors for neck pain: a longitudinal study in the general population. PAIN 2001;93:317–25.
[9] Derogatis LR, Melisaratos N. The brief symptom inventory: an introductory report. Psychol Med 1983;13:595–605.
[10] Freimann T, Coggon D, Merisalu E, Animägi L, Pääsuke M. Risk factors for musculoskeletal pain amongst nurses in Estonia: a cross-sectional study. BMC Musculoskelet Disord 2013;14:334.
[11] Harcombe H, McBride D, Derrett S, Gray A. Prevalence and impact of musculoskeletal disorders in New Zealand nurses, postal workers and office workers. Aust New Zealand J Pub Health 2009;33:437–41.
[12] Harcombe H, McBride D, Derrett S, Gray A. Physical and psychosocial risk factors for musculoskeletal disorders in New Zealand nurses, postal workers and office workers. Inj Prev 2010;16:96–100.
[13] Hartvigsen J, Davidsen M, Hestbaek L, Roos EM. Patterns of musculoskeletal pain in the population: a latent class analysis using a nationally representative interviewer-based survey of 4817 Danes. Eur J Pain 2013;17:452–60.
[14] Haukka E, Leino-Arjas P, Solovieva S, Ranta R, Viikari-Juntura E, Riihimäki H. Co-occurrence of musculoskeletal pain among female kitchen workers. Int Arch Occup Environ Health 2006;80:141–8.
[15] Hoe VC, Kelsall HL, Urquhart DM, Sim MR. Risk factors for musculoskeletal symptoms of the neck or shoulder alone or neck and shoulder among hospital nurses. Occup Environ Med 2012;69:198–204.
[16] Kraatz S, Lang J, Kraus T, Münster E, Ochsmann E. The incremental effect of psychosocial workplace factors on the development of neck and shoulder disorders: a systematic review of longitudinal studies. Int Arch Occup Environ Health 2013;86:375–95.
[17] Lawrence JS. Disc degeneration: its frequency and relationship to symptoms. Ann Rheum Dis 1969;28:121–37.
[18] Mäkelä M, Heliövaara M, Sievers K, Impivaara O, Knekt P, Aromaa A. Prevalence, determinants, and consequences of chronic neck pain in Finland. Am J Epidemiol 1991;134:1356–67.
[19] Matsudaira K, Palmer KT, Reading I, Hirai M, Yoshimura N, Coggon D. Prevalence and correlates of regional pain and associated disability in Japanese workers. Occup Environ Med 2011;68:191–6.
[20] Mayer J, Kraus T, Ochsmann E. Longitudinal evidence for the association between work-related physical exposures and neck and/or shoulder complaints: a systematic review. Int Arch Occup Environ Health 2012;85: 587–603.
[21] McLean SM, May S, Klaber-Moffett J, Sharp DM, Gardiner E. Risk factors for the onset of non-specific neck pain: a systematic review. J Epidemiol Community Health 2010;64:565–72.
[22] Oha K, Animägi L, Pääsuke M, Coggon D, Merisalu E. Individual and work-related risk factors for musculoskeletal pain: a cross-sectional study among Estonian computer users. BMC Musculoskelet Disord 2014;15:181.
[23] Paksaichol A, Janwantanakul P, Purepong N, Pensri P, van der Beek AJ. Office workers' risk factors for the development of non-specific neck pain: a systematic review of prospective cohort studies. Occup Environ Med 2012;69:610–18.
[24] Palmer KT, Smedley J. Work relatedness of chronic neck pain with physical findings—a systematic review. Scand J Work Environ Health 2007;33:165–91.
[25] Parot-Schinkel E, Descatha A, Ha C, Petit A, Leclerc A, Rocquelaure Y. Prevalence of multi-site musculoskeletal symptoms: a French cross-sectional working population-based study. BMC Musculoskelet Disord 2012,13:122.
[26] Sadeghian F, Raei M, Ntani G, Coggon D. Predictors of incident and persistent neck/shoulder pain in Iranian workers: a cohort study. Plos One 2013;8:e57544.
[27] Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. PAIN 1993;52:157–68.
[28] Waersted M, Hanvold TN, Veiersted KB. Computer work and musculoskeletal disorders of the neck and upper extremity: a systematic review. BMC Musculoskelet Disord 2010;11:79.
[29] Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). Med Care 1992;30:473–83.
[30] Warnakulasuriya SSP, Peiris-John RJ, Coggon D, Ntani G, Sathiakumar N, Wickremasinghe AR. Musculoskeletal pain in four occupational populations in Sri Lanka. Occup Med 2012;62:269–72.

Osteoarthritis and Cartilage 24 (2016) 117–123

Osteoarthritis and Cartilage

Prevalence of radiographic hip osteoarthritis and its association with hip pain in Japanese men and women: the ROAD study

T. Iidaka †, S. Muraki ‡ *, T. Akune §, H. Oka ||, R. Kodama ¶, S. Tanaka ¶, H. Kawaguchi #, K. Nakamura §, N. Yoshimura †

† Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan
‡ Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan
§ National Rehabilitation Center for Persons with Disabilities, Saitama, Japan
|| Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan
¶ Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan
# Department of Orthopaedic Surgery, Japan Community Healthcare Organization Tokyo Shinjuku Medical Center, Tokyo, Japan

ARTICLE INFO

Article history:
Received 26 January 2015
Accepted 20 July 2015

Keywords:
Osteoarthritis
Hip
Prevalence
Pain
Cross-sectional

SUMMARY

Objective: Although hip osteoarthritis (OA) is a major cause of hip pain and disability in elderly people, few epidemiologic studies have been performed. We investigated the prevalence of radiographic hip OA and its association with hip pain in Japanese men and women using a large-scale population of a nationwide cohort study, Research on Osteoarthritis/osteoporosis Against Disability (ROAD).
Methods: From the baseline survey of the ROAD study, 2975 participants (1043 men and 1932 women), aged 23–94 years (mean 70.2 years), living in urban, mountainous, and coastal communities were analyzed. The radiographic severity at both hips was determined by the Kellgren/Lawrence (K/L) grading system. Radiographic hip OA was defined as K/L $\geq$ 2, and severe radiographic hip OA as K/L $\geq$ 3.
Results: The crude prevalence of radiographic hip OA was 18.2% and 14.3% in men and women, respectively, that of severe radiographic hip OA was 1.34% and 2.54%, and that of symptomatic K/L $\geq$ 2 OA was 0.29% and 0.99%, respectively. The crude prevalence of hip OA, including severe OA, was not age-dependent in men or women. Male sex was a risk factor for radiographic hip OA, whereas female sex was a risk factor for severe radiographic hip OA and hip pain. Compared with K/L = 0/1, hip pain was significantly associated with K/L $\geq$ 3, but not with K/L = 2.
Conclusion: The present cross-sectional study revealed the prevalence of radiographic hip OA and severe hip OA in Japanese men and women. Hip pain was strongly associated with K/L $\geq$ 3.

Introduction

Hip osteoarthritis (OA) is a major public health issue causing chronic disability of elderly people in most developed countries[1,2]. Despite the urgent need for strategies to prevent and treat this condition, epidemiologic data on hip OA are sparse. The reported prevalence of radiographic hip OA differs considerably among previous population-based epidemiologic studies[1,3–8]. This may be due to limitations in sample size or variability in age, ethnicity, and radiologic acquisition. In particular, previous studies suggested that the prevalence of OA at other sites, such as the knee, differed among races. In addition, anthropometric measurements and environmental situations vary substantially in different countries. Thus findings in Caucasians cannot be applied to different ethnic groups. In Japan, our previous study in 1998 was the only population-based study to examine the prevalence of hip OA. With the aging population, there have been dramatic changes in number of elderly people; this aging may have affected the prevalence of hip OA. To the best of our knowledge, no population-based cohort studies for hip OA have been performed in Japan since our previous study.

* Address correspondence and reprint requests to: S. Muraki, Department of Clinical Motor System Medicine, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. Tel: 81-3-5800-9178; Fax: 81-3-5800-9179.
E-mail address: murakis-ort@h.u-tokyo.ac.jp (S. Muraki).

http://dx.doi.org/10.1016/j.joca.2015.07.017
1063-4584/

Hip pain is the principal clinical symptom of hip OA[9], but the reported prevalence of hip pain and symptomatic hip OA also differs among previous population-based epidemiologic studies[1,5–8]. In addition the impact of hip OA on pain remains controversial.

With the goal of establishing epidemiologic indices to evaluate clinical evidence for the development of disease-modifying treatment, we set up a large-scale nationwide cohort study for bone and joint disease called ROAD (Research on Osteoarthritis/osteoporosis Against Disability) in 2005. We have to date created a baseline database with detailed clinical and genetic information on three population-based cohorts in urban, mountainous, and coastal communities of Japan.

The objective of this study was to examine the prevalence of radiographic hip OA as well as hip pain and symptomatic hip OA by gender and age strata in Japanese men and women in a large-scale, population-based cohort from the ROAD study. We also examined the association of the severity of hip OA with the presence of hip pain.

Subjects and methods

The ROAD study is a nationwide prospective study of bone and joint diseases (with osteoarthritis and osteoporosis as the representative bone and joint diseases) constituting population-based cohorts established in several communities in Japan. As a detailed profile of the ROAD study has already been described elsewhere[10–12], a brief summary is provided here. From 2005 to 2007, we created a baseline database that included clinical and genetic information for 3040 inhabitants (1061 men, 1979 women) in the age range of 23–95 years (mean 70.6 years), recruited from listings of resident registrations in three communities: an urban region in Itabashi, Tokyo, with a population of 529,400/32 km^2 with 0.1, 25, and 75% of jobs in the primary industry (agriculture, forestry, fishing, and mining), the secondary industry (manufacturing and construction), and the tertiary industry (service industry), respectively, and residents $\geq$65 years constituted 19.1% of the population; a mountainous region in Hidakagawa, Wakayama, with a population of 11,300/330 km^2 with 29, 24, and 47% of jobs in the three industries above, and 30.5% were $\geq$65 years; and a coastal region in Taiji, Wakayama, with a population of 3500/6 km^2 with 13, 18, and 69% of jobs in the three industries, and those $\geq$65 years accounted for 34.9% of the total. Participants in the urban region were recruited from a cohort study[13] in which the participants were randomly drawn from the Itabashi-ward residents register database, and the response rate in the age groups of $\geq$60 years was 75.6%. Participants in the mountainous and coastal regions were recruited from listings of resident registration and the response rates in the age group of $\geq$40 years were 57.3% and 33.1%, respectively. However, those inhabitants aged <60 years in the urban region and <40 years in the mountainous and coastal regions who were interested in participating in the study were invited to be examined. The inclusion criteria, apart from residence in the communities mentioned above, were the ability to walk to the survey site, report data, and understand and sign an informed consent form. All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Geriatric Medical Center.

Participants completed an interviewer-administered questionnaire of 400 items that included lifestyle information such as smoking habits, alcohol consumption, family history, medical history, and previous hip injury history. Anthropometric measurements included height and weight, from which the body mass index (BMI) (weight [kg]/height [m^2]) was calculated. Furthermore, all participants were interviewed by well-experienced orthopedists regarding pain in both hips, who asked “Have you experienced right hip pain on most days in the past month, in addition to now?” and “Have you experienced left hip pain on most days in the past month, in addition to now?” Subjects who answered “yes” were defined as having hip pain. We defined an individual as having hip pain if at least one of the hip joints was affected.

Radiographic assessment

All participants underwent radiographic examination of both hips using an anteroposterior view with weight-bearing and feet internally rotated. Fluoroscopic guidance with a horizontal anteroposterior X-ray beam was used to properly visualize the joint space. Hip radiographs at baseline were read without knowledge of the participant's clinical status by a single, well-experienced orthopedist (TI), and the Kellgren/Lawrence (K/L) grade was defined using the K/L radiographic atlas for overall hip radiographic grades[14]. In the K/L grading system, radiographs are scored from grade 0 to grade 4, with higher grades being associated with more severe OA. To evaluate intraobserver variability of K/L grading, 100 randomly selected radiographs of the hip were scored by the same observer more than 1 month after the first reading. One hundred other radiographs were also scored by two experienced orthopedic surgeons (TI and SM) using the same atlas for interobserver variability. The intra- and intervariabilities evaluated for K/L grade (0–4) were confirmed by kappa analysis to be sufficient for assessment ($\kappa = 0.87$ and $\kappa = 0.85$, respectively).

Radiographic hip OA was defined as a K/L radiographic severity grade $\geq$2 (i.e., presence of at least probable joint space narrowing [JSN] in either the superolateral or superomedial hip joint, as well as presence of an osteophyte) and severe radiographic hip OA was defined as K/L $\geq$ 3. We defined an individual as having radiographic hip OA if at least one of the hip joints was affected. In addition, symptomatic hip OA was defined as having hip pain with corresponding radiographic OA in the same hip. Prevalence of total prevalence of hip OA (%) = (total number of subjects who were diagnosed as radiographic hip OA/total subjects who participated in the X-ray examination) $\times$ 100.

Individuals who had undergone a total hip arthroplasty (THA) were defined as having severe radiographic hip OA in that joint ($n = 13$ subjects; 18 hips). However at the time of analysis of the association with hip pain, we excluded all subjects who had undergone a THA.

Statistical analysis

Odds ratios (ORs) and 95% confidence intervals (95% CIs) are provided. Differences of age and BMI between men and women were examined by non-paired *t*-test. Differences in age, height, weight, and BMI among the urban, mountainous, and coastal communities were determined using one-way analysis of covariance and Tukey's honestly significant difference test. We used the chi square test to compare the prevalence of radiographic hip OA between men and women. Association of prevalence with age was determined by logistic-regression analysis after adjustment for BMI. Association of the variables such as age, BMI, gender, and community with radiographic hip OA was evaluated by multivariate logistic-regression analysis. Logistic-regression analyses were used to estimate OR and the associated 95% CI of K/L = 2 and K/L $\geq$ 3 hip OA for pain compared with K/L = 0/1 after adjustment for age, BMI, and community. Data analyses were performed using SAS version 11.0 (SAS Institute Inc., Cary, NC).

Results

Characteristics of participants

Of the 3040 subjects in the present study, 62 (2.0%) did not undergo plain radiography, 1 (0.03%) had just experienced bilateral hip fractures, and 2 (0.07%) could not read; these subjects were excluded. The remaining 2975 subjects (95.8%) (1043 men and 1932 women), aged 23–94 years (mean 70.2 years), were included in this study (Table I). Men were significantly older than women in the overall population and in the urban population. Although the coastal residents tended to show higher body height and weight than residents in the other two communities, BMI was comparable among the three communities.

Prevalence of radiographic hip OA, hip pain, and symptomatic hip OA

Table II shows the prevalence of radiographic hip OA, severe radiographic hip OA, including unilateral and bilateral hip OA, hip pain, and symptomatic hip OA in the overall population and subgroups classified by gender and community. In the overall population, the prevalence of radiographic hip OA was 15.7%, severe radiographic hip OA was 2.12%, and that of hip pain was 1.86%. The prevalence of K/L $\geq$ 2 and K/L $\geq$ 3 symptomatic hip OA was 0.75% and 0.64%. The prevalence of radiographic hip OA was significantly higher in men than in women, but that of severe radiographic hip OA, hip pain, and symptomatic hip OA was significantly higher in women than in men. The prevalence of radiographic hip OA and hip pain were not significantly associated with age in either gender [Fig. 1]. Table II also shows the prevalence of radiographic hip OA classified by the regions. In the urban region, the prevalence of K/L $\geq$ 2 hip OA was 9.4% in men and 6.0% in women, respectively, and that of K/L $\geq$ 3 was 0.89% and 2.13%, respectively. In the mountainous region, the prevalence of K/L $\geq$ 2 hip OA was 16.4% in men and 16.1% in women, respectively, and that of K/L $\geq$ 3 was 0.63% and 2.59%, respectively. In the coastal region, the prevalence of K/L $\geq$ 2 hip OA was 34.7% in men and 25.4% in women, respectively, and that of K/L $\geq$ 3 was 2.89% and 3.11%, respectively. In the urban and mountainous regions, the prevalence of K/L $\geq$ 2 hip OA was significantly higher in men than in women, and in the coastal region, the prevalence of K/L $\geq$ 3 hip OA was significantly higher in women than in men.

Characteristics of participants classified by presence or absence of hip OA and hip pain

Mean age of subjects with and without radiographic hip OA was 70.4 ± 10.4 and 70.2 ± 11.2 years, respectively ($P = 0.68$). Mean age of subjects with and without severe radiographic hip OA was 72.5 ± 9.3 and 70.1 ± 11.1 years, respectively ($P = 0.05$), and that of subjects with and without hip pain was 67.6 ± 13.6 and 70.2 ± 11.1 years, respectively ($P = 0.16$).

Association of radiographic hip OA with hip pain

Table III shows the association of age, BMI, gender, and community with radiographic hip OA, severe radiographic hip OA, and hip pain. BMI was classified as normal (18.5 $\leq$ BMI < 25.0), thin (BMI < 18.5), obesity (25.0 $\leq$ BMI < 27.5), and high obesity (BMI $\geq$ 27.5). BMI $\geq$ 27.5, female sex, and community were significantly associated with radiographic hip OA. Female sex and coastal community were significantly associated with severe radiographic hip OA. Only female sex was significantly associated with hip pain. We then determined independent associated factors using a multiple logistic regression analysis that included the above significant

Table I
Characteristics of participants

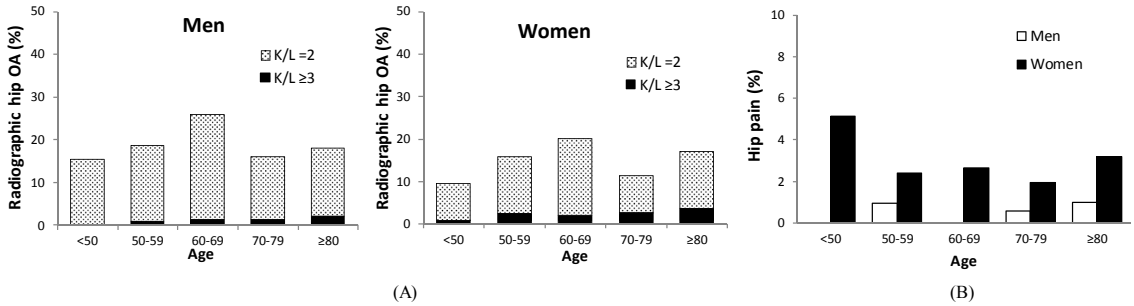
	Men				Women			
	Overall	Urban	Mountainous	Coastal	Overall	Urban	Mountainous	Coastal
Number of subjects	1043	449	317	277	1932	845	540	547
Age (years)	71.0 ± 10.7	77.2 ± 4.2	69.5 ± 9.1†	62.6 ± 13.2†	69.8 ± 11.3*	76.3 ± 5.0*	68.6 ± 10.4†	60.8 ± 12.5†
Height (cm)	162.5 ± 6.7	161.3 ± 5.9	161.3 ± 6.9	165.8 ± 6.8†	149.8 ± 6.5*	148.6 ± 5.6*	148.2 ± 6.7*	153.2 ± 6.2*,†
Weight (kg)	61.3 ± 10.0	60.1 ± 8.7	60.0 ± 10.2	64.8 ± 11.0†	51.5 ± 8.6*	50.7 ± 8.4*	50.5 ± 8.6*	53.5 ± 8.8*,†
BMI (kg/m^2)	23.2 ± 3.1	23.1 ± 2.9	23.0 ± 3.0	23.5 ± 3.4	22.9 ± 3.5*	23.0 ± 3.5	23.0 ± 3.3	22.8 ± 3.6*

Data are means ± SD.
BMI, body mass index.
* $P < 0.05$ vs men in the corresponding group by non-paired *t*-test.
† $P < 0.05$ vs urban residents in the corresponding group by Tukey's honestly significant difference test.

Table II
Number (percentage) of participants with radiographic hip osteoarthritis, hip pain, and their combination

	Total ($n = 2975$)	Men ($n = 1043$)				Women ($n = 1932$)			
		Overall	Urban	Mountainous	Coastal	Overall	Urban	Mountainous	Coastal
K/L $\geq$ 2 hip OA									
Total	467 (15.7)	190 (18.2)	42 (9.4)	52 (16.4)	96 (34.7)	277 (14.3)*	51 (6.0)*	87 (16.1)	139 (25.4)*
Unilateral	278 (9.3)	103 (9.9)	29 (6.5)	30 (9.5)	44 (15.9)	175 (9.1)	36 (4.3)	55 (10.2)	84 (15.4)
Bilateral	189 (6.4)	87 (8.4)	13 (2.9)	22 (7.1)	52 (19.0)	102 (5.3)*	15 (1.8)	32 (6.0)	55 (10.1)*
K/L $\geq$ 3 hip OA									
Total	63 (2.12)	14 (1.34)	4 (0.89)	2 (0.63)	8 (2.89)	49 (2.54)*	18 (2.13)	14 (2.59)*	17 (3.11)
Unilateral	37 (1.24)	7 (0.67)	2 (0.45)	1 (0.32)	4 (1.44)	30 (1.55)*	13 (1.54)	10 (1.85)	7 (1.28)
Bilateral	26 (0.88)	7 (0.68)	2 (0.45)	1 (0.32)	4 (1.46)	19 (0.99)	5 (0.60)	4 (0.75)	10 (1.84)
Hip pain	55 (1.86)	6 (0.58)	3 (0.68)	0	3 (1.08)	49 (2.56)*	23 (2.77)*	11 (2.05)*	15 (2.75)
Symptomatic hip OA									
K/L $\geq$ 2	22 (0.75)	3 (0.29)	1 (0.23)	0	2 (0.72)	19 (0.99)*	8 (0.96)	5 (0.93)	6 (1.10)
K/L $\geq$ 3	19 (0.64)	2 (0.20)	1 (0.23)	0	1 (0.36)	17 (0.89)*	6 (0.72)	5 (0.93)	6 (1.10)

* $P < 0.05$ vs men in the corresponding group by chi-squared test.



(A) (B)

Fig. 1. (A) Prevalence (percentage) of subjects with radiographic hip osteoarthritis in each age stratum (<50, 50–59, 60–69, 70–79, ≥80). (B) Prevalence (percentage) of subjects overall with hip pain in each age stratum.

factors in the univariate model. The results were similar to the crude odds ratios.

When we considered hip pain in 5891 hips, we evaluated the association between K/L grade and hip pain in the designated hip. Figure 2 shows the percentage of subjects with hip pain in subgroups classified by radiographic hip OA severity: K/L = 0/1, K/L = 2, and K/L ≥ 3. In the overall population, the percentage of K/L = 0/1 was 0.75% (0.17% in men and 1.05% in women, respectively), that of K/L = 2 was 0.71% (0.78% and 0.64%, respectively), and that of K/L ≥ 3 was 36.2% (11.1% and 45.1%, respectively). In the urban community, the percentage of K/L = 0/1 was 0.79% (0.24% in men and 1.07% in women, respectively), that of K/L = 2 was 2.17% (0% and 4.65%, respectively), and that of K/L ≥ 3 was 42.1% (25.0% and 46.7%, respectively). In the mountainous community, the percentage of K/L = 0/1 was 0.40% (0% in men and 0.63% in women, respectively), that of K/L = 2 was 0%, and that of K/L ≥ 3 was 33.3% (0% and 40.0%, respectively). In the coastal community, the percentage of K/L = 0/1 was 1.08% (0.25% in men and 1.45% in women, respectively), that of K/L = 2 was 0.66% (1.47% and 0%, respectively), and that of K/L ≥ 3 was 34.4% (9.1% and 47.6%, respectively). Although the percentage with pain was positively correlated with radiographic severity, the difference between K/L = 2 and K/L ≥ 3 appeared to be greater than that between K/L = 0/1 and K/L = 2 in the overall population and all communities. Compared with K/L = 0/1, the OR of K/L ≥ 3 hip OA for hip pain was high, whereas that of K/L = 2 was not significantly associated with hip pain, even after adjustment for age, BMI, and community (Table IV).

Discussion

This is the first large-scale, population-based study to examine the prevalence of radiographic hip OA in Japanese men and women. We found that 15.7% of subjects had radiographic hip OA, 2.12% of subjects had severe radiographic hip OA, and 0.75% of subjects had symptomatic hip OA in at least one hip. We also examined the relation between the prevalence of radiographic hip OA, sex, and age. The present study showed factors associated with hip OA and the association of hip OA with hip pain.

Few studies have examined the prevalence of radiographic hip OA in Japan[3,15]. In 2000, Inoue *et al.* estimated the prevalence of K/L ≥ 3 hip OA among Japanese aged 20–79 years to be 1.4% in men and 3.5% in women, but their subjects were patients who underwent intravenous urography, who may not be representative of a general Japanese population. To the best of our knowledge, our previous study was the only population-based study to estimate the prevalence of hip OA among Japanese subjects; results showed that the prevalence of Croft grade ≥3 hip OA was 0% in men and 2% (95% CI 0.04–4.0) in women aged 60–79 years, but this study was published in 1998[3,16]. Because of the increasing number of elderly subjects in Japan, it is likely that these data have changed since our previous study. Furthermore, in Japan, previous studies showed only the prevalence of severe radiographic hip OA, but the prevalence of radiographic hip OA (e.g., K/L ≥ 2) was not reported. In the present study, we examined the prevalence of radiographic hip OA and severe radiographic hip OA using a large-scale, population-

Table III
Association factor for radiographic hip osteoarthritis and hip pain*

	Radiographic hip OA						Hip pain	
	K/L grade ≥2			K/L grade ≥3				
	No. of subjects (%)	Crude OR (95%Cl)	Adjust OR (95%Cl)	No. of subjects (%)	Crude OR (95%Cl)	Adjust OR (95%Cl)	No. of subjects (%)	Crude OR (95%Cl)
Age (+1 years)	–	1.00 (0.99–1.01)	–	–	0.98 (0.95–1.004)	–	–	1.02 (0.996–1.04)
BMI (kg/m^2)								
18.5≤, <25.0	297 (14.9)	Reference	Reference	37 (1.86)	Reference	–	33 (1.66)	Reference
<18.5	28 (13.1)	0.86 (0.56–1.28)	0.80 (0.51–1.22)	5 (2.34)	1.26 (0.43–2.97)	–	5 (2.34)	1.42 (0.48–3.36)
25.0≤, <27.5	74 (16.3)	1.11 (0.83–1.45)	1.09 (0.81–1.45)	9 (1.98)	1.07 (0.48–2.13)	–	12 (2.64)	1.61 (0.79–3.05)
≥27.5	66 (23.0)	1.70 (1.25–2.29)	1.83 (1.32–2.51)	10 (3.48)	1.91 (0.89–3.73)	–	5 (1.75)	1.06 (0.36–2.50)
Sex								
Men	189 (18.2)	Reference	Reference	13 (1.25)	Reference	Reference	6 (0.58)	Reference
Women	276 (14.5)	0.76 (0.62–0.93)	0.76 (0.62–0.95)	48 (2.51)	2.03 (1.13–3.92)	2.11 (1.17–4.09)	49 (2.57)	4.53 (2.09–11.85)
Community								
Urban	91 (7.18)	Reference	Reference	20 (1.58)	Reference	Reference	26 (2.06)	Reference
Mountainous	139 (16.2)	2.51 (1.90–3.32)	3.45 (2.59–4.62)	16 (1.87)	1.19 (0.60–2.30)	1.62 (0.81–3.19)	11 (1.29)	0.62 (0.29–1.23)
Coastal	235 (28.6)	5.16 (3.99–6.74)	10.08 (7.48–13.68)	25 (3.04)	1.95 (1.08–3.58)	3.47 (1.78–6.74)	18 (2.19)	1.07 (0.57–1.95)

* Adjusted odds ratios (ORs) were calculated by multiple logistic regression analysis after adjustment for all other variables in addition to regions. We show all variables we analyzed in the present study. K/L = Kellgren/Lawrence; 95%CI = 95% confidence interval; BMI = body mass index.

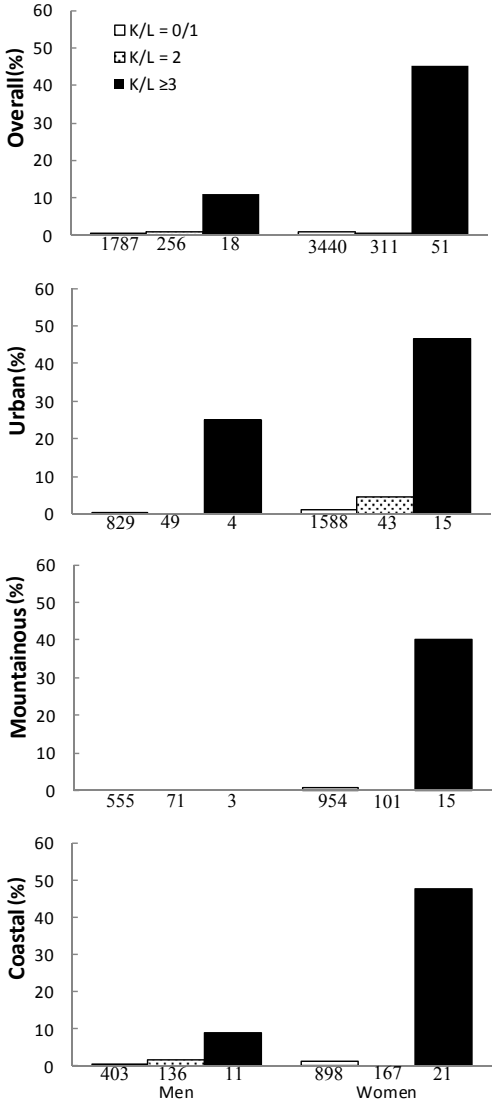


Fig. 2. Proportion (percentage) of subjects with hip pain in each subgroup classified by K/L grade in the overall population and communities. The number of subjects in each subgroup is shown under the bars. K/L, Kellgren/Lawrence.

based study and found that the prevalence of radiographic hip OA was 18.2% in men and 14.3% in women.

Although strict comparisons may be limited because the definitions of hip OA differ among studies and interobserver reliability for categorical methods is not good, the prevalence of hip OA in the present study is much lower than that seen in studies of Caucasians. In the Framingham study, the prevalence of K/L $\geq$ 2 hip OA was 24.7% (95% CI 20.6–28.7) and 13.6% (95%CI 10.7–16.5) in men and women, respectively[8]. The Johnston County prevalence study, a prevalence survey of a rural community in the United States, reported that the prevalence of K/L $\geq$ 2 hip OA was 27.6% (95%CI 26.3–28.9) and that of severe radiographic hip OA was 2.5% (95%CI 2.2–3.0)[6]; African Americans had a higher prevalence of hip OA than Caucasians. In the Rotterdam study, the prevalence of K/L $\geq$ 2 hip OA was 15.0% and that of K/L $\geq$ 3 hip OA was 4.3%[1]. In contrast, in a Beijing study, the prevalence of radiographic hip OA was 1.1% in men and 0.9% in women, which are similar or lower than values in the present study[7]. It is thought that the prevalence of hip OA is low in Asia[3,7,15,17,18], and that of severe radiographic hip OA is lower in Asians than in Caucasians; however, the presence of radiographic hip OA was not as low in the present study. These findings suggest that some ethnic factors may affect hip OA.

In the present study, coastal residency was significantly associated with radiographic hip OA, including severe radiographic hip OA, even after adjustment for age and BMI, indicating the involvement of environmental factors like nutrition or occupation. Both rural community backgrounds and farming have long been documented to be risk factors for hip OA. In England and India, rural male farmers were shown to have a higher risk of hip OA compared to rural male non-farmers[19,20]. The principle industries in the coastal residency are farming and fishing, which demand physical activity and repetitive laborious use of the hip joints, which may partly explain the higher prevalence of hip OA in the coastal region. We also found that the prevalence of radiographic hip OA was not associated with age in either gender. In the Copenhagen study, the prevalence of radiographic hip OA was age-dependent in both genders[4], whereas in the Beijing study, it slightly increased with age in men, but it did not increase with age in women[7]. These findings may also indicate a distinct etiology of hip OA among races. In addition, we also found that the prevalence of lumbar spondylosis (LS) and knee OA was significantly associated with age in the ROAD study[10,21], which may indicate that the etiology of hip OA may be different from that of LS and knee OA.

The association of gender with hip OA is controversial. Several studies in Caucasians showed that radiographic hip OA was more prevalent in men than in women[8,22], whereas in the Johnston County study and Rotterdam study, it was more prevalent in women than in men[1,6]. Previous studies in Japan showed that hip OA was significantly more prevalent in women than in men[3,15]. In the present study, interestingly, radiographic hip OA was more prevalent in men than in women, whereas, severe radiographic hip OA was more prevalent in women. In addition, the prevalence of radiographic hip OA was much higher than that of severe radiographic hip OA in the present study. This may be because a greater number of subjects in this study had osteophytosis than JSN. We have reported that osteophytosis of the lumbar spine was more

Table IV
Association of Kellgren/Lawrence grade with hip pain*

	Overall			Men			Women		
	No. of subjects (%)	Crude OR (95%CI)	Adjust OR (95%CI)	No. of subjects (%)	Crude OR (95%CI)	Adjust OR (95%CI)	No. of subjects (%)	Crude OR (95%CI)	Adjust OR (95%CI)
K/L grade									
0/1	39 (0.75)	Reference	Reference	3 (0.17)	Reference	Reference	36 (1.05)	Reference	Reference
2	4 (0.71)	0.9 (0.28–2.35)	1.36 (0.40–3.53)	2 (0.78)	4.68 (0.61–28.38)	4.50 (0.53–31.15)	2 (0.64)	0.6 (0.10–2.01)	0.79 (0.13–2.68)
≥3	25 (37.3)	80 (43.7–141.9)	123.4 (62.1–250.5)	2 (11.1)	74.3 (9.33–478.6)	57.3 (6.06–476.9)	23 (46.9)	83 (43.4–160.3)	129.1 (61.7–279.4)

We show all variables we analyzed in the present study. K/L = Kellgren/Lawrence; 95%CI = 95% confidence interval; BMI = body mass index.
* Adjusted odds ratios (ORs) were calculated by multiple logistic regression analysis after adjustment for age, BMI and communities.

prevalent in elderly Japanese men in the ROAD study[21]. In Japan, it appears that osteophytosis is more common in men than women[23,24]. However, this may indicate that the etiology of hip OA may be different from that of LS, because of the prevalence and the association were different between hip OA and LS. BMI was associated with radiographic hip OA, but not with severe radiographic hip OA in the present study. Several studies have reported that obesity has a low association with hip OA[7,18], whereas a multi-institutional study in Japan showed that obesity was a major cause for hip OA in women[25]. The discrepancy regarding the effect of obesity on hip OA may partly explain the distinct prevalence of various severities of hip OA[26].

Like the prevalence of severe radiographic hip OA, that of hip pain and symptomatic hip OA were low in both genders in the present study compared with previous studies, which showed that prevalence of hip pain was 7–40%, and that of symptomatic hip OA was 3–11%[1,5–8]. The present study also showed that the percentage of subjects with hip pain was less than 1% in subjects with K/L = 0/1 and 2, whereas it was more than 10% in men and more than 40% in women with K/L $\geq$ 3 hip OA. Furthermore, the OR of K/L $\geq$ 3 hip OA for hip pain was approximately 80 in both genders, which is much higher that of knee OA for knee pain in our previous study (K/L $\geq$ 3, OR 8.55, 95% CI 5.00–14.84 vs K/L = 0/1)[10]. This finding suggests that the prevalence of severe radiographic hip OA, hip pain, and symptomatic hip OA is low, but the association of hip pain with hip OA is much stronger than that for the knee.

Although the prevalence of radiographic hip OA was much higher than that of severe radiographic hip OA in the present study, the prevalence of symptomatic K/L $\geq$ 2 and K/L $\geq$ 3 hip OA was very low, and the difference in prevalence rates was small (0.75% and 0.64%, respectively). This finding indicates that subjects with K/L = 2 hip OA mostly did not have hip pain. This finding suggests that JSN, rather than osteophytosis, was associated with hip pain. We think that it is important to clarify the association of hip OA and hip pain to examine the prevalence of both K/L $\geq$ 2 and K/L $\geq$ 3.

There are several limitations to this study. First, regarding the selection bias of all participants of the ROAD study, we have already confirmed that participants of the ROAD study are representative of the Japanese population after comparison of anthropometric measurements and frequency of smoking and alcohol drinking between the participants and the general Japanese population. Thus, the values for the general population were obtained from the report on the 2005 National Health and Nutrition Survey conducted by the Ministry of Health, Labour and Welfare, Japan. No significant differences were identified between our participants and the total Japanese population, except that the male participants aged 70–74 years in the ROAD study were significantly smaller in terms of body structure than the overall Japanese population[12]. Unfortunately, we could not avoid the difference in the selection methods used in the three areas including the urban area, and both mountainous and coastal areas, performed during surveys in the ROAD study. Therefore, although coastal residency was significantly associated with radiographic hip OA in the present study, this factor might be affected by selection bias. Second, in the present report, we described the prevalence of hip OA with no mention of acetabular dysplasia.

In conclusion, this cross-sectional study using a large-scale population from the ROAD study clarified the prevalence of radiographic hip OA in Japanese men and women. The prevalence of radiographic hip OA was significantly higher in men than in women, but that of severe radiographic hip OA was significantly higher in women than in men and was not age-dependent in either gender. In addition, hip pain was strongly associated with K/L $\geq$ 3 hip OA. Further progress, along with continued longitudinal surveys of the ROAD study, will elucidate the backgrounds of hip OA and its relation with hip pain.

Author contributions

All authors have made substantial contributions to all three of the following sections:

(1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data;
(2) drafting the article or revising it critically for important intellectual content; and
(3) final approval of the version to be submitted.

Conflict of interests

There are no conflicts of interest.

Role of funding source

This work was supported by a Grant-in-Aid for H17-Men-eki-009 (Director, Kozo Nakamura), H20-Choujyu-009 (Director, Noriko Yoshimura), H23-Choujyu-002 (Director, Toru Akune), H-25-Choujyu-007 (Director, Noriko Yoshimura), and H25-Nanchitou (Men)-005 (Director, Sakae Tanaka) of the Ministry of Health, Labour and Welfare; and Scientific Research B23390172, B20390182, and Challenging Exploratory Research 24659317 to Noriko Yoshimura; B23390356, C20591774, and Challenging Exploratory Research 23659580 to Shigeyuki Muraki; Challenging Exploratory Research 24659666 and 21659349 and Young Scientists A18689031 to Hiroyuki Oka; B23390357 and C20591737 to Toru Akune; and Collaborating Research with NSF 08033011-00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. This study also was supported by grants from the Japan Osteoporosis Society (2006-1) (Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, and Toru Akune), and research aid from the Japanese Orthopaedic Association (2010-2) (JOA-Subsidized Science Project Research 2006-1 & 2010-2; Director, Hiroshi Kawaguchi).

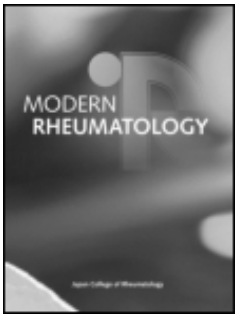
Acknowledgments

The authors wish to thank Dr Takako Nojiri and Mr Kazuhiro Hatanaka of the Gobo Public Health Centre; Dr Naoki Hirabayashi of the Kawakami Clinic in Hidakagawa Town; Mrs Tomoko Takijiri, Mrs Rie Takiguchi, Mrs Kyoko Maeda, and other members of the public office in Hidakagawa Town; Dr Shinji Matsuda of the Shingu Public Health Centre; and Mrs Tamako Tsutsumi, Mrs Kanami Maeda, Mrs Megumi Takino, Mrs Shuko Okada, Mrs Kazuyo Setoh, Mrs Chise Ryouno, Mrs Miki Shimosaki, Mrs Chika Yamaguchi, Mrs Yuki Shimoji, and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. We also thank Ms Kyoko Yoshimura, Mrs Toki Sakurai, Mrs Saeko Sahara, and Mr Noriyuki Oe for their assistance in data reduction and administration.

References

1. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57:203–8.
2. Lane NE, Nevitt MC, Hochberg MC, Hung YY, Palermo L. Progression of radiographic hip osteoarthritis over eight years in a community sample of elderly white women. Arthritis Rheum 2004;50:1477–86.

3. Yoshimura N, Campbell L, Hashimoto T, Kinoshita T, Okayasu T, Wilman C, *et al*. Acetabular dysplasia and hip osteoarthritis in Britain and Japan. Br J Rheumatol 1998;37: 1193–7.
4. Jacobsen S, Sonne-Holm S, Søballe K, Gebuhr P, Lund B. Radiographic case definitions and prevalence of osteoarthrosis of the hip: a survey of 4151 subjects in the Osteoarthritis Substudy of the Copenhagen City Heart Study. Acta Orthop Scand 2004;75:713–20.
5. Reijman M, Hazes JM, Pols HA, Bernsen RM, Koes BW, Bierma-Zeinstra SM. Validity and reliability of three definitions of hip osteoarthritis: cross sectional and longitudinal approach. Ann Rheum Dis 2004;63:1427–33.
6. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, *et al*. Prevalence of hip symptoms and radiographic and symptomatic hip osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol 2009;36:809–15.
7. Nevitt MC, Xu L, Zhang Y, Lui LY, Yu W, Lane NE, *et al*. Very low prevalence of hip osteoarthritis among Chinese elderly in Beijing, China, compared with whites in the United States: the Beijing Osteoarthritis Study. Arthritis Rheum 2002;46:1773–9.
8. Kim C, Linsenmeyer KD, Vlad SC, Guermazi A, Clancy MM, Niu J, *et al*. Prevalence of radiographic and symptomatic hip osteoarthritis in an urban United States community: the Framingham osteoarthritis study. Arthritis Rheumatol 2014;66:3013–7.
9. van Dijk GM, Dekker J, Veenhof C, van den Ende CH, Carpa Study Group. Course of functional status and pain in osteoarthritis of the hip or knee: a systematic review of the literature. Arthritis Rheum 2006;55:779–85.
10. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, *et al*. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthritis Cartilage 2009;17:1137–43.
11. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-yo Y, Yoshida M, *et al*. Prevalence of knee osteoarthritis, lumbar spondylosis and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Min Metab 2009;27:620–8.
12. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on Osteoarthritis/osteoporosis Against Disability (ROAD) study. Int J Epidemiol 2010;39: 988–95.
13. Shimada H, Lord SR, Yoshida H, Kim H, Suzuki T. Predictors of cessation of regular leisure-time physical activity in community-dwelling elderly people. Gerontology 2007;53: 293–7.
14. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
15. Inoue K, Wicart P, Kawasaki T, Huang J, Ushiyama T, Hukuda S, *et al*. Prevalence of hip osteoarthritis and acetabular dysplasia in French and Japanese adults. Rheumatology (Oxford) 2000;39:745–8.
16. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132:514–22.
17. Lau EM, Lin F, Lam D, Silman A, Croft P. Hip osteoarthritis and dysplasia in Chinese men. Ann Rheum Dis 1995;54:965–9.
18. Felson DT, Zhang Y. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41:1343–55.
19. Croft P, Coggon D, Cruddas M, Cooper C. Osteoarthritis of the hip: an occupational disease in farmers. BMJ 1992;304: 1269–72.
20. Thelin A, Holmberg S. Hip osteoarthritis in a rural male population: a prospective population-based register study. Am J Ind Med 2007;50:604–7.
21. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, *et al*. Prevalence of radiographic lumbar spondylosis and its association with low back pain in elderly subjects of population-based cohorts: the ROAD study. Ann Rheum Dis 2009;68:1401–6.
22. Jacobsen S, Sonne-Holm S, Søballe K, Gebuhr P, Lund B. Hip dysplasia and osteoarthrosis: a survey of 4151 subjects from the Osteoarthrosis Substudy of the Copenhagen City Heart Study. Acta Orthop 2005;76:149–58.
23. O'Neill TW, McCloskey EV, Kanis JA, Bhalla AK, Reeve J, Reid DM, *et al*. The distribution, determinants, and clinical correlates of vertebral osteophytosis: a population based survey. J Rheumatol 1999;26:842–8.
24. Yoshimura N, Dennison E, Wilman C, Hashimoto T, Cooper C. Epidemiology of chronic disc degeneration and osteoarthritis of the lumbar spine in Britain and Japan: a comparative study. J Rheumatol 2000;27:429–33.
25. Jingushi S, Ohfuji S, Sofue M, Hirota Y, Itoman M, Matsumoto T, *et al*. Multiinstitutional epidemiological study regarding osteoarthritis of the hip in Japan. J Orthop Sci 2010;15:626–31.
26. Cooper C, Inskip H, Croft P, Campbell L, Smith G, McLaren M, *et al*. Individual risk factors for hip osteoarthritis: obesity, hip injury, and physical activity. Am J Epidemiol 1998;147: 516–22.



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

The effect of cartilage degeneration on ultrasound speed in human articular cartilage

Satoru Ohashi, Isao Ohnishi, Hiroyuki Oka, Takuya Matsumoto, Masahiko Bessho, Kozo Nakamura & Sakae Tanaka

To cite this article: Satoru Ohashi, Isao Ohnishi, Hiroyuki Oka, Takuya Matsumoto, Masahiko Bessho, Kozo Nakamura & Sakae Tanaka (2016) The effect of cartilage degeneration on ultrasound speed in human articular cartilage, Modern Rheumatology, 26:3, 426-434, DOI: 10.3109/14397595.2015.1097012

To link to this article: http://dx.doi.org/10.3109/14397595.2015.1097012

Accepted author version posted online: 22 Sep 2015.
Published online: 03 Nov 2015.

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://informahealthcare.com/mor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2016; 26(3):426–434

DOI: 10.3109/14397595.2015.1097012

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

The effect of cartilage degeneration on ultrasound speed in human articular cartilage

Satoru Ohashi[1,2], Isao Ohnishi[1], Hiroyuki Oka[3], Takuya Matsumoto[1], Masahiko Bessho[1], Kozo Nakamura[1], and Sakae Tanaka[1]

[1]*Department of Sensory & Motor System Medicine, Faculty of Medicine, The University of Tokyo, Tokyo, Japan,* [2]*Department of Orthopaedic Surgery, Sagamihara Hospital, National Hospital Organization, Sagamihara, Japan, and* [3]*Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*

Abstract

Objectives: We investigated the effect of cartilage degeneration on ultrasound speed in human articular cartilage *in vitro*.
Methods: Ultrasound speed was calculated by the time-of-flight method for 22 femoral condyle osteochondral blocks obtained from osteoarthritis patients. In parallel, histological evaluation of specimens was performed using the modified Mankin and OARSI scores.
Results: The mean ultrasound speed was 1757 ± 109 m/s. Ultrasound speed showed significant negative correlation with OARSI score, and a decreasing tendency with high Mankin scores. Good correlation was found between the optically measured and the calculated cartilage thickness.
Conclusion: Our results show that articular cartilage degeneration has relatively little influence on ultrasound speed. In addition, morphological evaluation of articular cartilage using a preset value of ultrasound speed seems to offer relatively accurate results.

Keywords

Cartilage degeneration, Cartilage thickness, Osteoarthritis, Ultrasonography, Ultrasound speed

History

Received 19 February 2015
Accepted 14 September 2015
Published online 30 October 2015

Introduction

Osteoarthritis (OA) of the knee is a condition characterized by morphological, biochemical, molecular, and biomechanical changes in both cells and the extracellular matrix, resulting in softening, fibrillation, ulceration, and eventual loss of articular cartilage [1]. In clinical practice, plain radiography is typically used to evaluate the stage of OA [2,3]. However, this method does not allow direct imaging of the cartilage, because it only evaluates the distance between the femoral and tibial bone surfaces, and the presence of osteophytes and sclerosis of the subchondral bone. Direct imaging of cartilage has been achieved using magnetic resonance imaging (MRI), which allows morphological evaluation of articular cartilage, including the determination of cartilage thickness and volume [4], and identification of cartilage degeneration [5].

In addition to MRI, ultrasonography has also been investigated for applications allowing the direct evaluation of articular cartilage, including degenerative changes in cartilage [6] and cartilage surface roughness [7]. Ultrasonography was also used in previous investigations to visualize articular cartilage and evaluate cartilage thickness, either directly on the surface of cartilage [8,9] or percutaneously [10–12]. In these studies, the set-up speed value of the diagnostic ultrasound device (1540 m/s) was used for the calculation of cartilage thickness [13]. Theoretically, however, for quantification of cartilage thickness or volume using ultrasonography, the actual ultrasound speed in each articular cartilage should be measured, since the speed of sound might differ among tissues, and thus affect the calculations [14].

Studies have been performed in articular cartilage to investigate the effect of degeneration and other factors on ultrasound speed, mostly using animal samples [15]. These studies have shown that the speed of sound in cartilage can be affected by composition [16,17], material properties [17,18], or mechanical strain [19–21], as well as by orientation of collagen fibrils [22] or anisotropy [23] of articular cartilage. Cartilage ultrasound speed can also be affected by external factors, such as the ultrasound beam angle against the cartilage surface [24], and temperature or saline concentration [23].

Some studies have investigated ultrasound speed in human articular cartilage. Based on experimental results on bovine cartilage and the results of a previous study on human cartilage, Toyras et al. [17] performed simulations investigating the relationship between the speed of sound, cartilage thickness, and the error in dynamic modulus; they suggested that a constant speed of sound can be utilized to obtain a clinically acceptable accuracy for cartilage thickness and modulus. However, relatively variable mean values for ultrasound speed have been reported in human articular cartilage (1658 m/s [25], 1892 m/s [26], ca. 1580 m/s [20]). In bovine cartilage, ultrasound speed decreases as the cartilage degenerates through chemical treatment [17]. In addition, ultrasound speed in cartilage of OA patients was reported to be lower than in normal cartilage [25]. Since it would be difficult to measure a patient-specific value of ultrasound speed in cartilage and apply this value for each patient during clinical morphological

Correspondence to: Satoru Ohashi, Department of Sensory & Motor System Medicine, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan and Department of Orthopaedic Surgery, Sagamihara Hospital, National Hospital Organization, 18-1 Sakura-dai, Minami-ku, Sagamihara, Kanagawa 252-0315, Japan. Tel: +81 427428311. Fax: +81 427425314.
E-mail: soohashi-tky@umin.ac.jp

evaluation of cartilage, the relationship between ultrasound speed and the degree of degeneration in human cartilage warrants further investigation.

The aim of this study was to perform measurements of ultrasound speed, histologically score the degeneration in human cartilage samples, and to investigate the correlation between cartilage degeneration and ultrasound speed, in order to investigate the feasibility of using a constant value of speed in morphological evaluation of articular cartilage by ultrasound.

Methods

Cartilage sample preparation

All procedures used in this investigation were approved by the institutional review board at our university. Subjects comprised 11 OA patients who planned to undergo total knee arthroplasty and provided written informed consent prior to participation in the study. All the patients were female, with an average age of 73.2 ± 8.0 years (range: 56–83 years). Pre-operation plain radiographs showed that the Kellgren–Lawrence score [27] of all the patients was grade 4. Osteochondral blocks removed from the medial and lateral femoral condyles during operation were wrapped in gauze moistened with normal saline, packed in plastic bags, manually degassed, hermetically sealed, and stored at $-60\,^{\circ}$C. A total number of 22 osteochondral blocks from femoral condyles were acquired from the patients through operations. On the day of the experiment, the osteochondral samples were thawed in normal saline solution (Otsuka Pharmaceutical, Tokyo, Japan) at room temperature (20 °C). Osteochondral blocks from the femoral condyle were trimmed by a band saw (SWD-250; Fuijiwara Sangyo, Miki, Japan), achieving a surface size of approximately 18 mm × 18 mm for cartilage samples. Trimming was performed to obtain a sample containing sufficient quantities of cartilage for the acoustic and microscopic measurements, preferably from the part of the block closest to the weight-bearing area. Samples were then fixed on a custom-made acryl sample holder (30 mm × 30 mm × 13 mm; Murai & Co., Tokyo, Japan) with resin (GC-Ostron; GC Corporation, Tokyo, Japan) (Figure 1). During preparation, samples were continuously cooled at 20 °C and moistened using normal saline solution.

Acoustic measurements

Acoustic measurements were performed using a custom-made apparatus (Figure 1). The acryl holder with the human osteochondral block affixed was positioned in a water tank containing normal saline (20 °C), so that the cartilage surface faced upward. A stage underneath, with three micrometers (accuracy, 10 μm), allowed horizontal movement of the sample. Two micrometers, perpendicular to each other in the horizontal plane, were used for position adjustment by linear movement (*x*- and *y*-axes). The third micrometer enabled circular movement in the horizontal plane (rotation movement in the *x*–*y*-plane). An ultrasound transducer was placed over the sample in the water tank, and the holder of the transducer had a *z*-adjustment device so that the distance between the cartilage surface and the transducer could be kept at the transducer focus distance (2.5″=63.5 mm).

Ultrasound measurements were performed using the A-mode pulse-echo method and a focused non-contact ultrasound transducer (V311-SU; Olympus NDT, Waltham, MA) (center frequency = 7.3 MHz, 3.4–11.2 MHz, −3 dB; transducer tip diameter = 16 mm; element diameter = 13 mm; radius of curvature = 63.5 mm). Acoustic pulses were excited electrically using a pulser/receiver board (NDT-5800; Olympus NDT). Echoes of the transmitted pulse were recorded with the transducer and pulser/receiver board. A bandpass filter (1.0–20.0 MHz) was used to enhance the ultrasound signal-to-noise ratio. The signal was digitized at a 1000-MHz sampling frequency using an oscilloscope (DPO4034; Tektronix Japan, Tokyo, Japan).

For acoustic measurements, the edges and the center point of the 30 mm × 30 mm acryl sample holder surface were first identified by moving the stage horizontally under the fixed ultrasound beam. The cartilage surface was then scanned with the ultrasound transducer by moving the stage to identify the top cartilage surface point (point C) (Figure 2). The ultrasound beam was, theoretically, perpendicular to the cartilage surface at this point, as the cartilage of the femoral condyle has a convex surface. After identifying the coordinates for this point as (a, b), two additional points at 1 mm apart on each side of this point were set as radiofrequency signal acquisition points, along with point *C*. The *x*–*y* coordinates of the two points were thus (a+1, b), (a−1, b) using units of 1 mm.

RF signals at these three points were acquired and output from the oscilloscope device as comma-separated values data. Time of flight was measured in each sample using the peak envelope method previously described [28] (Figure 3). The envelope of each RF signal was calculated using a Hilbert transform [29]. Peaks of the envelope signal were attributed to reflections occurring at the cartilage surface and at the cartilage-bone interface. Time of flight was defined as the duration (Δt) between peaks, corresponding to the travel time of the ultrasound pulse back and forth between the cartilage surface and the cartilage-bone interface of the specimen.

Microscopic optical thickness measurement and calculation of ultrasound speed

In order to measure cartilage thickness, direct optical measurement using microscopy was performed on a cross-section of the sample. The acryl holder with the osteochondral sample was attached to the holding arm of a diamond saw device (Minitom; Struers, Cleveland, OH) such that the saw blade was vertical to the holder top surface, that is, vertical to the *x*-*y* plane of the sample coordinates and parallel to the *y*-axis. By adjusting the position of the arm within an accuracy of 10 μm, cut planes were created, each containing 3 RF signal acquisition points. Subsequently, each cut sample was mounted on a glass slide and covered with a cover glass after dripping normal saline onto the sample surface, to keep the cartilage moist and inhibit deformation due to drying during measurement.

Cartilage thickness [4] was measured using an optical measuring microscope (×30 magnification) (MM-400; Nikon, Tokyo, Japan) containing an eyepiece with adjustable crosshairs, and an adjustable stage system (MHS 2 × 2; Nikon) (Figure 4). With the optical measuring microscope and the stage, the center point of the sample holder could be identified by measuring the distance from both edges of the sample holder, and then the RF signal acquisition points could be determined in a similar manner. The microscope could also align the sides of the sample holder, which were parallel to the direction of the ultrasound beam in RF signal acquisition, to the direction of thickness measurement. After these adjustments, cartilage thickness (d_C) along the beam direction was measured at each RF acquisition point, and the speed of sound in cartilage (SOS_C) at each point was calculated as follows:

$$SOS_C = \frac{2d_C}{\Delta t}. \quad (1)$$

Histological evaluation

Each osteochondral sample was fixed in 4% paraformaldehyde phosphate buffer solution (Wako Pure Chemical Industries, Osaka,

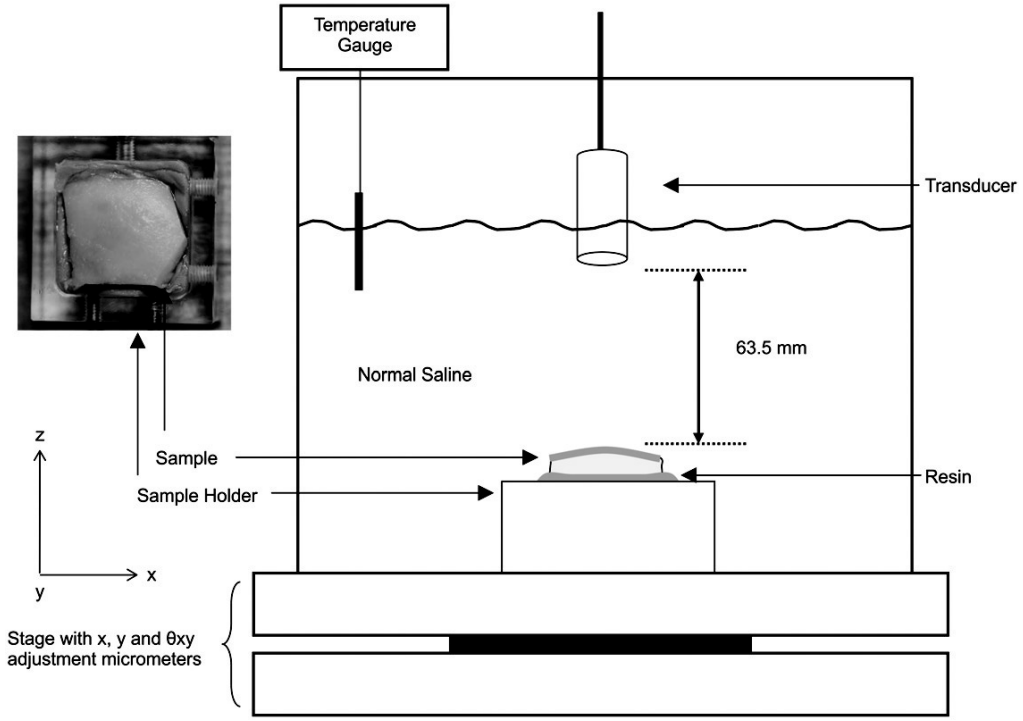


Figure 1. Custom-made apparatus for acoustic measurements. A human cartilage sample with subchondral bone was immersed in normal saline and fixed on the sample holder by resin. The water tank has a stage underneath with three micrometers (*x*-, *y*- directions and rotation movement in the *x*–*y* plane) to allow horizontal movement of the sample.

Japan) for 4 days, followed by decalcification with Plank-Rychlo's Solution composed of 0.3 M aluminum chloride, 3% hydrochloric acid, and 5% formic acid for 36 h. After decalcification, all specimens were dehydrated with ethanol, embedded in paraffin and sectioned by microtome with a thickness of 4 μm. Fast Green and Safranin O stainings were performed, and specimens were histologically evaluated using the modified Mankin score [30,31] and the Osteoarthritis Research Society International (OARSI) score [32] by two well-experienced examiners (Tables 1 and 2) (Figure 5). Histological evaluation was carried out twice by each examiner with an interval of two weeks and the mean score was used for statistical analysis.

Statistical analysis

SOS_C was defined as the mean ultrasound speed of the three acoustic measurement points in each sample. In order to assess the reliability of the histological evaluation, intraclass correlation coefficients (ICCs) comparing the first and second histological scores of each examiner were evaluated for intraobserver reliability. In addition, ICC calculation and linear regression analysis were performed to assess interobserver reliability, comparing the mean of the first and second histological scores of the specimens between the two examiners.

Spearman's rank correlation coefficient between SOS_C and the histological scores of the first examiner's first scoring as well as the correlation coefficient between SOS_C and d_C were calculated to investigate the influence of cartilage degeneration and cartilage thickness on ultrasound speed. Correlation analysis was also performed between d_C and histological scores to investigate the degree of confounding between them. In addition, to investigate the feasibility of using a preset value of ultrasound speed in thickness measurements of articular cartilage using ultrasound, linear regression analysis and Bland–Altman plot analysis were performed between optical thickness measurement values

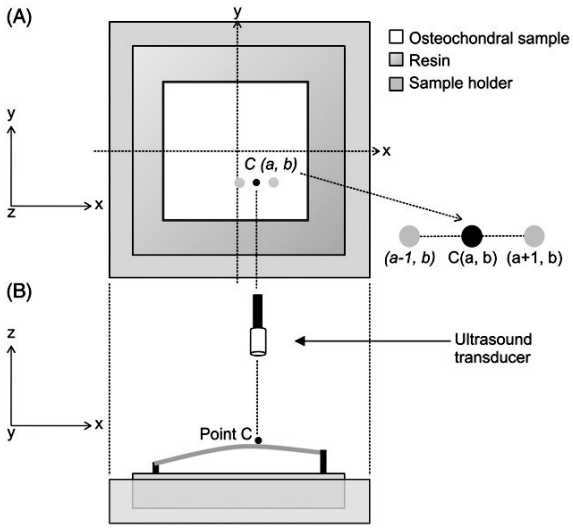


Figure 2. (A) The cartilage surface point closest to the transducer (point *C*) was acoustically identified. (B) With point *C* as the center point, radiofrequency signals were acquired at three points, each 1 mm apart. Units in the figure are 1 mm.

(d_C-optical) and thickness values calculated from time of flight using the average ultrasound speed of this study (d_C-US).

Statistical analysis was performed using IBM SPSS Statistics version 21.0 software (IBM, Armonk, NY), and results were considered significant for values of $p < 0.05$.

Results

In all RF signals, peaks of the reflected ultrasound wave envelopes from the cartilage surface and the cartilage-bone interface were clear enough to be identified. The mean SOS_C of all articular

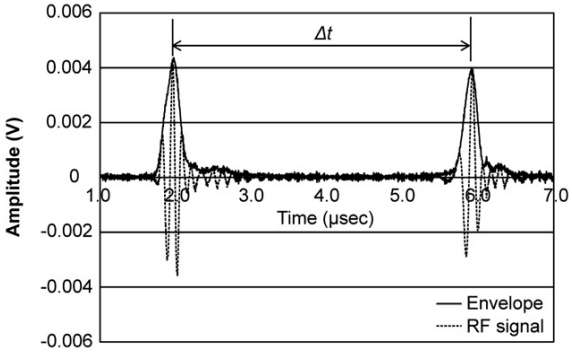


Figure 3. The graph shows an example of the radiofrequency [30] signal wave and the envelope wave calculated from the RF signal. Time of flight (Δt) was defined as the duration between peaks of the envelope wave.

cartilage samples was 1757 ± 109 m/s. The mean standard deviation calculated from the standard deviation of the three-point ultrasound speed values of individual samples was 55.2 m/s. The mean coefficient of variance calculated from each sample's SOS_C and standard deviation of the three-point ultrasound speed values was 3.2%.

ICCs for intraobserver reliability of examiner 1 and examiner 2 were 0.888 [95% confidence interval (CI), 0.753–0.952] and 0.914 (95% CI, 0.807–0.963) [overall, 0.904 (95% CI, 0.832–0.947)] for the modified Mankin score, and 0.927 (95% CI, 0.834–0.969) and 0.945 (95% CI, 0.874–0.977) [overall, 0.935 (95% CI, 0.885–0.964)] for the OARSI score, respectively. ICCs for interobserver reliability were 0.717 (95% CI, 0.438–0.871) for the modified Mankin score, and 0.965 (95% CI, 0.919–0.985) for the OARSI score. Significant linear correlation was noted between the histological scores of the two examiners by linear regression analysis ($r = 0.783$; root mean square error, 1.87; $p < 0.01$; slope, 1.24 for the modified Mankin score and $r = 0.967$; root mean square error, 0.310; $p < 0.01$; slope, 1.05 for the OARSI score).

The scatter plots for SOS_C and histological scores are shown as Figure 6. SOS_C showed a decreasing tendency with high modified Mankin scores ($r = -0.330$; $p = 0.134$), and significantly correlated with the OARSI score ($r = -0.483$; $p < 0.05$). In addition, SOS_C showed a significant positive correlation with cartilage thickness ($r = 0.484$, $p < 0.05$). There were no significant correlations between cartilage thickness and the modified Mankin score ($r = -0.253$; $p = 0.256$) or OARSI score ($r = -0.420$; $p = 0.052$).

Using the average SOS_C value, linear regression analysis showed a significant correlation between cartilage thickness measured optically and cartilage thickness calculated by time of flight (Figure 7A) ($r = 0.959$; root mean square error, 0.194 mm; $p < 0.01$; slope, 1.053). Bland–Altman plots showed a mean difference of 0.0478 mm with a standard deviation of 0.188 mm between d_C-optical and d_C-US (Figure 7B).

Discussion

Several studies have measured ultrasound speed in human articular cartilage, reporting a relatively wide range of values (1658 and 1581 m/s for normal and OA femoral cartilage, respectively [25]; 1892 m/s for the ankle joint and hip joint cartilage of one patient [26]; and ca. 1580 m/s for patellar cartilage [20]). Since cartilage degeneration has been reported to influence ultrasound speed in articular cartilage in animal studies, degeneration might be one of the reasons behind the observed differences [17,33]. We performed ultrasound speed measurements in human articular cartilage and investigated the influence of cartilage degeneration on ultrasound speed. As a result, we obtained a mean ultrasound speed of 1757 m/s, which is comparable to values reported for human articular cartilage in a previous study [26], but is higher than measurements in two other studies [20,25], including one conducted on femoral cartilage. A possible reason accounting for this discrepancy could be swelling of the cartilage during cross-sectioning. Moreover, the cartilage sample preparation steps, such as freezing, storage, thawing, and immersion in saline, could also have contributed to the discrepancy. Although we confirmed that cartilage thickness did not change after cross-sectioning, by covering the cross-section surface with a cover glass and performing the same procedures on all the samples, we cannot exclude the possibility that swelling of the cartilage during cross-sectioning, or change of propagation properties through sample preparation, might have occurred, resulting in higher ultrasound speed values.

Since we wanted to evaluate the reliability of our method on human cartilage, we performed cartilage thickness measurements by acquisition of RF signals at three points. The mean standard deviation and the mean coefficient of variance calculated for each osteochondral sample were relatively low (55.2 m/s and 3.2%, respectively) compared with the coefficient of variance of this method published for animal cartilage (3.4% for a 6-month-old pig and 6.4% for a 3-year-old pig) [28]. However, although we validated the accuracy of the cartilage thickness measurements by cross-sectioning using the custom-made devices described in a previous study involving micro-CT [28], it would be ideal to use a less invasive method, such as the needle probe method [33–35] or the custom-made ultrasound probe method [20,21], with which more ultrasound speed measurement points can be acquired and SOS_C could be more accurate.

Ultrasound speed showed a significant negative correlation with OARSI scores used for the histological evaluation, decreasing with higher degrees of cartilage degeneration. The present study is the first to report these findings in human cartilage samples. Ultrasound speed also decreased with cartilage degeneration assessed by the modified Mankin score, although the trend was not significant. The trend between the ultrasound speed and cartilage degeneration was compatible with results of previous studies on animal cartilage [17,33]. Treatment of bovine articular cartilage with trypsin for 4 h, resulting in the digestion of proteoglycan and minor cleavage of collagen, decreased ultrasound speed [33]. In bovine articular cartilage samples obtained from different locations, ultrasound speed decreased with Mankin score and water content but increased with uronic acid and hydroxyproline levels [17]. Nevertheless, a constant speed of sound was suggested to provide a clinically acceptable accuracy for cartilage thickness (error: 7.8%) in that study.

Several factors could have affected ultrasound speed in the present study. Uronic acid and hydroxyproline levels have been reported to be lower in degenerated cartilage than in normal cartilage [36]. Amide I-rich areas in the superficial layer and carbohydrate-rich areas in the whole layer have been observed to be decreased in the human OA samples [37]. These factors might have caused changes in the acoustic properties of cartilage with age, as was also observed in a study using rat articular cartilage [38].

Instead of evaluating individual components of cartilage degeneration, we performed histological scoring, in order to ensure that we investigate the overall effect of cartilage degeneration on ultrasound speed. Mankin score has been previously reported to negatively correlate with the uronic acid and hyaluronic acid content of articular cartilage [39]. In the present study, however, the OARSI score showed a better correlation with ultrasound speed than the modified Mankin score, which we believe is an interesting finding of the two different histological

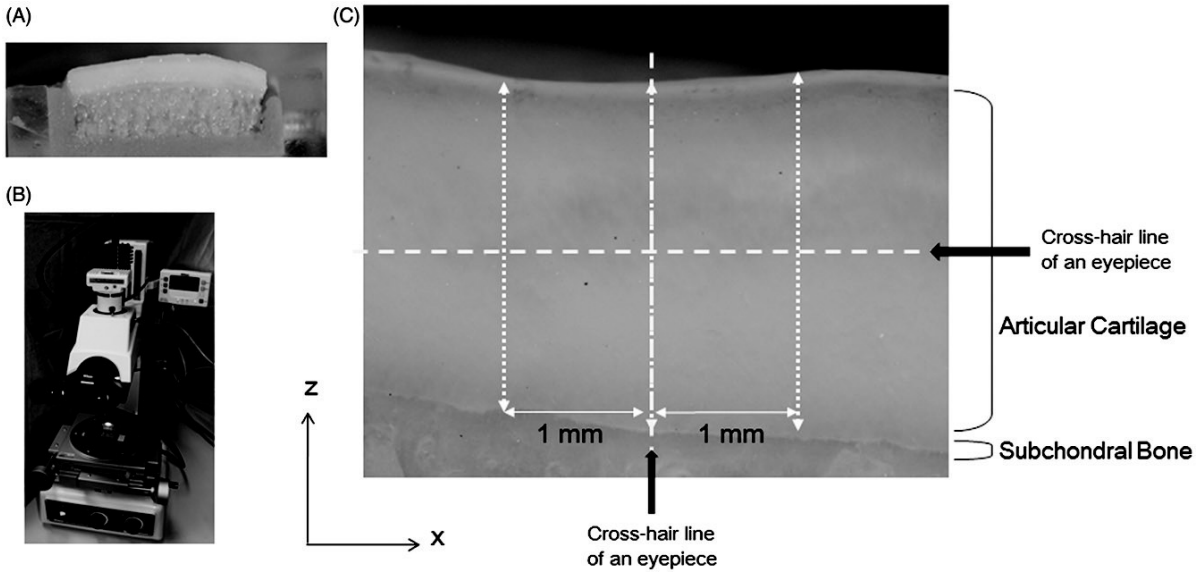


Figure 4. Images showing cartilage thickness measurement using a microscope. After registration of the RF signal acquisition points on the articular cartilage, the cut plane was created (A), containing three measurement points (B). Cartilage thickness was measured optically using a microscope (C) at the RF signal acquisition points.

Table 1. Modified Mankin score.

		Grade
I	Structure	
	Normal	0
	Surface irregularities	1
	Pannus and surface irregularities	2
	Clefts to transitional zone	3
	Clefts to radial zone	4
	Clefts to calcified zone	5
	Complete disorganization	6
II	Cells	
	Normal	0
	Diffuse hypercellularity	1
	Cloning	2
	Hypocellularity	3
III	Safranin-O staining	
	Normal	0
	Slight reduction	1
	Moderate reduction	2
	Severe reduction	3
	No dye noted	4

evaluations. A possible reason for this discrepancy could be that the OARSI score comprises not only a qualitative evaluation of articular cartilage, but also evaluation of morphological damage, a feature of advanced cartilage degeneration. In contrast, the modified Mankin score does not contain evaluation of morphological change and captures relatively early degenerative changes of articular cartilage. Indeed, we found that ultrasound speed showed a significant positive correlation with cartilage thickness, and that cartilage thickness did not correlate with the histological scores. We assume that not only cartilage degeneration, but also cartilage wear, which generally occurs in advanced OA, could have influenced the ultrasound speed. However, articular cartilage thickness can differ even between healthy individuals [40]. Thus, we assumed that cartilage wear or decrease in cartilage thickness could not be quantified in the patients in the present study because the original cartilage thickness (i.e. before OA had started) is unknown in each patient, and the positive correlation between the ultrasound speed and cartilage thickness in this study did not prove the correlation between ultrasound speed and cartilage wear.

In the present study, we found that both the modified Mankin score and the OARSI score were precise and reliable, as judged by intraobserver and interobserver reliability values, corroborating the findings of previous studies [31,41–43]. The correlation coefficient between the two scoring systems was 0.942 ($p < 0.001$), but ICCs for both intraobserver reliability and interobserver reliability were lower for the modified Mankin score than for the OARSI score. The OARSI score covers a relatively wide range of cartilage change, from early to advanced degeneration, while the modified Mankin score evaluates relatively early degenerative changes of articular cartilage. Thus, samples showing advanced degeneration might have resulted in a lower reliability for the modified Mankin score.

In a study using animal cartilage samples [17], a constant speed of sound was shown to provide a clinically acceptable accuracy for cartilage thickness. In addition, a good correlation ($r = 0.78$) was observed between the cartilage thickness calculated acoustically and the thickness measured optically in a study using human osteochondral samples [25]. Our results show an even better correlation ($r = 0.959$) between these values, although this might be due to differences in patient populations. Ultrasound intensity of the cartilage surface has been reported to significantly decrease as degeneration or OA develops, both in animals and in humans, and is suggested to have the potential to detect early osteoarthritic changes at the preclinical stage [37]. In the present study, ultrasound speed had a significant correlation with the OARSI score but not with the modified Mankin score. In addition, since it is technically difficult to measure the ultrasound speed in cartilage and apply this value for each patient during clinical morphological evaluation of cartilage, using a specific preset value of ultrasound speed seems justifiable based on our findings.

An MRI study on OA patients with OARSI grade 1, 2, and 3 medial joint space narrowing (JSN) has shown a reduced cartilage thickness (with differences of 0.190, 0.630, and 1.560 mm in the respective groups) in weight-bearing medial femorotibial compartments compared to cartilage in knees without JSN [44]. In addition, the mean annual loss of cartilage thickness in the center of the medial femoral condyle was over 0.180 mm in the grade 2 and 3 patient groups [45]. Clinical morphological evaluation of

Table 2. OARSI score.

Grade (key feature)	Subgrade
Grade 0: surface intact, cartilage morphology intact	No subgrade
Grade 1: surface intact	1.0 Cells Intact
	1.5 Cell death
Grade 2: surface discontinuity	2.0 Fibrillation through superficial zone
	2.5 Surface abrasion with matrix loss within superficial zone
Grade 3: vertical fissures (clefts)	3.0 Simple fissures
	3.5 Branched/complex fissures
Grade 4: erosion	4.0 Superficial zone delamination
	4.5 Mid zone excavation
Grade 5: denudation	5.0 Bone surface intact
	5.5 Reparative tissue surface present
Grade 6: deformation	6.0 Joint margin osteophytes
	6.5 Joint margin and central osteophytes

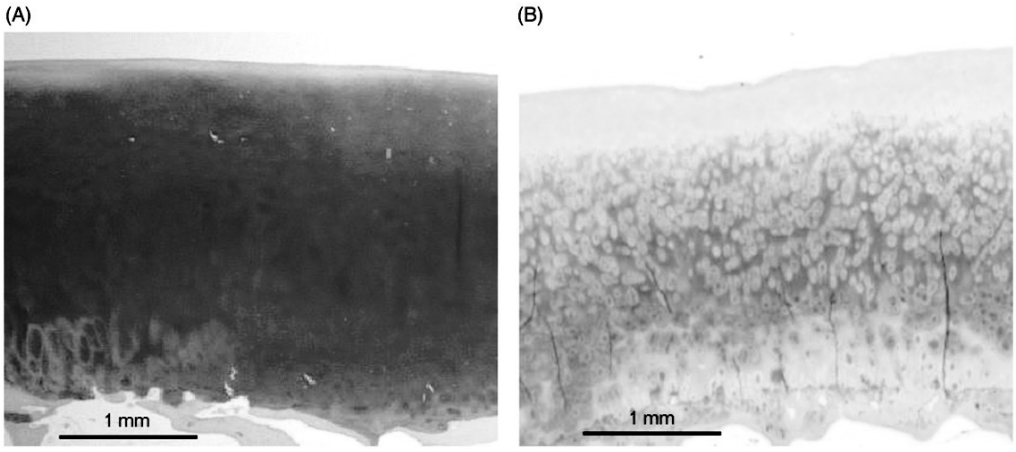


Figure 5. Representative images of histological sections stained with Fast Green and Safranin O. (A) Relatively healthy cartilage exhibits slight reduction in Safranin O staining. Histological scores were graded as 2 based on the modified Mankin score and 1 based on the OARSI score. (B) Moderately degenerated cartilage exhibits pannus/surface irregularities, diffuse hypercellularity, and moderate reduction in Safranin O staining. Histological scores were graded as 5 based on the modified Mankin score and 2.5 based on the OARSI score.

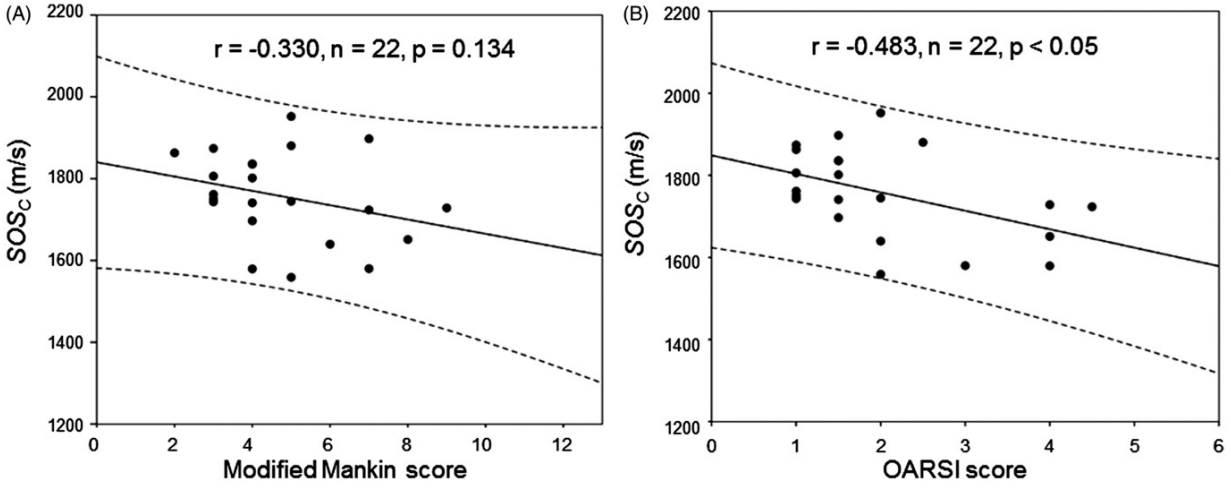


Figure 6. Scatter plots of ultrasound speed (SOS_C) and histological scores. (A) Modified Mankin score; (B) OARSI score. Spearman's rank correlation coefficients (r) are shown. The regression line and the 95% CIs for the population (dashes) are also shown.

articular cartilage using ultrasound is performed either percutaneously [46,47] or arthroscopically [48,49]. The ultrasound frequency used in our study is relatively close to the ultrasound frequency used clinically (5–15 MHz), and we believe that our results could be applied to both percutaneous and arthroscopic evaluation of cartilage thickness. The mean and standard deviation (0.0478 ± 0.188 mm) of the differences between ultrasonic and optical thickness in the present study assures that cartilage

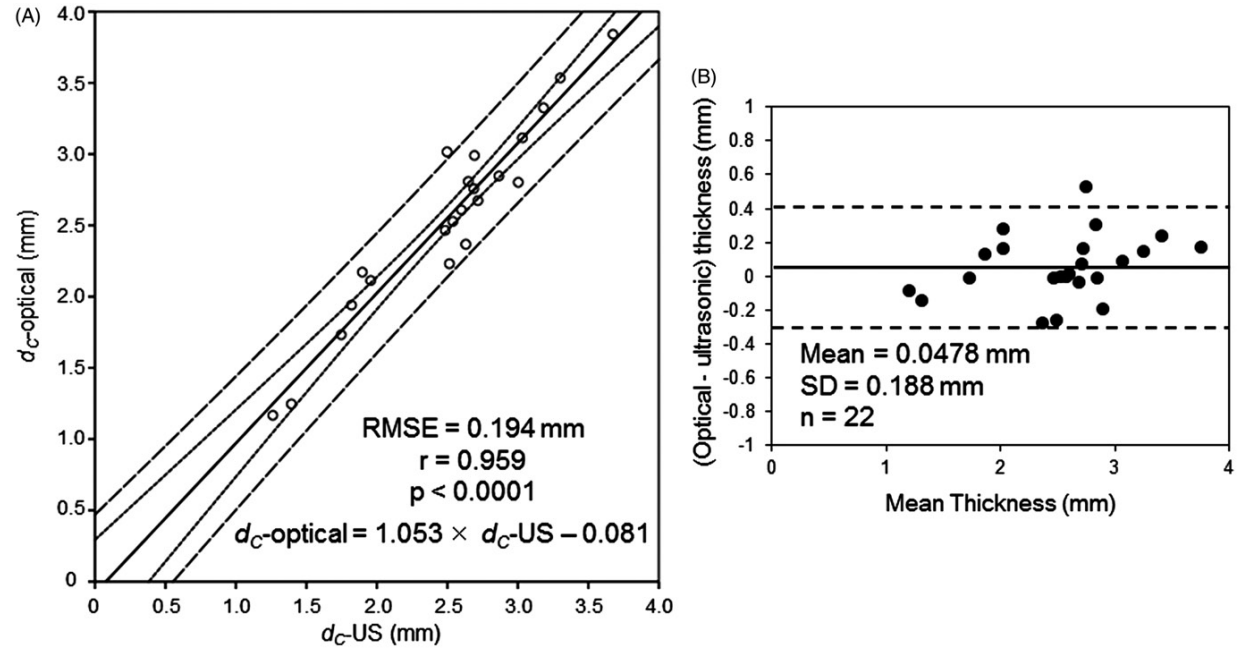


Figure 7. (A) Correlations between optical thickness measurement values (d_C-optical) and thickness values calculated using the average ultrasound speed (d_C-US). The 95% CIs for regression (short dashes) and the population (long dashes) are shown. Root mean square errors (RMSE) and Pearson's correlation coefficients (r) are also shown. (B) Bland–Altman plot analysis of the difference in d_C-optical and d_C-US compared with the mean thickness of d_C-optical and d_C-US. The line corresponding to the mean difference of d_C-optical and d_C-US and lines for mean ± 1.96 × SD (dashes) are shown.

evaluation using a specific ultrasound speed can detect clinically important differences or changes in articular cartilage thickness, considering the results of the past MRI studies.

We are aware of several limitations of our study that will require further exploration. First, we were able to collect specimens only from OA patients who underwent total joint arthroplasty. Although we performed measurements on samples with various degrees of degeneration, from relatively normal areas to degenerated lesions on the femoral condyles, probably none of the samples could be considered fully normal cartilage in this study. Ideally, normal cartilage samples are acquired from cadavers without OA of the knee. Second, we performed evaluation only on samples acquired from the knees, but not from other joints. In animal studies, ultrasound speed could differ among samples obtained from different sites [17,50]. Thus, our results cannot be automatically extrapolated to ultrasonic evaluation of cartilage of other joints, although we assume that the effect of degeneration on ultrasound speed will be similar. Finally, we did not perform a biochemical evaluation of cartilage degeneration. As mentioned before, our aim was to investigate the overall effect of cartilage degeneration on ultrasound speed. Nevertheless, performing biochemical evaluations could reveal which component of the cartilage affects ultrasound speed.

The present study has several strengths. To our knowledge, this is the first study investigating the effect of the degree of cartilage degeneration on ultrasound speed using human samples. We believe that a relatively broad range of samples, representing different degrees of degeneration, was covered in our study and that the findings of the present study support the usage of a preset ultrasound speed value in clinical morphological evaluations of cartilage. In conclusion, our results show that cartilage degeneration has relatively little influence on ultrasound speed in articular cartilage. In addition, morphological evaluation of articular cartilage using a preset value of ultrasound speed seems to offer relatively accurate values of cartilage thickness.

Acknowledgments

The authors thank Mr. Koichi Miyasaka and Mr. Ryoichi Sakai from the Research Laboratory at Hitachi-Aloka Medical Co., Ltd. (Tokyo, Japan) for their technical support.

Conflict of interest

This work was funded by MEXT/JSPS KAKENHI Grant Number 24689030 and H25-Nanchi-009 (Director, Sakae Tanaka) of the Ministry of Health, Labour and Welfare, Japan.

References

1. Sharma L, Kapoor D. Epidemiology of osteoarthritis. In: Moskowitz R, Altman R, Hochberg M, Buckwalter J, Goldberg V, eds. Osteoarthritis: diagnosis and medical/surgical management. 4th ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2007:3–26.
2. Momohara S, Okada N, Ikari K, Mizuno S, Okamoto H. Dermatan sulfate in the synovial fluid of patients with knee osteoarthritis. Mod Rheumatol. 2007;17(4):301–5.
3. Oka H, Akune T, Muraki S, Tanaka S, Kawaguchi H, Nakamura K, et al. The mid-term efficacy of intra-articular hyaluronic acid injections on joint structure: a nested case control study. Mod Rheumatol. 2013;23(4):722–8.
4. Eckstein F, Buck RJ, Burstein D, Charles HC, Crim J, Hudelmaier M, et al. Precision of 3.0 Tesla quantitative magnetic resonance imaging of cartilage morphology in a multicentre clinical trial. Ann Rheum Dis. 2008;67(12):1683–8.
5. Multanen J, Rauvala E, Lammentausta E, Ojala R, Kiviranta I, Hakkinen A, et al. Reproducibility of imaging human knee cartilage by delayed gadolinium-enhanced MRI of cartilage (dGEMRIC) at 1.5 Tesla. Osteoarthr Cartil. 2009;17(5):559–64.
6. Kaleva E, Saarakkala S, Toyras J, Nieminen HJ, Jurvelin JS. *In-vitro* comparison of time-domain, frequency-domain and wavelet ultrasound parameters in diagnostics of cartilage degeneration. Ultrasound Med Biol. 2008;34(1):155–9.
7. Chiang EH, Adler RS, Meyer CR, Rubin JM, Dedrick DK, Laing TJ. Quantitative assessment of surface roughness using backscattered

ultrasound: the effects of finite surface curvature. Ultrasound Med Biol. 1994;20(2):123–35.
8. Adam C, Eckstein F, Milz S, Schulte E, Becker C, Putz R. The distribution of cartilage thickness in the knee-joints of old-aged individuals – measurement by A-mode ultrasound. Clin Biomech (Bristol, Avon). 1998;13(1):1–10.
9. Aisen AM, McCune WJ, MacGuire A, Carson PL, Silver TM, Jafri SZ, et al. Sonographic evaluation of the cartilage of the knee. Radiology. 1984;153(3):781–4.
10. Castriota-Scanderbeg A, De Micheli V, Scarale MG, Bonetti MG, Cammisa M. Precision of sonographic measurement of articular cartilage: inter- and intraobserver analysis. Skeletal Radiol. 1996;25(6):545–9.
11. McCune WJ, Dedrick DK, Aisen AM, MacGuire A. Sonographic evaluation of osteoarthritic femoral condylar cartilage. Correlation with operative findings. Clin Orthop Relat Res. 1990; 254:230–5.
12. Ohashi S, Ohnishi I, Matsumoto T, Bessho M, Matsuyama J, Tobita K, et al. Measurement of articular cartilage thickness using a three-dimensional image reconstructed from B-mode ultrasonography mechanical scans feasibility study by comparison with MRI-derived data. Ultrasound Med Biol. 2012;38(3):402–11.
13. Shin HC, Prager R, Gomersall H, Kingsbury N, Treece G, Gee A. Estimation of average speed of sound using deconvolution of medical ultrasound data. Ultrasound Med Biol. 2010;36(4): 623–36.
14. Torp-Pedersen S, Bartels EM, Wilhjelm J, Bliddal H. Articular cartilage thickness measured with US is not as easy as it appears: a systematic review of measurement techniques and image interpretation. Ultraschall Med (Stuttgart, Germany: 1980). 2011; 32(1):54–61.
15. Nieminen HJ, Zheng Y, Saarakkala S, Wang Q, Toyras J, Huang Y, et al. Quantitative assessment of articular cartilage using high-frequency ultrasound: research findings and diagnostic prospects. Crit Rev Biomed Eng. 2009;37(6):461–94.
16. Suh JK, Youn I, Fu FH. An *in situ* calibration of an ultrasound transducer: a potential application for an ultrasonic indentation test of articular cartilage. J Biomech. 2001;34(10):1347–53.
17. Toyras J, Laasanen MS, Saarakkala S, Lammi MJ, Rieppo J, Kurkijarvi J, et al. Speed of sound in normal and degenerated bovine articular cartilage. Ultrasound Med Biol. 2003;29(3):447–54.
18. Saarakkala S, Korhonen RK, Laasanen MS, Toyras J, Rieppo J, Jurvelin JS. Mechano-acoustic determination of Young's modulus of articular cartilage. Biorheology. 2004;41(3–4):167–79.
19. Ling HY, Zheng YP, Patil SG. Strain dependence of ultrasound speed in bovine articular cartilage under compression *in vitro*. Ultrasound Med Biol. 2007;33(10):1599–608.
20. Nieminen HJ, Toyras J, Laasanen MS, Jurvelin J.S. Acoustic properties of articular cartilage under mechanical stress. Biorheology. 2006;43(3–4):523–35.
21. Nieminen HJ, Julkunen P, Toyras J, Jurvelin JS. Ultrasound speed in articular cartilage under mechanical compression. Ultrasound Med Biol. 2007;33(11):1755–66.
22. Agemura DH, O'Brien WD Jr, Olerud JE, Chun LE, Eyre DE. Ultrasonic propagation properties of articular cartilage at 100 MHz. J Acoust Soc Am. 1990;87(4):1786–91.
23. Patil SG, Zheng YP. Measurement of ultrasound speed of articular cartilage in variable conditions. Conf Proc IEEE Eng Med Biol Soc. 2004;2:1341–4.
24. Barthez PY, Bais RJ, Vernooij JC. Effect of ultrasound beam angle on equine articular cartilage thickness measurement. Vet Radiol Ultrasound. 2007;48(5):457–9.
25. Myers SL, Dines K, Brandt DA, Brandt KD, Albrecht ME. Experimental assessment by high frequency ultrasound of articular cartilage thickness and osteoarthritic changes. J Rheumatol 1995;22(1):109–16.
26. Yao JQ, Seedhom BB. Ultrasonic measurement of the thickness of human articular cartilage *in situ*. Rheumatology (Oxford). 1999;38(12):1269–71.
27. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16(4):494–502.
28. Ohashi S, Ohnishi I, Matsumoto T, Matsuyama J, Bessho M, Tobita K, et al. Comparison of ultrasound speed in articular cartilage measured by different time-of-flight methods. J Med Ultrason. 2011;38(4):225–34.
29. Patwardhan A, Moghe S, Wang K, Cruise H, Leonelli F. Relation between ventricular fibrillation voltage and probability of defibrillation shocks. Analysis using Hilbert transforms. J Electrocardiol. 1998;31(4):317–25.
30. Mankin HJ, Dorfman H, Lippiello L, Zarins A. Biochemical and metabolic abnormalities in articular cartilage from osteo-arthritic human hips. II. Correlation of morphology with bio-chemical and metabolic data. J Bone Joint Surg Am. 1971;53(3): 523–37.
31. van der Sluijs JA, Geesink RG, van der Linden AJ, Bulstra SK, Kuyer R, Drukker J. The reliability of the Mankin score for osteoarthritis. J Orthop Res. 1992;10(1):58–61.
32. Pritzker KP, Gay S, Jimenez SA, Ostergaard K, Pelletier JP, Revell PA, et al. Osteoarthritis cartilage histopathology: grading and staging. Osteoarthritis Cartilage. 2006;14(1):13–29.
33. Nieminen HJ, Toyras J, Rieppo J, Nieminen MT, Hirvonen J, Korhonen R, et al. Real-time ultrasound analysis of articular cartilage degradation *in vitro*. Ultrasound Med Biol. 2002;28(4): 519–25.
34. Jurvelin JS, Rasanen T, Kolmonen P, Lyyra T. Comparison of optical, needle probe and ultrasonic techniques for the measurement of articular cartilage thickness. J Biomech. 1995;28(2):231–5.
35. Brommer H, Laasanen MS, Brama PA, van Weeren PR, Barneveld A, Helminen HJ, et al. Influence of age, site, and degenerative state on the speed of sound in equine articular cartilage. Am J Vet Res. 2005;66(7):1175–80.
36. Shanbhogue AK, Sandhu MS, Singh P, Ojili V, Khandelwal N, Sen R. Real time spatial compound ultrasound in the evaluation of meniscal injuries: a comparison study with conventional ultrasound and MRI. Knee. 2009;16(3):191–5.
37. Nishitani K, Kobayashi M, Kuroki H, Mori K, Shirai T, Satake T, et al. Ultrasound can detect macroscopically undetectable changes in osteoarthritis reflecting the superficial histological and biochemical degeneration: *ex vivo* study of rabbit and human cartilage. PLoS One. 2014;9(2):e89484.
38. Cherin E, Saied A, Pellaumail B, Loeuille D, Laugier P, Gillet P, et al. Assessment of rat articular cartilage maturation using 50-MHz quantitative ultrasonography. Osteoarthritis Cartilage. 2001;9(2): 178–86.
39. Rizkalla G, Reiner A, Bogoch E, Poole AR. Studies of the articular cartilage proteoglycan aggrecan in health and osteoarthritis. Evidence for molecular heterogeneity and extensive molecular changes in disease. J Clin Invest. 1992;90(6):2268–77.
40. Eckstein F, Winzheimer M, Hohe J, Englmeier KH, Reiser M. Interindividual variability and correlation among morphological parameters of knee joint cartilage plates: analysis with three-dimensional MR imaging. Osteoarthritis Cartilage. 2001;9(2): 101–11.
41. Custers RJ, Creemers LB, Verbout AJ, van Rijen MH, Dhert WJ, Saris DB. Reliability, reproducibility and variability of the traditional Histologic/Histochemical Grading System vs the new OARSI Osteoarthritis Cartilage Histopathology Assessment System. Osteoarthritis Cartilage. 2007;15(11):1241–8.
42. Pearson RG, Kurien T, Shu KS, Scammell BE. Histopathology grading systems for characterisation of human knee osteoarthritis–reproducibility, variability, reliability, correlation, and validity. Osteoarthritis Cartilage. 2011;19(3):324–31.
43. Rutgers M, van Pelt MJ, Dhert WJ, Creemers LB, Saris DB. Evaluation of histological scoring systems for tissue-engineered, repaired and osteoarthritic cartilage. Osteoarthr Cartilage. 2010;18(1): 12–23.
44. Eckstein F, Wirth W, Hunter DJ, Guermazi A, Kwoh CK, Nelson DR, et al. Magnitude and regional distribution of cartilage loss associated with grades of joint space narrowing in radiographic osteoarthritis – data from the osteoarthritis initiative (OAI). Osteoarthritis Cartilage. 2010;18(6):760–8.
45. Wirth W, Benichou O, Kwoh CK, Guermazi A, Hunter D, Putz R, et al. Spatial patterns of cartilage loss in the medial femoral condyle in osteoarthritic knees: data from the osteoarthritis initiative. Magn Reson Med. 2010;63(3):574–81.
46. Abraham AM, Goff I, Pearce MS, Francis RM, Birrell F. Reliability and validity of ultrasound imaging of features of knee osteoarthritis in the community. BMC Musculoskelet Disord. 2011;12:70.

47. Saarakkala S, Waris P, Waris V, Tarkiainen I, Karvanen E, Aarnio J, et al. Diagnostic performance of knee ultrasonography for detecting degenerative changes of articular cartilage. Osteoarthritis Cartilage. 2012; 20(5):376–81.
48. Hattori K, Takakura Y, Ishimura M, Habata T, Uematsu K, Ikeuch K. Quantitative arthroscopic ultrasound evaluation of living human cartilage. Clin Biomech (Bristol, Avon). 2004;19(2):213–16.
49. Kuroki H, Nakagawa Y, Mori K, Kobayashi M, Yasura K, Okamoto Y, et al. Ultrasound properties of articular cartilage in the tibiofemoral joint in knee osteoarthritis: relation to clinical assessment (International Cartilage Repair Society grade). Arthritis Res Ther. 2008;10(4):R78.
50. Patil SG, Zheng YP, Chen X. Site dependence of thickness and speed of sound in articular cartilage of bovine patella. Ultrasound Med Biol. 2010;36(8):1345–52. doi: 10.1186/ar2452.

PLOS | ONE

RESEARCH ARTICLE

Metabolic Syndrome Components Are Associated with Intervertebral Disc Degeneration: The Wakayama Spine Study

Masatoshi Teraguchi[1], Noriko Yoshimura[2], Hiroshi Hashizume[1]*, Shigeyuki Muraki[2], Hiroshi Yamada[1], Hiroyuki Oka[3], Akihito Minamide[1], Yuyu Ishimoto[1], Keiji Nagata[1], Ryohei Kagotani[1], Sakae Tanaka[4], Hiroshi Kawaguchi[5], Kozo Nakamura[6], Toru Akune[6], Munehito Yoshida[1]

1 Department of Orthopaedic surgery, Wakayama Medical University, 811–1 Kimiidera, Wakayama, 641–8509, Japan, **2** Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113–8655, Japan, **3** Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113–8655, Japan, **4** Department of Orthopaedic surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113–8655, Japan, **5** Japan Community Healthcare Organization Tokyo Shinjuku Medical Center, 5–1 Tsukudo-chome, Shinjuku-ku, Tokyo, 162–8543, Japan, **6** Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, 1 Namiki 4-chome, Tokorozawa City, Saitama, 359–8555, Japan

* hashizum@wakayama-med.ac.jp

Citation: Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, et al. (2016) Metabolic Syndrome Components Are Associated with Intervertebral Disc Degeneration: The Wakayama Spine Study. PLoS ONE 11(2): e0147565. doi:10.1371/journal.pone.0147565

Editor: Masaru Katoh, National Cancer Center, JAPAN

Received: September 23, 2015

Accepted: January 5, 2016

Published: February 3, 2016

Data Availability Statement: The present study used resident data from two communities in Wakayama prefecture. It is impossible for us to provide and upload these data in a public repository because we have confirmed with these municipalities and residents that data will remain confidential. We will provide anonymized data on request after discussing the contents with the municipalities, as long as researchers are qualified to request these data. Data requests can be made to the corresponding author at hashizum@wakayama-med.ac.jp.

Abstract

Objective

The objective of the present study was to examine the associations between metabolic syndrome (MS) components, such as overweight (OW), hypertension (HT), dyslipidemia (DL), and impaired glucose tolerance (IGT), and intervertebral disc degeneration (DD).

Design

The present study included 928 participants (308 men, 620 women) of the 1,011 participants in the Wakayama Spine Study. DD on magnetic resonance imaging was classified according to the Pfirrmann system. OW, HT, DL, and IGT were assessed using the criteria of the Examination Committee of Criteria for MS in Japan.

Results

Multivariable logistic regression analysis revealed that OW was significantly associated with cervical, thoracic, and lumbar DD (cervical: odds ratio [OR], 1.28; 95% confidence interval [CI], 0.92–1.78; thoracic: OR, 1.75; 95% CI, 1.24–2.51; lumbar: OR, 1.87; 95% CI, 1.06–3.48). HT and IGT were significantly associated with thoracic DD (HT: OR, 1.54; 95% CI, 1.09–2.18; IGT: OR, 1.65; 95% CI, 1.12–2.48). Furthermore, subjects with 1 or more MS components had a higher OR for thoracic DD compared with those without MS components

Funding: This study was supported by H23-Choujyu-002 (Director, Toru Akune), H-25-Choujyu-007 (Director, Noriko Yoshimura), H25-Nanchitou (Men)-005 (Director, Sakae Tanaka), 201417014A (Director, Noriko Yoshimura), and H22-Choujyu-Wakate-007 (Director, Shigeyuki Muraki) from the Ministry of Health, Labour and Welfare; a Grant-in-Aid for Scientific Research (B26293139, B23390172 to Noriko Yoshimura, B2629333, C20591774 to Shigeyuki Muraki, C26462249 to Hiroshi Hashizume, C25462305 to Hiroshi Yamada) and a Grant-in-Aid for Young Researcher (B25860448 to Yuyu Ishimoto, B26861286 to Masatoshi Teraguchi, B26860419 to Ryohei Kagotani, B15K20013 to Hiroki Iwahashi), and Grant-in-Aid for Challenging Exploratory Research (15K15219 to Noriko Yoshimura, 26670307 to Shigeyuki Muraki, 24659666 to Hiroyuki Oka, 25670293 to Toru Akune) of JSPS KAKENHI grant; a Grant from the Japanese Orthopaedics and Traumatology Foundation, Inc. (No. 287) to Masatoshi Teraguchi; and Collaborating Research with NSF 08033011- 00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. This study was also supported by grants from the Japan Osteoporosis Society (Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, and Toru Akune), a grant from JA Kyosai Research Institute (Hiroyuki Oka), a grant from Mitsui Sumitomo Insurance Welfare Foundation (Shigeyuki Muraki), and research aid from the Japanese Orthopaedic Association (JOA-Subsidized Science Project Research 2006-1 & 2010-2; Director, Hiroshi Kawaguchi).

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: DD, disc degeneration; MS, metabolic syndrome; OW, overweight; HT, hypertension; DL, dyslipidemia; IGT, impaired glucose tolerance; ROAD, Research on Osteoarthritis/Osteoporosis Against Disability.

(vs. no component; 1 component: OR, 1.58; 95% CI, 1.03–2.42; 2 components: OR, 2.60; 95% CI, 1.62–4.20; ≥3 components: OR, 2.62; 95% CI, 1.42–5.00).

Conclusion

MS components were significantly associated with thoracic DD. Furthermore, accumulation of MS components significantly increased the OR for thoracic DD. These findings support the need for further studies of the effects of metabolic abnormality on DD.

Introduction

Intervertebral disc degeneration (DD) is generally considered as the first step of spinal change and undergoes destructive changes with age. It is typically followed by the loss of water and proteoglycan content of the nucleus, annulus tears, gradual formation of osteophytes, disc narrowing, and spinal canal stenosis [1, 2], and low back pain [3–6], is a major public health problem that negatively influences activities of daily living and quality of life in those affected. The number of patients with degenerative disease of the spine is increasing [6], thereby causing medical expenses to rise. In spite of these situation, the cause of DD is not fully understood. Because the etiology of DD exclude aging remains poorly understood. Accordingly, we need to clarify which risk factors promote DD to establish preventive measures against DD. In the present study, we focused on metabolic syndrome (MS) component, such as overweight (OW), hypertension (HT), dyslipidemia (DL), and impaired glucose tolerance (IGT), because MS component has some influence on atherosclerosis [7] and accumulation of MS component increase the risk of atherosclerosis events [8]. MS may increase not only the risk of cardiovascular events but also the risk of DD in the whole body [9], because intervertebral discs, which are structures with precarious nutrient supply at tissue level throughout the whole body, may suffer and gradually degenerate as a consequence of failure of nutrient supply to disc cells [10, 11]. However the association between MS component and DD remains controversial [9]. In some previous epidemiologic studies, OW [6, 12–16], DL [16], and IGT [17] were found to be associated with DD in the lumbar region. Other studies, however, have found no clear associations between hypertension (HT) [9], IGT [9, 16], and DL [9, 18], and DD in the lumbar region. This may be due to the limitation of potential biases related to patient selection and the consequences of disease on behavior. Furthermore, the majority of epidemiologic investigations have focused only on the lumbar spine. We believe that analysis of DD in the entire spine would provide more useful data than that of DD in only the lumbar region. Since the cervical and lumbar regions comprise mobile segments, the intervertebral discs in these regions are easily affected by mechanical and motion stress; thus, the effects of certain factors imposed on all intervertebral discs equally, such as age and endogenic factors, might be masked. In contrast, the thoracic region is stabilized by the thoracic cage, which reduces mechanical stress imposed on the intervertebral discs. We conducted a thorough literature review and found no studies of associations between component of MS and DD that focused on a population-based analysis using whole-spine magnetic resonance imaging (MRI).

The purpose of the present study was to examine the association of each MS component, such as OW, HT, DL, and IGT, with DD in the cervical, thoracic, and lumbar regions of the entire spine in a large population. We also examined the relationship between accumulation of MS components and DD.

Methods

Participants

The present study design was approved by the Wakayama medical university ethics committee. All participants provided their written informed consent. The present study, entitled the Wakayama Spine Study, was a population-based study of DD performed using a subcohort of the large-scale population-based cohort study called Research on Osteoarthritis/Osteoporosis Against Disability (ROAD). The ROAD study is a nationwide, prospective study of bone and joint diseases consisting of population-based cohorts established in several communities in Japan [19, 20]. A second visit of the ROAD study to the mountainous region of H town and the seacoast region of T town was performed between 2008 and 2010. From inhabitants participating in the second visit of the ROAD study, 1,063 volunteers were recruited for MRI examinations. Among the 1,063 volunteers, 52 declined to attend the examination; therefore, 1,011 inhabitants were recruited for registration in the Wakayama Spine Study. Among the 1,011 participants, those who had an MRI-sensitive implanted device (e.g., pacemaker) or other disqualifiers were excluded. Consequently, 980 individuals underwent whole-spine MRI. One participant who had undergone a previous cervical operation and 4 participants who had undergone previous posterior lumbar fusion were excluded from the analysis. Whole-spine MRI results were available for 975 participants (324 men, 651 women) with an age range of 21 to 97 years (mean, 67.2 years for men, 66.0 years for women). Thirty participants with incomplete anthropometric measurements and 17 participants without blood measurements were excluded. Finally, the present study comprised 928 participants (308 men, 620 women) with a mean age of 67.4 years.

The participants completed an interviewer-administered questionnaire of 400 items that included lifestyle information, such as smoking habit, alcohol consumption, family history, past history, occupation, physical activity, and health-related quality of life. Anthropometric measurements included height, weight, and body mass index (BMI) (weight [kg]/height $[m]^2$). An experienced public health nurse measured systolic and diastolic blood pressure (BP) using a mercury sphygmomanometer.

MRI

A mobile MRI unit (Excelart 1.5 T; Toshiba, Tokyo, Japan) was used in the present study, and whole-spine MRI was performed for all participants on the same day as the questionnaire and anthropometric examination. The participants were supine during MRI, and those with rounded backs used triangular pillows under their head and knees. The imaging protocol included sagittal T2-weighted fast-spin echo (FSE) (repetition time [TR], 4000 ms/echo; echo time [TE], 120 ms; field of view [FOV], 300 × 320 mm) and axial T2-weighted FSE (TR, 4000 ms/echo; TE, 120 ms; FOV, 180 × 180 mm).

Sagittal T2-weighted images were used to assess the intervertebral space from C2/3 to L5/S1. C2/3 to C7/T1, T1/2 to T12/L1, and L1/2 to L5/S1 were defined as the cervical, thoracic, and lumbar region, respectively. Grading of DD was performed by a board certified orthopedic surgeon (M.T.) who was blinded to the background of the subjects. The degree of DD on MRI was classified into 5 grades based on the Pfirrmann system [21], with grades 4 and 5 indicating DD. The signal intensity for grade 4 is intermediate to hypointense to cerebrospinal fluid (dark gray), while the structure is inhomogeneous. The signal intensity for grade 5 is hypointense to cerebrospinal fluid (black), and the structure is likewise inhomogeneous. In addition, the disc space is collapsed. It has been reported that loss of signal intensity is significantly associated

with morphologic level of DD and also with water and proteoglycan content in a disc [22]. Therefore, we used a grading system based on signal intensity and disc height.

For evaluating intraobserver variability, 100 randomly selected whole-spine magnetic resonance images were rescored by the same observer (M.T.) more than 1 month after the first reading. Furthermore, to evaluate interobserver variability, 100 other magnetic resonance images were scored by 2 board certified orthopedic surgeons (M.T. and R.K.) using the same classification system. The intra- and interobserver variability for DD, as evaluated by kappa analysis, were 0.94 and 0.94, respectively.

Blood examination

All blood and urine samples were extracted between 9:00 AM and 3:00 PM. Some samples were extracted under fasting conditions. After centrifugation of the blood samples, sera were immediately placed in dry ice, and transferred to a deep freezer within 24 hours. These samples were stored at −80°C until assayed. For the samples of participants in the baseline study, the following items were measured: blood counts, hemoglobin, hemoglobin A1c (HbA1c), blood sugar, total protein, aspartate aminotransferase, alanine aminotransferase, γ-glutamyl transpeptidase, high-density lipoprotein cholesterol (HDL-C), total cholesterol, triglycerides (TGs), blood urea nitrogen, uric acid, and creatinine. These analyses were performed at the same laboratory within 24 hours after extraction (Osaka Kessei Research Laboratories Inc., Osaka, Japan).

Definitions of MS components were based mainly on the criteria of the Examination Committee of Criteria for MS in Japan [23]. According to the consensus, an abdominal circumference $\geq$85 cm in men and $\geq$90 cm in women is a necessary condition for MS. HT was diagnosed as systolic BP $\geq$130 mm Hg and/or diastolic BP $\geq$85 mm Hg; DL, as serum TG level $\geq$150 mg/dL and/or serum HDL-C level $<$40 mg/dL; and IGT, as fasting serum glucose level $\geq$100 mg/dL. Recently, the National Cholesterol Education Program's Adult Treatment Panel III report proposed a new set of criteria to define MS without central obesity, as indicated by waist circumference, as the core feature [24]. Furthermore, compared with BMI, measurement of waist circumference is less reproducible due to lack of uniformity in measurement methods [25, 26]. By contrast, measurement of BMI is more user-friendly and widely practiced. In this study, we decided to use BMI $\geq$25 kg/m^2 as an indicator of OW, based on the criteria of the Japan Society for the Study of Obesity [25].

In addition, because not all blood samples were obtained under fasting conditions, we did not use participants' data concerning serum levels of glucose and TGs because of their large variation depending on hours after eating. Instead, we used serum HDL-C level $<$40 mg/dL to indicate DL, and serum HbA1c level $\geq$5.5% to indicate IGT (the value for HbA1c (National Glycohemoglobin Standardization Program (NGSP)) (%) is estimated as an NGSP-equivalent value calculated by the formula HbA1c (%) = HbA1c (Japan Diabetes Society (JDS)) (%) + 0.4%) [27]. These are indices used in the National Health and Nutrition Survey in Japan, which were adopted as criteria for MS in this national screening based on the difficulty of collecting samples under fasting conditions [28].

Statistical analysis

All statistical analyses were performed using JMP version 8 (SAS Institute Japan, Tokyo, Japan). Differences between the groups depending on the presence or absence of DD were tested using a variance analysis. Multivariable logistic regression analysis was performed to determine the association of OW, HT, DL, and IGT with DD. The DD in the cervical, thoracic, or lumbar region was separately served as an objective variable. Then, to clarify the association

Table 1. Background characteristics of the participants.

	Overall	Men	Women
No. of participants	**928**	**308**	**620**
Mean (SD) selected characteristics			
Age (years)	67.4 (12.3)	68.5 (12.4)	66.8 (12.2)
Height (cm)	155.8 (9.4)	160.5 (8.0)	153.4 (9.1)
Weight (kg)	56.7 (11.5)	60.2 (11.4)	55.0 (11.2)
Body mass index (kg/m^2)	23.3 (3.6)	23.7 (3.3)	23.1 (3.7)
Systolic BP, mmHg	139.5 (19.6)	141.3 (18.5)	138.7 (20.0)
Diastolic BP, mmHg	76.0 (11.5)	78.1 (12.5)	74.9 (10.9)
Serum levels of HDL-C, mg/dl	63.2 (16.2)	56.0 (14.8)	66.8 (15.7)
Serum levels of HbA1c, %	5.3 (0.7)	5.3 (0.9)	5.2 (0.6)
Prevalence of selected characteristics, %			
Smoking habit	10.1	23.3	3.5
Alcohol consumption	31.2	57.5	18.2
Prevalence of each metabolic abnormality, %			
Obesity	29.4	32.5	27.9
Hypertension	74.7	78.9	72.6
Dyslipidemia	4.5	10.1	1.8
Impaired glucose tolerance	23.3	27.3	21.3

Values are the means ± standard deviation. HDL-C = high density lipoprotein cholesterol, HbA1c = glycosylated haemoglobin, ABI = ankle brachial index, SD = standard deviation

doi:10.1371/journal.pone.0147565.t001

between accumulation of MS components and DD, logistic regression analysis was repeated using presence of DD in the cervical, thoracic, and lumbar region, respectively, as the objective variable and number of MS components present as the explanatory variable, after adjusting for age, sex, regional difference, smoking habit, and alcohol consumption. P value of <0.05 was treated as significant.

Results

Table 1 shows selected characteristics of the participants, including age, height, weight, BMI, systolic and diastolic BP, and serum levels of HDL-C and HbA1c, classified by sex. Table 1 also shows the proportion of subjects who smoked (regularly or more than once a month) and consumed alcohol (regularly or more than once a month), and the prevalence of OW, HT, DL, and IGT. In the total population, the MS component with the highest prevalence was HT, followed by OW, IGT, and DL.

Table 2 shows the mean value of each MS component according to absence and presence of DD in the cervical, thoracic, and lumbar region, respectively. Mean values of age, BMI, systolic BP, and HbA1c were significantly higher, while those of HDL-C were significantly lower, in subjects with DD than in those without DD.

To determine the associations of DD with OW, HT, DL, and IGT, multivariable logistic regression analysis was performed (Table 3). OW was significantly associated with presence of DD in the cervical, thoracic, and lumbar regions. In addition, HT, DL, and IGT were significantly associated with presence of DD in the thoracic region, but not with DD in the cervical and lumbar regions.

Next, to determine the effect of accumulation of MS components on DD in the thoracic region, we examined the association of number of MS components present with DD after

Table 2. Mean value (SD) of each demographic characteristics and measurements in the absence and presence of disc degeneration in the cervical, thoracic, and lumbar region, respectively.

	Cervical			Thoracic			Lumbar		
	Presence of DD	Absence of DD	p-value	Presence of DD	Absence of DD	p-value	Presence of DD	Absence of DD	p-value
No. of participants	592	336		578	350		839	89	
Demographic characteristics and measurements									
Age (years)	70.8 (11.3)	61.4 (11.7)	0.0001	71.7 (10.3)	60.1 (11.9)	0.0001	67.7 (11.8)	64.0 (16.0)	0.007
Body mass index (kg/m^2)	23.5 (3.6)	23.0 (3.6)	0.0552	23.6 (3.7)	22.9 (3.3)	0.0058	23.4 (3.6)	22.2 (3.3)	0.0028
Systolic BP, mmHg	141.3 (19.1)	136.6 (19.9)	0.0004	142.5 (18.8)	134.7 (19.8)	0.0001	140.0 (19.8)	134.9 (16.9)	0.0188
Diastolic BP, mmHg	75.5 (11.2)	76.9 (12.1)	0.0663	75.6 (11.2)	76.7 (12.0)	0.1707	76.1 (11.5)	74.9 (12.1)	0.3362
Serum levels of HDL-C, mg/dl	61.8 (15.2)	65.7 (17.5)	0.0004	61.9 (15.4)	65.3 (17.3)	0.0022	63.2 (16.1)	63.5 (17.3)	0.853
Serum levels of HbA1c, %	5.3 (0.7)	5.2 (0.7)	0.0011	5.4 (0.8)	5.1 (0.5)	0.0001	5.3 (0.7)	5.1 (0.5)	0.0067

DD = disc degeneration, BMI = body mass index, BP = blood pressure, HDL-C = high density lipoprotein in cholesterol, HbA1c = hemoglobin A1c, SD = standard deviation

Differences between the groups depending on the presence or absence of DD were tested using a variance analysis.

doi:10.1371/journal.pone.0147565.t002

adjusting for age, sex, regional difference, smoking habit, and alcohol consumption. Fig 1. shows the odds ratio (OR) of number of MS components for presence of DD in the thoracic region. Subjects with 1 or more MS components had a higher OR for presence of DD compared with those without MS components (vs. no component; 1 component: OR, 1.58; 95% confidence interval [CI], 1.03–2.42; p = 0.0353; 2 components: OR, 2.60; 95% CI, 1.61–4.20; $p < 0.0001$; $\geq$3 components: OR, 2.62; 95% CI, 1.42–5.00; p = 0.0021).

Table 3. Association of OW, HT, DL and IGT in cervical, thoracic and lumbar region, respectively.

		Cervical		Thoracic		Lumbar	
		OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value
Overweight	**Yes vs no**	1.28 (0.92–1.78)	0.1397	1.75 (1.24–2.51) *	0.0016	1.87 (1.06–3.48) *	0.0306
Hypertension	**Yes vs no**	1.19 (0.85–1.66)	0.286	1.54 (1.09–2.18) *	0.0138	0.88 (0.52–1.45)	0.6189
Dyslipidemia	**Yes vs no**	1.06 (0.49–2.17)	0.8853	0.42 (0.21–0.86) *	0.0176	0.87 (0.34–2.70)	0.7963
Impaired Glucose Tolerance	**Yes vs no**	1.27 (0.88–1.85)	0.1943	1.65 (1.12–2.48) *	0.0115	1.48 (0.80–2.95)	0.2211
Age	**over 65 vs under 65**	2.98 (2.16–4.11) * * *	<0.0001	5.76 (4.10–8.16) * * *	<0.0001	4.72 (2.47–9.69) * * *	<0.0001
Sex	**Women vs men**	1.32 (0.92–1.90)	0.1342	1.02 (0.70–1.48)	0.9198	0.95 (0.56–1.66)	0.8691
Regional difference	**Mountainous town vs seacoast town**	1.75 (1.13–2.78) *	0.012	1.20 (0.77–1.89)	0.4353	0.14 (0.07–0.27) * * *	<0.0001
Smoking habit	**Yes vs no**	0.88 (0.54–1.44)	0.6002	0.90 (0.54–1.49)	0.6707	0.56 (0.29–1.12)	0.0984
Alcohol consumption	**Yes vs no**	0.94 (0.67–1.33)	0.7332	0.96 (0.67–1.37)	0.8332	0.92 (0.55–1.55)	0.7456

Multivariable logistic regression analysis was performed to determine the association of OW, HT, DL, and IGT with DD. The DD in the cervical, thoracic, or lumbar region was separately served as an objective variable. DD = disc degeneration, OW = Overweight, HT = Hypertension, DL = Dyslipidemia, IGT = Impaired Glucose Tolerance. Overweight was diagnosed as BMI $\geq$ 25, Hypertension was diagnosed as systolic BP $\geq$ 130 mm Hg and/or diastolic BP $\geq$ 85 mm Hg, DL was diagnosed as serum HDL-C level < 40 mg/dl, Impaired Glucose Intolerance was diagnosed as serum HbA1c level $\geq$ 5.5%

OR = odds ratio, 95% CI = 95% confidence interval

*p value < 0.001

***p value < 0.0001

doi:10.1371/journal.pone.0147565.t003

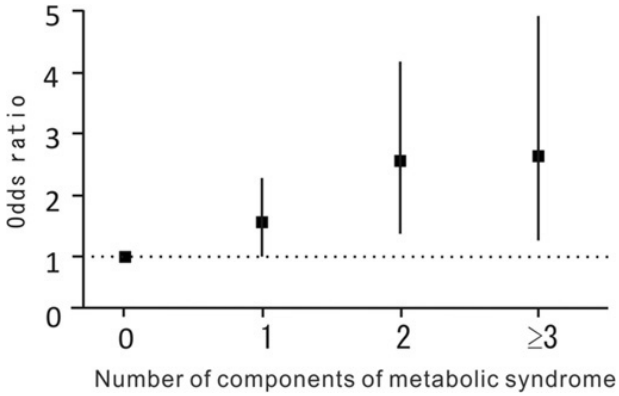


Fig 1. ORs of the number of MS components for the presence of DD in the thoracic region, compared with no components present. Subjects with 1 or more MS components had a higher OR for presence of DD compared with those without MS components.

doi:10.1371/journal.pone.0147565.g001

Discussion

The present study was the first to determine the associations between MS components and DD in the entire spine using whole-spine MRI in a large population. We elucidated that OW was significantly associated with presence of DD in the entire spine, including the cervical, thoracic, and lumbar regions. HT, DL, and IGT were significantly associated with presence of DD in the thoracic region, but not with DD in the cervical and lumbar regions. Furthermore, we also found that accumulation of MS components was significantly associated with presence of DD in the thoracic region.

Regarding the association between degenerative musculoskeletal disease and metabolic risk factors, Yoshimura et al. clarified the association between accumulation of metabolic risk factors and presence and occurrence of knee osteoarthritis (OA) [29, 30]. Hart et al. found that metabolic risk factors, such as high blood glucose level, hypercholesterolemia, and even treated HT, were associated with development of knee OA [31]. Furthermore, Anekstein et al. clarified the association between diabetes mellitus and lumbar spinal stenosis in the patients [32]. However, to our knowledge, there has been no report concerning the association between MS components and DD in the spine, especially the entire spine, using whole-spine MRI in a large population.

In the present study, OW was significantly associated with presence of DD in the cervical, thoracic, and lumbar regions. The association between OW and DD has been previously reported, and Liuke et al. found that past OW was more strongly associated with DD than present OW [13]. Samartzis et al. reported that DD in the lumbar region was significantly associated with OW and obesity [14]. On the other hand, according to Okada et al. and Matsumoto et al., DD in the cervical and thoracic regions did not have significant correlation with BMI [32, 33]. Therefore, the association remains controversial. The present study is the first to determine the association of OW with DD in the entire spine using a population-based design, and found that OW was significantly associated with DD in not only the lumbar region but also the cervical and thoracic regions. DD is influenced by inflammatory cytokines, such as adipokines, known as key metabolism mediators [34–37]. Inflammatory cytokines, such as leptin, adiponectin, and resistin, have more addressed in body fat [34, 38]. Thus, OW may lead to an increase in adipokine secretion of proinflammatory cytokines and metabolic mediators; thus, all intervertebral discs in the entire spine may be influenced by inflammatory cytokines.

Further research is needed to elucidate the mechanism through which OW affects DD since both direct mechanical stress and indirect factors affect the intervertebral discs.

To our knowledge, there has been less report regarding the association of HT with DD or lumbar spinal stenosis as spinal disorder [9, 39]. HT is a well-known risk factor for development of atherosclerosis [40]. Thus, HT might lead to vascular insufficiency to the disc, due to atherosclerosis, which can affect nutrient and metabolite transport into the disc.

The present study also confirmed the significant association between IGT and DD. In one study, it was reported that diabetic sand rats had more dehydrated discs compared with a control group [17]. In the Nurses' Health Study, IGT increased the risk of lumbar disc herniation [41]. However, several previous reports on DD also showed a weak association with IGT [9, 16, 42]; this may be due to their investigation of DD in only the lumbar region. In this study, we found an association between IGT and presence of DD in the thoracic region. Therefore, IGT, which is well known for causing microangiopathy throughout the whole body, also might be a predisposing factor for development of DD. Furthermore, advanced glycation end products accumulate in the intervertebral discs with aging, particularly when the concentration of serum glucose is high, such as in IGT [43]. Therefore, IGT might be associated with DD.

In this study, we found a negative association between DL and DD. The association of DL and DD also remains controversial in previous reports [9, 18, 44]. We believe that DD might be the result of decreased blood supply, caused by DL, to the already poorly vascularized discs [45, 46]. The mean HDL-C was higher in women than in men, as shown in Table 2. Because women in Japan use health services more frequently compared with men [28], the proportion of patients with DL in women was higher than that in men. This might have influenced the negative association between DD and DL. In a follow-up study, we will further investigate the association between DL and DD.

We found no associations of HT, DL, and IGT with DD in the cervical and lumbar regions. Since the cervical and lumbar regions comprise mobile segments, the intervertebral discs are easily affected by mechanical and motion stress, while the effect of endogenous factors might be masked. In contrast, in the thoracic region, mechanical stress on the intervertebral discs is reduced because the region is stabilized by the thoracic cage. Distinct associations among DD in the cervical, thoracic, and lumbar regions might indicate the effects of HT, DL, and IGT on DD are due to endogenous factors. To clarify risk factors for DD, particularly endogenous risk factors, it may be useful to examine associations in not only the cervical and lumbar regions, but in the thoracic region as well.

Study limitations

This study has several limitations. First, this was a cross-sectional study; thus, the causal relationships between MS components and DD remain unclear. These can only be ascertained by a follow-up study that clarifies the incidence and/or progression rates of DD in the same cohort. Second, the participants included in the present study may not represent the general population since they were recruited from only 2 local areas. To confirm whether the participants are representative of the Japanese population, we compared anthropometric measurements and frequencies of smoking and alcohol consumption between the general Japanese population and the study participants. No significant difference in BMI was observed (men: 24.0 kg/m^2 vs. 23.7 kg/m^2, $p = 0.33$; women: 23.5 kg/m^2 vs. 23.1 kg/m^2, $p = 0.07$). Further, the proportion of men who smoked and who consumed alcohol (those who regularly smoked or consumed alcohol more than once per month) and the proportion of women who consumed alcohol were significantly higher in the general Japanese population than in the study population, whereas there was no significant difference in the proportion of women who smoked (men who smoked: 32.6% vs. 23.3%,

p = 0.015; women who smoked: 4.9% vs. 3.5%, p = 0.50; men who consumed alcohol: 73.9% vs. 57.5%, $p < 0.0001$; women who consumed alcohol: 28.1% vs. 18.2%, $p < 0.0001$). These results suggest the likelihood that, in this study, participants had healthier lifestyles than those of the general Japanese population [28]. This "healthy" selection bias should be taken into consideration when generalizing the results obtained from the Wakayama Spine Study. In addition, since the blood samples obtained were not always from participants under fasting conditions, we used serum HDL-C level <40 mg/dL, and not TG level, to indicate DL, and serum HbA1c level ≥5.5%, and not blood glucose level, to indicate IGT, which are indices used by the National Health and Nutrition Survey in Japan [28]. These differences in the definition of MS might have skewed the true association between MS and DD.

Conclusions

We investigated the associations between MS components and DD in the cervical, thoracic, and lumbar regions in a large population of individuals ranging in age from 21 to 97 years. We revealed that OW was significantly associated with presence of DD in the entire spine, and that HT and IGT were significantly associated with presence of DD in the thoracic region. We also found that subjects with 1 or more MS components had a higher OR for presence of DD compared with those without MS components. The prevention of MS may be useful for avoiding DD. Further investigations, along with continued longitudinal surveys of the Wakayama Spine Study, will elucidate the associations between MS components and occurrence or progression of DD.

Acknowledgments

The authors wish to thank Mrs. Tomoko Takijiri and other members of the Public Office in Hidakagawa Town, and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, and other members of the Public Office in Taiji Town for their assistance in the location and scheduling of participants for examinations.

Author Contributions

Conceived and designed the experiments: MT NY HH SM HY HO AM YI K. Nagata RK ST HK K. Nakamura TA MY. Performed the experiments: MT NY HH SM HY HO YI K. Nagata RK TA. Analyzed the data: MT NY HH SM HO. Contributed reagents/materials/analysis tools: MT NY HH SM. Wrote the paper: MT.

References

1. Boos N, Weissbach S, Rohrbach H, Weiler C, Spratt KF, Nerlich AG. Classification of age-related changes in lumbar intervertebral discs: 2002 Volvo Award in basic science. Spine. 2002; 27(23):2631–44. PMID: 12461389
2. Katz JN, Harris MB. Lumbar spinal stenosis. N Engl J Med 2008; 358(8):818–25. PMID: 18287604 doi: 10.1056/NEJMcp0708097
3. Luoma K, Riihimäki H, Luukkonen R, Raininko R, Viikari-Juntura E, Lamminen A. Low back pain in relation to lumbar disc degeneration. Spine 2000; 25(4):487–92. PMID: 10707396
4. Borenstein DG, O'Mara JW Jr, Boden SD, Lauerman WC, Jacobson A, Platenberg, et al. The value of magnetic resonance imaging of the lumbar spine to predict low-back pain in asymptomatic subjects: a seven-year follow-up study. J Bone Joint Surg Am 2001 2001; 83-A(9):1306–11. PMID: 11568190
5. de Schepper EI, Damen J, van Meurs JB, Ginai AZ, Popham M, Hofman A, et al. The association between lumbar disc degeneration and low back pain: the influence of age, gender, and individual radiographic features. Spine 2010; 35(5):531–6. doi: 10.1097/BRS.0b013e3181aa5b33 PMID: 20147869

6. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, et al. Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study. Osteoarthritis Cartilage 2014; 22(1):104–10. doi: 10.1016/j.joca.2013.10.019 PMID: 24239943
7. Vitale C, Marazzi G, Volterrani M, Aloisio A, Rosano G, Fini M: Metabolic syndrome. Minerva Med. 2006, 97 (3): 219–29. PMID: 16855517
8. Wilson PW, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic Syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. Circulation. 2005; 112(20):3066–72. PMID: 16275870
9. Suri P, Hunter DJ, Rainville J, Guermazi A, Katz JN. Quantitative assessment of abdominal aortic calcification and associations with lumbar intervertebral disc height loss: the Framingham Study. Spine J 2012; 12(4): 315–23. doi: 10.1016/j.spinee.2012.03.033 PMID: 22561175
10. Urban JP, Smith S, Fairbank JC. Nutrition of the intervertebral disc. Spine 2004; 29(23):2700–9. PMID: 15564919
11. Urban JP, Roberts S. Degeneration of the intervertebral disc. Arthritis Res Ther 2003; 5(3):120–30. PMID: 12723977
12. Yoshimura N, Dennison E, Wilman C, Hashimoto T, Cooper C. Epidemiology of chronic disc degeneration and osteoarthritis of the lumbar spine in Britain and Japan: a comparative study. J Rheumatol 2000; 27(2): 429–33. PMID: 10685810
13. Liuke M, Solovieva S, Lamminen A, Luoma K, Leino-Arjas P, Luukkonen R, et al. Disc degeneration of the lumbar spine in relation to overweight. Int J Obes (Lond) 2005; 29(8): 903–8. PMID: 15917859
14. Samartzis D, Karppinen J, Chan D, Luk KD, Cheung KM. The association of lumbar intervertebral disc degeneration on magnetic resonance imaging with body mass index in overweight and obese adults. Arthritis Rheum 2012; 64(5):1488–96. doi: 10.1002/art.33462 PMID: 22287295
15. Takatalo J, Karppinen J, Taimela S, Niinimäki J, Laitinen J, Sequeiros RB, et al. Association of abdominal obesity with lumbar disc degeneration—a magnetic resonance imaging study. PLoS One 2013; 8 (2): e56244. doi: 10.1371/journal.pone.0056244 PMID: 23418543
16. Hangai M, Kaneoka K, Kuno S, Hinotsu S, Sakane M, Mamizuka N, et al. Factors associated with lumbar intervertebral disc degeneration in the elderly. Spine J 2008; 8(5): 732–40. PMID: 18037353
17. Ziv I, Moskowitz RW, Kraise I, Adler JH, Maroudas A. Physicochemical properties of the aging and diabetic sand rat intervertebral disc. J Orthop Res 1992; 10(2): 205–10. PMID: 1740738
18. Kauppila LI, Mikkonen R, Mankinen P, Pelto-Vasenius K, Mäenpää I. MR aortography and serum cholesterol levels in patients with long-term nonspecific lower back pain. Spine 2004; 29(19): 2147–52. PMID: 15454707
19. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study. Int J Epidemiol 2010; 39(4): 988–95. doi: 10.1093/ije/dyp276 PMID: 19749026
20. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 2009; 27(5): 620–28. doi: 10.1007/s00774-009-0080-8 PMID: 19568689
21. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine 2001; 26(17): 1873–8. PMID: 11568697
22. Benneker LM, Heini PF, Anderson SE, Alini M, Ito K. Correlation of radiographic and MRI parameters to morphological and biochemical assessment of intervertebral disc degeneration. Eur Spine J 2005; 14 (1): 27–35. PMID: 15723249
23. The Examination Committee of Criteria for Metabolic Syndrome. The definition and criteria of metabolic syndrome [in Japanese]. J Jpn Soc Intern Med 2005; 94: 794–809.
24. Grundy SM, Brewer HB Jr, Cleeman JI, Smith SC Jr, Lenfant C. Definition of metabolic syndrome: report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. Circulation 2004; 109(3): 433–8. PMID: 14744958
25. Examination Committee of Criteria for 'Obesity Disease' in Japan; Japan Society for the Study of Obesity. New criteria for 'obesity disease' in Japan. Circ J 2002; 66(11): 987–92. PMID: 12419927
26. Shibata K, Suzuki S, Sato J, Ohsawa I, Goto S, Hashiguchi M, et al. Abdominal circumference should not be a required criterion for the diagnosis of metabolic syndrome. Environ Health Prev Med 2010; 15 (4): 229–35. doi: 10.1007/s12199-009-0132-7 PMID: 21432550
27. Japan Diabetes Society Committee on the Diagnostic Criteria of Diabetes Mellitus. Report of the Committee on the classification and diagnostic criteria of diabetes mellitus. J. Jpn. Diabetes Soc. 2010; 53: 450–67.

28. Ministry of Health, Labour and Welfare. The outline of the results of National Health and Nutrition Survey 2008 [in Japanese]. Available at: http://www.mhlw.go.jp/houdou/2008/04/dl/h0430-2c.pdf.

29. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Association of knee osteoarthritis with the accumulation of metabolic risk factors such as overweight, hypertension, dyslipidemia, and impaired glucose tolerance in Japanese men and women: the ROAD study. J Rheumatol 2011; 38(5): 921–30. doi: 10.3899/jrheum.100569 PMID: 21324967

30. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, et al. Accumulation of metabolic risk factors such as overweight, hypertension, dyslipidaemia, and impaired glucose tolerance raises the risk of occurrence and progression of knee osteoarthritis: a 3-year follow-up of the ROAD study. Osteoarthritis Cartilage 2012; 20(11): 1217–26. doi: 10.1016/j.joca.2012.06.006 PMID: 22796312

31. Hart DJ, Doyle DV, Spector TD. Association between metabolic factors and knee osteoarthritis in women: the Chingford study. J Rheumatol 1995; 22(6):1118–23. PMID: 7674240

32. Anekstein Y, Smorgick Y, Lotan R, Agar G, Shalmon E, Floman Y, et al. Diabetes mellitus as a risk factor for the development of lumbar spinal stenosis. Isr Med Assoc J 2010; 12(1): 16–20. PMID: 20450123

33. Okada E, Matsumoto M, Ichihara D, Chiba K, Toyama Y, Fujiwara H, et al. Aging of the cervical spine in healthy volunteers: a 10-year longitudinal magnetic resonance imaging study. Spine 2009; 34(7):706–12. doi: 10.1097/BRS.0b013e31819c2003 PMID: 19333104

34. Matsumoto M, Fujimura Y, Suzuki N, Nishi Y, Nakamura M, Yabe Y, et al. MRI of cervical intervertebral discs in asymptomatic subjects. J Bone Joint Surg Br 1998; 80(1):19–24. PMID: 9460946

35. Balistreri CR, Caruso C, Candore G. The role of adipose tissue and adipokines in obesity-related inflammatory diseases. Mediators Inflamm 2010; 2010:802078. doi: 10.1155/2010/802078 PMID: 20671929

36. Zhao CQ, Liu D, Li H, Jiang LS, Dai LY. Expression of leptin and its functional receptor on disc cells: contribution to cell proliferation. Spine 2008;; 33(23):E858–64. doi: 10.1097/BRS.0b013e31818338e5 PMID: 18978578

37. Gruber HE, Ingram JA, Hoelscher GL, Hanley EN Jr. Leptin expression by annulus cells in the human intervertebral disc. Spine J 2007; 7(4): 437–43. PMID: 17433782

38. Li Z, Shen J, Wu WK, Yu X, Liang J, Qiu G, et al. The role of leptin on the organization and expression of cytoskeleton elements in nucleus pulposus cells. J Orthop Res 2013; 31(6): 847–57. doi: 10.1002/jor.22308 PMID: 23335226

39. Lotan R, Oron A, Anekstein Y, Shalmon E, Mirovsky Y. Lumbar stenosis and systemic disease: is there any relevance? J Spinal Disord Tech 2008; 21(4): 247–51. doi: 10.1097/BSD.0b013e31813707af PMID: 18525484

40. Alexander RW. Theodore Cooper Memorial Lecture. Hypertension and the pathogenesis of atherosclerosis. Oxidative stress and the mediation of arterial inflammatory response: a new perspective. Hypertension 1995; 25(2):155–61. PMID: 7843763

41. Jhawar BS, Fuchs CS, Colditz GA, Stampfer MJ. Cardiovascular risk factors for physician-diagnosed lumbar disc herniation. Spine J 2006; 6(6): 684–91. PMID: 17088199

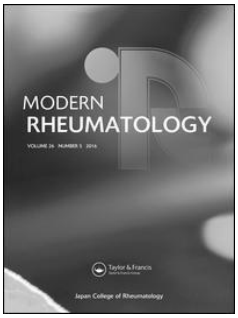
42. Videman T, Battié MC, Gibbons LE, Kaprio J, Koskenvuo M, Kannus P, et al. Disc degeneration and bone density in monozygotic twins discordant for insulin-dependent diabetes mellitus. J Orthop Res 2000; 18(5): 768–72. PMID: 11117299

43. Tsai TT, Ho NY, Lin YT, Lai PL, Fu TS, Niu CC, et al. Advanced glycation end products in degenerative nucleus pulposus with diabetes. J Orthop Res 2014; 32(2):238–44. doi: 10.1002/jor.22508 PMID: 24151186

44. Leino-Arjas P, Kaila-Kangas L, Solovieva S, Riihimäki H, Kirjonen J, Reunanen A. Serum lipids and low back pain: an association? A follow-up study of a working population sample. Spine 2006; 31(9): 1032–7. PMID: 16641781

45. Kauppila LI. Prevalence of stenotic changes in arteries supplying the lumbar spine. A postmortem angiographic study on 140 subjects. Ann Rheum Dis 1997; 56(10): 591–5. PMID: 9389219

46. Kauppila LI, McAlindon T, Evans S, Wilson PW, Kiel D, Felson DT. Disc degeneration/back pain and calcification of the abdominal aorta. A 25-year follow-up study in Framingham. Spine 1997; 22(14): 1642–7. PMID: 9253101



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

Prevalence of hand osteoarthritis and its relationship to hand pain and grip strength in Japan: The third survey of the ROAD study

Rie Kodama, Shigeyuki Muraki, Hiroyuki Oka, Toshiko Iidaka, Masatoshi Teraguchi, Ryohei Kagotani, Yoshiki Asai, Munehito Yoshida, Yutaka Morizaki, Sakae Tanaka, Hiroshi Kawaguchi, Kozo Nakamura, Toru Akune & Noriko Yoshimura

To cite this article: Rie Kodama, Shigeyuki Muraki, Hiroyuki Oka, Toshiko Iidaka, Masatoshi Teraguchi, Ryohei Kagotani, Yoshiki Asai, Munehito Yoshida, Yutaka Morizaki, Sakae Tanaka, Hiroshi Kawaguchi, Kozo Nakamura, Toru Akune & Noriko Yoshimura (2016) Prevalence of hand osteoarthritis and its relationship to hand pain and grip strength in Japan: The third survey of the ROAD study, Modern Rheumatology, 26:5, 767-773, DOI: 10.3109/14397595.2015.1130673

To link to this article: http://dx.doi.org/10.3109/14397595.2015.1130673

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://www.tandfonline.com/imor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2016; 26(5):767–773

DOI: 10.3109/14397595.2015.1130673

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Prevalence of hand osteoarthritis and its relationship to hand pain and grip strength in Japan: The third survey of the ROAD study

Rie Kodama[1], Shigeyuki Muraki[2], Hiroyuki Oka[3], Toshiko Iidaka[4], Masatoshi Teraguchi[5], Ryohei Kagotani[5], Yoshiki Asai[5], Munehito Yoshida[5], Yutaka Morizaki[1], Sakae Tanaka[1], Hiroshi Kawaguchi[6], Kozo Nakamura[7], Toru Akune[7], and Noriko Yoshimura[4]

[1]Department of Orthopedic Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, [2]Department of Clinical Motor System Medicine, [3]Department of Medical Research and Management for Musculoskeletal Pain, [4]Department of Joint Disease Research, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan, [5]Department of Orthopedic Surgery, Wakayama Medical University, Wakayama, Japan, [6]Department of Orthopedic Surgery, Japan Community Health care Organization Tokyo Shinjuku Medical Center, Tokyo, Japan, and [7]National Rehabilitation Center for Persons with Disabilities, Saitama, Japan

Abstract

Objectives: To examine the prevalence and pattern of hand osteoarthritis (HOA), and determine its relationship with grip strength and hand pain.

Methods: Among the participants of the third survey of the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study, 507 Japanese men and 1028 Japanese women were included. Radiographs of both hands were graded for osteoarthritis (OA) using the modified Kellgren–Lawrence (KL) scale. HOA was defined as the presence of at least one affected joint. The absence or presence of subchondral erosion was also scored.

Results: The prevalence of HOA (KL grade $\geq$2) was 89.9% in men and 92.3% in women ($p = 0.11$), and it was significantly associated with age. OA in the distal interphalangeal (DIP) joint was the highest overall. After adjusting for age, sex, body mass index, and the residing area, both severity (KL grade $\geq$3) and erosion were significantly related to low grip strength and hand pain. With regard to the joint groups, severe OA in the DIP and first carpometacarpal joints were related to hand pain.

Conclusion: This study showed a high prevalence of radiographic HOA and a significant relationship between hand pain and the severity of HOA, in addition to erosion.

Keywords

Erosion; Hand, Osteoarthritis, Population-based, Prevalence

History

Received 28 September 2015
Accepted 4 December 2015
Published online 15 February 2016

Introduction

Hand osteoarthritis (HOA) is one of the most common degenerative joint diseases in the elderly throughout the world. It causes chronic pain and functional disabilities that lead to serious problems in one's daily life [1–3]. Moreover, joint swelling and deformities, such as Heberden or Bouchard nodes, may cause serious cosmetic issues, especially in middle-aged women. However, HOA has not received attention until recently because its clinical burden was extremely underestimated. As a result, the pathogenesis of HOA has largely remained unknown. The reported prevalence of radiographic or symptomatic HOA differs considerably among previous population-based epidemiologic studies [2–10]. This may be due to a limitation in the sample size or variabilities in age, ethnicity, or the definition of HOA. In addition, there are few large-scale population-based cohort studies on the prevalence of HOA in Asia [7].

Hand pain is one of the main symptoms of HOA, yet the association of HOA with pain remains controversial [1,11]. In addition, the relationship between HOA and grip strength remains unclear [1,12]. One of the reasons for this may be that in previous studies, the severity of HOA was not examined, despite the fact that the severity of HOA may be important for pain or grip strength. Thus, in the present study, we examined the association of hand pain or grip strength using a Kellgren–Lawrence (KL) grade $\geq$2 HOA and a KL $\geq$3 HOA.

In addition, the idea of erosive HOA has received more attention in recent years. Erosive HOA was defined as a specific subgroup of HOA with subchondral erosion and cortical destruction [13]. This was first described by Peter et al. in 1966 [14], but there were few population-based studies on erosive HOA [10,15–17]. It is still unknown whether erosive HOA is a separate disease entity or a severe form of HOA [18].

This study aimed to (1) examine the prevalence, pattern, and severity of radiographic HOA in addition to erosive HOA in the general Japanese population and (2) determine the associations between the severity of HOA and hand pain or grip strength, as well as erosive HOA.

Materials and methods

Participants

The Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study was a nationwide prospective study on bone and

Correspondence to: Dr. Noriko Yoshimura, Department of Joint Disease Research, 22nd Century Medical and Research Center, Graduate School of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. Tel: +81 3 5800 9178. Fax: +81 3 5800 9179. E-mail: yoshimuran-ort@h.u-tokyo.ac.jp

joint diseases that consisted of population-based cohorts in three communities in Japan: an urban region in Itabashi, Tokyo; a mountainous region in Hidakagawa, Wakayama; and a coastal region in Taiji, Wakayama. Details of the study have been described previously [19,20]. Briefly, residents of these regions were recruited from the resident registration lists. Three thousand forty inhabitants (1061 men and 1979 women) with a mean age of 70.3 years were included in the first survey, which was administered from 2005 to 2007. The study was approved by the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology. Written informed consent was provided by all participants.

The third survey of the ROAD study was administered during 2012–2013. All participants who attended the first and the second surveys were invited to the follow-up survey. In addition, inhabitants aged ≥60 years and residing in the urban area and those aged ≥40 years and residing in the mountainous and coastal areas who were willing to respond to the survey were included. As a result, 2566 subjects participated in the third visit. In the present study, we used data from the 1535 participants in the mountainous and coastal regions who underwent radiography for both hands after excluding 27 people who had a history of rheumatoid arthritis (RA) or took medication for RA, based on their responses to an interviewer-administered questionnaire. All of the participants included in the study were Japanese.

All participants completed an interviewer-administered questionnaire that included information on their medical history, family history, physical activity, joint pain, etc. Those who responded that they had pain in any part of their right or left hand ≥ 1 day were regarded as people who had hand pain. We also measured the participants' height and weight, and their body mass index (BMI) was calculated. Grip strength was measured using a Toei Light handgrip dynamometer (Toei Light Co., Ltd., Saitama, Japan). Grip strength of both hands was measured, and the dominant hand value was used in the analyses. The dominant hand was defined as the hand that is mainly used in daily life (e.g. the hand that commonly does the writing or uses chopsticks and scissors). When we measured grip strength, we asked the participants which hand was their dominant hand.

Radiographic assessment

Anterior–posterior radiographs of both hands were taken for each patient by licensed radiography technicians using standard radiographic techniques. The radiographs were read by one orthopedist (RK). The second to fifth distal interphalangeal (DIP), proximal interphalangeal (PIP), first to fifth metacarpophalangeal (MCP), thumb interphalangeal (IP), and first carpometacarpal (CMC) joints for each hand were graded for osteoarthritis (OA) using the modified KL scale [21], which was used in Framingham and other studies, to assess the existence and severity of osteophytes, joint space narrowing, sclerosis, and erosion [10]. The modified KL scale was graded from 0 to 4, where 0 is no OA; 1 is questionable osteophytes (OPs) and/or joint space narrowing (JSN); 2 is definite small OPs and/or mild JSN; 3 is moderate OPs and/or moderate JSN, sclerosis, and erosions may be present; and 4 is large OPs and/or severe JSN, sclerosis, and erosions may be present. Subchondral erosion, the characteristic central erosion and associated pseudowidening, was also scored for its absence or presence in the DIP, PIP, and first CMC joints according to the atlas by Altman et al. [22].

Radiographic OA was defined as a KL grade ≥2. Severe OA was defined as a KL grade ≥3. HOA or severe HOA at the individual level was defined by the presence of at least one affected joint. Hand-joint groups were similarly defined by the presence of at least one affected joint. The presence of subchondral erosion was defined as erosive HOA.

To investigate the intra-observer reliability of the scale, 20 randomly selected hand radiographs were scored by the same reader, and two orthopedists (RK and HO) also scored the 20 radiographs to assess the inter-observer reliability. The intra- and inter-observer reliability was assessed by the κ statistic, and they were 0.78 and 0.77, respectively.

Statistical analysis

The prevalence of HOA for each joint, the joint groups, and the entire hand was compared between sexes using the Chi-square test. Univariate and multivariate logistic regression analyses were used to examine the association between age, sex, BMI, the residing area, grip strength, and hand pain with HOA, severe HOA, or erosive HOA. Univariate and multivariate regression models were used to analyze the effects of each joint group on hand pain or grip strength. A p value <0.05 was used to indicate a significant difference. All the analyses were performed using JMP, version 11.0 (SAS Institute Inc., Cary, NC).

Results

The characteristics and clinical outcomes of the 1535 participants are shown in Table 1. The mean (standard deviation) age was not significantly different between the sexes. More than 90% of subjects were right-hand dominant. The prevalence of radiographic HOA (KL grade ≥2) in at least one joint among all the joints was >90% in both sexes, and there was no significant difference between the sexes. The prevalence of severe HOA (KL grade ≥3) was approximately 40%, but there was no difference between the sexes.

Figure 1 shows the prevalence of HOA and the severity of HOA classified by age and sex. Both were significantly associated with age ($p<0.05$) in both sexes, and nearly 100% of men and women >70 years had at least one radiographic HOA joint.

The prevalence of OA at the joint level is shown in Table 1. The prevalence of OA was the highest in the DIP joints, followed by the thumb IP, PIP, first CMC, and MCP joints. Figure 2 shows the prevalence of OA observed in each hand joint in men and women in detail. Numbers on the left side show the prevalence of OA in men, and those of the right side show the prevalence of OA in women. OA of the DIP joint occurred more frequently in women than in men ($p<0.05$). OA in the PIP joint also tended to occur more in women than in men, but only the right fifth, left fourth, and fifth PIP joints had significantly more OA in women than in men ($p<0.05$). In contrast, regarding the MCP joints, OA in the right third and fourth as well as the left second and third MCP joints was more common in men than in women ($p<0.01$).

The prevalence of HOA classified by the dominant and non-dominant hands is shown in Table 2. The first to third MCP joints with OA were significantly more frequent in the dominant hand (all, $p<0.0001$). In contrast, the prevalence of the fifth PIP and first CMC joints with OA was higher in the non-dominant hand than in the dominant hand ($p=0.0055$ and 0.0016, respectively).

Regarding erosive HOA, all the people who had more than one joint of erosive HOA were included in the group of severe HOA. They accounted for 10.6% of people who had severe HOA. The prevalence of erosive HOA was higher in women than in men ($p<0.01$) (Table 1).

Table 1 also shows the prevalence of hand pain. When comparing the high prevalence of HOA, the prevalence of hand pain was smaller in both sexes, and there were no significant differences between the sexes. Hand pain was not significantly

Table 1. Participants' characteristics and the clinical outcomes of hand osteoarthritis.

	Total ($n = 1535$)	Men ($n = 507$)	Women ($n = 1028$)	p value (Men versus Women)
Age, years	65.6 ± 13.0	66.3 ± 13.7	65.3 ± 12.6	0.19
BMI, kg/m^2	23.0 ± 3.6	23.6 ± 3.5	22.7 ± 3.6	<0.0001
Residing in the coastal area, %	54.4	48.3	44.3	0.14
Dominant hand (right), %	94.2	92.3	95.1	0.030
Dominant hand grip strength, kg	30.4 ± 9.9	40.5 ± 9.0	25.5 ± 5.8	<0.0001
Radiographic HOA (≥1 joint)	91.5	89.9	92.3	0.11
DIP (2–5) OA	85.5	84	86.3	0.24
PIP (2–5) OA	57.5	50.9	60.7	0.0003
Thumb IP OA	64.2	63.5	64.6	0.68
MCP (1–5) OA	38.2	39.8	37.5	0.36
First CMC OA	50.2	50.5	50	0.86
Severe HOA (≥1 joint)	39.8	37.3	41.1	0.16
Erosive HOA (≥1 joint)	4.2	2.2	5.3	0.0048
Number of HOA joints (KL ≥2)	9.3	8.7	9.6	0.014
Hand pain, %	7.4	5.7	8.2	0.096

BMI, body mass index; HOA, hand osteoarthritis; DIP, distal interphalangeal; PIP, proximal interphalangeal; IP, interphalangeal; MCP, metacarpophalangeal; CMC, carpometacarpal.

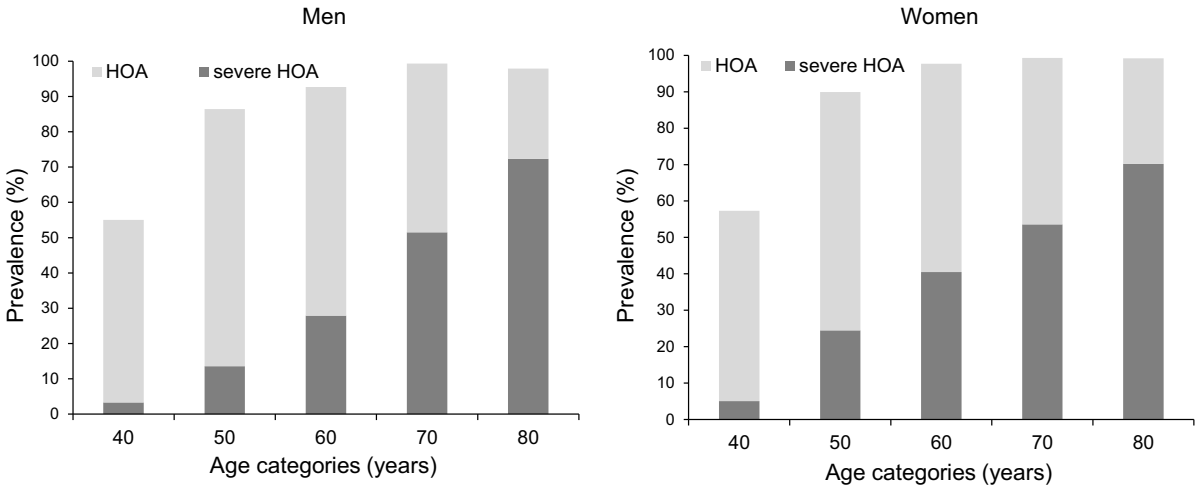


Figure 1. Prevalence of radiographic hand osteoarthritis (HOA) and severe HOA in both sexes among the age categories.

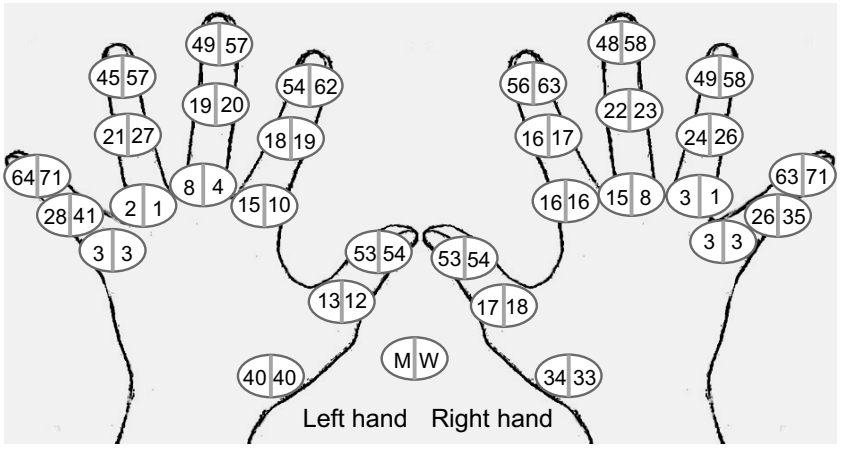


Figure 2. Prevalence (%) of radiographic hand osteoarthritis in each joint in men (M, left) and women (W, right).

different between the dominant and non-dominant hands (Table 2).

Table 3 shows the factors related to the prevalence of HOA, severe HOA, and erosive HOA. Multivariate logistic regression analysis showed that age was significantly associated with HOA and severe HOA; however, sex had no association with them. A higher BMI was significantly related to HOA, but it had no relationship with severe HOA. After adjusting for age, sex, BMI, and the residing area, HOA had no relationship with grip strength or hand pain. Conversely, severe HOA was significantly related to a low grip strength and hand pain compared with non-severe HOA ($KL \leq 2$). To clarify the relationship between hand pain and the severity of HOA, we categorized HOA in three groups: $KL \leq 1$, $KL = 2$, and $KL \geq 3$ HOA. Only severe HOA ($KL \geq 3$) was significantly related to hand pain (Table 4).

Table 2. Comparison of the prevalence of hand osteoarthritis (HOA) and hand pain between the dominant and non-dominant hands.

	Dominant hand	Non-dominant hand	*p* value
Radiographic HOA (≥1 joint), %	87.8	86.7	0.30
DIP5	69.1	68.4	0.67
DIP4	55.3	52.6	0.14
DIP3	55.2	53.8	0.44
DIP2	61.1	58.6	0.17
PIP5	32.5	37.3	0.0055
PIP4	25.3	24.9	0.77
PIP3	22.3	20.1	0.13
PIP2	16.6	19.0	0.080
MCP5	2.7	3.1	0.59
MCP4	1.8	1.4	0.31
MCP3	10.8	4.9	<0.0001
MCP2	16.0	11.1	<0.0001
Thumb IP	53.9	53.1	0.66
MCP1	17.7	11.7	<0.0001
CMC1	34.0	39.5	0.0016
Hand pain	5.2	4.0	0.12

DIP, distal interphalangeal; PIP, proximal interphalangeal; MCP, metacarpophalangeal; IP, interphalangeal; CMC, carpometacarpal.

Age, sex, and BMI were not significantly related to erosive HOA in multivariate logistic regression analysis (Table 3). However, there was a significant relationship between erosive HOA and low grip strength, and hand pain. As severe HOA and erosive HOA were related to hand pain, we separated those who had a $KL \geq 3$ with erosive HOA and those who had a $KL \geq 3$ without erosive HOA. Compared to those with a $KL \leq 1$, both erosive HOA and a $KL \geq 3$ without erosive HOA were significantly related to hand pain after adjusting for age, sex, BMI, the residing area, and grip strength ($p = 0.0001$). Moreover, the odds ratio (OR) for erosive HOA was much higher (OR: 10.25; 95% confidence interval [CI]: 3.07–41.28) than that of a $KL \geq 3$ without erosive HOA (OR: 4.13; 95% CI: 1.42–15.27). When classified based on the KL grade and the presence of erosive HOA, the prevalence of hand pain in the erosive HOA group was much higher than that of the other types ($p < 0.0001$) (Figure 3).

We also examined factors associated with hand pain, and found that hand pain was related to low grip strength ($p = 0.0051$) (Table 4). The association of hand pain with OA in the DIP, PIP,

Table 3. Factors related to the prevalence of hand osteoarthritis (HOA), severe HOA, and erosive HOA.

	HOA ($KL \geq 2$)		Severe HOA ($KL \geq 3$)		Erosive HOA	
	Crude OR (95% CI)	Adjusted OR (95% CI)	Crude OR (95% CI)	Adjusted OR (95% CI)	Crude OR (95% CI)	Adjusted OR (95% CI)
Age (+1 year)	1.14** (1.12–1.16)	1.15** (1.12–1.17)	1.09** (1.08–1.10)	1.08** (1.06–1.09)	1.05** (1.03–1.08)	1.02 (0.99–1.05)
Sex (reference: men)	1.34 (0.93–1.94)	3.18** (1.45–6.94)	1.17 (0.94–1.46)	0.93 (0.62–1.39)	2.50** (1.35–5.08)	0.77 (0.33–1.92)
BMI (+1 kg/m^2)	1.09** (1.04–1.16)	1.11** (1.04–1.19)	1.00 (0.97–1.03)	1.02 (0.99–1.05)	0.93 (0.86–1.01)	0.95 (0.88–1.03)
Residing area (reference: mountainous)	0.78 (0.54–1.12)	1.34 (0.86–2.08)	0.53** (0.43–0.65)	0.68** (0.53–0.85)	0.81 (0.49–1.33)	0.96 (0.57–1.62)
Grip strength (+1 kg)	0.95** (0.93–0.97)	1.03 (1.00–1.07)	0.94** (0.93–0.95)	0.97* (0.95–0.99)	0.89** (0.86–0.92)	0.90** (0.86–0.95)
Hand pain	2.65* (1.09–8.74)	1.85 (0.69–6.54)	2.01** (1.37–2.96)	2.23** (1.45–3.45)	3.80** (1.97–6.93)	3.56** (1.79–6.73)

The adjusted odds ratios (ORs) were calculated using multivariate logistic regression analysis.
BMI, body mass index; CI, confidence interval; KL, Kellgren–Lawrence grade.
*$p < 0.05$.
**$p < 0.01$.

Table 4. Factors related to hand pain.

	Crude OR (95% CI)	*p* value	Adjusted OR (95% CI)	*p* value
Age (+1 year)	1.01 (0.99–1.02)	0.36	0.97 (0.95–1.00)	0.020
Sex (reference: men)	1.47 (0.96–2.30)	0.078	0.67 (0.35–1.34)	0.025
BMI (+1 kg/m^2)	1.06 (1.00–1.11)	0.032	1.07 (1.01–1.12)	0.014
Residing area (reference: mountainous)	1.24 (0.84–1.84)	0.28	1.36 (0.91–2.05)	0.13
Grip strength (+1 kg)	0.97 (0.95–0.99)	0.0027	0.95 (0.92–0.98)	0.0051
HOA				
$KL \leq 1$ HOA	1 (reference)		1 (reference)	
$KL = 2$	1.94 (0.77–6.51)	0.17	2.30 (0.85–8.14)	0.11
$KL \geq 3$ HOA	3.62 (1.46–12.08)	0.0035	4.82 (1.67–17.69)	0.0024

The adjusted odds ratios (ORs) were calculated using multivariate logistic regression analysis.
BMI, body mass index; HOA, hand osteoarthritis; CI, confidence interval.

MCP, thumb IP, and first CMC joints was examined separately, and the DIP and first CMC joints with severe OA were found to have a significant association with hand pain ($p = 0.010$ and $p = 0.0043$) (Table 5). With regard to grip strength, severe OA of any joint group was significantly related to low grip strength after adjusting for age, sex, BMI, the residing area, and hand pain (all, $p < 0.01$).

Discussion

This is the first population-based study to examine the prevalence and patterns of HOA in Japanese men and women in detail. This study showed a high prevalence of radiographic HOA in Japanese elderly and a relationship between the HOA severity and hand pain, and low grip strength. We also found that age and BMI were related to HOA. With regard to the joint groups, the DIP and first CMC joints with severe OA were related to hand pain. We also showed the prevalence of erosive HOA in this Japanese population and found that hand pain was significantly related to erosive HOA. Hand pain was also related to low grip strength.

In the present study, the prevalence of radiographic HOA was 89.9% in men and 92.3% in women, which are much higher than those reported in previous studies in the United States, Europe, and Asia [2,5–10]. As there may be some differences in the participants' number, age distribution, or sex ratio, we cannot compare the prevalence among cohorts directly. However, it may be suggested that ethnicity is the main cause of the difference in the prevalence among the studies. We previously showed that the prevalence of knee OA in our cohorts was much higher than that in Caucasians [20]. Japanese individuals may have a high prevalence of OA in the knee and hand joints.

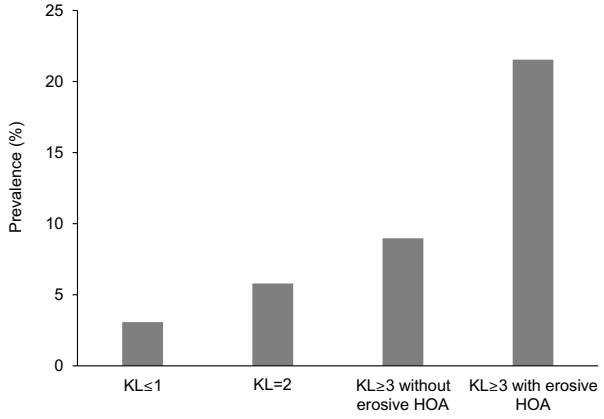


Figure 3. Prevalence (%) of hand pain in each subgroup classified by the Kellgren–Lawrence (KL) grade and the presence of erosive hand osteoarthritis (HOA).

The association between HOA and hand pain remains controversial. Zhang et al. reported that the severity of HOA is related to hand pain, although the details were not described [1]. Dahaghin et al. reported that radiographic HOA was a poor explanation for hand pain ($R^2 = 0.005$) in a multivariate model from the Rotterdam study [11]. In the present study, we focused on KL grade ≥ 2 HOA and severe HOA (KL ≥ 3), and we found that although there was no significant relationship between KL grade ≥ 2 HOA and hand pain, or grip strength, severe HOA had a significant relationship with these variables according to the multivariate logistic model. The severity of HOA may be important for hand pain or low grip strength. Few studies have focused on the relationship between the severity of HOA and hand pain or disabilities [1]. We also determined the relationship between hand pain and low grip strength after adjusting for other factors. Zhang et al. reported that subjects with symptomatic HOA had reduced maximal grip strength (by 10%) [1]. As our study was a cross-sectional one, we could not confirm causality between low grip strength, hand pain, and the severity of HOA. Our future longitudinal study will be able to clarify the relationship among these factors.

We also examined erosive HOA, which was defined as the presence of subchondral erosion. To the best of our knowledge, this is the first large-scale cohort study to show the prevalence of erosive HOA in Asia, and we found discrepancies between erosive HOA and HOA, as defined by the KL grade. First, although the prevalence of KL grade ≥ 2 HOA in the present study was much higher than that in previous studies, the prevalence of erosive HOA in the present study was similar to that in previous studies [10,15,16]. This indicates that erosive HOA may have a distinct aetiology from severe HOA because if erosive HOA is a severe form of HOA, the percentage of erosive HOA of HOA in our cohort should be larger than those of other cohorts. Second, although the prevalence of severe HOA was significantly associated with age, we could not find a significant relationship between age and erosive HOA, which may also indicate the distinct aetiology between erosive HOA and severe HOA. In fact, erosive HOA was associated with several genetic factors [23,24]. Third, we also found a strong relationship between erosive HOA and hand pain, which may indicate that erosion was more strongly related to hand pain than the severity of HOA. Previous studies have shown that the prevalence of hand pain was significantly higher in subjects with erosive HOA than in those without erosive HOA [10,15,16], but no study has analyzed the severity of HOA and erosive HOA separately. As all the erosive HOA cases were included in severe HOA, we could not thoroughly conclude that erosive HOA was not a severe form of HOA; however, there may be different aetiologies between severe HOA with and without erosion.

The pattern of joint involvement in our cohorts was different from that of other studies. Caucasians had the highest prevalence of OA in the first CMC joints, but the prevalence of first CMC joints with OA was the second lowest in the present study [10]. A former population-based cohort study in Japan with a small sample

Table 5. Association between hand pain and severe hand osteoarthritis in the joint groups.

	Crude OR (95% CI)	*p* value	Adjusted OR (95% CI)	*p* value
DIP joint	1.78 (1.20–2.63)	0.0041	1.76 (1.15–2.70)	0.010
PIP joint	1.75 (0.98–2.94)	0.058	1.58 (0.86–2.76)	0.13
MCP joint	0.29 (0.015–1.85)	0.059	2.51 (0.96–5.81)	0.059
Thumb IP joint	0.88 (0.45–1.57)	0.68	0.85 (0.42–1.58)	0.62
First CMC joint	2.41 (1.42–3.94)	0.0016	2.31 (1.30–3.92)	0.0043

Adjusted odds ratios (ORs) were calculated using multiple logistic regression analysis after adjusting for age, sex, body mass index, the residing area, and dominant hand grip strength.
DIP, distal interphalangeal; PIP, proximal interphalangeal; MCP, metacarpophalangeal; IP, interphalangeal; CMC, carpometacarpal; CI, confidence interval.

size also showed that the prevalence of OA at the first CMC joint was less in the Japanese cohort than in American women (OR: 0.15; 95% CI: 0.11–0.22) [4]. This may be partly explained by environmental factors (e.g. using chopsticks, which is part of the Japanese lifestyle). In fact, Hunter et al. reported that the prevalence of OA in the thumb IP and second and third PIP and MCP joints in the hands that used the chopsticks was higher than that in hands that did not use the chopsticks [25]. Hand pain was related to severe OA in the DIP and first CMC joints. This result is comparable to that of the Framingham study, which noted that the prevalence of symptomatic OA was higher in the DIP, PIP, and first CMC joints [1].

We also found that BMI was related to HOA. As hand joints are not weight-bearing joints, the relationship between BMI and HOA cannot only be explained by mechanical factors; thus, some metabolic factors may influence HOA. In fact, there were connections between adipokines and knee OA, independently of animal weight [26]. Yusuf et al. reported that adiponectin levels were associated with the progression of HOA [27]. HOA may also be explained by these adipokines, which influence the joint cartilage. As we have collected the participants' blood samples and the interviewer-administered questionnaires that include nutritional information, further analyses may reveal the relationship between HOA and BMI.

The present study has several limitations. First, there was some selection bias in our cohort because we excluded those who could not come to the survey site and those who could not understand or sign informed consent form [19]. Second, although we used the same radiographic atlas and definition that former studies had used to read our radiographs, strict comparisons among our results and other studies may be limited because of the differences in readers [1,21,22]. Third, we assessed the first CMC joints using anterior–posterior radiographs of the hand. Strictly speaking, it may have been better to use lateral view radiographs to assess OA in the first CMC joints [28]. However, the intra- and inter-observer reliability for the first CMC joints in our study was not bad (0.70 and 0.64, respectively). Furthermore, in previous cohort studies, the anterior–posterior view was used to assess OA in all of the hand joints [1,2,4,5,7,8]; thus, comparing the prevalence of OA among them might be useful.

In conclusion, the present study showed a high prevalence of radiographic HOA in the Japanese elderly. Severe HOA defined as KL grade $\geq$3 was significantly related to grip strength and hand pain. In addition, hand pain had a relationship with severe HOA, particularly in the DIP and first CMC joints. We also showed the prevalence of erosive HOA and found a strong relationship between erosive HOA and hand pain. Further studies, along with continued longitudinal surveys from the ROAD study, will help to elucidate the environmental backgrounds of HOA, including erosive HOA and its relationship to hand pain or grip strength.

Acknowledgments

The authors wish to thank Dr. Takako Nojiri and Mr. Kazuhiro Hatanaka of the Gobo Public Health Center; Dr. Naoki Hirabayashi of the Kawakami Clinic in Hidakagawa Town; Mmes. Tomoko Takijiri, Rie Takiguchi, and Kyoko Maeda and other members of the public office in Hidakagawa Town; Dr. Shinji Matsuda of the Shingu Public Health Center; and Mmes. Tamako Tsutsumi, Kanami Maeda, Megumi Takino, Shuko Okada, Kazuyo Setoh, Chise Ryouno, Miki Shimosaki, Chika Yamaguchi, and Yuki Shimoji and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. We would also like to thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, Mrs. Saeko Sahara, and Mr. Noriyuki Oe for their assistance in data reduction.

Conflict of interest

This work was supported by a Grant-in-Aid for H17-Men-eki-009 (Director Kozo Nakamura), H20-Choujyu-009 (Director Noriko Yoshimura), H23-Choujyu-002 (Director Toru Akune), H-25-Choujyu-007 (Director Noriko Yoshimura), and H25-Nanchitou (Men)-005 (Director Sakae Tanaka) from the Ministry of Health, Labor, and Welfare; Scientific Research B26293139, B23390172, and B20390182 and Challenging Exploratory Research 24659317 to Noriko Yoshimura; B23390356 and C20591774, and Challenging Exploratory Research 23659580 to Shigeyuki Muraki; Challenging Exploratory Research 24659666 and 21659349 and Young Scientists A18689031 to Hiroyuki Oka; B23390357 and C20591737 to Toru Akune; and Collaborating Research with NSF 08033011-00262 (Director Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan. This study was also supported by grants from the Japan Osteoporosis Society (Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, and Toru Akune) and research aid from the Japanese Orthopedic Association (Subsidized Science Project Research 2006-1 and 2010-2; Director Hiroshi Kawaguchi).

References

1. Zhang Y, Niu J, Kelly-Hayes M, Chaisson CE, Aliabadi P, Felson DT. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly: the Framingham Study. Am J Epidemiol. 2002;156:1021–7.
2. Dahaghin S, Bierma-Zeinstra SM, Ginai AZ, Pols HA, Hazes JM, Koes BW. Prevalence and pattern of radiographic hand osteoarthritis and association with pain and disability (the Rotterdam study). Ann Rheum Dis. 2005;64:682–7.
3. Dillon CF, Hirsch R, Rasch EK, Gu Q. Symptomatic hand osteoarthritis in the United States: prevalence and functional impairment estimates from the third US National Health and Nutrition Examination Survey, 1991–1994. Am J Phys Med Rehabil. 2007;86:12–21.
4. Yoshida S, Aoyagi K, Felson DT, Aliabadi P, Shindo H, Takemoto T. Comparison of the prevalence of radiographic osteoarthritis of the knee and hand between Japan and the United States. J Rheumatol. 2002;29:1454–8.
5. Haara MM, Manninen P, Kroger H, Arokoski JP, Karkkainen A, Knekt P, et al. Osteoarthritis of finger joints in Finns aged 30 or over: prevalence, determinants, and association with mortality. Ann Rheum Dis. 2003;62:151–8.
6. Mannoni A, Briganti MP, Di Bari M, Ferrucci L, Costanzo S, Serni U, et al. Epidemiological profile of symptomatic osteoarthritis in older adults: a population based study in Dicomano, Italy. Ann Rheum Dis. 2003;62:576–8.
7. Zhang Y, Xu L, Nevitt MC, Niu J, Goggins JP, Aliabadi P, et al. Lower prevalence of hand osteoarthritis among Chinese subjects in Beijing compared with white subjects in the United States: the Beijing Osteoarthritis Study. Arthritis Rheum. 2003;48:1034–40.
8. Bernard TE, Wilder FV, Aluoch M, Leaverton PE. Job-related osteoarthritis of the knee, foot, hand, and cervical spine. J Occup Environ Med. 2010;52:33–8.
9. Kalichman L, Li L, Batsevich V, Malkin I, Kobyliansky E. Prevalence, pattern and determinants of radiographic hand osteoarthritis in five Russian community-based samples. Osteoarthritis Cartilage. 2010;18:803–9.
10. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis. 2011;70:1581–6.
11. Dahaghin S, Bierma-Zeinstra SM, Reijman M, Pols HA, Hazes JM, Koes BW. Prevalence and determinants of one month hand pain and hand related disability in the elderly (Rotterdam study). Ann Rheum Dis. 2005;64:99–104.

12. Chaisson CE, Zhang Y, Sharma L, Kannel W, Felson DT. Grip strength and the risk of developing radiographic hand osteoarthritis: results from the Framingham Study. Arthritis Rheum. 1999;42:33–8.
13. Zhang W, Doherty M, Leeb BF, Alekseeva L, Arden NK, Bijlsma JW, et al. EULAR evidence-based recommendations for the diagnosis of hand osteoarthritis: report of a task force of ESCISIT. Ann Rheum Dis England. 2009;68:8–17.
14. Peter JB, Pearson CM, Marmor L. Erosive osteoarthritis of the hands. Arthritis Rheum. 1966;9:365–88.
15. Kwok WY, Kloppenburg M, Rosendaal FR, van Meurs JB, Hofman A, Bierma-Zeinstra SM. Erosive hand osteoarthritis: its prevalence and clinical impact in the general population and symptomatic hand osteoarthritis. Ann Rheum Dis. 2011;70:1238–42.
16. Kwok WY, Kloppenburg M, Marshall M, Nicholls E, Rosendaal FR, van der Windt DA, et al. Comparison of clinical burden between patients with erosive hand osteoarthritis and inflammatory arthritis in symptomatic community-dwelling adults: the Keele clinical assessment studies. Rheumatology (Oxford). 2013;52:2260–7.
17. Magnusson K, Hagen KB, Osteras N, Nordsletten L, Natvig B, Haugen IK. Diabetes is associated with increased hand pain in erosive hand osteoarthritis: data from a population-based study. Arthritis Care Res (Hoboken). 2015;67:187–95.
18. Kloppenburg M, Kwok WY. Hand osteoarthritis – a heterogeneous disorder. Nat Rev Rheumatol. 2012;8:22–31.
19. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: Research on Osteoarthritis/ Osteoporosis Against Disability study. Int J Epidemiol England. 2010;39:988–95.
20. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. Osteoarthritis Cartilage. 2009;17:1137–43.
21. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16:494–502.
22. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage. 2007;15(Suppl A):A1–56.
23. Stern AG, de Carvalho MR, Buck GA, Adler RA, Rao TP, Disler D, et al. Association of erosive hand osteoarthritis with a single nucleotide polymorphism on the gene encoding interleukin-1 beta. Osteoarthritis Cartilage. 2003;11:394–402.
24. Ramonda R, Musacchio E, Campana C, Frigato M, Frallonardo P, Barbieri V, et al. Immunogenetic aspects of erosive osteoarthritis of the hand in patients from northern Italy. Scand J Rheumatol. 2011;40:139–44.
25. Hunter DJ, Zhang Y, Nevitt MC, Xu L, Niu J, Lui LY, et al. Chopstick arthropathy: the Beijing Osteoarthritis Study. Arthritis Rheum. 2004;50:1495–500.
26. Berenbaum F, Eymard F, Houard X. Osteoarthritis, inflammation and obesity. Curr Opin Rheumatol. 2013;25:114–8.
27. Yusuf E, Ioan-Facsinay A, Bijsterbosch J, Klein-Wieringa I, Kwekkeboom J, Slagboom PE, et al. Association between leptin, adiponectin and resistin and long-term progression of hand osteo-arthritis. Ann Rheum Dis. 2011;70:1282–4.
28. Eaton RG, Lane LB, Littler JW, Keyser JJ. Ligament reconstruction for the painful thumb carpometacarpal joint: a long-term assessment. J Hand Surg Am. 1984;9:692–9.

Patient Satisfaction with Posterior Decompression Surgery for Cervical Ossification of the Posterior Longitudinal Ligament: Prognostic Radiographic Factors and Patient-Reported Outcomes for the Effectiveness of Surgical Treatment

Junichi Ohya[1], Yasushi Oshima[1], Hiroyuki Oka[2], Fumiko Saiki[1], Yuki Taniguchi[1], Yoshitaka Matsubayashi[1], Sakae Tanaka[1], Hirotaka Chikuda[1], Katsushi Takeshita[1,3]

OBJECTIVE: **To identify the prognostic factors associated with patient satisfaction after double-door laminoplasty for cervical compression myelopathy due to ossification of the posterior longitudinal ligament (OPLL).**

METHODS: **The study group comprised 44 patients (30 males and 14 females) with OPLL who underwent double-door laminoplasty at our institution with a minimum follow-up of 1 year. The mean patient age was 63.8 years (range, 48–86 years). We evaluated the patients' postoperative satisfaction using a questionnaire and divided them into 2 groups, satisfied and dissatisfied. We assessed various radiographic parameters. The patient-reported outcomes, including the Short Form-36 Physical Component Summary (SF-36 PCS), Neck Disability Index, neck pain, arm pain, and Japanese Orthopedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ), were assessed and used to evaluate the effectiveness of surgical treatment according to the concept of minimum clinically important difference (MCID).**

RESULTS: **The satisfied group comprised 29 patients (65.9%). The dissatisfied group had a higher percentage of hill-shaped ossifications compared with the satisfied group (46.7% vs 17.2%; $P = 0.04$). The satisfied group had a higher proportion of patients with SF-36 PCS reaching the MCID threshold value (81.8% vs 14.3%; $P < 0.01$) and with effective surgical treatment as evaluated by the JOACMEQ lower extremity function domain (61.5% vs 10.0%; $P < 0.01$).**

CONCLUSION: **Patient satisfaction after laminoplasty was insufficient in patients with a hill-shaped ossification. The patients with OPLL who were able to recognize a difference in their clinical physical function, especially lower extremity function, were satisfied after laminoplasty.**

INTRODUCTION

Ossification of the posterior longitudinal ligament (OPLL), characterized by progressive heterotopic ossification within the spinal ligament, has been recognized as an important pathology causing cervical compressive myelopathy not only in Asian populations, but also in Europe and North America.[1,2] Although the main treatment for cervical myelopathy caused by OPLL is surgery, several unique characteristics of this disorder make the appropriate surgical approach controversial. The advantages of an anterior approach are that the excision of OPLL enables direct decompression of the spinal cord, and that spinal fusion can immobilize dynamic factors to maintain a suitable

Key words
- Cervical myelopathy
- Laminoplasty
- Minimum clinically important difference
- Ossification of the posterior longitudinal ligament
- Patient satisfaction
- Surgical outcome

Abbreviations and Acronyms
BF: Bladder function
CF: Cervical spine function
HRQOL: Health-related quality of life
JOACMEQ: Japanese Orthopedic Association Cervical Myelopathy Evaluation Questionnaire
LEF: Lower extremity function
MCID: Minimum clinically important difference
NDI: Neck Disability Index
OPLL: Ossification of the posterior longitudinal ligament
PRO: Patient-reported outcome
QOL: Quality of life
SCB: Substantial clinical benefit
SF-36 PCS: Short Form-36 Physical Component Summary
UEF: Upper extremity function

From the Departments of [1]Orthopedic Surgery and [2]Medical Research and Management of Musculoskeletal Pain, The University of Tokyo, Tokyo; and [3]Department of Orthopedic Surgery, Jichi Medical University, Tochigi, Japan

To whom correspondence should be addressed: Junichi Ohya, M.D. [E-mail: jun.ohya@gmail.com]

Citation: World Neurosurg. (2016) 96:272-279.
http://dx.doi.org/10.1016/j.wneu.2016.09.011

Journal homepage: www.WORLDNEUROSURGERY.org

Available online: www.sciencedirect.com

cervical alignment; however, this approach is technically demanding and has a higher potential risk of complications, such as nonunion, graft dislodgment, dural tear, and neurologic deterioration.[3-5] In contrast, a posterior approach, such as laminoplasty, is a relatively safe technique that can cope with multisegmental OPLL and comorbid developmental spinal canal stenosis. Nonetheless, there have been some cases with limited surgical outcomes after laminoplasty, and several risk factors for a poor surgical outcome also have been reported, including kyphotic cervical alignment,[6,7] severe occupation of the spinal canal,[6-10] hill-shaped ossification,[9] negative K-line,[11] intramedullary high signal intensity,[6,10,12] and dynamic factors, such as hypermobility of the cervical spine.[12-14]

Patient-reported outcomes (PROs) and health-related quality of life (HRQOL) are becoming the main outcomes used in rigorous studies of the efficacy of treatment after spinal surgery. Recent studies have focused on the patients' viewpoints and the finite medical resources. Research using the concept of the minimum clinically important difference (MCID) has been increasingly performed to reassess the efficacy of spinal surgery.[15-21] Because the MCID indicates the smallest change in an outcome measure that reflects a clinically meaningful improvement for patients,[15,22] surgical treatment for spinal disorders able to reach MCID threshold values justifies its incorporation into clinical practice.[19] Therefore, based on this trend, the efficacy of laminoplasty for cervical compressive myelopathy due to OPLL should be reevaluated based on assessment using PROs, HRQOL, and the MCID.

Although several studies on self-reported postoperative satisfaction after laminoplasty in patients with cervical compressive myelopathy have been reported to date,[23-26] none of these has investigated prognostic radiographic factors, which are reportedly important factors affecting the choice of surgical options for patients with OPLL,[7-14] with regard to PROs. The aim of this retrospective study was to identify the prognostic factors associated with patient satisfaction after double-door laminoplasty for cervical compression myelopathy due to OPLL.

METHODS

Data Source

We retrospectively reviewed all patients with OPLL who underwent double-door laminoplasty between April 2003 and November 2013 in our institution and who completed both preoperative and postoperative questionnaires. The minimum duration of follow-up was 1 year. Patient characteristics and perioperative surgical data were obtained from medical charts. The patients without myelopathy who underwent surgical treatment with an expectation of improved radiculopathy were excluded from the study. Patients with acute spinal cord injury were excluded as well. In addition, 2 patients with OPLL undergoing laminoplasty during the observation period were excluded because they were lost to follow-up. Informed consent was obtained from each patient, and the study was approved by the Institutional Review Board of The University of Tokyo.

Regarding surgical treatment for the patients with OPLL, surgeons used their preferred surgical method according to individual patient characteristics.

Double-Door Laminoplasty

We performed double-door laminoplasty as described previously.[26,27] The cervical laminae were exposed laterally to the medial aspect of the facet joints, and the interspinous ligaments were removed. The spinous processes were split sagittally. Once bilateral gutters for the hinges were carefully created at the transitional area between the facet joint and laminae, spinal canal enlargement was achieved via a bilateral opening of the laminae. HA spacers (Boneceram; Olympus Terumo Biomaterials, Tokyo, Japan) were placed between the opened laminae and fixed with nonabsorbable sutures. The patient wore a soft cervical orthosis for approximately 3 weeks after the surgery.

Radiographic Parameters

All patients underwent radiography, computed tomography (CT), and magnetic resonance imaging (MRI) preoperatively. In the postoperative period, radiographs were obtained routinely at each follow-up visit. Postoperative MRI was performed several weeks after surgery once patient consent was obtained.

Based on the preoperative findings from standard lateral radiographs, the OPLL type, OPLL shape, C2–C7 angle, and K-line were evaluated. OPLL was classified into 4 types: continuous, segmental, mixed, and other (circumscribed or localized).[28-30] According to a previous report on the sagittal shape of the ossified lesion,[9] the presence of a hill-shaped OPLL was specified. When the C2–C7 angle had a positive value, this meant that the cervical alignment was lordotic. A C2–C7 angle $<-5^\circ$ was considered to indicate the presence of kyphosis.[10] The K-line was defined as the straight line connecting the midpoints of the spinal canal at C2 and C7; a negative K-line meant that the OPLL exceeded the K-line and grew beyond it.[11] We used preoperative sagittal CT of the cervical spine to measure the ratio of OPLL occupying the canal, which was calculated as the ratio of the maximum anteroposterior thickness of the OPLL to the anteroposterior diameter of the spinal canal at the corresponding level. A value >60% was defined as a severe occupying ratio.[9] The presence of intramedullary high intensity in the spinal cord was assessed using preoperative T2-weighted MRI of the cervical spine.

Postoperative radiographic evaluations, including the C2–C7 angle and residual anterior compression of the spinal cord, were examined as well. Postoperative kyphosis was defined as a C2–C7 angle $<-5^\circ$. Anterior compression of the spinal cord after laminoplasty was evaluated using postoperative MRI. According to previous reports,[31] the criteria for defining significant were as follows: 1) effacement of anterior cerebral spinal fluid buffer on the T2 sagittal and axial images, and 2) evidence of anterior compression of cord substance on the T1 sagittal and axial images. If both criteria were satisfied, then the MRI findings were considered to indicate the presence of anterior compression of the spinal cord.

PROs

We assessed the preoperative PROs from the questionnaires administered before surgery during hospital admission. The questionnaires included several PROs, including the Neck Disability Index (NDI),[32] the Short Form-36 Physical Component Summary (SF-36 PCS),[33] numeric rating scales of pain in the neck

and arms, and the Japanese Orthopedic Association Cervical Myelopathy Evaluation Questionnaire (JOACMEQ).[34] The JOACMEQ was used to evaluate patient-reported neurologic function and HRQOL in 5 areas: cervical spine function (CF), upper extremity function (UEF), lower extremity function (LEF), bladder function (BF), and quality of life (QOL). Postoperatively, questionnaires that included the aforementioned PROs, in addition to the original satisfaction scales that assessed postoperative outcome, were sent to each patient.

Patient satisfaction was evaluated based on a 7-point scale as reported previously[26]: very satisfied, satisfied, slightly satisfied, neither satisfied nor dissatisfied, slightly dissatisfied, dissatisfied, and very dissatisfied. Based on this evaluation, the patients were divided into 2 groups: satisfied (comprising very satisfied, satisfied, and slightly satisfied) and dissatisfied (comprising neither satisfied nor dissatisfied, slightly dissatisfied, dissatisfied, and very dissatisfied). Postoperative PROs were evaluated using the questionnaires administered at the latest follow-up examination.

Assessment of the Effectiveness of Surgical Treatment

We considered that all of the PROs (except the JOACMEQ) that reached the MCID threshold values indicated that the surgical treatment was effective. Based on a previous report,[15] the MCID threshold values for each PRO were set as follows: 7.5 for the NDI, 4.1 for the SF-36 PCS, and 2.5 for arm and neck pain. According to this MCID concept, we defined patients with an 8-point decrease in the NDI, a 4.1-point increase in the PCS, and a 3-point decrease in arm or neck pain as having undergone effective surgical treatment.

Evaluation of the therapeutic effect using the JOACMEQ has been described previously.[34] According to this evaluation, we defined effective surgical treatment for each domain of the JOACMEQ as follows: 1) the posttreatment score was higher than the pretreatment score by $\geq$20 points, and 2) the pretreatment score was $<$90 and the posttreatment score was $\geq$90. Patients with both preoperative and postoperative scores $>$90 were excluded from this analysis.

Assessment of Objective Neurologic Function

We evaluated preoperative baseline cervical compressive myelopathy using the conventional doctor-based Japanese Orthopedic Association (JOA) score.[35]

Statistical Analysis

All PROs and the effectiveness of surgical treatment were compared between the satisfied and dissatisfied groups. Continuous outcomes were compared using a 1-factor analysis of variance, and categorical outcomes were compared using the χ^2 test and Fisher's exact test. The Jonckheere–Terpstra test was used to identify associations between the duration of follow-up and the satisfaction rating. All statistical analyses except the Jonckheere–Terpstra test were performed using JMP PRO version 11 (SAS Institute Japan, Tokyo, Japan). The Jonckheere–Terpstra test was performed using SPSS version 23 (IBM Japan, Tokyo, Japan). The threshold for significance was $P < 0.05$.

RESULTS

The study cohort comprised 44 consecutive patients (30 males and 14 females), with a mean age of 63.8 years (range, 48–86 years). The mean duration of follow-up was 23.8 months (range, 12–89 months). The mean preoperative JOA score was 10.9 points (range, 6–15 points). The most common surgical decompression level was C3–C7 (in 20 patients), followed by C2–C7 (in 16 patients), and C3–T1 (in 3 patients). The patient characteristics and radiographic parameters are presented in Table 1. The mean C2–C7 angle was 6.4° (range, 27° to −27°). The most common type of OPLL was the localized type (34.1%), followed by the continuous type (25.0%) and the mixed type (22.7%). A hill-shaped ossification was seen in 12 patients (27.3%). A negative K-line was seen in 14 patients (31.8%). The mean occupying ratio of ossification was 46.1% (range, 26%–73%). An intramedullary high signal intensity on cervical MRI was seen in 39 patients (88.7%). Postoperatively, the mean C2–C7 angle was 6.8° (range,

Table 1. Baseline and Radiographic Characteristics

Characteristic	Value
Total number of patients	44
Age, years, mean ± SD (range)	63.8 ± 8.3 (48–86)
Sex, male, *n* (%)	30 (68.2)
Follow-up, months, mean ± SD (range)	23.8 ± 14.4 (12–89)
Preoperative JOA score, mean ± SD (range)	10.9 ± 2.2 (6–15)
Decompression level, *n* (%)	
C3–C7	20 (45.5)
C2–C7	16 (35.6)
C3–T1	3 (6.8)
Other	5 (11.4)
Preoperative radiographic parameters	
C2–C7 angle, degrees, mean ± SD (range)	6.4 ± 11.2 (27 to −27)
Type of ossification, *n* (%)	
Continuous	11 (25.0)
Segmental	8 (18.2)
Mixed	10 (22.7)
Others	15 (34.1)
Hill-shaped ossification, *n* (%)	12 (27.3)
K-line (−), *n* (%)	14 (31.8)
Occupation ratio, %, mean ± SD (range)	46.1 ± 10.9 (26–73)
Presence of IHSI on MRI, *n* (%)	39 (88.7)
Postoperative radiographic parameters	
C2–C7 angle, degrees, mean ± SD (range)	6.8 ± 12.1 (33 to −32)
Residual ACS on MRI, *n* (%)	14 (33.3)

JOA, Japanese Orthopedic Association; IHSI, intramedullary high signal intensity; MRI, magnetic resonance imaging; ACS, anterior compression of the spinal cord.

33° to −32°). Among the 42 patients who underwent postoperative MRI, 14 (33.3%) showed residual anterior compression of the spinal cord.

One patient exhibited neurologic deterioration owing to collapse of the cervical lamina at 7 months after the index surgery, necessitating further decompression surgery. No patient experienced any other perioperative complications.

Preoperative and postoperative PROs are summarized in Table 2. Several postoperative PROs, including the SF-36 PCS, arm pain, and all JOACMEQ domains except bladder function, improved significantly compared with preoperative values (from 24.5 ± 17.5 to 34.3 ± 15.4 [$P < 0.01$], from 4.1 ± 3.3 to 2.8 ± 2.6 [$P < 0.01$], from 55.5 ± 35.1 to 64.9 ± 30.6 [$P < 0.01$], from 73.6 ± 21.3 to 84.8 ± 13.3 [$P < 0.01$], from 56.6 ± 25.9 to 72.8 ± 19.6 [$P < 0.001$], and from 44.4 ± 18.0 to 52.5 ± 16.9 [$P = 0.01$], respectively). The mean change between the preoperative and postoperative periods in the SF-36 PCS reached the MCID, whereas the changes in the NDI, arm pain, and neck pain did not reach the MCID.

Overall, the satisfied group comprised 29 patients (65.9%), including 7 patients who were very satisfied, 16 who were satisfied, and 6 who were slightly satisfied. The dissatisfied group comprised the remaining 15 patients (34.1%), including 13 patients who were neither satisfied nor dissatisfied, 1 who was slightly dissatisfied, and 1 who was dissatisfied. No patients reported being very dissatisfied.

A comparison of baseline and radiographic characteristics between the satisfied and dissatisfied groups is shown in Table 3. The baseline data, including age, sex, follow-up period, and preoperative JOA score, did not differ significantly between the 2 groups. Hill-shaped ossifications were present in 5 patients (17.2%) in the satisfied group, compared with 7 patients (46.7%) in the dissatisfied group ($P = 0.04$). The satisfied group had an insignificantly higher proportion of continuous-type OPLL compared with the dissatisfied group (34.5% vs. 6.7%; $P = 0.13$). None of the other radiographic factors, including kyphosis, a negative K-line, a severe occupying ratio, an intramedullary high signal intensity, or anterior compression of the spinal cord, was significantly different between the 2 groups.

We also analyzed the effectiveness of surgical treatment based on each PRO between the 2 groups, as shown in Table 4. Compared with the dissatisfied group, the satisfied group had a higher proportion of patients with the SF-36 PCS reaching the MCID ($P < 0.01$) and a higher level of effective surgical treatment based on the JOACMEQ LEF ($P < 0.01$). The satisfied group had a higher portion of patients with the NDI reaching the MCID compared with the dissatisfied group (50.0% vs. 18.2%), but the difference was not significant ($P = 0.14$). In addition, the satisfied

Table 2. Preoperative and Postoperative Patient-Reported Outcomes

Patient-Reported Outcome	Preoperative	Postoperative	*P* Value	MCID
SF-36 PCS, mean ± SD	24.5 ± 17.5	34.3 ± 15.4	<0.01	Reached
NDI, mean ± SD	30.2 ± 16.1	25.1 ± 16.3	0.09	Not reached
Neck pain, mean ± SD	3.2 ± 3.3	2.8 ± 2.7	0.47	Not reached
Arm pain, mean ± SD	4.1 ± 3.3	2.8 ± 2.6	<0.01	Not reached
JOACMEQ, mean ± SD				
CF	55.5 ± 35.1	64.9 ± 30.6	<0.01	
UEF	73.6 ± 21.3	84.8 ± 13.3	<0.01	
LEF	56.6 ± 25.9	72.8 ± 19.6	<0.001	
BF	70.8 ± 21.2	76.4 ± 19.1	0.15	
QOL	44.4 ± 18.0	52.5 ± 16.9	0.01	
Satisfaction, *n* (%)				
Very satisfied		7 (15.9)		
Satisfied		16 (36.4)		
Slightly satisfied		6 (16.7)		
Neither satisfied nor dissatisfied		13 (29.5)		
Slightly dissatisfied		1 (2.3)		
Dissatisfied		1 (2.3)		
Very dissatisfied		0 (0.0)		

MCID, minimum clinically important difference; SF-36 PCS, Short Form 36 Physical Component Summary; NDI, Neck Disability Index; JOACMEQ, Japanese Orthopedic Association Cervical Myelopathy Evaluation Questionnaire; CF, cervical spine function; UEF, upper extremity function; LEF, lower extremity function; BF, bladder function; QOL, quality of life.

Table 3. Comparison of Baseline and Radiographic Characteristics Between the Satisfied and Dissatisfied Groups

Characteristic	Satisfied ($n = 29$)	Dissatisfied ($n = 15$)	*P* Value
Average age, years, mean ± SD	62.8 ± 8.5	65.6 ± 7.7	0.29
Male sex, *n* (%)	22 (75.9)	8 (53.3)	0.13
Follow-up, months, mean ± SD	21.5 ± 14.5	28.1 ± 13.5	0.15
Preoperative JOA score, mean ± SD	10.7 ± 2.4	11.4 ± 1.7	0.34
Preoperative kyphosis, *n* (%)	3 (10.3)	2 (13.3)	1.00
Type of ossification, *n* (%)			0.13*
Continuous	10 (34.5)	1 (6.7)	
Segmental	5 (17.2)	3 (20.0)	
Mixed	7 (24.1)	3 (20.0)	
Others	7 (24.1)	8 (53.3)	
Hill-shaped ossification, *n* (%)	5 (17.2)	7 (46.7)	0.04
K-line (−), *n* (%)	8 (27.6)	6 (40.0)	0.41
Severe occupying ratio, *n* (%)	4 (13.8)	3 (20.0)	0.85
Presence of IHSI on MRI, *n* (%)	27 (96.4)	12 (80.0)	0.11*
Postoperative kyphosis, *n* (%)	15 (51.7)	11 (73.3)	0.16
Residual ACS on MRI, *n* (%)	7 (25.9)	7 (46.7)	0.17

JOA, Japanese Orthopedic Association; IHSI, intramedullary high signal intensity; MRI, magnetic resonance imaging; ACS, anterior compression of the spinal cord.
*Fisher's exact test.

Table 4. Comparison of the Effectiveness of Surgical Treatment According to Patient-Reported Outcomes Between the Satisfied and Dissatisfied Groups

Parameters	Satisfied ($n = 29$)	Dissatisfied ($n = 15$)	*P* Value
SF-36 PCS (MCID reached)	18 (81.8)	1 (14.3)	<0.01*
NDI (MCID reached)	13 (50.0)	2 (18.2)	0.14*
Neck pain (MCID reached)	6 (21.4)	3 (23.1)	0.91
Arm pain (MCID reached)	11 (37.9)	6 (46.2)	0.62
JOACMEQ (effectiveness)			
CF	12 (46.2)	2 (22.2)	0.26*
UEF	12 (50.0)	3 (30.0)	0.45*
LEF	16 (61.5)	1 (10.0)	<0.01*
BF	9 (33.3)	0 (0.0)	0.09*
QOL	9 (32.1)	1 (7.7)	0.13

SF-36 PCS, Short Form 36 Physical Component Summary; MCID, minimum clinically important difference; NDI, Neck Disability Index; JOACMEQ: Japanese Orthopedic Association Cervical Myelopathy Evaluation Questionnaire; CF, cervical spine function; UEF, upper extremity function; LEF, lower extremity function; BF, bladder function; QOL, quality of life.
*Fisher's exact test.

group had a higher proportion of patients who considered the surgical treatment effective based on the JOACMEQ BF (33.3% vs. 0.0%), but again the difference also did not reach statistical significance (P = 0.09).

According to the Jonckheere–Terpstra test, there was no statistically significant association between the duration of follow-up and patient satisfaction rating (P = 0.15).

DISCUSSION

Several predictors of poor outcome after laminoplasty in patients with OPLL have been reported.[7-14] Because the concept of laminoplasty through a posterior approach as treatment for patients with OPLL is based on indirect decompression, radiographic parameters related to a larger ossification and cervical kyphotic alignment were considered risk factors for a poor surgical outcome; for instance, physicians could use the K-line as an index to evaluate the cervical alignment and OPLL size at the same time.[11] Fujiyoshi et al[11] reported that patients with a negative K-line demonstrated poorer neurologic improvement compared with those with a positive K-line. Other parameters related to OPLL size and cervical alignment, such as kyphotic cervical alignment, an occupying ratio >60° in the spinal canal, and a hill-shaped ossification also have been identified as predictors of poor surgical outcome of laminoplasty for cervical compressive myelopathy due to OPLL.[6-10] The assessment of the surgical outcome in these reports was based on conventional doctor-reported functional outcomes, such as the JOA score, however.

We used PROs to evaluate HRQOL and pain during the preoperative and postoperative periods. To our knowledge, this is the first report to examine the association between several known prognostic radiographic factors and patient-reported satisfaction after laminoplasty for cervical compressive myelopathy caused by OPLL. Iwasaki et al[13] previously reported that a hill-shaped OPLL was a predictor of a poor outcome following laminoplasty, as evaluated using the JOA recovery rate. The outcome in the present study, showing a smaller proportion of patients with a hill-shaped OPLL in the satisfied group compared with the dissatisfied group, was consistent with that in the previous study. For patients with a hill-shaped OPLL, the indications for laminoplasty should be thoroughly considered, and alternative surgical methods, such as anterior decompression and fusion surgery or posterior decompression surgery with additional fusion, should be considered as well.

We also examined the association between the postoperative radiographic parameters and patient satisfaction. Based on the surgical results of laminoplasty in patients with cervical spondylotic myelopathy, postoperative kyphosis and residual anterior compression of the spinal cord were thought to cause worse outcomes[34,36]; however, few studies have examined the association between postoperative kyphosis and the surgical results of laminoplasty in patients with OPLL.[37,38] In the present study of postoperative satisfaction in patients with OPLL, we found no

association between postoperative kyphosis and dissatisfaction. Although our sample size may be too small to support this conclusion, some previous studies of the surgical results of laminoplasty in patients with OPLL have reported similar findings. For example, Iwasaki et al,[37] in a study of the long-term (>10 years) results of laminoplasty in patients with OPLL, reported that postoperative deterioration of cervical alignment to kyphosis did not cause neurologic deterioration.[37] Moreover, Lee et al[38] recently reported that cervical laminoplasty increased the risk of cervical kyphotic change, but that this postoperative radiographic change was not related to clinical outcomes, including JOA, SF-36, neck pain, or NDI scores.[38]

Although sagittal imbalance has been identified as a major source of pain and disability after cervical fusion surgery,[39,40] the negative impact of postoperative kyphosis on the surgical results of laminoplasty for cervical myelopathy due to OPLL remains controversial, despite a common perception that postoperative kyphosis results in a less indirect decompression effect after laminoplasty. One potential explanation for this is that some patterns of ossification fused to the spinal column might affect both the postoperative alignment change and the surgical outcome. Because we were unable to identify potential reasons to explain the relationship between postoperative kyphotic alignment and surgical outcome in this small study, further large-scale analyses designed to investigate the mechanism of clinical outcomes in patients with OPLL and postoperative kyphosis are warranted.

Several reports on patient satisfaction after laminoplasty in patients with cervical compressive myelopathy have been published. Kimura et al[24] reported that postoperative satisfaction after laminoplasty for cervical compressive myelopathy, including spondylosis and OPLL, was associated with some preoperative domains of the SF-36. Using almost the same inclusion criteria, we also previously analyzed the associations between patient satisfaction and several PROs, including the preoperative and postoperative outcomes, and the effectiveness of surgical treatment using the JOACMEQ.[26] We found that patients reporting effective surgical treatment of lower extremity function were satisfied. In a study focused on patients with OPLL, Fujimori et al[23] demonstrated that the JOA score and several postoperative PROs were associated with patient satisfaction. Their stepwise logistic regression analysis revealed that the SF-36 PF, JOACMEQ QOL, JOACMEQ LEF, and maximum recovery rate were correlated with patient satisfaction. The results of these previous studies, indicating that postoperative physical functions, especially in the lower extremities, were associated with satisfaction in patients with OPLL who underwent laminoplasty, support our present findings. One advantage of the present study is that it provides more practical clinical research on the effectiveness of surgical treatment using the concept of the MCID and the definition of the JOACMEQ.

In this study, we used the MCID to evaluate the effectiveness of surgical treatment with regard to HRQOL and pain. The number of studies evaluating the clinical impact of spinal surgery on HRQOL has been increasing, and several PROs have been used to assess the effects of treatments. The most common metrics used to assess the true clinical impact of surgery are the MCID and the substantial clinical benefit (SCB). The MCID is the smallest change in the score necessary for the patients to be able to discern improvement,[15,16] whereas the SCB is a realistic target value indicating the amount of improvement necessary for the patient to feel that he or she is doing much better.[41] It is important to recognize that a statistically significant difference does not necessarily mean a clinically meaningful difference in HRQOL or a pain scale.[42,43] In this study, although the changes in arm pain and the SF-36 PCS were considered statistically significant, a comparison of the preoperative and postoperative values for the overall population in this study showed that only the SF-36 PCS reached the MCID (and also reached 6.5 points as the SCB).[15]

Studies planned to evaluate the impact of surgery on HRQOL and pain should be focused on expressing a clinically important difference resulting from the surgical intervention, not a statistically significant difference. Our present results suggest that patients with OPLL undergoing laminoplasty could perceive a difference in clinical physical function, but could not recognize any difference in arm pain, neck pain, or neck disability status. Moreover, the satisfied group had a higher proportion of patients reaching the MCID on the SF-36 PCS compared with the dissatisfied group, suggesting that patients with OPLL who were able to recognize a difference in clinical physical function after laminoplasty were satisfied with the surgery.

The JOACMEQ, a patient-based method for evaluating cervical compressive myelopathy, consists of 24 questions.[36] Five functional domains, including cervical spine function, upper extremity function, lower extremity function, bladder function, and QOL, are calculated separately according to the formulas provided. Each functional score ranges from 0 to 100, with higher scores indicating a better condition. Therapeutic effectiveness also can be evaluated using this scoring system. The advantage of using this patient-reported scoring system is that physicians can evaluate each function, such as upper extremity, lower extremity, and bladder functions, based on the patient-reported function and HRQOL. Several reports on cervical compressive myelopathy using this scoring system have been published to date.[23,26,44-49] Further research examining the association between the definition of therapeutic effectiveness using this scoring system and the change in values between the preoperative and postoperative periods using the concept of the MCID will be of interest.

Postoperative complications or the need for further surgery may affect patient satisfaction. Although a previous study examining PROs after spinal column osteotomy reported an influence on PROs in requiring a major reoperation,[50] the influence of complications and further surgery on postoperative satisfaction remains unclear. In this study, presumably owing to the relatively short duration of follow-up, no patients required further surgery because of OPLL progression. In addition, only 1 patient underwent further surgery for collapse of the cervical lamina, at 7 months after the index surgery. Because of the small number of adverse events, we were unable to include the cases with postoperative complications or further surgery as potential predictors in our analysis. If the number of patients with postoperative complications increases as the follow-up period progresses, then further studies examining the association between patient satisfaction and postoperative complications may be conducted.

This study has several limitations, starting with the relatively small number of cases. Because the small number of patients did not allow for a multivariate analysis with several variables, we were unable to avoid the impact of important confounding factors among each variable on our results. Thus, our findings, without adjusting for confounding factors, could be merely hypotheses driven by analyses using the data obtained from a small number of patients. Consequently, according to the potential factors identified in this hypothesis-generating study, we plan to conduct further large-scale, multicenter studies in patients with OPLL. Second, there may have been both selection bias and recall bias. Patients with known risk factors for a poor surgical outcome, such as kyphosis or a negative K-line, might have been less likely to be included in this study, because the surgeons may have deemed them contraindicated for laminoplasty. In addition, clinical outcome studies based on PROs have the potential to be affected by recall bias.

Finally, the timing of the follow-up when patient satisfaction and PROs were evaluated varied, ranging from 12 to 89 months. OPLL progression after posterior decompression surgery has been reported to occur with an incidence of 56.5% at 2 years and 71% at 5 years,[51] although we could not evaluate this complication in the present study owing to the lack of postoperative CT images in the same operative period among this cohort. This late-onset postoperative feature in patients with OPLL may have led to a poorer outcome in patients with a longer follow-up. Despite the foregoing limitations, however, we believe that this study provides valuable information that is of clinical importance.

In conclusion, patient satisfaction after laminoplasty for cervical OPLL was insufficient in patients with a hill-shaped ossification. The indications for laminoplasty in patients with a hill-shaped OPLL must be carefully considered, and the procedure should be avoided when appropriate. Moreover, patient satisfaction was related to postoperative clinical improvement in physical function, especially lower extremity function. We believe that our findings provide useful information that can help improve our understanding of the selection of treatment for cervical compressive myelopathy caused by OPLL.

REFERENCES

1. Fujimori T, Le H, Hu SS, Chin C, Pekmezci M, Schairer W, et al. Ossification of the posterior longitudinal ligament of the cervical spine in 3161 patients: a CT-based study. *Spine (Phila Pa 1976)*. 2015;40:E394-E403.

2. Kalb S, Martirosyan NL, Perez-Orribo L, Kalani MY, Theodore N. Analysis of demographics, risk factors, clinical presentation, and surgical treatment modalities for the ossified posterior longitudinal ligament. *Neurosurg Focus*. 2011;30:E11.

3. Cardoso MJ, Koski TR, Ganju A, Liu JC. Approach-related complications after decompression for cervical ossification of the posterior longitudinal ligament. *Neurosurg Focus*. 2011;30:E12.

4. Kimura A, Seichi A, Hoshino Y, Yamazaki M, Mochizuki M, Aiba A, et al. Perioperative complications of anterior cervical decompression with fusion in patients with ossification of the posterior longitudinal ligament: a retrospective, multi-institutional study. *J Orthop Sci*. 2012;17:667-672.

5. Sakai K, Okawa A, Takahashi M, Arai Y, Kawabata S, Enomoto M, et al. Five-year follow-up evaluation of surgical treatment for cervical myelopathy caused by ossification of the posterior longitudinal ligament: a prospective comparative study of anterior decompression and fusion with floating method versus laminoplasty. *Spine (Phila Pa 1976)*. 2012;37:367-376.

6. Liu H, Li Y, Chen Y, Wu W, Zou D. Cervical curvature, spinal cord MRIT2 signal, and occupying ratio impact surgical approach selection in patients with ossification of the posterior longitudinal ligament. *Eur Spine J*. 2013;22:1480-1488.

7. Matsumoto M, Chiba K, Toyama Y. Surgical treatment of ossification of the posterior longitudinal ligament and its outcomes: posterior surgery by laminoplasty. *Spine (Phila Pa 1976)*. 2012;37:E303-E308.

8. Fujimori T, Iwasaki M, Okuda S, Takenaka S, Kashii M, Kaito T, et al. Long-term results of cervical myelopathy due to ossification of the posterior longitudinal ligament with an occupying ratio of 60% or more. *Spine (Phila Pa 1976)*. 2014;39:58-67.

9. Iwasaki M, Okuda S, Miyauchi A, Sakaura H, Mukai Y, Yonenobu K, et al. Surgical strategy for cervical myelopathy due to ossification of the posterior longitudinal ligament, part 1: clinical results and limitations of laminoplasty. *Spine (PhilaPa 1976)*. 2007;32:647-653.

10. Kim B, Yoon DH, Shin HC, Kim KN, Yi S, Shin DA, et al. Surgical outcome and prognostic factors of anterior decompression and fusion for cervical compressive myelopathy due to ossification of the posterior longitudinal ligament. *Spine J*. 2015;15:875-884.

11. Fujiyoshi T, Yamazaki M, Kawabe J, Endo T, Furuya T, Koda M, et al. A new concept for making decisions regarding the surgical approach for cervical ossification of the posterior longitudinal ligament: the K-line. *Spine (Phila Pa 1976)*. 2008;33:E990-E993.

12. Li H, Jiang LS, Dai LY. A review of prognostic factors for surgical outcome of ossification of the posterior longitudinal ligament of cervical spine. *Eur Spine J*. 2008;17:1277-1288.

13. Maruo K, Moriyama T, Tachibana T, Inoue S, Arizumi F, Daimon T, et al. The impact of dynamic factors on surgical outcomes after double-door laminoplasty for ossification of the posterior longitudinal ligament of the cervical spine. *J Neurosurg Spine*. 2014;21:938-943.

14. Masaki Y, Yamazaki M, Okawa A, Aramomi M, Hashimoto M, Koda M, et al. An analysis of factors causing poor surgical outcome in patients with cervical myelopathy due to ossification of the posterior longitudinal ligament: anterior decompression with spinal fusion versus laminoplasty. *J Spinal Disord Tech*. 2007;20:7-13.

15. Carreon LY, Glassman SD, Campbell MJ, Anderson PA. Neck Disability Index, Short Form-36 Physical Component Summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. *Spine J*. 2010;10:469-474.

16. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J*. 2008;8:968-974.

17. Parker SL, Adogwa O, Mendenhall SK, Shau DN, Ankjlderson WN, Cheng JS, et al. Determination of minimum clinically important difference (MCID) in pain, disability, and quality of life after revision fusion for symptomatic pseudoarthrosis. *Spine J*. 2012;12:1122-1128.

18. Parker SL, Adogwa O, Paul AR, Anderson WN, Aaronson O, Cheng JS, et al. Utility of minimum clinically important difference in assessing pain, disability, and health state after transforaminal lumbar interbody fusion for degenerative lumbar spondylolisthesis. *J Neurosurg Spine*. 2011;14:598-604.

19. Parker SL, Godil SS, Shau DN, Mendenhall SK, McGirt MJ. Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion: clinical article. *J Neurosurg Spine*. 2013;18:154-160.

20. Parker SL, Mendenhall SK, Shau D, Adogwa O, Cheng JS, Anderson WN, et al. Determination of minimum clinically important difference in pain, disability, and quality of life after extension of fusion for adjacent-segment disease. *J Neurosurg Spine*. 2012;16:61-67.

21. Parker SL, Mendenhall SK, Shau DN, Adogwa O, Anderson WN, Devin CJ, et al. Minimum clinically important difference in pain, disability, and quality of life after neural decompression and

fusion for same-level recurrent lumbar stenosis: understanding clinical versus statistical significance. *J Neurosurg Spine*. 2012;16:471-478.

22. Theodore BR. Methodological problems associated with the present conceptualization of the minimum clinically important difference and substantial clinical benefit. *Spine J*. 2010;10: 507-509.

23. Fujimori T, Iwasaki M, Okuda S, Nagamoto Y, Sakaura H, Oda T, et al. Patient satisfaction with surgery for cervical myelopathy due to ossification of the posterior longitudinal ligament. *J Neurosurg Spine*. 2011;14:726-733.

24. Kimura A, Endo T, Inoue H, Seichi A. Preoperative predictors of patient satisfaction with outcome after cervical laminoplasty. *Global Spine J*. 2014;4:77-82.

25. Neo M, Fujibayashi S, Takemoto M, Nakamura T. Clinical results of and patient satisfaction with cervical laminoplasty for considerable cord compression with only slight myelopathy. *Eur Spine J*. 2012;21:340-346.

26. Ohya J, Oshima Y, Takeshita K, Oka H, Chikuda H, Taniguchi Y, et al. Patient satisfaction with double-door laminoplasty for cervical compression myelopathy. *J Orthop Sci*. 2015;20: 64-70.

27. Seichi A, Takeshita K, Ohishi I, Kawaguchi H, Akune T, Anamizu Y, et al. Long-term results of double-door laminoplasty for cervical stenotic myelopathy. *Spine (Phila Pa 1976)*. 2001;26:479-487.

28. Iwasaki M, Yonenobu K. Ossification of the posterior longitudinal ligament. In: Herkowitz HN, Garfin SR, Eismont FJ, Bell GR, Balderston RA, eds. *Rothman-Simeone The Spine*. 6th ed. Philadelphia, PA: Saunders; 2006:791-807.

29. Saetia K, Cho D, Lee S, Kim DH, Kim SD. Ossification of the posterior longitudinal ligament: a review. *Neurosurg Focus*. 2011;30:E1.

30. Tsuyama N. Ossification of the posterior longitudinal ligament of the spine. *Clin Orthop Relat Res*. 1984;184:71-84.

31. Hirai T, Kawabata S, Enomoto M, Kato T, Tomizawa S, Sakai K, et al. Presence of anterior compression of the spinal cord after laminoplasty inhibits upper extremity motor recovery in patients with cervical spondylotic myelopathy. *Spine (Phila Pa 1976)*. 2012;37:377-384.

32. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther*. 1991;14:409-415.

33. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30: 473-483.

34. Fukui M, Chiba K, Kawakami M, Kikuchi S, Konno S, Miyamoto M, et al. JOA Back Pain Evaluation Questionnaire (JOABPEQ)/JOA Cervical Myelopathy Evaluation Questionnaire (JOACMEQ). The report on the development of revised versions. April 16, 2007. The Subcommittee of the Clinical Outcome Committee of the Japanese Orthopaedic Association on Low Back Pain and Cervical Myelopathy Evaluation. *J Orthop Sci*. 2009; 14:348-365.

35. Yonenobu K, Abumi K, Nagata K, Taketomi E, Ueyama K. Interobserver and intraobserver reliability of the Japanese Orthopaedic Association scoring system for evaluation of cervical compression myelopathy. *Spine (Phila Pa 1976)*. 2001;26:1890-1894 [discussion: 1895].

36. Kim TH, Lee SY, Kim YC, Park MS, Kim SW. T1 slope as a predictor of kyphotic alignment change after laminoplasty in patients with cervical myelopathy. *Spine (Phila Pa 1976)*. 2013;38: E992-E997.

37. Iwasaki M, Kawaguchi Y, Kimura T, Yonenobu K. Long-term results of expansive laminoplasty for ossification of the posterior longitudinal ligament of the cervical spine: more than 10 years follow up. *J Neurosurg*. 2002;96(2 Suppl):180-189.

38. Lee CK, Shin DA, Yi S, Kim KN, Shin HC, Yoon do H, et al. Correlation between cervical spine sagittal alignment and clinical outcome after cervical laminoplasty for ossification of the posterior longitudinal ligament. *J Neurosurg Spine*. 2016;24: 100-107.

39. Tang JA, Scheer JK, Smith JS, Deviren V, Bess S, Hart RA, et al. The impact of standing regional cervical sagittal alignment on outcomes in posterior cervical fusion surgery. *Neurosurgery*. 2012;71: 662-669 [discussion: 669].

40. Tang JA, Scheer JK, Smith JS, Deviren V, Bess S, Hart RA, et al. The impact of standing regional cervical sagittal alignment on outcomes in posterior cervical fusion surgery. *Neurosurgery*. 2015; 76(Suppl 1):S14-S21 [discussion: S21].

41. Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY. Defining substantial clinical benefit following lumbar spine arthrodesis. *J Bone Joint Surg Am*. 2008;90:1839-1847.

42. Mattei TA. "Statistically significant" does not necessarily mean "clinically different" on pain/quality of life scales: opportune remarks on clinical outcomes measures in cervical spondylotic myelopathy [letter]. *Neurosurgery*. 2012;71: E518-E521 [author reply: E521-E522].

43. Mattei TA, Fassett DR. Minimum clinically important difference [letter]. *J Neurosurg Spine*. 2011;15:690-691 [author reply: 691].

44. Matsumoto M, Watanabe K, Hosogane N, Tsuji T, Ishii K, Nakamura M, et al. Impact of lamina closure on long-term outcomes of open-door laminoplasty in patients with cervical myelopathy: minimum 5-year follow-up study. *Spine (Phila Pa 1976)*. 2012;37:1288-1291.

45. Nakashima H, Yukawa Y, Ito K, Machino M, Kanbara S, Morita D, et al. Validity of the 10-s step test: prospective study comparing it with the 10-s grip and release test and the 30-m walking test. *Eur Spine J*. 2011;20:1318-1322.

46. Nakashima H, Yukawa Y, Ito K, Machino M, Kanbara S, Morita D, et al. Prediction of lower limb functional recovery after laminoplasty for cervical myelopathy: focusing on the 10-s step test. *Eur Spine J*. 2012;21:1389-1395.

47. Nikaido T, Kikuchi S, Yabuki S, Otani K, Konno S. Surgical treatment assessment using the Japanese orthopedic association cervical myelopathy evaluation questionnaire in patients with cervical myelopathy: a new outcome measure for cervical myelopathy. *Spine (Phila Pa 1976)*. 2009;34: 2568-2572.

48. Oshima K, Iwasaki M, Sakaura H, Fujimori T, Nagamoto Y, Yoshikawa H. Comparison of the Japanese Orthopaedic Association score and the Japanese Orthopaedic Association cervical myelopathy evaluation questionnaire scores: time-dependent changes in patients with cervical spondylotic myelopathy and posterior longitudinal ligament. *Asian Spine J*. 2015;9:47-53.

49. Yoon ST, Raich A, Hashimoto RE, Riew KD, Shaffrey CI, Rhee JM, et al. Predictive factors affecting outcome after cervical laminoplasty. *Spine (Phila Pa 1976)*. 2013;38(22 Suppl 1): S232-S252.

50. O'Neill KR, Lenke LG, Bridwell KH, Neuman BJ, Kim HJ, Archer KR. Factors associated with long-term patient-reported outcomes after three-column osteotomies. *Spine J*. 2015;15:2312-2318.

51. Chiba K, Yamamoto I, Hirabayashi H, Iwasaki M, Goto H, Yonenobu K, et al. Multicenter study investigating the postoperative progression of ossification of the posterior longitudinal ligament in the cervical spine: a new computer-assisted measurement. *J Neurosurg Spine*. 2005;3:17-23.

Conflict of interest statement: This study was funded by grants from the Investigation Committee on the Ossification of the Spinal Ligaments of the Japanese Ministry of Health, Labor and Welfare, which played no role in the design, data collection, analysis, decision to publish, or preparation of the manuscript.

Received 21 June 2016; accepted 6 September 2016

Citation: World Neurosurg. (2016) 96:272-279.
http://dx.doi.org/10.1016/j.wneu.2016.09.011

Journal homepage: ***www.WORLDNEUROSURGERY.org***

Available online: ***www.sciencedirect.com***

Spine

SPINE Volume 41, Number 21, pp E1265–E1270

CERVICAL SPINE

Effect of Preoperative Sagittal Balance on Cervical Laminoplasty Outcomes

Yasushi Oshima, MD, PhD,* Katsushi Takeshita, MD, PhD,† Yuki Taniguchi, MD, PhD,* Yoshitaka Matsubayashi, MD,* Toru Doi, MD,* Junichi Ohya, MD,* Kazuhito Soma, MD,* So Kato, MD,* Hiroyuki Oka, MD,‡ Hirotaka Chikuda, MD, PhD,* and Sakae Tanaka, MD, PhD*

Study Design. Retrospective case series.

Objective. To clarify how preoperative global sagittal imbalance influences outcomes in patients with cervical compression myelopathy undergoing cervical laminoplasty.

Summary of Background Data. The influence of sagittal balance on outcomes of cervical laminoplasty remains uncertain.

Methods. The authors retrospectively reviewed data of 106 patients who underwent double-door cervical laminoplasty between 2004 and 2011 and investigated the influence of the C7 sagittal vertical axis (SVA) on outcome scores. Primary outcomes used were Japanese Orthopedic Association (JOA) scores, Numerical Rating Scale for neck or arm pain, the Short Form 36 Health Survey (physical and mental component summary scores), and the Neck Disability Index (NDI).

Results. Ninety-two patients with complete data were eligible for inclusion. The preoperative C7 SVA was $\leq$5 cm in 64 patients (69.6%) and $>$ 5 cm in 28 (30.4%). We compared each parameter by the magnitude of spinal sagittal balance (preoperative C7 SVA $>$ 5 cm *vs.* C7 SVA $\leq$ 5 cm) after adjusting for age via the least square mean analysis because the average age was significantly higher in patients with C7 SVA $>$ 5 cm. As for the radiographic parameters, both C2–7 SVA and C7 SVA were larger in patients when the C7 SVA was $>$ 5 cm. Numerical Rating Scale for postoperative arm pain, postoperative JOA scores, and both pre- and postoperative physical component summary and NDI were worse in patients with C7 SVA $>$ 5 cm; however, the JOA score recovery rate and changes in physical component summary and NDI were not significantly different.

Conclusion. Postoperative functional outcome scores were significantly lower in patients with C7 SVA $>$ 5 cm, although the improvement after cervical laminoplasty was not greatly affected. The involvement of global sagittal balance and cervical regional alignment should be considered in evaluating surgical outcomes for patients undergoing cervical laminoplasty.

Key words: alignment, cervical myelopathy, global balance, laminoplasty, outcome, sagittal balance, sagittal vertical axis.

Level of Evidence: 4

Spine 2016;41:E1265–E1270

Cervical laminoplasty is an established procedure for the treatment of cervical compression myelopathy, and relatively good long-term surgical results have been reported.[1,2] One of the advantages of this technique is that it can be applied to multisegmental cord compression cases without fusing the vertebral bodies; this is particularly suitable for the elderly who often have multiple segments involved.[3–5] The possible postoperative complications related to this procedure include posterior axial neck and scapular pain,[6,7] decreased range of motion (ROM),[8,9] and postoperative kyphosis.[10,11]

The known causes of poor surgical outcomes after cervical laminoplasty include cervical regional kyphosis [12] and large anterior factors such as ossification of the longitudinal ligament (OPLL),[13] which appears logical given that the concept of this procedure is a posterior shift of the spinal cord.[14] Increased age at surgery is also reported to affect postoperative neurological recovery.[5,15,16] However, only few reports have investigated the influence of global sagittal spinal balance on surgical outcomes. Because sagittal spinal imbalance can influence health-related quality of life (HRQOL),[17,18] it is reasonable to speculate that not only general outcome scores utilized for the evaluation of

From the *Department of Orthopedic Surgery, University of Tokyo, Tokyo, Japan; †Department of Orthopedics, Jichi Medical University; and ‡Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo.

Acknowledgment date: December 14, 2015. First revision date: February 11, 2016. Acceptance date: March 16, 2016.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work.

Relevant financial activities outside the submitted work: expert testimony, grants, and payment for development of educational presentations.

Address correspondence and reprint requests to Yasushi Oshima, MD, PhD, Department of Orthopedic Surgery, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; E-mail: yoo-tky@umin.ac.jp

DOI: 10.1097/BRS.0000000000001615

HRQOL such as the Short-Form 36 (SF-36) Health Survey but also disease-specific outcome scores such as the Japanese Orthopedic Association (JOA) score and Neck Disability Index (NDI) can be influenced by the existence of sagittal imbalance itself, regardless of the degree of myelopathy. Furthermore, patients with global sagittal spinal imbalance may be urged to keep their cervical spine in a hyperlordotic alignment to maintain an upright position of the head, which may not be necessary for those with good sagittal spinal balance.[19,20] As cervical alignment tends to change from lordosis to kyphosis after cervical laminoplasty, such patients may have more difficulty in maintaining their HRQOL with a slight kyphotic change of postoperative cervical alignment. In other words, HRQOL in patients with sagittal spinal imbalance may be influenced more by postoperative regional kyphotic change of the cervical spine after cervical laminoplasty. Therefore, it is necessary to consider the influence of preoperative global sagittal imbalance on the impairment of HRQOL before surgery, even if the cervical region does not appear to be impaired.

The purpose of this study was to clarify how preoperative global sagittal imbalance influences outcome scores in patients with cervical myelopathy undergoing cervical laminoplasty.

MATERIALS AND METHODS

The study protocols were approved by the institutional review board of the authors' institution. We retrospectively reviewed 106 patients with cervical compression myelopathy who underwent double-door laminoplasty between 2004 and 2011 and replied to the patient-reported outcomes questionnaires both pre- and postoperatively at our institution. Patients with rheumatoid arthritis, disc herniation, tumor, trauma, severe lumbar spinal canal stenosis, or previous surgery were not included. Fourteen patients were excluded because they did not undergo radiographic examinations of the whole spine.

Radiological parameters included the measurement of Cobb angles between the C2 and C7 vertebrae with cervical lateral radiographs and the C7 slope and C7 sagittal vertical axis (C7 SVA) with whole-spine standing lateral radiographs. C7 SVA was the distance between the C7 plumb line and posterior corner of the sacrum. The materials were sub-analyzed by the magnitude of sagittal imbalance (preoperative C7 SVA >5 cm vs C7 SVA ≤ 5 cm). The degree of cervical spinal cord compression was evaluated using mid-sagittal T2-weighted magnetic resonance imaging (MRI), as previously described, by comparing the sagittal diameter of the spinal cord at the maximum compression level with that of C1 and C7 [maximum spinal cord compression (MSCC)].[21] A higher MSCC means that the patient has more severe cervical spinal cord compression.

Primary outcome measures used were JOA scores, Numerical Rating Scale (NRS) for each part of the body, the SF-36 Health Survey [physical and mental component summary scores (PCS and MCS, respectively)], and NDI.

SPSS v18 software (SPSS, Chicago, IL) was used for the Wilcoxon signed rank test, Mann–Whitney *U* test, and Spearman rank correlation coefficient. Least square mean analysis was used to compare each variable by the magnitude of sagittal imbalance (C7 SVA >5 cm *vs.* C7 SVA ≤ 5 cm) after adjusting for age using the SAS procedure PROC GLM (SAS Institute, Inc., Cary, NC). A P value of <0.05 was considered statistically significant.

RESULTS

Ninety-two patients with complete data were eligible for inclusion. There were 61 males and 31 females with a mean age of 64 years (range, 34–82 yrs), and the mean follow-up period was 27 months (range, 12–60 months) (Table 1). The pre- and postoperative radiographic parameters were not significantly different, except for cervical ROM (Table 2).

The mean preoperative C7 SVA was 3.4 cm (−4.0–20.0 cm), which had a positive correlation with age ($P=0.36$) (Figure 1). The preoperative C7 SVA was ≤ 5 cm in 64 patients (69.6%) and >5 cm in 28 (30.4%). C7 SVA worsened and increased to >5 cm in four of the 64 patients (6.3%) with preoperative C7 SVA ≤ 5 cm, whereas C7 SVA improved and reduced to <5 cm in two the 27 patients (7.4%) with preoperative C7 SVA >5 cm. However, the average pre- and postoperative C7 SVA measurements were not significantly different (Table 2) and were strongly correlated ($P=0.86$).

Then, we compared demographic data and each parameter between the two groups by the magnitude of spinal sagittal balance (preoperative C7 SVA >5 cm *vs.* C7 SVA ≤ 5 cm; Table 3). As expected, the average age was significantly different between the two groups. Therefore, further comparison was performed after adjusting for age *via* least square mean analysis. As for the radiographic and MRI parameters, there were no statistical differences between the two groups, except for C2–7 SVA and lumbar lordosis (Table 4). NRS scores for postoperative arm pain were significantly higher in patients with C7 SVA >5 cm. Similarly, the JOA showed worse scores only postoperatively in patients with C7 SVA >5 cm ($P=0.01$); however, the preoperative JOA score tended to be lower in those with C7 SVA >5 cm ($P=0.06$) and the recovery rates of the JOA scores were not significantly different between the groups. Both pre- and postoperative PCS and NDI were inferior for

TABLE 1. Demographic Data of Patients (N = 92)

Age (yrs)	**64 (range, 34–82)**
Sex (Males/Females)	61/31
Follow-up period (months)	27 (range, 12–60)
CSM/OPLL	54/38

CSM indicates cervical spondylotic myelopathy; OPLL, ossification of posterior longitudinal ligament.

TABLE 2. Pre- and Postoperative Radiographic Measurements

	Preoperative		Postoperative		
	Average	SD	Average	SD	*P*
C2C7 Cobb (degrees)	9.0	9.5	8.8	12.0	0.83
ROM (degrees)	40.6	15.5	25.4	12.7	0.00
C7 SVA (cm)	3.4	4.5	3.6	4.5	0.23
C7 slope (degrees)	26.0	8.8	25.6	9.4	0.69
C2–7 SVA (cm)	2.5	3.3	2.2	1.5	0.42

ROM indicates range of motion; SD, standard deviation; SVA, sagittal vertical axis.

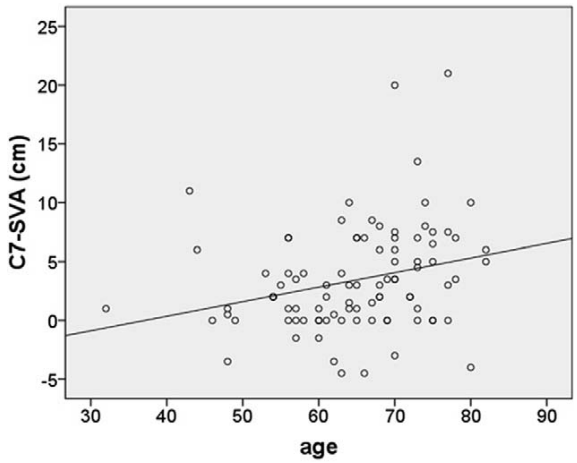


Figure 1. Correlation between age and C7 SVA. C7 SVA was positively correlated with age ($P = 0.36$).

patients with C7 SVA > 5 cm. Changes in PCS and NDI were not significantly different.

DISCUSSION

We sought to clarify the influence of global sagittal imbalance on outcome scores in patients with cervical myelopathy who underwent laminoplasty and found that each of our outcomes measurements showed worse scores in patients with C7 SVA > 5 cm. NRS for postoperative arm pain, postoperative JOA scores, and both pre- and postoperative PCS and NDI were worse in patients with C7 SVA > 5 cm; however, the JOA score recovery rate and changes in PCS and NDI were not significantly different. Considering the above results, we speculate that postoperative outcome scores tend to be lower in patients with global sagittal imbalance, although the degree of neurological improvement after cervical laminoplasty may not be greatly affected by the presence of spinal sagittal imbalance. To the best of our knowledge, this is the first study to report an influence of global sagittal balance on HRQOL outcomes in patients undergoing cervical laminoplasty.

There are two possibilities for the worse outcome scores in patients with C7 SVA > 5 cm. The first possibility is the influence of coexisting global sagittal imbalance. It is reasonable to speculate that patients with global sagittal imbalance have worse general HRQOL outcomes such as SF-36 PCS. Many authors have reported the correlation between sagittal imbalance, defined as a C7 plum line of ≥5 cm anterior to the posterior superior sacral margin, and inferior functional outcomes in patients with thoracolumbar disorders.[18,20,22,23] In addition, such patients exhibit walking disability as a result of the sagittal imbalance itself, which can affect the lower extremity motor scores on the JOA. NDI is also influenced by sagittal imbalance because those patients may have difficulty gazing, which affects regional alignment and neck functions. The fact that both pre- and postoperative outcome scores were inferior in patients with C7 SVA > 5 cm indicates possible involvement of sagittal imbalance itself.

One important factor that can affect global sagittal balance is age, which is reported to influence surgical outcomes after cervical laminoplasty.[3–5] In our study, C7 SVA correlated with age. However, we showed that patients with C7 SVA > 5 cm had worse pre- and postoperative functional

TABLE 3. Demographic Data of Patients According to Preoperative C7 SVA

	C7 SVA ≤ 5 cm	C7 SVA > 5 cm	
	N = 65	N = 27	*P*
Age (yrs)	63.4	68.2	0.03
Sex (Males/Feamles)	44/21	17/10	0.81
Follow-up period (months)	27	24	0.38
CSM / OPLL	38/27	16/11	0.94

CSM indicates cervical spondylotic myelopathy; OPLL, ossification of posterior longitudinal ligament; SVA, sagittal vertical axis.

TABLE 4. Outcome Scores and Imaging Parameters According to Preoperative C7SVA After Adjusting for Age

		C7 SVA ≤ 5 cm		C7 SVA > 5 cm		
		N = 65		N = 27		
		Average	SE	Average	SE	*P*
Outcome values						
NRS						
Neck pain	preoperative	3.3	0.4	3.8	0.7	0.57
	postoperative	2.7	0.3	3.5	0.5	0.19
Arm pain	preoperative	4.1	0.4	4.2	0.6	0.91
	postoperative	2.4	0.3	3.8	0.5	0.02
JOA score	preoperative	10.7	0.3	9.6	0.5	0.06
	postoperative	13.9	0.2	12.6	0.4	0.01
	recorery rate	47.8	3.8	43.7	6.5	0.58
NDI	preoperative	32.6	2.3	43.0	3.7	0.02
	postoperative	25.7	1.8	33.5	3.0	0.03
	change	6.9	2.3	7.5	3.7	0.58
PCS	preoperative	25.1	2.3	12.2	3.7	0.00
	postoperative	36.8	2.0	19.8	3.4	<0.001
	change	11.7	2.4	7.6	3.9	0.29
MCS	preoperative	48.2	1.6	49.9	2.5	0.57
	postoperative	49.4	1.3	51.9	2.1	0.32
	change	1.2	1.7	2.0	2.7	0.79
Imaging parameters						
C2C7 Cobb	preoperative	8.8	1.2	9.0	1.9	0.90
(degrees)	postoperative	9.0	1.5	7.2	2.5	0.52
ROM	preoperative	41.9	2.0	34.9	3.5	0.09
(degrees)	postoperative	26.1	1.6	22.0	2.8	0.20
C7 SVA	preoperative	1.3	0.3	8.6	0.6	<0.001
(cm)	postoperative	1.6	0.4	8.6	0.6	<0.001
C7 slope	preoperative	25.3	1.1	28.4	1.8	0.16
(degrees)	postoperative	24.6	1.2	28.7	1.9	0.07
C2–7 SVA	preoperative	2.1	0.4	4.1	0.6	0.01
(cm)	postoperative	2.0	0.2	3.2	0.3	0.01
Thoracic kyphosis	preoperative	35.7	1.4	35.3	2.3	0.90
(degrees)	postoperative	37.4	1.5	35.2	2.6	0.47
Lumbar ordosis	preoperative	45.7	1.3	33.6	2.2	<0.001
(degrees)	postoperative	44.9	1.4	32.3	2.3	<0.001
Sacral slope	preoperative	32.0	0.8	29.5	1.8	0.11
(degrees)	postoperative	32.2	0.8	29.1	1.4	0.06
Pelvic tilt	preoperative	14.7	0.9	17.1	1.5	0.17
(degrees)	postoperative	14.8	1.0	18.3	1.7	0.09
Pelvic incidence	preoperative	46.7	1.3	46.6	2.0	0.96
(degrees)	postoperative	47.0	1.2	47.3	2.1	0.89
MSCC on MRI	preoperative	45.2	2.1	47.9	3.5	0.51

JOA indicates Japanese Orthopedic Association; MCS, mental component summary; MRI, magnetic resonance imaging; MSCC, maximum spinal cord compression; NDI, Neck Disability Index; NRS, Numerical Rating Scale; PCS, physical component summary; ROM, range of motion; SE, standard error; SVA, sagittal vertical axis.

outcomes, even after adjusting for age, which indicates that the relatively low functional outcome scores in patients with C7 SVA > 5 cm mainly resulted from the presence of sagittal imbalance itself and not from age. The relatively poorer surgical outcomes reported in the past for the elderly who have undergone cervical laminoplasty may be partly because of the influence of the global sagittal imbalance of such patients. Nevertheless, we should consider the influence of sagittal balance when treating the elderly with cervical myelopathy.

Another possible reason for the low functional outcome scores in patients with global sagittal imbalance is the influence of regional alignment, particularly of the C2–7 SVA. In this study, the subjects with C7 SVA > 5 cm also had regional sagittal imbalance in their neck. There have been several reports about the influence of C2–7 SVA on cervical myelopathy or surgical results. Tang *et al*[24] reported a relationship between postoperative C2–7 SVA and postoperative NDI and PCS of SF-36 in patients undergoing cervical fusion surgery. Smith *et al*[25] reported that C2–7 SVA correlated with myelopathy severity, as measured by modified JOA scores. Roguski *et al*[26] also investigated the correlation between C2–7 SVA and HRQOL outcome scores in 21 patients with anterior surgery and in 28 patients with posterior surgery for cervical myelopathy. They showed statistical correlation between pre- and postoperative C2–7 SVA and improvement in the PCS; the majority of patients with C2–7 SVA > 40 mm did not improve after surgical intervention. In our cases, C2–7 SVA was significantly larger in patients with C7 SVA > 5 cm, which was surprising to a certain extent because we expected that cervical hyperlordosis would have compensated for the malposition of the head in patients with global sagittal imbalance. It is unclear why such patients did not show hyperlordosis of the neck to keep their head position; however, we speculate that muscle atrophy resulting from cervical myelopathy may have made the compensation impossible. Here, we classified patients into two groups according to the magnitude of C7 SVA because the basic concept of this study was to investigate influence of global sagittal balance on outcome scores. Therefore, multivariable analysis would be necessary to evaluate the true impact of global sagittal balance and regional cervical alignment on cervical laminoplasty outcomes by setting one outcome as a dependent variable. Nevertheless, we believe that both global sagittal imbalance and regional sagittal imbalance affected the functional outcomes scores in our case. Cervical fusion surgery with correction of cervical alignment instead of cervical laminoplasty may lead to better surgical outcomes in such patients, although the patients in this study did show improvements after cervical laminoplasty to some extent.

Postoperative kyphosis of the cervical spine has been reported as a possible adverse event after cervical laminoplasty.[10,11] It is reasonable to speculate that patients with global sagittal imbalance would have more difficulty gazing if such regional kyphosis occurs after laminoplasty. Fortunately, we did not note any severe kyphotic changes of the cervical spine in our cases. Techniques such as laminoplasty or fusion surgery can preserve posterior structure and should be considered in patients with severe global sagittal imbalance.

There were several limitations in this study. First, this is not a comparative study and we could not compare laminoplasty with other procedures such as anterior or posterior fusion surgery. Second, although we classified patients according to the preoperative C7 SVA, some patients showed improvement or deterioration in their sagittal balance postoperatively. However, because the number of such patients was not so large and the pre- and postoperative C7 SVA was well correlated, we believe that the results were not greatly affected by the postoperative changes in sagittal balance. Third, we did not include cases with severe sagittal imbalance. Treatment for the deformity should be paramount when treating such patients. Finally, the small number of patients made multivariable analysis of the imaging parameters impossible. A prospective study with a larger number of patients will elucidate these problems.

In conclusion, in patients undergoing cervical laminoplasty for cervical myelopathy, postoperative functional outcome scores appear to be lower in those with C7 SVA > 5 cm, although the improvement after cervical laminoplasty is not greatly affected. The involvement of global sagittal balance and cervical regional alignment should be considered when evaluating surgical outcomes for cervical laminoplasty, particularly for the elderly.

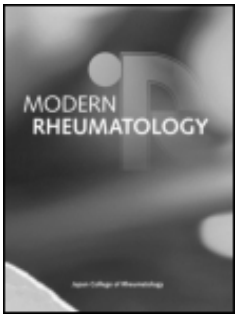
➢ Key Points

- ❑ We investigated the influence of global sagittal balance on outcomes of 92 patients undergoing cervical laminoplasty.
- ❑ Postoperative functional outcome scores were significantly lower in patients with C7 SVA > 5 cm, although the improvement after cervical laminoplasty was not greatly affected.
- ❑ The involvement of global sagittal balance and cervical regional alignment should be considered in evaluating surgical outcomes for patients undergoing cervical laminoplasty.

References

1. Chiba K, Ogawa Y, Ishii K, et al. Long-term results of expansive open-door laminoplasty for cervical myelopathy: average 14-year follow-up study. *Spine* 2006;31:2998–3005.
2. Seichi A, Takeshita K, Ohishi I, et al. Long-term results of double-door laminoplasty for cervical stenotic myelopathy. *Spine* 2001;26:479–87.
3. Machino M, Yukawa Y, Imagama S, et al. Age-related and degenerative changes in the osseous anatomy, alignment, and range of motion of the cervical spine: a comparative study of radiographic data from 1016 patients with cervical spondylotic myelopathy and 1230 asymptomatic subjects. *Spine (Phila Pa 1976)* 2016;41:476–82.
4. Maeno T, Okuda S, Yamashita T, et al. Age-related surgical outcomes of laminoplasty for cervical spondylotic myelopathy. *Global Spine J* 2015;5:118–23.
5. Oshima Y, Miyoshi K, Mikami Y, et al. Long-Term Outcomes of Cervical Laminoplasty in the Elderly. *BioMed Res Internat* 2015;2015:75–86.
6. Hosono N, Yonenobu K, Ono K. Neck and shoulder pain after laminoplasty. A noticeable complication. *Spine* 1996;21:1969–73.
7. Moon MS. Neck and shoulder pain after laminoplasty. *Spine* 1997;22:1674–6.
8. Fujimori T, Le H, Ziewacz JE, Chou D, et al. Is there a difference in range of motion, neck pain, and outcomes in patients with ossification of posterior longitudinal ligament *versus* those with

cervical spondylosis, treated with plated laminoplasty? *Neurosurg Focus* 2013;35:E9.
9. Hyun SJ, Riew KD, Rhim SC. Range of motion loss after cervical laminoplasty: a prospective study with minimum 5-year follow-up data. *Spine J* 2013;13:384–90.
10. Liu J, Ebraheim NA, Sanford CG Jr, et al. Preservation of the spinous process-ligament-muscle complex to prevent kyphotic deformity following laminoplasty. *Spine J* 2007;7:159–64.
11. Suk KS, Kim KT, Lee JH, et al. Sagittal alignment of the cervical spine after the laminoplasty. *Spine* 2007;32:E656–60.
12. Suda K, Abumi K, Ito M, et al. Local kyphosis reduces surgical outcomes of expansive open-door laminoplasty for cervical spondylotic myelopathy. *Spine* 2003;28:1258–62.
13. Fujimori T, Iwasaki M, Okuda S, et al. Long-term results of cervical myelopathy due to ossification of the posterior longitudinal ligament with an occupying ratio of 60% or more. *Spine* 2014;39:58–67.
14. Sodeyama T, Goto S, Mochizuki M, et al. Effect of decompression enlargement laminoplasty for posterior shifting of the spinal cord. *Spine* 1999;24:1527–31.
15. Tanaka J, Seki N, Tokimura F, et al. Operative results of canal-expansive laminoplasty for cervical spondylotic myelopathy in elderly patients. *Spine* 1999;24:2308–12.
16. Matsuda Y, Shibata T, Oki S, et al. Outcomes of surgical treatment for cervical myelopathy in patients more than 75 years of age. *Spine* 1999;24:529–34.
17. Lafage V, Schwab F, Patel A, et al. Pelvic tilt and truncal inclination: two key radiographic parameters in the setting of adults with spinal deformity. *Spine* 2009;34:E599–606.
18. Jackson RP, Hales C. Congruent spinopelvic alignment on standing lateral radiographs of adult volunteers. *Spine* 2000;25:2808–15.
19. Djurasovic M, Glassman SD. Correlation of radiographic and clinical findings in spinal deformities. *Neurosurg Clin N Am* 2007;18:223–7.
20. Glassman SD, Berven S, Bridwell K, et al. Correlation of radiographic parameters and clinical symptoms in adult scoliosis. *Spine* 2005;30:682–8.
21. Rao SC, Fehlings MG. The optimal radiologic method for assessing spinal canal compromise and cord compression in patients with cervical spinal cord injury. Part I: an evidence-based analysis of the published literature. *Spine* 1999;24:598–604.
22. Glassman SD, Bridwell K, Dimar JR, et al. The impact of positive sagittal balance in adult spinal deformity. *Spine* 2005;30:2024–9.
23. Schwab FJ, Smith VA, Biserni M, et al. Adult scoliosis: a quantitative radiographic and clinical analysis. *Spine* 2002;27:387–92.
24. Tang JA, Scheer JK, Smith JS, et al. The impact of standing regional cervical sagittal alignment on outcomes in posterior cervical fusion surgery. *Neurosurgery* 2012;71:662–9.
25. Smith JS, Lafage V, Ryan DJ, et al. Association of myelopathy scores with cervical sagittal balance and normalized spinal cord volume: analysis of 56 preoperative cases from the AOSpine North America Myelopathy study. *Spine* 2013;38:S161–70.
26. Roguski M, Benzel EC, Curran JN, et al. Postoperative cervical sagittal imbalance negatively affects outcomes after surgery for cervical spondylotic myelopathy. *Spine* 2014;39:2070–7.



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

The impact of joint disease on the Modified Health Assessment Questionnaire scores in rheumatoid arthritis patients: A cross-sectional study using the National Database of Rheumatic Diseases by iR-net in Japan

Kumiko Ono, Satoru Ohashi, Hiroyuki Oka, Yuho Kadono, Tetsuro Yasui, Yasunori Omata, Jinju Nishino, Sakae Tanaka & Shigeto Tohma

To cite this article: Kumiko Ono, Satoru Ohashi, Hiroyuki Oka, Yuho Kadono, Tetsuro Yasui, Yasunori Omata, Jinju Nishino, Sakae Tanaka & Shigeto Tohma (2016) The impact of joint disease on the Modified Health Assessment Questionnaire scores in rheumatoid arthritis patients: A cross-sectional study using the National Database of Rheumatic Diseases by iR-net in Japan, Modern Rheumatology, 26:4, 529-533, DOI: 10.3109/14397595.2015.1106640

To link to this article: http://dx.doi.org/10.3109/14397595.2015.1106640

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://informahealthcare.com/mor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2016; 26(4):529–533

DOI: 10.3109/14397595.2015.1106640

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

The impact of joint disease on the Modified Health Assessment Questionnaire scores in rheumatoid arthritis patients: A cross-sectional study using the National Database of Rheumatic Diseases by iR-net in Japan

Kumiko Ono[1], Satoru Ohashi[2], Hiroyuki Oka[3], Yuho Kadono[1], Tetsuro Yasui[1,4], Yasunori Omata[1], Jinju Nishino[1], Sakae Tanaka[1], and Shigeto Tohma[5]

[1]Department of Orthopaedic Surgery, Faculty of Medicine, University of Tokyo, Tokyo, Japan, [2]Department of Orthopaedic Surgery, Sagamihara Hospital, National Hospital Organization, Kanagawa, Japan, [3]Department of Medical Research and Management for Musculoskeltal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, University of Tokyo, Tokyo, Japan, [4]Department of Orthopaedic Surgery, Teikyo University Mizonokuchi Hospital, Kanagawa, Japan, and [5]Clinical Research Center for Allergy and Rheumatology, Sagamihara National Hospital, National Hospital Organization, Kanagawa, Japan

Abstract

Objectives: To investigate the effect of bilateral and unilateral joint disease on the Modified Health Assessment Questionnaire (MHAQ) scores and the differences in joint weighting in rheumatoid arthritis patients.
Methods: A total of 9212 subjects from the Japanese nationwide cohort database NinJa, 2011, were analyzed. The presence or absence of disease in each joint, including swelling and/or tenderness, was investigated. The correlations between bilateral and unilateral disease in each joint and MHAQ scores were investigated using multivariable logistic regression analysis.
Results: The patients' mean age and disease duration was 63.2 and 12.2 years, respectively. The Disease Activity Score-28 was 3.3. The odds ratios of physical impairment according to the MHAQ using multivariable logistic regression models for bilateral and unilateral joints, respectively, were: shoulder, 4.0 and 1.8; elbow, 2.6 and 1.8; wrist, 1.9 and 1.5; hip, 1.7 and 3.0; knee, 2.6 and 1.9; ankle, 2.3 and 2.0, finger, 1.4 and 1.2; and toe, 1.0 and 1.1. The shoulder, elbow, wrist, knee, and ankle had a significant effect on physical impairment.
Conclusions: The MHAQ score was significantly affected by shoulder, elbow, wrist, knee, and ankle joint disease. Furthermore, bilateral disease tended to have a greater effect on physical impairment than unilateral disease.

Keywords

Cohort study, Joint involvement, Physical function, Rheumatoid arthritis, Scoring system

History

Received 5 August 2015
Accepted 30 September 2015
Published online 17 November 2015

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disease caused by activation of the osteoclast pathway, resulting in local bone destruction around joints and systemic osteoporosis. In addition, reduction of the joint range of motion and the destruction of joints lead to reduced daily physical activity [1]. With the development of pharmacotherapies, it has become possible to control the disease activity of RA and avoid worsening of physical function and activity. One of the major goals of RA treatment is to avoid disability. The Disease Activity Score (DAS) 28 is widely used for the assessment in RA mainly focusing on joint inflammation, and consists of swollen joint counts, erythrocyte sedimentation rate (ESR), and patient's global assessment [2]. The DAS28 mainly focuses on the upper extremities, and evaluates joints including the shoulder, elbow, wrist, finger, and knee. On the other hand, the Health Assessment Questionnaire (HAQ) measures functional disability [3]. Its modified version, the modified HAQ (MHAQ), was developed to assess patients' functional capacity in daily activities, such as dressing and grooming, standing up, eating, walking, hygiene, reach, grip, and common daily activities, in clinical trials and daily practice [4].

Previous studies showed that the HAQ score was associated with disease activity, swollen and tender joints, and laboratory data [5,6]. Accounting of joint diseases is important to evaluate the physical function of RA patients. On the other hand, a few studies showed that the impact of the impaired joint on the functional disability differed among individual joints [7]. However, no previous study has examined how joint disease affects physical function when the disorder is bilateral or unilateral. In addition, a consensus has not been reached for determining an appropriate weighted score system for joint evaluation in RA patients.

Using the National Database of Rheumatic Diseases by iR-net in Japan (NinJa), a multicenter, rheumatic disease database, the present study investigated the effect of bilateral and unilateral joint disease in the shoulder, elbow, wrist, hip, knee, ankle, finger, and toe on the MHAQ scores and the differences in joint weighting in RA patients.

Correspondence to: Satoru Ohashi, MD, PhD, Department of Orthopaedic Surgery, Sagamihara Hospital, National Hospital Organization, 18-1 Sakura-dai, Minami-ku, Sagamihara, Kanagawa 252-0315, Japan. Tel: +81-42-742-8311. Fax: +81-42-742-5314. E-mail: soohashi-tky@umin.ac.jp

Materials and methods

Data source

The study protocol was reviewed and approved by the Research Ethics Committees of the National Hospital Organization and each participating hospital, and all patients enrolled provided written informed consent. NinJa is a nationwide, multicenter, observational cohort database of rheumatic diseases that was established in 2002 in Japan [8]. The collected data consist of two components: one is the patient information over the course of the year (outcome, death, hospitalization, operation, number of total joint arthroplasties in large joints [hip, knee, shoulder, and elbow], malignancy, and tuberculosis), and the other is the information collected on an arbitrary day in the daily clinical practice (tender-joint and swollen-joint count, MHAQ Steinbrocker functional classification, Steinbrocker stage, patient global and pain visual analog scales [VAS], doctor VAS, ESR, C-reactive protein [CRP], DAS28-ESR, DAS28-CRP, and use of corticosteroids, methotrexate, and nonsteroidal anti-inflammatory drugs).

Patients

The subjects were the 9212 patients (1766 men, 7446 women) with complete medical records among the 10,367 patients registered in NinJa in fiscal year 2011 (from April 2011 to March 2012). The presence or absence of disease in each joint (swelling and/or tenderness were considered as disease) and whether the disease was bilateral or unilateral were investigated. The presence of disease in the finger or toe joints was defined as swelling and/or tenderness in at least one metacarpophalangeal joint, metatarsophalangeal joint, or proximal interphalangeal joint. In addition, if a joint had been treated with a surgical procedure, it was defined as having absence of disease.

Statistical analyses

Descriptive statistics were used to analyze the clinical information, demographic factors, and other test data. Continuous variables were expressed as means and standard deviation (SD). Furthermore, the correlation between the MHAQ and DAS28-ESR scores were examined by Spearman correlation coefficient.

The MHAQ median score of this study was 0.25. Functional impairment was defined as a MHAQ score ≥ 0.25. The MHAQ score of the functional impairment group was evaluated using univariate logistic regression, and the odds ratio was calculated. In the next step, the variables with a p value of <0.2 in the univariate analyses were included in a stepwise multivariable logistic regression model for functional impairment.

Receiver operating characteristic (ROC) curve analysis was performed to develop a support tool of functional impairment. Discriminatory power is the ability to identify which patients are likely to have a functional impairment, and it was determined using ROC curve analysis, in which an area under the ROC curve (AUC) of 1.0 indicated perfect discrimination, and an AUC ≥ 0.7 was, in general, considered to indicate acceptable discrimination [9]. Finally, to examine the performance of the support tool, we calculated the sensitivity and specificity. Statistical analyses were conducted using the JMP 10.0.2 software program (SAS, Cary, NC). All statistical tests were 2-tailed, and a significance level of 0.05 was used.

Results

The clinical features of the 9212 patients with RA are shown in Table 1. The patients had a mean ± SD age and disease duration of 63.2 ± 12.9 years and 12.2 ± 10.7 years, respectively. The majority of subjects had moderate disease activity (mean DAS28-ESR score, 3.3 ± 1.3). The mean MHAQ score was 0.48 ± 0.64.

Table 1. Baseline demographic and clinical characteristics of 9212 patients with rheumatoid arthritis.

Age, years	63.2 ± 12.9
Female, %	80.1
RA disease duration, years	12.2 ± 10.7
C-reactive protein, mg/100 mL	0.7 ± 1.3
Erythrocyte sedimentation rate, mm/h	30.3 ± 25.7
Patient's pain VAS, mm	27.1 ± 23.9
Patient's general VAS, mm	27.9 ± 23.9
Physician's general VAS, mm	17.9 ± 16.6
DAS28-ESR score	3.3 ± 1.3
DAS28-CRP score	2.6 ± 1.1
MTX use, %	61.6
MTX dosage, mg/week	5.2 ± 4.7
Corticosteroid use, %	46.7
Corticosteroid dosage, mg/day	4.2 ± 2.7
MHAQ score	0.48 ± 0.64
MHAQ score, median	0.25

Values are mean ± standard deviation unless otherwise indicated.
CRP, C-reactive protein; DAS, Disease Activity Score; ESR, erythrocyte sedimentation rate; MHAQ, Modified Health Assessment Questionnaire; MTX, methotrexate; RA, rheumatoid arthritis; VAS, visual analog scale.

The two most frequently affected joints were the finger joints (42.2%) and wrist (36.6%), followed by the knee (21.2%), ankle (20.9%), toe joints (18.7%), elbow (17.8%), and shoulder (11.5%). In contrast, the frequency of hip joint involvement was small (2.0%) (Figure 1).

There was a moderate correlation between the MHAQ and DAS28 scores by Spearman correlation coefficient ($r = 0.52$, $p < 0.01$). Significant associations were observed between the MHAQ scores and bilateral and unilateral disease of all joints except for bilateral disease of the hip and bilateral and unilateral disease of the toes. The odds ratios [95% confidence intervals] using multivariable logistic regression models for bilateral and unilateral joint, respectively, were as follows: shoulder, 4.0 [2.9–5.6] and 1.8 [1.5–2.1]; elbow, 2.6 [2.1–3.4] and 1.8 [1.5–2.1]; wrist, 1.9 [1.7–2.2] and 1.5 [1.3–1.7]; hip, 1.7 [0.7–4.7] and 3.0 [2.0–4.7]; knee, 2.6 [2.2–3.2] and 1.9 [1.7–2.2]; ankle, 2.3 [1.9–3.0] and 2.0 [1.8–2.4]; finger, 1.4 [1.2–1.5] and 1.2 [1.0–1.3]; and toe, 1.0 [0.8–1.3] and 1.1 [0.9–1.3] (Figure 2).

To develop a weighted scoring system from the results of this analysis, an integer score derived from the β-coefficient was assigned to each identified factor [10]. In this study, for each patient, all applicable score values were summed up to attain a total score. Regarding the odds ratio, an integer score was assigned to each identified bilateral and unilateral joint disease, respectively, as follows: shoulder, 4 points and 2 points; elbow, 3 points and 2 points; wrist, 2 points and 2 points; hip, 0 points and 3 points; knee, 3 points and 2 points; ankle 2 points and 2 points; and finger, 1 point and 1 point (Table 2). The total scores for each patient ranged from 0 to 18. The mean ± SD of the total scores was 2.5 ± 2.8. The results of ROC analysis were as follows: cut-off value, 3 points; AUC, 0.709; sensitivity, 58.6%; and specificity, 72.8% (Figure 3). We divided the patients into two groups by using a weighted scoring system cut-off value of 3 points. The distribution of patients and the mean MHAQ score of each group based on a weighted scoring system cut-off value of 3 points are shown in Table 3.

Discussion

In our study, the effects of disease of various joints on physical function were examined, and the findings showed that joint disease of the shoulders, elbows, knees, and ankles had a

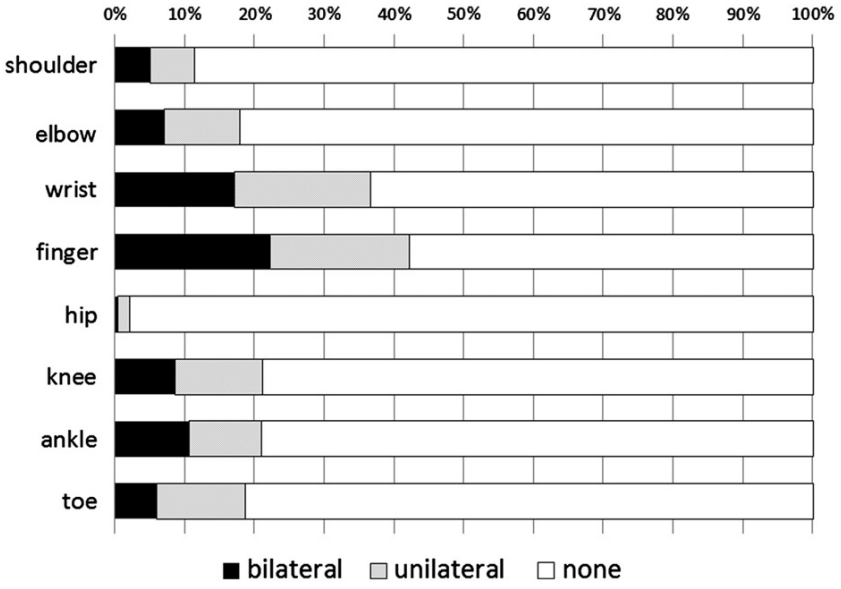


Figure 1. Percentage of involvement of each joint in rheumatoid arthritis patients. Black bars indicate bilaterally affected joints. Gray bars indicate unilaterally affected joints.

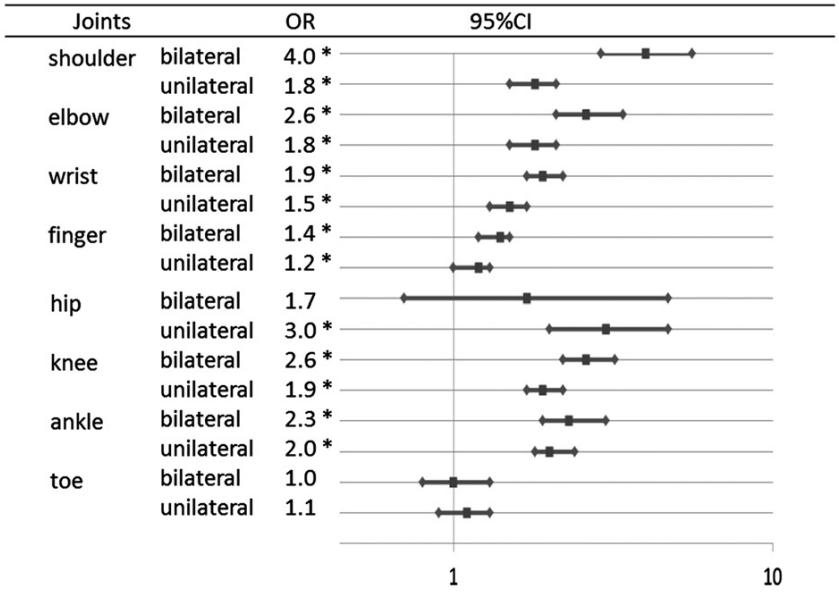


Figure 2. The odds ratios (ORs) and 95% confidence intervals (CIs) of physical impairment using multivariable logistic regression models for each joint with bilateral and unilateral disease. *$p < 0.05$.

significant effect on physical function. The findings also showed that deterioration of physical function was more severe when the disease affected the joints bilaterally than when the disease was unilateral. In addition, a clinically useful scoring system was developed based on these findings, allowing for weighting the scores, depending on whether each joint was bilaterally or unilaterally affected.

Thus far, a number of studies have shown that disease activity and physical function are correlated in RA [11–13]. In addition, painful and swollen joints can particularly lead to deterioration of physical function, regardless of disease duration [5,14,15].

Physical function assessments of joints in RA patients have previously been conducted on 68 joints [16] and 28 joints [2,17], and all joints were scored in exactly the same manner. However, the percentages of morbidity were different for every joint. In our study, the morbidity rates were high in the following joints in descending order: finger, wrist, and knee joints; our findings were consistent with those of previous reports [18,19]. Thus, the morbidity rate was different in each joint type, and the impact on physical function may be different depending on the affected joint. Reports from cross-sectional studies have mentioned that the type of joint such as the shoulder, knee, elbow, wrist, or ankle had a

Table 2. Multivariable predictors of physical impairment according to the MHAQ score and joint scoring system as a support tool.

Characteristic	Regression β-coefficient	95% CI	Score*
Shoulder bilateral	4.0	2.9–5.6	4
Shoulder unilateral	1.8	1.5–2.1	2
Elbow bilateral	2.6	2.1–3.4	3
Elbow unilateral	1.8	1.5–2.1	2
Wrist bilateral	1.9	1.7–2.2	2
Wrist unilateral	1.5	1.3–1.7	2
Hip unilateral	3.0	2.0–4.7	3
Knee bilateral	2.6	2.2–3.2	3
Knee unilateral	1.9	1.7–2.2	2
Ankle bilateral	2.3	1.9–3.0	2
Ankle unilateral	2.0	1.8–2.4	2
Finger bilateral	1.4	1.2–1.5	1
Finger unilateral	1.2	1.0–1.3	1

*The score was obtained by rounding the raw score to one decimal place if the coefficient was statistically significant.
CI, confidence interval; MHAQ, Modified Health Assessment Questionnaire.

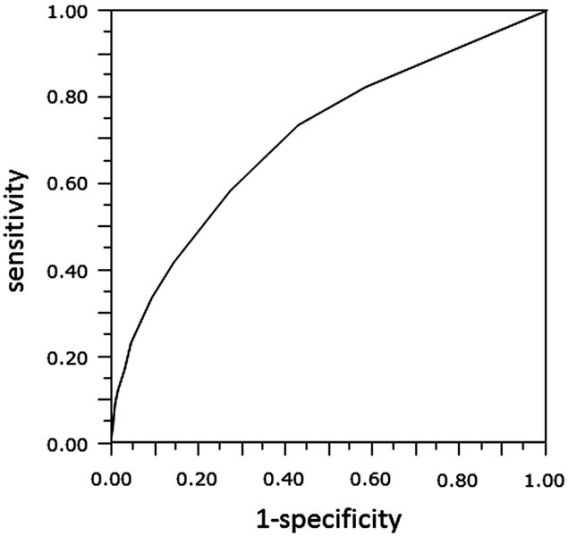


Figure 3. Receiver operating characteristic (ROC) curve for the weighted scoring system. Cut-off value, 3 points; area under the ROC curve, 0.709; sensitivity, 58.6%; specificity, 72.8%.

Table 3. Distribution of patients and the mean MHAQ score of each group based on a weighted scoring system cut-off value of 3 points.

	Score of the weighted scoring system	
	0–2	3–18
Number of patients	5247	3965
%	57.0	43.0
MHAQ score, mean ± SD	0.32 ± 0.53	0.70 ± 0.70
MHAQ score, 95% CI	0.30–0.33	0.68–0.73

CI, confidence interval; MHAQ, Modified Health Assessment Questionnaire; SD, standard deviation.

significant effect on the HAQ score [18]. In addition, a longitudinal study conducted on the same cohort showed that the group, which showed aggravation of the HAQ score during a 3-year period, was significantly affected by diseases of the shoulder, wrist, knee, and ankle joints [19]. A previous study showed that the HAQ score was not affected by joint damage due to radiography, but was affected by pain and range-of-motion limitations of the knee, shoulder, and wrist joints [20].

Our study showed that the MHAQ score was significantly affected by large joints (shoulder, elbow, knee, and ankle, but not the hip) as well as wrist joints, and this was consistent with previous findings [18]. In addition, the effect of ankle joint disease on physical function should also be considered in the daily practice; however, ankle joints are not assessed in the DAS28-ESR. On the other hand, one of the reasons why the presence of bilateral hip joint disease did not significantly affect the MHAQ score could have been the low number of subjects with bilateral hip disease.

The present study showed that in all joints, physical function impairment was more severe when the joints were affected bilaterally than when they were affected unilaterally. No previous report has examined the differences in the effect of bilateral or unilateral joint disease, according to individually affected joints, on physical function. The effect of bilateral and unilateral disorders on physical function tended to be different between upper limb joints and lower limb joints. The difference in the odds ratio of physical impairment between bilateral and unilateral disease was relatively greater in upper limb joints than in lower limb joints; when an upper limb joint is affected unilaterally, the function can be compensated by using the contralateral side.

Applying the weighted scoring system described in this study will account for disease in several joints in RA patients; in those with a cut-off score of 3 or higher, a greater impact of the joint disease on functional disability was predicted. When shoulders, elbows, and knee joints are affected bilaterally, aggressive treatment might be preferable. In addition, using the weighted scoring system may allow for predicting the MHAQ score on the basis of the sites and number of diseased joints. We divided the patients into two groups by using the cut-off value of the weighted scoring system, and the difference in the mean MHAQ score between the two groups was 0.3.

Because physical function and disease activity are correlated [11–13], using the appropriate medication and inhibiting the disease activity are of primary importance in order to maintain physical function. In addition, using the scoring system described in this study to elucidate the joints that are most deeply involved with physical functions may facilitate prioritizing the treatment of diseased joints. This will allow for confirmation of the presence of functional impairments in carrying out activities of daily living and improvement in the inhibition of systemic inflammation through pharmacological treatment, recommendation of self-help devices, rehabilitation therapy, and surgical treatments such as arthrodesis and prosthetic joint replacement surgery, and these measures could reduce the load on joints, and prevent the exacerbation of symptoms and functional impairments.

This study has some limitations. First, this was a cross-sectional study using a cohort database. Second, we did not classify joint disease by tenderness and swelling. We needed to binarize joint disease for the multivariate analysis. Third, we did not evaluate the number affected fingers and toes because we considered that the additional 20 joints might cause confusion in the results of the multivariate analysis. This may have caused unclear differences between the effects of other major joints on the MHAQ; therefore, we evaluated all fingers and toes as a single unit. These two limitations may have affected our scoring system. Fourth, the validation of our scoring system was conducted on a single population; hence, validation of the evidence on another population group will be needed in the future. However, this study also had certain advantages. A large database, in which the MHAQ and DAS28-ESR scores were moderately correlated, was used, and it did not deviate from the general knowledge of the RA population.

Our scoring system has the advantage of providing a cut-off value of 3 points. For examples, according to this weighted scoring system, a patient with bilateral tenderness on the shoulders (4 points) might have worse MHAQ scores than another patient with unilateral tenderness on the knee (2 points). By using this scoring system, we believe that rheumatologists can predict functional disability in a simpler way by examining each joint. Moreover, the scoring system developed in this study will be validated if its use leads to a more accurate prediction of functional disability.

While the MHAQ scores were significantly affected by disease in almost all joints, a greater effect was exerted by the following major joints, in increasing order: ankle, knee, elbow, and shoulder. Bilateral disease tended to have a greater effect on physical impairment than unilateral disease in these major joints and the wrist. We believe that the use of weighted scoring system in the clinical setting will improve the accuracy of predicting functional disability in RA patients.

Acknowledgments

The authors thank Mayumi Yokoyama, who provided expert technical assistance. The authors acknowledge the assistance of the following clinicians who referred patients to NinJa: Kumiko Akiya, Tomiaki Asai, Noriyuki Chiba, Yoshito Eto, Kenji Ichikawa, Atsushi Kaneko, Yojiro Kawabe, Toshihiro Matusi, Kinori Matsumori, Ryutaro Matsumura, Satoru Motokawa, Kunikazu Ogawa, Akira Okamoto, Koichiro Omura, Yusuke Ota, Yukihiko Saeki, Koichiro Saisho, Tomotaro Sato, Yoshiki Shinohara, Makoto Sueishi, Eichi Suematu, Yasuo Suenaga, Shoji Sugii, Takao Sugiyama, Koichiro Takashi, Norio Tamura, Kaeharu Tonai, Issaku Toyohara, Hiroshi Tsutani, Hajime Yamagata, Hidetoshi Yanagida, Masyuki Yasuda, Yasuhiko Yoshinaga, and Shigeru Yoshizawa.

Conflict of interest

This work was supported in part by Health and Labour Sciences Research Grants from the Ministry of Health, Labour and Welfare of Japan to Shigeto Tohma.

References

1. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet. 2010;376(9746):1094–108.
2. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38(1):44–8.
3. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137–45.
4. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum. 1983;26(11):1346–53.
5. Sokka T, Kankainen A, Hannonen P. Scores for functional disability in patients with rheumatoid arthritis are correlated at higher levels with pain scores than with radiographic scores. Arthritis Rheum. 2000;43(2):386–9.
6. Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LALM, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum. 2001;44(9):2009–17.
7. Bandeira M, Falcone A, Pistorio A, Ruperto N, Magni-Manzoni S, Buoncompagni A, et al. Weighting improves the information provided by joint counts on the severity of arthritis and its impact on patients' well-being in juvenile idiopathic arthritis. Rheumatology (Oxford). 2006;45(3):343–7.
8. Yamanaka H, Tohma S. Potential impact of observational cohort studies in Japan on rheumatoid arthritis research and practice. Mod Rheumatol. 2006;16(2):75–6.
9. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Hosmer DW, Lemeshow S, eds. Applied logistic regression. 2nd ed. New York: Wiley; 2000:143–202.
10. Yamada H, Oka H, Iwasaki H, Endo T, Kioka M, Ishimoto Y, et al. Development of a support tool for the clinical diagnosis of symptomatic lumbar intra- and/or extra-foraminal stenosis. J Orthop Sci. 2015;20(5):811–17.
11. Hernández-Hernández V, Ferraz-Amaro I, Díaz-González F. Influence of disease activity on the physical activity of rheumatoid arthritis patients. Rheumatology (Oxford). 2014;53(4):722–31.
12. Jansen LM, van Schaardenburg D, van Der Horst-Bruinsma IE, Bezemer PD, Dijkmans BA. Predictors of functional status in patients with early rheumatoid arthritis. Ann Rheum Dis. 2000;59(3):223–6.
13. Boyd TA, Bonner A, Thorne C, Boire G, Hitchon C, Haraoui BP, et al. The relationship between function and disease activity as measured by the HAQ and DAS28 varies over time and by rheumatoid factor status in early inflammatory arthritis (EIA). Results from the CATCH cohort. Open Rheumatol J. 2013;7:58–63.
14. Sarzi-Puttini P, Fiorini T, Panni B, Turiel M, Cazzola M, Atzeni F. Correlation of the score for subjective pain with physical disability, clinical and radiographic scores in recent onset rheumatoid arthritis. BMC Musculoskelet Disord. 2002;3:18.
15. Molenaar ET, Voskuyl AE, Dijkmans BA. Functional disability in relation to radiological damage and disease activity in patients with rheumatoid arthritis in remission. J Rheumatol. 2002;29(2):267–70.
16. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum. 1995;38(6):727–35.
17. Smolen JS, Breedveld FC, Eberl G, Jones I, Leeming M, Wylie GL, et al. Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. Arthritis Rheum. 1995;38(1):38–43.
18. Tanaka E, Saito A, Kamitsuji S, Yamada T, Nakajima A, Taniguchi A, et al. Impact of shoulder, elbow, and knee joint involvement on assessment of rheumatoid arthritis using the American College of Rheumatology Core Data Set. Arthritis Rheum. 2005;53(6):864–71.
19. Shidara K, Inoue E, Hoshi D, Tanaka E, Seto Y, Nakajima A, et al. The influence of individual joint impairment on functional disability in rheumatoid arthritis using a large observational database of Japanese patients. J Rheumatol. 2012;39(3):476–80.
20. Hakkinen A, Kautianinen H, Hannonen P, Ylinen J, Arkela-Kautiainen M, Sokka T. Pain and joint mobility explain individual subdimensions of the health assessment questionnaire (HAQ) disability index in patients with rheumatoid arthritis. Ann Rheum Dis. 2005;64(1):59–63.

PLOS | ONE

RESEARCH ARTICLE

Classification of High Intensity Zones of the Lumbar Spine and Their Association with Other Spinal MRI Phenotypes: The Wakayama Spine Study

Masatoshi Teraguchi[1], Dino Samartzis[2]*, Hiroshi Hashizume[1]*, Hiroshi Yamada[1], Shigeyuki Muraki[3], Hiroyuki Oka[4], Jason Pui Yin Cheung[2], Ryohei Kagotani[1], Hiroki Iwahashi[1], Sakae Tanaka[5], Hiroshi Kawaguchi[6], Kozo Nakamura[7], Toru Akune[7], Kenneth Man-Chee Cheung[2], Noriko Yoshimura[3], Munehito Yoshida[1]

1 Department of Orthopaedic Surgery, Wakayama Medical University, 811-1 Kimiidera, Wakayama, Japan, 641–8509, **2** Department of Orthopaedics and Traumatology, The University of Hong Kong, Professorial Block, 5th Floor 102 Pokfulam Road, Pokfulam, Hong Kong, SAR, China, **3** Department of Joint Disease Research, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan, 113-8655, **4** Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan, 113-8655, **5** Department of Orthopaedic surgery, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyoku, Tokyo, 113–8655, Japan, **6** Japan Community Healthcare Organization Tokyo Shinjuku Medical Center, 5–1 Tsukudo-chome, Shinjuku-ku, Tokyo, Japan, 162–8543, **7** Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, 1 Namiki 4-chome, Tokorozawa City, Saitama, Japan, 359–8555

* dsamartzis@msn.com (DS); hashizum@wakayama-med.ac.jp (HH)

Citation: Teraguchi M, Samartzis D, Hashizume H, Yamada H, Muraki S, Oka H, et al. (2016) Classification of High Intensity Zones of the Lumbar Spine and Their Association with Other Spinal MRI Phenotypes: The Wakayama Spine Study. PLoS ONE 11(9): e0160111. doi:10.1371/journal.pone.0160111

Editor: Hiroyuki Tsuchiya, Kanazawa University, JAPAN

Received: May 11, 2016

Accepted: July 13, 2016

Published: September 20, 2016

Data Availability Statement: The present study used resident data from two communities in Wakayama prefecture. It is impossible for us to provide and upload these data in a public repository because we have confirmed with these municipalities and residents that data will remain confidential. We will provide anonymized data on request after discussing the contents with the municipalities, as long as researchers are qualified to request these data. Data requests can be made

Abstract

Introduction

High intensity zones (HIZ) of the lumbar spine are a phenotype of the intervertebral disc noted on MRI whose clinical relevance has been debated. Traditionally, T2-weighted (T2W) magnetic resonance imaging (MRI) has been utilized to identify HIZ of lumbar discs. However, controversy exists with regards to HIZ morphology, topography, and association with other MRI spinal phenotypes. Moreover, classification of HIZ has not been thoroughly defined in the past and the use of additional imaging parameters (e.g. T1W MRI) to assist in defining this phenotype has not been addressed.

Materials and Methods

A cross-sectional study of 814 (69.8% females) subjects with mean age of 63.6 years from a homogenous Japanese population was performed. T2W and T1W sagittal 1.5T MRI was obtained on all subjects to assess HIZ from L1-S1. We created a morphological and topographical HIZ classification based on disc level, shape type (round, fissure, vertical, rim, and enlarged), location within the disc (posterior, anterior), and signal type on T1W MRI (low, high and iso intensity) in comparison to the typical high intensity on T2W MRI.

to the corresponding author at hashizum@wakayama-med.ac.jp.

Funding: The author(s) received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: HIZ, high intensity zone; DD, disc degeneration; MRI, magnetic resonance imaging; ROAD, Research on Osteoarthritis/Osteoporosis Against Disability.

Results

HIZ was noted in 38.0% of subjects. Of these, the prevalence of posterior, anterior, and both posterior/anterior HIZ in the overall lumbar spine were 47.3%, 42.4%, and 10.4%, respectively. Posterior HIZ was most common, occurring at L4/5 (32.5%) and L5/S1 (47.0%), whereas anterior HIZ was most common at L3/4 (41.8%). T1W iso-intensity type of HIZ was most prevalent (71.8%), followed by T1W high-intensity (21.4%) and T1W low-intensity (6.8%). Of all discs, round types were most prevalent (anterior: 3.6%, posterior: 3.7%) followed by vertical type (posterior: 1.6%). At all affected levels, there was a significant association between HIZ and disc degeneration, disc bulge/protrusion and Modic type II ($p<0.01$). Posterior HIZ and T1W high-intensity type of HIZ were significantly associated with disc bulge/protrusion and disc degeneration ($p<0.01$). In addition, posterior HIZ was significantly associated with Modic type II and III. T1W low-intensity type of HIZ was significantly associated with Modic type II.

Conclusions

This is the first large-scale study reporting a novel classification scheme of HIZ of the lumbar spine. This study is the first that has utilized T2W and T1W MRIs in differentiating HIZ sub-phenotypes. Specific HIZ sub-phenotypes were found to be more associated with specific MRI degenerative changes. With a more detailed description of the HIZ phenotype, this scheme can be standardized for future clinical and research initiatives.

Introduction

Since the advent of magnetic resonance imaging (MRI), there has been a tremendous interest to identify unique spinal phenotypes (e.g. patterns of intervertebral disc degeneration (DD), Modic changes, endplate abnormalities) that may be representative of the degenerative disc process and that may provide insight into determining the painful disc level(s) [1–7]. High-intensity zones (HIZ) of the lumbar spine are an example of a disc phenotype that have gathered widespread interest since their initial report in 1992 by Aprill and Bogduk [8]. Based on their report, HIZ was described as a hyperintense signal in the posterior annulus fibrosus of the disc on T2-weighted (T2W) MRI using only a relatively low-strength 0.6 Tesla scanner in patients with low back pain (LBP) undergoing discography. Since then, numerous reports have surfaced attempting to address the clinical relevance of HIZ and its relationship with LBP, but the significance of this association remains under heated debate [8–16]. Some studies have suggested that lumbar HIZ is related to a concordant pain response on discography and have concluded it to be a significant MRI biomarker for the diagnosis of LBP [8–11]. Alternatively, others studies have not found any association between HIZ with LBP [12–16]. To further complicate this issue, the prevalence of HIZ in symptomatic and asymptomatic populations has varied greatly between reported studies [8–16]. Besides symptomatology, additional controversies exist with regards to its pathology, natural history, and morphology/topography [1, 8–16]. This may be attributed to the lack of a strict phenotype definition of HIZ, proper sampling of the study samples with appropriate demographics, standardized imaging assessment methods, insufficient statistical analyses and consideration of occupational/lifestyle factors, limited

knowledge regarding its relationship with other spinal phenotypes, and the poor imaging resolution of particular MRI sequences [1, 8–16].

Understanding the pathogenesis of HIZ is necessary to clearly define its clinical significance with regards to LBP. Previous reports suggested that HIZ was an effect of annular tears leading to an accumulation of disc material that is toxic to the disc and surrounding neural structures, and may cause further degenerative changes within the intervertebral disc [9, 10, 13, 17, 18]. Alternatively, annular tears were also reported to appear in the early stages of DD [19]. Therefore, the relationship between HIZ and DD remains unclear. Traditionally, annular tears require discography, an invasive examination, in order to determine the type of tear that produces degenerative changes and pain. The MRI is a non-invasive method used to characterize HIZ but there is currently no standardized classification system for researchers to phenotype HIZ and most descriptions are based solely upon T2W MRI. As such, these concerns need to be addressed since they are an important initial step to better understand the pathobiology, prevalence, etiology, and clinical significance of HIZ. In addition to the lack of standardized phenotyping, the role of varying morphological/topographical traits of HIZ remains unknown and demand attention.

Coupling of T2W and T1W MRI sequences have been found useful to elaborate upon various spinal phenotypes, such as Modic changes and their classification, and have shed light upon their clinical relevance and decision-making [1, 20–27]. However, to date, no such approach has been adopted for HIZ. Therefore, utilizing a multimodal MRI approach to better characterize the HIZ phenotype is imperative to assist communication between study centers and aid large scale cross-cohort and cross-ethnic analyses. Furthermore, better understanding of HIZ may contribute to more sensitive identification of symptomatic disc levels, prediction and progression of disc or adjacent endplate changes, and potential use for patient selection for regenerative therapies for the disc. It also has potential to be a marker for identifying patients at risk for adjacent segment degeneration/disease in relation to a fusion or arthroplasty procedure.

Due to the limitations as addressed, better classification and understanding of HIZ is needed. Thus, this current study's objectives are three-fold and are based on a large-scale, population-based study. Firstly, we aimed to address the prevalence and morphological/topographical variations of HIZ throughout the lumbar spine using both T2W and T1W MRI. This imaging mapping further facilitated the creation of a novel classification of HIZ. Secondly, we aimed to assess the association of HIZ with other MRI spinal phenotypes.

Methods

Participants

This was a cross-sectional study based on the *Wakayama Spine Study* [28–34], a large population-based study created to address the etiology of common spinal disorders in Japan. Our study population was a sub-cohort of the large-scale population-based cohort study called *Research on Osteoarthritis/Osteoporosis Against Disability* (ROAD). The ROAD study was a nationwide, prospective study of bone and joint diseases consisting of population-based cohorts established in three communities in Japan [35–38]. The participants of ROAD study were recruited from listings of resident registrations in three communities that have different characteristics based on three geographical regions: an urban region in "I town" (Tokyo); a mountainous region in "H town" (Wakayama); and a coastal region in "T town" (Wakayama). *The Wakayama Spine Study* started in mountainous region H town and coastal region T town in Wakayama from 2008 as a population-based cohort [28–34]. For the current study, recruited

subjects were 20 years of age or older, irrespective of gender residing in T town who were willing to respond to a survey distributed in 2013.

The inclusion criteria were the ability to walk to the survey site, report data, and sign an informed consent form. Subjects with spinal tumors, infections, chronic inflammatory conditions, previous posterior spinal fusion operation, contraindicated to MRI (e.g., pacemakers) and/or other disqualifiers (e.g., pregnant) were excluded. In total, 857 individuals underwent MRI of the lumbar spine. However, 43 MRI results were not available due to incomplete T1W and T2W sagittal lumbar spine MRI or of quality too poor to read for HIZ. The *Wakayama Spine Study* obtained approval from the local ethics committee of the University of Tokyo, the Tokyo Metropolitan Institute of Gerontology, and Wakayama Medical University. All participants provided their own written informed consent.

MRI Assessment

Lumbar MRI were performed using a mobile MRI unit (Achieva 1.5 T; Philips Medical Systems, Best, The Netherlands) for all participants. On the same day of imaging assessment, participants also completed standardized questionnaires and underwent anthropometric examination, which accounted for height (meters) and weight (kilograms) as well as additional subject demographics (e.g. age [years], sex-type). All participants underwent MRI in the supine position. The imaging protocol included sagittal T2W fast-spin echo (FSE), with a repetition time (TR) of 3000 ms/echo and an echo time (TE) of 120 ms. The field of view (FOV) was 270 × 270 mm. The sagittal T1W FSE was with a TR of 540 ms/echo, a TE of 10 ms and a FOV of 270 × 270 mm. All cuts were 5mm thick and 11 total slices were available.

Evaluation of MRI

HIZ was defined as a bright white signal located in the substance of the annulus fibrosus, clearly dissociated from the signal of the nucleus pulposus, which was surrounded by a low-intensity (black) signal of the annulus fibrosus and in turn was appreciably brighter than the cerebrospinal fluid signal at the same level on T2W sagittal MR images of L1-S1 [8, 13]. Our novel classification of HIZ was created based on the disc level, shape (round type, fissure type, vertical type, rim type, and enlarge type), and location within disc (posterior or anterior) (Table 1, Fig 1). We also included details regarding the signal type on either T1W MRI (low-intensity, high-intensity, and iso-intensity signal) and T2W MRI (high-intensity signal) (Table 1, Fig 2). The novel classification scheme was developed based on empirical evidence and observational variants as noted between both imaging modalities in the context of HIZ, further agreed to by a panel of experts on spinal phenotyping.

Sagittal T2W and T1W MRI were used to assess the intervertebral space from L1/L2 to L5/S1. HIZ assessment was performed by a board certified orthopedic surgeon (MT) who was blinded to the background of the subjects. For evaluating intra-observer variability, 20 randomly selected lumbar MRIs were rescored by the same observer (MT) more than 1 month after the first reading, again blinded to the patient details. For inter-observer variability, another 20 MRIs (100 discs) were scored by 2 board certified orthopedic surgeons (MT and HI) using the same classification system. The intra- and inter-observer reliabilities for HIZ on T2W MRI were evaluated by kappa analysis and were 0.92 and 0.84 ($p<0.0001$, 95% confidence interval (CI): 0.96–1.06), respectively. As for the intensity of HIZ on T1W-MRI, kappa analysis of the intra- and inter-observer reliabilities were 0.90 and 0.82 ($p<0.0001$, 95% CI: 0.83–0.95). Kappa statistics >0.90 were considered excellent, 0.80–0.90 were considered good, 0.60–0.80 were considered fair, and <0.60 were considered poor [39, 40]. Any disagreements

Table 1. Assessment of lumbar High Intensity Zones on MRI.

Variables	Definition
Shape	
Round	Concentric or oval cavity
Fissure	Parallel and transverse layer to the adjacent endplate
Vertical	Vertical layer to the adjacent endplate
Rim	Oblique radiating layer from the adajacent endplate
Enlarged	Greater concentric area than typical round HIZ
Horizontal location within disc	
Posterior	HIZ located in the posterior annulus fibrosus
Anteriror	HIZ located in the anterior annulus fibrosus
Signal type on T1W and T2W HIZ image	
T1W low-intensity type of HIZ	Decreased signal than the bone marrow on T1W sagittal MRI
T1W high-intensity type of HIZ	Increased signal than the bone marrow on T1W sagittal MRI
T1W iso-intensity type of HIZ	Same signal than the bone marrow on T1W sagittal MRI

HIZ: high intensity zones, MRI: magnetic resonance imaging, T1W: T1-weighted, T2W: T2-weighted, MRI: magnetic resonance imaging

doi:10.1371/journal.pone.0160111.t001

in classification were settled by consensus after the reliability assessments were completed. The final classification of HIZ was agreed upon by both observers and DS.

Other spinal MRI phenotypes, such as DD, disc displacement, Modic changes, and Schmorl's node (SN) were also assessed by two board certified orthopedic surgeons (MT and

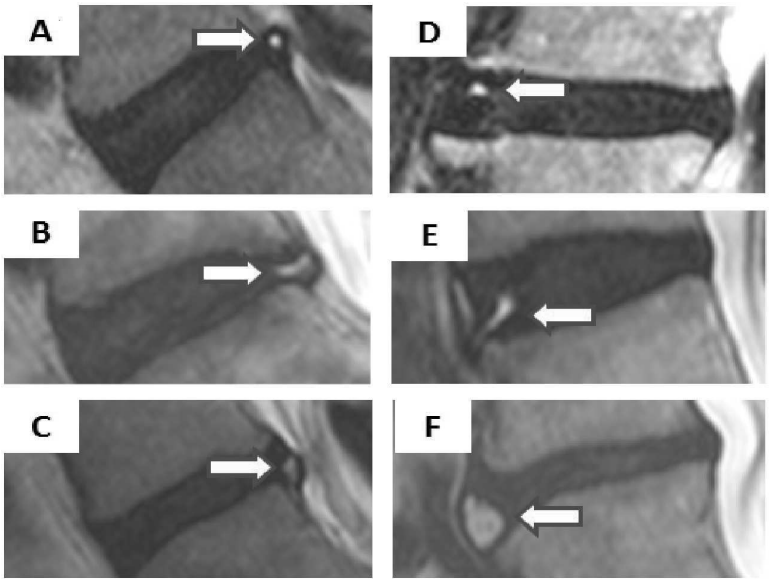


Fig 1. Classification of High Intensity Zones based on morphology and topography. High Intensity Zones (HIZ) were defined as a high intensity signal (white) surrounded by low intensity (black) located in the annulus fibrosus on T2-weighted sagittal MRI. Six types of HIZs were created based on the shape (round type, fissure type, vertical type, rim type, and giant type), and location within the disc (posterior or anterior). The images represent **(A)** posterior round type, **(B)** posterior fissure type, **(C)** posterior vertical type, **(D)** anterior round type, **(E)** anterior rim type, and **(F)** anterior enlarged type.

doi:10.1371/journal.pone.0160111.g001

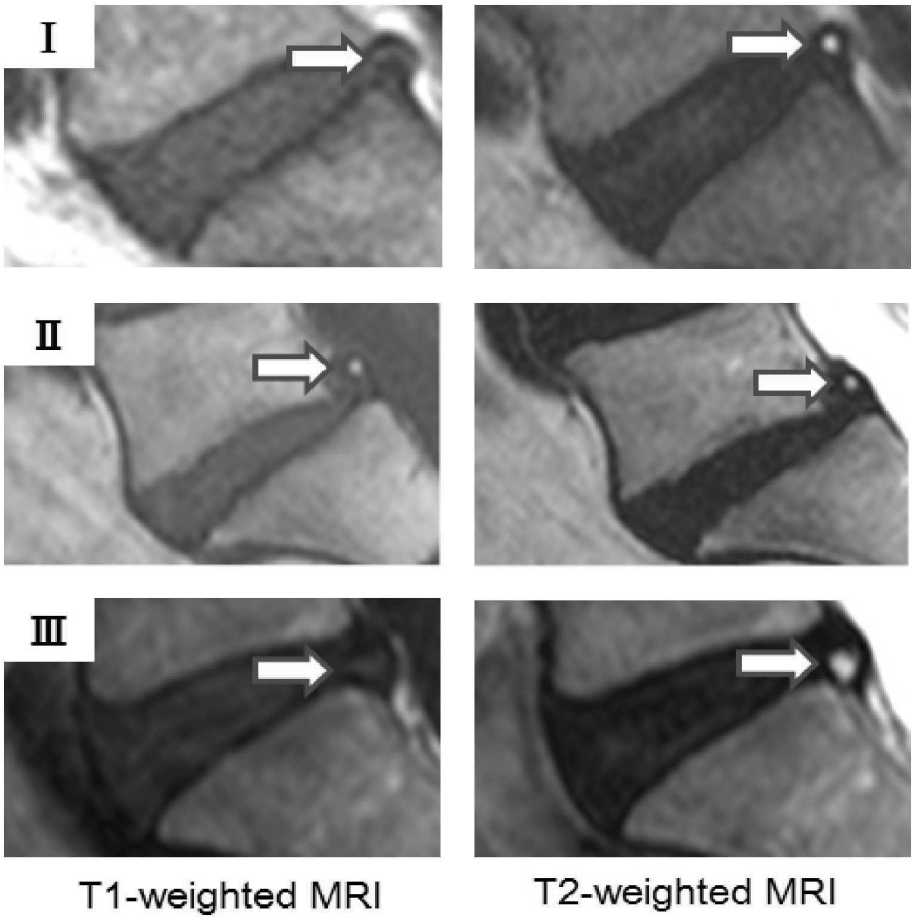


Fig 2. High Intensity Zones based on signal types on T1- and T2- weighted MRI. Three types of High Intensity Zones (HIZ) were created based on the signal type on T1-weighted MRI (low-intensity, high-intensity, and iso-intensity signal) and T2-weighted MRI (high-intensity signal). (I) T1-weighted low-intensity and T2-weighted high-intensity image, (II) T1-weighted high-intensity and T2-weighted high-intensity image, and (III) T1-weighted iso-intensity and T2-weighted high-intensity.

doi:10.1371/journal.pone.0160111.g002

RK). DD was classified by grade 4 or 5 on sagittal T2W MRI based on Pfirrmann's classification [41]. Disc displacement was evaluated as a disc bulge, protrusion, or extrusion. Disc bulge was defined as a disc displacement posteriorly beyond the line of the posterior edges of the adjacent vertebral bodies. Disc protrusion was noted as the nucleus displacement beyond the confines of the annulus fibrosus. Disc extrusion was recognized when the distance between the edges of the disc material beyond the disc space was greater than the distance between the edges of the base of the disc material beyond the disc space [42, 43]. Modic change was defined as diffuse areas of signal change along the endplates, and always parallel to the vertebral end plates on sagittal T1 and T2W images. Modic classification was based on the description originally proposed by Modic *et al* [44] on MRI: Type I was defined as decreased signal intensity on T1W and increased signal intensity on T2W, Type II change was defined as increased signal intensity on both T1W and T2W, and Type III change was defined as decreased signal intensity on both T1W and T2W. Endplate abnormality in any rostral or caudal endplate were assessed as SN defined as a local vertebral endplate defect/abnormality in deviation of the typical

concavity or flattened continuous shape of the endplate [30,45]. The intra- and inter-observer reliabilities of these additional MRI phenotypes have been previously reported to be good to excellent [30, 39, 40].

Statistical analysis

All statistical analyses were performed using JMP version 8 (SAS Institute Japan, Tokyo, Japan). Prevalence of HIZ was examined both per subjects and per disc level. Presence of HIZ was defined as having at least one HIZ in the lumbar region. Moreover, we assessed the prevalence of HIZ regarding shape (round type, fissure type, vertical type, rim type, and enlarge type), location within disc (posterior or anterior), and signal types on T1W MRI of HIZ in the lumbar region and at each affected lumbar disc level, respectively. Pearson χ^2 test and ANOVA (analysis of variance) with within group Tukey post-hoc tests were used to assess the association between HIZ and no HIZ, between posterior HIZ and anterior HIZ, and among T1W low-, high-, and iso- intensity type of HIZ where applicable. Non-paired student t-test was performed to compare continuous Pfirrmann grade at HIZ affected disc level. The threshold for statistical significance was a p-value less than 0.05.

Results

There were 814 individuals who underwent lumbar MRI assessment, of which 246 were males (30.2%) and 568 were females (69.8%). The mean age of the subjects was 63.6 years (SD: ±13.1 years). The mean age of males was 63.1 years (SD: ±14.0 years) and the mean age of females was 63.8 years (SD: ±12.7 years). The mean height was 166.8 cm (SD: ±6.7 cm) in males and 153.3 cm (SD: ±6.4 cm) in females. The mean weight was 66.8kg (SD: ±11.0kg) in males and 53.1 kg (SD: ±9.0 kg) in females. In addition, the mean body mass index (BMI) was 24.0 kg/m^2 (SD: ±3.6 kg/m^2) in males and 22.6 kg/m^2 (SD: ±3.6 kg/m^2) in females.

Prevalence of HIZ

HIZ were noted in 38.0% (n = 309) of all participants, and within these subjects the prevalence of posterior HIZ, anterior HIZ, and both posterior/anterior HIZ in the overall lumbar spine were 47.3% (n = 146), 42.4% (n = 131), and 10.4% (n = 32), respectively. Of the 309 HIZ subjects, 26.0% had single HIZ (n = 212), 8.6% had 2 HIZs (n = 70), 2.7% had 3 HIZs (n = 22) and 6.1% had 4 HIZs (n = 5). Of these subjects, involved discs only had a single HIZ. In addition, of the 97 multilevel HIZ subjects, 71.1% had consecutive level HIZs (n = 69) and 26.9% had skipped level HIZs (n = 28). The overall percentage prevalence of posterior and anterior HIZ per lumbar levels is illustrated in Fig 3. Posterior HIZ was most common at L5/S1 followed by L4/5. Alternatively, anterior HIZ had the highest prevalence at L3/4 followed by L2/3. As such, region-specific variations between upper (L1-L4) and lower (L4-S1) lumbar spine HIZ were noted.

Morphology and topography of HIZ

Table 2 illustrated the morphological distributions of HIZs of the lumbar discs. Round type HIZ (Fig 1A and 1D) were most common in both posterior and anterior discs. Furthermore, round type HIZ in the posterior disc was more common at L4/L5 and L5/S1, whereas round type HIZ in the anterior disc was most common at L2/L3 and L3/L4. Fissure type and vertical type HIZ in the posterior disc (Fig 1B and 1C) was most common at L5/S1 and L4/L5. Rim type and enlarged type HIZ in the anterior disc (Fig 1E and 1F) were most common at L3/L4 and L4/L5. In addition, of the 309 subjects with HIZ, 222 (71.8%) had T1W iso-intensity type

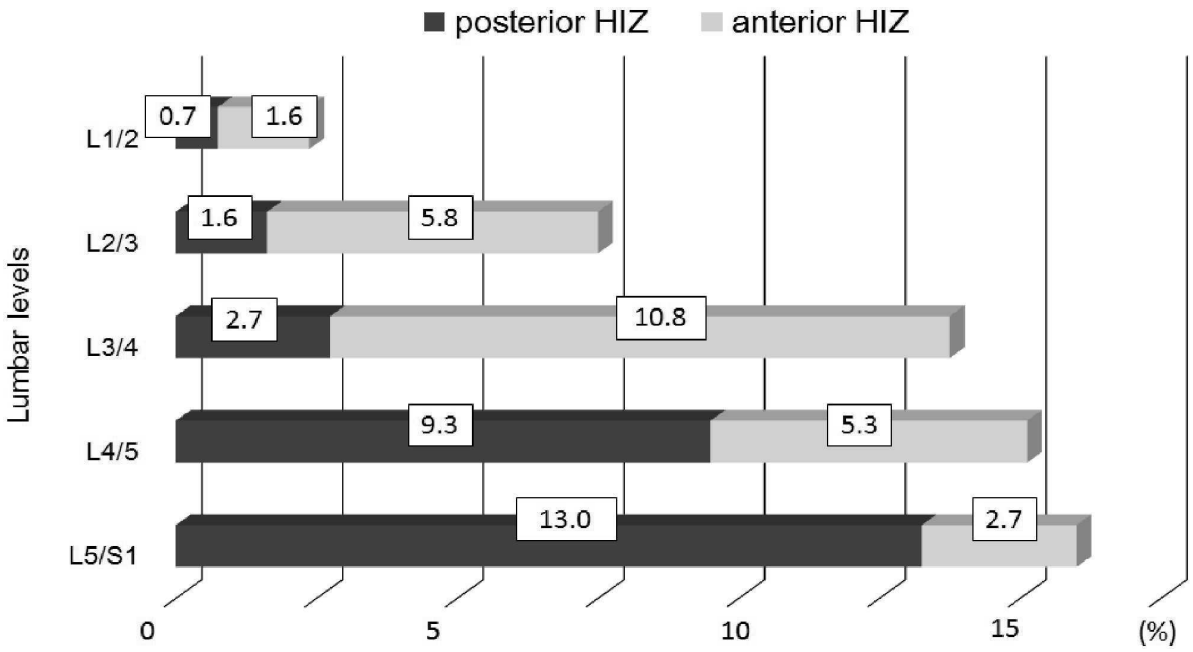


Fig 3. Bar chart showing the overall percent prevalence of anterior and posterior High Intensity Zones per lumbar level. Posterior HIZ was most common at L5/S1 followed by L4/5. Alternatively, anterior HIZ had the highest prevalence at L3/4 followed by L2/3.

doi:10.1371/journal.pone.0160111.g003

of HIZ (Fig 2 type III), followed by 66 (21.4%) with T1W high-intensity type of HIZ (Fig 2 type II) and 21 (6.8%) with T1W low-intensity type of HIZ (Fig 2 type I). As for disc level, T1W iso-intensity type of HIZ was most common at L4/L5 (11.5%, n = 94) followed by L5/S1 (11.4%, n = 93), T1W high-intensity type of HIZ was the highest at L5/S1 (3.7%, n = 30) followed by L4/L5 (2.9%, n = 24), and T1W low-intensity type of HIZ was the highest at L4/L5 (1.0%, n = 8) followed by L3/L4 (0.9%, n = 7).

Association of other spinal MRI phenotypes

As Table 3 illustrates the presence of HIZ was a clear determinant whether that disc level had other spinal MRI phenotypes or not. Disc levels with HIZ had significantly more disc bulges/protrusions (37.9% vs 29.3%, $p<0.01$) and DD (median 3.8, SD: ± 0.7 vs. 3.7, SD: ± 0.7, $p<0.001$), but not extrusions (1.1% vs 1.3%, p = 0.97). Modic type II change was significantly

Table 2. Distribution of shapes of High Intensity Zones at lumbar levels (n: 814 subjects).

Disc level	Posterior round, n (%)	Posterior fissure, n (%)	Posterior vertical, n (%)	Anterior round, n (%)	Anterior rim, n (%)	Anterior enlarged, n (%)
L1/L2	6 (4.0)	0 (0)	0 (0)	11 (7.5)	2 (3.7)	0 (0)
L2/L3	9 (6.0)	0 (0)	4 (6.2)	42 (28.8)	5 (9.3)	0 (0)
L3/L4	16 (10.6)	0 (0)	6 (9.2)	61 (41.8)	21 (38.9)	6 (40.0)
L4/L5	49 (32.5)	3 (42.9)	24 (36.9)	20 (13.7)	17 (30.9)	8 (53.3)
L5/S1	71 (47.0)	4 (57.1)	31 (47.7)	12 (8.2)	9 (16.4)	1 (6.7)
Total	151(100)	7 (100)	65 (100)	146 (100)	54 (100)	15 (100)

Note, every disc level from L1/L2 to L5/S1 has been individually evaluated.

doi:10.1371/journal.pone.0160111.t002

Table 3. Associated variables with High Intensity Zones at affected lumbar levels.

Variables	HIZ	No HIZ	p-value	Posterior HIZ	Anterior HIZ	p- value	T1W low-intensity type of HIZ	T1W high-intensity type of HIZ	T1W iso-intensity type of HIZ	p-value
Total discs; 4070	438	3632		223	215		22	80	339	
HIZ affected disc level										
Disc bulges/ protrusions, n (%)	166 (37.9)	1065 (29.3)	<0.01	96 (43.0)	70 (32.6)	<0.001	9 (41.0)	36 (45.0)	123 (36.3)	<0.01
Extrusions, n (%)	5 (1.1)	48 (1.3)	0.97	4 (1.8)	1(0.5)	0.33	0 (0)	2 (2.6)	3 (0.9)	0.52
Disc degeneration (mean ±SD)	3.8± 0.7	3.7±0.7	<0.001	3.8±0.7	3.6 ± 0.7	<0.001	3.7±0.8	3.9±0.7	3.8±0.6	<0.01
HIZ affected vertebral body adjacent to the end plate (total endplates;4070)										
Modic type I, n (%)	24 (5.5)	176 (4.9)	0.29	13 (5.8)	11 (5.1)	0.32	0 (0)	7 (8.8)	17 (5.0)	0.18
Modic type II, n (%)	122 (27.9)	779 (21.4)	<0.01	75 (33.6)	47 (21.9)	<0.001	7 (31.8)	22 (27.5)	94 (27.7)	<0.05
Modic type III, n (%)	14 (3.2)	88 (2.4)	0.18	13 (5.8)	1 (0.5)	<0.0001	0 (0)	3 (3.8)	11 (3.2)	0.4
Schmorl's node, n (%)	101 (23.1)	707 (19.5)	0.075	50 (22.4)	51 (23.7)	0.19	3 (13.6)	21 (26.3)	74 (21.8)	0.75

Pearson χ test and ANOVA (analysis of variance) with within group Tukey post-hoc tests were used to assess the association between HIZ and no HIZ, between posterior HIZ and anterior HIZ, and among T1W low-, high-, and iso- intensity type of HIZ where applicable. Non-paired student t-test was performed to compare continuous Pfirrmann grade at HIZ affected disc level. High-intensity zones (HIZ), T1W: T1-weighted, SD: standard deviation, %: percentage, n: number of subjects.

doi:10.1371/journal.pone.0160111.t003

associated with HIZ at the affected vertebral body adjacent to the end plate (27.9% vs. 21.4%, $p<0.01$).

Posterior HIZ had more bulges/protrusions (43.0% vs. 32.6%, $p<0.001$) and DD (median: 3.8, SD: ± 0.7 vs. 3.6, SD: ± 0.7, $p<0.001$) than anterior HIZ. Modic type II change was more significantly associated with posterior HIZ at each affected vertebral body (33.6% vs. 21.9%, $p<0.001$), Modic type III change was in comparison more significantly associated with posterior HIZ (5.8% vs. 0.5%, $p<0.0001$).

When comparing T1W low-intensity, T1W high-intensity and T1W iso-intensity types of HIZ, T1W low-intensity and high- intensity types of HIZ had more bulges/protrusions as compared with T1W iso-intensity type of HIZ (41.0% vs. 46.8% vs. 36.0% $p<0.01$) and DD (median: high 3.7, SD: ± 0.8 vs. low 3.9, SD: ± 0.7, vs iso 3.8, SD: ± 0.6, $p<0.01$). Modic type II change was significantly associated with T1W low-intensity type of HIZ than T1W iso- and high intensity types of HIZ (31.8% vs. 28.6% vs. 27.4%, $p<0.05$). (Table 3)

Discussion

Our large-scale population-based study presents a novel classification scheme of HIZ based upon evaluation of the morphology, topography, and the relationship of T1W and T2W MRI signal changes of HIZ. This classification is more precise and comprehensive than what has been traditionally reported and can be utilized for any future analysis regarding phenotype association and clinical relevance. Furthermore, to our knowledge, this study is also the first to address HIZ and their association of the other MRI spinal phenotypes based on both T1W and T2W MRI.

Since the original description of the HIZ on T2W sagittal MRI in 1992 [8], the prevalence of HIZ has varied greatly between reported studies in spite of the subjects with or without LBP. The prevalence of posterior HIZ was reported to be from 28.6% to 59% in symptomatic patients [8–11, 13] as compared to 3.2% to 24% in asymptomatic subjects [13–16]. Our large-scale population study in comparison showed that the prevalence of posterior HIZ was 21.9% (179/814 subjects). We also found posterior HIZ to be most common at L5/S1 (13.0%) followed by L4/L5 (9.3%), which was supported by a few studies [8, 10, 15]. However, we also report anterior HIZ to commonly occur at L3/L4 (10.8%) followed by L2/L3 (5.8%). This finding underscores the fact that region-specific variations of HIZ exists within the lumbar spine, with distinction between the upper (i.e. L1-L4) and lower (i.e. L4-S1) lumbar discs. Recent studies have noted more of a developmental origin or predisposition of upper lumbar segment phenotypes [46]. Nonetheless, the fact that HIZ is frequently found to be at the anterior of the disc is contrary to the traditional belief that HIZ must be posterior [8–18]. Hence, the lack of standardization for classifying HIZ for including anterior HIZ may be a likely reason for the discrepancies in the current literature regarding the reported prevalence.

Provocation discography has been utilized for assessment of annular tears and LBP [17, 18, 47, 48]. However, discography remains controversial due to the associated risks. For example, the procedure is invasive and complications include infection (epidural abscess, discitis), neurological injury, and possible contrast medium reaction [48, 49]. There is also the possibility of increased progression of DD and herniation after the examination [48, 49]. Therefore, to allow for future non-invasive HIZ research, Yu *et al* [17] reported the sensitivity of HIZ to diagnose annular tears on MRI with discography and cadavers and concluded that HIZ demonstrated some radial tears of annulus in 1989. With our more thorough MRI study with advanced sequences and imaging technique, our findings and classification of morphological/topographical variants of HIZ will further enhance our understanding of the pathology of intervertebral disc disorder. This allows us to have a more sensitive and non-invasive method of identifying symptomatic disc levels, predicting disc changes, and potential use for patient selection for disc regenerative therapies. This also has potential to be a marker for identifying patients at risk for adjacent segment degeneration/disease in relation to a fusion or arthroplasty procedure.

Various proposals have been put forward to explain the discrepancy between the presence of HIZs in asymptomatic and symptomatic individuals [8–16]. Six years after the initial paper [8], Bogduk postulated that annular tears may be present in asymptomatic subjects as low-intensity zones on T2W MRI, and these may become painful and assume a brighter signal to become an "activated" HIZ [50]. Indeed, the present study is in concurrence as we did not find low intensity zones on T2W MRI. Bogduk also reported an inability to detect HIZ on T1W MRI [8]. However, this is disputed in our study as we observed variable intensity types of HIZ on T1W MRI. Hence, we believe that coupling of T2W and T1W MRI sequences is necessary to define the HIZ phenotype. HIZ has been defined as collections of mucoid fluid within the annulus tear and thus have a bright signal on T2W MRI in the pathological studies [8, 10, 14]. However, HIZ may also change and represent a reflection in the pathological process, which may convert from one type to another, for example, neovascularization of annulus, a healing annular tear, and fluid or mucoid material filled in the inflamed annular tear. These processes may express as different signals on T1W MRI.

We found in this study significant associations between the presence of HIZ and DD, disc bulge/protrusion and Modic type II changes at all affected levels. These results support the view that degenerative findings and HIZ co-exist. Some investigators have suggested that HIZ was a part of the degenerative process as HIZs occurred in association with degenerative changes within the disc [9,10,13], whereas others disagreed [14]. This discrepancy is partly explained by the sample population, presence or not of symptoms and how clinical parameters

are defined, small sample size, and/or insufficient statistical analyses. However this large-scale, population-based study identified a strong association between HIZ with DD and disc bulge/protrusion. We also found that Modic Type II changes were more associated with the presence of HIZ, especially posterior HIZ, T1W low-intensity type of HIZ. In addition, Modic type III change was more associated with posterior HIZ than anterior HIZ. These relationships are understandable and can be attributed to the altered biomechanics associated with endplate failure caused by HIZ or as a reverse causality of Modic changes leading to HIZ. Furthermore, Schmidt *et al* [51] showed that HIZ was associated with instability of the intervertebral disc which caused fluid to move through annular tear into the outer annulus [15]. Subsequently, the unstable motion of intervertebral disc increased the stress and strain at adjacent disc segments, leading to Modic change [52]. Thus, HIZ and its sub-phenotypes may have potential as imaging biomarkers to identify those patients at risk for DD, instability of disc, and adjacent segment degeneration/disease. In general, studies have noted that Modic changes are highly associated with LBP; however, different degrees of pain severity and disability may exist [4–6]. There are also subjects with Modic changes and no HIZ. As such, being able to identify clinically relevant HIZ associated with Modic changes may shed additional light into identifying more problematic disc levels.

These results of our study may be influenced by the high age groups of our cohort (mean age over 60 years); thus, additional study is necessary to further assess HIZ among different age strata. Moreover, as with all population-based studies, there may be an effect of ethnic variability that should be addressed in future studies [53]. In addition, due to the availability of scanning units at the initiation of our study, we utilized a mobile 1.5 T MRI unit to facilitate the assessment of our subjects. Although a higher field strength, such as 3T MRI, may theoretically have a higher sensitivity in detecting specific HIZs; there have been no studies that have addressed such a concern to date to gauge the extent of the variation and it was not an aim of our current study. However, it is also important to consider that all subjects in our current study were assessed via the 1.5T MRI, representing a consistency in assessment. Our work raises awareness of the variation of HIZs that may exist in the lumbar spine and we hope will form the much needed foundation for future studies to explore upon this research platform to a much greater extent. Finally, the current study did not address an association of HIZ with LBP due to the limited pain profile assessment available in the cohort. Importantly, the strength of the present study is the size of the study population and the novel in-depth multi-parametric phenotype profiling on MRI that could serve as the basis for future HIZ study and phenotype standardization in the future. Such a foundation can then be utilized to assess more in-depth clinical relevance and utility.

Conclusions

This is the first large-scale, population-based study to systematically assess the epidemiology of HIZ on 1.5T MRI and report upon a novel classification of this phenotype in the lumbar spine. In addition, this study is also the first to utilize a multi-parametric imaging approach to assess the different variants of HIZ by the use of T2W and T1W MRI. Hence, with such alternative imaging in mind, it may be appropriate in the future to not refer to the HIZ phenotype as representing "high" intensity zones but rather "intensity zones". Such a nomenclature may be more apropos given that some HIZ on T1W MRI are not "high" intensity. Although HIZ is frequently found to be posterior, as traditionally believed, they do occur anteriorly in the disc, and numerous morphological variants exist that are disc-level and region-specific, and distinguishable via a multi-parametric imaging approach. Furthermore, HIZ are highly associated with specific MRI spinal phenotypes, such as DD, disc bulges/protrusions, and Modic changes. In

an age whereby various "omics" approaches and large data set cohorts are becoming more commonplace, a standardized phenotype classification of HIZ is imperative. Such a scheme can be further utilized to assess the clinical profile of patients, identify problematic discs, prognosticate outcomes and help tailor specific spine treatments. Additional, large-scale, comparative prospective studies are needed to further validate our findings and address their clinical impact.

Acknowledgments

The authors wish to thank Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, and other members of the Public Office in T Town for their assistance in the location and scheduling of participants for examinations.

Author Contributions

Conceived and designed the experiments: DS MY.

Performed the experiments: MT HH HY SM HO RK HI ST TA NY.

Analyzed the data: MT DS.

Contributed reagents/materials/analysis tools: MT DS JC.

Wrote the paper: MT DS.

Responsible for critical revisions: DS JC. Administrative support: KN KMC HK.

References

1. Samartzis D, Borthakur A, Belfer I, Bow CH, Lotz JC, Wang H-Q, et al. Novel diagnostic and therapeutic methods for intervertebral disc degeneration and low back pain. Spine J 2015; 15 (9): 1919–32.
2. Luk KDK, Samartzis D. Intervertebral disc "dysgeneration". Spine J 2015; 15(9):1915–8. doi: 10.1016/j.spinee.2014.07.020 PMID: 26303177
3. Wang HQ, Samartzis D. Clarifying the nomenclature of intervertebral disc degeneration and displacement: from bench to bedside. Int J Clin Exp Pathol 2014; 7 (4): 1293–8. PMID: 24817926
4. Määttä JH, Karppinen JI, Luk KD, Cheung KM, Samartzis D. Phenotype profiling of Modic changes of the lumbar spine and its association with other MRI phenotypes: a large-scale population-based study. Spine J 2015; 15(9):1933–42. doi: 10.1016/j.spinee.2015.06.056 PMID: 26133258
5. Määttä JH, Karppinen J, Paananen M, Bow C, Luk KDK, Cheung KMC, et al. Refined phenotyping of Modic changes: potential imaging biomarkers of prolonged severe low back pain and disability. Medicine (in Press).
6. Mok F, Samartzis D, Karppinen J, Fong DYT, Luk KD, Cheung KM. Modic changes of the lumbar spine: prevalence, risk factors and association with disc degeneration and low back pain in a large-scale population-based cohort. Spine J 2016; 16(1):32–41. doi: 10.1016/j.spinee.2015.09.060 PMID: 26456851
7. Samartzis D, Mok FPS, Karppinen J, Fong DYT, Luk KD, K., et al. Classification of Schmorl's nodes of the lumbar spine and association with disc degeneration: a large-scale population-based MRI study. Osteoarthritis Cartilage (In Press).
8. Aprill C, Bogduk N. High-intensity zone: A diagnostic sign of painful lumbar disc on magnetic resonance imaging. Bri J Radiol 1992; 65 (773): 361–9. PMID: 1535257
9. Chen JY, Ding Y, Lv RY, Liu QY, Huang JB, Yang ZH, et al. Correlation between MR imaging and discography with provocative concordant pain in patients with low back pain. Clin J Pain 2011; 27 (2):125–30. PMID: 21268300
10. Lam KS, Carlin D, Mulholland RC. Lumbar disc high-intensity zone: the value and significance of provocative discography in the determination of the discogenic pain source. Eur Spine J 2000; 9 (1): 36–41. PMID: 10766075
11. Schellhas KP, Pollei SR, Gundry CR, Heithoff KB. Lumbar disc high-intensity zone: Correlation of magnetic resonance imaging and discography. Spine 1996; 21(1): 79–86. PMID: 9122767

12. Rankine JJ, Gill KP, Hutchinson CE, Ross ER, Williamson JB. The clinical significance of the high-intensity zone on lumbar spine magnetic resonance imaging. Spine 1999; 24 (18): 1913–9. PMID: 10515016

13. Carragee EJ, Paragioudakis SJ, Khurana S. 2000 Volvo Award winner in clinical studies: Lumbar high-intensity zone and discography in subjects without low back problems. Spine 2000; 25 (23): 2987–92. PMID: 11145809

14. Ricketson R, Simmons JW, Hauser BO. The prolapsed intervertebral disc. The high-intensity zone with discography correlation. Spine 1996; 21 (23): 2758–62. PMID: 8979322

15. Park KW, Song KS, Chung JY, Choi JM, Lee JH, Lee CK, et al. High-intensity zone on L-spine MRI: Clinical relevance and association with trauma history. Asian Spine Journal 2007; 1(1):38–42. doi: 10.4184/asj.2007.1.1.38 PMID: 20411151

16. Takatalo J, Karppinen J, Niinimäki J, Taimela S, Mutanen P, Sequeiros RB, et al. Association of Modic changes, Schmorl's nodes, spondylolytic defects, high-Intensity zone lesions, disc herniations, and radial tears with low back symptom severity among young Finnish adults. Spine 2012; 37(14):1231–9. doi: 10.1097/BRS.0b013e3182443855 PMID: 22166927

17. Yu SW, Haughton VM, Sether LA, Wagner M. Comparison of MR and discography in detecting radial tears of the annulus: A postmortem study. AJNR Am J Neuroradiol 1989; 10 (5): 1077–81. PMID: 2505523

18. Osti OL, Vernon-Roberts B, Moore R, Fraser RD. Annular tears and disc degeneration in the lumbar spine. A post-mortem study of 135 discs. J Bone Joint Surg Br 1992; 74 (5): 678–82. PMID: 1388173

19. Sharma A, Pilgram T, Wippold FJ II. Association between annular tears and disk degeneration: a longitudinal study. AJNR Am J Neuroradiol 2009; 30(3):500–6. doi: 10.3174/ajnr.A1411 PMID: 19147713

20. Määttä JH, Karppinen JI, Luk KD, Cheung KM, Samartzis D. Phenotype profiling of Modic changes of the lumbar spine and its association with other MRI phenotypes: a large-scale population-based study. Spine J 2015; 15(9):1933–42. doi: 10.1016/j.spinee.2015.06.056 PMID: 26133258

21. Määttä JH, Kraatari M, Wolber L, Niinimaki J, Wadge S, Karppinen J, et al. Vertebral endplate change as a feature of intervertebral disc degeneration: a heritability study. Eur Spine J 2014; 23(9):1856–62. doi: 10.1007/s00586-014-3333-8 PMID: 24828957

22. Määttä JH, Wadge S, MacGregor A, Karppinen J, Williams FM. ISSLS Prize Winner: Vertebral Endplate (Modic) Change is an Independent Risk Factor for Episodes of Severe and Disabling Low Back Pain. Spine 2015; 40(15):1187–93. doi: 10.1097/BRS.0000000000000937 PMID: 25893353

23. Määttä JH, Karppinen J, Paananen M, Bow C, Luk KDK, Cheung KMC, et al. Refined phenotyping of Modic changes: potential imaging biomarkers of prolonged severe low back pain and disability. Medicine (in Press).

24. Martinez-Quinones JV, Aso-Escario J, Gonzalez-Garcia L, Consolini F, Arregui-Calvo R. Are Modic Changes Able to Help us in Our Clinical Practice? A Study of the Modic Changes in Young Adults During Working Age. J Spinal Disord Tech 2014. (in Press) PMID: 25340321

25. Wang Y, Videman T, Battie MC. Modic changes: prevalence, distribution patterns, and association with age in white men. Spine J 2012; 12(5):411–6. doi: 10.1016/j.spinee.2012.03.026 PMID: 22515998

26. Dudli S, Fields AJ, Samartzis D, Karppinen J, Lotz JC. Pathobiology of Modic changes. Eur Spine J 2016 (in Press). PMID: 26914098

27. Mok F, Samartzis D, Karppinen J, Fong DYT, Luk KD, Cheung KM. Modic changes of the lumbar spine: prevalence, risk factors and association with disc degeneration and low back pain in a large-scale population-based cohort. Spine J 2016; 16(1):32–41. doi: 10.1016/j.spinee.2015.09.060 PMID: 26456851

28. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, et al. Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study. Osteoarthritis Cartilage 2014; 22(1):104–10. doi: 10.1016/j.joca.2013.10.019 PMID: 24239943

29. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, et al. Metabolic syndrome components are associated with intervertebral disc degeneration: the Wakayama Spine Study. PLoS One 11: e0147565. doi: 10.1371/journal.pone.0147565 PMID: 26840834

30. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Oka H, et al. The association of combination of disc degeneration, end plate signal change, and Schmorl node with low back pain in a large population study: the Wakayama Spine Study. Spine J 2015; 15(4):622–8. doi: 10.1016/j.spinee.2014.11.012 PMID: 25433277

31. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Associations between radiographic lumbar spinal stenosis and clinical symptoms in the general population: the Wakayama

Spine Study. Osteoarthritis Cartilage 2013; 21(6):783–8. doi: 10.1016/j.joca.2013.02.656 PMID: 23473979

32. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012; 20(10):1103–8. doi: 10.1016/j.joca.2012.06.018 PMID: 22796511

33. Nagata K, Yoshimura N, Hashizume H, Muraki S, Ishimoto Y, Yamada H, et al. The prevalence of cervical myelopathy among subjects with narrow cervical spinal canal in a population-based magnetic resonance imaging study: the Wakayama Spine Study. Spine J 2014; 14(12):2811–7. doi: 10.1016/j.spinee.2014.03.051 PMID: 24709229

34. Nagata K, Yoshimura N, Muraki S, Hashizume H, Ishimoto Y, Yamada H, et al. Prevalence of Cervical Cord Compression and Its Association With Physical Performance in a Population-Based Cohort in Japan: The Wakayama Spine Study. Spine 2012; 37(22):1892–8. doi: 10.1097/BRS.0b013e31825a2619 PMID: 22565382

35. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on osteoarthritis/osteoporosis against disability (ROAD) study. Int J Epidemiol 2010; 39(4):988–95. doi: 10.1093/ije/dyp276 PMID: 19749026

36. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 2009; 27(5):620–8. doi: 10.1007/s00774-009-0080-8 PMID: 19568689

37. Muraki S, Akune T, Oka H, Ishimoto Y, Nagata K, Yoshida M, et al. Incidence and risk factors for radiographic lumbar spondylosis and lower back pain in Japanese men and women: the ROAD study. Osteoarthritis Cartilage 2012; 20(7):712–8. doi: 10.1016/j.joca.2012.03.009 PMID: 22484574

38. Muraki S, Oka H, Akune T, Mabuchi A, En-yo Y, Yoshida M, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: The ROAD study. Osteoarthritis Cartilage 2009; 17(9):1137–43. doi: 10.1016/j.joca.2009.04.005 PMID: 19410032

39. Vangeneugden T, Laenen A, Geys H, Renard D, Molenberghs G. Applying concepts of generalizability theory on clinical trial data to investigate sources of variation and their impact on reliability. Biometrics 2005; 61 (1):295–304. PMID: 15737106

40. Vavken P, Ganal-Antonio AKB, Shen FH, Chapman J, Samartzis D. Fundamentals of clinical outcomes assessment for spinal disorders: study designs, methodologies, and analyses. Global Spine J 2015; 5 (2): 156–64. doi: 10.1055/s-0035-1547525 PMID: 25844291

41. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine 2001; 26 (17): 1873–8. PMID: 11568697

42. Cheung KM, Karppinen J, Chan D, Ho DW, Song YQ, Sham P, et al. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine 2009; 34(9):934–40. doi: 10.1097/BRS.0b013e3181a01b3f PMID: 19532001

43. Samartzis D, Karppinen J, Mok F, Fong DY, Luk KD, Cheung KM. A population-based study of juvenile disc degeneration and its association with overweight and obesity, low back pain, and diminished functional status. J Bone Joint Surg Am 2011; 93(7):662–70. doi: 10.2106/JBJS.I.01568 PMID: 21471420

44. Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disc disease: assessment of changes in vertebral body marrow with MR imaging. Radiology 1988; 166 (1 Pt 1):193–9. PMID: 3336678

45. Mok FP, Samartzis D, Karppinen J, Luk KD, Fong DY, Cheung KM. ISSLS prize winner: prevalence, determinants, and association of Schmorl nodes of the lumbar spine with disc degeneration: a population-based study of 2449 individuals. Spine 2010; 35(21):1944–52. doi: 10.1097/BRS.0b013e3181d534f3 PMID: 20838277

46. Li Y, Samartzis D, Campbell D, Cherny S, Cheung KMC, Luk KD, K., et al. Two subtypes of intervertebral disc degeneration distinguished by a large-scale population-based study. Spine J (In Press).

47. Hirsch C, Schajowicz F. Studies on structural changes in the lumbar annulus fibrosus. Acta Orthop Scand 1952; 22 (1–4):184–231. PMID: 13079741

48. Carragee EJ, Don AS, Hurwitz EL, Cuellar JM, Carrino J, Herzong R. 2009 ISSLS Prize Winner: Does discography cause accelerated progression of degeneration changes in the lumbar disc; A ten- year matched cohort study. Spine 2009; 21 (34): 2338–45. doi: 10.1097/BRS.0b013e3181ab5432 PMID: 19755936

49. Carragee EJ, Barcohana B, Alamin T, van den Haak E. Prospective Controlled Study of the Development of Lower Back Pain in Previously Asymptomatic Subjects Undergoing Experimental Discography. Spine 2004; 29 (19): 1112–7. PMID: 15131439

50. Bogduk N. Point of View. Spine 2008; 33:1298.

51. Schmidt TA, An HS, Lim TH, Nowicki BH, Haughton VM. The stiffness of lumbar spinmotion segments with a high- intensity zone in the annulus fibrosus. Spine 23 (20):2167–73. PMID: 9802156

52. Hayashi T, Daubs MD, Suzuki A, Scott TP, Phan KH, Ruangchainikom M, et al. Motion characteristics and related factors of Modic changes in the lumbar spine. J Neurosurg Spine 2015; 22 (5): 511–7. doi: 10.3171/2014.10.SPINE14496 PMID: 25700242

53. Williams R, Cheung JP, Goss B, Rajasekaran S, Kawaguchi Y, Acharya S, et al. An International Multicenter Study Assessing the Role of Ethnicity on Variation of Lumbar Facet Joint Orientation and the Occurrence of Degenerative Spondylolisthesis in Asia Pacific: A Study from the AOSpine Asia Pacific Research Collaboration Consortium. Global Spine J. 2016 Feb; 6(1):35–45. doi: 10.1055/s-0035-1555655 Epub 2015 Jul 16. PMID: 26835200

Nakamura et al. BMC Geriatrics (2016) 16:166
DOI 10.1186/s12877-016-0339-6

BMC Geriatrics

RESEARCH ARTICLE Open Access

Locomotive syndrome is associated with body composition and cardiometabolic disorders in elderly Japanese women

Misa Nakamura[1*], Yosuke Kobashi[2], Hiroshi Hashizume[3], Hiroyuki Oka[4], Ryohei Kono[5], Sachiko Nomura[5], Akihiro Maeno[6], Munehito Yoshida[3] and Hirotoshi Utsunomiya[5]

Abstract

Background: A concept referred to as locomotive syndrome (LS) was proposed by the Japanese Orthopaedic Association in order to help identify middle-aged and older adults who may be at high risk of requiring healthcare services because of problems associated with locomotion. Cardiometabolic disorders, including obesity, hypertension, diabetes, and dyslipidemia, have a high prevalence worldwide. The purpose of this study was to determine the associations between LS and both body composition and cardiometabolic disorders.

Methods: The study participants were 165 healthy adult Japanese women volunteers living in rural areas. LS was defined as a score ≥16 on the 25-question Geriatric Locomotive Function Scale (GLFS-25). Height, body weight, body fat percentage, body mass index (BMI), and bone status were measured. Bone status was evaluated by quantitative ultrasound (i.e., the speed of sound [SOS] of the calcaneus) and was expressed as the percent of Young Adult Mean of the SOS (%YAM). Comorbid conditions of hypertension, hyperlipidemia, and diabetes were assessed using self-report questionnaires.

Results: Twenty-nine participants (17.6 %) were classed as having LS. The LS group was older, shorter, and had a higher body fat percentage, a higher BMI, and lower bone status than the non-LS group. Multiple logistic regression analysis showed that participants with a BMI ≥23.5 kg/m^2 had a significantly higher risk for LS than those with a BMI <23.5 kg/m^2 (odds ratio [OR] = 3.78, $p < 0.01$). Furthermore, GLFS-25 scores were higher in participants with than those without hypertension, diabetes, or obesity, and significantly increased with the number of present disorders.

Conclusions: These findings suggest that BMI may be a useful screening tool for LS. Furthermore, because hypertension and diabetes were associated with LS, the prevention of these disorders accompanied by weight management may help protect against LS.

Keywords: Body composition, Locomotive syndrome, Bone mass index, Cardiometabolic disorders

Background

Japan is rapidly transforming into a super-aged society. The Japanese Statistics Bureau reported that as of 2015, individuals aged 65 years or older comprised 26.2 % of the Japanese population [1]. Parallel with this transformation is an increase in the incidence of health issues such as stroke, senility, dementia, falls, fractures, and joint disorders, and in turn, the number of individuals requiring nursing care [2].

Maintaining a healthy locomotive system, which includes the bone, cartilage, muscle, and nervous systems, is the foundation of increased disability-free life expectancy. It follows that, from a public health perspective, preventing the deterioration of motor function is an issue that requires urgent attention. Therefore, an epidemiological concept referred to as locomotive syndrome (LS) [2, 3] has been proposed by the Japanese Orthopaedic Association (JOA). LS primarily affects elderly individuals who currently require nursing care services owing to problems involving the locomotive system or those who have risk conditions that will likely necessitate such services in the

* Correspondence: nakamuram@kawasakigakuen.ac.jp
[1]Department of Rehabilitation, Osaka Kawasaki Rehabilitation University, 158 Mizuma, Kaizuka, Osaka 597-0104, Japan
Full list of author information is available at the end of the article

future [4]. LS is caused by the reduced muscle strength and balance associated with aging and locomotive pathologies including osteoporosis, osteoarthritis (OA), and sarcopenia [2]. In females, LS may also be caused by the decreasing levels of physical activity and bone mineral density (BMD) that tend to occur after menopause. The incidence of LS increases with age, and is significantly higher in women (35.6 %) than in men (21.2 %) [5, 6]. As beneficial locomotive exercises for the prevention of LS, the JOA recommends performing "half-squats" and "unipedal standing balance exercises with open eyes" [3].

The 25-question Geriatric Locomotive Function Scale (GLFS-25), a quantitative and evidence-based screening tool, can be used to identify individuals with LS [7, 8]. A previous study reported finding a Spearman's correlation coefficient of 0.85 ($p < 0.001$) for the association between GLFS-25 scores and the European Quality of Life 5 Dimensions Index (EQ-5D), indicating that the GLFS-25 had excellent concurrent validity [7].

Identifying factors associated with the development of LS is crucial for its prevention. Results from a number of recent studies suggest that GLFS-25 scores strongly correlate with several physical performance measures, including the unipedal standing balance and Timed Up And Go tests [5, 9–13]. However, only a few reports [5, 14] have focused on the association between the development of LS and body composition, even though numerous studies have reported that weight, body fat percentage, and BMD are associated with cardiovascular disease, various cancers, osteoporosis, hypertension, diabetes and hyperlipidemia [15–21].

An association was recently reported between abdominal obesity and LS in elderly Japanese women, suggesting that waist circumference may be useful measure to assess the risk for LS [22].

Metabolic syndrome comprises a combination of medical disorders, including increased fasting plasma glucose, abdominal obesity, high triglyceride levels, and high blood pressure, that increase the risk of developing metabolic conditions and cardiovascular disease [23]. In conjunction with metabolic syndrome, obesity, hypertension, diabetes, and dyslipidemia, which are known as the "deadly quartet," have a high prevalence worldwide [24]. Many studies have reported that physical activity and body components are associated with metabolic syndrome [17–21]. The proportion of the Japanese population with LS (47 million) is estimated to be more than twice that with metabolic syndrome (20 million) [25, 26].

In the present study, we evaluated body composition using body mass index (BMI), body fat percentage, and bone status, and hypothesized that these variables would be predictive of GLFS-25 scores. Confirmation of this hypothesis would indicate that, in addition to sports performance training, control of body composition could be used to prevent LS. Therefore, the purpose of this study was to determine the association between body composition and LS, the threshold values of body composition measures for discriminating between individuals with and without LS, and the OR of LS according to body composition above or below these thresholds in elderly Japanese women living in rural areas. An additional objective was to determine the association between LS and cardiometabolic disorders.

Methods

Participants

This study was conducted in a rural area (Tanabe city, Wakayama Prefecture, Japan) between January 2013 and March 2015. The study inclusion criteria were as follows: 1) Japanese women, age > 60 years; 2) ability to walk independently; and 3) living at home and being capable of self-care. All participants initially underwent body composition measurements, in the order of height, weight, and body fat percentage, followed by an evaluation of LS status at a public hall where a "Lecture meeting and checkup for health" was held with support from the local government. Afterwards, 198 women were asked to complete a self-report questionnaire at home regarding comorbid conditions and then to return the questionnaire by mail. A stamped envelope was provided to encourage the return of the questionnaire. A total of 165 women underwent the measurements and returned the questionnaire (mean age ± standard deviation, 68.8 ± 6.1 years; range, 60–83 years).

Measurement of variables

Body weight and body fat percentage were measured using the KaradaScan Body Composition Monitor with Scale (HBF-362; Omron Co., Kyoto, Japan) while participants were wearing normal indoor clothing. The procedure was performed with the participants standing barefoot on a metal surface in accordance with the manufacturer's instructions. BMI was calculated as weight divided by the square of height, and obesity was defined as BMI ≥25 kg/m^2 in accordance with the guidelines of the Japan Society for the Study of Obesity [27]. Bone status was assessed using speed of sound (SOS) measured using a quantitative ultrasound (QUS) device (Canon Life Care Solutions Inc., CM-200, Osaka, Japan) at the calcaneus of the dominant foot while the participants were barefoot and seated. QUS, which has a number of advantages, including portability, low cost, and a lack of exposure to radiation [28], enables the evaluation of bone quality, especially the microarchitecture at the calcaneus, and is useful for predicting risk for future fracture [29]. Bone status was shown as the percent of Young Adult Mean of the SOS (%YAM) [30]. It has been reported that calcaneal QUS parameters reflect the characteristics of the trochanteric area of the proximal

hip, although these values are not specifically reflective of values of the femoral neck or shaft [31].

Evaluation of LS

LS status (presence and degree) was evaluated based on GLFS-25 scores (Additional file 1). The GLFS-25 is a self-report questionnaire composed of the following 25 items focusing on the month before completing the measure: four questions on pain; 16 on activities of daily living; three on social function; and two on mental health status [7]. These 25 items are scored from 0 (no impairment) to 4 (severe impairment), with a total score range from 0 to 100. Higher scores indicate worse locomotive function. The cutoff score for LS, as determined by receiver operating characteristic (ROC) analysis, is 16 points [7].

Comorbid conditions

Comorbid conditions of hypertension, hyperlipidemia, and diabetes were assessed using the following question on the self-report questionnaire: "Do you presently take medication for hypertension, diabetes, or hyperlipidemia?"

Statistical analysis

Participants were classified as LS (≥16) or non-LS (<16) based on GLFS-25 scores, and then independent variables were compared between groups. For numerical variables, normality of distribution and homogeneity of variance were tested before across-group comparisons. When the assumptions of normal distribution and homogeneity of variance were met in both groups, we performed the student t-test, and when the assumption of normal distribution was met, but not the assumption of homogeneity of variance, we performed Welch's t-test. When the data were non-normally distributed, the Wilcoxon signed-rank test was used.

ROC analysis was used to evaluate the threshold of each body composition measure (BMI, body fat percentage, and %YAM) in order to discriminate the LS from the non-LS group. An area under the ROC (AUC-ROC) curve of 1.00 was taken to indicate perfect discrimination, whereas an AUC-ROC of 0.50 was taken to indicate the complete absence of discrimination.

Multiple logistic regression analysis was performed to evaluate the age-adjusted significance of the prevalence of LS. The chi-square test was used for comparison of prevalence or number of cardiometabolic disorders between non-LS and LS. The Wilcoxon signed-rank test was used for comparison of GLFS-25 scores classified by with and without cardiometabolic disorders, as well as by the number of present disorders. Statistical analysis was conducted using JMP 11 (SAS Institute, Cary, NC). All statistical tests were 2-tailed, and a significance level of 0.05 was used.

Results

Age, height, body weight, body fat percentage, BMI, bone status, GLFS-25 score, and the prevalence of components of cardiometabolic disorder are shown in Table 1. Twenty-nine participants (17.5 %) had a GLFS-25 score ≥16 and were thereby classified as LS (Table 2). The LS group was older and shorter than the non-LS group, and had a higher body fat percentage, a higher BMI, and lower bone status (Table 2).

ROC analysis was conducted for each body composition measure, and the threshold for discriminating the non-LS and LS groups was identified. This threshold was 37.3 % for body fat percentage, 23.5 kg/m^2 for BMI, and 73 % for %YAM (Table 3). ORs for the prevalence of LS according to the threshold values are shown in Table 4. High BMI was a significant risk factor for LS, with an OR of 3.78 as determined by multiple logistic regression analysis.

Figure 1 shows GLFS-25 scores classified by the presence or absence of each metabolic syndrome component (hypertension, diabetes, hyperlipidemia, and obesity). GLFS-25 scores were higher in participants with than without hypertension or diabetes, and in obese than in non-obese participants. Figure 2 shows GLFS-25 scores classified by the number of present cardiometabolic disorders. The results showed that GLFS-25 scores significantly increased with the number of cardiometabolic disorders ($p < 0.01$). Table 5 shows a comparison of the prevalence or number of cardiometabolic disorders between non-LS and LS subjects. The prevalence of LS was higher in participants with than without hypertension ($p < 0.05$) and obesity ($p < 0.01$).

Discussion

Association between body composition and LS

LS was proposed by the JOA in 2007 in order to identify individuals at high risk of requiring nursing care owing

Table 1 Characteristics of the study participants

Variables for components	Mean (SD[a])
Age (years)	68.8 (6.1)
Height (cm)	150.4 (11.9)
Weight (kg)	52.9 (8.3)
Body fat percentage (%)	33.9 (4.5)
BMI (kg/m^2)	23.1 (3.6)
%YAM[b] (%)	69.8 (11.0)
GLFS-25 score (points)	10.0 (10.3)
Components of cardiometabolic disorders	Prevalence (%)
Obesity	45 (27.3)
Hypertension	59 (35.8)
Diabetes mellitus	12 (7.3)
Hyperlipidemia	23 (13.9)

[a]*SD* standard deviation
[b]*YAM* percent of Young Adult Mean of the speed of sound of the calcaneus

Table 2 Comparison of characteristics between non-locomotive and locomotive syndrome[a]

Variables	Non-LS[b] (*n* = 136)	LS[c] (*n* = 29)	*p* value
Age (years)	68.1(5.9)	72.1(6.0)	0.0014[e]
Height (cm)	151.9(5.0)	143.7(25.5)	0.0015[e]
Weight (kg)	52.3(8.3)	55.4(8.3)	0.0730[f]
Body fat percentage (%)	33.4(4.3)	36.3(4.6)	0.0020[e]
BMI (kg/m^2)	22.7(3.1)	25.2(3.7)	0.0007[e]
%YAM[d] (%)	70.6(11.4)	65.7(8.3)	0.0288[f]

[a]Locomotive Syndrome: GLFS-25 score ≥16 points
[b]*Non-LS* non-locomotive syndrome
[c]*LS* locomotive syndrome
[d]*YAM* percent of Young Adult Mean of the speed of sound of the calcaneus
[e]Wilcoxon signed-rank test was applied for age, height, body fat percentage, and BMI
[f]Student's t-test was applied for weight and YAM

to problems associated with the locomotive system [2]. The GLFS-25 was subsequently developed to measure the presence and degree of LS in Japanese individuals [7]. However, since its implementation, the GLFS-25 cutoff value for identifying individuals with LS has been determined in accordance with health-related quality of life [7, 11]; therefore, information on the association between GLFS-25 scores and body composition is limited. Therefore, the primary purpose of this study was to determine the association between LS as defined by GLFS-25 scores and body composition measures in elderly Japanese women.

Our results showed that participants with LS were shorter, had a higher body fat percentage, a higher BMI, and lower bone status than participants without LS. Previous studies have reported similar results in middle-aged and elderly Japanese women [5, 14]. Muramoto et al. found that GLFS-25 scores had a significant positive correlation with body fat percentage and BMI, a negative correlation with body height and BMD, and no correlation with body weight according to correlation analysis [5].

Based on comparative analysis, participants with LS have been shown to have significantly greater BMI and body fat percentage and lower height than those without LS, whereas no significant difference has been observed in body weight or BMD [14].

Table 3 Threshold values of age and body composition for locomotive syndrome

	Threshold values	AUC[a]	Sensitivity (%)	Specificity (%)
Body fat percentage (%)	37.3	0.68	51.72	68.13
BMI (kg/m^2)	23.5	0.70	72.41	67.29
%YAM[b] (%)	73.0	0.61	86.21	79.23

[a]*AUC* area under the curve
[b]*YAM* percent of Young Adult Mean of the speed of sound of the calcaneus

Table 4 Evaluation of odds ratios for locomotive syndrome according to body composition

	Above or below the threshold value	Odds ratio (95 % CI[a])	*P* value
Body fat percentage (%)	<37.3	1	0.3584
	≥37.3	1.62 (0.58–5.00)	
BMI (kg/m^2)	<23.5	1	0.0087
	≥23.5	3.78 (1.39–11.07)	
%YAM[b] (%)	<73	1.68 (0.65–4.73)	0.2900
	≥73	1	

[a]*CI* confidence interval
[b]*YAM* percent of Young Adult Mean of the speed of sound of the calcaneus
Data were adjusted by age

In the present study, we found that participants with LS were shorter than those without LS. Shorter height has been reported to be significantly associated with fear of falling in elderly Japanese individuals [32]. Shorter height may be caused by an age-related change in the curvature of the spine or atrophy of trunk extension muscles, which can decrease postural control. A reduction in postural control can cause fear of falling or a decline in the amount of physical activity [33]. Therefore, we propose that the LS group included more participants that had lost height due to a change in the curvature of the spine or atrophy of trunk extension muscles than the non-LS group, and therefore had less postural control and engaged in fewer activities in daily life, which increased their risk for developing LS.

The present results showed that participants with LS had a higher body fat percentage than those without LS. Increased body fat causes more mechanical stress in weight-bearing joints and promotes the degeneration of joint tissue through the production and release of adipokines [34]. Adipokines are derived from adipocytes and may upregulate receptor activators of nuclear kappa B ligand, leading to increased bone resorption and reduced BMD [35]. Participants with a higher body fat percentage may have secreted more adipocytes, and this may have had a negative influence on the movement of the joints, thereby increasing the risk of LS.

The present study showed that a BMI ≥23.5 k/m^2 was significantly associated with LS, with an OR of 3.78 as identified by multiple logistic regression analysis. The Japan Society for the Study of Obesity defines the cutoff for obesity as a BMI 25 kg/m^2 [8]. In the present study, the mean BMI of the participants with LS was ≥25 kg/m^2. Furthermore, GLFS-25 scores were higher in obese than in non-obese participants (Fig. 2d). LS is closely associated with age-related skeletal disorders such as osteoporosis, OA, lumbar spinal stenosis (LSS), degenerative spinal disease and sarcopenia [4]. Furthermore, obesity is a risk factor for these disorders because mechanical overload on weight-bearing joints can activate chondrocytes, accelerate the degeneration of cartilage, and increase static compressive loading and

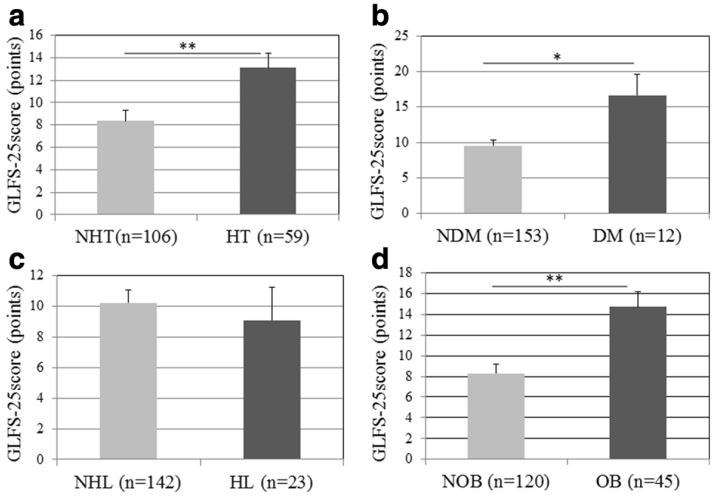


Fig. 1 GLFS-25 scores in participants with and without hypertension (**a**), diabetes (**b**), hyperlipidemia (**c**), and obesity (**d**). HT, hypertension; NHT, non-hypertension; DM, diabetes mellitus; NDM, non-diabetes mellitus; HL, hyperlipidemia; NHL, non-hyperlipidemia; OB, obesity (BMI ≥25 kg/m^2); NOB, non-obesity (BMI <25 kg/m^2); *$p < 0.05$; **$p < 0.01$; error bars, standard deviation

pressures associated with postures that damage disc integrity [36–38]. Moreover, it has been proposed that metabolic factors, including inflamed adipose tissue, dyslipidemia, oxidative stress, endothelial dysfunction and leptin dysregulation, as well as the clustering of these factors in metabolic syndrome, may play a crucial role in obesity-induced OA [39–41]. These findings support the present results regarding the association between obesity and LS.

Association between LS and cardiometabolic disorders

Obesity, hypertension, diabetes, and dyslipidemia are known as the "deadly quartet" [24]. Numerous studies have reported that BMD is associated with these disorders [17–21]. The second purpose of the present study was to determine the association between LS and cardiometabolic disorders. Our results showed that LS is associated with hypertension, diabetes, and overweight, as well as with higher BMI. Furthermore, GLFS-25 scores significantly increased with the number of present cardiometabolic disorders.

There are some reports on the association between cardiometabolic disorders and OA. Some evidence suggests

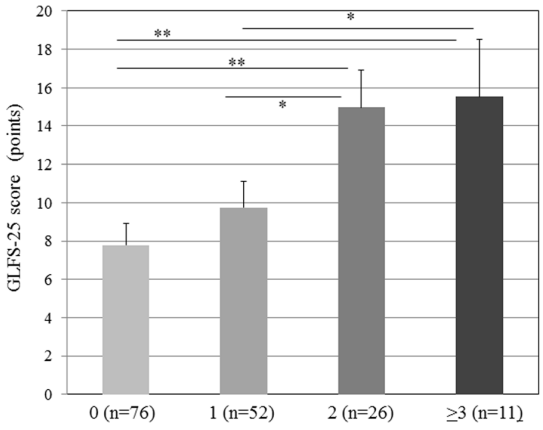


Fig. 2 GLFS-25 scores in participants with 0, 1, 2, and ≥3 present cardiometabolic disorders. A significant main effect was observed for the number of disorders ($p < 0.01$; analysis of variance). *$p < 0.05$; **$p < 0.01$; error bars, standard deviation

Table 5 Comparison of prevalence or number of present cardiometabolic disorders between non-locomotive and locomotive syndrome[a]

		Non-LS[b]	LS[c]	*P* value
NHT[d] (153)		128 (83.7 %)	25 (16.3 %)	0.0482
HT[e] (12)		8 (66.7 %)	4 (33.3 %)	
NDM[f] (142)		116 (81.7 %)	26 (18.3 %)	0.1364
DM[g] (23)		20 (87.0 %)	3 (13.0 %)	
NHL[h] (106)		92 (86.8 %)	14 (13.2 %)	0.5382
HL[i] (59)		44 (74.6 %)	15 (25.4 %)	
NOB[j] (120)		106 (88.3 %)	14 (11.7 %)	0.0011
OB[k] (45)		30 (66.7 %)	15 (33.3 %)	
Number of cardiometabolic disorders	0 (76)	67 (88.2 %)	9 (11.8 %)	0.0526
	1 (52)	44 (84.6 %)	8 (15.4 %)	
	2 (26)	18 (69.2 %)	8 (30.8 %)	
	≥3 (11)	7 (63.6 %)	4 (36.4 %)	

[a]Locomotive Syndrome: GLFS-25 score ≥16 points; [b]*Non-LS* non-locomotive syndrome, [c]*LS* locomotive syndrome, [d]*NHT* non-hypertension, [e]*HT* hypertension, [f]*NDM* non-diabetes, [g]*DM* diabetes, [h]*NHL* non-hyperlipidemia, [i]*HL* hyperlipidemia, [j]*NOB* non-obesity (BMI <25 kg/m^2), [k]*OB* obesity (BMI ≥25 kg/m^2)

that metabolic factors such as type 2 diabetes mellitus and elevated glucose concentration are associated with the development and progression of OA [40, 42]. In particular, the advanced glycation end products in cartilage collagen seem to be associated with both the senescent cartilage matrix and reduced chondrocyte function [43]. The presence of advanced glycation end products associated with the expression of advanced glycation end-product receptors in the cartilage collagen results in the increased production of matrix metalloproteinase and the modulation of the chondrocyte phenotype to hypertrophy and OA [44, 45].

OA and hypertension have been shown to frequently coexist [46]. The proposed mechanism of the development of OA with hypertension is as follows: narrow and/or constricted vessels restrict blood flow to subchondral bone, impairing circulation and nutritional supply to overlying articular cartilage, which ultimately contributes to the deterioration of cartilage in OA [47].

Mutual relations exist between the occurrence and presence of musculoskeletal diseases, particularly knee OA and cardiometabolic disorders [48]. Yoshimura et al. suggested that metabolic risk factors such as overweight, hypertension, hyperlipidemia, and impaired glucose tolerance increase the risk of occurrence and progression of knee OA [49, 50]. Recent reports have indicated that waist circumference, back muscle strength, and spinal inclination angle are important risk factors for LS [22]. In the present study, we demonstrated that LS is associated with hypertension and obesity, as well as a higher BMI. Furthermore, GLFS-25 scores significantly increased with the number of present cardiometabolic disorders. These findings suggest a close relationship between the locomotive system and cardiometabolic organs.

The proportion of adults with BMI >25 kg/m^2 has significantly increased worldwide [51]. The present findings contribute to the identification of factors that may prevent locomotive disorder and metabolic syndrome, particular in Western societies, in which many patients have metabolic syndrome. Although the concept of LS is currently used only in Japan, we believe it will become more common worldwide as the population continues to age.

The results of the present study suggest that BMI might be a useful measure for the simple detection of LS. Furthermore, hypertension and diabetes were found to be associated with LS. Weight management and prevention of these disorders may help protect against LS in elderly women. Elderly men should be included in future studies.

Limitations and future research

This study did have several limitations. First, the sample size of 165 was small; this number only represents about 1.2 % of all women aged between 60 and 83 years in Tanabe city. Furthermore, no significant relationship was found between LS and dyslipidemia; this may have been due in part to a lack of statistical power. Second, because the participants in this study were all Japanese women, care should be taken in generalizing the results to men or other ethnic groups. Third, data from a cross-sectional study are not sufficient to determine whether a causal relationship exists between BMI, LS, and cardiometabolic disorders. LS may cause obesity or hypertension and diabetes because it limits physical activity. Conversely, these cardiometabolic disorders may lead to the development of LS. It is therefore crucial to perform longitudinal studies to clarify the causal relationships among these factors. Fourth, comorbid conditions were only assessed using self-report questionnaires; therefore, blood pressure, blood glucose concentration, and blood lipid concentration measurements were not controlled. Thus, untreated participants with comorbid conditions may have been excluded from analysis; however, this possibility is low because participants attending the "Lecture meeting" would have been expected to have relatively high health awareness. Fifth, further research in larger-sized studies should measure lean mass because it is an important component of BMI. It is possible that BMI underestimates body fat percentage in clinical populations [52].

Conclusion

BMI, body fat percentage, and bone status were significantly associated with LS. In particular, a BMI ≥23.5 k/m^2 was significantly associated with LS. Moreover, GLFS-25 scores were higher in participants with a BMI ≥25 kg/m^2, hypertension, and diabetes than in the respective comparison groups. These results suggest that BMI is an important measure for the detection of LS. Furthermore, weight management and the prevention of metabolic syndrome may reduce the risk for LS.

Additional file

Additional file 1: The 25-question Geriatric Locomotive Function Scale [7]. (DOCX 154 kb)

Abbreviations

%YAM: Percent of Young Adult Mean of the SOS; AUC: area under the curve; BMD: Bone mineral density; BMI: Body mass index; GLFS-25: 25-question Geriatric Locomotive Function Scale; JOA: Japanese Orthopaedic Association; LS: Locomotive syndrome; OA: Osteoarthritis; OR: odds ratio; QUS: qualitative ultrasound; ROC: Receiver operating characteristic; SOS: speed of sound

Acknowledgements

We would like to thank Dr. Hiroshi Kameda, Mr. Shota Okumi, Mr. Nobuhiro Koike, Mr. Daiki Kanata, Mr. Sho Tachibana, Mr. Shodai Tanaka, Ms. Sakiko Enomoto, Mr. Takuma Nishimae, Mr. Kenta Higashi, Mr. Yusuke Saeki, Mr. Ryosuke Hashikaku, Mr. Taichi Takemoto, and Mr. Yoshiki Kushi at Osaka Kawasaki Rehabilitation University for their cooperation.

Funding

None.

Availability of data and materials

The datasets analyzed in the current study are available from the corresponding author on reasonable request.

Authors' contributions
MN participated in the study design, performed the statistical analysis, and drafted the manuscript; YK provided assistance in the statistical analysis; RK, SN, AM, and HU performed measurements of variables; HH, HO, and MY provided assistance in the literature review and revised the manuscript; All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
All subjects provided written informed consent to use their data in the study. The study protocol was approved by the Ethics Committee of Wakayama Medical University (Reference No. 1005). This study was performed in accordance with the Declaration of Helsinki.

Author details
[1]Department of Rehabilitation, Osaka Kawasaki Rehabilitation University, 158 Mizuma, Kaizuka, Osaka 597-0104, Japan. [2]Department of Rehabilitation Medicine, Wakayama Medical University, 811-1 Kimiidera, Wakayama, Wakayama 641-8510, Japan. [3]Department of Orthopedic Surgery, Wakayama Medical University, 811-1 Kimiidera, Wakayama 641-8510, Japan. [4]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan. [5]Department of Strategic Surveillance for Functional Food and Comprehensive Traditional Medicine, Wakayama Medical University, Kimiidera 811-1, Wakayama, Wakayama 641-0012, Japan. [6]Laboratory of Chemistry, Kansai Medical University, 2-5-1 Shinnmachi, Hirakata 573-1010, Japan.

Received: 18 November 2015 Accepted: 19 September 2016
Published online: 27 September 2016

References

1. Statistics Bureau. Population estimates (March 2015). http://www.e-stat.go.jp/SG1/estat/ListE.do?lid=000001131732. Accessed 22 Sept 2016.
2. Nakamura K. A "super-aged" society and the "locomotive syndrome". J Orthop Sci. 2008;3(1):1–2. doi:10.1007/s00776-007-1202-6.
3. Japanese Orthopaedic Association. Guidebook on locomotive syndrome. Tokyo: Bunkodo; 2010 (in Japanese).
4. Nakamura K. The concept and treatment of locomotive syndrome: its acceptance and spread in Japan. J Orthop Sci. 2011;16(5):489–91. doi:10.1007/s00776-007-1202-6.
5. Muramoto A, Imagama S, Ito Z, Hirano K, Ishiguro N, Hasegawa Y. Physical performance tests are useful for evaluating and monitoring the severity of locomotive syndrome. J Orthop Sci. 2012;17(6):782–8. doi:10.1007/s00776-013-0382-5.
6. Sasaki E, Ishibashi Y, Tsuda E, Ono A, Yamamoto Y, Inoue R, et al. Evaluation of locomotive disability using loco-check: a cross-sectional study in the Japanese general population. J Orthop Sci. 2013;17(6):782–8. doi:10.1007/s00776-012-0329-2.
7. Seichi A, Hoshino Y, Doi T, Akai M, Tobimatsu Y, Iwaya T. Development of a screening tool for risk of locomotive syndrome in the elderly: the 25-question Geriatric Locomotive Function Scale. J Orthop Sci. 2012;17(6):782–8. doi:10.1007/s00776-011-0193-5.
8. Japanese Orthopaedic Association 2013. Locomotive syndrome pamphlet. P. 08. https://locomo-joa.jp/en/index.pdf. Accessed 22 Sept 2016.
9. Yoshimura N, Oka H, Muraki S, Akune T, Hirabayashi N, Matsuda S, et al. Reference values for hand grip strength, muscle mass, walking time, and one-leg standing time as indices for locomotive syndrome and associated disability: the second survey of the ROAD study. J Orthop Sci. 2011;16(6):768–77. doi:10.1007/s00776-011-0160-1.
10. Hirano K, Hasegawa Y, Imagama S, Wakao N, Muramoto A, Ishiguro N. Impact of spinal imbalance and back muscle strength on locomotive syndrome in community-living elderly people. J Orthop Sci. 2012;17(5):532–7. doi:10.1007/s00776-012-0266-0.
11. Seichi A, Hoshino Y, Doi T, Akai M, Tobimatsu Y, Kita K, et al. Determination of the optimal cutoff time to use when screening elderly people for locomotive syndrome using the one-leg standing test (with eyes open). J Orthop Sci. 2014;19(4):620–6. doi:10.1007/s00776-014-0581-8.
12. Fukumori N, Yamamoto Y, Takegami M, Yamazaki S, Onishi Y, Sekiguchi M, et al. Association between hand-grip strength and depressive symptoms: locomotive syndrome and health outcomes in Aizu cohort study (LOHAS). Age Ageing. 2015;44(4):592–8. doi:10.1093/ageing/afv013.
13. Nakamura M, Hashizume H, Oka H, Okada M, Takakura R, Hisari A, et al. Physical performance measures associated with locomotive syndrome in middle-aged and older Japanese women. J Geriatr Phys Ther. 2015;38(4):202–7. doi:10.1519/JPT.0000000000000033.
14. Muramoto A, Imagama S, Ito Z, Hirano K, Tauchi R, Ishiguro N, et al. Threshold values of physical performance tests for locomotive syndrome. J Orthop Sci. 2013;18(4):618–26. doi:10.1007/s00776-013-0382-5.
15. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health. 2009;9:88. doi:10.1186/1471-2458-9-88.
16. Bang E, Tanabe K, Yokoyama N, Chijiki S, Kuno S. Relationship between thigh intermuscular adipose tissue accumulation and number of metabolic syndrome risk factors in middle-aged and older Japanese adults. Exp Gerontol. 2016;16(79):26–30. doi:10.1016/j.exger.2016.03.010.
17. Hwang DK, Choi HJ. The relationship between low bone mass and metabolic syndrome in Korean women. Osteoporos Int. 2010;21:425–31. doi:10.1007/s00198-009-0990-2.
18. Kim HY, Choe JW, Kim HK, Bae SJ, Kim BJ, Lee SH, et al. Negative association between metabolic syndrome and bone mineral density in Koreans, especially in men. Calcif Tissue Int. 2010;86(5):350–8. doi:10.1007/s00223-010-9347-2.
19. Xue P, Gao P, Li Y. The association between metabolic syndrome and bone mineral density: a meta-analysis. Endocrine. 2012;42(3):546–54. doi:10.1007/s12020-012-9684-1.
20. Muka T, Trajanoska K, Kiefte-de Jong JC, Oei L, Uitterlinden AG, Hofman A, et al. The association between metabolic syndrome, bone mineral density, hip bone geometry and fracture risk: the Rotterdam study. PLoS One. 2015;10(6):e0129116. doi:10.1371/journal.pone.0129116.
21. Ebron K, Andersen CJ, Aguilar D, Blesso CN, Barona J, Dugan CE, et al. A larger body mass index is associated with increased atherogenic dyslipidemia, insulin resistance, and low-grade inflammation in individuals with metabolic syndrome. Metab Syndr Relat Disord. 2015;13(10):458–64. doi:10.1089/met.2015.0053.
22. Muramoto A, Imagama S, Ito Z, Hirano K, Tauchi R, Ishiguro N, et al. Waist circumference is associated with locomotive syndrome in elderly females. J Orthop Sci. 2014;19(4):612–9. doi:10.1007/s00776-013-0382-5.
23. Alberti KG, Zimmet P, Shaw J. Metabolic syndrome–a new world-wide definition. A consensus statement from the International Diabetes Federation. Diabet Med. 2006;23:469–80.
24. Kaplan NM. The deadly quartet. Upper-body obesity, glucose intolerance, hypertriglyceridemia, and hypertension. Arch Intern Med. 1989;149:1514–20.
25. Yoshimura N, Muraki S, Oka H, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab. 2009;27:620–8.
26. Ministry of Health Labour and Welfare. Measures for national health Promotion, Ref-70. http://www.mhlw.go.jp/english/wp/wp-hw2/part2/p3_0024.pdf. Accessed 22 Sept 2016.
27. Examination Committee of Criteria for 'Obesity Disease' in Japan; Japan Society for the Study of Obesity. New criteria for 'obesity disease' in Japan. Circ J. 2002;66:987–92.
28. Kohri T, Kaba N, Murakami T, Narukawa T, Yamamoto S, Sakai T, Sasaki S. Search for promotion factors of ultrasound bone measurement in Japanese males and pre/post-menarcheal females aged 8–14 years. J Nutr Sci Vitaminol (Tokyo). 2012;58(4):263–71. 10.3177/jnsv.58.263.
29. Chin KY, Ima-Nirwana S. Calcaneal quantitative ultrasound as a determinant of bone health status: what properties of bone does it reflect? J Med Sci. 2013;10(12):1778–83. doi:10.7150/ijms.6765.
30. Iizuka Y, Iizuka H, Mieda T, Tajika T, Yamamoto A, Takagishi K. Population-based study of the association of osteoporosis and chronic musculoskeletal pain and locomotive syndrome: the Katashina study. J Orthop Sci. 2015;20(6):1085–9. doi:10.1007/s00776-015-0774-9.
31. Zhang L, Lv H, Zheng H, Li M, Yin P, Peng Y, Gao Y, Zhang L, Tang P. Correlation between parameters of calcaneal quantitative ultrasound and

hip structural analysis in osteoporotic fracture patients. PLoS One. 2015; 10(12):e0145879. doi:10.1371/journal.pone.0145879.
32. Nishimura A, Ikezoe T, Kitase S, et al. The factor influences fear of falling in elderly people. Phys Ther Kyoto. 2006;35:98–9 (in Japanese).
33. Ogaya S, Ikezoe T, Tateuchi H, et al. The relationship of fear of falling and daily activity to postural control in the elderly. Phys Ther Sci. 2010;37:78–84 (in Japanese).
34. Conde J, Scotece M, López V, Gómez R, Lago F, Pino J, et al. Adipokines: novel players in rheumatic diseases. Discov Med. 2013;15:73–83.
35. Hofbauer LC, Schoppet M. Clinical implications of the osteoprotegerin/ RANKL/RANK system for bone and vascular diseases. JAMA. 2004;292:490–5.
36. Halade GV, El Jamali A, Williams PJ, Fajardo RJ, Fernandes G. Obesity-mediated inflammatory microenvironment stimulates osteoclastogenesis and bone loss in mice. Exp Gerontol. 2011;46(1):43–52. doi:10.1016/j.exger.2010.09.014.
37. Knutsson B, Sandén B, Sjödén G, Järvholm B, Michaëlsson K. Body mass index and risk for clinical lumbar spinal stenosis: a cohort study. Spine (Phila Pa 1976). 2015;40(18):1451–6. doi:10.1097/BRS.0000000000001038.
38. Ou CY, Lee TC, Lee TH, Huang YH. Impact of body mass index on adjacent segment disease after lumbar fusion for degenerative spine disease. Neurosurgery. 2015;76(4):396–401. doi:10.1227/NEU.0000000000000627.
39. Flamme CH. Obesity and low back pain–biology, biomechanics and epidemiology. Orthopade. 2005;34(7):652–7.
40. Musumeci G, Aiello FC, Szychlinska MA, Di Rosa M, Castrogiovanni P, Mobasheri A. Osteoarthritis in the XXIst century: risk factors and behaviours that influence disease onset and progression. Int J Mol Sci. 2015;16(3):6093–112. doi:10.3390/ijms16036093.
41. Thijssen E, van Caam A, van der Kraan PM. Obesity and osteoarthritis, more than just wear and tear: Pivotal roles for inflamed adipose tissue and dyslipidaemia in obesity-induced osteoarthritis. Rheumatology (Oxford). 2014;54(4):588–600. doi:10.1093/rheumatology/keu464.
42. Zhuo Q, Yang W, Chen J, Wang Y. Metabolic syndrome meets osteoarthritis. Nat Rev Rheumatol. 2012;8:729–37. doi:10.1038/nrrheum.2012.135.
43. Verzijl N, DeGroot J, Ben ZC, Brau-Benjamin O, Maroudas A, Bank RA. Crosslinking by advanced glycation end products increases the stiffness of the collagen network in human articular cartilage: a possible mechanism through which age is a risk factor for osteoarthritis. Arthritis Rheum. 2002;46:114–23.
44. Yammani RR, Carlson CS, Bresnick AR, Loeser RF. Increase in production of matrix metalloproteinase 13 by human articular chondrocytes due to stimulation with S100A4: role of the receptor for advanced glycation end products. Arthritis Rheum. 2006;54:2901–11.
45. Cecil DL, Johnson K, Rediske J, Lotz M, Schmidt AM, Terkeltaub R. Inflammation-induced chondrocyte hypertrophy is driven by receptor for advanced glycation end products. J Immunol. 2005;175:8296–302.
46. Singh G, Miller JD, Lee FH, Pettitt D, Russell MW. Prevalence of cardiovascular disease risk factors among US adults with self-reported osteoarthritis: data from the Third National Health and Nutrition Examination Survey. Am J Manag Care. 2002;8(15 Suppl):S383–91.
47. van den Berg WB. Osteoarthritis year 2010. In review: pathomechanisms. Osteoarthritis Cartilage. 2011;19(4):338–41. doi:10.1016/j.joca.2011.01.022.
48. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, et al. Mutual associations among musculoskeletal diseases and metabolic syndrome components: a 3-year follow-up of the ROAD study. Mod Rheumatol. 2015;25(3):438–48. doi:10.3109/14397595.2014.972607.
49. Yoshimura N, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, et al. Accumulation of metabolic risk factors such as overweight, hypertension, dyslipidaemia, and impaired glucose tolerance raises the risk of occurrence and progression of knee osteoarthritis: a 3-year follow-up of the ROAD study. Osteoarthritis Cartilage. 2012;20(11):1217–26. doi:10.1016/j.joca.2012.06.006.
50. Yoshimura N, Muraki S, Oka H, Tanaka S, Ogata T, Kawaguchi H, et al. Association of knee osteoarthritis with the accumulation of metabolic risk factors such as overweight, hypertension, dyslipidemia, and impaired glucose tolerance in Japanese men and women: the ROAD study. J Rheumatol. 2011;16(6):768–77. doi:10.1007/s00776-015-0741-5.
51. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2014;384(9945):766–81. doi:10.1016/S0140-6736(14)60460-8.
52. Shah NR, Braverman ER. Measuring adiposity in patients: the utility of body mass index (BMI), percent body fat, and leptin. PLoS One. 2012;7(4):e33308. doi:10.1371/journal.pone.0033308.

General Hospital Psychiatry 45 (2017) 7–11

Contents lists available at ScienceDirect

General Hospital Psychiatry

journal homepage: http://www.ghpjournal.com

Development of a Japanese version of the Somatic Symptom Scale-8: Psychometric validity and internal consistency

Ko Matsudaira [a,*], Hiroyuki Oka [a], Mika Kawaguchi [b], Masato Murakami [c], Shin Fukudo [d,e], Makoto Hashizume [f], Bernd Löwe [g,h]

[a] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan
[b] Clinical Study Support, Inc., Daiei Bldg., 2F, 1-11-20 Nishiki, Naka-ku, Nagoya 460-0003, Japan
[c] Department of Psychosomatic Internal Medicine, Sanno Hospital, International University of Health and Welfare, 8-10-16 Akasaka, Minato-ku, Tokyo 107-0052, Japan
[d] Department of Behavioral Medicine, Tohoku University Graduate School of Medicine, 2-1 Seiryo, Aoba, Sendai 980-8575, Japan
[e] Department of Psychosomatic Medicine, Tohoku University Hospital, 1-1 Seiryo, Aoba, Sendai 980-8574, Japan
[f] Hashizume Clinic, 101, 3-4-5 Miyakojimanakadori, Miyakojima, Osaka 534-0022, Japan
[g] Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, Gebäude O25, 20246 Hamburg, Germany
[h] Department of Psychosomatic Medicine and Psychotherapy, Schön Clinic Hamburg Eilbek, Dehnhaide 120, 22081 Hamburg, Germany

ARTICLE INFO

Article history:
Received 18 May 2016
Revised 15 November 2016
Accepted 2 December 2016
Available online xxxx

Keywords:
Japanese version
Psychometric validation
Reliability
Somatic symptom burden
Somatic Symptom Scale-8

ABSTRACT

Objective: We aimed to psychometrically validate the Japanese version of the Somatic Symptom Scale-8 (SSS-8) in Japanese individuals.
Method: Data were collected from Japanese individuals aged 20–64 years, who were recruited online, in February 2015. The scale reliability and validity were analyzed.
Results: Data from 52,353 individuals were analyzed. Cronbach's alpha for the assessment of internal consistency reliability was 0.86 for the total score. The concurrent validity results showed strong correlations with three domains of the Profile of Mood States-Brief form ($\rho = 0.51$–0.61) and the EuroQol 5 Dimension ($\rho = -0.54$). The known-group validity results indicated a linear trend in the severity of depression stratified using the Patient Health Questionnaire-2 (Jonckheere-Terpstra test, $p < 0.001$). Regarding convergent and discriminant validities, all items correlated most strongly with their own domains (coefficients ≥ 0.5), except for one item (headaches). Scores on perceived stress, pain, and general health differed across five SSS-8 severity groups (Steel-Dwass test, $p < 0.001$), expect for one group pair in health.
Conclusion: The Japanese version of the SSS-8 was valid with good internal consistency. This questionnaire could help detect somatic symptom burdens of chronic and severe musculoskeletal pain for primary prevention.

1. Introduction

Somatic symptoms are generally considered manifestations of an underlying psychiatric illness, such as anxiety, depression, or common mental disorders [1]. Common symptoms include various types of pain (e.g., back pain, joint pain, headache), gastrointestinal symptoms (e.g., food intolerance, regurgitation of food, bloating), cardiopulmonary symptoms (e.g., sweating, palpitation, breathlessness), and excessive tiredness [2]. Somatic symptoms are associated with deterioration of quality of life and psychological distress and increased use of health care services [3,4].

The Somatic Symptom Scale-8 (SSS-8) is a self-administered questionnaire assessing somatic symptom burden [5]. The SSS-8 consists of eight items that assess the following symptoms: stomach or bowel problems; back pain; pain in your arms, legs, or joints; headaches; chest pain or shortness of breath; dizziness; feeling tired or having low energy; and trouble sleeping. These items comprise the four symptom domains of gastrointestinal, pain, cardiopulmonary, and fatigue. Respondents rate how much each symptom has bothered them during the previous 7 days and score each item from 0 to 4: not at all (0), a little bit (1), somewhat (2), quite a bit (3), and very much (4), with no reverse-coded items included. The total score, ranging from 0 to 32, is a simple sum of each item score: a higher score indicates more severe somatic symptom burden.

The SSS-8 was originally developed in English as an abbreviated version of the Patient Health Questionnaire-15 (PHQ-15) [4], a questionnaire used worldwide to assess the presence and severity of somatic symptoms [6–11]. The PHQ-15 was used as a reference measure in the Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) (DSM-5) field trials to facilitate the diagnosis of somatic symptom disorder [12]. The German version of the SSS-8 has been linguistically and

* Corresponding author at: Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8654, Japan.
E-mail address: kohart801@gmail.com (K. Matsudaira).

http://dx.doi.org/10.1016/j.genhosppsych.2016.12.002
0163-8343/

psychometrically validated and has shown good reliability and validity for the general German population [5] as well as psychometric equivalence to the PHQ-15 [13]. Whereas the PHQ-15 has been used internationally, the SSS-8 is newly developed but is expected to be a useful tool in busy medical settings because it requires little time to complete and score.

To make the SSS-8 available in Japan, we translated the English version into Japanese and conducted a linguistic validation of the scale [14]. To ensure conceptual equivalence between the original and its translation, the translation and linguistic validation were conducted as follows: (i) forward-translation by two independent Japanese translators; (ii) back translation by a native English speaker; and (iii) pilot testing for comprehension in five patients with a history of musculoskeletal pain and somatic symptoms. Through a step-by-step process, a linguistically validated Japanese version of the SSS-8 was successfully developed, yet its psychometric properties have not been assessed.

Thus, the objective of the present study was to psychometrically validate the Japanese version of the SSS-8 for assessment of somatic symptoms in Japanese individuals.

2. Methods

2.1. Study population

Members of the Japanese general population aged 20–64 years were included in the analysis. Participants were recruited by an Internet research company, United Inc. (Tokyo, Japan), with which >1.37 million individuals across Japan have voluntarily registered. Out of approximately 1.25 million individuals aged 20–64 years selected as eligible participants, 270,000 individuals were randomly selected and invited by e-mail to complete an online questionnaire.

Participation in the online surveys was voluntary and no personally identifiable information (e.g., name and address) was collected. All participants gave their consent and were compensated. After obtaining ethical approval from The University of Tokyo, the questionnaire administration was conducted in February 2015.

2.2. Measures

The administered questionnaire included questions on demographic and clinical characteristics, the SSS-8, the EuroQol 5 Dimension (EQ-5D) [15], the brief form of the Profile of Mood States (POMS) [16,17], the Patient Health Questionnaire-2 (PHQ-2) [18], and questions on perceived stress, subjective health, and perceived general health.

The EQ-5D is an instrument developed to measure general health status [15]. It contains five questions assessing mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [19]. All responses are converted into a single index score of general health status ranging from −0.11 to 1.00: a score of 1 indicates “perfect health” and a score of 0 indicates “death.” The Japanese version of the EQ-5D, which was approved by the EuroQol Group in 1997, has been widely used in research [20].

The POMS is a 65-item questionnaire that assesses the mood of individuals based on the following six mood construct domains: tension/anxiety, depression/dejection, anger/hostility, vigor, fatigue, and confusion. The present study used the POMS-Brief form, which consists of 30 items assessing the same six domains. Each item is rated on a five-point scale, and each domain score ranges from 0 to 20, with higher scores indicating more disturbances, except for the vigor domain. A Japanese version of the POMS-Brief form was shown to be reliable and valid [21].

The PHQ-2 is a questionnaire comprising two questions extracted from the universally used original Patient Health Questionnaire-9 [18]. The questions assess whether the respondent has experienced depression and anhedonia over the past 2 weeks. Although each item is rated on a scale of 0–3 in the original PHQ-2, the present study used the National Center of Neurology and Psychiatry version of the Japanese PHQ-2, which gives each item a binary response of yes or no [22]. Individuals who answer yes to at least one question are suspected of experiencing depression and a closer assessment of the individual is recommended.

Participants' perceived stress and subjective pain (including numbness) during the past 4 weeks were rated using a numerical rating scale (NRS) ranging from 0 to 10, with higher scores indicating greater stress/pain (numbness). Participants' current perceived health was also rated using an NRS on a scale of 0 to 10; a score of 0 indicates the worst health status and a score of 10 indicates the best health status.

2.3. Statistical analysis

Participants' demographic and clinical characteristics were analyzed descriptively. For descriptive statistics of the Japanese SSS-8, the total score and individual item scores were calculated to examine missing data and floor and ceiling effects (>60%).

To assess the psychometric properties of the Japanese version of the SSS-8, we evaluated its reliability and validity. Internal consistency was assessed to evaluate the reliability of the SSS-8. The extent to which items in the SSS-8 correlated with each other was evaluated using Cronbach's alpha coefficient. A Cronbach's alpha coefficient was computed for both total and symptom domain scores. A Cronbach's alpha of 0.7 or higher would indicate that the SSS-8 is internally consistent [23].

The validity was evaluated by assessing concurrent validity, known-group validity, and convergent and discriminant validity. Concurrent validity was assessed by examining associations between the SSS-8 and external reference questionnaires (EQ-5D and POMS) using Spearman's correlation coefficient. Scales that measure similar concepts should be strongly correlated; those measuring different concepts should be weakly correlated. The correlation coefficient was interpreted according to Cohen's criteria: 0.1 is considered a weak correlation; 0.3, moderate; and 0.5, strong [24].

For known-group validity, scores among different groups of participants based on the results of the PHQ-2 were examined. It was hypothesized that participant groups with more affirmative responses would have higher SSS-8 scores. To test whether there was such a linear trend across groups with different levels of depression, the Jonckheere-Terpstra test was performed [25,26]. The Jonckheere-Terpstra test is a non-parametric test, which tests if the SSS-8 scores increase as the number of affirmative response in the PHQ-2 increases, based on a hypothesis that the response distribution does not differ by the number of affirmative response.

Convergent and discriminant validity examined whether an individual item fits in its own domain (convergent validity) while the individual item does not fit in the other domains aside from its own domain (divergent validity). To assess convergent and divergent validity, the correlations between items and the symptom domains that those items assess (gastrointestinal, pain, cardiopulmonary, and fatigue domains) were calculated. It was hypothesized that items would strongly correlate with other items assumed to belong to the same symptom domain, and would weakly correlate with items assumed to belong to different symptom domains. The item-total correlations were evaluated using Spearman's correlation coefficient.

Additionally, we assessed the relationships between the Japanese SSS-8 severity groups and the participants' clinical status. Based on the SSS-8 total score, participants were categorized into five severity groups using the German version of the severity thresholds: a score of 0–3 was categorized as “no to minimal severity,” 4–7 as “low,” 8–11 as “medium,” 12–15 as “high,” and ≥16 as “very high” [5]. Among these five severity groups, all pairwise comparisons were conducted by the Steel-Dwass non-parametric test [27,28] to identify any pairs with statistically significant differences in perceived stress, subjective pain, and perceived general health.

All statistical tests were two-sided with a significance level of 0.05. Unanswered questionnaire items were treated as missing data. All analyses were performed using SAS software version 9.3 (SAS Institute, Inc., Cary, NC, USA).

3. Results

Data from 52,353 individuals who responded to the questionnaires were analyzed. Participants' demographic and clinical characteristics are summarized in Table 1. The median age was 43 years with a range of 20 to 64 years, and 50.0% of the participants were male. Median perceived stress, subjective pain, and perceived general health scores were 4 (3–6), 2 (1–4), and 6 (5–8), respectively.

Table 2 summarizes descriptive statistics of the Japanese SSS-8. To complete the questionnaire, respondents must answer all SSS-8 items; therefore, there was no missing data. No remarkable floor or ceiling effects were observed for the total scores. No ceiling effects were observed for the individual item scores; however, floor effects were observed for the following six items: stomach or bowel problems; pain in the arms, legs, or joints; headaches; chest pain or shortness of breath, dizziness; and trouble sleeping.

For reliability, internal consistency of the Japanese SSS-8 was evaluated using Cronbach's alpha coefficients. The Cronbach's alpha for the total score was 0.86, which demonstrates good consistency. The Cronbach's alpha coefficients within each symptom domain were 0.69 for the pain domain, 0.77 for the cardiopulmonary domain, and 0.77 for the fatigue domain.

To assess concurrent validity, Spearman's correlation coefficients for the associations between the SSS-8 and the two external criteria (EQ-5D, POMS) were calculated. A strong correlation was observed with the EQ-5D ($\rho = -0.54$), and also with the three POMS domains: 0.61 for the POMS-fatigue, 0.55 for the POMS-tension/anxiety, and 0.51 for the POMS-depression/dejection ($p < 0.001$ for all). For the remaining POMS domains, moderate to weak correlations were observed with the POMS-anger/hostility ($\rho = 0.46$), POMS-confusion ($\rho = 0.46$), and POMS-vigor domains ($\rho = -0.01$) ($p < 0.001$ for all).

To examine known-group validity, SSS-8 total scores were compared among groups categorized based on their responses to the PHQ-2 items: 71.6% of patients made no affirmative responses, 14.9% made one affirmative response, and 13.6% made two affirmative responses. As hypothesized, the median SSS-8 total scores and its interquartile ranges (25th–75th percentile) were higher in the groups with more affirmative responses to the PHQ-2 items: 2 (0–5) in the group with no affirmative responses, 5 (2–9) in the group with one affirmative response, and 8 (4–12) in the group with two affirmative responses. The statistical test results showed a linear increasing trend in the SSS-8 total score across these three PHQ-2 categories (which indicate suspected depression levels) (Jonckheere-Terpstra test, $p < 0.001$).

Table 1
Participants' demographic and clinical characteristics ($n = 52353$).

Characteristics	
Age, years (median, range)	43 (20–64)
Sex, male (n, %)	26,191 (50.0)
Educational qualification (n, %)	
Junior high school	1293 (2.5)
High school	16,105 (30.8)
Vocational school	7105 (13.6)
Technical college	855 (1.6)
Junior college	5302 (10.1)
University	19,102 (36.5)
Graduate school	2191 (4.2)
Others	400 (0.8)
Employment status (n, %)	
Full-time employee	20,565 (39.3)
Part-time/contract employee	9945 (19.0)
Temporary staff	1783 (3.4)
Business executive	2903 (5.6)
Family business	765 (1.5)
Work at home	1267 (2.4)
Students	1861 (3.6)
Do housework	7843 (15.0)
Without job	4363 (8.3)
Others	1058 (2.0)
Perceived stress (NRS) (median, IQR)	4 (3–6)
Subjective pain (NRS) (median, IQR)	2 (1–4)
Perceived general health (NRS) (median, IQR)	6 (5–8)

Values are median (range), n (%), or median (IQR). IQR = interquartile range (25th–75th percentile); NRS = numerical rating scale (score range: 0 to 10).

To test convergent and discriminant validity, item-total correlations were examined. Table 3 shows the Spearman's correlation coefficients between each SSS-8 item and other items belonging to the same or different symptom domains (gastrointestinal, pain, cardiopulmonary, or fatigue domains). Shaded cells in Table 3 indicate correlations between each item and the other items in the same symptom domain. Boldface text indicates each item's highest correlation, to show the domain with which it was most strongly associated. All the items except for headaches (item 4) showed the highest correlation with items within their own domain, and most of them demonstrated strong correlations ($\rho \geq 0.5$) ($p < 0.001$ for all correlations). Item 4 showed the highest correlation with the cardiopulmonary domain ($\rho = 0.48$) instead of with its own pain domain ($\rho = 0.38$).

The plausibility of the Japanese SSS-8 severity groups based on the SSS-8 total score was examined by comparing the median scores on perceived stress, subjective pain, and perceived general health in each severity group (Table 4). For both perceived stress and subjective pain, median NRS scores were higher for more severe category groups, and the score differences between any pair of severity groups were significant (Steel-Dwass test, $p < 0.001$). For perceived general health, median scores were lower for more severe category groups. As with the stress and pain scores, this indicates that participants in more severe groups experience greater symptom burden. The score differences were significant between all pairs of severity groups (Steel-Dwass test, $p < 0.001$) except for between the “high” and “very high” groups (Steel-Dwass test, $p = 0.13$).

4. Discussion

This study used data collected online from 52,353 individuals to assess the psychometric properties of the Japanese SSS-8, which had been linguistically validated previously. Overall, the results demonstrated that the Japanese SSS-8 had good internal consistency, and acceptable to good concurrent validity, known-group validity, and convergent and discriminant validities.

Although the descriptive statistics of the Japanese SSS-8 revealed no ceiling or floor effects for the Japanese SSS-8 total scores, a floor effect was observed for six individual items. This is probably because the present study sampled members of the Japanese general population, who did not necessarily have any somatic symptoms, and because item scores were between 0 and 4. In fact, over 50% of the participants obtained total scores of between 0 and 5. Therefore, these floor effects were not considered critical.

The internal consistency of the Japanese SSS-8 was evaluated here using Cronbach's alpha coefficient [29]. The coefficient exceeded a generally acceptable level of 0.7 for psychometric scales and reached over 0.8, which is regarded as a good level. These levels are similar to the Cronbach's alpha of 0.81 found for the German SSS-8 [5].

The validity of the Japanese SSS-8 was evaluated by examining concurrent validity, known-group validity, and convergent and discriminant validity. The concurrent validity analysis showed strong correlations exceeding 0.5 (or −0.5) between the Japanese SSS-8 and the measures of self-reported health status, fatigue, anxiety, and depression. Known-group validity was also found: there was a statistically significant trend for patients with more depression symptoms to report higher Japanese SSS-8 scores. Both the concurrent validity and known-group validity results found here indicate relationships between

Table 2
SSS-8 total scores and distributions of individual item scores.

	Mean	SD	Median	Range Min–Max	Floor effect (%)	Ceiling effect (%)
Total score	4.5	5.2	3	0–32	24.9	0.4
1. Stomach or bowel problems	1.6	0.9	1	1–5	64.4	1.5
2. Back pain	1.8	1.0	1	1–5	53.6	2.7
3. Pain in the arms, legs, or joints	1.5	0.9	1	1–5	67.0	1.9
4. Headaches	1.6	0.9	1	1–5	66.6	2.0
5. Chest pain or shortness of breath	1.3	0.7	1	1–5	83.9	1.1
6. Dizziness	1.3	0.7	1	1–5	81.0	1.3
7. Feeling tired or having low energy	1.9	1.1	1	1–5	50.1	3.9
8. Trouble sleeping	1.6	1.0	1	1–5	67.4	3.0

SSS-8 = Somatic Symptom Scale-8; SD = standard deviation; Min = minimum; Max = maximum.

somatic burden, anxiety, and depression, which supports previous research showing that somatic, anxiety, and depression symptoms are highly comorbid and partially overlap [30,31].

For the convergent and discriminant validities, the highest correlations were between each question item and the domain to which it belonged, with the exception of the headaches item. Headaches had the highest correlation with the cardiopulmonary domain (0.48), followed by the fatigue domain (0.45). Although headaches are a type of pain, they differ from back pain and pain in your arms, legs, or joints which are classified as musculoskeletal pain. Earlier research grouped headache into the category of general symptoms or head-and-gastrointestinal symptoms along with symptoms such as dizziness and fatigue as a result of factor analysis in somatic symptoms [32–34]. In fact, when looking into correlation between each items rather domains, headaches in the present study indicated the highest correlation with dizziness ($\rho = 0.45$) followed by feeling tired or having low energy ($\rho = 0.44$). Therefore, headaches showing such stronger correlations with the cardiopulmonary and fatigue domains may be accountable. Similarly in the German version of the SSS-8, confirmatory factor analysis results revealed coefficients of between 0.61 and 0.84; the lowest was for headaches and the highest for pain in the arms, legs, or joints in the pain domain [5]. This may have resulted from differences in sampling methods; however, the observed lowest coefficient for headaches in the present study is consistent with the German findings.

Furthermore, five severity thresholds from the German SSS-8 are applicable to the Japanese general population. For the German version of the SSS-8, the severity increased as levels of perceived stress, pain, and general health increased [5]. The differences between all pairs were statistically significant, except for the pair of high and very high in perceived general health. However, as the medians in high and very high were the same for perceived stress and general health, further research is warranted to determine whether the category cutoff points for high and very high are appropriate for the Japanese general population.

There are several limitations of the present study. First, generalization of these results is limited. As recruitment was conducted online, some demographic groups may have been under-represented (e.g., those without access to the Internet) and some over-represented (e.g., those with a greater motivation to participate). In addition, the recruitment targets were limited to registered individuals between the ages of 20 and 64 years (considered to be the working age population). However, the present study obtained a large sample from the general Japanese population and this sample reflected the age and sex composition ratio of the Japanese population. Therefore, such under- or over-represented groups may not be a critical problem in the present study. Second, misclassifications of response and recall bias are concerns. Response misclassification is inevitable when using subjective measures. Recall bias toward retrospective questions might also have distorted participants' responses. Therefore, these need to be interpreted with caution. Third, the present study did not evaluate the responsiveness of the Japanese SSS-8. The ability of the questionnaire to detect changes if the condition changes (e.g., responsiveness to treatment) needs to be evaluated prior to its use in longitudinal studies. Further assessment of responsiveness is thus necessary. Fourth, as the present study targeted the general population residing in Japan, use of the Japanese SSS-8 in a clinical setting may produce results that differ from the present results. The English version of the SSS-8, a short form of the PHQ-15 [4], was originally developed for the DSM-5 field trials [12], and its German version has been psychometrically validated for the German general population, suggesting that the SSS-8 could be applicable to both clinical and general populations [5]. However, the relevance of the Japanese SSS-8 for patients in Japan needs to be demonstrated.

Table 3
Correlations[a] among each item and other items belonging to the same or different symptom domains.

	Domain			
SSS-8 Item	Gastrointestinal Item #1	Pain Item #2–4	Cardiopulmonary Item #5–6	Fatigue Item #7–8
1. Stomach or bowel problems	**1.00**	0.46	0.41	0.45
2. Back pain	0.39	**0.51**	0.38	0.44
3. Pain in your arms, legs, or joints	0.32	**0.47**	0.38	0.38
4. Headaches	0.39	0.38	**0.48**	0.45
5. Chest pain or shortness of breath	0.38	0.43	**0.50**	0.43
6. Dizziness	0.36	0.43	**0.50**	0.44
7. Feeling tired or having low energy	0.43	0.52	0.47	**0.57**
8. Trouble sleeping	0.38	0.43	0.43	**0.57**

[a]Spearman's correlation coefficient.

SSS-8 = Somatic Symptom Scale-8. Item 1 comprises thegastrointestinal symptoms domain, items 2–4 comprise the pain domain, items 5–6 comprise the cardiopulmonary domain, and items 7–8 comprise the fatigue domain. All the correlations were $p < 0.001$.

Table 4
Stress, pain, and overall health NRS scores for each SSS-8 severity category.

SSS-8 severity category (SSS-8 score)	n (%)	Perceived stress (NRS) Median (IQR)	Subjective pain (NRS) Median (IQR)	Perceived general health (NRS) Median (IQR)
No to minimal (0–3)	29,294 (56.0)	3 (2–5)	1 (0–3)	7 (5–8)
Low (4–7)	12,243 (23.4)	5 (3–7)	3 (2–5)	6 (4–7)
Medium (8–11)	5731 (10.9)	6 (4–7)	4 (3–6)	5 (4–6)
High (12–15)	2725 (5.2)	7 (5–8)	5 (3–7)	4 (3–6)
Very high (≥16)	2360 (4.5)	7 (5–8)	6 (4–7)	4 (3–6)

Score differences between severity groups were tested ($p < 0.001$ for all pairs of severity groups in perceived stress and subjective pain and $p < 0.001$ for all pairs in perceived general health, except for a pair of high and very high groups, $p = 0.13$).
SSS-8 = Somatic Symptom Scale-8; NRS = numerical rating scale; IQR = interquartile range (25th–75th percentile).

In sum, the present study demonstrated that our linguistically validated version of the Japanese SSS-8 was valid with a good internal consistency. Our results also suggested that the somatic symptom burdens determined by the SSS-8 severity thresholds were in proportion to individuals' perceptions of stress and pain levels and inverse to their perception of health status. This brief questionnaire could be useful in a medical setting and could help to detect the somatic symptom burden of chronic and severe musculoskeletal pain for primary prevention.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure

All authors have declared potential conflicts of interest as follows: KM received the research grant from Ministry of Health, Labour and Welfare for the submitted work; grant support including an endowed chair from Sumitomo Dainippon Pharma Co., Ltd. and OKAMURA CORPORATION; grant support including an endowed chair and lecture fees from AYUMI Pharmaceutical Corporation, Nippon Zoki Pharmaceutical Co., Ltd., ONO PHARMACEUTICAL CO., LTD., Eli Lilly Japan K.K., Astellas Pharma Inc., TOTO LTD., and Eisai Co., Ltd.; lecture fees from Pfizer Japan Inc., HISAMITSU PHARMACEUTICAL CO., INC., Janssen Pharmaceutical K.K., KAKEN PHARMACEUTICAL CO., LTD., TEIJIN PHARMA LIMITED; and lecture fees and advisory fees from Shionogi & Co., Ltd., outside the submitted work. HO received the research grant from JA Kyosai Research Institute, and had endowed chairs at Showa Yakuhin Kako Co., Ltd., Nippon Zoki Pharmaceutical Co., Ltd., and ONO PHARMACEUTICAL CO., LTD., outside the submitted work. SF received the research grant from Kao Corporation; grant support and lecture fees from Astellas Pharma Inc.; and lecture fees from ABBOTT JAPAN CO., LTD. and KISSEI PHARMACEUTICAL CO., LTD., outside the submitted work. These entities did not have any roles in the study design; data collection, analysis, and interpretation; manuscript writing; and/or decision to submit for publication. MK, MM, MH, and BL had no competing interest to report.

References

[1] Desai G, Chaturvedi SK, Dahale A, Marimuthu P. On somatic symptoms measurement: the scale for assessment of somatic symptoms revisited. Indian J Psychol Med 2015;37:17–9.
[2] Hiller W, Rief W, Brähler E. Somatization in the population: from mild bodily misperceptions to disabling symptoms. Soc Psychiatry Psychiatr Epidemiol 2006;41:704–12.
[3] Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. Arch Gen Psychiatry 2005; 62:903–10.
[4] Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64:258–66.
[5] Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, et al. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. JAMA Intern Med 2014;174:399–407.
[6] Han C, Pae CU, Patkar AA, Masand PS, Kim KW, Joe SH, et al. Psychometric properties of the Patient Health Questionnaire-15 (PHQ-15) for measuring the somatic symptoms of psychiatric outpatients. Psychosomatics 2009;50:580–5.
[7] Ros Montalbán S, Comas Vives A, Garcia-Garcia M. Validation of the Spanish version of the PHQ-15 questionnaire for the evaluation of physical symptoms in patients with depression and/or anxiety disorders: DEPRE-SOMA study. Actas Esp Psiquiatr 2010;38:345–57.
[8] Lee S, Ma YL, Tsang A. Psychometric properties of the Chinese 15-item patient health questionnaire in the general population of Hong Kong. J Psychosom Res 2011;71:69–73.
[9] Yazici Güleç M, Güleç H, Simşek G, Turhan M, Aydin Sünbül E. Psychometric properties of the Turkish version of the Patient Health Questionnaire - somatic, anxiety, and depressive symptoms. Compr Psychiatry 2012;53:623–9.
[10] Karekla M, Pilipenko N, Feldman J. Patient Health Questionnaire: Greek language validation and subscale factor structure. Compr Psychiatry 2012;53:1217–26.
[11] Nordin S, Palmquist E, Nordin M. Psychometric evaluation and normative data for a Swedish version of the Patient Health Questionnaire 15-Item Somatic Symptom Severity Scale. Scand J Psychol 2013;54:112–7.
[12] Narrow WE, Clarke DE, Kuramoto SJ, Kraemer HC, Kupfer DJ, Greiner L, et al. DSM-5 field trials in the United States and Canada, part III: development and reliability testing of a cross-cutting symptom assessment for DSM-5. Am J Psychiatry 2013;170:71–82.
[13] Gierk B, Kohlmann S, Toussaint A, Wahl I, Brünahl CA, Murray AM, et al. Assessing somatic symptom burden: a psychometric comparison of the patient health questionnaire-15 (PHQ-15) and the somatic symptom scale-8 (SSS-8). J Psychosom Res 2015;78:352–5.
[14] Matsudaira K, Kawaguchi M, Murakami M, Fukudo S, Hashizume M, Oka H, et al. Development of a linguistically validated Japanese version of the Somatic Symptom Scale-8 (SSS-8). Shinshin Irgaku 2016;56:931–7 (in Japanese).
[15] EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
[16] McNair DM, Lorr M. An analysis of mood in neurotics. J Abnorm Soc Psychol 1964; 69:620–7.
[17] DM MN, Lorr M, Droppleman LF. Manual for the profile of mood states. San Diego, CA: Educational and Industrial Testing Services; 1971.
[18] Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003;41:1284–92.
[19] Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ 2002;11:341–53.
[20] The Japanese EuroQol Translation Team. The development of the Japanese EuroQol Instrument. J Health Care Soc 1998;8:109–23 (in Japanese).
[21] Yokoyama K. POMS shortened version - manual and commentary on cases. Tokyo: Kaneko Syoboh; 2005(in Japanese).
[22] Muramatsu K. Usability of the Japanese version of the Patient Health Questionnaire (PHQ)-2. The 46th congress of the Japanese Society of psychosomatic medicine. Nara; 2005.
[23] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297–334.
[24] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Philadelphia: Lawrence Erlbaum Associates; 1988.
[25] Terpstra TJ. The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. Indag Math 1952;14:327–33.
[26] Jonckheere AR. A distribution-free k-sample test against ordered alternatives. Biometrika 1954;41(1/2):133–45.
[27] Steel RGD. A rank sum test for comparing all pairs of treatments. Technometrics 1960;2.2:197–207.
[28] Dwass M. Some k-sample rank-order tests. Contributions to probability and. Statistics 1960:198–202.
[29] Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2nd ed. Chichester, UK: John Wiley & Sons; 2007.
[30] Löwe B, Spitzer RL, Williams JB, Mussell M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. Gen Hosp Psychiatry 2008;30:191–9.
[31] Kroenke K, Spitzer RL, Williams JB, Löwe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. Gen Hosp Psychiatry 2010;32:345–59.
[32] Fink P, Toft T, Hansen MS, Ørnbøl E, Olesen F. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. Psychosom Med 2007;69:30–9.
[33] Rosmalen JGM, Tak LM, De Jonge P. Empirical foundations for the diagnosis of somatization: implications for DSM-5. Psychol Med 2011;41:1133–42.
[34] Tsai CH. Factor analysis of the clustering of common somatic symptoms: a preliminary study. BMC Health Serv Res 2010;10:1.

Journal of Pain Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

The relationship between findings on magnetic resonance imaging and previous history of low back pain

This article was published in the following Dove Press journal:
Journal of Pain Research
29 December 2016
Number of times this article has been viewed

Juichi Tonosu[1]
Hiroyuki Oka[2]
Ko Matsudaira[2]
Akiro Higashikawa[1]
Hiroshi Okazaki[1]
Sakae Tanaka[3]

[1]Department of Orthopedic Surgery, Kanto Rosai Hospital, Kawasaki, [2]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, [3]Department of Orthopedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

Abstract: The objective of this study was to evaluate the relationship between magnetic resonance imaging (MRI) findings and previous low back pain (LBP) in participants without current LBP. Current LBP was defined as LBP during the past month. Previous LBP was defined as a history of medical consultation for LBP. Ninety-one participants without current LBP were included. Sagittal T2-weighted MRI was used to assess the intervertebral space from T12/L1 to L5/S1. These images were classified into five grades based on the Pfirrmann grading system. Furthermore, we evaluated the presence of disk bulging, high-intensity zone, and spondylolisthesis. We compared the MRI findings between groups with (27 participants) and without (64 participants) previous LBP without current LBP. Intraobserver and interobserver kappa values were evaluated. Participants had an average age of 34.9 years; 47 were female and 44 were male; and their average body mass index was 21.8 kg/m^2. Compared to the group of participants without previous LBP, the group of participants with previous LBP had a significantly higher incidence of disk degeneration such as a Pfirrmann grade ≥3, disk bulging, and high-intensity zone in the analyses adjusted by age and sex. There were no significant differences in spondylolisthesis between the groups. An odds ratio of >10 was only found for Pfirrmann grade ≥3, ie, a Pfirrmann grade ≥3 was strongly associated with a history of previous LBP in participants without current LBP.

Keywords: disk bulging, low back pain, magnetic resonance imaging, MRI, Pfirrmann grading, previous history, high-intensity zone

Introduction

Low back pain (LBP) affects most adults at some point in their lives. Approximately 85%–90% of cases are classified as nonspecific LBP.[1] In the last decade, LBP was continuously found to be the top leading cause of years lived with disability globally.[2] Similarly, in Japan, LBP is one of the most common causes of health disability, as in other industrialized countries, with a reported lifetime prevalence of >80%.[3] Especially in the workplace, LBP is an important and costly medical problem that leads to decreased employee health and productivity.[4]

Magnetic resonance imaging (MRI) can identify underlying pathologies of LBP. However, the importance of MRI findings is unclear and controversial. Some reports have shown that disk degeneration was associated with LBP,[5–7] while others have shown that there was no relationship between disk degeneration and LBP.[8,9] Although these reports focused on the relationship between disk degeneration and current LBP, there are a few reports on the relationship between MRI findings, including disk degeneration

Correspondence: Juichi Tonosu
Department of Orthopedic Surgery, Kanto Rosai Hospital, 211-8510, 1-1, Kizukisumiyoshicho, Nakahara-ku, Kawasaki, Kanagawa, Japan
Tel +81 44 411 3131
Fax +81 44 433 3150
Email juichitohnosu@yahoo.co.jp

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/JPR.S122380

and previous LBP.[5,10] It has been suggested that symptoms of chronic LBP are often fluctuating, and this is a condition with a pattern of exacerbation and remission.[11] Some individuals have chronic LBP, whereas others have intermittent pain. We anticipate that if physicians know about the predictive MRI findings of recurrent severe LBP, we can selectively educate patients about preventing LBP. Therefore, we hypothesized that people whose lumbar MRI showed disk degeneration would be prone to developing severe LBP, unless they did not have current severe LBP. The purpose of this study was to evaluate the relationship between MRI findings and previous LBP symptoms in participants without current LBP.

Materials and methods

Study participants

From September 2005 to March 2006, we recruited volunteers who were personnel at Kanto Rosai Hospital to participate in the study. Ninety-one participants without current LBP were included. We administered a questionnaire to determine whether they had previous LBP symptoms. According to previous reports, current LBP was defined as pain localized between the costal margin and the inferior gluteal folds depicted in a diagram with or without lower extremity pain in the past 1 month.[1,12] The area was shown diagrammatically on the questionnaire according to a previous study.[12] Previous LBP was defined as a history of medical consultation for LBP. Medical consultation for LBP is one of the standards for evaluating the severity of LBP.[13] This indicated that the LBP was not mild. Then, we classified the participants into two groups, those with previous LBP and those without previous LBP. The study was approved by the review board of the Minister of Labor, Health, and Welfare of Japan. Written informed consent was obtained from all individual participants included in the study.

Image assessment

MRI was performed using a 1.5 T Siemens Symphony scanner (Siemens Healthcare, Erlangen, Germany). The imaging protocol included sagittal T2-weighted fast spin echo (repetition time: 3,500 ms/echo, echo time: 120 ms, and field of view: 300 × 320 mm). Sagittal T2-weighted images were used to assess the intervertebral space from T12/L1 to L5/S1. Assessment of the MRI scans was performed by an orthopedist (J.T.) who was blinded to the participants' backgrounds. We evaluated the degree of disk degeneration, disk bulging, the high-intensity zone (HIZ), and spondylolisthesis at each level of the spine. The degree of disk degeneration on MRI was classified into five grades based on the Pfirrmann classification system.[14] In the analysis, we divided Pfirrmann grading into two categories, grades 1–2 and grades 3–5. Disk bulging was defined as displacement of the disk material, usually by >50% of the disk circumference and <3 mm beyond the edges of the disk space in the axial plane.[15] As we were only able to evaluate the sagittal planes of MRI scans, we defined disk bulging as posterior disk displacement <3 mm and equivalent to the anterior disk displacement in the sagittal plane. We defined HIZ as an area of brightness or high signal intensity located in the posterior annulus on T2-weighted images based on previous literature.[16] We defined spondylolisthesis as vertebral slips of >5 mm. To evaluate intraobserver variability, 20 randomly selected MRI scans of the lumbar spine were rescored by the same observer (J.T.) >1 month after the first reading. Furthermore, to evaluate interobserver variability, 20 other MRI scans were scored by two orthopedists (J.T. and A.H.) using the same classification.

Finally, we focused on comparing the relationship between the MRI findings and previous LBP.

Statistical analysis

The kappa statistic was used to summarize the intrareader and interreader reliability of the ratings. The kappa statistics were calculated with linear weights to give less importance to disagreements closer together on an ordinal scale. The schema of Landis and Koch[17] was used to interpret the strength of agreement based on the following values: 0, poor; 0–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, almost perfect. Between-group differences in baseline characteristics were evaluated using the Fisher's exact test for categorical variables and the Student's *t*-test for continuous variables. We compared the MRI findings between groups with and without previous LBP that did not have current LBP by using the Fisher's exact test. Furthermore, we determined the odds ratios of each item using univariate analyses and adjusting the analyses by age and sex. The statistical analyses were performed using the JMP 11.0 software program (SAS Institute, Cary, NC, USA). A p-value of <0.05 was considered to be significant.

Results

Of 91 participants, 27 had a history of LBP, which was indicated during medical consultation. The remaining 64 participants did not have any history of LBP. Participants' average age was 34.9 ± 10.6 years; 47 were female and 44 were male; and their average body mass index (BMI) was 21.8 ± 3.0 kg/m^2. The average ages of those who did and did

Table 1 Demographic data of the participants

Backgrounds	Total, n = 91	Previous LBP (+) group, n = 27	Previous LBP (–) group, n = 64	*p*-value
Age (years)	34.9 ± 10.6	38.3 ± 10.7	33.5 ± 10.4	0.0486*
Sex				
Female	47	12 (25.5)	35 (74.5)	0.3718
Male	44	15 (34.1)	29 (65.9)	
BMI (kg/m^2)	21.8 ± 3.0	21.8 ± 0.6	21.7 ± 0.4	0.9639

Notes: Data are shown as mean ± standard deviation or the number of participants (%). *$p < 0.05$.
Abbreviations: –, negative; +, positive; LBP, low back pain; BMI, body mass index.

Table 2 Details of the intraobserver and interobserver reliability of Pfirrmann grading, disk bulging, the high-intensity zone, and spondylolisthesis on magnetic resonance imaging reading

MRI findings	MRI (n)	Kappa	95% CI
Pfirrmann grading			
Intraobserver reliability	20 vs 20	0.66	0.55–0.77
Interobserver reliability	20 vs 20	0.64	0.52–0.76
Disk bulging			
Intraobserver reliability	20 vs 20	0.60	0.39–0.81
Interobserver reliability	20 vs 20	0.67	0.48–0.87
High-intensity zone			
Intraobserver reliability	20 vs 20	0.85	0.64–1.06
Interobserver reliability	20 vs 20	0.93	0.79–1.07
Spondylolisthesis			
Intraobserver reliability	20 vs 20	NA	NA
Interobserver reliability	20 vs 20	NA	NA

Abbreviations: CI, confidence interval; MRI, magnetic resonance imaging; NA, not applicable.

Table 3 Magnetic resonance imaging findings at any spinal level in groups with and without previous LBP that did not have current LBP

MRI findings	Total, n = 91	Previous LBP (+) group, n = 27	Previous LBP (–) group, n = 64	*p*-value
Pfirrmann grade ≥3	69 (75.8)	26 (96.3)	43 (67.2)	0.0026*
Disk bulging +	48 (52.3)	21 (77.8)	27 (42.2)	0.0019*
High-intensity zone +	19 (20.9)	9 (33.3)	10 (15.6)	0.0883
Spondylolisthesis +	4 (4.4)	3 (11.1)	1 (1.6)	0.0766

Notes: Data are shown as the number of participants (%). *$p < 0.05$.
Abbreviations: –, negative; +, positive; LBP, low back pain; MRI, magnetic resonance imaging.

not have a history of LBP were 38.3 and 33.5 years, respectively, which were significantly different ($p = 0.0486$). There were no significant differences in sex and BMI between the groups (Table 1).

The intraobserver and interobserver variabilities for Pfirrmann grading on MRI were 0.66 and 0.64, respectively. Those for disk bulging were 0.60 and 0.67, respectively. Those for the HIZ were 0.85 and 0.93, respectively. In 20 randomly selected MRIs, one observer did not identify spondylolisthesis at all, while the other observer identified spondylolisthesis in two levels of one participant. Thus, the intraobserver and interobserver variabilities for spondylolisthesis could not be calculated (Table 2).

Compared to the group without previous LBP, the group with previous LBP had a significantly higher incidence of disk degeneration such as a Pfirrmann grade ≥3 in at least one spinal level ($p = 0.0026$). In addition, the group with previous LBP had a significantly higher incidence of disk bulging in at least one spinal level than the group without previous LBP ($p = 0.0019$). There were no significant differences in HIZ ($p = 0.0883$) and spondylolisthesis ($p = 0.0766$) between the two groups according to the results of the Fisher's exact test (Table 3). Regarding the findings for each spinal level, compared to the group without previous LBP, the group with previous LBP had a significantly higher incidence of disk degeneration such as a Pfirrmann grade ≥3 at the T12/L1 ($p = 0.0350$), L3/4 ($p = 0.0232$), L4/5 ($p = 0.0005$), and L5/S1 ($p = 0.0026$) levels; and disk bulging at the L2/3 ($p = 0.0277$), L3/4 ($p = 0.0113$), L4/5 ($p = 0.0018$), and L5/S1 levels ($p = 0.0081$; Table 4). The findings of HIZ were almost all observed at the L4/5 and L5/S1 levels. Spondylolisthesis was only observed at the L4/5 and L5/S1 levels. In univariate analyses, the odds ratios of a Pfirrmann grade ≥3, disk bulging, HIZ, and spondylolisthesis were 12.7, 4.8, 2.7, and 7.9, respectively. There were significant differences for a Pfirrmann grade ≥3 ($p = 0.0009$) and disk bulging ($p = 0.0015$) in univariate analyses. In the adjusted analyses by age and sex, the odds ratios of the aforementioned four items were 10.5, 4.2, 3.1, and 6.6, respectively, and there were significant differences for a Pfirrmann grade ≥3 ($p = 0.0065$), disk bulging ($p = 0.0047$), and HIZ ($p = 0.0405$; Table 5).

Discussion

Among the participants in this study, ~30% had previous LBP, which was determined during the medical consultation. As in many industrialized countries, LBP is one of the most common health disabilities in Japan. In a population-based

Table 4 Pfirrmann grade and disk bulging at each spinal level in groups with and without previous LBP that did not have current LBP

MRI findings	Level	Total, n = 91	Previous LBP (+) group, n = 27	Previous LBP (–) group, n = 64	*p*-value
Pfirrmann grade ≥3	T12/L1	18 (19.8)	9 (33.3)	9 (14.1)	0.0350*
	L1/2	22 (24.2)	9 (33.3)	13 (20.3)	0.1851
	L2/3	30 (33.0)	10 (37.0)	20 (31.3)	0.5917
	L3/4	44 (48.4)	18 (66.7)	26 (40.6)	0.0232*
	L4/5	56 (61.5)	24 (88.9)	32 (50.0)	0.0005*
	L5/S1	56 (61.5)	23 (85.2)	33 (51.6)	0.0026*
Disk bulging (+)	T12/L1	2 (2.2)	1 (3.7)	1 (1.6)	0.5245
	L1/2	1 (1.1)	1 (3.7)	0 (0.0)	0.1216
	L2/3	2 (2.2)	2 (7.4)	0 (0.0)	0.0277*
	L3/4	5 (5.5)	4 (14.8)	1 (1.6)	0.0113*
	L4/5	35 (38.5)	17 (63.0)	18 (28.1)	0.0018*
	L5/S1	35 (38.5)	16 (59.3)	19 (29.7)	0.0081*

Notes: Data are shown as the number of participants (%). *$p < 0.05$.
Abbreviations: –, negative; +, positive; LBP, low back pain; MRI, magnetic resonance imaging.

Table 5 Odds ratio, 95% CI, and *p*-value from univariate analyses and analyses adjusted by age and sex for magnetic resonance imaging findings of groups with and without previous LBP that did not have current LBP

MRI findings	Univariate analyses			Age-adjusted and sex-adjusted analyses		
	Odds ratio	95% CI	*p*-value	Odds ratio	95% CI	*p*-value
Pfirrmann grade ≥3	12.7	2.43–234.18	0.0009*	10.5	1.78–202.09	0.0065*
Disk bulging	4.8	1.79–14.55	0.0015*	4.2	1.54–13.15	0.0047*
High-intensity zone	2.7	0.94–7.78	0.0652	3.1	1.05–9.42	0.0405*
Spondylolisthesis	7.9	0.96–163.50	0.0551	6.6	0.74–141.71	0.0923

Note: *$p < 0.05$.
Abbreviations: CI, confidence interval; LBP, low back pain; MRI, magnetic resonance imaging.

survey, the lifetime and 4-week LBP prevalence was 83% and 36%, respectively.[3] Therefore, LBP is one of the common causes of disability. In the current study, we precisely defined the region of LBP, which seemed to be important for standardizing the study protocol for LBP.[1,12] We also defined previous LBP as a history of medical consultation for LBP, which can exclude mild previous LBP. There was a significant difference in age between the two groups. Considering that disk degeneration progresses with advancing age,[6] the analyses performed in our study can be considered as appropriate.

The intraobserver and interobserver variabilities for each MRI finding were greater than moderate for all evaluated items.

MRI findings consistent with Pfirrmann grade ≥3, especially at the lower lumbar disk level, disk bulging, and HIZ were associated with previous LBP. Spondylolisthesis was not associated with previous LBP. There were significant differences between the groups in terms of a Pfirrmann grade ≥3, disk bulging, and HIZ according to the analyses adjusted by age and sex. The odds ratio of only the Pfirrmann grade ≥3 was >10, ie, a Pfirrmann grade ≥3 is strongly associated with a history of previous LBP in those without current LBP.

Pfirrmann grading indicates the degree of disk degeneration.[14] We divided the grading into two groups for the purpose of analysis. We regarded those with grades 1–2 as having no or little disk degeneration and those with grades 3–5 as having some degree of disk degeneration. There have been many reports on the relationship between current LBP and disk degeneration;[5–7] however, none have reported on the relationship between previous LBP and Pfirrmann grading. Videman et al[10] showed that disk height narrowing was associated with previous LBP, but they did not use Pfirrmann grading. Since disk height narrowing was classified as Pfirrmann grade 5,[14] this can be interpreted as implying that severe disk degeneration was associated with previous LBP. Although we included Pfirrmann grades 5, 3, and 4, which did not indicate severe disk height narrowing, our findings were almost consistent with the previous study's findings in terms of disk degeneration.

Pfirrmann grade ≥3 at T12/L1, L3/4, L4/5, and L5/S1 was associated with previous LBP. A large population study showed that disk degeneration was most commonly affected at L5-S1 and L4-L5,[6] which corresponds with our findings. A mechanical study showed that the range of motion in the

lower lumbar segments was significantly smaller than that in the upper segments.[18] The small range of motion at the intervertebral disk space can cause the load to increase at the disk, which can easily cause disk degeneration. This may be a reason why disk degeneration was more prominent at the lower lumbar disk levels than at the upper disk levels in the current study.

Disk bulging was associated with previous LBP. Regarding each spinal level, disk bulging at the L2/3, L3/4, L4/5, and L5/S1 levels was associated with previous LBP. Although the *p*-values were inclined to be smaller at lower disk levels than at upper disk levels, previous LBP was associated with disk bulging at almost all the lumbar disk levels. Some studies have shown that disk bulging was frequently observed in asymptomatic subjects, and it was concluded that there was no relationship between disk bulging and current LBP,[19,20] whereas another study of a meta-analysis showed a strong relationship.[7] As for previous LBP, Videman et al[10] showed that disk bulging was not associated with previous LBP. Our findings were not consistent with previous findings in terms of disk bulging.

A systematic review of the relationship between MRI findings and current LBP showed that disk degeneration and disk bulging are associated with current LBP, especially in younger adults, and this relationship disappears in older populations.[7] Although the study did not mention previous LBP, we can assume that older adults with disk degeneration or disk bulging who do not have current LBP may have had LBP when they were younger. These results correspond with our findings.

The HIZ was often observed at the level of L4/5 and L5/S1, and it was associated with previous LBP. There was a significant difference in the analyses adjusted by age and sex ($p = 0.0405$), although no significant relationship was found using the Fisher's exact test and univariate analyses. Aprill and Bogduk[16] reported a strong correlation between the annular high signal intensity zone and positive provocative discography. Some study has shown that the HIZ was associated with current LBP.[21] Dongfeng et al[22] performed a histological study on excised disks with a HIZ, and they concluded that the HIZ may be a specific signal for the inflammatory reaction of a painful disk. Conversely, other studies have shown that the HIZ was frequently observed in asymptomatic subjects.[7,19,20] As for previous LBP, Videman et al[10] showed that annular tear on axial MRI scans was associated with previous LBP. However, there has been no report on the relationship between the HIZ and previous LBP.

Spondylolisthesis was considered to be one of the findings of lumbar spine instability.[23] Considering that instability of the lumbar spine can cause LBP, it was assumed that those who had spondylolisthesis were inclined to have LBP.[24] However, some reports identified no significant relationship between spondylolisthesis and current LBP.[7,25] Furthermore, Hasegawa et al[26] showed that the radiological findings of spondylolisthesis cannot indicate instability. However, there has been no report on the relationship between spondylolisthesis and previous LBP. In our study, only four participants who did not have current LBP had spondylolisthesis. Three of these had previous LBP, and only one did not have previous LBP. There was no significant relationship between spondylolisthesis and previous LBP; however, this may be attributed to the small number of spondylolisthesis cases in our study.

One systematic review showed that HIZ and spondylolisthesis are not associated with current LBP, even in younger adults.[7] Therefore, the aforementioned information about disk degeneration or disk bulging does not correspond with HIZ and spondylolisthesis.

While some chronic LBP patients show continuous pattern, others have intermittent pattern.[11] Therefore, there was a possibility that the participants in our study who had previous LBP without current LBP had chronic LBP as intermittent pain. They did not have LBP at the time of participation; however, they may suffer recurrent LBP in the future as a natural course in the intermittent LBP pattern. Based on the results of the current study, MRI findings consistent with Pfirrmann grade ≥3, disk bulging, and HIZ may be one of the predictive signs of recurrent severe LBP. Thus, we can selectively educate patients about preventing LBP.

There were some limitations to the current study. First, we did not evaluate end plate changes because we only analyzed sagittal T2-weighted images and T1-weighted images were unavailable, even though Modic change has been considered to be associated with LBP.[5] In a population-based study on 975 participants, Teraguchi et al[27] showed that the combination of disk degeneration and end plate changes was highly associated with current LBP, whereas disk degeneration alone was not associated with current LBP. There is no previous report on the relationship between end plate changes and previous LBP, and we did not assess this relationship in our study. Second, we only analyzed sagittal images. Disk bulging and the HIZ can sometimes be visible at the posterolateral sides; however, these can be underestimated. Third, there was selection bias among our study participants, as they were volunteers from all types of employment at the hospital and did not represent the general population.

Conclusion

MRI findings consistent with Pfirrmann grading ≥3, especially at the lower lumbar disk level, disk bulging, and HIZ

were associated with previous LBP. In addition, spondylolisthesis was not associated with previous LBP. These findings may be one of the predictive signs of recurrent severe LBP.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Krismer M, van Tulder M, Low Back Pain Group of the Bone and Joint Health Strategies for Europe Project. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). *Best Pract Res Clin Rheumatol*. 2007;21(1):77–91.
2. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2163–2196.
3. Fujii T, Matsudaira K. Prevalence of low back pain and factors associated with chronic disabling back pain in Japan. *Eur Spine J*. 2013;22(2):432–438.
4. Feldman JB. The prevention of occupational low back pain disability: evidence-based reviews point in a new direction. *J Surg Orthop Adv*. 2004;13(1):1–14.
5. Kjaer P, Leboeuf-Yde C, Korsholm L, Sorensen JS, Bendix T. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine (Phila Pa 1976)*. 2005;30(10):1173–1180.
6. Cheung KM, Karppinen J, Chan D, et al. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. *Spine*. 2009;34(9):934–940.
7. Brinjikji W, Diehn FE, Jarvik JG, et al. MRI findings of disc degeneration are more prevalent in adults with low back pain than in asymptomatic controls: a systematic review and meta-analysis. *AJNR Am J Neuroradiol*. 2015;36(12):2394–2399.
8. Berg L, Hellum C, Gjertsen Ø, et al; Norwegian Spine Study Group. Do more MRI findings imply worse disability or more intense low back pain? A cross-sectional study of candidates for lumbar disc prosthesis. *Skeletal Radiol*. 2013;42(11):1593–1602.
9. Endean A, Palmer KT, Coggon D. Potential of magnetic resonance imaging findings to refine case definition for mechanical low back pain in epidemiological studies: a systematic review. *Spine (Phila Pa 1976)*. 2011;36(2):160–169.
10. Videman T, Battié MC, Gibbons LE, Maravilla K, Manninen H, Kaprio J. Associations between back pain history and lumbar MRI findings. *Spine (Phila Pa 1976)*. 2003;28(6):582–588.
11. Tamcan O, Mannion AF, Eisenring C, Horisberger B, Elfering A, Müller U. The course of chronic and recurrent low back pain in the general population. *Pain*. 2010;150(3):451–457.
12. Dionne CE, Dunn KM, Croft PR, et al. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. *Spine (Phila Pa 1976)*. 2008;33(1):95–103.
13. Mikkonen P, Heikkala E, Paananen M, et al. Accumulation of psychosocial and lifestyle factors and risk of low back pain in adolescence: a cohort study. *Eur Spine J*. 2016;25(2):635–642.
14. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine (Phila Pa 1976)*. 2001;26(17):1873–1878.
15. Fardon DF, Milette PC; Combined Task Forces of the North American Spine Society; American Society of Spine Radiology; American Society of Neuroradiology. Nomenclature and classification of lumbar disc pathology. Recommendations of the combined task forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. *Spine (Phila Pa 1976)*. 2001;26(5):E93–E113.
16. Aprill C, Bogduk N. High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. *Br J Radiol*. 1992;65(773):361–369.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
18. Lee SH, Daffner SD, Wang JC. Does lumbar disk degeneration increase segmental mobility in vivo? Segmental motion analysis of the whole lumbar spine using kinetic MRI. *J Spinal Disord Tech*. 2014;27(2):111–116.
19. Stadnik TW, Lee RR, Coen HL, Neirynck EC, Buisseret TS, Osteaux MJ. Annular tears and disk herniation: prevalence and contrast enhancement on MR images in the absence of low back pain or sciatica. *Radiology*. 1998;206(1):49–55.
20. Weishaupt D, Zanetti M, Hodler J, Boos N. MR imaging of the lumbar spine: prevalence of intervertebral disk extrusion and sequestration, nerve root compression, end plate abnormalities, and osteoarthritis of the facet joints in asymptomatic volunteers. *Radiology*. 1998;209(3):661–666.
21. Schellhas KP, Pollei SR, Gundry CR, Heithoff KB. Lumbar disc high-intensity zone. Correlation of magnetic resonance imaging and discography. *Spine (Phila Pa 1976)*. 1996;21(1):79–86.
22. Dongfeng R, Hou S, Wu W, et al. The expression of tumor necrosis factor-α and CD68 in high-intensity zone of lumbar intervertebral disc on magnetic resonance image in the patients with low back pain. *Spine (Phila Pa 1976)*. 2011;36(6):E429–E433.
23. McGregor AH, McCarthy ID, Doré CJ, Hughes SP. Quantitative assessment of the motion of the lumbar spine in the low back pain population and the effect of different spinal pathologies of this motion. *Eur Spine J*. 1997;6(5):308–315.
24. Alfieri A, Gazzeri R, Prell J, Röllinghoff M. The current management of lumbar spondylolisthesis. *J Neurosurg Sci*. 2013;57(2):103–113.
25. Kalichman L, Kim DH, Li L, Guermazi A, Berkin V, Hunter DJ. Spondylolysis and spondylolisthesis: prevalence and association with low back pain in the adult community-based population. *Spine (Phila Pa 1976)*. 2009;34(2):199–205.
26. Hasegawa K, Kitahara K, Shimoda H, et al. Lumbar degenerative spondylolisthesis is not always unstable: clinicobiomechanical evidence. *Spine (Phila Pa 1976)*. 2014;39(26):2127–2135.
27. Teraguchi M, Yoshimura N, Hashizume H, et al. The association of combination of disk degeneration, end plate signal change, and Schmorl node with low back pain in a large population study: the Wakayama Spine Study. *Spine J*. 2015;15(4):622–628.

Journal of Orthopaedic Science xxx (2016) 1–6

Contents lists available at ScienceDirect

Journal of Orthopaedic Science

journal homepage: http://www.elsevier.com/locate/jos

Original Article

The Japanese version of the STarT Back Tool predicts 6-month clinical outcomes of low back pain

Ko Matsudaira [a, *], Hiroyuki Oka [a], Norimasa Kikuchi [b, c], Yuri Haga [b], Takayuki Sawada [b, c], Sakae Tanaka [d]

[a] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8654, Japan
[b] Clinical Study Support, Inc., Daiei Bldg., 2F, 1-11-20 Nishiki, Naka-ku, Nagoya, 460-0003, Japan
[c] Department of Public Health, Aichi Medical University, 1-1 Yazakokarimata, Nagakute, Aichi 480-1195, Japan
[d] Department of Orthopaedic Surgery, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8654, Japan

ARTICLE INFO

Article history:
Received 6 July 2016
Received in revised form
11 November 2016
Accepted 2 December 2016
Available online xxx

ABSTRACT

Background: The STarT Back Tool classifies patients into low-, medium-, or high-risk groups according to risk for chronic low back pain. The Japanese version of the STarT Back Tool (STarT-J) has been translated and psychometrically validated. The present analysis investigated the predictive ability of the STarT-J.
Methods: Baseline data were collected through an online survey conducted with Japanese patients with low back pain. Long-term outcomes were assessed in a 6-month follow-up survey. Clinical outcomes at 6 months were evaluated with a pain numerical rating scale, the Roland–Morris Disability Questionnaire, and the EuroQol 5 Dimension. Differences in these scores among the three STarT-J risk groups were analyzed. Participants' perceived changes in low back pain and overall health status were examined to determine associations between the chronicity of low back pain at 6 months and STarT-J risk groups.
Results: Data of 1228 volunteers who responded to the baseline and follow-up surveys were included in this analysis. Mean ± standard deviation (SD) scores for the pain numerical rating scale and the Roland–Morris Disability Questionnaire were highest in the high-risk group (5.6 ± 1.9 and 9.6 ± 7.5) and lowest in the low-risk group (3.9 ± 1.6 and 2.1 ± 3.5). Mean ± SD EuroQol 5 Dimension index scores were lowest in the high-risk group (0.66 ± 0.20) and highest in the low-risk group (0.86 ± 0.14). A small percentage of high-risk patients (5.3%) perceived improvement in low back pain at the 6-month follow-up.
Conclusions: The STarT-J predicted 6-month pain and disability outcomes. The STarT-J is an easy-to-use tool to screen for patients who are more likely to have chronic low back pain, and may be useful to initiate stratified care in primary care settings.

1. Introduction

Stratified care to provide targeted treatment suitable for specific groups of patients has become a dominant approach in the management of low back pain (LBP) [1]. Stratification can be determined in several ways (e.g., based on underlying causes, prognostic factors, and patients' responses to treatment), but stratification based on prognostic factors is a prominent approach in primary care [1]. Evidence-based guidelines recommend that prognostic factors should be identified in deciding the management of LBP [2,3].

In assessing prognostic factors, early identification of risk for persistent LBP is particularly important [4], as recovery from chronic LBP is less likely when pain and disability are prolonged [5]. It is widely recognized that psychological factors such as depression, pain catastrophizing, and fear-avoidance beliefs contribute to the chronicity of LBP [6]. Therefore, it is especially important to screen for the presence of psychological factors in the early stages of LBP, to help predict the risk for chronicity and determine the most appropriate future management strategy.

The STarT Back Tool (STarT) is a multidimensional screening measure that identifies risk for chronic LBP by assessing physical and psychological prognostic factors [7]. The STarT classifies patients into three risk groups based on scores on nine overall items and five psychosocial subscale items (items 5–9) (Fig. 1) [8].

* Corresponding author. Fax: +81 3 5800 9549.
E-mail address: kohart801@gmail.com (K. Matsudaira).

http://dx.doi.org/10.1016/j.jos.2016.11.023
0949-2658/

Patients with a total score of 0–3 are classified as low-risk, those with a total score of ≥4 but a psychosocial subscore of ≤3 are medium-risk, and those with a psychosocial subscore of ≥4 are classified as high-risk. Providing education and support for self-management may be suitable for low-risk patients, and physical therapy may reduce pain and disability for medium-risk patients. For patients in the high-risk group, a combination of cognitive-behavioral therapy and physical therapy would be appropriate to manage psychological obstacles [9]. Stratified care based on the STarT risk classification has been demonstrated to be clinically and economically beneficial for patients with LBP [9,10]. Because of its benefits and usefulness, the STarT has been translated into various languages and is used worldwide [11–18].

In Japan, Matsudaira et al. translated and linguistically validated the Japanese version of the STarT (STarT-J) [19], and assessed its psychometric properties using cross-sectional data [20]. For concurrent validity, an overall moderate correlation was found between the STarT-J and external reference questionnaires. Known-groups validity was demonstrated by assessing the relationships between the STarT-J total scores or risk groups and the LBP-associated disability. The Cronbach's alpha coefficient showed the overall scale of the STarT-J had good internal consistency. The analysis demonstrated that the STarT-J was valid and reliable for the assessment of LBP in the Japanese population [20]. However, clinical outcomes of patients who were classified using the STarT-J have not yet been investigated.

Therefore, the present analysis assessed the relationships between the STarT-J risk groups and clinical outcomes for the stratified patients with LBP. We used longitudinal data to investigate whether the STarT-J high-risk group would have more chronic LBP at a 6-month follow-up.

2. Materials and methods

2.1. Study population

The present analysis is a part of a larger study consisting of a series of online surveys on LBP in the Japanese population. The baseline survey was conducted in January to February 2014, to investigate the physical and psychological aspects of patients with LBP. Participants were volunteers who were aged 20–64 years and recently had LBP, recruited through an online panel provided by an Internet research company, UNITED, Inc. (Tokyo, Japan). According to the standard definition of LBP by Dionne et al. [21], LBP was defined as pain in the low back that was experienced in the last 4 weeks and that lasted for more than 1 day. Pain associated with menstruation or pregnancy and pain during a feverish illness were excluded. Responses were obtained from 2000 individuals. To assess long-term clinical outcomes, a follow-up survey was conducted 6 months later and responses were received from 1228 individuals. In the present analysis, we used data of these 1228 participants for whom baseline and follow-up survey data were

The Keele STarT Back Screening Tool

Patient name: _____ Date: _____

Thinking about the **last 2 weeks** tick your response to the following questions:

		Disagree 0	Agree 1
1	My back pain has **spread down my leg(s)** at some time in the last 2 weeks	☐	☐
2	I have had pain in the **shoulder** or **neck** at some time in the last 2 weeks	☐	☐
3	I have only **walked short distances** because of my back pain	☐	☐
4	In the last 2 weeks, I have **dressed more slowly** than usual because of back pain	☐	☐
5	It's not really safe for a person with a condition like mine to be physically active	☐	☐
6	**Worrying thoughts** have been going through my mind a lot of the time	☐	☐
7	I feel that **my back pain is terrible** and **it's never going to get any better**	☐	☐
8	In general I have **not enjoyed** all the things I used to enjoy	☐	☐

9. Overall, how **bothersome** has your back pain been in the **last 2 weeks**?

Not at all	Slightly	Moderately	Very much	Extremely
☐ 0	☐ 0	☐ 0	☐ 1	☐ 1

Total score (all 9): _____ **Sub Score (Q5-9):**_____

© Keele University 01/08/07
Funded by Arthritis Research UK

Fig. 1. The STarT Back Tool. Items 5–9 constitute the psychosocial subscale.

available, and investigated the 6-month clinical status of participants in each STarT-J risk group. A 6-month follow-up period was chosen because the predictive validity of the original STarT developed in UK was tested using 6-month follow-up data [8].

The present analysis received approval from the medical/ethics review board of the Japan Labour Health and Welfare Organization, Kanto Rosai Hospital. Participation was voluntary, and no personal information was collected. Because of the nature of the study (online surveys), no written informed consent was obtained. As potential participants first read an explanation of the aim of the survey and only those who were willing to participate could proceed to the questionnaire, submission of a completed questionnaire was considered as evidence of consent. As an incentive, reward points for online shopping were given to the survey respondents from the Internet research company.

2.2. Measures

The following measures were included in the 6-month follow-up survey to assess participants' long-term clinical outcomes.

2.2.1. Numerical rating scale (NRS)

A NRS was used to assess the degree of pain related to LBP. The scale ranged from 0 (no pain at all) to 10 (the worst pain imaginable).

2.2.2. The Roland–Morris Disability Questionnaire (RDQ)

The RDQ was used to assess LBP-associated disability participants experienced in their daily lives. The RDQ consists of 24 Yes/No questions, and the total score ranges from 0 to 24. A higher score indicates greater disability. The validity and reliability of the Japanese version of the RDQ have been previously confirmed [22].

2.2.3. The EuroQol 5 Dimension (EQ-5D)

The EQ-5D [23] provides a simple descriptive profile and a single index score for general health status. We used this scale to measure participants' general health status. A converted index score ranges from −0.11 to 1.00: a score of 1 means "perfect health" and a score of 0 means "death."

To examine the changes in LBP status and overall health status over the 6-month period in each STarT-J risk group, we included two questions to assess the subjectively-perceived changes in health status: 1) How did your LBP status change compared with 6 months ago? and 2) How did your overall health status change compared with 6 months ago? Participants were asked to respond to each question on a 7-point scale: completely recovered, greatly improved, a little improved, not changed, a little worsened, greatly worsened, or became worst.

2.3. Statistical analysis

Demographic and clinical characteristics of participants at baseline were summarized for each STarT-J risk group using descriptive statistics. Demographic data included age, sex, body mass index, working status, smoking habits, and exercise habits.

For clinical outcomes, we calculated summary statistics for the pain NRS, RDQ, and EQ-5D scores at baseline and at 6-month follow-up for each STarT-J risk group. We performed an analysis of variance to test whether or not there were differences in scores between the three STarT-J risk groups (low-, medium-, and high-risk) at the 6-month follow-up.

Subjectively-perceived changes in health status were analyzed using a Cochran-Armitage trend test, by which we investigated whether there was a linear trend in the rate of improvement in LBP or overall health status across the STarT-J risk groups. In the analysis, the LBP or overall health status reported on the 7-point scale was interpreted as either improved or not improved: the responses "completely recovered" and "greatly improved" were considered improved, and the remaining 5 responses ("a little improved," "not changed," "a little worsened," "greatly worsened," and "became worst") as not improved.

All statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA). The level of significance was set at 0.05.

3. Results

3.1. Participant characteristics

The present analysis included baseline and 6-month follow-up data of 1228 Japanese individuals. Table 1 presents a summary of the baseline demographic and clinical characteristics of the participants in each risk group. The mean (standard deviation [SD]) age of 1228 participants was 47.9 years (9.1 years) and 55.6% of participants were male. Based on STarT-J scores at baseline, participants were classified into three risk groups: 958 participants (78.0%) in the low-risk group, 176 participants (14.3%) in the medium-risk group, and 94 participants (7.7%) in the high-risk group. At baseline, the mean (SD) pain NRS score in the low-risk group was 3.8 (1.6), that in the medium-risk group was 5.2 (1.8), and that in the high-risk group was 6.2 (1.9). The mean (SD) RDQ score at baseline was 2.6 (3.2) in the low-risk group, 8.1 (4.9) in the medium-risk group, and 11.6 (6.7) in the high-risk group. Overall, four participants were considered to have LBP with specific pathologic change, 185 participants had LBP with radiating leg pain, and the remaining 1039 participants were considered to have LBP with non-specific causes.

3.2. Scores at 6-month follow-up

At the 6-month follow-up, the mean (SD) pain NRS score in the low-risk group was 3.9 (1.6), that in the medium-risk group was 5.0 (1.9), and in the high-risk group was 5.6 (1.9) (Fig. 2). Higher mean scores were observed in higher risk groups, and there were significant differences in scores among the three risk groups (the analysis of variance, $p < 0.0001$). The mean (SD) RDQ score at 6-month follow-up in the low-risk group was 2.1 (3.5), in the medium-risk group was 6.3 (5.6), and in the high-risk group was 9.6 (7.5) (Fig. 3). Again, the higher risk groups had higher scores, and the differences in scores among the three risk groups were significant (the analysis of variance, $p < 0.0001$). The mean (SD) EQ-5D index scores at 6-month follow-up were 0.86 (0.14), 0.73 (0.13), and 0.66 (0.20) for the low-, medium-, and high-risk groups respectively (Fig. 4). The higher risk groups had lower EQ-5D index scores, meaning that those who were classified in a higher risk group tended to report poorer health status. The between-group differences in EQ-5D index scores were also significant (the analysis of variance, $p < 0.0001$).

3.3. Changes in LBP and health status over 6 months

To investigate the actual chronicity of LBP at 6 months in each risk group, we evaluated participants' perception of their improvement in LBP and overall health status over 6 months. In total, 18.3% of participants in the low-risk group perceived improvement in LBP at 6-month follow-up, whereas 10.2% in the medium-risk group and 5.3% in the high-risk group perceived improvement. The statistical analysis showed a decreasing linear trend (the Cochran-Armitage trend test, $p < 0.0001$) (Fig. 5a). A similar trend was observed in the assessment of overall health

Please cite this article in press as: Matsudaira K, et al., The Japanese version of the STarT Back Tool predicts 6-month clinical outcomes of low back pain, Journal of Orthopaedic Science (2016), http://dx.doi.org/10.1016/j.jos.2016.11.023

Table 1
Baseline characteristics of participants in each STarT-J risk group (N = 1228).

Characteristics	Low-risk group (N = 958)	Medium-risk group (N = 176)	High-risk group (N = 94)
Age, years	48.0 (9.1)	47.4 (9.0)	48.1 (9.4)
Sex, male (n, %)	543 (56.7)	96 (54.5)	44 (46.8)
BMI ≥ 25 (kg/m^2) (n, %)	234 (24.4)	47 (26.7)	29 (30.9)
Job (n, %)			
With	714 (74.5)	132 (75.0)	53 (56.4)
Without	244 (25.5)	44 (25.0)	41 (43.6)
Current smoking habits (n, %)			
Yes	286 (29.9)	66 (37.5)	29 (30.9)
No	672 (70.1)	110 (62.5)	65 (69.1)
Exercise habits (n, %)			
Yes	236 (24.6)	37 (21.0)	16 (17.0)
No	722 (75.4)	139 (79.0)	78 (83.0)
Duration of low back pain (n, %)			
Less than 3 months	352 (36.7)	54 (30.7)	21 (22.3)
3 months or longer	606 (63.3)	122 (69.3)	73 (77.7)
Number of recurrence (n, %)			
First time	79 (8.2)	11 (6.3)	5 (5.3)
Second time	58 (6.1)	12 (6.8)	7 (7.4)
3 or 4 times	167 (17.4)	18 (10.2)	20 (21.3)
5 to 9 times	149 (15.6)	25 (14.2)	10 (10.6)
10 times or more	505 (52.7)	110 (62.5)	52 (55.3)
Pain NRS score	3.8 (1.6)	5.2 (1.8)	6.2 (1.9)
RDQ score	2.6 (3.2)	8.1 (4.9)	11.6 (6.7)
EQ-5D index score	0.82 (0.14)	0.66 (0.13)	0.61 (0.20)

Values are n (%) or mean (standard deviation). STarT-J: Japanese version of the STarT Back Tool. BMI: body mass index. NRS: numerical rating scale. RDQ: The Roland–Morris Disability Questionnaire. EQ-5D: The EuroQol 5 Dimension.

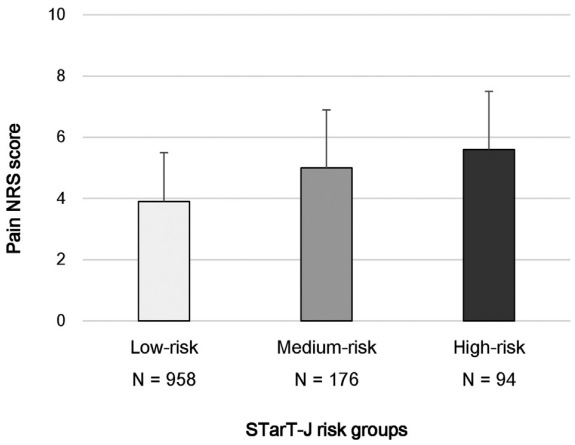


Fig. 2. Mean pain NRS scores at 6-month follow-up. Score differences among STarT-J risk groups were tested using analysis of variance ($p < 0.0001$). NRS: numerical rating scale. STarT-J: Japanese version of the STarT Back Tool.

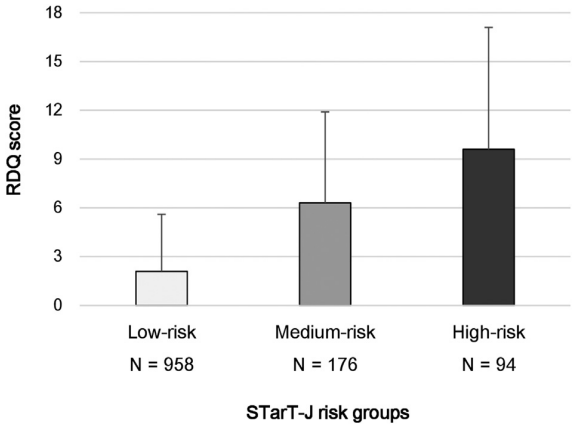


Fig. 3. Mean RDQ scores at 6-month follow-up. Score differences among STarT-J risk groups were tested using analysis of variance ($p < 0.0001$). RDQ: The Roland–Morris Disability Questionnaire. STarT-J: Japanese version of the STarT Back Tool.

status, with improvement perceived in overall health status at 6 months by 17.3%, 6.3%, and 4.3% of participants in the low-, medium-, and high-risk groups respectively. There was a decreasing linear trend in the improvement rate across the risk groups (the Cochran-Armitage trend test, $p < 0.0001$) (Fig. 5b).

4. Discussion

As psychological factors such as depression, fear-avoidance beliefs and behaviors, pain catastrophizing, and anxiety are known to be associated with chronicity of LBP [6], it is important to target and modify these prognostic factors at an early stage to improve the outcomes of LBP. A previous study showed that higher levels of fear-avoidance beliefs, kinesiophobia, and depressive symptoms were associated with non-recovery of LBP at 6 months [24]. The STarT is one tool that can assess these psychological factors in addition to LBP-related symptoms and physical impairment, and it stratifies patients into three risk groups based on the modifiable prognostic factors of their LBP. Considering that the STarT helps clinicians decide an appropriate therapeutic approach, it is important that the stratification is appropriate. Therefore in the present analysis, we assessed the 6-month clinical outcomes of patients stratified into each STarT-J risk group to evaluate whether the STarT-J appropriately predicted a poor prognosis of LBP (e.g. high-risk patients would be more likely to develop chronic LBP).

The results of the present analysis showed associations between the STarT-J high risk group and the poor 6-month clinical outcomes. The mean pain NRS scores were highest in the high-risk group and lowest in the low-risk group at the 6-month follow-up. This indicates that the STarT-J identified patients who would have greater pain at 6 months. The original STarT was able to predict pain at both 6 and 12 months in patients with non-specific LBP [25]. In addition, the mean RDQ scores at the 6-month follow-up were significantly

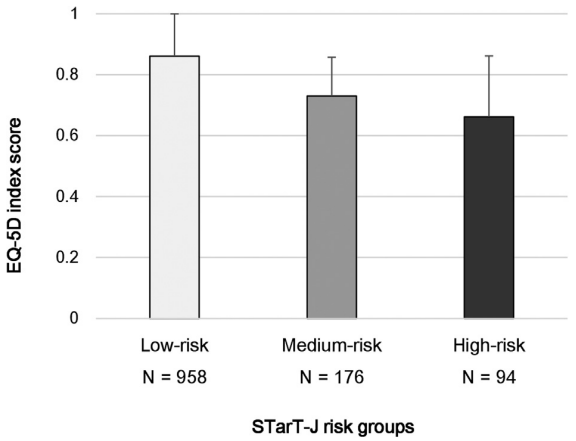


Fig. 4. Mean EQ-5D index scores at 6-month follow-up. Score differences among STarT-J risk groups were tested using analysis of variance ($p < 0.0001$). EQ-5D: The EuroQol 5 Dimension. STarT-J: Japanese version of the STarT Back Tool.

higher in the higher risk groups, indicating that similar to the original STarT, the STarT-J also detected patients with greater disability at 6 months [25]. Associations between the STarT risk groups and the long-term disability outcomes were reported in an analysis of the Danish version of the STarT [26]. In that analysis, the proportion of patients with a poor outcome (a RDQ score >30 on a 0–100 scale) at 3 months was highest in the high-risk group and lowest in the low-risk group. Although direct comparison is not appropriate given the different study designs, similar results were observed in the present analysis, with higher risk groups reporting poorer long-term disability outcomes.

Chronic LBP may be strongly related to patients' poor overall health status. In the present analysis, nearly 95% of participants in the high-risk group did not perceive sufficient improvement in LBP over 6 months, indicating a high possibility that these participants were suffering from chronic LBP. This risk group reported the least improvement in overall health status. Poor health status in this group was also shown by the lowest mean EQ-5D index score at the 6-month follow-up. Although the present analysis was unable to assess causal relationships between LBP and these variables, the results imply that chronic LBP may have a negative impact on patients' perception of their overall health status. Chronic LBP has been reported to have a negative impact on patients' psychological health, as well as their physical health [27]. Considering the potential negative impact of chronic LBP, our results highlight the importance of starting stratified care at an early stage, allowing modifiable risk factors (especially psychological factors) to be targeted in controlling the chronicity of LBP.

The major advantage of the STarT is in its simplicity. The tool is a 9-item self-report questionnaire for patients with non-specific LBP, which was developed to aid primary care decision making. As it is a quick tool, it would be helpful for clinicians, especially in primary care settings, to stratify patients according to their prognostic factors and decide which treatment strategy would be most appropriate (e.g., cognitive-behavioral therapy for high-risk patients). The STarT-J may therefore contribute to early stratified care in Japanese primary care settings.

The present analysis has some limitations. First, we observed a high attrition rate: in the follow-up survey, responses were received from 1228 of the 2000 participants who responded to the baseline survey. However, participants' characteristics in both analyses were similar and this was considered to represent a natural decline. A possible reason may be that participation in the survey was voluntary and no action was taken to achieve a high follow-up rate (e.g., e-mail reminders). Second, although participants were recruited according to the standard definition of LBP [21], the present analysis might have included patients not targeted by the STarT such as those with specific LBP causes. According to the diagnostic triage of LBP [28], responses indicated that four participants were probable “red flag,” who had LBP with specific pathologic change, but the remaining 1224 participants probably fit into the STarT target group. However, as these responses were based on participants' self-report, we cannot exclude the possibility of misclassification or misdiagnosis. Third, as this was an observational study and we did not interfere in participants' choice of treatment for LBP, outcomes at 6 months might have been influenced by a treatment that participants had received. Some participants reported to have received some treatment during the period, but others reported to have received no treatment. However, as information on the types of treatment was not collected in the

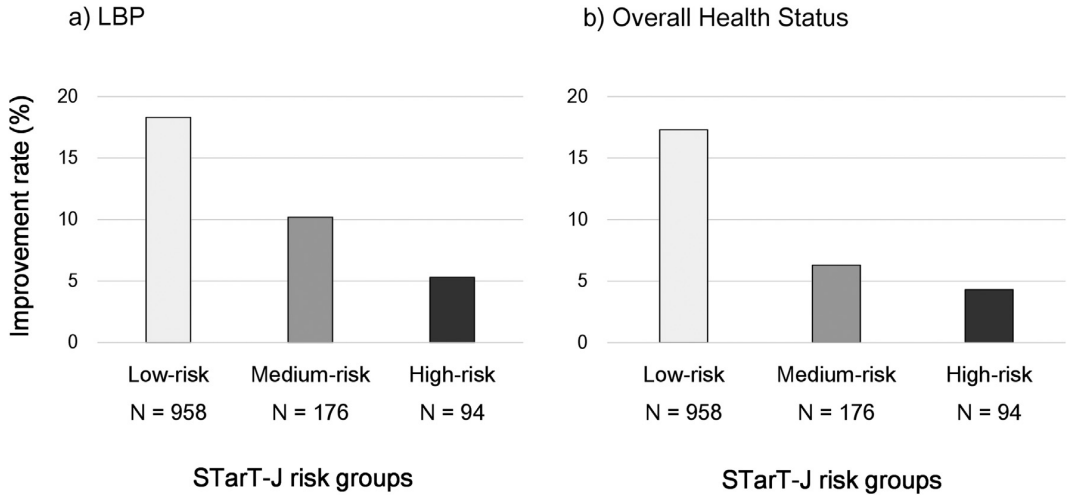


Fig. 5. Improvement rate of a) LBP and b) overall health status. Cochran-Armitage trend tests were performed to investigate the improvement rates ($p < 0.0001$ for both). LBP: low back pain. STarT-J: Japanese version of the STarT Back Tool.

Please cite this article in press as: Matsudaira K, et al., The Japanese version of the STarT Back Tool predicts 6-month clinical outcomes of low back pain, Journal of Orthopaedic Science (2016), http://dx.doi.org/10.1016/j.jos.2016.11.023

survey, actual treatment status of the individual participants is unknown. If participants had received effective treatments targeting modifiable indicators of LBP, it might have resulted in better outcomes compared to the natural course of LBP as predicted by the STarT-J at baseline. Therefore, the results of the present analysis need to be interpreted with care.

In conclusion, the STarT-J may help predict the 6-month prognosis of LBP, which allows classification of patients according to their risk for chronic LBP. Chronic LBP causes great disability and negatively affects patients' overall health status; therefore, it is important to start stratified care at an early stage in the management of non-specific LBP. The STarT-J is a simple, quick screening tool appropriate for use in primary care, which would enable and further promote early stratified care.

Conflict of interest statement

KM received grants from Ministry of Health, Labour and Welfare during the conduct of the study; grants from Sumitomo Dainippon Pharma Co., Ltd. and Okamura Corporation; grants and personal fees from AYUMI Pharmaceutical Corporation, Nippon Zoki Pharmaceutical Co., Ltd., Ono Pharmaceutical Co., Ltd., Eli Lilly Japan K.K., Astellas Pharma Inc., TOTO Ltd., and Eisai Co., Ltd.; personal fees and advisory fees from Shionogi & Co., Ltd.; and personal fees from Pfizer Japan Inc., Hisamitsu Pharmaceutical Co., Inc., Janssen Pharmaceutical K.K., Kaken Pharmaceutical Co., Ltd., and Teijin Pharma Limited, outside the submitted work.

HO received personal fees from Chugai Pharmaceutical Co., Ltd. and grants from Teijin Pharma Limited and Pfizer Japan Inc., outside the submitted work.

NK is a board member of Clinical Study Support, Inc.

YH and TS are employed by Clinical Study Support, Inc.

ST received personal fees from Amgen Inc., Asahi Kasei Pharma Corporation, Amgen Astellas BioPharma K.K., Ono Pharmaceutical Co., Ltd., KYOCERA Medical Corporation, Teijin Pharma Limited, and Eli Lilly Japan K.K.; endowments from Astellas Pharma Inc., AYUMI Pharmaceutical Corporation, Bristol-Myers Squibb, and Chugai Pharmaceutical Co., Ltd.; personal fees and endowments from Pfizer Japan Inc. and Daiichi Sankyo Company, Limited; and grants from The Japan Agency for Medical Research and Development (AMED), Grant-in-Aid for Scientific Research (A), and Grant-in-Aid for Exploratory Research, outside the submitted work.

Acknowledgments

This study was supported by a grant for clinical research on workers' diseases from the Ministry of Health, Labour and Welfare (14020301-01).

References

[1] Foster NE, Hill JC, O'Sullivan P, Hancock M. Stratified models of care. Best Pract Res Clin Rheumatol 2013 Oct;27(5):649–61.
[2] van Tulder M, Becker A, Bekkering T, Breen A, del Real MT, Hutchinson A, Koes B, Laerum E, Malmivaara A. COST B13 working group on guidelines for the management of acute low back pain in primary care. Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. Eur Spine J 2006 Mar;15(Suppl. 2):S169–91.
[3] Chou R, Qaseem A, Snow V, Casey D, Cross Jr JT, Shekelle P, Owens DK. Clinical efficacy assessment subcommittee of the American College of Physicians; American College of Physicians; American pain society low back pain guidelines panel. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Int Med 2007 Oct 2;147(7):478–91. Erratum in: Ann Intern Med. 2008 Feb 5;148(3):247–8.
[4] Melloh M, Elfering A, Egli Presland C, Roeder C, Barz T, Rolli Salathé C, Tamcan O, Mueller U, Theis JC. Identification of prognostic factors for chronicity in patients with low back pain: a review of screening instruments. Int Orthop 2009 Apr;33(2):301–13.
[5] Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. BMJ 2006 Jun 17;332(7555):1430–4.
[6] Pincus T, McCracken LM. Psychological factors and treatment opportunities in low back pain. Best Pract Res Clin Rheumatol 2013 Oct;27(5):625–35.
[7] Beneciuk JM, Robinson ME, George SZ. Subgrouping for patients with low back pain: a multidimensional approach incorporating cluster analysis and the STarT Back Screening Tool. J Pain 2015 Jan;16(1):19–30.
[8] Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, Hay EM. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum 2008 May 15;59(5):632–41.
[9] Foster NE, Mullis R, Hill JC, Lewis M, Whitehurst DG, Doyle C, Konstantinou K, Main C, Somerville S, Sowden G, Wathall S, Young J, Hay EM. IMPaCT back study team. Effect of stratified care for low back pain in family practice (IMPaCT Back): a prospective population-based sequential comparison. Ann Fam Med 2014 Mar-Apr;12(2):102–11.
[10] Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, Konstantinou K, Main CJ, Mason E, Somerville S, Sowden G, Vohora K, Hay EM. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet 2011 Oct 29;378(9802):1560–71.
[11] Piironen S, Paananen M, Haapea M, Hupli M, Zitting P, Ryynänen K, Takala EP, Korniloff K, Hill JC, Häkkinen A, Karppinen J. Transcultural adaption and psychometric properties of the STarT back screening tool among Finnish low back pain patients. Eur Spine J 2016 Jan;25(1):287–95.
[12] Abedi M, Manshadi FD, Khalkhali M, Mousavi SJ, Baghban AA, Montazeri A, Parnianpour M. Translation and validation of the Persian version of the STarT back screening tool in patients with nonspecific low back pain. Man Ther 2015 Dec;20(6):850–4.
[13] Luan S, Min Y, Li G, Lin C, Li X, Wu S, Ma C, Hill JC. Cross-cultural adaptation, reliability, and validity of the Chinese version of the STarT back screening tool in patients with low back pain. Spine (Phila Pa 1976) 2014 Jul 15;39(16): E974–9.
[14] Pilz B, Vasconcelos RA, Marcondes FB, Lodovichi SS, Mello W, Grossi DB. The Brazilian version of STarT back screening tool – translation, cross-cultural adaptation and reliability. Braz J Phys Ther 2014 Sep-Oct;18(5):453–61.
[15] Bruyère O, Demoulin M, Beaudart C, Hill JC, Maquet D, Genevay S, Mahieu G, Reginster JY, Crielaard JM, Demoulin C. Validity and reliability of the French version of the STarT back screening tool for patients with low back pain. Spine (Phila Pa 1976) 2014 Jan 15;39(2):E123–8.
[16] Aebischer B, Hill JC, Hilfiker R, Karstens S. German translation and cross-cultural adaptation of the STarT back screening tool. PLoS One 2015 Jul 10;10(7):e0132068.
[17] Azimi P, Shahzadi S, Azhari S, Montazeri A. A validation study of the Iranian version of STarT back screening tool (SBST) in lumbar central canal stenosis patients. J Orthop Sci 2014 Mar;19(2):213–7.
[18] Morsø L, Albert H, Kent P, Manniche C, Hill J. Translation and discriminative validation of the STarT back screening tool into Danish. Eur Spine J 2011 Dec;20(12):2166–73.
[19] Matsudaira K, Kikuchi N, Kawaguchi M, Inuzuka K, Arisaka M, Hara N, Isomura T. Development of a Japanese version of the STarT (subgrouping for targeted treatment) back screening tool: translation and linguistic validation. J Musculoskelet Pain Res 2013;5(1):11–9. in Japanese.
[20] Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S. Psychometric properties of the Japanese version of the STarT back tool in patients with low back pain. PLoS One 2016 Mar 22;11(3):e0152019.
[21] Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, Wyatt M, Cassidy JD, Rossignol M, Leboeuf-Yde C, Hartvigsen J, Leino-Arjas P, Latza U, Reis S, Gil Del Real MT, Kovacs FM, Oberg B, Cedraschi C, Bouter LM, Koes BW, Picavet HS, van Tulder MW, Burton K, Foster NE, Macfarlane GJ, Thomas E, Underwood M, Waddell G, Shekelle P, Volinn E, Von Korff M. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976) 2008 Jan 1;33(1):95–103.
[22] Suzukamo Y, Fukuhara S, Kikuchi S, Konno S, Roland M, Iwamoto Y, Nakamura T. Committee on science project, Japanese Orthopaedic Association. Validation of the Japanese version of the Roland-Morris disability questionnaire. J Orthop Sci 2003;8(4):543–8.
[23] EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199–208.
[24] George SZ, Beneciuk JM. Psychological predictors of recovery from low back pain: a prospective study. BMC Musculoskelet Disord 2015 Mar 7;16(1):49.
[25] Pagé I, Abboud J, O Shaughnessy J, Laurencelle L, Descarreaux M. Chronic low back pain clinical outcomes present higher associations with the STarT back screening tool than with physiologic measures: a 12-month cohort study. BMC Musculoskelet Disord 2015 Aug 19;16(1):201.
[26] Morsø L, Kent P, Albert HB, Hill JC, Kongsted A, Manniche C. The predictive and external validity of the STarT back tool in Danish primary care. Eur Spine J 2013 Aug;22(8):1859–67.
[27] Mathew J, Singh SB, Garis S, Diwan AD. Backing up the stories: the psychological and social costs of chronic low-back pain. Int J Spine Surg 2013 Dec 1;7(1):e29–38.
[28] Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. Spine (Phila Pa 1976) 2001 Nov 15;26(22):2504–13. discussion 2513–4.

Please cite this article in press as: Matsudaira K, et al., The Japanese version of the STarT Back Tool predicts 6-month clinical outcomes of low back pain, Journal of Orthopaedic Science (2016), http://dx.doi.org/10.1016/j.jos.2016.11.023

SPINE An International Journal for the study of the spine Publish Ahead of Print
DOI : 10.1097/BRS.0000000000001956

EPIDEMIOLOGICAL DIFFERENCES BETWEEN LOCALISED AND NON-LOCALISED LOW BACK PAIN

David Coggon, FMedSci[1,2], Georgia Ntani, MSc[1,2], Karen Walker-Bone, PhD[1,2], Keith T Palmer, DM, MSc[1,2], Vanda E Felli, PhD[3], Raul Harari, MD, PhD[4], Lope H Barrero, ScD[5], Sarah A. Felknor, DrPH[6,7], David Gimeno, PhD[6], Anna Cattrell, PhD[8], Sergio Vargas-Prada, PhD[9,10,11], Matteo Bonzini, PhD[12], EleniSolidaki, PhD[13], EdaMerisalu, PhD[14], Rima R. Habib, PhD[15], FaridehSadeghian, MSc[16], M Masood Kadir, MPH[17], Sudath SP Warnakulasuriya, PhD[18], KoMatsudaira, PhD[19], BusisiweNyantumbu, MSc[20,21], Malcolm R Sim, PhD[22], Helen Harcombe, PhD[23], Ken Cox, ONC[1], Leila M MSarquis, PhD[24],Maria H Marziale, PhD[25], FlorenciaHarari, MD, PhD[4], Rocio Freire, MSc[4], Natalia Harari, MSc[4], Magda V Monroy, MSc[5], Leonardo A Quintana, PhD[5], Marianela Rojas, PhD[26], E Clare Harris, PhD[1,2], Consol Serra, MD, PhD[9,10,11,27], J Miguel Martinez, PhD[28], George Delclos, MD[6,9,10,11], Fernando G Benavides, MD, PhD[9,10,11], Michele Carugno, PhD[29], Marco M Ferrario, MD[12], Angela C Pesatori, PhD[29,30], Leda Chatzi, MD, PhD[13], PanosBitsios, PhD[31], ManolisKogevinas, PhD[10,11,32], KristelOha, MSc[33], TiinaFreimann, MSc[34], Ali Sadeghian, MD[35], Roshini J Peiris-John, PhD[36,37], NaliniSathiakumar, MD, DrPH[38], A RajithaWickremasinghe, PhD[39], Noriko Yoshimura, PhD[40], Helen L Kelsall, PhD[22], Victor C W Hoe, PhD[41], Donna M Urquhart, PhD[22], Sarah Derrett, PhD[42], David McBride, PhD[23], Peter Herbison, DSc[23], Andrew Gray, BCom (Hons) BA[23]Eduardo J. Salazar Vega, PhD[43].

[1]Medical Research Council Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK

[2]Arthritis Research UK/MRC Centre for Musculoskeletal Health and Work, University of Southampton, Southampton, UK

[3]School of Nursing, University of São Paulo, São Paulo, Brazil

[4]Corporación para el Desarrollo de la Producción y el Medio Ambiente Laboral – IFA (InstitutefortheDevelopment of Production and theWorkEnvironment), Quito, Ecuador

[5]Department of Industrial Engineering, School of Engineering, Pontificia Universidad Javeriana, Bogotá, Colombia

[6]Southwest Center for Occupational and Environmental Health, The University of Texas Health Science Center at Houston School of Public Health, Houston, Texas, USA

[7]Center for Disease Control and Prevention/National Institute for Occupational Safety and Health, Atlanta, USA

[8]North East London NHS Foundation Trust, Goodmayes Hospital, Ilford, Essex, UK
[9]Center for Research in Occupational Health (CiSAL), UniversitatPompeuFabra, Barcelona, Spain
[10]CIBER of Epidemiology and Public Health, Barcelona, Spain
[11]IMIM (Hospital del Mar ResearchInstitute), Barcelona, Spain
[12]Epidemiology and Preventive Medicine Research Centre, Department of Clinical and Experimental Medicine, University of Insubria, Varese, Italy
[13]Department of Social Medicine, Medical School, University of Crete, Heraklion, Greece
[14]Institute of Technology, Estonian University of Life Sciences,Tartu, Estonia
[15]Department of Environmental Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon
[16]Department of Occupational Health, School of Public Health, Shahroud University of Medical Sciences, Shahroud, Iran
[17]Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan
[18]Department of Allied Health Sciences,Faculty of Medical Sciences,University of Sri Jayawardenepura, Gangodawila, Nugegoda, Sri Lanka
[19]Department for Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo Hospital, Tokyo, Japan
[20]National Institute for Occupational Health, National Health Laboratory Service, Johannesburg, South Africa
[21]Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa
[22]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
[23]Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand
[24]Federal University of Paraná, Curitiba-PR, Brazil
[25]School of Nursing of RibeirãoPreto, University of São Paulo, São Paulo, Brazil
[26]Program Health, Work and Environment in Central America, Institute for Studies on Toxic Substances (IRET), National University of Costa Rica, Heredia, Costa Rica
[27]Occupational Health Service, Parc de Salut MAR, Barcelona, Spain
[28]Servicio de Investigación y Análisis IT/EP, Departamento de Investigación y Análisis de Prestaciones, MC Mutual, Barcelona, Spain
[29]Department of Clinical Sciences and Community Health, UniversitàdegliStudi di Milano, Milan, Italy
[30]Fondazione Ca' GrandaOspedale Maggiore Policlinico, Milan, Italy
[31]Department of Psychiatry, Medical School, University of Crete, Heraklion, Greece

[32]Centre forResearch in EnvironmentalEpidemiology(CREAL), Barcelona, Spain
[33]North Estonia Medical Centre, Tallinn, Estonia
[34]Tartu University Hospital, Tartu, Estonia
[35]Klinikum Leverkusen, Leverkusen, Germany
[36]Department of Physiology, Faculty of Medical Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka
[37]Section of Epidemiology and Biostatistics, School of Population Health, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand
[38]Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA
[39]Faculty of Medicine, University of Kelaniya, Kelaniya, Sri Lanka
[40]Department of Joint Disease Research, 22nd Century Medical and Research Center, University of Tokyo, Tokyo, Japan
[41]Centre for Occupational and Environmental Health, Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[42]Injury Prevention Research Unit, Department of Preventive and Social Medicine, University of Otago, Dunedin, New Zealand
[43]Health Safety and Environment Department, AkzoNobel, Houston, Texas, USA

Correspondence to:
Professor David Coggon
MRC Lifecourse Epidemiology Unit
Southampton General Hospital
Southampton
SO16 6YDUK

Tel: #44 2380 777624
Fax: #44 2380 704021
Email: dnc@mrc.soton.ac.uk

Acknowledgement: December 14, 2015
1st Revise: April 14, 2016
2nd Revise: June 17, 2016
3rd Revise: August 2, 2016
Accept: August 31, 2016

The manuscript submitted does not contain information about medical device(s)/drug(s).

The Medical Research Council, Arthritis Research UK, NHMRC, and The Ministry of Higher Education in Malaysia funds were received in support of this work. In addition, a research training grant to Southwest Center for Occupational and Environmental Health at the University of Texas Health Science Center from the NIH Fogarty International Center, Monash University, Health Research Council of New Zealand, The Deputy for Training and Research, Shahroud University of Medical Sciences, ISCII, and The Colt Foundation funded data collection.

Relevant financial activities outside the submitted work:consultancy, grants, employment, travel/accommodations/meeting expenses.

ABSTRACT

Study design: Cross-sectional survey with longitudinal follow-up

Objectives: To test the hypothesis that pain which is localised to the low back differs epidemiologically from that which occurs simultaneously or close in time to pain at other anatomical sites

Summary of Background Data: Low back pain (LBP) often occurs in combination with other regional pain, with which it shares similar psychological and psychosocial risk factors. However, few previous epidemiological studies of LBP have distinguished pain that is confined to the low back from that which occurs as part of a wider distribution of pain.

Methods: We analysed data from CUPID, a cohort study that used baseline and follow-up questionnaires to collect information about musculoskeletal pain, associated disability and potential risk factors, in 47 occupational groups (office workers, nurses and others) from 18 countries.

Results: Among 12,197 subjects at baseline, 609 (4.9%) reported localised LBP in the past month, and 3,820 (31.3%) non-localised LBP. Non-localised LBP was more frequently associated with sciatica in the past month (48.1% vs. 30.0% of cases), occurred on more days in the past month and past year, was more often disabling for everyday activities (64.1% vs. 47.3% of cases), and had more frequently led to medical consultation and sickness absence from work. It was also more often persistent when participants were followed up after a mean of 14 months (65.6% vs. 54.1% of cases). In adjusted Poisson regression analyses, non-localised LBP was differentiallyassociated with risk factors, particularly female sex, older age and somatising tendency. There were also marked differences in the relative prevalence of localised and non-localised LBP by occupational group.

Conclusions: Future epidemiological studies should distinguish where possible between pain that is limited to the low back and LBP which occurs in association with pain at other anatomical locations.

Key Words: Low back pain; diagnostic classification; epidemiology; disability; medical consultation; sickness absence; sciatica; risk factors; somatising; occupation; prognosis

Level of Evidence: 2

INTRODUCTION

Low back pain (LBP) is a major cause of disability among people of working age [1], but investigation of its causes has been hindered by challenges in case definition. In most people with LBP, there is no clearly demonstrable underlying spinal pathology, and even where the pain occurs in association with structural abnormalities such as disc herniation or nerve root compression, only a minority of cases are attributable to theobserved pathology [2]. In the absence of more objective diagnostic criteria, most epidemiological studies have defined cases according to report of symptoms and/or accompanying disability,and this approach has given useful insights. For example, we know that LBP is associated with heavy lifting and other physical activities which subject the spine to mechanical stresses [3], although disappointingly, ergonomic interventions in the workplace to reduce such exposures have failed to prevent back problems [4]. Associations have also been found with psychological characteristics such as low mood [5-7], tendency to worry about common somatic symptoms (somatising tendency) [5,7], adverse health beliefs about musculoskeletal pain [6], and (to a lesser extent) psychosocial aspects of work [8].

The same psychological and psychosocial risk factors have been linked alsowith other regional musculoskeletal pain, for example in the upper limb [8,9] and knee [10]; and somatising tendency has shown particularly strong associations with multi-site pain [11]. Moreover, LBP frequently occurs in combination with pain at other anatomical sites, either simultaneously or close in time [12-15]. This raises the possibility that the observed associations of LBP with psychological and psychosocial risk factors might reflect effects on musculoskeletal pain more generally, and that pain which is limited only to the low back is epidemiologically distinct from that which occurs as part of a widerdistribution of pain. If this were the case, studies that failed to distinguish localised from non-localised LBP might miss associations with preventable causes, or incorrectly assess the impacts of treatment.

To test the hypothesis that localised and non-localised LBP are epidemiologically distinct, we analysed data from CUPID (Cultural and Psychosocial Influences on Disability), a large, multinational cohort study of musculoskeletal pain and associated disability in selected occupational groups [16], looking for differences in severity, associations with risk factors, and prognosis of LBP, according to whether or not pain was limited to the low back.

MATERIALS AND METHODS

The study sample for CUPID comprisedmen and women from 47 occupational groups (mainly nurses, office staff and workers carrying out repetitive manual tasks with their hands or arms) in 18 countries. Each of the 12,426 participants (overall response rate 70%) completed a baseline questionnaire, either by self-administration or at interview. The questionnaire was originally drafted in English and then translated into local languages as necessary, accuracy being checked by independent back-translation. Among other things, it asked about demographic characteristics, smoking habits, whether an average working day entailed lifting weights ≥25 kg, various psychosocial aspects of work, somatising tendency, mental health, beliefs about back pain, and experience of musculoskeletal pain during the past 12 months.

Somatising tendency was ascertained through questions taken from the Brief Symptom Inventory [17], and classified according to how many of five commonsomatic symptoms (faintness or dizziness, pains in the heart or chest, nausea or upset stomach, trouble getting breath and hot or cold spells) had caused at least moderate distress during the past week. Mental health was assessed through the relevant section of the Short Form 36 (SF-36) questionnaire [18], and scores were graded to three levels (good, intermediate or poor) representing approximate thirds of the distribution across the study sample. Participants were classed as having adverse beliefs about the work-relatedness of back pain if they completely agreed that such pain is commonly caused by work; about its relationship to physical activity if they completely agreed that for someone with

back pain, physical activity should be avoided as it might cause harm, and that rest is needed to get better; and about its prognosis if they completely agreed that neglecting such problems can cause serious harm, and completely disagreed that such problems usually get better within three months.

The questions about musculoskeletal pain used diagrams to define 10 anatomical regions of interest (low back; neck; and right and left shoulder, elbow, wrist/hand and knee). Participants were asked whether during the past 12 months, they had experienced pain lasting for a day or longer at these sites, and those who reported LBP were also asked whether the pain had occurred in the past month, whether it had spread down the leg to below the knee (sciatica), how long in total it had been present during the past month and past 12 months, whether during the past month it had made it difficult or impossible to cut toe nails, get dressed or do normal jobs around the house(disabling pain), whether it had led to medical consultation during the past 12 months, the total duration of any resultant sickness absence from work during the past 12 months, and whether the most recent episode had started suddenly while at work, suddenly while not at work or gradually (an episode of pain was defined as occurring after a period of at least one month without the symptom).

After an interval of approximately 14 months, participants from 45 of the occupational groups were asked to complete a short follow-up questionnaire, which again asked about LBP in the past month.

Further details of the methods of data collection, specification of variables, and characteristics of the study sample have been reported elsewhere [16]. Approval for the study was provided by the relevant research ethics committees in each participating country [16].

Statistical analysis was carried out with Stata software (Stata Corp LP 2012, Stata Statistical Software: Release 12.1,College Station TX, USA). From the baseline questions about pain, we distinguished participants who reported: LBP in the past month but no pain at any other site during

the past 12 months ("localised LBP"); LBP in the past month with pain at one or more other sites during the past 12 months ("non-localised LBP"); and no LBP at any time during the past 12 months. We used simple descriptive statistics to compare the features of localised and non-localised LBP, including the prevalence of continuing LBP (i.e. present in the past month) at follow-up. Associations with risk factors were explored by Poisson regression, and summarised by prevalence rate ratios (PRRs) with 95% confidence intervals (CIs) based on robust standard errors.To account for possible clustering by occupational group, we fitted random-intercept models. A scatter plot was used to explore the correlation of localised and non-localised LBP across the 47 occupational groups after adjustment for other risk factors. To derive adjusted prevalence rates, we took no LBP in the past 12 months as a comparator, and first estimated PRRs for the two pain outcomes in each occupational group relative to a reference (office workers in the UK),using Poisson regression models that included the other risk factors. We then calculated the "adjusted numbers" of participants in each occupational group with the two pain outcomes that would give crude PRRs equal to those estimated from the regression model. Finally, we used these adjusted numbers to calculate adjusted prevalence rates.

RESULTS

From the total of 12,426 participants who completed the baseline questionnaire, we excluded 149 because of missing information about LBP in the past month (122), 12 months (2) or both (25), and a further 80 who did not provide full responses regarding pain at other anatomical sites in the past 12 months. Among the remaining 12,197 subjects (35% men), 609 (5.0%) reported localised LBP in the past month, and 3,820 (31.3%) non-localised LBP.

Table 1 compares the characteristics of the pain in these two groups of people with low back symptoms. Non-localisedLBP was more frequently associated with sciatica (48.1% vs. 30.0% in past month), occurred on more days in the past month and past year, was more often disabling for

everyday activities (64.1% vs. 47.3%), and had more frequently led to medical consultation and sickness absence from work during the past year. However, there was no difference between the categories of LBP in the prevalence of sudden as opposed to gradual onset.

Table 2 summarises the associations of localised and non-localised LBP with various risk factors.The comparator in this analysis was no LBP at any time in the past 12 months (n = 5,501). Non-localised LBP was significantly more common in women than men, and at older ages, whereas the prevalence of localised LBP was significantly higher in men, and varied little with age. Somatising tendency was much more strongly related to non-localised LBP (PRR 1.7, 95%CI 1.5-1.8 for report of distress from two or more somatic symptoms) than localised LBP (PRR 1.1, 95%CI 0.9-1.4). Associations with non-localised pain were stronger also for poor mental health and report of time pressure at work. Direct comparison of participants with localised and non-localised LBP in a single Poisson regression model (effectively taking those with non-localised LBP as cases and those with localised LBP as controls) indicated that the differences in associations with sex, age and somatising tendency were allhighly significantstatistically ($p < 0.001$).

Figure 1 shows the one-month prevalence of localised and non-localised LBP by occupational group, after adjustment for all of the risk factors in Table 2. Rates of localised LBP ranged from zero among postal workers in New Zealand and 1.0% in office workers in Nicaragua to 11.9% in Sri Lankan nurses, and 12.6% in Brazilian sugar cane cutters. For non-localised LBP, the absolute variation in prevalence was even greater – from 3.9% in Brazilian sugar cane cutters and 6.8% among office workers in Pakistan to 28.1% in Brazilian office workers and 28.8% in Brazilian nurses. However there was no clear relationship between the two categories of LBP. Thus, as illustrated in Figure 2, the proportion of all back pain cases that were localisedvaried substantially, but did not consistently rise or fall as the overall prevalence of LBP increased (Spearman correlation coefficient = –0.37).

Among the 11,764 participants from whom follow-up data were sought, 9,188 (78%) provided satisfactory information about LBP at a mean of14 months (range 3-35 months, 84% within 11-19 months) after baseline.Table 3 shows the prevalence of continuing LBP at follow-up according to the features of pain at baseline. Overall, persistence of pain was more frequent when initially it was non-localised (65.6%) than when it was localised (54.1%). Moreover, both categories of pain were more likely to be persistent if there was associated sciatica at baseline.

DISCUSSION

In this large international study, we found that most LBP (86%) was non-localised. In comparison with localised LBP, non-localised LBP tended to be more troublesome, disabling and persistent, and showed distinctive associations with risk factors. In addition, the two categories of LBP differed markedly in their relative prevalence across the 47 occupational groups that were studied.

Apart from occupational group, all of the information that was analysed came from questionnaires. Pain, somatising tendency, mental health and health beliefs are all best assessed through self-report. However, it is possible that reliance on participants' recall led to inaccuracies in other variables such as smoking habitsand exposure to heavy lifting at work. If so, the impact on risk estimates will have depended on whether errors differed systematically according to report of pain. If they were non-differential with respect to pain, then any resultant bias will have been towards the null. On the other hand, if they varied by pain status (e.g. if participants with LBP tended to report heavy lifting more completely than those who were pain-free), then risk estimates could have been spuriously exaggerated. However, even if such biases occurred, it seems unlikely that they would have differed importantly according to whether or not LBP was localised.

A particular methodological challenge in the CUPID study was the possibility that despite our efforts to minimise errors in translation of the questionnaires, terms for pain might be understood

differently in different cultures. However, misunderstandings are less likely to have occurred in determining the anatomical location of symptoms, which was assisted by the use of diagrams. Thus, while some of the differences between occupational groups in the overall prevalence of LBP may have been a linguistic artefact, variations in the proportion of LBP that was localised are likely to be more reliable.

It seemsunlikely that the differences which we found between localised and non-localised LBP could be explained by selective participation in the study. Eligibility for inclusion depended only on participants' employment in designated jobs and being in the specified age range, and response rates were relatively high both at baseline and at follow-up. Moreover, we can think of no reason why responders should differ from non-responders differentially in relation to associations with non-localised as compared with localised LBP.

In comparison with localised LBP, non-localised LBP was more persistent and more often a cause of disability, sickness absence from work and medical consultation.This accords with the observation in a Dutch study that among industrial workers with LBP, those whose pain was disabling or had lasted for longer than three months were more likely to have musculoskeletal co-morbidity [14], although in that investigation rates of sickness absenceand medical care-seeking were only marginally higher in subjects whose LBP was accompanied by pain in the upper extremity. Also, in a community-based Norwegian investigation, functional ability was better among participants with localised LBP than in those who reported LBP as part of widespread pain [12]. These differences may occur because people who report pain at multiple sites have a generally lower threshold for awareness and intolerance of symptoms.

Before performing our analysis, we speculated that sudden onset and associated sciatica might be indications that LBP arises from acute injury or other localised spinal pathology, and therefore

would be more common among people with localised LBP. However, we found no evidence for such a relationship. On the contrary, sciatica was more prevalent among participants with non-localised LBP than in those whose LBP was localised.

Previous analysis of the CUPID dataset has indicated that multi-site musculoskeletal pain is more common in women than men, and at older ages [15]. It is therefore unsurprising that non-localised LBP showed similar associations. In marked contrast, however, localised LBP was more frequent among men than women, and tended to have higher prevalence at younger ages. This is consistent with findings from a community-based survey in Norway [12].

After adjustment for sex and age, both localised and non-localised LBP were associated with smoking, heavy lifting, somatising tendency, poor mental health, adverse beliefs about occupational causation and the prognosis of LBP, and less clearly with some psychosocial aspects of work (Table 2). Because the analysis was cross-sectional, these associations cannot necessarily be interpreted as causal, although they are consistent with findings from other studies [3,5-8,19,20]. Of greater interest are the differences in the strength of the relationships according to whether LBP was localised or associated with pain at other anatomical sites. As well as somatising tendency, poor mental health and several psychosocial aspects of work showed significantly stronger associations with non-localised LBP. This could occur if the psychological risk factors were associated with proneness to pain more generally, and not specifically in the low back.

We are aware of only one other study that has compared the epidemiology of localised and non-localised LBP [12], and that did not investigate multiple risk factors as we have done. However, a prospective cohort study in Germany of patients whoconsulted general practitionerswith chronic LBP, but in whom pain was not at the time widespread, found that transition to chronic widespread pain at follow-up was associated with female sex and a high rate of psychosomatic symptoms

[21,22]. Non-localised LBP, as we defined it, would not necessarily be classed as chronic widespread pain – the pain may have occurred at only one other anatomical site in addition to the low back, and may have been only short-lived. Moreover, we do not know whether the onset of pain in the low backpreceded or followed that at other anatomical sites. Nevertheless, our observation that non-localised LBP was differentially associated with female sex and somatising tendency is consistent with the results of the German study.

When the risk factors in Table 2 were taken into account, there were also marked differences in the relative prevalence of localised and non-localised LBP by occupational group. Thus the proportion of LBP that was localised varied from zero in New Zealand postal workers to 76.4% among sugar cane cutters in Brazil, with a tendency to be lower when the overall prevalence of LBP was higher (Figure 2). This again is an indication that localised LBP is epidemiologicallydistinct.

Our study sample was limited to men and women in employment, and we cannot be certain that the differences which were found between localised and non-localised LBP in severity, associations with risk factors, and prognosis, would be the same in all populations. However, their observation in a large sample of workers from 18 countries across five continents is sufficient to demonstrate that potentially important epidemiological differences do occur.Thissuggests that where possible, epidemiological studies on the causes andprognosis of LBPshould distinguish pain that is limited to the low back from that which occurs in association with pain at other anatomical locations.

ACKNOWLEDGEMENTS

We thank: Pietro Muñoz, PatricioOyos, Gonzalo Albuja, MaríaBelduma and Francisco Lara for their assistance with data collection in Ecuador; PatricaMonge, MelaniaChaverrri and Freddy Brenes, who helped with data collection in Costa Rica; Aurora Aragón, Alberto Berríos, Samaria Balladaresand Martha Martínez who helped with data collection in Nicaragua; Alfredo José Jirón

who assisted with data entry in Nicaragua; Catalina Torres for translation and piloting of the questionnaire in Spain; Ben and Marie Carmen Coggon for back translation of the Spanish questionnaire; Cynthia Alcantara, Xavier Orpella, Josep Anton Gonzalez, Joan Bas, Pilar Peña, Elena Brunat, Vicente San José, Anna Sala March, Anna Marquez, Josefina Lorente, Cristina Oliva, Montse Vergara and Eduard Gaynés for their assistance with data collection in Spain; NataleBattevi,Lorenzo Bordini, Marco Conti and Luciano Riboldi who carried out data collection in Italy; Paul Maurice Conway for back translation of the Italian questionnaire; TuuliSirk who helped with data collection in Estonia; Asad Ali Khan for supervision of data collection and checking in Pakistan; Khalil Qureshi for training of field workers and supervision of data collection and checking in Pakistan; and Masami Hirai, Tatsuya Isomura, Norimasa Kikuchi, Akiko Ishizuka and Takayuki Sawada for their help with data collection and management in Japan.

We are particularly grateful to the Colt Foundation, which funded data collection in Brazil, Ecuador, Costa Rica, Nicaragua, UK, Greece, Estonia, Lebanon, Pakistan and South Africa; all of the organisations that allowed us to approach their employees; and all of the workers who kindly participated in the study.

References

1. Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH et al. Disabling musculoskeletal pain in working populations: Is it the job, the person or the culture? Pain 2013;154:856-63.
2. Endean A, Palmer KT, Coggon D. Potential of MRI findings to refine case definition for mechanical low back pain in epidemiological studies: A systematic review. Spine2011;36:160-9.
3. Lötters F, Burdorf A, Kuiper J, Miedema H. Model for the work-relatedness of low back pain. ScandJWork Environ Health 2003;29:431-40.
4. Driessen MT, Proper KI, van Tulder MW, Anema JR, Bongers PM, van der Beek AJ. The effectiveness of physical and organisational ergonomic interventions on low back pain and neck pain: a systematic review. Occup Environ Med 2010;67:277-85.
5. Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back pain. Spine 2002;5:E109-E120.
6. Ramond A, Bouton C, Richard I, Roquelaire Y, Baufreton C, Legrand E, Huez JF. Psychosocial risk factors for chronic low back pain in primary care – a systematic review. Fam Pract 2011;28:12-21.
7. Vargas-Prada S, Serra C, Martinez J, Ntani G, Delclos G, Palmer K, Coggon D, Benavides F. Psychological and culturally-influenced risk factors for the incidence and persistence of low back pain and associated disability in Spanish workers: findings from the CUPID study. Occup Environ Med 2013;70:57-62.
8. Lang J, Ochsmann E, Kraus T, Lang JW. Psychosocial work stressors as antecedents of musculoskeletal problems: a systematic review and meta-analysis of stability-adjusted longitudinal studies. SocSci Med 2012;75:1163-74.

9. Palmer KT, Reading I, Linaker C, Calnan M, Coggon D. Population-based cohort study of incident and persistent arm pain: role of mental health, self-rated health and health beliefs. Pain 2008;136:30-37.

10. Palmer KT, Reading I, Calnan M, Linaker C, Coggon D. Does knee pain in the community behave like a regional pain syndrome? Prospective cohort study of incidence and persistence. Ann Rheum Dis 2007;66:1190-1194.

11. Vargas-Prada S, Coggon D. Psychological and psychosocial determinants of musculoskeletal pain and associated disability. Best Pract Res ClinRheumatol 2015;29:374-90.

12. Natvig B, Bruusgaard D, Eriksen W. Localised low back pain and low back pain as part of widespread musculoskeletal pain: two different disorders? A cross-sectional population study. J Rehab Med 2001;33:21-5.

13. Haukka E, Leino-Arjas P, Solovieva S, Ranta R, Viikari-Juntura E, Riihimäki H. Co-occurrence of musculoskeletal pain among female kitchen workers. Int Arch Occup Environ Health 2006;80:141-8.

14. IJzelenberg W, Burdorf A. Impact of musculoskeletal co-morbidity of neck and upper extremities on healthcare utilisation and sickness absence for low back pain. Occup Environ Med 2004;61:806-10.

15. Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH et al. Patterns of multisite pain and associations with risk factors. Pain 2013;154:1769-77.

16. Coggon D, Ntani G, Palmer KT et al. The CUPID (Cultural and Psychosocial Influences on Disability) Study: Methods of Data Collection and Characteristics of Study Sample. PLoS ONE 2012;7:1-12.

17. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983;13:595-605.

18. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). Med Care 1992;30:473-83.

19. Shiri R, Karppinen J, Leino-Arjas P, Solovieva S, Viikari-Juntura E. The association between smoking and low back pain: a meta-analysis. Am J Med 2010;123:87.e7-87.e35.

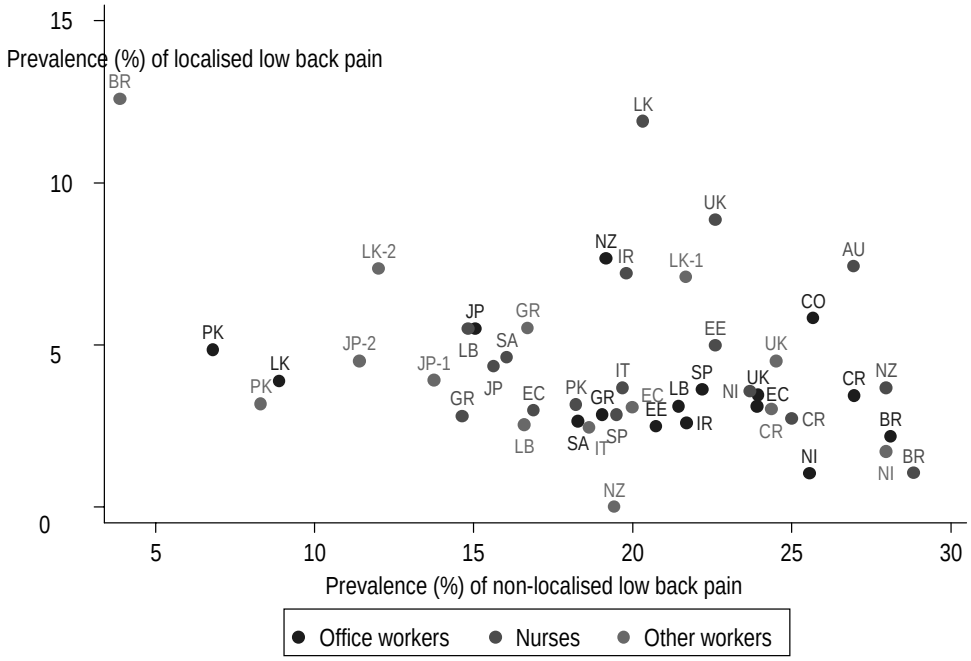
20. Main CJ, Foster N, Buchbinder R. How important are back pain beliefs and expectations for satisfactory recovery from back pain? Best Pract Res ClinRheumatol 2010;24:205-17.

21. Viniol A, Jegan N, Leonhardt C et al. Study protocol: Transition from localized low back pain to chronic widespread pain in general practice: Identification of risk factors, preventive factors and key elements for treatment – a cohort study. BMC Musculoskeletal Disorders 2012;13:77.

22. Viniol A, Jegan N, Brugger M et al. Even worse – risk factors and protective factors for transition from chronic localized low back pain to chronic widespread pain in general practice: a cohort study. Spine 2015;40:E890-9.

Figure 1 One-month prevalence of localised and non-localised low back pain by occupational group

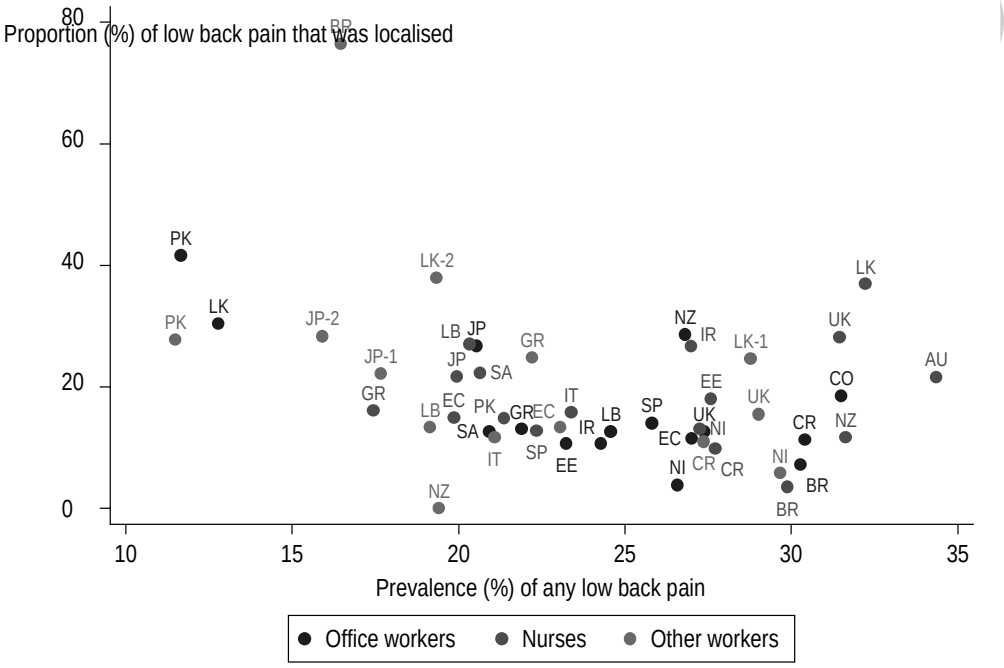
Prevalence rates are adjusted for all of the risk factors in Table 2



Key to countries: AU Australia; BR Brazil; CO Colombia; CR Costa Rica; EC Ecuador; EE Estonia; GR Greece; IR Iran; IT Italy; JP Japan; LB Lebanon; LK Sri Lanka; NI Nicaragua; NZ New Zealand; PK Pakistan; SA South Africa; SP Spain; UK United Kingdom

Figure 2 Proportion of low back pain that was localised according to overall prevalence of low back pain in each occupational group

Prevalence rates are adjusted for all of the risk factors in Table 2



Key to countries: AU Australia; BR Brazil; CO Colombia; CR Costa Rica; EC Ecuador; EE Estonia; GR Greece; IR Iran; IT Italy; JP Japan; LB Lebanon; LK Sri Lanka; NI Nicaragua; NZ New Zealand; PK Pakistan; SA South Africa; SP Spain; UK United Kingdom

Table 1 Characteristics of localised and non-localised low back pain

Characteristic	Localised low back pain (n = 609)			Non-localised low back pain (n = 3,820)		
	N	%	(95%CI)	N	%	(95%CI)
Sciatica in past month	183	30.0	(26.4,33.9)	1,836	48.1	(46.5,49.7)
Sciatica in past 12 months	233	38.3	(34.4,42.3)	2,238	58.6	(57.0,60.2)
Total duration in past month						
1-6 days	369	60.6	(56.6,64.5)	2,067	54.1	(52.5,55.7)
1-2 weeks	123	20.2	(17.1,23.6)	783	20.5	(19.2,21.8)
>2 weeks	112	18.4	(15.4,21.7)	947	24.8	(23.4,26.2)
Not known	5	0.8		23	0.6	
Total duration in past 12 months						
1-6 days	180	29.6	(26.0,33.4)	740	19.4	(18.1,20.7)
1-4 weeks	263	43.2	(39.2,47.2)	1,661	43.5	(41.9,45.1)
1-12 months	162	26.6	(23.1,30.3)	1,403	36.7	(35.2,38.3)
Not known	4	0.7		16	0.4	
Disabling in past month	288	47.3	(43.3,51.3)	2,447	64.1	(62.5,65.6)

Led to medical consultation in past 12 months	255	41.9	(37.9,45.9)	1,974	51.7	(50.1,53.3)
Attributed sickness absence in past 12 months (days)						
0	475	78.0	(74.4,81.2)	2,707	70.9	(69.4,72.3)
1-5	83	13.6	(11.0,16.6)	674	17.6	(16.4,18.9)
6-30	29	4.8	(3.2,6.8)	238	6.2	(5.5,7.0)
>30	10	1.6	(0.8,3.0)	85	2.2	(1.8,2.7)
Not known	12	2.0		116	3.0	
Onset of most recent episode						
Sudden while at work	167	27.4	(23.9,31.2)	1,176	30.8	(29.3,32.3)
Sudden not while at work	110	18.1	(15.1,21.4)	530	13.9	(12.8,15.0)
Gradual	318	52.2	(48.2,56.2)	2,015	52.7	(51.2,54.3)
Not known	14	2.3		99	2.6	

Table 2 **Associations of localised and non-localised low back pain with personal and occupational risk factors**

Risk factor	No low back pain in past 12 months	Localised low back pain			Non-localised low back pain		
	N	N	PRR[a]	(95%CI)	N	PRR[a]	(95%CI)
Sex							
Male	2,265	292	1		943	1	
Female	3,236	317	0.8	(0.6,0.9)	2,877	***1.2	(1.1,1.3)
Age (years)							
20-29	1,502	175	1		783	1	
30-39	1,737	208	1.0	(0.8,1.2)	1,189	**1.1	(1.1,1.2)
40-49	1,446	147	0.9	(0.7,1.1)	1,203	***1.2	(1.1,1.4)
50-59	816	79	0.9	(0.7,1.1)	645	***1.2	(1.1,1.4)

Smoking							
Never smoked	3,631	339	1		2,349	1	
Ex-smoker	727	91	1.3	(1.0,1.7)	579	1.1	(1.1,1.2)
Current smoker	1,124	176	1.3	(1.0,1.7)	885	1.1	(1.1,1.3)
Not known	19	3			7		
Activity in average working day							
Lifting weights ≥25 kg	1,684	266	1.4	(1.2,1.7)	1,599	1.2	(1.1,1.3)
Psychosocial aspects of work							
Work for >50 hours per week	1,394	176	1.0	(0.8,1.3)	601	†1.0	(0.9,1.1)
Time pressure at work	3,948	456	1.0	(0.8,1.2)	3,046	*1.2	(1.1,1.3)
Incentives at work	1,605	168	0.9	(0.7,1.1)	1,054	1.0	(0.9,1.1)

Lack of support at work	1,104	126	1.0	(0.8,1.3)	1,190	**1.1	(1.0,1.2)
Job dissatisfaction	1,087	128	0.9	(0.8,1.1)	817	1.0	(0.9,1.2)
Lack of job control	1,136	134	1.1	(0.9,1.3)	864	1.0	(1.0,1.1)
Job insecurity	1,652	220	1.1	(1.0,1.3)	1,277	1.1	(1.0,1.2)
Number of distressing somatic symptoms in past week							
0	3,871	406	1		1,631	1	
1	983	127	1.2	(1.0,1.5)	943	***1.4	(1.3,1.5)
2+	596	70	1.1	(0.9,1.4)	1,200	***1.7	(1.5,1.8)
Missing	51	6			46		
Mental health							
Good	2,417	225	1		1,137	1	
Intermediate	1,628	181	1.1	(0.9,1.3)	1,157	1.2	(1.1,1.3)

Poor	1,418	198	1.2	(1.0,1.5)	1,504	**1.4	(1.3,1.5)
Missing	38	5			22		
Adverse beliefs about back pain							
Work-relatedness	1,472	215	1.3	(1.1,1.5)	1,617	*1.3	(1.2,1.3)
Physical activity	999	119	0.9	(0.7,1.1)	669	0.9	(0.9,1.0)
Prognosis	598	86	1.2	(1.0,1.4)	709	**1.2	(1.1,1.3)

[a]Prevalence rate ratios relative to no low back pain in past 12 months derived from a single Poisson regression model for each pain outcome, with random intercept modelling to allow for clustering by occupational group

*Risk significantly higher for non-localised when compared directly with localised low back pain ($p<0.05$)

**Risk significantly higher for non-localised when compared directly with localised low back pain ($p<0.01$)

***Risk significantly higher for non-localised when compared directly with localised low back pain ($p<0.001$)

†Risk significantly lower for non-localised when compared directly with localised low back pain ($p<0.01$)

Table 3 One-month prevalence of low back pain at follow-up according to localisation of low back pain at baseline

Analysis was restricted to the 9,188 cases with satisfactory information about low back pain at follow-up

Category of low back pain at baseline	Number of cases at baseline	Low back pain in past month at follow-up	
		Number of cases	Prevalence % (95%CI)
Localised with no sciatica in past 12 months	282	144	51.1 (45.1,57.0)
Localised with sciatica in past 12 months	158	94	59.5 (51.4,67.1)
All localised low back pain	440	238	54.1 (49.3,58.8)
Non-localised with no sciatica in past 12 months	1,199	718	59.9 (57.0,62.6)
Non-localised with sciatica in past 12 months	1,695	1,181	69.7 (67.4,71.8)
All non-localised low back pain	2,894	1,899	65.6 (63.8,67.4)

ORIGINAL ARTICLE

Development of the Japanese Version of the Leeds Assessment of the Neuropathic Symptoms and Signs Pain Scale: Diagnostic Utility in a Clinical Setting

Tatsuya Isomura, PhD, MSc*,†; Masahiko Sumitani, MD, PhD‡; Ko Matsudaira, MD, PhD§; Mika Kawaguchi, MSc*; Reo Inoue, MD‡; Jun Hozumi, MD‡; Takeyuki Tanaka, MD¶; Hirofumi Oshima, MD**; Kanto Mori, MD¶; Shuji Taketomi, MD††; Hiroshi Inui, MD††; Keitaro Tahara, MD††; Ryota Yamagami, MD††; Kazuhiro Hayakawa‡‡

*Clinical Study Support Inc., Nagoya; †Institute of Medical Science, Tokyo Medical University, Tokyo; ‡Department of Pain and Palliative Medicine, The University of Tokyo Hospital, Tokyo; §Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo Hospital, Tokyo; ¶Orthopaedic Surgery in Sensory and Motor System Medicine, Division of Surgical Science, Graduate School of Medicine, The University of Tokyo, Tokyo; **Division of Science for Joint Reconstruction, Graduate School of Medicine, The University of Tokyo, Tokyo; ††Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo; ‡‡Pfizer Japan Inc., Tokyo, Japan

■ **Abstract**

Objective: We aimed to assess the diagnostic utility of the linguistically validated Japanese version of the Leeds Assessment of Neuropathic Symptoms and Signs Pain Scale (LANSS-J) as a screening tool for neuropathic pain in the clinical setting.

Methods: Patients with neuropathic pain or nociceptive pain who were 20 to 85 years of age were included. Sensitivity and specificity using the original cutoff value of 12 were assessed to evaluate the diagnostic utility of the LANSS-J. Sensitivity and specificity with possible cutoff values were calculated, along with area under the receiver operating characteristic curve. We then evaluated agreement regarding assessment of the LANSS-J by two investigators. We used the intraclass correlation coefficient (ICC) for the total score and Cohen's kappa coefficient for each item.

Results: Data for patients with neuropathic pain (n = 30) and those with nociceptive pain (n = 29) were analyzed. With a cutoff of 12, the sensitivity was 63.3% (19/30) and the specificity 93.1% (27/29). Sensitivity improved substantially with a cutoff of $\leq$ 11 ($\geq$ 83.3%, 25/30). High specificity (93.1%, 27/29) was sustained with a cutoff of 9 to 12. The ICC for the total score was 0.85, indicating sufficient agreement. Kappa coefficients ranged from 0.68 to 0.84.

Conclusions: The LANSS-J is a valid screening tool for detecting neuropathic pain. Our results suggest that

Address correspondence and reprint requests to: Tatsuya Isomura, MSc, Clinical Study Support Inc., 2F Daiei Bldg., 1-11-20 Nishiki, Naka-ku, Nagoya 460-0003, Japan. E-mail: tatsuya_isomura@jp-css.com.

Submitted: May 19, 2016; Revised August 1, 2016;
Revision accepted: September 4, 2016
DOI. 10.1111/papr.12528

Pain Practice, Volume ••, Issue •, 2016 ••–••

employing the original cutoff value provides high specificity, although a lower cutoff value of 10 or 11 (with its high specificity maintained) may be more beneficial when pain attributed to neuropathic mechanisms is suspected in Japanese patients. ■

Key Words: neuropathic pain, Leeds Assessment of Neuropathic Symptoms and Signs Pain Scale, Japanese version, screening tools, diagnostic utility

INTRODUCTION

Neuropathic pain is defined as "pain caused by a lesion or disease of the somatosensory nervous system."[1] Neuropathic pain negatively affects physical functioning, emotional functioning (eg, depression, anxiety), sleep, and role and social functioning.[2] Unsurprisingly, health-related quality of life is lower in patients with chronic neuropathic pain than in those with chronic non-neuropathic pain.[3,4]

Although an appropriate diagnosis is essential for successful management of neuropathic pain, the diagnosis is challenging because neuropathic pain often coexists with other types of pain and symptoms.[5] Neuropathic pain mostly presents at and is managed in a primary care setting or in hospital clinics by nonspecialists.[6] Hence, a reliable, quick screening tool could help these nonspecialists identify patients with neuropathic pain.

The Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) Pain Scale was developed as a screening questionnaire to differentiate patients with neuropathic pain from those with nociceptive pain.[7] It consists of seven items: five items for assessing pain and two items for sensory examination. For pain assessment, clinicians interview patients with questions on dysesthesia, autonomic dysfunction, evoked pain, paroxysmal pain, and thermal pain. For the sensory examination, the clinician tests for the presence of allodynia and for an altered pinprick threshold (PPT). The total score (sum of the 7-item scores) ranges from 0 to 24 points. A total score of ≥ 12 indicates that neuropathic mechanisms are likely contributing to the patient's pain.[7]

The original English language version of the LANSS Pain Scale is known to have high diagnostic accuracy.[7,8] It has been translated and widely used in several languages, including Turkish, Spanish, Swedish, Chinese, and Brazilian Portuguese.[9–14] In a previous study, we translated the LANSS Pain Scale into Japanese (LANSS-J) and validated it linguistically, after obtaining development permission from the original developer, Dr. Michael I. Bennett.[15] Its diagnostic utility as a screening tool, however, has not yet been assessed. Therefore, in this study we evaluated the diagnostic utility of the LANSS-J to determine whether it can be used as a screening tool in the clinical setting in Japan.

METHODS

The study was approved by the Ethics Committee (an investigational review board) of the University of Tokyo in January 2015. Written informed consent was obtained from each eligible participant.

Participants

Patients with neuropathic pain or nociceptive pain who were 20 to 85 years of age were included in this study. Patients with neuropathic pain were included only if their chief complaint was diagnosed as pain of neuropathic origin, which included diabetic peripheral neuropathy, postherpetic neuralgia, trigeminal neuralgia, and postchemotherapy neuropathy. Patients with nociceptive pain were included only if their chief complaint was diagnosed as nociceptive pain, such as osteoarthritis of the knee or hip.

Patients with neuropathic pain were excluded if they had clear comorbidity-related nociceptive symptoms, including bruises or joint pain derived from osteoarthritis. Patients with nociceptive pain were excluded if they had clear comorbidity-related neuropathic symptoms such as that derived from diabetes under treatment, intervertebral disk herniation (positive straight leg raising test, $< 70°$), or lumbar spinal stenosis (positive Kemp test). Patients who had mixed pain, a psychiatric disorder, dementia, fever, or menstrual pain, those who were incapable of understanding and completing the questionnaires by themselves, and those who were deemed inappropriate for participation by the investigators were also excluded.

Data Collection

After ethical approval was obtained, participants were recruited from March through July 2015 at two departments of the University of Tokyo Hospital. Patients with neuropathic pain were recruited at the Department of Anesthesiology and Pain Relief Center and patients with nociceptive pain at the Department of Orthopaedic Surgery and Spinal Surgery.

The attending doctor (primary investigator) interviewed each patient to assess pain using the LANSS-J. The patient then completed two self-administered questionnaires: the Japanese version of the painDETECT questionnaire (PDQ-J) and the EuroQol 5 Dimension (EQ-5D). Another doctor (secondary investigator) then administered the same LANSS-J to the patient on the same day.

The primary investigator collected the demographic and clinical characteristics of each patient. For the original LANSS Pain Scale, the PPT was assessed using a 23-gauge needle. In this study, however, a partially extended paper clip (instead of the needle) was used to avoid injuring the skin.

The PDQ-J is a reliable, valid screening tool for identifying neuropathic pain. It was originally developed in Germany to detect neuropathic pain components in patients with chronic low back pain.[16] A total PDQ-J score ranges from 0 to 38. Scores of ≤ 12 indicate that it is unlikely that neuropathic pain is present. Scores of ≥ 19 indicate that it is highly likely that neuropathic pain is present.

The EQ-5D is a 5-item, self-administered questionnaire that provides a single index value for the general health status of the respondent.[17] The Japanese version of the EQ-5D has been widely used in research. The index score produced by conversion of the assessed health status ranges from −0.11 to 1.00. A score of 1 indicates "perfect health," and a score of 0 indicates "death."

Statistical Analysis

We performed descriptive analyses of demographic and clinical characteristics of patients. Summary statistics on age, sex, diagnosis, time since diagnosis, body mass index (BMI), PDQ-J score, and EQ-5D score were calculated for patients with neuropathic pain and those with nociceptive pain.

The sensitivity and specificity of the LANSS-J were assessed using data collected by the primary investigators to evaluate its diagnostic utility. The scoring method of the original LANSS Pain Scale was utilized. Using the same cutoff value as for the original LANSS Pain Scale, the sensitivity and specificity of the LANSS-J were computed. Sensitivity was the percentage of patients with a LANSS-J score of ≥ 12 among those with a diagnosis of neuropathic pain. Specificity was the percentage of patients with a LANSS-J score of < 12 among those with a diagnosis of nociceptive pain, along with the area under the receiver operating characteristic curve (AUC). To assess changes in the screening results of the LANSS-J that depended on a cutoff value, we calculated the sensitivity and specificity with possible cutoff values and AUCs.

Subsequently, the intraclass correlation coefficient (ICC) value for the LANSS-J total score was calculated to evaluate agreement of the assessments by the primary and secondary investigators using data collected on the same day. In addition, Cohen's kappa was calculated for each item to assess agreement of the assessments by investigators. An ICC of ≥ 0.7 was considered the minimum required.[18] The kappa coefficients were interpreted according to the following criteria: poor, < 0.20; fair, 0.21 to 0.40; moderate, 0.41 to 0.60; good, 0.61 to 0.80; and very good, 0.81 to 1.00.[19,20]

All of the statistical tests were 2-sided, with a significance level of 0.05. All analyses were performed using SAS software version 9.3 (SAS Institute, Inc. Cary, NC, USA).

RESULTS

A total of 60 patients were included in the study. Among them, one patient had missing responses to the LANSS-J. Therefore, our final study group was composed of 59 patients (Figure 1). In all, 30 patients (50.8%) were diagnosed with neuropathic pain and 29 patients (49.2%) with nociceptive pain while waiting for knee or hip replacement surgery. Demographic and clinical characteristics of the patients are summarized by pain type in Table 1. The neuropathic pain patients were younger, included more men, had a longer interval since diagnosis, and had a higher average PDQ-J score than the nociceptive pain group. BMI did not differ greatly between the two groups. Detailed etiologies of the diagnoses are shown in Table 2.

Employing the cutoff value of 12, as suggested by the original developer, the sensitivity of the LANSS-J for diagnosing neuropathic pain was 63.3% (19/30), and the specificity of the scale for diagnosing nociceptive pain was 93.1% (27/29) (Table 3). The AUC for the cutoff value was 0.782.

Table 3 shows the sensitivity, specificity, and AUC for each possible LANSS-J cutoff value. The sensitivity of the scale substantially improved with a cutoff value of 11 (83.3%, 25/30), whereas the specificity was unchanged using cutoff values of 12 descending to 9 (93.1%, 27/29).

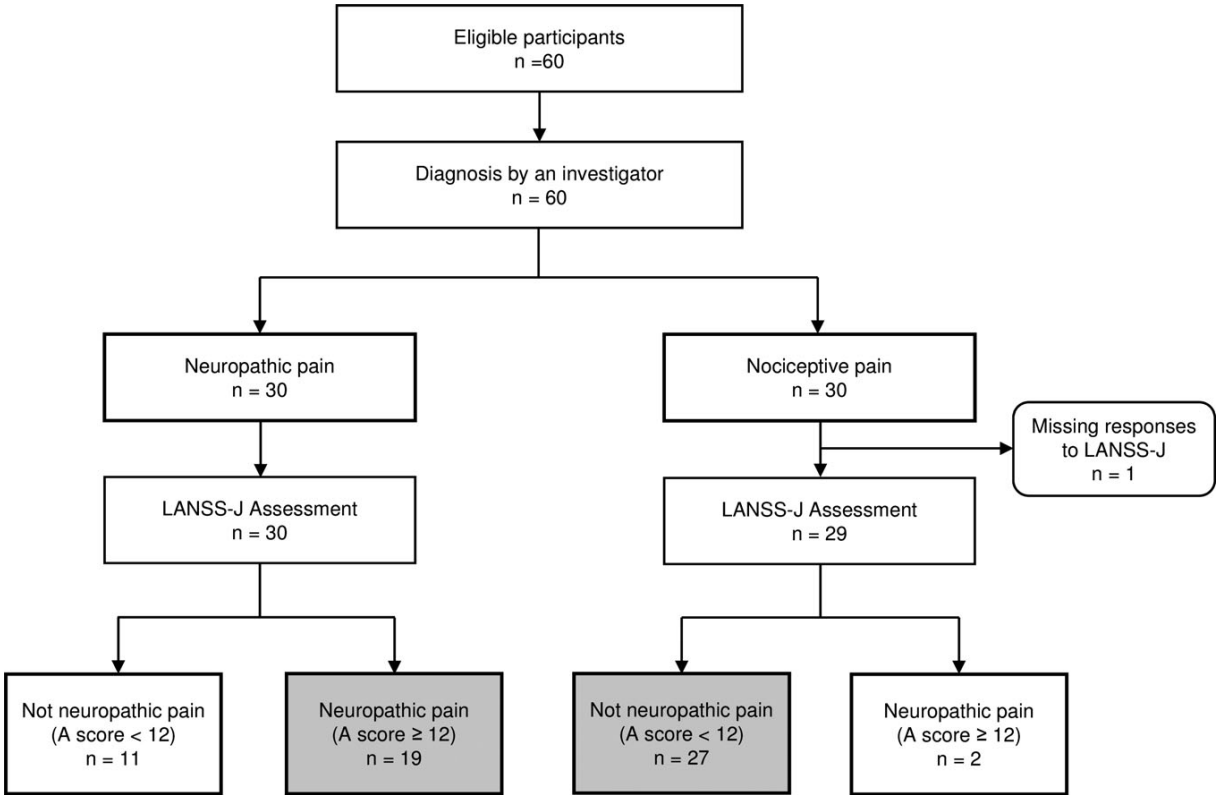


Figure 1. Flow diagram of participants in the study. The diagnostic flow diagram shows the case when using a cutoff value of 12 for the Japanese version of the Leeds Assessment of Neuropathic Symptoms and Signs Pain Scale (LANSS-J). A LANSS-J score of ≥ 12 indicates the probable pain originating from neuropathic mechanisms. A score of < 12 indicates the probable pain not originating from neuropathic mechanisms.

Table 1. Demographic and Clinical Characteristics of Participants, by Pain Type

Characteristics	Neuropathic Pain Group (*n* = 30)	Nociceptive Pain Group (*n* = 29)
Mean years of age (SD)	56.4 (12.9)	70.5 (8.6)
Female, *n* (%)	10 (33.3)	26 (89.7)
Mean months since diagnosis (SD)	94.4 (89.9)	67.6 (50.8)
Mean BMI (kg/m^2) (SD)	23.2 (3.7)	25.0 (4.1)
Mean PDQ-J score (SD)	18.1 (5.4)	5.0 (5.6)
Mean EQ-5D score (SD)	0.4 (0.3)	0.6 (0.1)

BMI, body mass index, PDQ-J, Japanese version of the painDETECT Questionnaire; EQ-5D, EuroQol 5 Dimension.

Agreement of the Assessments by Investigators

Agreement in the assessment of the LANSS-J by the primary and secondary investigators was reflected in an ICC of 0.85 using data from 51 patients (28 neuropathic pain patients, 23 nociceptive pain patients) in whom the assessments were conducted on the same day. The kappa coefficient for agreement between investigators for individual items was 0.71 for dysesthesia, 0.84 for autonomic dysfunction, 0.69 for evoked pain, 0.76 for paroxysmal pain, 0.80 for thermal pain, 0.68 for allodynia, and 0.80 for altered PPT.

DISCUSSION

We assessed the diagnostic utility of the LANSS-J using data collected from Japanese patients with neuropathic or nociceptive pain. The results suggest that the LANSS-J had debatable sensitivity when it employed the cutoff value used for the original LANSS Pain Scale but good specificity and agreement of the assessments.

The sensitivity using the original cutoff value was lower (63.3%) than that for the original LANSS (85.0%) or for LANSS versions in other languages (80.0% to 89.9%), whereas the specificity was higher (93.1%) than that for the original LANSS (80.0%) and Spanish LANSS (89.4%) but lower than for the Turkish (94.2%) and Chinese (97.1%) versions.[7,9,10,12] The

Table 2. Etiology of Patients' Diagnoses

Etiology	Neuropathic Pain Group (n = 30)	Nociceptive Pain Group (n = 29)
Complex regional pain syndrome II	1	
Failed back surgery syndrome	5	
Diabetic polyneuropathy	1	
Chemotherapy-induced neuropathy	1	
Traumatic radial nerve injuries	1	
Syringomyelia	1	
Cervical spondylotic myelopathy	2	
Cervical spondylotic radiculopathy	1	
Cervical radiculopathy	1	
Vascular polyneuropathy	1	
Phantom limb pain	1	
Thalamic pain	3	
Spinal cord injuries induced by metastatic tumor	1	
Diabetic neuropathy	1	
Postoperative neuropathy (mammary gland)	1	
Brachial plexus injury	7	
Brachial plexus palsy	1	
Knee osteoarthritis		15
Hip osteoarthritis		14

Table 3. Sensitivity and Specificity of Possible Cutoff Values with the AUC

Cutoff Value	Sensitivity % (n/N Patients)	Specificity % (n/N Patients)	AUC
≥ 0	100 (30/30)	0 (0/29)	0.500
≥ 3	100 (30/30)	65.5 (19/29)	0.828
≥ 6	100 (30/30)	72.4 (21/29)	0.862
≥ 9	90.0 (27/30)	93.1 (27/29)	0.916
≥ 10	86.7 (26/30)	93.1 (27/29)	0.899
≥ 11	83.3 (25/30)	93.1 (27/29)	0.882
≥ 12	63.3 (19/30)	93.1 (27/29)	0.782
≥ 13	63.3 (19/30)	93.1 (27/29)	0.782
≥ 14	63.3 (19/30)	93.1 (27/29)	0.782
≥ 15	60.0 (18/30)	93.1 (27/29)	0.766
≥ 18	46.7 (14/30)	96.6 (28/29)	0.716
≥ 21	16.7 (5/30)	96.6 (28/29)	0.566

AUC, area under the receiver operating characteristic curve.

sensitivity was lower in the present study probably because patients with severe traumatic nerve injury were included in the neuropathic pain group. Because of the severe nerve damage in these patients, evoked pain and allodynia could barely be perceived. In fact, among the 30 patients, 15 and 17 patients provided a negative response to the descriptors of evoked pain and allodynia, respectively. Traumatic nerve injury does not entail changes in the color of the skin attributable to impaired blood flow or hyperpigmentation, unlike diabetic neuropathy and postherpetic neuralgia. Among the 30 patients in the neuropathic pain group, 17 reported a negative response to the descriptors of autonomic dysfunctions. It is legitimate that a score of evoked pain in the PDQ-J was also low in the neuropathic pain group —the third lowest score in seven pain categories—but as low as the second lowest pain caused by slight pressure. Hence, lower scores for evoked pain, allodynia, and autonomic dysfunction in patients with neuropathic pain presumably resulted in the lower sensitivity.

Table 4. Agreement of LANSS-J Assessments Between the Two Investigators, by Cutoff Values

Cutoff Value	Kappa	No. of Patients Correctly Identified by the LANSS-J in Two Assessments	
		Neuropathic Pain (n = 28)	Nociceptive Pain (n = 23)
≥ 12	0.65	17	21
≥ 11	0.80	23	21
≥ 10	0.84	24	21

Cohen's kappa coefficients were computed to evaluate the agreement in screening results (either neuropathic or non-neuropathic pain) between the two assessments based on the LANSS-J total scores for a cutoff value of 12, 11, and 10.
LANSS-J, Japanese version of the Leeds Assessment of Neuropathic Symptoms and Signs Pain Scale.

Another possible explanation for the lower sensitivity is that the patients did not openly express their feelings about the pain they felt. In traditional Japanese culture, stoicism and the desire to conceal pain and emotions are expected—unlike in European and American cultures, where expressing personal feelings is encouraged.[21]

The sensitivity was lower when using the original cutoff value of 12, whereas specificity was favorable. Exploration of a possible cutoff value shows that using a value of 10 or 11 alone improved sensitivity while leaving the specificity unchanged (sensitivity 86.7% and 83.3%, respectively; specificity 93.1% for both values; AUC 0.899 and 0.882, respectively). Given that the lower sensitivity in the LANSS-J, compared with that of the LANSS in other languages, results not only from the number of patients with traumatic nerve injuries included in the present study but also from cultural influences on their verbal expression, physicians conducting screening should suspect neuropathic components in the pain in patients whose LANSS-J score is 10 or 11.

For agreement of the LANSS-J assessments by two investigators, the ICC for the total LANSS-J score exceeded the sufficient level of 0.7.[22] The Spanish and Brazilian Portuguese versions indicated relatively higher ICCs (0.92 and 0.97, respectively) than were seen in the present results.[10,13] Regarding individual items, the

kappa coefficients, ranging from 0.68 for allodynia to 0.84 for autonomic dysfunction, indicated good to very good levels of agreement,[19] which is equivalent to the results of the original LANSS Pain Scale (0.6 for dysesthesia, 0.88 for autonomic dysfunction).[7] When looking at the ICCs in each pain group, however, compared with the ICC in the neuropathic pain group, the ICC in the nociceptive pain group was lower (0.81 vs. 0.22). Therefore, we further evaluated whether score changes between the two assessments resulted in a change in the screening results by the LANSS-J or if the LANSS-J screening results remained the same for the two assessments. The results were presented in Table 4. When employing the original cutoff value of 12, the kappa coefficient was 0.65, which is regarded as a good level, with 17 of 28 patients in the neuropathic pain group remaining positive, whereas 21 of 23 patients in the nociceptive pain group remained negative.

The results of further examinations depended on the cutoff point employed. The results suggest that a lower cutoff value may be more helpful for detecting neuropathic pain in a Japanese population. Similarly, a lower cutoff value yielded a higher kappa coefficient at the almost very good level or very good level. In contrast, the number of patients in the nociceptive pain group who remained negative stayed consistent: 0.80 for the cutoff value of 11 and 0.84 for the cutoff value of 10. When using the original cutoff value, the kappa coefficient was at the same level as with the original LANSS (0.65),[7] although higher kappa coefficients were observed in the Turkish (0.84) and Spanish (0.70) LANSS versions.[9,10] Regardless of the lower ICC in the nociceptive pain group, there was a good level of agreement of the LANSS-J screening results between the two assessments. Thus, the diagnostic utility with regard to agreement in the LANSS-J screening results was determined to be reasonably good. A large sample size, however, is needed for further assessment.

There are several limitations in the present study. First, generalization of the results of the present study is limited due to the relatively small sample size collected at a single institution. We prioritized patient recruitment with accurate diagnosis, and its feasibility as accurate diagnosis is essential in diagnostic utility. Thus, the findings should be considered exploratory in nature due to the limited number of the sample size. Our sample size may have resulted in lower sensitivity using the original cutoff value of 12. However, it should be kept in mind that the result may have resulted from the inclusion of patients with severe traumatic nerve injury in the sample. In addition, recruitment was conducted in a university hospital setting. As patients visiting a university hospital may differ from those receiving nonspecialized primary care, a particular group of patients were underrepresented. Further investigations with a large sample in various settings may be needed for more generic features of the LANSS-J, especially for nonspecialized primary care settings. For these limitations on generalizability, results need to be interpreted with care. Second, patients with neuropathic pain and those with nociceptive pain were recruited separately from the Department of Anesthesiology and Pain Relief Center and the Department of Orthopaedic Surgery and Spinal Surgery for feasibility reasons. Although the investigators administered LANSS-J in an interview format, the fact that not a single investigator assessed patients with both types of pain might have influenced the results to some extent. Results may differ if the LANSS-J is administered to patients experiencing pain of unknown origin at a nonspecialist, primary care level. Third, a partially stretched out paper clip was used to test for altered PPT to avoid skin cuts and bleeding because concerns were raised about using a 23-gauge needle for the pinprick (as the original LANSS Pain Scale instructed).[23] Patients' responses toward an altered PPT may differ if a needle were utilized in the present study (although pinprick has been commonly performed with a paper clip as an alternative method worldwide). Finally, to evaluate agreement in the LANSS-J assessments, we included only data that were obtained on the same day to retain the same evaluation time period in "as stable a condition as possible" by the primary and secondary investigators deemed to be equally capable of assessing patient conditions. It should be noted, however, that conducting the same-day assessment of the LANSS-J may not guarantee the same conditions for the two assessments regardless of a good level of agreement in the LANSS-J screening results based on kappa coefficients in patients whose assessments were conducted on the same day.

CONCLUSION

The present study indicated a sufficient level of diagnostic utility for the LANSS-J, demonstrating that the Japanese version of the LANSS Pain Scale is a valid screening tool for detecting pain originating from a neuropathic mechanism. The results suggest that employing the original cutoff value of 12 provides high specificity, allowing it to filter out patients with

non-neuropathic pain. A lower cutoff value of 11 or 10 (which maintains the high specificity) may be more beneficial when evaluating Japanese patients whose pain was suspected to be caused by a neuropathic mechanism.

ACKNOWLEDGEMENTS

This study was funded by Pfizer Japan Inc. and Eisai Co., Ltd.

DISCLOSURES

This study was funded by Pfizer Japan Inc. and Eisai Co., Ltd. TI is a founder and the chief executive of Clinical Study Support, Inc. KM received the grant support including an endowed chair outside the submitted work from AYUMI Pharmaceutical Corporation, Nippon Zoki Pharmaceutical Co., Ltd., Ono Pharmaceutical Co., Ltd., Eli Lilly Japan K.K., Sumitomo Dainippon Pharma Co., Ltd., Astellas Pharma Inc., Toto Ltd., Okamura Corporation, and Eisai Co., Ltd.; honoraria for lecturing from AYUMI Pharmaceutical Corporation, Nippon Zoki Pharmaceutical Co., Ltd., Ono Pharmaceutical Co., Ltd., Pfizer Japan Inc., Shionogi & Co., Ltd., Eli Lilly Japan K.K., Astellas Pharma Inc., Hisamitsu Pharmaceutical Co., Inc., Janssen Pharmaceutical K.K., Kaken Pharmaceutical Co., Ltd., Teijin Pharma Limited, Eisai Co., Ltd., and Toto Ltd.; and advisory fees from Shionogi & Co., Ltd. These entities did not have any role in the study design; data collection, analysis, and interpretation; manuscript writing; and/or decision to submit for publication. MK is an employee of Clinical Study Support, Inc. HO is employed by KYOCERA Medical Corporation. ST has received grants from Smith & Nephew, Zimmer Biomet Holdings, and Stryker and has received personal fees from Smith & Nephew, Depuy Synthes, and Arthrex, outside the submitted work, and is also an editorial board member of Journal of Orthopaedic Science. HI has received personal fees from Smith & Nephew, Stryker, and Zimmer-Biomet, outside the submitted work. KH is an employee of Pfizer Japan Inc. MS, RI, JH, TT, KM, KT and RY have no financial relationships to disclose.

REFERENCES

1. International Association for the Study of Pain. *IASP Taxonomy*. Washington, DC: International Association for the Study of Pain; 2015. http://www.iasp-pain.org/Taxonomy#Neuropathicpain (accessed February 9, 2016)

2. Jensen MP, Chodroff MJ, Dworkin RH. The impact of neuropathic pain on health-related quality of life: review and implications. *Neurology*. 2007;68:1178–1182.

3. Smith BH, Torrance N, Bennett MI, Lee AJ. Health and quality of life associated with chronic pain of predominantly neuropathic origin in the community. *Clin J Pain*. 2007;23:143–149.

4. Torrance N, Smith BH, Lee AJ, Aucott L, Cardy A, Bennett MI. Analysing the SF-36 in population-based research. A comparison of methods of statistical approaches using chronic pain as an example. *J Eval Clin Pract*. 2009;15:328–334.

5. Dworkin RH, O'Connor AB, Backonja M, et al. Pharmacologic management of neuropathic pain: evidence-based recommendations. *Pain*. 2007;132:237–251.

6. Smith BH, Torrance N. Epidemiology of neuropathic pain and its impact on quality of life. *Curr Pain Headache Rep*. 2012;16:191–198.

7. Bennett M. The LANSS pain scale: the Leeds assessment of neuropathic symptoms and signs. *Pain*. 2001;92:147–157.

8. Cruccu G, Truini A. Tools for assessing neuropathic pain. *PLoS Med*. 2009;6:e1000045.

9. Yucel A, Senocak M, Kocasoy Orhan E, Cimen A, Ertas M. Results of the Leeds assessment of neuropathic symptoms and signs pain scale in Turkey: a validation study. *J Pain*. 2004;5:427–432.

10. Pérez C, Gálvez R, Insausti J, Bennett M, Ruiz M, Rejas J, Group for the study of Spanish validation of LANSS. [Linguistic adaptation and Spanish validation of the LANSS (Leeds Assessment of Neuropathic Symptoms and Signs) scale for the diagnosis of neuropathic pain]. *Med Clin (Barc)*. 2006;127:485–491.

11. Hallström H, Norrbrink C. Screening tools for neuropathic pain: can they be of use in individuals with spinal cord injury? *Pain*. 2011;152:772–779.

12. Li J, Feng Y, Han J, et al. Linguistic adaptation, validation and comparison of 3 routinely used neuropathic pain questionnaires. *Pain Physician*. 2012;15:179–186.

13. Schestatsky P, Félix-Torres V, Chaves ML, et al. Brazilian Portuguese validation of the Leeds Assessment of Neuropathic Symptoms and Signs for patients with chronic pain. *Pain Med*. 2011;12:1544–1550.

14. Kaki AM, El-Yaski AZ, Youseif E. Identifying neuropathic pain among patients with chronic low-back pain: use of the Leeds Assessment of Neuropathic Symptoms and Signs pain scale. *Regional Anesth Pain Med*. 2005;30:422–428.

15. Isomura T, Sumitani M, Matsudaira K, et al. Development of a Japanese version of the Leeds Assessment of Neuropathic Symptoms and Signs Pain Scale: translation and linguistic validation. *Pain Clin*. 2014;35:933–940.

16. Matsubayashi Y, Takeshita K, Sumitani M, et al. Validity and reliability of the Japanese version of the painDETECT questionnaire: a multicenter observational study. *PLoS One*. 2013;8:e68013.

17. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199–208.

18. Streiner D, Norman G, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 5th ed. Oxford, UK: Oxford University Press; 2014.

19. Altman D. *Practical Statistics for Medical Research*. London: Chapman and Hall/CRC; 1990.

20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.

21. Chambers D, Thompson S, Narayanasamy A. Engendering cultural responsive care: a reflective model for nurse education. *J Nurs Educ Pract*. 2013;3:70–81.

22. Fayers P, Machin D. *Quality of Life: The Assessment, Analysis and Interpretation of Patient-reported Outcomes*. 2nd ed. Chichester, UK: John Wiley & Sons; 2007.

23. Backonja M. Need for differential assessment tools of neuropathic pain and the deficits of LANSS pain scale. *Pain*. 2002;98:229–230; author reply 230–231.

J Anesth
DOI 10.1007/s00540-016-2303-1

ORIGINAL ARTICLE

Fear-avoidance beliefs are independently associated with the prevalence of chronic pain in Japanese workers

Kenta Wakaizumi[1] · Keiko Yamada[2,3] · Hiroyuki Oka[4] · Shizuko Kosugi[1] · Hiroshi Morisaki[1] · Masahiko Shibata[3,5] · Ko Matsudaira[4]

Received: 26 October 2016 / Accepted: 19 December 2016

Abstract
Purpose Pain is a global public health problem with implications for both personal and social heath. Fear-avoidance beliefs (FABs) have been demonstrated to negatively impact and prolong pain in many Western countries, but little is known about the association between FABs and chronic pain (CP) in Asian countries, including Japan. We examined the relationship between FABs and CP in Japanese white-collar workers, a growing population with a high prevalence of CP.
Methods Questionnaires and company records were used to gather data from 433 Japanese white-collar workers. Data were related to experience of pain, participant sociodemographic/health/lifestyle characteristics, fear-avoidance beliefs [Tampa Scale for Kinesiophobia (TSK)], work-related psychosocial factors (Brief Job Stress Questionnaire), and depressive illness [Psychological Distress Scale (K6)]. Analysis of covariance and multilevel logistic regression modeling were used to analyze associations between the data while controlling for factors known to influence CP prevalence.
Results Prevalence rate of CP was 11.1% (48 of 433 persons). Adjusted odds ratios for participants with CP significantly increased in participants with high TSK scores, even after adjusting for factors known to influence CP prevalence.
Conclusion We found a significant association between high TSK scores and CP in Japanese white-collar workers when controlling for other known factors that influence CP such as work-related psychosocial characteristics and depressive conditions. This finding suggests that FABs are independently associated with prevalence of CP.

Keywords Chronic pain · Fear of movement · Fear-avoidance belief · TSK · White-collar workers · Occupational medicine

✉ Kenta Wakaizumi
k.wakamail@keio.jp

[1] Department of Anesthesiology, Keio University School of Medicine, 35, Shinanomachi, Shinjyuku-ku, Tokyo 160-0016, Japan

[2] Public Health, Department of Social Medicine, Osaka University Graduate School of Medicine, Suita-Shi, Osaka, Japan

[3] Center for Pain Management, Osaka University Hospital, Suita-Shi, Osaka, Japan

[4] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

[5] Department of Pain Medicine, Osaka University Graduate School of Medicine, Suita-Shi, Osaka, Japan

Introduction

Pain is a health problem that dramatically effects the global population [1, 2]. In particular, chronic pain (CP) can impair the daily lives of its sufferers, as well as placing a substantial economic burden on a country's resources [3]. A number of large-scale surveys suggest that CP is prevalent in approximately 20–25% of the Japanese population [4, 5], with prevalence differing between different work-related occupations [6]. For example, prevalence of chronic musculoskeletal pain is typically higher within "white-collar" professional, office, and technical employees, despite lower levels of demanding physical tasks. In contrast, employees within the "blue-collar" agricultural, forestry, and fisheries industries report lower rates of such chronic pain [6].

Published online: 03 January 2017

To be able to prevent CP in white-collar workers (WCWs) would have substantial benefits in the lives of individuals who experience pain in their everyday lives. Furthermore, preventing CP effectively would also have economic and social benefits because of the association between CP and employment sick leave, absence, and poor productivity [7, 8]. The urgency to develop effective treatments and prevention strategies for CP is becoming increasingly more salient as there is a growing number of WCWs in the workforce [9], which thus increases the number of individuals experiencing CP.

Biological disorders are often considered as the primary causal factor for CP; however, it is also important to consider the effect of psychological and social factors, such as work-related stress and depression. In particular, dysfunctional beliefs relating to pain itself, and fear of pain, play a key role in the development of CP [10, 11]. Fear of pain can lead to avoidance of activities that patients associate with the occurrence or exacerbation of pain, even after physical recovery from the associated condition has occurred. The avoidance of physical activities based on fear of movement, known as kinesophobia, can also lead to a cycle of further fear and avoidance [12]. Moreover, excessive avoidance of activities believed to cause pain can reduce muscle strength and flexibility, which may delay recovery from a painful condition. This cycle of pain and avoidance behavior can be explained by a fear-avoidance model in which fear-avoidance beliefs (FABs) represent typical cognitive and emotional responses that can lead people with pain into a cycle of avoidance. Research has demonstrated that FABs contribute to long-term work-related disabilities; furthermore, low levels of FABs are a useful predictor of early recovery from an impairment [13, 14].

Although a number of studies in the United States and Europe have investigated the relationship between FABs and CP, very little research of a similar nature has taken place in Asian countries, including Japan. As the influence of FABs differs depending on culture and ethnicity [15], it is not possible to generalize the results from studies in Western populations to an Asian population. Our study aims to address this issue by investigating the prevalence of CP among Japanese WCWs, and furthermore, by analyzing the association between FABs and CP.

Methods

Data were collected from WCWs in the technology development division of a company listed on the First Section of the Tokyo Stock Exchange. Questionnaires were used to collect data relating to participants' pain, fear of pain, work-related psychosocial factors, and depression. We also collected data relating to participant demographics and lifestyle from the company's most recent employee health survey conducted within the year preceding the current study. The questionnaires were distributed to participants on February 10, 2015, and the survey was closed on February 29, 2015.

All procedures were approved by both the Keio University School of Medicine Ethics Committee (approval no. 20140296) and the Health and Safety Committee within the participating company. Participants were informed about the nature of the survey, and the use of demographic data from the annual health check, through the company's intranet.

Participants

In total, 517 full-time employees were asked to take part, with 433 returning completed questionnaires (83.8% response rate). Respondents were aged from 20 to 65 years old (mean = 41.5; standard deviation = 10.8). Male participants comprised the majority of the sample (375 participants, 86.8%). According to the company's data, all participants were recognized as WCWs who were engaged predominantly in deskwork.

Measures

Data were collected from the company's health check related to the participant's age, sex, body mass index (BMI, kg/m^2; participants categorized in quartiles), height, smoking habit (participants categorized as never, ex-smoker, or current smoker), daily alcohol intake [1 glass of sake (180 ml) was coded as 23 g ethanol; participants were categorized as consuming 0, 1–23, 24–45, or ≥46 g ethanol/day], highest education achieved (high school graduate or junior college graduate, bachelor's degree, master's degree, or doctorate), exercise routine, and daily working hours. Sleep patterns were evaluated for quantity of sleep in the past 4 weeks (participants categorized as having <5, 5, 6, 7, 8, 9, or >9 h/day) [16], with a reported sleep duration of less than 5 h coded as 'short sleep.' A measurement of participants' subjective evaluation of their exercise routine was also collected and used in the subsequent analyses.

Participants were asked to provide specific details of pain they had experienced during the previous 4 weeks relating to pain location(s), intensity, duration, and frequency. Location of pain was marked on an illustration by the participants (see Fig. 1), with multiple answers allowed. Pain intensity was scored on a numeric rating scale (NRS) comprising 11 points (0 = no pain to 10 = worst pain imaginable). Participants were coded as having CP when the following criteria were met: (1) NRS score of 5 or more, (2) pain persisted for at least 3 months, and (3) pain experienced at least two times a week [17].

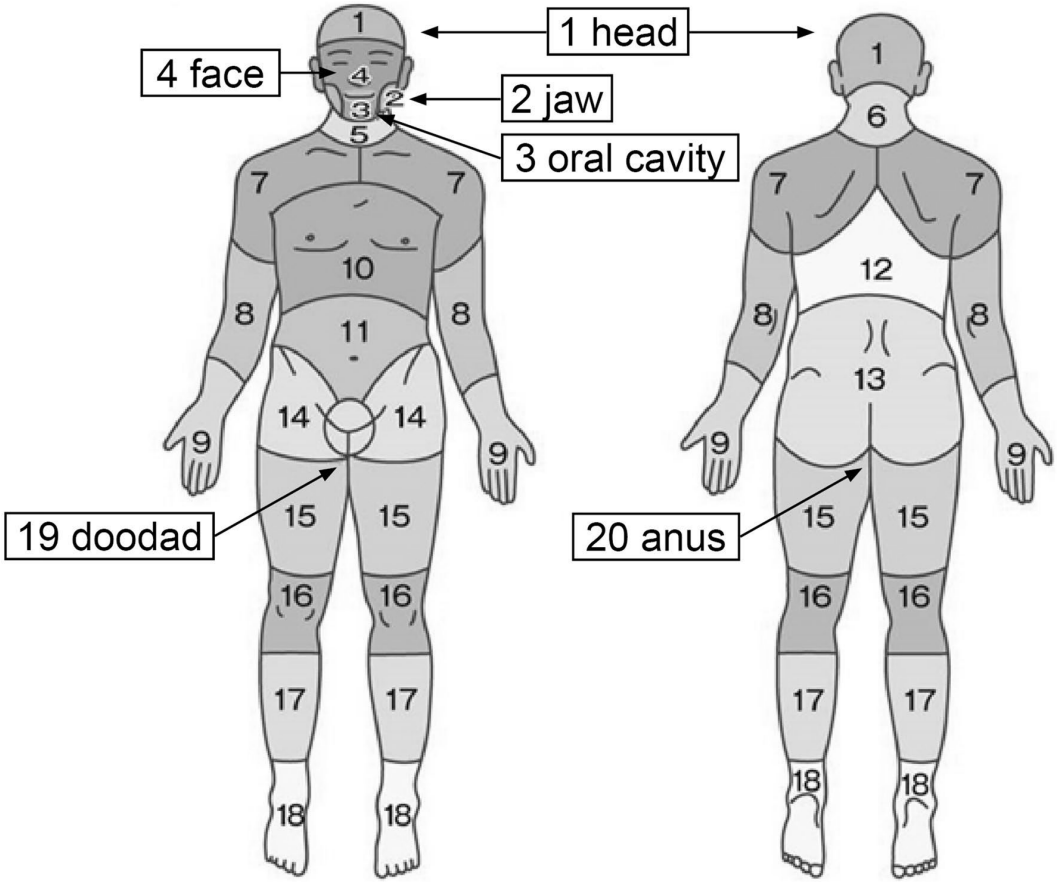


Fig. 1 Full-body manikin divided into 20 areas for marking of pain sites. Shoulder, back neck, low back, and head were defined as areas numbered *7*, *6*, *13*, and *1*, respectively

To evaluate kinesophobia within participants, we used the Japanese short version of the Tampa Scale of Kinesophobia (TSK-J11) developed and validated by Matsudaira et al. [18, 19]. Although a longer version of the TSK is available, we deemed the short version suitable for the study because of its good psychometric properties (Cronbach's $\alpha = 0.92$) and the advantage of brevity it offered [18, 20]. Reliability and validity of the TSK has been confirmed in several patient populations, including patients with chronic musculoskeletal pain [21], low back pain [22, 23], whiplash injury pain [24], shoulder pain [25], temporomandibular disorder [26], sciatica [27], and fibromyalgia [28]. The TSK-J11 comprises 11 items with each scored on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree). The total score is obtained by summing the scores for the 11 items and ranges from 11 to 44. Higher scores indicate a greater degree of kinesophobia within participants.

Work-related psychosocial factors were measured in five different aspects (job demand, job control, social support from supervisors and co-workers, and job satisfaction) using subscales of the Brief Job Stress Questionnaire (BJSQ) [29]. All items, except job satisfaction, were respectively rated on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree). Job demand was calculated by summing the item scores for psychological job overload within the BJSQ (three items). Job control was calculated by summing the item scores for subjective adjustability of work within the BJSQ (three items). The questionnaire section on social support from supervisors and co-workers consisted of three items, respectively, with the total score calculated by summing the three items, and ranging from 3 to 12 (lower scores indicating greater levels of support). Job satisfaction was rated on a 4-point scale ranging from 1 (satisfied) to 4 (unsatisfied).

Depressive condition was measured using the Kessler Psychological Distress Scale (K6). The Japanese version was developed in 2008, and then reliability and validity were confirmed by Furukawa et al. (Cronbach's α was 0.85) [30]. The K6 was developed in 2002 as a short-form

version of the K10 [31] and consists of six items related to depression and anxiety, with each rated on a 5-point scale. In accordance with Kawakami, participants with a K6 score of 10 points or more were defined as having a depressive condition.

Statistical analysis

An analysis of Dunnett's test was conducted to test for differences in the age- and sex-adjusted means and proportions of participants' demographic and lifestyle characteristics. To investigate any association between kinesophobia and chronic pain, logistic regression was conducted to calculate multivariable-adjusted odds ratios (ORs) and 95% confidence intervals (95% CI). The data were fitted to three different regression models, with each adjusting for increasing numbers of variables. Model 1 adjusted for age, sex, BMI, smoking status, daily alcohol intake, highest education achieved, exercise habit, sleeping time, and working time. Model 2 also adjusted for job demands (categorized in tertiles), job control (categorized in tertiles), social support from supervisors and co-workers (categorized in quartiles), and job satisfaction (four categories) in addition to the control variables of model 1. Model 3 adjusted for depressive condition (K6 score $\geq$10) in addition to the control variables of model 2.

p values <0.05 for two-tailed tests were considered statistically significant. All statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA).

Results

A total of 259 participants (60.0%) reported experiencing pain during the 4 weeks preceding the study, with 48 participants (11.1%) meeting the criteria for experiencing CP. The 211 participants who reported experiencing no pain in the previous 4 weeks were categorized into a non-CP group. The most frequently reported location of CP was in the shoulders (64.6%), followed by the back of the neck (54.2%), the low back (41.7%), and the head (31.3%).

Table 1 shows demographic characteristics of both the CP group and the non-CP group. Although no significant differences were observed between the groups for any of the items, there was a trend for higher BMI, poorer job control, poorer job satisfaction, and a higher rate of depressive people within the CP group when compared with the non-CP group.

Characteristics of participants who reported pain according to five categories of TSK-J11 scores are reported in Table 2. They were classified in quintiles as extremely low (Q1, 11–18), low (Q2, 19–20), intermediate (Q3, 21–23), high (Q4, 24–25), or extremely high (Q5, 26–44). Higher TSK-J11 scores were significantly associated with a greater prevalence of CP. In addition, a greater proportion of participants with high TSK-J11 scores reported poorer job control, poorer support from supervisors and co-workers, poorer job satisfaction, and a higher rate of depressive conditions (K6 score $\geq$10).

Table 3 illustrates the age- and sex-adjusted OR values for the CP versus non-CP groups according to TSK score. Significantly higher age- and sex-adjusted OR values of CP versus non-CP were observed among participants with extremely high TSK scores (Q5) compared to the participants with extremely low TSK scores. The OR values gradually increased for participants in the Q5 groups (3.13). In the first model, which adjusted for additional demographics and lifestyle variables, the Q5 group also exhibited significantly high OR values (3.13), with the difference remaining significant after adjusting for both the work-related psychosocial variables in model 2 (OR = 4.07) and the depressive symptom variables in model 3 (OR = 4.09). The stepwise increase in OR values with TSK score found in the columns of Q4 and Q5 also remained in all the models.

Table 1 Age- and sex-adjusted mean values and proportions of chronic pain risk factors

	Non-chronic pain	Chronic pain
n	211	48
Age, years (SE)	42.0 (0.7)	44.5 (1.5)
Men, *n*, %	185, 87.7%	40, 83.3%
Body mass index $\geq$25, *n*, %	42, 19.9%	12, 25.0%
Current smoker, *n*, %	15, 7.1%	4, 8.3%
Alcohol intake >46 g/day, *n*, %	32, 15.2%	9, 18.8%
Master's degree, *n*, %	156, 73.9%	37, 77.1%
Exercise >30 min twice a week, *n*, %	141, 66.8%	29, 60.4%
Sleep time 5 h or less, *n*, %	11, 5.2%	3, 6.3%
Working time $\geq$10 h, *n*, %	117, 55.5%	25, 52.1%
High job demands, *n*, %	66, 31.3%	17, 35.4%
Poor job control, *n*, %	71, 33.7%	23, 47.9%
Poor support from supervisor, *n*, %	65, 30.8%	16, 33.3%
Poor support from co-workers, *n*, %	59, 28.0%	11, 22.9%
Job dissatisfaction, *n*, %	51, 24.2%	15, 31.3%
K6 $\geq$10 points, *n*, %	11, 5.2%	4, 8.3%

No significant differences were observed between the groups for any of the items

SE standard error, *K6* Kessler Psychological Distress Scale

Table 2 Age- and sex-adjusted mean values and proportions of chronic pain risk factors according to the Tampa Scale for Kinesiophobia

	Tampa Scale for Kinesiophobia (TSK)				
	Q1 TSK = 11–18	Q2 TSK = 19–20	Q3 TSK = 21–23	Q4 TSK = 24–25	Q5 TSK = 26–44
n	52	37	61	38	63
Age, years (SE)	41.7 (1.5)	40.6 (1.7)	41.7(1.4)	44.2 (1.7)	43.9(1.3)
Men, *n*, %	42, 80.8%	31, 83.8%	51, 83.6%	36, 94.7%	57, 90.5%
Body mass index ≥ 25, *n*, %	11, 21.2%	5, 13.5%	14, 23.0%	6, 15.8%	16, 25.4%
Current smoker, *n*, %	2, 3.8%	4 10.8%	3, 4.9%	4, 10.5%	6, 9.5%
Alcohol intake more than 46 g/day, *n*, %	9, 17.3%	2, 5.4%	6, 9.8%	7, 18.4%	13, 20.6%
Master's degree, *n*, %	39, 75.0%	28, 75.7%	43, 70.5%	26, 68.4%	50, 79.4%
Exercise >30 min twice a week, *n*, %	38, 73.1%	21, 56.8%	44, 72.1%	30, 78.9%	33, 52.4%*
Sleep time 5 h or less, *n*, %	2, 3.8%	1, 2.7%	3, 4.9%	3, 7.9%	8, 12.7%
Working time ≥10 h, *n* (%)	33, 63.5%	20, 54.1%	35, 57.4%	23, 60.5%	27, 42.9%*
High job demands, *n*, %	24, 46.2%	9, 24.3%	17, 27.9%	12, 31.6%	17, 27.0%
Poor job control, *n*, %	8, 15.4%	16, 43.2%*	21, 34.4%	15, 39.5%*	32, 50.8%‡
Poor support from supervisor, *n*, %	9, 17.3%	14, 37.8%	13, 21.3%	15, 39.5%	28, 44.4%‡
Poor support from co-worker, *n*, %	9, 17.3%	6, 16.2%	18, 29.5%	10, 26.3%	24, 38.1%*
Job dissatisfaction, *n*, %	7, 13.5%	5, 13.5%	15, 24.6%	12, 31.6%	25, 39.7%†
K6 ≥10 points, *n*, %	3, 5.8%	4, 10.8%	10, 16.4%	7, 18.4%	15, 23.8%†
Chronic pain, *n*, %	6, 11.5%	4, 10.8%	6, 9.8%	10, 26.3%	20, 31.7%*

Test for significance from the category of Q1: * $p < 0.05$, † $p < 0.01$, ‡ $p < 0.001$

SE standard errors

Table 3 Odds ratios (ORs, 95% CI) of chronic pain versus non-chronic pain according to Tampa Scale for Kinesiophobia

	Tampa scale for kinesiophobia: TSK				
	Q1 TSK = 11–18	Q2 TSK = 19–20	Q3 TSK = 21–23	Q4 TSK = 24–25	Q5 TSK = 26–44
No. of subjects	52	37	61	38	63
No. of subjects with chronic pain	6	4	6	10	20
Age-adjusted mean values	1.00	0.80 (0.22–2.89)	0.70 (0.23–2.17)	2.43 (0.85–7.00)	3.09 (1.22–7.82)*
Model 1 OR (95% CI)	1.00	0.73 (0.18–2.89)	0.68 (0.21–2.20)	2.46 (0.82–7.42)	3.13 (1.17–8.37)*
Model 2 OR (95% CI)	1.00	0.76 (0.18–3.26)	0.65 (0.19–2.26)	2.73 (0.81–9.19)	4.07 (1.35–12.23)*
Model 3 OR (95% CI)	1.00	0.79 (0.18–3.40)	0.64 (0.18–2.22)	2.66 (0.79–8.98)	4.09 (1.35–12.42)*

Test for significance from the category of Q1: * $p < 0.05$, † $p < 0.01$, ‡ $p < 0.001$

SE standard error

Discussion

This is the first study demonstrating that kinesophobia adversely affects CP in Japanese employees. Because the biopsychosocial model makes a substantial contribution in explaining the complicated mechanisms that underpin CP [32], psychosocial factors should also be taken into account to identify the independent relationship between kinesophobia and CP. As such, we performed multiple logistic regression analyses controlling biological characteristics, work-related factors (psychosocial factors), and depression (psychological factor).

In model 1, we adjusted for demographic and lifestyle factors comprising age, sex, BMI, smoking status, daily alcohol intake, highest education achieved, exercise routine, sleeping time, and working time. Although exercise routine indicates low pain sensitivity [8, 33], high TSK scores were significantly associated with higher prevalence of CP, even after adjusting for factors including exercise. Although an exercise routine may alleviate kinesophobia and improve TSK score, the influence was statistically low in the present study.

In model 2, we investigated the effect of psychosocial factors on the relationship between kinesophobia and CP.

The demand-control model posits "high strain jobs" [jobs that combine high demand within a job and low job control (low decision latitude)] as having adverse effects on employees' health [34]. Furthermore, social support by supervisors and co-workers is argued to also be an influential psychosocial factor in the workplace [34], with the demand-control-support model suggesting that workplace support (as well as job control) can reduce a job stress induced by job demand. Another influential psychosocial factor is job satisfaction, referring to the overall well-being an employee feels toward their job [35]. Conversely, job dissatisfaction refers to the negative emotions elicited through a reciprocal deficit in effort-reward, illustrated in the imbalance model [36]. A number of studies have highlighted the importance of these factors, with poor work-related psychosocial factors associated with a higher prevalence of CP among European and North American employees [37–39], and job satisfaction significantly associated with CP in Japanese employees [40–42]. As such, we adjusted for work-related psychosocial factors in model 2 of the current study. The results indicated that, even after adjusting for work-related psychosocial factors, OR values for participants with extremely high TSK scores remained significant, suggesting that FABs influenced the prevalence of CP independently of the psychosocial factors.

In model 3, we also adjusted for depressive condition because previous research has demonstrated depression as an independent factor that adversely affects CP [42]. Job stress is also an indicator of depression [42]. As in model 2, the OR values in model 3 for participants with extremely high TSK scores remained significantly high after adjusting for depression, further supporting the view that kinesophobia is an independent risk factor of CP.

The prevalence of FABs in acute, or subacute, phases of painful conditions can prolong pain and lead to intractable conditions [43, 44]. As shown in this study, the independent association between FABs and the prevalence of CP suggests the similar negative impact of FABs on pain. Therefore, tackling FABs is an important therapeutic approach for reducing CP. The introduction of a psychosocial flag system for chronic musculoskeletal pain is one such approach and is strongly recommended in Europe and the United States. Under this system, FABs indicate a yellow flag, which requires the employee to receive treatment from clinical physicians working in collaboration with the workplace [45]. To effectively contend with CP, it is necessary for Japanese physicians to understand the complex nature of occupational health and CP and provide interventions that target pain in the earliest stages of onset.

There were a number of limitations in the current study. First, it is likely that selection bias influenced the results to some degree. The fact that the participating company positively accepted to take part in the study perhaps suggests that they have a strong interest in occupational health. However, even within a company that supports its employee's health, the present study demonstrated that more than one tenth of their employees suffered from severe chronic pain and FABs associated with CP. Unfortunately, it was impossible to infer the degree to which selection bias may have impacted the results because of the lack of other research focusing on the relationship between CP and occupational environments in Japan. More investigations will be expected to identify risk factors of CP in the work site. Second, data relating to the cause of pain were not collected. It is possible that classifications of pain may have influenced the present results; however, nociceptive and neuropathic pain classification of pain overlap in most patients who experience pain [46], which suggests that classifying pain in such a way is not necessary. Finally, because of the cross-sectional nature of the study, causality regarding the direction of influence between TSK and CP cannot be inferred. It is possible that long-lasting severe pain elevated TSK scores in participants who experienced pain. To clarify causality, future research should include panel data analysis.

In conclusion, we found a significant association between high TSK score and CP in Japanese white-collar workers when controlling for other known factors that influence CP, such as work-related psychosocial characteristics and depressive conditions. This result suggests that FABs are independently associated with the prevalence of CP.

Acknowledgements This research is partially supported by a grant for The Research Project on Elucidation of Chronic Pain from Japan Agency for Medical Research and Development, AMED (16ek0610004h0003). Kenta Wakaizumi is supported as a research assistant in part by a Grant-in-Aid for the Program for Leading Graduate School for "Science for Development of Super Mature Society" from the Ministry of Education, Culture, Sport, Science, and Technology in Japan.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to declare.

References

1. Rice AS, Smith BH, Blyth FM. Pain and the global burden of disease. Pain. 2016;157:791–6.
2. Goldberg DS, McGee SJ. Pain as a global public health priority. BMC Public Health. 2011;11:770.
3. Gaskin DJ, Richard P. The economic costs of pain in the United States. J Pain. 2012;13:715–24.
4. Matsudaira K, Kunoki J, Yamazaki R, Yamada K, Takagi A. Pain associated cross-sectional epidemiological (PACE) survey 2009 (in Japanese). Pain Clin. 2011;32:1345–56.
5. Yabuki S, Ushida T, Takeshita K, Saura R, Ogawa S, Katsumata A, Hatanaka S. A nationwide survey of chronic pain sufferers in Japan (in Japanese). Clin Orthop. 2012;47:127–34.

6. Nakamura M, Nishiwaki Y, Ushida T, Toyama Y. Prevalence and characteristics of chronic musculoskeletal pain in Japan. J Orthop Sci. 2011;16:424–32.
7. McDonald M, DiBonaventura Md, Ullman S. Musculoskeletal pain in the workforce: the effects of back, arthritis, and fibromyalgia pain on quality of life and work productivity. J Occup Environ Med. 2011;53:765–70.
8. Takura T, Ushida T, Kanchiku T, Ebata N, Fujii K, DiBonaventura Md, Taguchi T. The societal burden of chronic pain in Japan: an internet survey. J Orthop Sci. 2015;20:750–60.
9. Ministry of Internal Affairs and Communications in Japan. Labour force survey (2015 yearly average results). 2016.
10. Asmundson GJG, Vlaeyen JWS, Crombez G. Understanding and treating fear of pain. New York: Oxford University Press; 2004.
11. Leeuw M, Goossens ME, Linton SJ, Crombez G, Boersma K, Vlaeyen JW. The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. J Behav Med. 2007;30:77–94.
12. Vowles KE, Gross RT. Work-related beliefs about injury and physical capability for work in individuals with chronic pain. Pain. 2003;101:291–8.
13. Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. Pain. 2001;94:7–15.
14. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? JAMA. 2010;303:1295–302.
15. Portenoy RK, Ugarte C, Fuller I, Haas G. Population-based survey of pain in the United States: differences among white, African American, and Hispanic subjects. J Pain. 2004;5:317–28.
16. Takahashi M, Matsudaira K, Shimazu A. Disabling low back pain associated with night shift duration: sleep problems as a potentiator. Am J Ind Med. 2015;58:1300–10.
17. Dureja GP, Jain PN, Shetty N, Mandal SP, Prabhoo R, Joshi M, Goswami S, Natarajan KB, Iyer R, Tanna DD, Ghosh P, Saxena A, Kadhe G, Phansalkar AA. Prevalence of chronic pain, impact on daily life, and treatment practices in India. Pain Pract. 2014;14:51–62.
18. Kikuchi N, Matsudaira K, Sawada T, Oka H. Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury pain and/or low back pain. J Orthop Sci. 2015;20:985–92.
19. Matsudaira K, Inuzuka K, Kikuchi N, Sakae C, Arisaka M, Isomura T. Development of a Japanese version of the Tampa Scale for Kinesiophobia (TSK-J): translation and linguistic validation (in Japanese). Seikei Geka (Orthop Surg). 2013;48:13–9.
20. Woby SR, Roach NK, Urmston M, Watson PJ. Psychometric properties of the TSK-11: a shortened version of the Tampa Scale for Kinesiophobia. Pain. 2005;117:137–44.
21. Koho P, Aho S, Kautiainen H, Pohjolainen T, Hurri H. Test-retest reliability and comparability of paper and computer questionnaires for the finnish version of the Tampa Scale of Kinesiophobia. Physiotherapy. 2014;100:356–62.
22. Rusu AC, Kreddig N, Hallner D, Hülsebusch J, Hasenbring MI. Fear of movement/(re)injury in low back pain: confirmatory validation of a German version of the Tampa Scale for Kinesiophobia. BMC Musculoskelet Disord. 2014;19:280.
23. Swinkels-Meewisse EJ, Swinkels RA, Verbeek AL, Vlaeyen JW, Oostendorp RA. Psychometric properties of the Tampa Scale for Kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. Man Ther. 2003;8:29–36.
24. Bunketorp L, Carlsson J, Kowalski J, Stener-Victorin E. Evaluating the reliability of multi-item scales: a non-parametric approach to the ordered categorical structure of data collected with the Swedish version of the Tampa Scale for Kinesiophobia and the Self-Efficacy Scale. J Rehabil Med. 2005;37:330–4.
25. Mintken PE, Cleland JA, Whitman JM, George SZ. Psychometric properties of the fear-avoidance beliefs questionnaire and Tampa Scale of Kinesiophobia in patients with shoulder pain. Arch Phys Med Rehabil. 2010;91:1128–36.
26. Visscher CM, Ohrbach R, van Wijk AJ, Wilkosz M, Naeije M. The Tampa Scale for Kinesiophobia for temporomandibular disorders (TSK-TMD). Pain. 2010;150:492–500.
27. Haugen AJ, Grøvle L, Keller A, Grotle M. Cross-cultural adaptation and validation of the Norwegian version of the Tampa Scale for Kinesiophobia. Spine (Phila Pa 1976). 2008;33:E595–601.
28. Burwinkle T, Robinson JP, Turk DC. Fear of movement: factor structure of the Tampa Scale of Kinesiophobia in patients with bromyalgia syndrome. J Pain. 2005;6:384–91.
29. Shimomitsu T, Haratani T, Nakamura K, Kawakami N, Hayashi T, Hiro H, Arai M, Miyazaki S, Furuki K, Ohya Y, Odagiri Y. Final development of the Brief Job Stress Questionnaire mainly used for assessment of the individuals (in Japanese). In: Kato M, editor. The Ministry of Labor sponsored grant for the prevention of work-related illness, FY 1999 report. Tokyo: Tokyo Medical University; 2000. p. 126–64.
30. Furukawa TA, Kawakami N, Saitoh M, Ono Y, Nakane Y, Nakamura Y, Tachimori H, Iwata N, Uda H, Nakane H, Watanabe M, Naganuma Y, Hata Y, Kobayashi M, Miyake Y, Takeshima T, Kikkawa T. The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res. 2008;17:152–8.
31. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, Walters EE, Zaslavsky AM. Short screening scales to monitor population prevalences and trends in nonspecific psychological distress. Psychol Med. 2002;32:959–76.
32. den Boer JJ, Oostendorp RA, Beems T, Munneke M, Oerlemans M, Evers AW. A systematic review of bio-psychosocial risk factors for an unfavourable outcome after lumbar disc surgery. Eur Spine J. 2006;15:527–36.
33. Landmark T, Romundstad PR, Borchgrevink PC, Kaasa S, Dale O. Longitudinal associations between exercise and pain in the general population–the HUNT pain study. PLoS One. 2013;8:e65279.
34. Karasek RA. Job demands, job decision latitude, and mental strain: implications for job design. Adm Sci Quart. 1979;24:285–308.
35. Johnson JV, Hall EM. Job strain, work place social support, and cardiovascular disease: a cross-sectional study of a random sample of the Swedish working population. In: Steptoe A, Wardle J, editors. Psychosocial processes and health: a reader. New York: Cambridge University Press; 1994. p. 25–42.
36. Sousa-Poza A, Sousa-Poza AA. Well-being at work: a cross-national analysis of the levels and determinants of job satisfaction. J Socio Econ. 2000;29:517–38.
37. Calnan M, Wainwright D, Almond S. Job strain, effort-reward imbalance and mental distress: a study of occupations in general medical practice. Work Stress. 2000;14:297–311.
38. Matsudaira K, Shimazu A, Fujii T, Kubota K, Sawada T, Kikuchi N, Takahashi M. Workaholism as a risk factor for depressive mood, disabling back pain, and sickness absence. PLoS One. 2013;8:e75140. doi:10.1371/journal.pone.0075140.
39. Fujii T, Matsudaira K, Yoshimura N, Hirai M, Tanaka S. Associations between neck and shoulder discomfort (Katakori) and job demand, job control, and worksite support. Mod Rheumatol. 2013;23:1198–204.
40. Yamada K, Matsudaira K, Imano H, Kitamura A, Iso H. Influence of work-related psychosocial factors on the prevalence of chronic pain and quality of life in patients with chronic pain. BMJ Open. 2016;6:e010356. doi:10.1136/bmjopen-2015-010356.
41. Boakye PA, Olechowski C, Rashiq S, Verrier MJ, Kerr B, Witmans M, Baker G, Joyce A, Dick BD. A critical review of neurobiological factors involved in the interactions between chronic pain, depression, and sleep disruption. Clin J Pain. 2016;32:327–36.

42. Bijl D, van Marwijk HWJ, de Haan M, van Tilburg W, Beekman AJTF. Effectiveness of disease management programmes for recognition, diagnosis and treatment of depression in primary care. a review. Eur J Gen Pract. 2004;10:6–12.
43. Wertli MM, Rasmussen-Barr E, Weiser S, Bachmann LM, Brunner F. The role of fear avoidance beliefs as a prognostic factor for outcome in patients with nonspecific low back pain: a systematic review. Spine J. 2014;14:816–36.
44. Wertli MM, Rasmussen-Barr E, Held U, Weiser S, Bachmann LM, Brunner F. Fear-avoidance beliefs: a moderator of treatment efficacy in patients with low back pain: a systematic review. Spine J. 2014;14:2658–78.
45. Kendall N, Burton K, Main C, Watson P. Tackling musculoskeletal problems: a guide for clinic and workplace. Identifying obstacles using the psychosocial flags framework. London: The Stationery Office; 2009. ISBN: 9780117037892
46. Cohen SP, Mao J. Neuropathic pain: mechanisms and their clinical implications. BMJ. 2014;348:f7656. doi:10.1136/bmj.f7656.

Journal of Pain Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Estimated risk for chronic pain determined using the generic STarT Back 5-item screening tool

This article was published in the following Dove Press journal:
Journal of Pain Research
24 February 2017
Number of times this article has been viewed

Hiroyuki Oka[1]
Ko Matsudaira[1]
Tomoko Fujii[1]
Norimasa Kikuchi[2,3]
Yuri Haga[2]
Takayuki Sawada[2,3]
Junji Katsuhira[4]
Takahiko Yoshimoto[5]
Kayo Kawamata[1]
Juichi Tonosu[6]
Masahiko Sumitani[7]
Satoshi Kasahara[7]
Sakae Tanaka[8]

[1]Department of Medical Research and Management for Musculoskeletal Pain, Faculty of Medicine, 22nd Century Medical & Research Center, University of Tokyo, Tokyo, [2]Clinical Study Support, Inc., [3]Department of Public Health, Aichi Medical University School of Medicine, Aichi, [4]Department of Prosthetics & Orthotics and Assistive Technology, Faculty of Medical Technology, Niigata University of Health and Welfare, Niigata, [5]Department of Rehabilitation, Kameda Medical Center, Chiba, [6]Department of Orthopaedic Surgery, Kanto Rosai Hospital, Kanagawa, [7]Department of Pain and Palliative Medicine, Faculty of Medicine, [8]Department of Orthopaedic Surgery, University of Tokyo, Tokyo, Japan

Correspondence: Hiroyuki Oka
Faculty of Medicine, 22nd Century Medical Center, University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo 113-8655, Japan
Tel +81 3 3815 5411ext34580
Fax +81 3 5800 9549
Email okah-tky@umin.ac.jp

Objective: The generic STarT Back 5-item screening tool (STarT-G) is used to manage chronic pain in the lower back and elsewhere. This study evaluated the validity of the Japanese version of this generic screening tool.

Materials and methods: Japanese participants between the ages of 20 and 64 years completed online surveys regarding pain. Survey reliability was assessed with internal consistency, as calculated using Cronbach's alpha coefficients. Spearman's correlation coefficients were used to evaluate concurrent validity between the STarT-G score and standard reference questionnaires. Associations between STarT-G scores and the presence of a disability due to chronic pain (DCP) were analyzed using receiver operator characteristic (ROC) curves.

Results: Analyses ultimately included data obtained from 52,842 Japanese participants (54.4% male) with a mean (standard deviation) age of 47.7 (9.4) years. Approximately 1.5% of participants had DCP, and the mean STarT-G score was 1.2 (1.4). The Cronbach's alpha coefficient was 0.71, indicating an acceptable reliability. The STarT-G score moderately correlated with the pain numerical rating scale (NRS) score (Spearman's correlation coefficient: $r = 0.34$). When the STarT-G threshold was set at 4, the sensitivity and specificity of the DCP predictive model were 65.8% and 82.4%, respectively, and the area under the ROC was 0.808.

Conclusion: The STarT-G was internally consistent and was able to distinguish between subjects with and without a DCP. Therefore, the STarT-G can reliably be used in the Japanese population to identify patients with DCP.

Keywords: chronic pain, disability, primary care, psychological factors, screening tool, somatic symptoms

Introduction

Disability due to chronic pain (DCP) results in absence from work and is a major public health concern in Japan and many Western countries.[1–4] Various screening tools have been developed to identify chronic pain subgroups and comorbid factors.[5–7] A widely used powerful tool is the STarT Back Tool (STarT), a 9-item screening tool that was developed as a prognostic indicator of lower back pain (LBP). Items 1–4 evaluate physical factors and items 5–9 assess psychosocial factors (Figure 1).[5,8] The STarT score is often used by primary care physicians in England to make clinical decisions.[5] Specifically, the STarT results indicate the subgroup that an LBP patient falls into, which helps determine which treatment strategies may be most effective. The STarT has been shown to be particularly effective for individual patient management in the physiotherapy setting. Patients who underwent STarT testing and subsequent targeted therapy had higher clinical and cost efficacy than patients who did not undergo STarT

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/JPR.S129585

Patient name: _____ Date: _____

Thinking about the **last 2 weeks** tick your response to the following questions:

		Disagree 0	Agree 1
1	My back pain has **spread down my leg(s)** at some time in the last 2 weeks	☐	☐
2	I have had pain in the **shoulder** or **neck** at some time in the last 2 weeks	☐	☐
3	I have only **walked short distances** because of my back pain	☐	☐
4	In the last 2 weeks, I have **dressed more slowly** than usual because of back pain	☐	☐
5	It's not really safe for a person with a condition like mine to be physically active	☐	☐
6	**Worrying thoughts** have been going through my mind a lot of the time	☐	☐
7	I feel that **my back pain is terrible** and **it's never going to get any better**	☐	☐
8	In general I have **not enjoyed** all the things I used to enjoy	☐	☐

9. Overall, how **bothersome** has your back pain been in the **last 2 weeks**?

Not at all	Slightly	Moderately	Very much	Extremely
☐ 0	☐ 0	☐ 0	☐ 1	☐ 1

Total score (all 9): _____ **Sub Score (Q5-9):**_____

© Keele University 01/08/07
Funded by Arthritis Research UK

Figure 1 The Keele STarT Back screening tool (9-item).
Note: Copyright ©2007. Reprinted from Keele University. STarT Back Screening Tool Website. Available from: https://www.keele.ac.uk/sbst/startbacktool/usingandscoring/.[8]

testing and were treated with usual care strategies.[5] We previously translated the STarT into Japanese,[9] and this version was linguistically validated in a general cross-cultural adaptation process.[10–12] We also evaluated the reliability and validity of "the STarT into Japanese" in a large number of Japanese patients with LBP.[13]

The lower back was the most common site of chronic pain and accounted for 65% of all cases of reported chronic pain in a Japanese epidemiological study.[1] However, chronic pain often originates in places other than the lower back, and a generic screening tool is needed to help effectively manage chronic pain from all sites. One such tool is the generic version of the STarT Back 5-item screening tool (STarT-G), a modified version of the 9-item STarT.[8] The STarT 9-item screening tool provides an easy way to stratify patients into three subgroups according to the probability of a poor prognosis or pain chronicity. These categories are defined as "low risk," "medium risk," and "high risk" (Figure 2).[8] On the other hand, the use of STarT-G (5-item) screening tool has not yet been established. The STarT-G has also not been validated for evaluating chronic pain in a large group of Japanese subjects. Therefore, the current study was performed to examine the validity of STarT-G in such a population using cross-sectional data obtained from STarT-G surveys administered online.

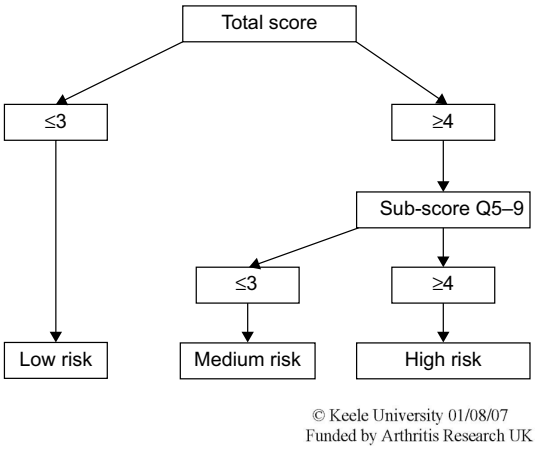


Figure 2 The STarT Back tool scoring system.
Notes: Scores were used to stratify patients into "low risk," "medium risk," and "high risk" groups. Copyright ©2007. Reprinted from Keele University. STarT Back Screening Tool Website. Available from: https://www.keele.ac.uk/sbst/startbacktool/usingandscoring/.[8]

Materials and methods

This study was reviewed and approved by the medical/ethics review board of the Japan Labour Health and Welfare Organization at Kanto Rosai Hospital (Kanagawa, Japan, approval number: 2012-22). All study procedures adhered to the tenets of the Declaration of Helsinki. Participation was voluntary, and no personal information was collected. Written informed consent was not obtained, but submitting the completed questionnaire was considered evidence of consent. Before completing the questionnaire, potential participants read an explanation of the survey's purpose and were informed that they should proceed to the questionnaire only if they agreed to participate in the study. As an incentive, participants received online shopping reward points from the Internet research company that helped conduct this study (UNITED, Inc., Tokyo, Japan).

Study population

Subject information was collected via surveys administered online in January and February 2014. Participants were recruited from an online panel conducted by an Internet research company (UNITED, Inc.). The all-Japanese study population consisted of ~1.25 million registered research volunteers between the ages of 20 and 64 years. From this volunteer pool, 965,919 individuals were randomly selected and invited by e-mail to complete an online questionnaire on health problems associated with pain. We ultimately obtained 52,842 online responses by January 31, 2014.

Study measures

The 5-item STarT-G tool is a modified version of the 9-item psychosocial subscale that specifically identifies distress in other conditions.[5] Questions address fear (one item from the Tampa Scale of Kinesiophobia), anxiety (one item from the Hospital Anxiety and Depression Scale), pessimistic patient expectations (one item from the Pain Catastrophizing Scale), low mood, (one item from the Hospital Anxiety and Depression Scale), and how bothersome pain is.[7] The first four items had possible responses of "agree" or "disagree," and the bothersome item had possible responses from 0 to 5 (Likert scale). We used the 5-item STarT back screening tool that is available from the Keele University website (March 2013, Figure 3).[8]

The study questionnaire investigated pain experienced over the past month in 20 different anatomical sites. All anatomical sites were illustrated on diagrams to ensure that participants correctly identified each area. Examined sites included the head, chin, teeth/mouth, face, throat, neck, shoulder, elbow, wrist/hand, chest, abdomen, back, low back, hip, thigh, knee, lower leg, ankle/foot, genitals, and anus. The degree of chronic pain experienced over the last 4 weeks was assessed using the numerical rating scale (NRS), with scores ranging from 0 (no pain at all) to 10 (the worst pain imaginable).

Somatizing tendency was assessed using a subset of items from a linguistically validated Japanese version of the Brief Symptom Inventory (BSI).[14,15] Seven somatic symptoms were assessed for severity, including faintness or dizziness, pain in the heart or chest, nausea or upset stomach, difficulty breathing, numbness or tingling in part of the body, weakness in part of the body, and hot or cold spells. All symptoms were assessed on a five-point scale that evaluated how much the participant was bothered by the symptom. Participants chose from the following response options: not at all (0), mildly (1), moderately (2), quite a bit (3), and extremely (4). For this test, participants were grouped by the number of somatic symptoms or pain sites. A participant was considered to have a symptom if he/she responded with a 2–4, which is indicative of somatization.[16,17]

The presence/absence of a DCP was also investigated. A DCP was considered present when the pain symptoms had continued for at least 6 months and the subject had withdrawn from social activities because of pain.

Statistical analyses

Data are presented as mean (standard deviation), where applicable. Participant demographic and clinical characteristics were summarized using descriptive statistics. To examine floor and ceiling effects, the percentages of respondents with total scores of 0 and 5 were calculated. Floor and ceiling effects were considered present when >15% of respondents had the lowest or highest possible score, respectively.[18] To examine STarT-G reliability, we evaluated

Patient name: _____ Date: _____

Thinking about the **last 2 weeks** tick your response to the following questions:

		Disagree 0	Agree 1
1	It's really not safe for a person with a condition like mine to be physically active	☐	☐
2	**Worrying thoughts** have been going through my mind a lot of the time in the last 2 weeks	☐	☐
3	I feel that **my problem is terrible** and that **it's never going to get any better**	☐	☐
4	In general in the last 2 weeks, I have **not enjoyed** all the things I used to enjoy	☐	☐

5. Overall, how **bothersome** has your condition been in the last 2 weeks?

Not at all	Slightly	Moderately	Very much	Extremely
☐	☐	☐	☐	☐
0	0	0	1	1

Score_____

© Keele University 01/08/07
Funded by Arthritis Research UK

Figure 3 The generic condition screening tool (5-items).
Note: Copyright ©2007. Reprinted from Keele University. STarT Back Screening Tool Website. Available from: https://www.keele.ac.uk/sbst/startbacktool/usingandscoring/.[8]

internal consistency by calculating Cronbach's alpha coefficients. An alpha index >0.70 indicates a satisfactory internal consistency.[19] Spearman's correlation coefficients were used to evaluate concurrent validity by examining correlations between STarT-G and NRS pain scores. Correlation coefficients were interpreted using Cohen's[20] criteria for correlation strength in psychometric validation (0.10 = weak, 0.30 = moderate, and 0.50 = strong).

The ability of STarT-G scores to differentiate between participants with known differences (known-group validity) was examined using the Jonckheere–Terpstra test. To do this, participants were categorized into the following groups according to the number of somatic symptoms present: no symptoms, one symptom, and two or more symptoms.

Associations between STarT-G scores and the presence of a DCP were examined using receiver operator characteristic (ROC) curves and the corresponding area under the curve (AUC). Accuracy was determined using the AUC. The following traditional academic point system for AUC values can be used as a rough guide for classifying diagnostic test accuracy: 0.90–1.00 = excellent, 0.80–0.90 = good, 0.70–0.80 = fair, 0.60–0.70 = poor, and 0.50–0.60 = fail.[21] Statistical analyses were performed using SPSS statistical software (version 20.0; SPSS, Inc., Chicago, IL, USA). All reported P values are two-sided, and statistical significance was defined as $P < 0.05$.

Results

A total of 52,842 participants were ultimately included in analyses. Mean subject age was 47.7 (9.4) years, and 54.4% of participants were male. Approximately 1.5% of participants claimed to have experienced a DCP. Table 1 summarizes participant demographic characteristics and overall pain survey results.

Mean STarT-G score was 1.2 (1.4). A remarkable ceiling effect was not observed, with only 2.3% of participants reporting the highest score of 5. However, a substantial floor effect was observed, with 41.0% of participants reporting the lowest score of 0. The Cronbach's alpha coefficient was

Table 1 Participant demographic and pain characteristics

Characteristics	
Sex, n (%)	
Male	28,769 (54.4)
Female	24,073 (45.6)
Age, years	47.7 (9.4)
BMI, kg/m^2	22.8 (3.8)
STarT-G score	1.2 (1.4)
NRS for pain	3.1 (2.4)
Pain sites, n (%)	
0	12,045 (22.8)
1	14,076 (26.6)
2	10,014 (19.0)
3	6,370 (12.1)
4–5	6,188 (11.7)
6–9	3,484 (6.6)
10+	665 (1.3)
Disability due to chronic pain, n (%)	
Present	818 (1.5)
Absent	52,024 (98.5)

Note: Data presented as mean (standard deviation) where applicable.
Abbreviations: BMI, body mass index; STarT-G, generic version of the STarT Back 5-item screening tool; NRS, numerical rating scale.

0.71, indicating good test reliability. Concurrent validity was examined by investigating the correlation between STarT-G score and pain NRS. The two pain measures were only moderately correlated ($r = 0.34$).

We examined the STarT-G scores among participants with known differences. As expected, participants with more somatic symptoms had significantly higher STarT-G scores. The mean score was 0.97 (1.12), 1.96 (1.42), and 2.74 (1.53) in participants with zero, one, and two or more somatic symptoms, respectively (Figure 4). This linear trend of increasing total STarT-G score with an increasing number of somatic symptoms was highly significant (Jonckheere–Terpstra test, $P < 0.0001$). Furthermore, participants with pain at a higher number of body sites had significantly higher STarT-G scores. The mean score was 0.63 (1.05), 1.05 (1.25), 1.27 (1.30), 1.50 (1.37), 1.80 (1.45), 2.23 (1.54), and 2.96 (1.57) in participants with zero, one, two, three, four-to-five, six-to-nine, and ≥10 pain sites, respectively (Figure 5). This linearly increasing trend in STarT-G score with an increasing number of bodily pain sites was highly significant (Jonckheere–Terpstra test, $P < 0.0001$).

The ability of the model to predict the presence of a DCP was also examined when the STarT-G threshold was set to 4. At this cutoff value, sensitivity and specificity for detecting a DCP were 65.8% and 82.4%, respectively. Additionally, area under the ROC curve was 0.808 for this STarT-G threshold, indicating that the model was good (Figure 6).

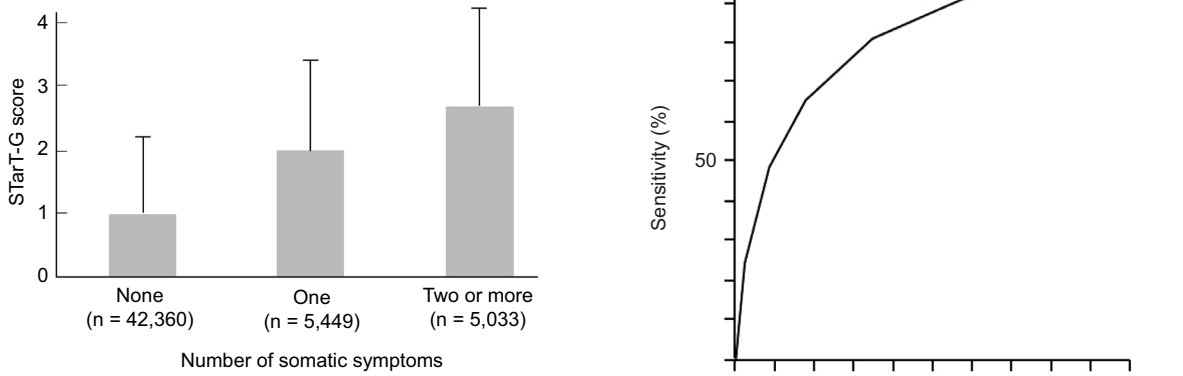


Figure 4 Mean STarT-G scores for participants with different numbers of somatic symptoms.
Notes: The linear trend was found to be highly significant (Jonckheere–Terpstra test, $P < 0.0001$). The STarT-G is the generic version of the STarT Back 5-item screening tool. The number of somatic symptoms was determined using the Brief Symptom Inventory somatization scale.

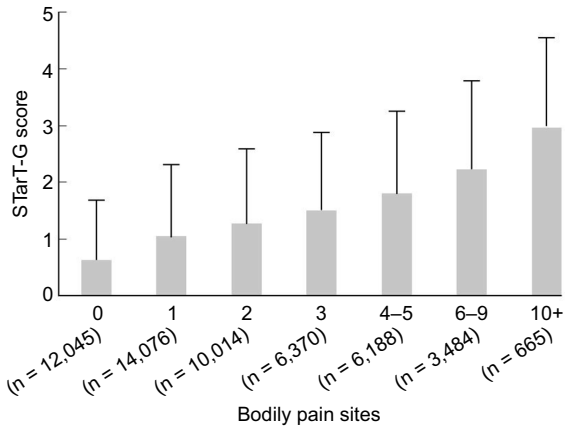


Figure 5 Mean STarT-G scores for participants with different numbers of pain sites.
Notes: The linear trend was found to be highly significant (Jonckheere–Terpstra test, $P < 0.0001$). The STarT-G is the generic versions of the STarT Back 5-item screening tool. The number of pain sites represents pain experienced during the past month in the head, chin, teeth/mouth, face, throat, neck, shoulder, elbow, wrist/hand, chest, abdomen, back, low back, hip, thigh, knee, lower leg, ankle/foot, genitals, and/or anus.

Figure 6 Receiver operating characteristics (ROC) curve of disability due to chronic pain, as assessed using a STarT-G score threshold value of 4.
Note: The area under the ROC curve was 0.808.

Discussion

Here, we evaluated psychometric properties of the STarT-G. We found that the survey was internally consistent and had acceptable concurrent and known-groups validity in the Japanese population. The Cronbach's alpha coefficient for the STarT-G was 0.71, indicating a good internal consistency. This value was similar to that obtained for the Japanese 9-item STarT scale (0.75).[13] Concurrent validity was assessed by analyzing correlations between the STarT-G and pain NRS scores, which were moderately correlated with each other ($r = 0.34$). Known-group validity was investigated by examining relationships between STarT-G scores and the number of somatic symptoms and body pain sites. These analyses showed that the STarT-G score increased as the number of somatic symptoms and pain sites increased. This suggests that the STarT-G is able to differentiate between patients with different levels of chronic pain and pain-related problems.

Yellow flags are useful in identifying patients with chronic LBP who have a poor prognosis.[22] The 5-item tool covers the minimal important psychological factors that are considered to be yellow flags for overall chronic LBP. This survey includes questions related to fear, anxiety, catastrophizing, depression, and bothersomeness, all of which are the most important predictors identified as yellow flags. For patients with high STarT-G scores, specific cognitive behavioral approaches are needed in addition to pain education, motivation, encouragement, exercise, medical therapy (minimal amounts), and physical treatment. This conclusion is based on previous reports that stated, "early intervention to yellow flag leads to better outcome."[23,24]

Finally, ~1.5% of participants reported having a DCP. At a STarT-G threshold value of 4 points, ROC analysis revealed that the sensitivity and specificity of DCP were 65.8% and 82.4%, respectively. Additionally, the AUC was 0.808, indicating a good capacity of the STarT-G to differentiate between patients with and without a DCP.

The STarT-G is a diagnosis-specific screening tool used for communication between primary care physicians and pain specialists in the care of chronic pain patients. Using the STarT-G threshold of 4 points, patients examined here were divided into the following two groups: those at risk for a DCP and those with minimal to no risk for a DCP. We recommend that patients at or beyond this threshold consult a pain specialist. The STarT-G is now planned to be used as a tool to identify patients for referral to one of 18 core facilities in Japan that provide cognitive behavioral therapy.

Our study had several limitations. First, our study population was selected from Internet research volunteers who have chronic pain. Given that 41% of participants had a STarT-G score of 0, many patients may have had chronic pain that was not severe enough to require medical care. This may have influenced our results. Second, Internet-based surveys can introduce a selection bias and may not be representative of the general population. Because our study population was selected from Internet research volunteers who may differ from general Internet users, caution is needed when interpreting our study findings. In particular, people living in large cities are overrepresented in Internet survey company volunteers. In addition, a higher proportion of respondents had completed university or graduate level education than the general population, particularly in older respondents.[25] Third, our study had a test reliability of >0.70.[19] However, Nunnally and Bernstein[26] recommend a minimum test reliability of >0.90 for making clinical decisions. Therefore, it is possible that test reliability was overestimated. Finally, this cross-sectional study did not assess the ability of the STarT-G to predict pain consistency. Future longitudinal studies are needed to better understand potential associations between risk groups and long-term pain outcomes. These should also examine whether or not the STarT-G score is predictive of DCP.

Conclusion

The STarT-G scale had acceptable internal consistency, reliability, and validity (concurrent and known groups) in Japanese patients with chronic pain. We hope that these analyses of the psychometric properties of STarT-G will enable Japanese clinicians to use this survey as a screening tool for detecting DCPs. The STarT-G is simple, fast, and suitable for use in primary care settings, all of which suggest that the STarT-G may facilitate screening for DCP in the primary care setting in Japan. We hope using the STarT-G will ultimately ease physical, social, and economical burdens of chronic pain in the Japanese population.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Nakamura M, Nishiwaki Y, Ushida T, Toyama Y. Prevalence and characteristics of chronic musculoskeletal pain in Japan. *J Orthop Sci*. 2011; 16(4):424–432.
2. Goldberg DS, McGee SJ. Pain as a global public health priority. *BMC Public Health*. 2011;11:770.

3. Guerriere DN, Choinière M, Dion D, et al. The Canadian STOP-PAIN project - Part 2: what is the cost of pain for patients on waitlists of multidisciplinary pain treatment facilities? *Can J Anaesth*. 2010;57(6):549–558.
4. Lynch ME. The need for a Canadian pain strategy. *Pain Res Manag*. 2011;16(2):77–80.
5. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*. 2011;378(9802):1560–1571.
6. Leboeuf-Yde C, Gronstvedt A, Borge JA, et al. The Nordic back pain subpopulation program: demographic and clinical predictors for outcome in patients receiving chiropractic treatment for persistent low back pain. *J Manipulative Physiol Ther*. 2004;27(8):493–502.
7. Dunn KM, Croft PR. Classification of low back pain in primary care: using "bothersomeness" to identify the most severe cases. *Spine (Phila Pa 1976)*. 2005;30:1887–1892.
8. STarT Back Screening Tool Website. Available from: https://www.keele.ac.uk/sbst/startbacktool/usingandscoring/. Accessed February 17, 2017.
9. Matsudaira K, Kikuchi N, Kawaguchi M, et al. Development of a Japanese version of the STarT (Subgrouping for Targeted Treatment) Back screening tool: translation and linguistic validation. *J Musculoskel Pain Res*. 2013;5:11–19. Japanese.
10. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417–1432.
11. Suzukamo Y, Kumano H. Psychometrics. In: Ikegami N, Fukuhara S, Shimozuma K, Ikeda S, editors. *QOL Evaluation Handbook for Clinical Diagnosis*. Tokyo: Igaku Shoin; 2001:8–13. Japanese.
12. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for translation and cultural adaptation. *Value Health*. 2005;8(2):94–104.
13. Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S. Psychometric properties of the Japanese version of the STarT Back Tool in patients with low back pain. *PLoS One*. 2016;11(3):e0152019.
14. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med*. 1983;13(3):595–605.
15. Matsudaira K, Inuzuka K, Kikuchi N, et al. Development of the Japanese version of the brief symptom inventory-somatization scale: translation and linguistic validation. *Orthop Surg*. 2012;63:149–153. Japanese.
16. Matsudaira K, Palmer KT, Reading I, Hirai M, Yoshimura N, Coggon D. Prevalence and correlates of regional pain and associated disability in Japanese workers. *Occup Environ Med*. 2011;68(3):191–196.
17. Derogatis LR, Melisoratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med*. 1983;13(3):595–605.
18. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42.
19. Nunnally JC. *Psychometric Theory*. 2nd ed. New York: McGraw-Hill; 1978.
20. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.
21. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Hosmer DW, Lemeshow S, editors. *Applied Logistic Regression*. 2nd ed. New York: Wiley; 2000:143–202.
22. Pincus T, McCracken LM. Psychological factors and treatment opportunities in low back pain. *Best Pract Res Clin Rheumatol*. 2013;27(5):625–635.
23. Nicholas MK, Linton SJ, Watson PJ, et al. Early identification and management of psychological risk factors ("yellow flags") in patients with low back pain: a reappraisal. *Phys Ther*. 2011;91(5):737–753.
24. Kendall NA, Linton SJ, Main CJ. *Guide to Assessing Psychosocial Yellow Flags in Acute Low Back Pain: Risk Factors for Long-term Disability and Work Loss*. Wellington, New Zealand: Accident Rehabilitation and Compensation Insurance Corporation of New Zealand and the National Health Committee; 1997.
25. Statistics Bureau Ministry of Internal Affairs and Communication [webpage on the Internet]. *Population Census and Labour Force Survey*. 2011. Available from: www.stat.go.jp; http://www.stat.go.jp/data/index.htm. Accessed October 4, 2011.
26. Nunnally JC, Bernstein IH. *Psychometric Theo*. 3rd ed. New York: McGraw-Hill; 1994.

Journal of Pain Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Sex-specific impact of early-life adversity on chronic pain: a large population-based study in Japan

This article was published in the following Dove Press journal:
Journal of Pain Research
16 February 2017
Number of times this article has been viewed

Keiko Yamada[1,2]
Ko Matsudaira[3,4]
Eizaburo Tanaka[1,5]
Hiroyuki Oka[3]
Junji Katsuhira[3,6]
Hiroyasu Iso[1]

[1]Public Health, Department of Social Medicine, Osaka University Graduate School of Medicine, Suita, Osaka, [2]Center for Pain Management, Osaka University Hospital, Suita, Osaka, [3]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, [4]Japan Labour Health & Welfare Organization, Tokyo, [5]Hyogo Institute for Traumatic Stress, Kobe, [6]Department of Prosthetics & Orthotics and Assistive Technology, Faculty of Medical Technology, Niigata University of Health and Welfare, Niigata, Japan

Background: Responses to early-life adversity may differ by sex. We investigated the sex-specific impact of early-life adversity on chronic pain, chronic multisite pain, and somatizing tendency with chronic pain.

Methods: We examined 4229 respondents aged 20–79 years who participated in the Pain Associated Cross-Sectional Epidemiological Survey in Japan. Outcomes were: 1) chronic pain prevalence, 2) multisite pain (≥3 sites) prevalence, and 3) multiple somatic symptoms (≥3 symptoms) among respondents with chronic pain related to the presence or absence of early-life adversity.

Multivariable-adjusted odds ratios (ORs) were calculated with 95% confidence intervals using a logistic regression model including age, smoking status, exercise routine, sleep time, body mass index, household expenditure, and the full distribution of scores on the Mental Health Inventory-5. We further adjusted for pain intensity when we analyzed the data for respondents with chronic pain.

Results: The prevalence of chronic pain was higher among respondents reporting the presence of early-life adversity compared with those reporting its absence, with multivariable ORs of 1.62 (1.22–2.15, $p<0.01$) in men and 1.47 (1.13–1.90, $p<0.01$) in women. Among women with chronic pain, early-life adversity was associated with multisite pain and multiple somatic symptoms; multivariable ORs were 1.78 (1.22–2.60, $p<0.01$) for multisite pain and 1.89 (1.27–2.83, $p<0.01$) for ≥3 somatic symptoms. No associations were observed between early-life adversity and chronic multisite pain or multiple somatic symptoms among men with chronic pain.

Conclusion: Early-life adversity may be linked to a higher prevalence of chronic pain among both sexes and to multisite pain and somatizing tendency among women with chronic pain.

Keywords: sex characteristics, early-life adversity, chronic pain, somatoform disorders, disaster

Introduction

Early-life adversity (ELA) is defined as traumatic experiences during childhood encompassing maltreatment, accidents, death of a close relative, and disaster, any of which could have an influence not only in childhood but also in later life in the form of difficulties such as posttraumatic stress disorder (PTSD) or irritable bowel syndrome.[1,2] Previous studies have also described the relationship between ELA events and chronic pain (e.g., low back pain or fibromyalgia), but most of these studies were small-scale or targeted to North American, European, Oceanian populations,[3–7] and once targeted to Japanese population.[8]

This study focuses on the effects of ELA as a broader concept in relation to chronic pain. We used a question about adverse life events in general, rather than specific adversities.

Correspondence: Hiroyasu Iso
Public Health, Department of Social Medicine, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
Tel +81 6 6879 3911
Fax +81 6 6879 3919
Email iso@pbhel.med.osaka-u.ac.jp

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/JPR.S125556

Various sex or gender differences in tolerance for stressful life events have been documented. For example, a meta-analytic review revealed that women reported more symptoms of depression and anxiety than did men, but that the sex difference in psychological symptoms accounted for only about 4% of the variance in sex differences in reports of stress.[9] Sex or gender differences as they relate to chronic pain have been discussed for decades. The prevalence of chronic pain among women is higher than that among men,[10,11] and somatic symptoms have been reported by women than by men.[12,13]

We hypothesized that ELA would have long-term adverse impact, which manifested as chronic pain on more women than men, so we investigated the sex-specific association between ELA and the prevalence of chronic pain, chronic multisite pain, and somatizing tendency complicated by chronic pain in a large population-based study of Japanese men and women aged 20–79 years.

Methods

Ethical provisions

All procedures followed were in accordance with the ethical standards of the Helsinki Declaration of 1975 as revised in 2000. The institutional review boards of Keio University and of the Japan Labour Health and Welfare Organization approved this study. All participants had given their written informed consent before responding to the questionnaire. A credit point for Internet shopping was given as an incentive to the respondents.

Study population

The Pain Associated Cross-Sectional Epidemiological (PACE) study was a web-based survey designed to evaluate pain in a large Japanese population using a self-reported questionnaire. The PACE survey was conducted from 10 to 18 January 2009. The data set was the same as in previous PACE studies, profiles of which have been reported elsewhere;[14,15] however, the aim of this study was completely different from that of previous studies. Figure 1 shows the sampling procedure that culminated in the sample analyzed in the present study. A total of 20,044 respondents (9,746 men and 10,298 women) aged 20–79 years and matching the Japanese demographic composition in 2007 (Japanese Ministry of Internal Affairs and Communications, 2007) were recruited by e-mail from 1,477,585 candidates who registered with an Internet survey company (Rakuten Research, Inc., Tokyo, Japan). Computer-generated invitational e-mails were sent with a link to the first questionnaire until the targeted sample number was achieved. Incomplete questionnaires were rejected automatically, so the response rate was not calculated. The first questionnaire included items on age, sex, and pain, and was completed by 20,044 respondents. Subsequently, detailed questionnaires about lifestyle and psychosocial factors were sent to 5,000 of these respondents. Half (2,500) were chosen from those who had reported pain on the first questionnaire; the other half had reported being pain-free. The profile of these 5,000 respondents was consistent with the Japanese demographic composition for sex and age in 2007.[16] A total of 5,000 participants responded to the second questionnaire. Of these, we drew the data on 4,229 individuals (1,729 with chronic pain and 2,500 without pain) in the analyses. Moreover, the respondents with chronic pain were included in some additional analyses.

Definitions and measures

Chronic pain

The first questionnaire included items on pain such as the pain sites, pain intensity at each site, the site of dominant pain, and the duration of dominant pain. Pain intensities were scored on an 11-point Numerical Rating Scale (NRS; 0=no pain, 10=worst pain imaginable). In accord with the definition of chronic pain from the International Association for the Study of Pain, participants reported persistent pain over 3 months.[17]

Early-life adversity

We used a simple yes/no question to detect ELA, "Did you have any mentally shocking events (e.g., accidents experienced by you or close relatives, death of close relatives or friends, disaster, injury from violence) when you were 14 years old or younger?".

Multisite pain

The questionnaire included a picture of a human form with its body parts numbered from 1 to 21, and respondents entered the number(s) that corresponded to their pain site(s). A count of pain sites is a simple and useful measure for the severity of chronic pain, and chronic multisite pain is a strong predictor of future disability.[18] We defined more than three pain sites as multisite pain in the current study.

Somatizing tendency

Somatic symptom disorder involves having physical symptoms such as fatigue or dizziness caused by major emotional distress and problems functioning.[19] The disorder decreases its sufferers' quality of life. The Brief Symptom Inventory (BSI) is a self-reported measure of somatic symptoms, in which respondents answer on a 5-point Likert-type scale,

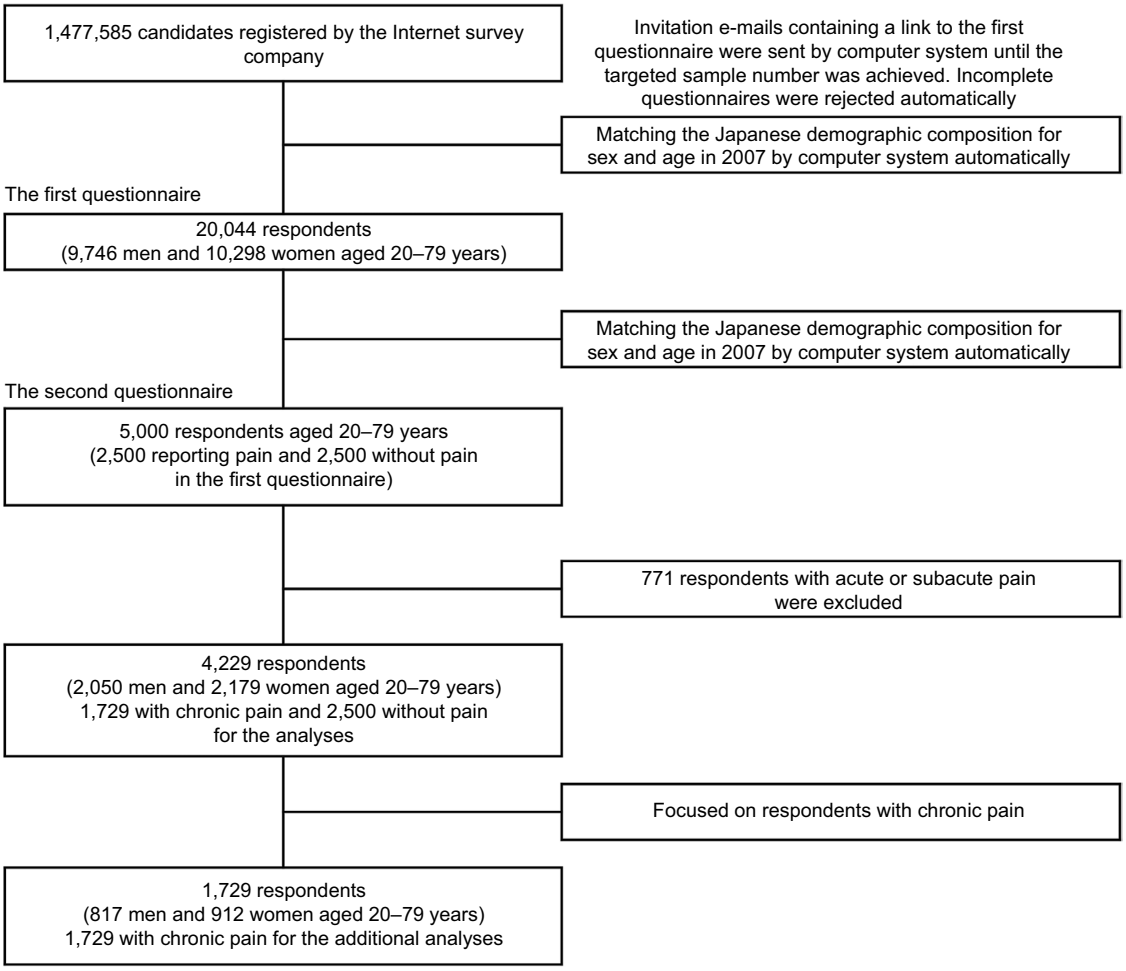


Figure 1 Flowchart of the sampling procedure ending in the sample being analyzed in the current study.

ranging from 0 (not at all) to 4 (extremely), regarding each of seven symptoms during the past 7 days: faintness or dizziness, pain in the heart or chest, nausea or upset stomach, trouble catching one's breath, numbness or tingling in parts of one's body, feeling weak in parts of one's body, and hot or cold spells.[20] Endorsing a response of 2, 3, or 4 was considered presence of the symptom. The number of symptoms with this level of response was counted; the totals ranged from zero to seven symptoms. We defined respondents with ≥3 symptoms, the highest tertile of the symptom count in our data, as existence of the somatizing tendency.

Mental status

We used the Mental Health Inventory (MHI-5), which is identical to the 36-item Short Form Health Survey (SF-36) "Mental Health" domain, to measure mental status.[21,22] The MHI-5 includes the following five questions: "How much of the time during the last month have you: 1) been a very nervous person?, 2) felt downhearted and blue?, 3) felt calm and peaceful?, 4) felt so down in the dumps that nothing could cheer you up?, and 5) been a happy person?". The respondents choose a number from 1 (all of the time) to 6 (none of the time).[21] The total score, which ranges from 5 to 30 points, is converted to a 100-point scale.[21] A previous Japanese study validated the cut point of <52 on the MHI-5 as screening for severe depressive symptoms.[21]

Statistical analysis

A Student's *t*-test was conducted to test for differences in age-adjusted mean values and proportions of risk factors for chronic pain. A chi-square test was performed to test for sex differences in the proportion of ELA.

Three outcomes were measured in the current study, 1) chronic pain prevalence among all respondents, 2) chronic multisite pain (≥3 sites) prevalence, and 3) multiple somatic symptoms (≥3 symptoms) among respondents with chronic pain, as these variables related to the presence or absence of ELA.

Multivariable-adjusted odds ratios (ORs) with 95% confidence intervals (CIs) were calculated using a logistic regression model to compare respondents with and without ELA.

p-Values <0.05 for two-tailed tests were considered statistically significant. All statistical analyses were performed using SAS version 9.4 (SAS Institute, Inc., Cary, NC, USA).

Confounding variables

We adjusted all analyses for the following confounding variables: age, smoking status (never, ex-smoker, or current smoker), have an exercise routine (exercise longer than 30 minutes more than twice a week; yes or no), sleep time (hours/day), body mass index (kg/m^2, categorized in quintiles), household expenditure (JPY/month), and the full distribution of scores on the MHI-5.

We further adjusted for pain intensity (the NRS that ranged from 0 to 10, i.e., 0=no pain, 10=worst pain imaginable) when we analyzed the data for respondents with chronic pain.

Results

Table 1 shows age-adjusted mean values of chronic pain risk factors according to the existence of ELA. Men with ELA were older (52.0 vs. 47.9 years), were more likely to have an exercise routine (45.9% vs. 33.7%), had a higher prevalence of body mass index ≥25 (31.8% vs. 24.8%), had higher household expenditures (380,000 vs. 293,000 JPY/month), had a higher proportion of severe depressive symptoms (27.3% vs. 19.5%), and had a higher prevalence of chronic pain (53.7% vs. 38.0%) compared with those who did not report ELA. Women with ELA had a higher prevalence of body mass index ≥25 (20.7% vs. 11.9%), severe depressive symptoms (38.5% vs. 20.2%), chronic pain (55.0% vs. 39.7%), and severe intensity of pain (6.5% vs. 6.1%) compared with those without it.

The prevalence of ELA was higher in women than in men (14.2% of women, 11.8% of men; $p<0.01$).

Multivariable-adjusted ORs of chronic pain prevalence of respondents with ELA are shown in Table 2. Multivariable-adjusted OR of chronic pain prevalence of men with ELA was 1.62 (1.22–2.15, $p<0.01$), and that of women with ELA was 1.47 (1.13–1.90, $p<0.01$).

Table 3 indicates multivariable-adjusted ORs of multisite pain (≥3 sites) among chronic pain sufferers with ELA. ELA was associated with higher risk for multisite pain among female chronic pain patients: multivariable-adjusted OR was 1.78 (1.22–2.60, $p<0.05$). However, there was no such association for men: multivariable-adjusted OR was 1.38 (0.88–2.16, $p=0.26$).

Table 1 Age-adjusted mean values and proportions of chronic pain risk factors

Chronic pain risk factors	Early-life adversity (–)	Early-life adversity (+)
Men		
n=2,050	1,808	242
Age, years, mean (SE)	47.9 (0.4)	52.0 (1.0)*
Current smoker, %	27.8	28.1
Have an exercise habit, %	33.7	45.9**
Sleep time <5 hours, %	3.4	5.0
Body mass index ≥25, %	24.8	31.8***
Household expenditure (*10,000 JPY/month)	29.3	38.0***
Severe depressive symptoms, %	19.5	27.3*
Chronic pain, % (no. of respondents with chronic pain=817)	38.0	53.7*
Intensity of pain among respondents with chronic pain (0–10 scale)	5.7	5.7
Women		
n=2,179	1,870	309
Age, years, mean (SE)	48.8 (0.4)	49.0 (0.9)
Current smoker, %	14.9	18.4
Have an exercise habit, %	29.0	33.3
Sleep time <5 hours, %	2.5	3.6
Body mass index ≥25, %	11.9	20.7*
Household expenditure (*10,000 JPY/month)	27.1	25.5
Severe depressive symptoms, %	20.2	38.5*
Chronic pain, % (no. of respondents with chronic pain=912)	39.7	55.0*
Intensity of pain among respondents with chronic pain (0–10 scale)	6.1	6.5***

Notes: Test for significance difference from the category of no early-life adversity: *$p<0.001$, **$p<0.01$, ***$p<0.05$.
Abbreviation: SE, standard error.

Table 2 ORs and 95% CIs of chronic pain prevalence of respondents with early-life adversity

	Early-life adversity (–)	Early-life adversity (+)
Men		
Number of respondents at risk	2,172	294
Number of respondents with chronic pain	687	130
Age-adjusted OR (95% CI)	1.00	1.86 (1.42–2.43)*
Multivariable-adjusted OR (95% CI)	1.00	1.62 (1.22–2.15)**
Women		
Number of respondents at risk	2,178	356
Number of respondents with chronic pain	742	170
Age-adjusted OR (95% CI)	1.00	1.86 (1.46–2.37)*
Multivariable-adjusted OR (95% CI)	1.00	1.47 (1.13–1.90)**

Notes: ORs are adjusted for age, smoking status, exercise routine, sleep time, body mass index, personal consumption expenditure, and the full distribution of scores on the Mental Health Inventory-5. Test for significant difference from the category of no early-life adversity: *$p<0.001$, **$p<0.01$.
Abbreviations: CI, confidence interval; OR, odds ratio.

Table 3 ORs and 95% CIs for multisite pain in chronic pain sufferers with early-life adversity

	Early-life adversity (–)	Early-life adversity (+)
Men		
Number of chronic pain sufferers	687	130
Number of chronic pain sufferers with multisite pain (≥3)	283	59
Age-adjusted OR (95% CI)	1.00	1.57 (1.06–2.34)***
Multivariable-adjusted OR (95% CI)	1.00	1.38 (0.88–2.16)
Women		
Number of chronic pain sufferers	742	170
Number of chronic pain sufferers with multisite pain (≥3)	379	117
Age-adjusted OR (95% CI)	1.00	2.27 (1.62–3.18)*
Multivariable-adjusted OR (95% CI)	1.00	1.78 (1.22–2.60)**

Notes: Adjusted for age, smoking status, exercise routine, sleep time, body mass index, household expenditure, the full distribution of scores on the Mental Health Inventory-5, and intensity of pain.Test for significant difference from the category of no early-life adversity: *p<0.001, **p<0.01, ***p<0.05.
Abbreviations: CI, confidence interval; OR, odds ratio.

Table 4 ORs and 95%CIs for multiple somatic symptoms among chronic pain sufferers with early-life adversity versus no early-life adversity

	Early-life adversity (–)	Early-life adversity (+)
Men		
Number of chronic pain sufferers	687	130
Number of multiple somatic symptoms (≥3)	283	59
Age-adjusted OR (95% CI)	1.00	1.57 (1.06–2.34)***
Multivariable-adjusted OR (95% CI)	1.00	1.27 (0.83–1.94)
Women		
Number of chronic pain sufferers	742	170
Number of multiple somatic symptoms (≥3)	379	117
Age-adjusted OR (95% CI)	1.00	2.10 (1.46–3.00)*
Multivariable-adjusted OR (95% CI)	1.00	1.89 (1.27–2.83)**

Notes: Adjusted for age, smoking status, exercise routine, sleep time, body mass index, household expenditure, the full distribution of scores on the Mental Health Inventory-5, and intensity of pain. Test for significant difference from the category of no early-life adversity: *p<0.001, **p<0.01, ***p<0.05.
Abbreviations: CI, confidence interval; OR, odds ratio.

In Table 4, ORs of multiple somatic symptoms (≥3 symptoms) for ELA among chronic pain sufferers are shown. The multivariable-adjusted OR of multiple somatic symptoms was 1.89 (1.27–2.83, p<0.01) for women with ELA. For men, ELA was not associated with somatic symptoms.

Discussion

The aim of this study was to examine the sex-specific impact of ELA on chronic pain, chronic multisite pain, and somatizing tendency with chronic pain. We hypothesized that ELA would have long-term adverse impact, which manifested as chronic pain on more women than men. The association of ELA with chronic multisite pain and with somatizing tendency among chronic pain sufferers supported our hypothesis. Although the significant associations were observed in women only, there was no sex difference in the association of ELA with the prevalence of chronic pain. Data from the Adverse Childhood Experience (ACE) study, which included 17,337 adults in the USA, also showed that ELA was associated with the prevalence of headache and with more frequent headaches in women than in men.[23]

ELA may reduce the volume of the hippocampus and prefrontal cortex; this reduction has been linked to major depression and to trait anxiety in adulthood, and predicts sensitivity to future stress events.[24,25] A magnetic resonance imaging study showed that 38 patients with chronic back pain and 30 patients with complex regional pain syndrome had a significantly smaller volume of bilateral hippocampal tissue than those of 50 healthy volunteers, whereas 20 patients with osteoarthritis did not.[26] Additionally, mice with neuropathic pain, in comparison with sham mice, showed more cellular and molecular changes linked to reduction of hippocampal function,[26] so reduction in the volume of the hippocampus due to ELA may actually cause chronic pain.

Moreover, sex differences in central sensitization could support our results. Central sensitization is the phenomenon in which nociceptive pain input from the peripheral nervous system triggers a prolonged but reversible synaptic change of pain pathways in the central nervous system.[27] Central sensitization contributes to the development of persistent pain hypersensitivity, spreads pain sensitivity across peripheral nerve territories without inflammation,[27] and amplifies pain from rheumatoid arthritis, osteoarthritis, fibromyalgia, and headache, as well as neuropathic pain, complex regional pain syndrome, and postsurgical pain.[27] Sex differences in enhanced pain sensitivity among patients with symptomatic knee osteoarthritis have been reported.[28]

In a psychological approach to chronic pain patients, especially women complaining of multisite pain or exhibiting somatizing tendencies, an intervention that addresses ELA should be considered.

Limitations

There were some limitations in this study. First, recall bias could exist because the current study was a cross-sectional design. The fact that people with persistent chronic pain are more likely to recall their ELA has been documented elsewhere.[29] Second, we used a simple and unvalidated question on ELA. A previous study of ELA among adolescents used a semi-structured interview that had good inter-rater reliability, and that study reported an association between ELA and depression.[30] Like that study, the current investigation concluded that respondents

with ELA showed a higher prevalence of depressive syndrome than did those without it. We believe that our single item on ELA was an appropriate proxy for the validated questionnaire. Third, the respondents may not be truly representative of the general population in Japan. The sampling issues with web-based surveys have been described previously.[31] Elderly people often have difficulty participating in such surveys. Moreover, the decision to respond to the survey may constitute selection bias, that is, the respondents who were suffering from chronic pain may have been particularly interested in pain research.

Conclusion

ELA was associated with a higher prevalence of chronic pain in both sexes, and with chronic multisite pain and somatizing tendency among women with chronic pain in the Japanese general population.

Acknowledgments

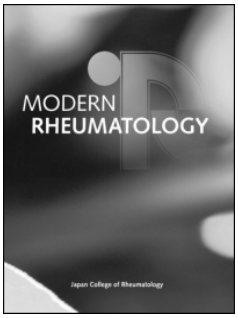
We are grateful to Dr. Yasuo Takagi, professor of Keio University, Japan, for his valuable help in conducting the survey. We also express our appreciation to all of the participants of this study, and we thank Dr. Masayuki Yao, psychiatrist of Ranryo Hospital, Japan, and Dr. Kenta Wakaizumi, Keio University, Japan, for their professional advice.

Disclosure

The authors report no conflict of interest in this work.

References

1. Chalavi S, Vissia EM, Giesen ME, et al. Abnormal hippocampal morphology in dissociative identity disorder and post-traumatic stress disorder correlates with childhood trauma and dissociative symptoms. *Hum Brain Mapp*. 2015;36(5):1692–1704.
2. Bradford K, Shih W, Videlock EJ, et al. Association between early adverse life events and irritable bowel syndrome. *Clin Gastroenterol Hepatol*. 2012;10(4):385–390.
3. Boisset-Pioro MH, Esdaile JM, Fitzcharles MA. Sexual and physical abuse in women with fibromyalgia syndrome. *Arthritis Rheum*. 1995; 38(2):235–241.
4. Filippon APM, Bassani DG, Aguiar RW de, Ceitlin LHF. Association between childhood trauma and loss of functionality in adult women with fibromyalgia. *Trends Psychiatry Psychother*. 2013;35(1):46–54.
5. Schofferman J, Anderson D, Hines R, Smith G, Keane G. Childhood psychological trauma and chronic refractory low-back pain. *Clin J Pain*. 1993;9(4):260–265.
6. Taylor ML, Trotter DR, Csuka ME. The prevalence of sexual abuse in women with fibromyalgia. *Arthritis Rheum*. 1995;38(2):229–234.
7. Jones GT, Power C, Macfarlane GJ. Adverse events in childhood and chronic widespread pain in adult life: results from the 1958 British Birth Cohort Study. *Pain*. 2009;143(1–2):92–96.
8. Stickley A, Koyanagi A, Kawakami N; WHO World Mental Health Japan Survey Group. Childhood adversities and adult-onset chronic pain: results from the world mental health survey, Japan. *Eur J Pain*. 2015;19(10):1418–1427.
9. Davis MC, Matthews KA, Twam/ey EW. Is life more difficult on Mars or Venus? a meta-analytic review of sex differences in major and minor life events. *Ann Behav Med*. 1999;21(1):83–97.
10. Craft RM, Mogil JS, Maria Aloisi A. Sex differences in pain and analgesia: the role of gonadal hormones. *Eur J Pain*. 2004;8:397–411.
11. Woodrow KM, Friedman GD, Siegelaub AB, Collen MF. Pain tolerance: differences according to age, sex and race. *Psychosom Med*. 1972;34(6): 548–556.
12. Yunus MB. The role of gender in fibromyalgia syndrome. *Curr Rheumatol Rep*. 2001;3(2):128–134.
13. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *J Gen Intern Med*. 2001;16(4):266–275.
14. Yamada K, Matsudaira K, Takeshita K, Oka H, Hara N, Takagi Y. Prevalence of low back pain as the primary pain site and factors associated with low health-related quality of life in a large Japanese population: a pain-associated cross-sectional epidemiological survey. *Mod Rheumatol*. 2013:1–8.
15. Yamada K, Matsudaira K, Imano H, Kitamura A, Iso H. Influence of work-related psychosocial factors on the prevalence of chronic pain and quality of life in patients with chronic pain. *BMJ Open*. 2016; 6(4):e010356.
16. Japanese Ministry of Internal Affairs and Communications. the Japanese demographic composition in 2007; 2007. Available from: http://www.stat.go.jp/data/jinsui/2007np/index.htm. Accessed March 14, 2015.
17. Treede R, Rief W, Barke A, et al. A classification of chronic pain for ICD-11. *Pain*. 2015;156(6):1003–1007.
18. Croft P, Blyth FM, van der Windt D. Number of pain sites-a simple measure of population risk? In: *Chronic Pain Epidemiology: from Aetiology to Public Health*. 1st ed. New York, NY: Oxford University Press; 2010:71–79.
19. Clinic M. Somatic symptom disorder – Mayo Clinic; 2015. Available from: http://www.mayoclinic.org/diseases-conditions/somatic-symptom-disorder/basics/definition/con-20124065. Accessed May 3, 2016.
20. Derogatis LR, Melisaratos N. The brief symptom inventory: an introductory report. *Psychol Med*. 1983;13(3):595–605.
21. Yamazaki S, Fukuhara S, Green J. Usefulness of five-item and three-item mental health inventories to screen for depressive symptoms in the general population of Japan. *Health Qual Life Outcomes*. 2005;3:48.
22. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 health survey for use in Japan. *J Clin Epidemiol*. 1998;51(11):1037–1044.
23. Anda R, Tietjen G, Schulman E, Felitti V, Croft J. Adverse childhood experiences and frequent headaches in adults. *Headache*. 2010;50(9): 1473–1481.
24. Frodl T, Reinhold E, Koutsouleris N, Reiser M, Meisenzahl EM. Interaction of childhood stress with hippocampus and prefrontal cortex volume reduction in major depression. *J Psychiatr Res*. 2010;44(13): 799–807.
25. Gorka AX, Hanson JL, Radtke SR, Hariri AR. Reduced hippocampal and medial prefrontal gray matter mediate the association between reported childhood maltreatment and trait anxiety in adulthood and predict sensitivity to future life stress. *Biol Mood Anxiety Disord*. 2014; 4:12.
26. Mutso AA, Radzicki D, Baliki MN, et al. Abnormalities in hippocampal functioning with persistent pain. *J Neurosci*. 2012;32(17):5747–5756.
27. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. *Pain*. 2011;152(3 Suppl):S2–S15.
28. Bartley EJ, King CD, Sibille KT, et al. Enhanced pain sensitivity among individuals with symptomatic knee osteoarthritis: potential sex differences in central sensitization. *Arthritis Care Res (Hoboken)*. 2016;68(4):472–480.
29. Croft P, Blyth FM, van der Windt D. Life-course influences on chronic pain in adults. In: *Chronic Pain Epidemiology: from Aetiology to Public Health*. 1st ed. New York, NY: Oxford University Press; 2010:177–183.
30. Rao U, Chen L-A, Bidesi AS, Shad MU, Thomas MA, Hammen CL. Hippocampal changes associated with early-life adversity and vulnerability to depression. *Biol Psychiatry*. 2010;67(4):357–364.
31. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. *J Epidemiol Community Health*. 2003;57(1):68–73.



Modern Rheumatology

ISSN: 1439-7595 (Print) 1439-7609 (Online) Journal homepage: http://www.tandfonline.com/loi/imor20

Relationship between roentgenographic joint destruction in the hands and functional disorders among patients with rheumatoid arthritis

Tetsuro Yasui, Hiroyuki Oka, Yasunori Omata, Yuho Kadono & Sakae Tanaka

To cite this article: Tetsuro Yasui, Hiroyuki Oka, Yasunori Omata, Yuho Kadono & Sakae Tanaka (2016): Relationship between roentgenographic joint destruction in the hands and functional disorders among patients with rheumatoid arthritis, Modern Rheumatology, DOI: 10.1080/14397595.2016.1254361

To link to this article: http://dx.doi.org/10.1080/14397595.2016.1254361

Accepted author version posted online: 10 Nov 2016.
Published online: 15 Nov 2016.

Japan College of Rheumatology
MODERN RHEUMATOLOGY

http://www.tandfonline.com/imor
ISSN 1439-7595 (print), 1439-7609 (online)

Mod Rheumatol, 2016; Early Online: 1–5

DOI: 10.1080/14397595.2016.1254361

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Relationship between roentgenographic joint destruction in the hands and functional disorders among patients with rheumatoid arthritis

Tetsuro Yasui[1,2], Hiroyuki Oka[3], Yasunori Omata[1], Yuho Kadono[4], and Sakae Tanaka[1]

[1]Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan, [2]Department of Orthopaedic Surgery, Teikyo University Mizonokuchi Hospital, Kanagawa, Japan, [3]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan, and [4]Department of Orthopaedic Surgery, Saitama Medical University, Saitama, Japan

Abstract

Objectives: Although a relationship between joint destruction and functional disorders seems apparent in patients with rheumatoid arthritis (RA), it has not been well proven in the literature. The aims of this study were to clarify the relationship between roentgenographic joint destruction in the hands and functional disorders in patients with RA, and to explore the appropriate assessment measures for functional disorders.

Methods: Cross-sectional data of the Genant-modified Total Sharp Score (Genant-mTSS), Health Assessment Questionnaire-Disability Index (HAQ-DI), Disabilities of the Arm, Shoulder, and Hand (DASH), and Michigan Hand Outcomes Questionnaire (MHQ) were collected from 50 consecutive RA patients and analyzed.

Results: HAQ-DI, DASH, and MHQ had close correlations with Genant-mTSS, with correlation coefficients of 0.69, 0.71, and −0.70, respectively, among patients with low disease activity (DAS28 < 3.2). A floor effect was observed in HAQ-DI, but neither floor nor ceiling effects were observed in DASH and MHQ. Both DASH and MHQ were strongly correlated with HAQ-DI, with correlation coefficients of 0.87 and 0.73, respectively.

Conclusions: Functional disorders had significant relationships with roentgenographic joint destruction in the hands among RA patients with low disease activity. As assessment measures of functional disorders, DASH and MHQ had no floor or ceiling effects, being different from HAQ-DI.

Keywords

Disabilities of the Arm, Shoulder, and Hand, Health Assessment Questionnaire, Michigan Hand Outcomes Questionnaire, Patient reported outcomes, Total Sharp Score

History

Received 23 May 2016
Accepted 24 October 2016

Introduction

Rheumatoid arthritis (RA) is characterized by polyarticular synovial inflammation, and chronic synovitis causes destruction of cartilage and bone [1]. The ''treat-to-target'' strategy using disease-modifying anti-rheumatic drugs (DMARDs) and/or biologics is widely accepted, and one of the important treatment goals is to prevent or reduce functional disorders caused by joint destruction [2].

Recently, strong bisphosphonates and anti-receptor activator of NF-kappa B ligand drugs have been proven to be very effective for osteoporosis and metastatic bone diseases [3,4], implying the possibility that these drugs may directly prevent or repair joint destruction in patients with RA and subsequently reduce functional disorders.

Before testing this therapeutic approach for functional disorders in RA patients, whether their joint destruction is actually related to functional disorders should be clarified. In addition, the choice of appropriate assessment measures for functional disorders needs to be discussed.

Correspondence to: Tetsuro Yasui, Department of Orthopaedic Surgery, Teikyo University Mizonokuchi Hospital, 3-8-3 Mizonokuchi, Takatsu-ku, Kawasaki, Kanagawa 213-8507, Japan. Tel: +81 44 844 3333. Fax: +81 44 844 3254. E-mail: yasuit@med.teikyo-u.ac.jp

The Total Sharp Score (TSS), which is calculated from X-ray films of the hands and feet, is widely used to quantify the joint destruction in RA patients [5–7]. For evaluation of functional disorders, the Health Assessment Questionnaire (HAQ) is commonly used for RA patients [8]. However, other measures, such as the Disabilities of the Arm, Shoulder, and Hand (DASH) [9,10] designed for assessment of the upper extremities and the Michigan Hand Outcomes Questionnaire (MHQ) [11,12] for specialized assessment of the hands, could be more appropriate than the HAQ when the hands are focused upon for assessment.

In the present study, we aimed to assess the relationship between roentgenographic joint destruction in the hands and functional disorders using several assessment measures. In addition, we assessed which measures are appropriate for the evaluation of functional disorders.

Methods

The study protocol was reviewed and approved by the research ethics committees of The University of Tokyo Hospital. Cross-sectional data were collected from 50 consecutive patients with RA at the clinic in The University of Tokyo Hospital in 2013 after receiving written informed consent. The subjects consisted of 46 women and four men, with a mean age of 62 years.

To quantify the joint destruction, X-ray examinations of both hands were performed and the Genant-modified TSS

Table 1. Demographic data of the patients.

Sex (female:male)	46:4
Age (years)	62 ± 11*
BMI (kg/m^2)	22 ± 3.5*
Disease duration (years)	16 ± 12*
No. of patients using DMARDs	46 (92%)
No. of patients using biologics	15 (30%)
No. of patients using steroids	22 (44%)
DAS28-ESR	3.2 ± 1.1*
Genant-mTSS	43 ± 37*
HAQ-DI	0.34 ± 0.47*
DASH	21 ± 19*
MHQ†	59 ± 16*
Grip strength (mmHg)	200 ± 79*
Pinch strength (kg)	4.0 ± 2.2*

*Data represent the mean ± SD.
†Mean of total scores of both hands.

(Genant-mTSS) was calculated. To evaluate the functional disorders, data for the validated Japanese translations of HAQ-Disability Index (HAQ-DI) [13], DASH [10], and MHQ [12] were collected. For the MHQ, the mean total scores of the MHQ in both hands were used for analysis. Japanese translation of MHQ had not been validated at the time of data collection, but we obtained and used the same version as the one which was later validated and published by Oda et al. [12].

Grip strength was measured with a mercury hand dynamometer (Acoma, Tokyo, Japan) and the mean strength of the right and left hands was used for analysis. Pinch strength was measured with a JAMAR hydraulic pinch gauge (Sammons Preston Rolyan, Bolingbrook, IL) and the mean strength of the right and left hands was used for analysis.

To test the relationship between functional disorder and joint destruction of the hand, correlations between Genant-mTSS and HAQ-DI, DASH, MHQ, grip strength, and pinch strength were analyzed. To test the usefulness of DASH and MHQ as assessment measures for functional disability, analyses were performed using HAQ-DI as the reference value.

For statistical assessments, we used SPSS Statistics version 21 (IBM Japan, Tokyo, Japan). Pearson's product–moment correlation analysis was employed to evaluate the relationship between two components, and a receiver-operating characteristic (ROC) analysis was used to determine cut-off scores. Values of $p < 0.01$ were considered to indicate statistical significance.

A ceiling or floor effect was defined to exist when the mean ± SD outranged the score range. Specifically, when the mean plus SD exceeded the highest score of the measure, it was referred to as a ceiling effect, and when the mean minus SD was below the lowest score of the measure, it was referred as a floor effect.

Results

The demographic data are shown in Table 1. The mean disease duration of RA was 16 years, and the proportions of patients using DMARDs, biologics, and steroids were 92%, 30%, and 44%, respectively. The mean values for the Disease Activity Score 28 (DAS28), Genant-mTSS, HAQ-DI, DASH, and MHQ were 3.2, 43, 0.34, 21, and 59, respectively. The mean grip strength and pinch strength were 200 mmHg and 4.0 kg, respectively.

Plots of the Genant-mTSS and individual functional scores are shown in Figure 1. HAQ-DI, DASH, and MHQ had moderate correlations with Genant-mTSS, with correlation coefficients of 0.51, 0.53, and −0.61, respectively. The correlation coefficient for MHQ had a negative value, because higher scores of MHQ indicated lower functional disability. Neither grip strength nor pinch strength had significant correlations with Genant-mTSS. No significant relationship was observed between DAS28 and HAQ-DI, DASH, or MHQ.

The subjects were divided into 20 patients with high disease activity (DAS28 ≥ 3.2) and 30 patients with low disease activity (DAS28 < 3.2) (Figure 1B,C). Among the patients with high disease activity, there were no significant correlations between Genant-mTSS and functional scores (Figure 1B). Meanwhile, close correlations between Genant-mTSS and all functional scores were found among patients with low disease activity (Figure 1C). A floor effect, or bottoming out of the score on zero, was observed in HAQ-DI. Neither floor nor ceiling effects were observed in DASH and MHQ.

As HAQ-DI is the established assessment measure for evaluating the functional disability of patients with RA, we assessed DASH and MHQ using HAQ-DI as the reference value. Plots of HAQ-DI (horizontal axis) and DASH and MHQ (vertical axis) are shown in Figure 2. Both DASH and MHQ were strongly correlated with HAQ-DI, with correlation coefficients of 0.87 and 0.73, respectively.

The cut-off score equivalent to HAQ-DI =0.5, which is widely accepted as the threshold for functional remission, was 26 and 55 in DASH and MHQ, respectively according to ROC analyses (Figure 3). The sensitivity and specificity were 86% and 100% for DASH, and 78% and 100% for MHQ, respectively. The area under the curve was 0.96 and 0.91 for DASH and MHQ, respectively.

Discussion

In patients with RA, functional disorders are dependent on disease activity and joint destruction [14]. Although a relationship between joint destruction and functional disorders seems apparent, it has not been well proven in the literature [15].

The present study revealed significant correlations between Genant-mTSS as a measure for joint destruction and functional measures such as HAQ-DI, DASH, and MHQ. The correlations were high among patients with low disease activity and low among patients with high disease activity, which supposedly reflects that functional disorders are affected by both disease activity and structural destruction.

HAQ-DI is a patient-reported measure of systemic functional disorders [8]. It is widely used as a measure to evaluate functional disorders in patients with RA, and HAQ-DI ≤0.5 is an accepted cut-off value for functional remission [16]. However, ceiling and floor effects have been pointed out [17].

DASH is a patient-reported assessment measure of the upper extremities created by the American Academy of Orthopaedic Surgeons [9]. MHQ is a patient-reported assessment measure established at Michigan University and is specialized for the hands [11]. These two measures have not been frequently applied for patients with RA, and their efficacies for evaluating functional disorders in RA patients have not been well discussed [18]. In the present study, we found that both DASH and MHQ had significant relationships with roentgenographic joint destruction in the hands, especially among patients with low disease activity. In addition, DASH and MHQ did not have floor or ceiling effects, being different from HAQ-DI. The results of our study indicate that DASH and MHQ are valuable assessment measures, and is superior to HAQ-DI, for evaluating functional disorders of the hands in patients with RA.

This study has some limitations. First, the patients included in our study were longstanding RA cases, whose mean duration of the disease was 16 years. We need further investigation to clarify if our findings in the study are independent of disease duration. Second, the explanation of the obtained results about patients with high disease activity is unclear. Although we guess functional disorder is explained in most part by disease activity and joint

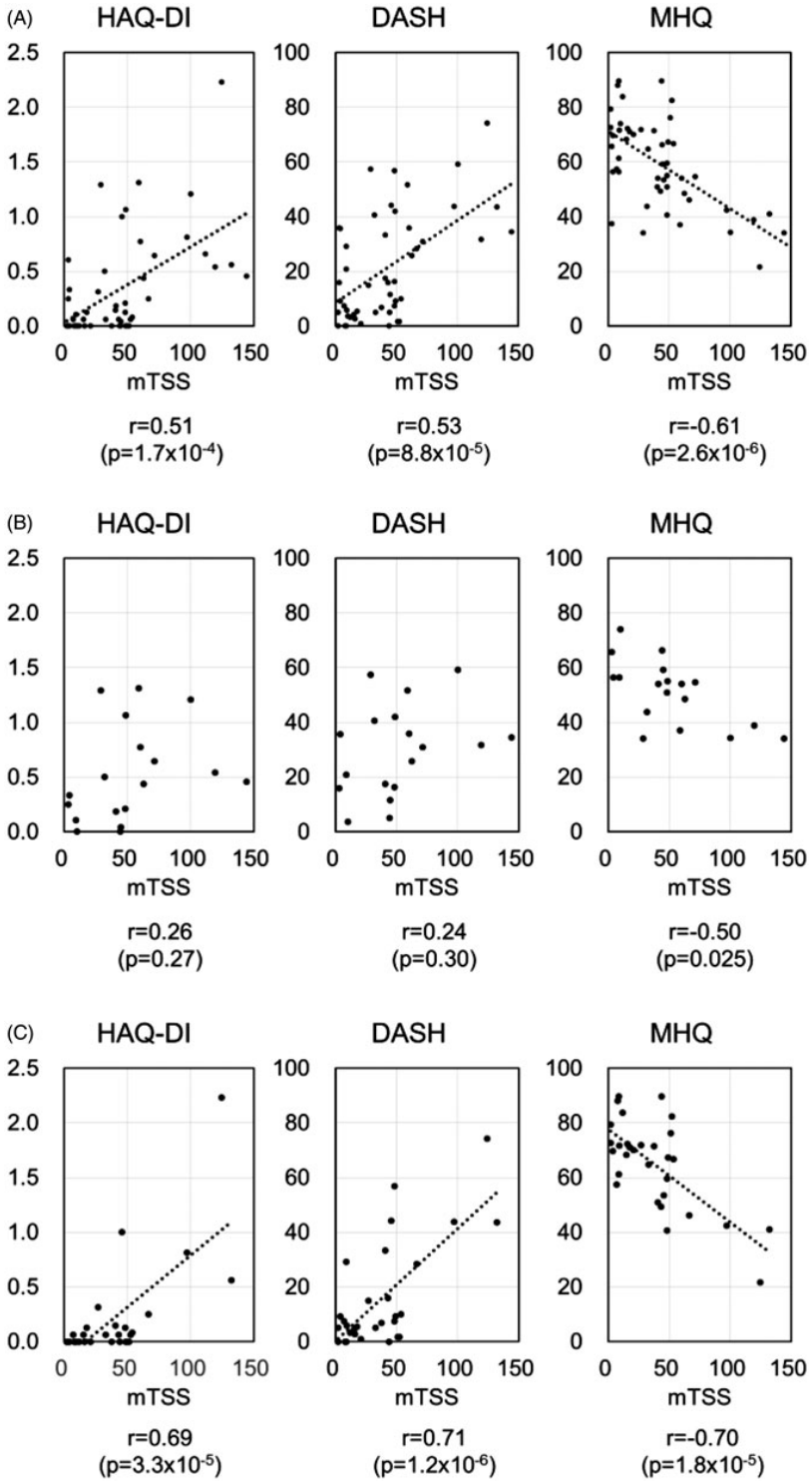


Figure 1. (A) Plots of Genant-mTSS and functional scores (HAQ-DI, DASH, and MHQ) for all patients. The dotted lines show the regression lines. The correlation coefficient of MHQ has a negative value, because high MHQ scores mean low functional disability. (B) Plots of Genant-mTSS and functional scores (HAQ-DI, DASH, and MHQ) for patients with high disease activity (DAS28 $\geq$ 3.2). It should be noted that no correlations are observed. (C) Plots of Genant-mTSS and functional scores (HAQ-DI, DASH, and MHQ) for patients with low disease activity (DAS28 $<$ 3.2). It should be noted that HAQ-DI, DASH, and MHQ are closely correlated with Genant-mTSS, and that a floor effect is observed for HAQ-DI while no floor and ceiling effects are observed for DASH and MHQ. The correlation coefficients are shown by the *r*-values.

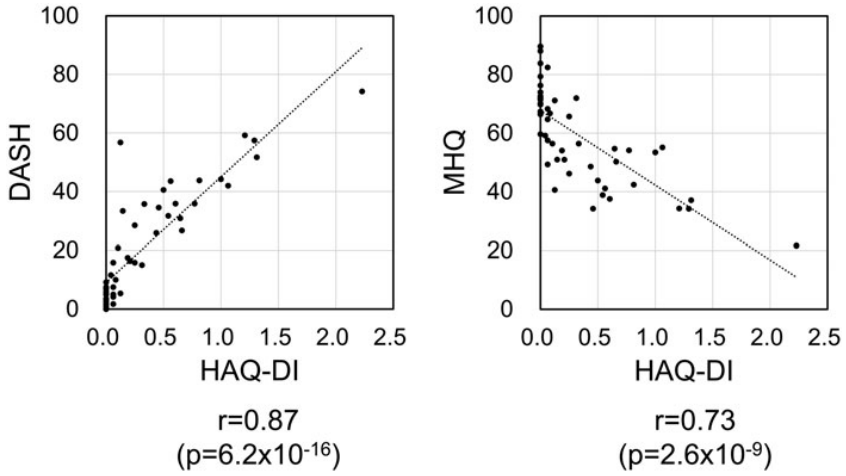


Figure 2. Plots of HAQ-DI with DASH and MHQ. It should be noted that both DASH and MHQ are strongly correlated with HAQ-DI.

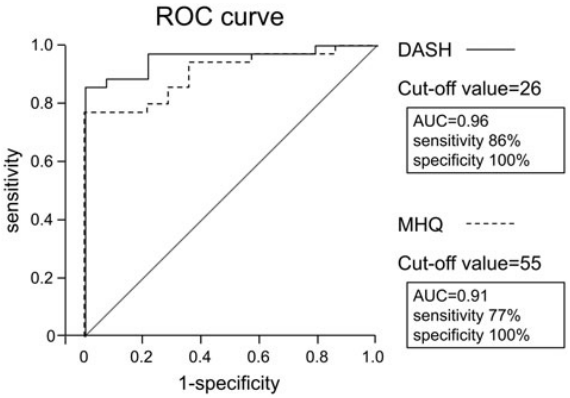


Figure 3. ROC curves for DASH and MHQ predicting HAQ-DI $\geq$0.5. The cut-off score equivalent for HAQ-DI $=$0.5 was 26 and 55 in DASH and MHQ, respectively. Note that AUC are as high as 0.96 and 0.91 in DASH and MHQ, respectively, indicating the high predictive accuracy.

destruction, the number of cases in our study is not enough for assessment.

Conclusions

Roentgenographic joint destruction in the hands had significant relationships with functional disorders among RA patients with low disease activity. DASH and MHQ are valuable assessment measures for evaluating functional disorders of the hands in patients with RA.

Acknowledgments

The authors thank Mrs. Orie Shiozuka for providing expert technical assistance.

Conflict of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

References

1. Harris ED. Rheumatoid arthritis. Pathophysiology and implications for therapy. N Engl J Med. 1990;322(18):1277–89.
2. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis. 2010;69(4):631–7.
3. Cummings SR, San Martin J, McClung MR, Siris ES, Eastell R, Reid IR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med. 2009;361(8):756–65.
4. Fizazi K, Carducci M, Smith M, Damiao R, Brown J, Karsh L, et al. Denosumab versus zoledronic acid for treatment of bone metastases in men with castration-resistant prostate cancer: a randomised, double-blind study. Lancet. 2011;377(9768):813–22.
5. Genant HK, Jiang Y, Peterfy C, Lu Y, Redei J, Countryman PJ. Assessment of rheumatoid arthritis using a modified scoring method on digitized and original radiographs. Arthritis Rheum. 1998;41(9):1583–90.
6. Sharp JT, Lidsky MD, Collins LC, Moreland J. Methods of scoring the progression of radiologic changes in rheumatoid arthritis. Correlation of radiologic, clinical and laboratory abnormalities. Arthritis Rheum. 1971;14(6):706–20.
7. van der Heijde D, Dankert T, Nieman F, Rau R, Boers M. Reliability and sensitivity to change of a simplification of the Sharp/van der Heijde radiological assessment in rheumatoid arthritis. Rheumatology (Oxford). 1999;38(10):941–7.
8. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137–45.
9. Hudak PL, Amadio PC, Bombardier C. The Upper Extremity Collaborative Group (UECG). Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. Am J Ind Med. 1996;29(6):602–8.
10. Imaeda T, Toh S, Nakao Y, Nishida J, Hirata H, Ijichi M, et al. Validation of the Japanese Society for Surgery of the Hand version of the Disability of the Arm, Shoulder, and Hand Questionnaire. J Orthop Sci. 2005;10(4):353–9.
11. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg Am. 1998;23(4):575–87.
12. Oda T, Abe Y, Katsumi Y, Ohi H, Nakamura T, Inagaki K. Reliability and validity of the Japanese Version of the Michigan Hand Outcomes Questionnaire: a comparison with the DASH and SF-36 Questionnaires. J Hand Surg Asian Pac Vol. 2016;21(1):72–7.
13. Kawai S. Rheumatoid arthritis and quality of life. Ryumachi. 1995;35(3):609–20.
14. Drossaers-Bakker KW, de Buck M, van Zeben D, Zwinderman AH, Breedveld FC, Hazes JM. Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. Arthritis Rheum. 1999;42(9):1854–60.
15. van der Heijde D, Landewe R, van Vollenhoven R, Fatenejad S, Klareskog L. Level of radiographic damage and radiographic

progression are determinants of physical function: a longitudinal analysis of the TEMPO trial. Ann Rheum Dis. 2008;67(9):1267–70.

16. Aletaha D, Machold KP, Nell VP, Smolen JS. The perception of rheumatoid arthritis core set measures by rheumatologists. Results of a survey. Rheumatology (Oxford). 2006;45(9):1133–9.
17. Fries J, Rose M, Krishnan E. The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. J Rheumatol. 2011;38(8):1759–64.
18. Aktekin LA, Eser F, Baskan BM, Sivas F, Malhan S, Oksuz E, et al. Disability of Arm Shoulder and Hand Questionnaire in rheumatoid arthritis patients: relationship with disease activity, HAQ, SF-36. Rheumatol Int. 2011;31(6):823–6.

Osteoporos Int (2017) 28:189–199
DOI 10.1007/s00198-016-3823-0

ORIGINAL ARTICLE

Is osteoporosis a predictor for future sarcopenia or vice versa? Four-year observations between the second and third ROAD study surveys

N. Yoshimura[1] · S. Muraki[1] · H. Oka[2] · T. Iidaka[1] · R. Kodama[3] · H. Kawaguchi[4] · K. Nakamura[5] · S. Tanaka[3] · T. Akune[5]

Received: 31 March 2016 / Accepted: 26 October 2016 / Published online: 24 November 2016

Abstract
Summary In a 4-year follow-up study that enrolled 1099 subjects aged ≥60 years, sarcopenia prevalence was estimated at 8.2%. Moreover, the presence of osteoporosis was significantly associated with short-term sarcopenia occurrence, but the reciprocal relationship was not observed, suggesting that osteoporosis would increase the risk of osteoporotic fracture and sarcopenia occurrence.
Introduction The present 4-year follow-up study was performed to clarify the prevalence, incidence, and relationships between sarcopenia (SP) and osteoporosis (OP) in older Japanese men and women.
Methods We enrolled 1099 participants (aged, ≥60 years; 377 men) from the second survey of the Research on Osteoarthritis/Osteoporosis against Disability (ROAD) study (2008–2010) and followed them up for 4 years. Handgrip strength, gait speed, skeletal muscle mass, and bone mineral density were assessed. SP was defined according to the Asian Working Group for Sarcopenia. OP was defined based on the World Health Organization criteria.
Results SP prevalence was 8.2% (men, 8.5%; women, 8.0%) in the second survey. In those with SP, 57.8% (21.9%; 77.6%) had OP at the lumbar spine L2–4 and/or femoral neck. SP cumulative incidence was 2.0%/year (2.2%/year; 1.9%/year). Multivariate regression analysis revealed that OP was significantly associated with SP occurrence within 4 years (odds ratio, 2.99; 95% confidence interval, 1.46–6.12; $p < 0.01$), but the reciprocal relationship was not significantly observed (2.11; 0.59–7.59; $p = 0.25$).
Conclusions OP might raise the short-term risk of SP incidence. Therefore, OP would not only increase the risk for osteoporotic fracture but may also increase the risk for SP occurrence.

Keywords Incidence · Osteoporosis · Osteosarcopenia · Population-based cohort study · Sarcopenia

✉ N. Yoshimura
yoshimuran-ort@h.u-tokyo.ac.jp

[1] Department of Joint Disease Research, 22nd Century Medical and Research Center, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8655, Japan

[2] Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo 113-8655, Japan

[3] Department of Orthopaedic Surgery, Sensory and Motor System Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

[4] JCHO Tokyo Shinjuku Medical Center, Tokyo 162-8542, Japan

[5] National Rehabilitation Center for Persons with Disabilities, Saitama 359-0042, Japan

Introduction

As the average age of the human population increases, there is an urgent need to develop strategies to prevent musculoskeletal disorders, which can impair activities of daily life (ADL) and quality of life (QOL) in the elderly. Sarcopenia (SP) and osteoporosis (OP) are major musculoskeletal diseases that impair ADL and QOL, leading to increased morbidity and mortality rates in the elderly. The recent National Livelihood Survey performed by the Ministry of Health, Labour, and Welfare in Japan [1] found that falls and osteoporotic fractures are ranked fourth, and frailty, to which muscle weakness and low physical performance contribute largely, was ranked third among the

causes of disabilities requiring support and long-term care. Therefore, developing approaches to prevent SP and OP could reduce ADL and QOL impairments and subsequent disabilities among the elderly.

In the elderly, SP is characterized by generalized loss of skeletal muscle mass and muscle strength and/or function, causing multiple adverse health outcomes, including physical disability, poor QOL, and death [2–7]. Although cross-sectional studies have investigated SP prevalence [8–14], the epidemiologic evidence of population-based samples remained insufficient. This might be because a widely accepted definition of SP was not established until the European Working Group on Sarcopenia in Older People (EWGSOP) developed a practical clinical definition and diagnostic criteria in 2010 [5]. There is a growing consensus that SP should not be defined based on muscle mass alone but also on muscle strength and function [5]. After publication of the EWGSOP consensus criteria, the Asian Working Group for Sarcopenia (AWGS) announced the appropriate diagnostic cutoff values for Asian populations [15]. In the AWGS consensus report, the reasons for creating different cutoff values from the European criteria were stated as follows: although the recommended approaches for measurements of muscle mass, muscle strength, and physical performance by AWGS were similar to the EWGSOP definition, the cutoff values of these measurements in Asian populations may differ from those in Caucasians because of ethnicities, body size, lifestyles, and cultural backgrounds. Therefore, developing an Asian consensus for sarcopenia diagnosis based on the evidence derived from Asian populations is essential for research and therapeutic approaches to sarcopenia in Asia [15]. This definition is now used widely for the assessment of SP in Asian countries.

The Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study is a prospective cohort study aimed at elucidating the environmental and genetic background of musculoskeletal diseases [16, 17]. The baseline data and that from the second survey of the ROAD study provided information on the prevalence and incidence of OP at the lumbar spine L2–4 and proximal femur [17, 18]. Furthermore, the prevalence of SP was evaluated using the EWGSOP definition in the second ROAD survey, giving an estimated prevalence in the general Japanese population of 13.8% in men and 12.4% in women [19]. The same study revealed that SP prevalence increased in an age-dependent manner in both sexes [19]. However, the SP incidence according to the AWGS criteria has not been investigated. Furthermore, the relationships of SP with other musculoskeletal diseases, especially OP, have not been determined, and it is not clear whether SP causes OP development, OP causes SP development, the conditions are comorbid, or if SP and OP represent concomitant modifications of one another.

In the present study, we completed the third ROAD study survey, a 4-year follow-up in which examinations identical to those in the second ROAD study survey were conducted. The aims of the present study were to clarify SP prevalence and incidence using the AWGS criteria, determine the co-existing proportions of SP and OP, and evaluate whether there was a significant contribution of SP to subsequent OP development, or vice versa, in elderly Japanese subjects.

Methods

Study participants

The present study was performed using the ROAD study cohorts that were established in 2005. The ROAD study is a national, prospective study of osteoarthritis that consists of population-based cohorts from several communities in Japan. Details of the cohort profiles have been reported elsewhere [16, 17]. In brief, between 2005 and 2007, a baseline database was created that included clinical and genetic information for 3040 residents (1061 men and 1979 women with a mean age of 70.3 (standard deviation [SD], 11.0) years; 71.0 (10.7) years in men, 69.9 (11.2) years in women). The subjects were recruited from resident registration listings in 3 communities with different characteristics: 1350 subjects from an urban region in Itabashi, Tokyo; 864 subjects from a mountainous region in Hidakagawa, Wakayama; and 826 subjects from a coastal region in Taiji, Wakayama.

After the baseline study, a second survey was performed in the same communities from 2008 to 2010 [20], and the third survey was followed from 2012 to 2013. In the second and third surveys, in addition to the OP assessment, examinations for the diagnosis of SP, including measurements such as gait speed, grip strength, and skeletal muscle mass were initiated in mountainous and coastal regions. In the present study, among the 1551 participants (521 men and 1030 women) in the second survey from mountainous and coastal regions who underwent all measurements for SP and OP, those aged ≥60 years were selected based on the AWGS criteria for SP [15]. As a result, 1099 (377 men and 722 women; mean age, 72.1 (7.4) years [72.7 (7.5) years in men, 71.8 (7.4) years in women]) participants were recruited as eligible subjects. A flow chart of subjects' recruitment and follow-up with reasons for dropout is shown in Fig. 1. The data obtained from these 1099 subjects was used to clarify mutual associations between SP and OP.

Examinations performed during the second ROAD study survey

Interviewer-administered questionnaire

Participants completed an interviewer-administered questionnaire that consisted of 200 questions related to lifestyle

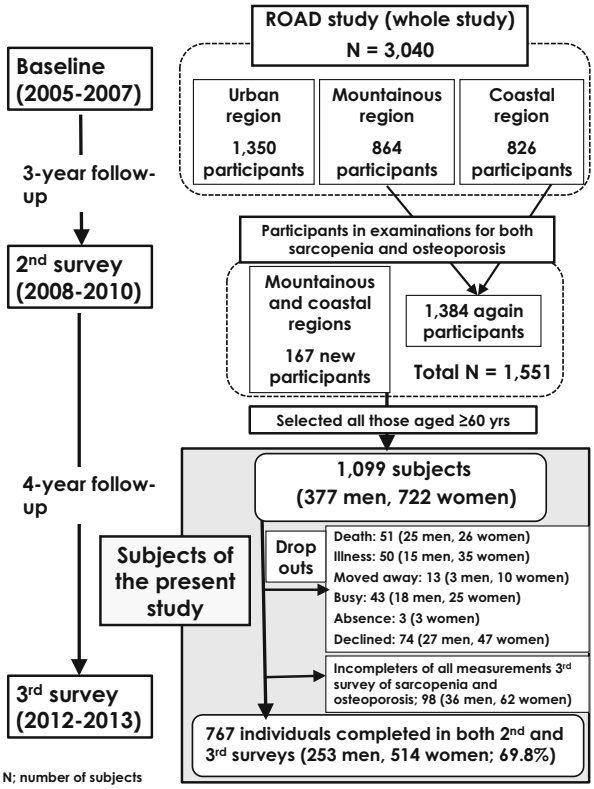


Fig. 1 Flow chart of the recruited participants in the present study

including occupation, smoking habits, alcohol consumption, family history, medical history, physical activity, reproductive history, and health-related QOL.

Anthropometric measurements and medical history

Anthropometric measurements, including height and weight, were measured in all participants. Body mass index (BMI; weight [kg]/height [m^2]) was calculated as weight in kilograms divided by height in meters squared. Experienced orthopedic surgeons collected medical information about pain, swelling, and the range of motion of the knee.

Skeletal muscle mass

Skeletal muscle mass was measured by bioimpedance analysis [21–25] using the Body Composition Analyzer MC-190 (Tanita Corp., Tokyo, Japan). The protocol was described by Tanimoto et al. [11, 13] and has been validated previously [26]. Appendicular skeletal muscle mass (ASM) was calculated as the sum of the muscle mass of the arms and legs. Absolute ASM was converted to an appendicular muscle mass index by dividing by height in meters squared (kg/m^2).

Muscle strength and walking speed

Handgrip strength was measured using a Toei Light handgrip dynamometer (Toei Light Co. Ltd., Saitama, Japan) to assess muscle strength. Both hands were tested and the largest value used to determine the maximum muscle strength. The usual walking speed was measured as an index of physical performance. The time taken (s) to walk 6 m at normal walking speed was recorded, and the usual gait speed was calculated.

BMD examination

Lumbar spine and proximal femur bone mineral density (BMD) values were determined using dual-energy X-ray absorptiometry (DXA; Hologic Discovery C; Hologic, Waltham, MA). To control DXA precision, the equipment was checked at every examination during the second and third surveys using the same phantom. The BMD of the phantom was regulated to within ±1.5% during all examinations. In addition, the same physician (N.Y.) examined all participants to control observer variability. Intraobserver variability of DXA using the Lunar DPX in vitro and in vivo had been measured by the same physician (N.Y.) in another study [27], and the coefficient of variance for L2–4 in vitro was 0.35%. The coefficients of variance for L2–4, the proximal femur, Ward's triangle, and the trochanter as examined in vivo in five male volunteers were 0.61–0.90, 1.02–2.57, 1.97–5.45, and 1.77–4.17%, respectively.

Definition of SP and OP

SP was determined according to the criteria defined by the AWGS [15]. AWGS criteria were as follows: (A) age ≥60 or ≥65 years; (B) low appendicular skeletal muscle mass, 7.0 kg/m^2 for men and 5.7 kg/m^2 for women, according to bioimpedance analysis; (C) low handgrip strength, <26 kg in men and <18 kg in women; and (D) low gait speed, with usual gait speed being ≤0.8 m/s. Subjects were diagnosed as having SP if they had criteria A and B and either of criteria C or D. Regarding age definition using the AWGS criteria, because of the different states of aging in Asia, not all countries use the same cutoff age to define elderly populations. Therefore, the AWGS recommends using either 60 or 65 years as the age for SP. In the present study, we defined subjects aged ≥60 years as potential subjects for SP.

In the present study, OP was defined according to the values of BMD. The World Health Organization criteria were used when BMD T-scores were lower than the peak bone mass by 2.5 SDs [28]. In Japan, the mean L2–4 BMD in young adults, as measured using the Hologic DXA, was 1.011 (0.119) g/cm^2 [29]. Therefore, lumbar spine osteoporosis was defined as an L2–4 BMD of <0.714 g/cm^2. In Japan, the mean femoral neck BMDs in young adult males and

females are 0.863 (0.127) and 0.787 (0.109) g/cm^2, respectively [28]. Therefore, OP at the femoral neck was defined as a BMD of <0.546 and <0.515 g/cm^2 for men and women, respectively.

Incidence of SP and OP

The cumulative incidences of SP and OP were determined based on changes in measurements between the second and third surveys. New cases of SP or OP were defined as when an individual did not meet criteria for SP or OP at the second survey but did meet them at the third survey.

Statistical analyses

All statistical analyses were performed using STATA statistical software (STATA Corp., College Station, TX). Differences in proportions were compared using the chi-square test. Differences in continuous variables were tested for significance using analysis of variance for comparisons among multiple groups or Scheffe's least significant difference test for group pairs.

Logistic regression analysis was used to test the association between the presence of SP and OP occurrence. OP occurrence was used as the objective variable, and the presence of SP (1: yes, 0: no) was used as the explanatory variable, after adjusting for age (years), sex (0: men, 1: women), and unconfirmed confounding factors. A second logistic regression analysis was conducted by replacing the objective and explanatory variables in the model mentioned above with SP occurrence and the presence of OP (1: yes, 0: no), respectively. The unconfirmed risk factors used for adjustment in the multivariate logistic analysis included residing area (0: mountainous area, 1: coastal area), emaciated stature (BMI <18.5 kg/m^2; [0: no, 1: yes]), current smoking habit (0: ex or never smoker, 1: current smoker), and alcohol consumption habit (0: ex or never drinker, 1: current drinker).

Results

SP prevalence

SP prevalence according to age group stratifications of 60–64, 65–69, 70–74, 75–79, and ≥80 years were 0.5, 0.0, 4.3, 11.2, and 27.0%, respectively (men, 1.5, 0.0, 4.7, 11.5, and 23.9%, for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively; women, 0.0, 0.0, 4.1, 10.9, and 28.7%, for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively). Above the age of 70 years, SP prevalence increased in an age-dependent manner, but there was no significant difference in prevalence according to sex.

OP prevalence

OP prevalence estimates were conducted on 1097 participants (376 men, 721 women) because the BMD at the spine L2–4 or femoral neck could not be measured in 2 individuals (1 man, 1 woman). OP prevalence according to age group stratifications of 60–64, 65–69, 70–74, 75–79, and ≥80 years were 10.8, 18.0, 19.5, 34.0, and 44.0%, respectively (men, 1.5, 1.5, 4.7, 13.8, and 11.3%, for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively; women, 15.4, 24.8, 26.9, 47.7, and 62.0%, for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively). OP prevalence increased in an age-dependent manner in women and was significantly higher in each age strata in women compared with that in men.

SP and OP co-existence

In the population aged ≥60 years, SP and OP co-existence was observed in 4.7%, SP alone was present in 3.5%, OP alone was noted in 20.2%, and 71.7% had neither SP nor OP.

In men, the prevalences of co-existing SP and OP, SP alone, OP alone, and neither SP nor OP were 1.9, 6.7, 5.1, and 86.4%, respectively, and in women, those were 6.2, 1.8, 28.0, and 63.9%, respectively. The difference in distribution in prevalences between men and women was most significant for OP. That is, prevalences of the co-existence of SP and OP and OP alone were significantly higher in women compared with men ($p < 0.001$).

Associated factors classified by the presence or absence of SP or OP

Table 1 shows a comparison of background characteristics for those with and without SP. Among subjects with SP, 57.8% had a concomitant diagnosis of OP, which was a significantly higher proportion than those without SP (22.0%, $p < 0.001$). Similarly, in those with OP, 19.1% had a concomitant diagnosis of SP, which was a significantly greater proportion than those without OP (4.6%, $p < 0.001$).

Diagnostic SP values such as grip strength and usual walking speed were significantly lower in the subjects with OP ($p < 0.001$). In addition, OP diagnostic values such as lumbar spine L2–4 and femoral neck BMD were significantly lower in the subjects with SP ($p < 0.0001$). Age and smaller stature were both significantly associated with SP and OP. Residing region was significantly associated with SP ($p = 0.005$). Being

Table 1 Comparison of characteristics at the second survey classified by the presence or absence of sarcopenia or osteoporosis

	Sarcopenia			Osteoporosis		
	Sarcopenia (−) (n = 1009)	Sarcopenia (+)	p value	Osteoporosis (−) (n = 824)	Osteoporosis (+) (n = 273)	p value
Mean values (SD) and percentage of selected characteristics						
Age (years)	71.3 (7.0)	81.1 (5.9)	<0.0001	71.0 (7.1)	75.5 (7.3)	<0.0001
Female sex (%)	65.8	64.4	0.794	57.5	90.5	<0.001
Height (cm)	154.1 (8.8)	149.0 (8.8)	<0.0001	155.5 (8.7)	148.1 (7.4)	<0.0001
Weight (kg)	55.8 (10.1)	45.6 (8.2)	<0.0001	57.6 (9.8)	47.1 (7.4)	<0.0001
BMI (kg/m^2)	23.4 (3.3)	20.4 (2.4)	<0.0001	23.8 (3.2)	21.5 (3.0)	<0.0001
Emaciation (BMI < 18.5 kg/m^2; %)	5.0	21.1	<0.001	3.3	15.4	<0.001
Residing in a coastal area (%)	47.5	32.2	0.005	46.4	45.8	0.870
Current smoking habit (%)	8.4	10.0	0.613	10.1	4.0	0.002
Current alcohol drinking habit (%)	31.2	25.6	0.268	35.3	16.5	<0.0001
Mean values (SD) of selected measurements for sarcopenia diagnosis						
Grip strength (maximum) (kg)	29.0 (8.6)	20.5 (7.4)	<0.0001	30.2 (8.8)	22.7 (6.0)	<0.0001
Usual walking speed (m/s)	1.10 (0.26)	0.75 (0.19)	<0.0001	1.10 (0.27)	0.99 (0.29)	0.035
Appendicular skeletal muscle mass adjusted by height (kg/m^2)	6.65 (0.99)	5.69 (0.63)	<0.0001	6.80 (0.99)	5.88 (0.69)	<0.001
Prevalence of sarcopenia (%)	0.0	100.0	–	4.6	19.1	<0.001
Measurements related to the presence of osteoporosis						
BMD (L2–4) (g/cm^2)	0.923 (0.207)	0.823 (0.205)	<0.0001	0.981 (0.186)	0.713 (0.129)	<0.0001
BMD (femoral neck) (g/cm^2)	0.640 (0.126)	0.552 (0.119)	<0.0001	0.679 (0.109)	0.492 (0.06)	<0.0001
Prevalence of osteoporosis (L2–4 or femoral neck; %)	22.0	57.8	<0.001	0.0	100.0	–

n number of subjects, *BMI* body mass index, *BMD* bone mineral density, *SD* standard deviation

female and drinking and smoking less were significantly associated with OP (Table 1).

Table 2 (A), (B) shows the mutual associations between the presence of SP and OP at the lumbar spine L2–4 and/or the femoral neck. After adjustment for potential confounding factors mentioned above, logistic regression analysis revealed that the presence of OP was significantly associated with SP presence (odds ratio, 2.86; 95% confidence interval, 1.59–5.13; $p < 0.001$; Table 2 (A)). Furthermore, the presence of SP was significantly associated with OP presence (odds ratio, 2.78; 95% confidence interval, 1.55–4.99; $p < 0.001$; Table 2 (B)).

Participants in both the second and third surveys

Among 1099 participants in the second survey who were aged ≥60 years at the assessment of SP, 865 individuals (78.7%, 289 men, 576 women) attended the third survey performed 4 years later. Therefore, 234 individuals (21.3%; 88 men, 146 women) dropped out in the third survey. The reasons for dropout are shown in Fig. 1. Among the 865 participants in both the second and third surveys, 98 (11.3%, 36 men, 62 women) did not have complete measurements for the diagnosis of SP and OP. Therefore, the data from 767 completers (69.8%, 253 men, 514 women) was used in the present study to assess the contribution of OP to the occurrence of SP, and vice versa.

Cumulative incidence of SP

Among 767 completers (253 men, 514 women) of the third survey of the ROAD study, 32 subjects (9 men and 23 women) were diagnosed with SP at the second survey. Therefore, the number of population at risk for SP occurrence was 735 (244 men, 491 women). The cumulative incidence of SP during the 4-year period between the surveys was 2.0%/year (men, 2.2%/year; women, 1.9%/year). Figure 2 shows the age-sex classified SP incidence. The cumulative SP incidences for the at-risk populations according to age group were 0.4, 0.5, 1.5, 4.2, and 6.9%/year for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively. The incidence increased in an age-dependent fashion ($p < 0.001$ for all subjects, $p = 0.005$ for men, and $p < 0.001$ for women), but there was no

Table 2 Mutual associations between presence of sarcopenia and osteoporosis among subjects at the second survey

Logistic regression model				
Objective variable	Reference			
A. Effect of the presence of osteoporosis on sarcopenia presence				
Sarcopenia	0: no, 1: yes			
Explanatory variables	Reference	OR	95% CI	*p* value
Osteoporosis at lumbar spine L2–4 or femoral neck	0: no, 1: yes	2.75	1.59–5.13	<0.001
Adjusted factors				
Age (years)	1+ year	1.22	1.17–1.28	<0.001
Sex	0: men, 1: women	0.69	0.37–1.29	0.247
Residing area	0: mountainous area, 1: coastal area	0.63	0.37–1.06	0.082
Emaciation (BMI < 18.5 kg/m^2)	0: no, 1: yes	3.19	1.58–6.44	0.001
Current smoking habit	0: ex or never smoker, 1: current smoker	1.90	0.77–4.67	0.162
Current alcohol drinking habit	0: ex or never drinker, 1: current drinker	0.97	0.53–1.78	0.930
B. Effect of the presence of sarcopenia on osteoporosis presence				
Osteoporosis at lumbar spine L2–4 or femoral neck	0: no, 1: yes			
Explanatory variables	Reference	OR	95% CI	*p* value
Sarcopenia	0: no, 1: yes	2.78	1.55–4.99	0.001
Adjusted factors				
Age (years)	1+ year	1.09	1.06–1.12	<0.001
Sex	0: men, 1: women	8.94	5.43–14.8	<0.001
Region	0: mountainous area, 1: coastal area	1.12	0.82–1.55	0.469
Emaciation (BMI < 18.5 kg/m^2)	0: no, 1: yes	5.28	2.86–9.76	<0.001
Current smoking habit	0: ex or never smoker, 1: current smoker	0.80	0.36–1.75	0.573
Current alcohol drinking habit	0: ex or never drinker, 1: current drinker	0.81	0.53–1.24	0.330

OR odds ratio, *95% CI* 95% confidence interval, *BMI* body mass index

significant difference in incidence according to sex ($p = 0.61$) (Fig. 2).

Cumulative incidence of OP

Among 767 completers (253 men, 514 women), 2 male subjects were excluded from the assessment of OP incidence at the third survey because their BMD measurements for both lumbar spine L2–4 and femoral neck could not be performed. Among the remaining 765 subjects (251 men, 514 women), 90 (2 men and 88 women) were diagnosed with OP at both the lumbar spine L2–4 and femoral neck at the second survey. Therefore, in the present study, the population at risk for OP at the lumbar spine L2–4 and/or femoral neck was 675 subjects (249 men, 426 women). The cumulative OP incidence during the 4-year period between the surveys was 1.9%/year (men, 0.8%/year; women, 2.5%/year). The cumulative OP incidences for the at-risk populations according to age group were 1.1, 2.3, 2.1, 1.7, and 2.7%/year for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively (men, 1.0, 1.0, 0.9, 0.0, and 1.5%/year, for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively; women, 1.1, 2.9, 2.8, 3.0, and 4.0%/year, for 60–64, 65–69, 70–74, 75–79, and ≥80 years, respectively.) OP incidence was not associated with age ($p = 0.38$ for total subjects, $p = 0.60$ for men, $p = 0.23$ for women), and it was significantly higher in women compared with that in men ($p = 0.001$).

Assessment of contribution of OP to the subsequent occurrence of SP, and vice versa

Table 3 shows the comparison of the background characteristics according to the occurrence or non-occurrence of SP during the 4-year follow-up. Among subjects without SP at the second survey, in addition to higher age, lower height, lower weight, and residing in a mountainous area, the presence of OP was significantly associated with future SP development ($p < 0.001$).

Table 3 also shows the comparison of the background characteristics according to the occurrence or non-occurrence of OP during the 4-year follow-up. In addition to female sex, lower height, lower weight, emaciation, and residing in a mountainous area, the presence of SP was significantly associated with the future incidence of OP ($p = 0.043$).

After adjustment for the potential confounding factors such as, age, sex, regional difference, emaciation (BMI < 18.5 kg/

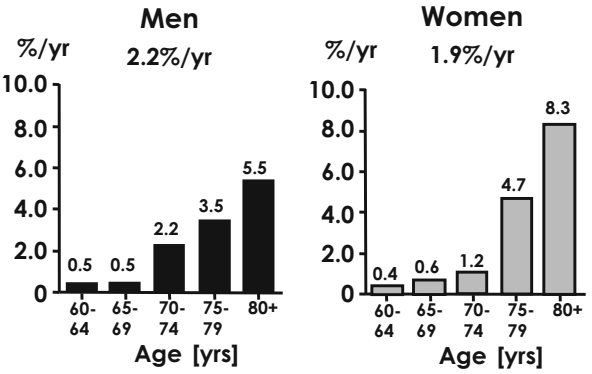


Fig. 2 Cumulative incidences (%/year) of sarcopenia stratified by age and sex

m^2), smoking habit, and alcohol drinking habit, logistic regression analysis revealed that the presence of OP was a significant predictive factor for SP occurrence in the near future (odds ratio, 2.99; 95% confidence interval, 1.46–6.12; $p = 0.003$; Table 4 (A)). This tendency was shown for both men and women when the logistic analysis was performed using identical adjustment factors except for sex; although, the association in men was diluted (men: odds ratio 6.92, 95% confidence interval 0.86–55.66, $p = 0.069$; women: odds ratio 2.58, 95% confidence interval 1.16–5.73, $p = 0.020$).

By contrast, the logistic regression analysis revealed that the presence of SP was no longer a significant predictive factor for OP occurrence in the near future (odds ratio, 2.11; 95% confidence interval, 0.59–7.59; $p = 0.253$; Table 4 (B)).

Discussion

In the present study, using information from the second and third surveys of the population-based ROAD cohort, we clarified the prevalence and characteristics of SP in Japan. We found that the prevalence of SP was significantly higher in those with OP compared to in those without OP. In addition, the prevalence of OP was significantly higher in those with SP compared to in those without SP. In the 4-year follow-up between the surveys, we estimated the SP incidence and found that the presence of OP significantly increased the future risk of SP, but the presence of SP did not increase the future risk of OP.

In the present study, SP prevalence was estimated using the AWGS definition because previous prevalence estimates on this cohort were conducted before the publication of the AWGS definitions [19]. The previous prevalences were higher compared with those noted in the present study, which is most likely because of the differences in cutoff values between the EWGSOP and AWGS definition criteria. According to the EWGSOP criteria, low handgrip strength was defined as <30 kg in men and <20 kg in women [5], whereas those for the AWGS definition are <26 and <18 kg, respectively. However, because all of the participants in the ROAD study were Japanese, we decided that the AWGS criteria would better reflect the SP prevalence of the cohort.

Besides previous reports using the EWGSOP definition [19], few studies have estimated the SP prevalence in the Japanese population. Applying the SP prevalence rate obtained in subjects aged ≥60 years in the present study (8.2%) to the Japanese 2010 census data [30] would indicate that in Japan, approximately 3,700,000 people (1,200,000 men and 2,500,000 women) aged ≥60 years might be affected by SP. Furthermore, in the present report, the degree of co-existence of SP and OP in those aged ≥60 years was clarified. The majority of patients with SP had OP, but patients with OP did not always have SP. Therefore, individuals with SP should be assessed for the potential co-existence of OP. Furthermore, not only was the presence of OP associated with the presence of SP, and vice versa, but also subjects with SP tended to have low BMD, whereas those with OP tended to have low physical performance and low muscle mass. Therefore, not only prevalence but also elements assessed for the diagnosis of SP and OP showed significant associations.

Regarding stature, emaciation is a well-known feature of OP. In a meta-analysis of prospective cohorts from >25 countries, including baseline BMI data from 398,610 women with an average age of 63 years and a follow-up of 2.2 million person-years, Johansson et al. reported that a high BMI was a protective factor for most fragility fracture sites [31]. Moreover, we reported previously that fast bone losers have significantly lighter body composition compared with healthy subjects [32]. However, regarding SP, despite consideration of sarcopenic obesity [33, 34], the association of emaciation and SP has received little attention. In the present study, emaciation was significantly associated with both OP and SP. Additionally, in the present study, none of the individuals with SP were obese (BMI >27.5 kg/m^2). In an overview of sarcopenic obesity, Cauley stated that obesity was usually defined by a high BMI, but some studies have relied on percentage body or visceral fat [35]. The findings of the present study suggest that high BMI might not be associated with the existence of SP. The definition of sarcopenic obesity should be incorporated from the view of the prevention of severe health illness of the elderly, such as cardiovascular diseases.

Regarding SP incidence, few reports have estimated SP incidence in not only Japan but also worldwide. In the present study, the cumulative SP incidence in Japanese subjects aged ≥60 years was 2.0%/year. Using the age-sex SP incidence against the Japanese 2010 census data [31], this suggests that approximately 1,050,000 people (350,000 men and 700,000 women) aged ≥60 years become newly affected by SP each year. The cumulative SP incidence increased with age, but there were no

Table 3 Comparison of characteristics of the subjects at the second survey classified by the occurrence or non-occurrence of sarcopenia or osteoporosis during a 4-year follow-up

	Sarcopenia (population at risk, n = 735)			Osteoporosis (population at risk, n = 675)		
	Non-occurrence (n = 677)	Occurrence (n = 58)	p value	Non-occurrence (n = 624)	Occurrence (n = 51)	p value
Mean values (SD) and percentage of selected characteristics						
Age (years)	69.5 (6.3)	76.2 (6.0)	<0.0001	70.1 (6.7)	71.3 (6.1)	0.223
Female sex (%)	67.1	63.8	0.612	61.4	84.3	<0.001
Height (cm)	154.9 (8.4)	152.2 (8.9)	0.0217	155.4 (8.5)	151.2 (7.9)	0.0008
Weight (kg)	56.7 (9.5)	49.2 (6.8)	<0.0001	57.2 (9.5)	49.9 (6.9)	<0.001
BMI (kg/m^2)	23.6 (3.1)	21.2 (2.2)	<0.0001	23.6 (3.1)	21.8 (2.8)	0.0001
Emaciation (BMI < 18.5 kg/m^2; %)	4.0	6.9	0.290	2.7	11.8	0.001
Residing in a coastal area (%)	51.3	32.8	0.007	50.0	37.3	<0.001
Current smoking habit (%)	7.6	8.8	0.758	8.6	4.1	0.269
Current alcohol drinking habit (%)	33.0	29.3	0.567	34.9	24.0	0.117
Mean values (SD) of selected measurements for sarcopenia diagnosis						
Grip strength (maximum) (kg)	30.3 (8.4)	25.9 (6.9)	0.0003	30.4 (8.6)	25.9 (6.8)	0.0003
Usual walking time (m/s)	1.16 (0.24)	1.02 (0.19)	<0.0001	1.15 (0.25)	1.08 (0.21)	0.0789
Appendicular skeletal muscle mass adjusted by height (kg/m^2)	6.70 (0.97)	6.00 (0.65)	<0.0001	6.75 (0.97)	6.17 (0.68)	<0.0001
Prevalence of sarcopenia (%)	0.0	0.0	–	2.7	7.8	0.043
Measurements related to the presence of osteoporosis						
BMD (L2–4) (g/cm^2)	0.929 (0.194)	0.884 (0.214)	0.104	0.968 (0.176)	0.797 (0.137)	<0.0001
BMD (femoral neck) (g/cm^2)	0.651 (0.119)	0.119) 0.596 (0.125)	0.0008	0.668 (0.116)	0.556 (0.042)	<0.0001
Prevalence of osteoporosis (L2–4 or femoral neck; %)	16.3	39.7	<0.001	0.0	0.0	–

n number of subjects, *BMI* body mass index, *BMD* bone mineral density, *SD* standard deviation

significant differences in the SP prevalence or incidence rates according to sex.

We reported the cumulative OP incidence previously, using the 3-year follow-up data from the baseline to second ROAD study surveys [36]. In that study, we estimated that the annual cumulative OP incidences were 0.76%/year at the lumbar spine and 1.83%/year at the femoral neck. In the present study, the annual OP incidence of subjects aged ≥60 years, between the second and third surveys, was estimated to clarify any associations with SP. The incidence of lumbar spine L2–4 OP in female subjects aged ≥60 years who participated in the baseline to the second survey [35] was compared with that in the present study (1.06 vs. 0.84%/year, respectively; $p < 0.01$; data not published). Similarly, the incidence of femoral neck OP in female subjects between these studies was significantly lower between the second and third surveys (2.49 vs. 1.87%/year, respectively; $p < 0.001$; data not published). We did not compare the OP incidence in men because the numbers were too low to provide statistical power for a comparison. This comparison shows that the OP incidence rate in women might be decreasing, although the reasons for this are unknown. Observation of the ROAD cohort is ongoing, and changes in incidence rates will be clarified after completion of the 10-year follow-up.

Finally, the logistic regression analysis revealed that the presence of OP significantly increased the risk of SP occurrence within 4 years. By contrast, the presence of SP did not predict OP occurrence within 4 years. However, as we noted, there was a significant proportion of patients with co-existent SP and OP (so-called 'osteosarcopenia'), suggesting that individuals with SP should be assessed for the presence of OP.

There are several limitations to the present study. First, although the ROAD study includes a large number of participants, the participants in the present study (second survey, individuals from the mountainous and coastal regions only) may not be completely representative of the general population. To address this issue, we compared the anthropometric measurements between the present study participants and the general Japanese population. The values for the general population were obtained from

Table 4 Mutual associations between the occurrence and presence of sarcopenia and osteoporosis

Logistic regression model				
Objective variable	Reference			
A. Effect of the presence of osteoporosis on sarcopenia occurrence				
Sarcopenia occurrence	0: no, 1: yes			
Explanatory variables	Reference	OR	95% CI	*p* value
Osteoporosis presence at the lumbar spine L2–4 or femoral neck	0: no, 1: yes	2.99	1.46–6.12	0.003
Adjusted factors				
Age (years)	1+ year	1.18	1.12–1.25	<0.001
Sex	0: men, 1: women	0.81	0.38–1.72	0.582
Residing area	0: mountainous area, 1: coastal area	0.45	0.24–0.85	0.013
Emaciation (BMI < 18.5 kg/m^2; %)	0: no, 1: yes	1.37	0.40–4.67	0.618
Current smoking habit	0: ex or never smoker, 1: current smoker	1.93	0.66–5.62	0.229
Current alcohol drinking habit	0: ex or never drinker, 1: current drinker	0.86	0.43–1.72	0.666
B. Effect of presence of sarcopenia for the occurrence of osteoporosis				
Osteoporosis occurrence at the lumbar spine L2–4 or femoral neck	0: no, 1: yes			
Explanatory variables	Reference	OR	95% CI	*p* value
Sarcopenia presence	0: no, 1: yes	2.11	0.59–7.59	0.253
Adjusted factors				
Age (years)	1+ year	1.03	0.98–1.08	0.280
Sex	0: men, 1: women	3.48	1.46–8.28	0.005
Region	0: mountainous area, 1: coastal area	0.51	0.27–0.85	0.033
Emaciation (BMI < 18.5 kg/m^2; %)	0: no, 1: yes	5.14	1.80–14.68	0.002
Current smoking habit	0: ex or never smoker, 1: current smoker	0.69	0.15–1.94	0.636
Current alcohol drinking habit	0: ex or never drinker, 1: current drinker	0.92	0.44–1.94	0.832

OR odds ratio, *95% CI* 95% confidence interval, *BMI* body mass index

the report on the 2008 National Health and Nutrition Survey conducted by the Ministry of Health, Labour, and Welfare, Japan [37] when the second ROAD survey began. For mean BMI, there were no significant differences between the second ROAD survey participants and the Japanese general population. In addition, among lifestyle factors, the proportion of current smokers and drinkers (those who regularly drink more than one drink/month) in the Japanese general population was compared with that in the present study population. The proportion of current smokers was lower in males in the present study population compared with the general Japanese population, but there was no significant difference in the number of female smokers (men, 19.1 vs. 25.6%, $p < 0.01$; women, 3.1 vs. 4.0%, $p = 0.28$). Moreover, the proportion of current drinkers was significantly lower in both men and women in our study population compared with the general Japanese population (men, 58.9 vs. 64.7%, $p < 0.05$; women, 16.0 vs. 21.0%, $p < 0.01$), suggesting that compared to the general Japanese population, the participants of the present study lead healthier lifestyles, at least in terms of smoking habits. This selection bias should be taken into consideration when generalizing the results obtained from the present study. Second, in the present study, handgrip strength, and 6-m walking tests were measured only once. Therefore, we could not exclude the effect of incidental changes in participants' performance around the examination date. Recurrent measurements should be taken into consideration to minimize fluctuation of measurements. However, we confirmed that none of the participants harbored hand or knee injuries that could have affected the measurements. Third, in the present study, OP was defined by BMD values using DXA alone. We might have to include participants who started medication for OP and/or those who developed new fractures. Although we had information regarding the medication and history of fractures, they were obtained from the self-report questionnaire leading to the possibility of recall bias. Therefore, in the present study, the incidence of OP might be underestimated. After the confirmation of medication by the interviewer, and assessment of fractures diagnosed by radiographic examinations performed in the ROAD study, the cumulative incidence of OP should be revised. Finally, the 4-year follow-up period might be too short to determine the causal relationship between SP and OP. Only a

small number of new OP and SP cases occurred during the 4-year observation period. However, the ROAD study continues, so determining the occurrence of OP and SP over an extended period will be possible in the future, enabling the validation of the causal relationship between SP and OP using the incidence rate, rather than cumulative incidence as an epidemiological index.

In conclusion, the prevalence of co-existing SP and OP were high, suggesting that subjects ≥60 years with SP should be assessed for concomitant OP. Moreover, the presence of OP was significantly associated with SP occurrence within 4 years. Therefore, treatment for OP might not only have clinical benefit for the treatment of OP itself but might also reduce the risk of subsequent SP development.

Acknowledgments This work was supported by Grant-in-Aid funding from the Ministry of Health, Labour and Welfare: H17-Men-eki-009 (Director, Kozo Nakamura), H20-Choujyu-009 (Director, Noriko Yoshimura), H23-Choujyu-002 (Director, Toru Akune), H25-Choujyu-007 (Director, Noriko Yoshimura), and H25-Nanchitou (Men)-005 (Director, Sakae Tanaka). The study was also supported by Scientific Research grants B26293139, B23390172, and B20390182 and Challenging Exploratory Research grants 15K15219 and 24659317 to Noriko Yoshimura; Scientific Research grants B23390356 and C20591774 and Challenging Exploratory Research grant 23659580 to Shigeyuki Muraki; Challenging Exploratory Research grants 24659666 and 21659349 and Young Scientists A18689031 to Hiroyuki Oka; Scientific Research grants B26293329, B23390357, C20591737 and Challenging Exploratory Research grant 25670293 to Toru Akune; and Collaborating Research with NSF from the Ministry of Education, Culture, Sports, Science, and Technology in Japan 08033011-00262 (Director, Noriko Yoshimura). The study was also supported by grants from the Japan Osteoporosis Society (Noriko Yoshimura, Shigeyuki Muraki, Hiroyuki Oka, and Toru Akune), Japan Osteoporosis Foundation (2015, Noriko Yoshimura), research aids from the Japanese Orthopaedic Association (JOA—Subsidized Science Project Research 2006-1 and 2010-2, Director, Hiroshi Kawaguchi; 2014-1, Director, Kozo Nakamura), and the Japanese Society for Musculoskeletal Medicine (2015, Director, Shigeyuki Muraki). The authors wish to thank Dr. Naoki Hirabayashi of the Kawakami Clinic, Hidakagawa Town; Mrs. Tomoko Takijiri, Mrs. Rie Takiguchi, Mrs. Kyoko Maeda, Ms. Ikuyo Ueyama, Mrs. Michiko Mori, Mrs. Hisayo Sugimoto, and other members of the public office in Hidakagawa Town; and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, Mrs. Megumi Takino, Mrs. Shuko Okada, Mrs. Kazuyo Setoh, Mrs. Chise Ryouno, Mrs. Miki Shimosaki, Mrs. Chika Yamaguchi, Mrs. Yuki Shimoji, and other members of the public office in Taiji Town for their assistance in locating and scheduling participants for examinations. We would also like to thank Ms. Kyoko Yoshimura, Mrs. Toki Sakurai, Mrs. Saeko Sahara, and Mr. Noriyuki Oe for their assistance with data reduction and administration.

Compliance with ethical standards

Conflicts of interest None.

Ethical approval All participants provided written informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo (Nos. 1264 and 1326) and the University of Wakayama Medical University (No. 373). All procedures were conducted in accordance with the ethical standards as described in the 1964 Declaration of Helsinki, and its later amendments.

References

1. Ministry of Health, Labour and Welfare. The outline of the results of the National Livelihood Survey (2013) available at http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa13/dl/16.pdf [in Japanese]
2. Rosenberg I (1989) Summary comments: epidemiological and methodological problems in determining nutritional status of older persons. Am J Clin Nutr 50:1231–1233
3. Rosenberg IH (1997) Sarcopenia: origins and clinical relevance. J Nutr 127(5 Suppl):990S–991S
4. Morley JE, Baumgartner RN, Roubenoff R, Mayer J, Nair KS (2001) Sarcopenia. J Lab Clin Med 137:231–243
5. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, Martin FC, Michel JP, Rolland Y, Schneider SM, Topinkova E, Vandewoude M, Zamboni M, European Working Group on Sarcopenia in Older People (2010) Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on Sarcopenia in Older People. Age Ageing 39:412–423. doi:10.1093/ageing/afq034
6. Delmonico MJ, Harris TB, Lee JS, Visser M, Nevitt M, Kritchevsky SB, Tylavsky FA, Newman AB, Health, Aging and Body Composition Study (2007) Alternative definitions of sarcopenia, lower extremity performance, and functional impairment with aging in older men and women. J Am Geriatr Soc 55:769–774
7. Goodpaster BH, Park SW, Harris TB, Kritchevsky SB, Nevitt M, Schwartz AV, Simonsick EM, Tylavsky FA, Visser M, Newman AB (2006) The loss of skeletal muscle strength, mass, and quality in older adults: the health, aging and body composition study. J Gerontol A Biol Sci Med Sci 61:1059–1064
8. Baumgartner RN, Koehler KM, Gallagher D, Romero L, Heymsfield SB, Ross RR, Garry PJ, Lindeman RD (1998) Epidemiology of sarcopenia among the elderly in New Mexico. Am J Epidemiol 147:755–763
9. Melton LJ III, Khosla S, Crowson CS, O'Connor MK, O'Fallon WM, Riggs BL (2000) Epidemiology of sarcopenia. J Am Geriatr Soc 48:625–630
10. Iannuzzi-Sucich M, Prestwood KM, Kenny AM (2002) Prevalence of sarcopenia and predictors of skeletal muscle mass in healthy, older men and women. J Gerontol A Biol Sci Med Sci 57:M772–M777
11. Tanimoto Y, Watanabe M, Sun W, Sugiura Y, Tsuda Y, Kimura M, Hayashida I, Kusabiraki T, Kono K (2012) Association between sarcopenia and higher-level functional capacity in daily living in community-dwelling elderly subjects in Japan. Arch Gerontol Geriatr 55:e9–e13
12. Patel HP, Syddall HE, Jameson K, Robinson S, Denison H, Roberts HC, Edwards M, Dennison E, Cooper C, Aihie Sayer A (2013) Prevalence of sarcopenia in community-dwelling older people in the UK using the European Working Group on Sarcopenia in Older People (EWGSOP) definition: findings from the Hertfordshire Cohort Study (HCS). Age Ageing 42:378–384. doi:10.1093/ageing/afs197
13. Tanimoto Y, Watanabe M, Sun W, Tanimoto K, Shishikura K, Sugiura Y, Kusabiraki T, Kono K (2013) Association of sarcopenia with functional decline in community-dwelling elderly subjects in Japan. Geriatr Gerontol Int 13:958–963. doi:10.1111/ggi.12037
14. Lin CC, Lin WY, Meng NH, Li CI, Liu CS, Lin CH, Chang CK, Lee YD, Lee CC, Li TC (2013) Sarcopenia prevalence and associated factors in an elderly Taiwanese metropolitan population. J Am Geriatr Soc 61:459–462. doi:10.1111/jgs.12129

15. Chen LK, Liu LK, Woo J, Assantachai P, Auyeung TW, Bahyah KS, Chou MY, Chen LY, Hsu PS, Krairit O, Lee JS, Lee WJ, Lee Y, Liang CK, Limpawattana P, Lin CS, Peng LN, Satake S, Suzuki T, Won CW, Wu CH, Wu SN, Zhang T, Zeng P, Akishita M, Arai H (2014) Sarcopenia in Asia: consensus report of the Asian Working Group for Sarcopenia. J Am Med Dir Assoc 15:95–101. doi:10.1016/j.jamda.2013.11.025
16. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T (2010) Cohort profile: research on osteoarthritis/osteoporosis against disability study. Int J Epidemiol 39: 988–995. doi:10.1093/ije/dyp276
17. Yoshimura N, Muraki S, Oka H, Mabuchi A, En-Yo Y, Yoshida M, Saika A, Yoshida H, Suzuki T, Yamamoto S, Ishibashi H, Kawaguchi H, Nakamura K, Akune T (2009) Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study. J Bone Miner Metab 27:620–628. doi:10.1007/s00774-009-0080-8
18. Yoshimura N, Muraki S, Oka H, Nakamura K, Kawaguchi H, Tanaka S, Akune T (2015) Serum levels of 25-hydroxyvitamin D and the occurrence of musculoskeletal diseases: a 3-year follow-up to the road study. Osteoporos Int 26:151–161. doi:10.1007/s00198-014-2844-9
19. Akune T, Muraki S, Oka H, Tanaka S, Kawaguchi H, Nakamura K, Yoshimura N (2014) Exercise habits during middle age are associated with lower prevalence of sarcopenia: the ROAD study. Osteoporos Int 25:1081–1088. doi:10.1007/s00198-013-2550-z
20. Yoshimura N, Oka H, Muraki S, Akune T, Hirabayashi N, Matsuda S, Nojiri T, Hatanaka K, Ishimoto Y, Nagata K, Yoshida M, Tokimura F, Kawaguchi H, Nakamura K (2011) Reference values for hand grip strength, muscle mass, walking time, and one-leg standing time as indices for locomotive syndrome and associated disability: the second survey of the ROAD study. J Orthop Sci 16: 768–777. doi:10.1007/s00776-011-0160-1
21. No authors listed (1996) Bioelectrical impedance analysis in body composition measurement: National Institutes of Health Technology Assessment Conference Statement. Am J Clin Nutr 64(3 Suppl):524S–532S
22. Janssen I, Heymsfield SB, Baumgartner RN, Ross, R (2000) Estimation of skeletal muscle mass by bioelectrical impedance analysis. J Appl Physiol (1985) 89:465–471
23. Kyle UG, Genton L, Slosman DO, Pichard C (2001) Fat-free and fat mass percentiles in 5225 healthy subjects aged 15 to 98 years. Nutrition 17:534–541
24. Kyle UG, Genton L, Karsegard L, Slosman DO, Pichard C (2001) Single prediction equation for bioelectrical impedance analysis in adults aged 20–94 years. Nutrition 17:248–253
25. Roubenoff R, Baumgartner RN, Harris TB, Dallal GE, Hannan MT, Economos CD, Stauber PM, Wilson PW, Kiel DP (1997) Application of bioelectrical impedance analysis to elderly populations. J Gerontol A Biol Sci Med Sci 52: M129–M136
26. Nemoto M, Yasbushita N, Kim M, Tomoaki M, Satoshi S, Jung S, Hiroyuki S, Kiyoji T (2012) Validity of the bioelectrical impedance method for assessing body composition in non-frail and pre-frail older adults. Int J Body Comps Res 10:225–262
27. Yoshimura N, Kakimoto T, Nishioka M, Kishi T, Iwasaki H, Niwa T, Morioka S, Sakata T, Hashimoto T (1997) Evaluation of reproducibility of bone mineral density measured by dual energy X-ray absorptiometry (Lunar DPX-L). J Wakayama Medical Society 48:461–466
28. World Health Organization (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. In: WHO Technical Report Series 843. WHO, Geneva
29. Orimo H, Hayashi Y, Fukunaga M, Sone T, Fujiwara S, Shiraki M, Kushida K, Miyamoto S, Soen S, Nishimura J, Oh-Hashi Y, Hosoi T, Gorai I, Tanaka H, Igai T, Kishimoto H, Osteoporosis Diagnostic Criteria Review Committee: Japanese Society for Bone and Mineral Research (2001) Diagnostic criteria for primary osteoporosis: year 2000 revision. J Bone Miner Metab 19:331–337
30. Portal site of Official Statistics of Japan (2010) Population Census 2010 available at http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&tclassID=000001034991&cycleCode=0&requestSender=search [In Japanese]
31. Johansson H, Kanis JA, Odén A, McCloskey E, Chapurlat RD, Christiansen C, Cummings SR, Diez-Perez A, Eisman JA, Fujiwara S, Glüer CC, Goltzman D, Hans D, Khaw KT, Krieg MA, Kröger H, LaCroix AZ, Lau E, Leslie WD, Mellström D, Melton LJ 3rd, O'Neill TW, Pasco JA, Prior JC, Reid DM, Rivadeneira F, van Staa T, Yoshimura N, Zillikens MC (2014) A meta-analysis of the association of fracture risk and body mass index in women. J Bone Miner Res 29:223–233. doi:10.1002/jbmr.2017
32. Yoshimura N, Takijiri T, Kinoshita H, Danjoh S, Kasamatsu T, Morioka S, Sakata K, Hashimoto T, Takeshita T (2004) Characteristics and course of bone mineral densities among fast bone losers in a rural Japanese community: the Miyama study. Osteoporos Int 15:139–144
33. Kohara K (2014) Sarcopenic obesity in aging population: current status and future directions for research. Endocrine 45:15–25
34. Vincent HK, Raiser SN, Vincent KR (2012) The aging musculoskeletal system and obesity-related considerations with exercise. Ageing Res Rev 11:361–373. doi:10.1016/j.arr.2012.03.002
35. Cauley JA (2015) An overview of sarcopenic obesity. J Clin Densitom 18:499–505. doi:10.1016/j.jocd.2015.04.013
36. Yoshimura N, Muraki S, Oka H, Nakamura K, Kawaguchi H, Tanaka S, Akune T (2015) Serum levels of 25-hydroxyvitamin D and occurrence of musculoskeletal diseases, such as osteoporosis, knee osteoarthritis and lumbar spondylosis: a three-year follow-up of the road study. Osteoporos Int 26:151–161
37. Ministry of Health, Labour and Welfare. The report of National Health and Nutrition Survey (2008) Available at http://www.mhlw.go.jp/bunya/kenkou/eiyou/h20-houkoku.html [In Japanese]

SPINE An International Journal for the study of the spine Publish Ahead of Print

DOI: 10.1097/BRS.0000000000001960

Association of Lumbar Spondylolisthesis with Low Back Pain and Symptomatic Lumbar Spinal Stenosis in a Population-basedCohort: The Wakayama Spine Study

Yuyu Ishimoto, MD, PhD[1]; Noriko Yoshimura, MD, PhD[2];Shigeyuki Muraki, MD, PhD[3];Hiroshi Yamada, MD, PhD[1];Keiji Nagata, MD[1]; Hiroshi Hashizume, MD, PhD[1];Noboru Takiguchi, MD[1];Akihito Minamide, MD, PhD[1];Hiroyuki Oka, MD[4];Sakae Tanaka, MD, PhD[5];Hiroshi Kawaguchi, MD, PhD[6];Kozo Nakamura, MD, PhD[7];Toru Akune, MD, PhD[7];Munehito Yoshida, MD, PhD[1]

[1]Wakayama Medical University, Wakayama, Japan

[2]Department of Joint Disease Research, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo 113-8655, Japan

[3]Department of Clinical Motor System Medicine, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo 113-8655, Japan

[4]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, The University of Tokyo, Tokyo 113-8655, Japan

[5]Department of Orthopaedic Surgery, Sensory and Motor System Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

[6]JCHO Tokyo Shinjuku Medical Center, Tokyo 162-8542, Japan

[7]National Rehabilitation Center for Persons with Disabilities, Saitama 359-0042, Japan

Address correspondence to

Munehito Yoshida, MD, PhD,

Wakayama Medical University,

811-1 Kimidera,

Wakayama City, Wakayama 641-8509, Japan;

TEL: +81-73-447-2300; FAX: +81-73-448-3008; E-mail: sekitui@wakayama-med.ac.jp

Acknowledgement: May 3, 2016

1st Revise: June 28, 2016

2nd Revise: August 6, 2016

3rd Revise: August 26, 2016

Accept: September 5, 2016

The manuscript submitted does not contain information about medical device(s)/drug(s).

Grants-in-Aid for Scientific Research (B20390182, B23390357, C20591737, C20591774), Young Scientists (A18689031), and Exploratory Research (19659305) from the Japanese Ministry of Education, Culture, Sports, Science and Technology; Grants-in-Aid, H17-Men-eki-009, H18-Choujyu-037, and H20-Choujyu-009 from the Ministry of Health, Labour and Welfare; Research Aid from the Japanese Orthopaedic Association; a grant from the Japanese Orthopaedics and Traumatology Foundation, Inc. (No. 166); and a Grant-in-Aid for Scientific Research, Scientific Research (C22591639) from the Japanese Society for the Promotion of Science; and the 2012 Wakayama Medical Award for Young Researchers funds were received in support of this work.

No relevant financial activities outside the submitted work.

ABSTRACT

Study Design

Cross-sectional study

Objective

To determine the association between lumbar spondylolisthesis and low back pain and symptomatic lumbar spinal stenosis (LSS)in a population-based cohort.

Summary of Background Data

The basic epidemiology of lumbar spondylolisthesis is not well known. There is little information regarding the association between lumbar spondylolisthesis and clinical symptoms such as low back pain and LSS symptoms.

Methods

This cross-sectional study included data from 938 participants (308 men, 630 women; mean age, 67.3 years; range, 40–93 years). Lumbar spondylolisthesis was defined as a slip of ≥5%. Diagnostic criteria for symptomatic LSS required the presence of both leg symptoms and radiographic LSS findings on magnetic resonance imaging. The prevalence of low back pain and symptomatic LSS was compared between those with or without spondylolisthesis.Furthermore, we determined the association between the amount of slippage and presence of symptomatic LSS.

Results

The prevalence of spondylolisthesis at any level was 15.8% in the total sample, 13.0% in men, and 17.1% in women; the prevalence was not significantly different between men and women (P = 0.09). In both, men and women, symptomatic LSS was related to spondylolisthesis (odds ratio [OR]: 2.07; 95% CI: 1.20–3.44); however, no such association was found for spondylolisthesis and presence of low back pain.The amount of slippage was not related to the presence of symptomatic LSS (P= 0.93).

Conclusions

This population-based cohort study revealed that lumbar spondylolisthesis had a closer association with leg symptoms than with low back pain. There was a significant difference in the presence of symptomatic LSS between participants with and without spondylolisthesis. However, the amount of slippage was not related to the presence of symptomatic LSS.

Keywords:lumbar spondylolisthesis,lumbar spinal stenosis, low back pain, magnetic resonance imaging,population-based cohort

Level of evidence: 3

Introduction

Lumbar spondylolisthesis is a disorder that causes one vertebral body to slipover the one below, withtwo main etiologies: spondylolytic and degenerative[1]. Despite the considerable number of surgeries performed for spondylolisthesis[2, 3], the epidemiology of lumbar spondylolisthesis is not well known. To the best of our knowledge, there are 4 reports on the prevalence of lumbar spondylolisthesis in the general population[4-7], including a study in Asia with 1242 urban taxi drivers in Taiwan, of whom 96% were men[4], anda study in Denmark that involved subjects of both the sexes[5].The prevalence of lumbar spondylolisthesis differs greatly among these reports, ranging from 3–31%in men [4-6] and 6–29%in women [5, 7]; this wide variation may be relatedto regional variancesand different sample sizes.

It is believed thatlumbar spondylolisthesis is a frequent cause of low back pain and leg symptoms. Although low back pain and neurogenic leg symptomswere originally considered to be the principal symptoms of lumbar spondylolisthesis, there is little information regarding the association between clinical symptoms and spondylolisthesis, withdiffering results in previous studies [5,7,8]. A relationship between posterior spondylolisthesis at L3 and low back painin white elderly women has been reported[7], while other reports concluded that the correlation was not as strong as expected[5,8]. In these studies, symptomswere assessed using a questionnaire; however, a specialist's clinical impression is needed forLSS diagnosis.We previouslyreported the prevalence of low back pain and symptomatic LSSin an elderly cohort,diagnosed by an orthopedic surgeon [9,10]. For symptomatic LSS, magnetic resonance imaging (MRI)finding that were consistent with the symptoms were confirmed.

The present study aimed to determine the epidemiological data, including the prevalence and distribution, for lumbar spondylolisthesisaccording to age andsexin the general population using the baseline survey of The Wakayama Spine Study as well asto evaluate the association oflumbar spondylolisthesis with low back pain and symptomatic LSS using mobile MRI in a population-based cohort.

Materials and Methods

Participants

The present study, entitled The Wakayama Spine Study, assessed a subcohort from the Research on Osteoarthritis/Osteoporosis Against Disability (ROAD) study, which is a large-scale, prospective study of bone and joint diseases among population-based cohorts established in several communities in Japan. As the detailed profile of the ROAD study is described elsewhere, only a brief summary is provided here[11-14]. A database including baseline clinical and genetic information for 3,040 inhabitants (1,061 men, 1,979 women) with a mean age of 70.6 years (range, 23–95 years) was created. We recruited individuals listed in resident registrations in 3 communities: an urban region in Itabashi, Tokyo; a mountainous region in Hidakagawa, Wakayama; and a coastal region in Taiji, Wakayama. All participants provided written, informed consent, and the study was conducted with the approval of the ethics committees of the University of Tokyo and the Tokyo Metropolitan Institute of Gerontology. Participants completed an interviewer-administered questionnaire that consists of 400 questions including those for lifestyle, and they underwent anthropometric measurements and assessments of physical performance. Blood and urine samples were collected for biochemical and genetic examinations. The ankle-brachial index (ABI) was measured for all participants. (OMRON Co.,Kyoto, Kyoto, Japan)

The ROAD study team made a second visit to Hidakagawa and Taiji, and the inhabitants who provided written, informed consent for the MRI examination were registered in the Wakayama Spine Study. Participants who had sensitive implanted devices (e.g., pacemaker), claustrophobia, or other contraindications were excluded, and 977 participants underwent the lumbar spine MRI in a mobile MRI unit (Excelart 1.5 T; Toshiba; Tokyo, Japan). Ten participants who underwent a previous lumbar operation for LSS were excluded, and 29 participants < 40 years old were excluded because LSS is a degenerative disease. Thus, MRI results were available for 938 participants (308 men, 630 women).

Assessment of spondylolisthesis

All participants underwent A-P and lateral radiographs of the lumbar spine, including intervertebral levels from L1-L2 to L5-S1. The %slip was calculated as the distance of sagittal translation between adjacent vertebral endplates. These lumbar spine radiographs were read without the knowledge of participant clinical status by a well-experienced orthopedic surgeon (YI).

A diagnosis of spondylolisthesis was established when %slip was ≥5% in the lateral views.Inter- and intra-observer reproducibility were assessed by having both the raters (YI, SM) independently evaluate spondylolisthesis on 150 levels of L3–L5 slipping from 50 randomly chosen images. The kappa statistic was computed as the measure of agreement. Both inter- and intra-observer agreements were excellent with respect to the presence of lumbar spondylolisthesis, with kappa values of 0.83 and 0.85, respectively.

Assessment of low back pain and symptomatic lumbar spinal stenosis

An experienced orthopedic surgeon (YI) collected the medical history and performed the physical testing for all the participants[9,10]. Under medical history, information aboutthe presence of low back, buttock, and leg pain; area of pain or otherdiscomfort;and presence of intermittent claudication and its distance was collected; and themodified Zurich Claudication Questionnaire (excluding six items about satisfaction and history of lumbar surgery for symptomatic LSS) was administered. Physical examinations included determination of the symptoms induced by lumbar extension;improvement or induction of symptoms with lumbar flexion; floor finger distance (cm); peripheral circulation, determinedby palpating the dorsalis pedis artery (good or poor); administration of thestraight leg raised test; manual muscle testing of both, the upper and lower extremities; tendon reflex testing for both, upper and lower extremities; and Babinski reflex testing.

Regarding low back pain, all participants were asked the following question by the same orthopedic surgeon: “In the past month, have you had pain that last on most days?”Those who answered “yes” were identified as having low back pain[11]. The diagnostic criteria for symptomatic LSS were based on the LSS definition from the North American Spine Society (NASS) guidelines[15].

The orthopedic surgeon (YI) established the diagnosis of LSS, which required one or more of the following symptoms that were induced or exacerbated with walking or prolonged standing and relieved with lumbar flexion, sitting, and recumbency: pain, numbness and neurological deficits in the lower extremities and buttocks, and bladder/bowel dysfunction. The severity of radiographic LSS was assessed qualitatively by a well-experienced orthopedic surgeon (YI). The severity of the central, lateral recess[16], and foraminal stenoses were rated as none, mild, moderate,or severe[17]; mild stenosis was defined as a maximum of 1/3 narrowing of the normal area, moderate stenosis as a 1/3 to 2/3

narrowing, and severe stenosis as more than 2/3 narrowing.[10]A diagnosis of radiographic LSS required more than moderate severity and radiographic findings consistent with the symptoms.

The same experienced orthopedic surgeon (YI) made the final diagnosis of symptomatic LSS, which required presentation of both, LSS symptoms and radiographic LSS. There were no participants with LSS symptoms due to tumor, inflammatory, or traumatic pathologies.

Statistical analysis

All statistical analyses were performed using JMP, version 8 (SAS Institute Japan; Tokyo, Japan). Differences in age, height, weight, and BMI between gendersas well as differences in the prevalence of spondylolisthesis based on age were examined using Student's *t*-tests. Chi-squared tests were used to compare low back pain and symptomatic LSS between genders,differences in spondylolisthesis based on radiographic LSS, and differences in low back pain based on backward slip at L3–4. Furthermore, logistic regression analysis was performed to determine ~~the~~the association of spondylolisthesis with symptomatic LSS, adjusted for age, gender, and BMI. To clarify the association between the amount of slippage and symptomatic LSS, we performed Student's *t*-test using anterior %slip of L4 (n = 86) and symptomatic LSS.

Results

Table 1 shows the demographic and clinical characteristics of the 938 participants (308 men and 630 women; mean age 67.3 years, range 40–93 years). Among the participants with symptomatic LSS (n =84), 5 presented with peripheral artery disease (ankle-brachial index<0.9). However, the leg symptoms of these 5 participants were dependent on position.

The prevalence of spondylolisthesis,including anterior and posterior at any level, was 15.8% in the total sample, 13.0% in men, and 17.1% in women; the prevalence of spondylolisthesis was not significantly different between the genders(P=0.09) (Figure 1). Spondylolisthesis was observed at L3/4, L4/5, and L5/S1, with the greatest prevalence at L4/5 in both genders (men, L3/4: 3.6%; L4/5: 7.5%; L5/S: 3.2%;women, L3/4: 4.5%; L4/5: 10.3%;L5/S: 2.9%).Only one vertebral level was involved in 95.3% of the participants with spondylolisthesis.Of the participants with spondylolisthesis, 16 had backward slip, with the majority at L3/4 (L3/4,n = 15; L4/5, n = 2; L5/S, n =

0). The presence of low back pain was notsignificantly different between those with and without backward slip at L3-4 (low back pain with backward slip at L3-4: 48.7% [19/39]; low back pain without backward slip at L3-4: 39.2% [352/899]; P= 0.23).

Bothgenderswith spondylolisthesis were more likely to have low back pain than those without spondylolisthesis, but this was not significant (men,P=0.55; women,P=0.11; Table 2). Theprevalence of symptomatic LSS wassignificantlyhigher in those with than in those without spondylolisthesis in both genders. The prevalence of symptomatic LSSin men with spondylolisthesis wasapproximately 3 times of that withoutspondylolisthesis.In the logistic regression analysis adjusted for age, sex, and BMI, spondylolisthesis was the significant risk factor for symptomatic LSS (odds ratio [OR]: 2.07;95% CI: 1.20–3.44).

The mean anterior L4%slipin the total sample was 14.1±4.3%, in participants with symptomatic LSS was 14.1±1.2%, and in participants without symptomatic LSS was 14.1±0.5%. There was no significant difference in the mean anterior L4% slip between symptomatic LSS and no symptomatic LSS (P=0.93).

Discussion

In this study, the prevalence of spondylolisthesis at any level was 15.8% in the total sample, 13.0% in men, and 17.1% in women; the difference between the genders was not significant. There were only 6 participants<50 years old with spondylolisthesis, and spondylolisthesis was observed with the greatest prevalence at L4/5 (59.5%). Spondylolisthesis was significantly associated withsymptomatic LSS, but not with low back pain. Furthermore, the mean L4% slip was not related tothe presence of symptomatic LSS.

As already mentioned, the prevalence of spondylolisthesis in previous studies varies (4-7): 3–31% in men and6–29% in women. Denard et al. [6]reported that the prevalence in 300 men recruited from 5,995 participants aged ≥65 years was 31% in the Osteoporotic Fractures in Men Study; however, it is widelyconsidered that the prevalence in women is significantly higher than that in men.The lack of a significant difference in the prevalencebetween genders in the present study might be related to the anatomical differences in the different racial groups;to the best of our knowledge, no other study has reported the prevalence of lumbar spondylolisthesisin the general populationof Japan.

In the present study,lumbar spondylolisthesis was not relatedto low back pain. Similarly,no relationship between spondylolisthesis and low back pain has beenreported. In The Copenhagen Osteoarthritis Study, which included 1533 men and 2618 women, degenerative spondylolisthesis assessed using computed tomography was not significantly associated with low back pain assessed using a questionnaire[5].In the Pittsburgh clinic of the Multicenter Study of the Osteoporotic Fractures, which included 788 white elderly women, only retrolisthesis at L3-4 was associated with low back pain[7].However, lumbar spondylolisthesis induces degenerated and subluxated facet joints, and segmental instability might cause tension of the facet joint capsule and ligaments.Therefore, the development ofspondylolisthesis mightbe related to low back symptoms; however,few studies have been conducted regarding the association between the development of spondylolisthesis and low back pain. In a 25-year longitudinal study in Framingham, which included 617 subjects, the development of degenerative lumbar spondylolisthesis was significantly related to low back pain [18]. However, because of the long follow-up period, the authors were unable to determine when spondylolisthesis or low back painoccurred or worsened. Further surveys of the Wakayama Spine Study thatare planned for 3-year intervals will help to clarify the association between progressivespondylolisthesis and low back symptoms.There are multiple factors associated with the occurrence of low back pain; therefore, we also aim to identify other contributing factors such as spinal stenosis, scoliosis, facet osteoarthritis,anddisc degeneration.

In the present study, the prevalence of symptomatic LSSwas significantly higher in those with than in those withoutspondylolisthesis, however, %slipof L4 was not related with the presence of symptomatic LSS. Lumbar spondylolisthesis has traditionally been considered a major cause of leg symptoms in LSS. However, previous studies have not been conducted regarding the association between lumbar spondylolisthesis and symptomatic LSS with a diagnosisbased on the presentation of both LSS symptoms and radiographic LSS; an association between lumbar spondylolisthesis and leg symptoms has been reported, although the leg symptoms were not diagnosed by a specialist and not confirmed using MRI.Radicular pain and lower extremity weaknessassessed using questionnaires occurred more frequently in men with spondylolisthesis thanin men without spondylolisthesis[8], and no relationship between any qualities of low back pain, including gluteal or radicular pain, and degenerative spondylolisthesis has also been reported[5].While the amount of slippage is of clinical concern because it can cause spinal stenosis and nerve root compression, the present study demonstrated that %slip was not related to the presence of symptomatic LSS. However, it is possible that the amount of slippage is related with the severity of clinical symptomsas well as the natural history of clinical symptoms, which will be explored in our longitudinal study.

There were several limitations of the present study. First, the participants were not randomly selected;however, approximately 1000 participants were included, and no significant differences in BMI were found between the participants and the general Japanese population (men: 23.71 [3.41] vs. 23.95 [2.64] kg/m^2; women: 23.06 [3.42] vs. 23.50 [3.69] kg/m^2)[19]. Hence, we think that the participants can represent the general Japanese population. The proportions of current smokers and drinkers (men) and current drinkers (women) were significantly higher in the general Japanese population than in the study population (smokers: men, 32.6% vs. 25.2%; women, 4.9% vs. 4.1%; drinkers: men, 73.9% vs. 56.8%; women, 28.1% vs. 18.8%), suggesting that the study participants likely led healthier lifestyles than the general Japanese. Second, conclusive evidence of any causal relationship could not be determined because this was a cross-sectional study. Third, this study investigated elderly participants who lived independently rather than those who lived in institutional settings, potentially resulting in underestimated prevalences. Finally, the exclusion of 10 subjects who already had surgery for LSS could have influenced the results.

Nevertheless, this is the first study to evaluate the association betweenlumbar spondylolisthesis and symptomatic LSS in the general population using MRI. In addition, the Wakayama Spine Study is a longitudinal survey; therefore, future results will help elucidate any causal relationships.

In conclusion, we described the prevalence of lumbar spondylolisthesis and its association with low back pain and symptomatic LSS. Lumbar spondylolisthesis more related to leg symptoms than low back pain. Although spondylolisthesis was significantly related with the presence of symptomatic LSS, the amount of slippage was not associated with symptomatic LSS.

Acknowledgments

The sponsors had no role in study design, data collection, data analysis, data interpretation, or in the writing of the report. The authors wish to thank Mrs. Tomoko Takijiri and other members of the Public Office in Hidakagawa Town, and Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, and other members of the Public Office in Taiji Town, for their assistance in locating and scheduling participants for examinations.

References

1. Kalichman L, Hunter DJ. Diagnosis and conservative management of degenerative lumbar spondylolisthesis. *Eur Spine J*2008;17:327-35.
2. Kuntz KM, Snider RK, Weinstein JN, et al. Cost-effectiveness of fusion with and without instrumentation for patients with degenerative spondylolisthesis and spinal stenosis. *Spine(Phila Pa 1976)*2000;25:1132-9.
3. Deyo R, Gray DT, Kreuter W, et al. United States trends in lumbar fusion surgery for degenerative conditions.*Spine(Phila Pa 1976)*2005;30:1441-7.
4. Chen JC, Chan WP, Katz JN, et al. Occupational and personal factors associated with acquired lumbar spondylolisthesis of urban taxi drivers. *Occup Environ Med*2004;61:992-8.
5. Jacobsen S, Sonne-Holm S, Rovsing H, et al. Degenerative lumbar spondylolisthesis: an epidemiological perspective: the Copenhagen Osteoarthritis Study. *Spine(Phila Pa 1976)*2007;32:120-5.
6. Denard PJ, Holton KF, Miller J, et al. Lumbar spondylolisthesis among elderly men: prevalence, correlates, and progression. *Spine(Phila Pa 1976)*2010;35:1072-8.
7. Vogt MT, Rubin D, Valentin RS, et al. Lumbar olisthesis and lower back symptoms in elderly white women. The Study of Osteoporotic Fractures. *Spine(Phila Pa 1976)*1998;23:2640-7.
8. Denard PJ, Holton KF, Miller J, et al. Back pain, neurogenic symptoms, and physical function in relation to spondylolisthesis among elderly men. *Spine J*2010;10:865-73.
9. Ishimoto Y, Yoshimura N, Muraki S, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. *Osteoarthritis Cartilage*2012;20:1103-8.
10. Ishimoto Y, Yoshimura N, Muraki S, et al. Associations between radiographic lumbar spinal stenosis and clinical symptoms in the general population: the Wakayama Spine Study. *Osteoarthritis Cartilage*2013;21:783-8.
11. Muraki S, Oka H, Akune T, et al. Prevalence of radiographic lumbar spondylosis and its association with low back pain in the elderly of population-based cohorts: the ROAD study. *Ann Rheum Dis* 2009;68:1401-6.
12. Muraki S, Oka H, Akune T, et al. Prevalence of radiographic knee osteoarthritis and its association with knee pain in the elderly of Japanese population-based cohorts: the ROAD study. *Osteoarthritis Cartilage* 2009;17:1137-43.
13. Yoshimura N, Muraki S, Oka H, et al. Prevalence of knee osteoarthritis, lumbar spondylosis, and osteoporosis in Japanese men and women: the research on osteoarthritis/osteoporosis against disability study.*J Bone Miner Metab* 2009;27:620-8.
14. Yoshimura N, Muraki S, Oka H, et al. Cohort profile: Research on Osteoarthritis/osteoporosis Against Disability study. *Int J Epidemiol* 2010;39:988-95.
15. North American Spine Society Clinical Guidelines. III. Definition and Natural History of Degenerative Lumbar Spinal Stenosis 2008 (11).

16. Fardon DF, Milette PC, Combined Task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. Nomenclature and classification of lumbar disc pathology. Recommendations of the combined task forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. *Spine* 2001;26:E93-113.
17. Suri P, Rainville J, Kalichman L, et al. Does this older adult with lower extremity pain have the clinical syndrome of lumbar spinal stenosis? *JAMA* 2010;304:2628-36.
18. Kauppila LI, Eustace S, Kiel DP, et al. Degenerative displacement of lumbar vertebrae. A 25-year follow-up study in Framingham. *Spine(Phila Pa 1976)*1998;23:1868-74.
19. Ministry of Health, Labour and Welfare. The Report of the National Health and Nutrition Survey. 2005 http://www.mhlw.go.jp/bunya/kenkou/eiyou07/01.html.

Figure 1. Prevalence of lumbar spondylolisthesis classified by age and gender from a community cohort in Japan

Fig. 1

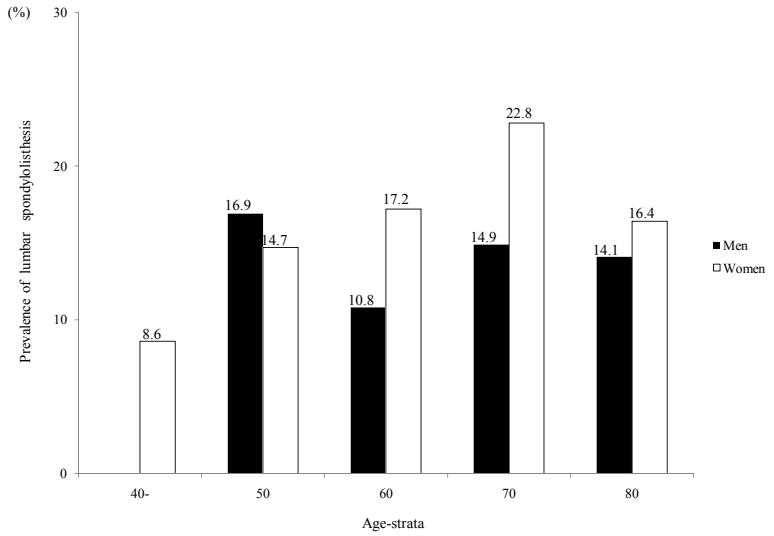


Fig. 1. Prevalence of lumbar spondylolisthesis classified by age and gender from a community cohort in Japan.

Table 1. Characteristics of participants

	Total	Men	Women
No. of participants	**938**	**308**	**630**
Age group (years)			
<49	96	26	70
50–59	175	59	116
60–69	222	65	157
70–79	258	87	171
≧80	187	71	116
Demographic characteristics			
Age, years	67.3 ± 12.4	68.3 ±12.5	66.9 ± 12.3
Height, cm	155.7 ± 9.3	164.4 ± 6.9**	151.4 ± 7.1
Weight, kg	56.7 ± 11.4	64.3 ± 11.3**	53.0 ± 9.4
Body mass index, kg/m^2	23.3 ± 3.6	23.7 ± 3.3*	23.1 ± 3.6
The state of participants			
Low back pain	371	111	260
Symptomatic LSS	84	29	55

LSS means lumbar spinal stenosis. A non-paired Student' s *t* test was used to determine differences in demographic characteristics between men and women. Chi-square test was used to determine differences in low back pain and symptomatic LSS between men and women. Values are the means ± standard deviation. * $p<0.05$, ** $p<0.01$,

Table2. Prevalence of low back pain and symptomatic LSS among spondylolisthesis or non-spondylolisthesis

	Low back pain			Symptomatic LSS		
	Total (N=371)	**Men (N=111)**	**Women (N=260)**	**Total (N=84)**	**Men (N=29)**	**Women (N=55)**
Spondylolisthesis (N=148)	69/148 (46.9%)	17/40 (42.5%)	52/108 (48.2%)	23/148** (15.5%)	8/40* (20.0%)	15/108* (13.9%)
Non-spondylolishesis (N=790)	302/790 (38.2%)	94/268 (35.1%)	208/522 (39.9%)	61/790 (7.2%)	21/268 (7.8%)	40/522 (7.7%)

Chi-square test was used to determine differences in low back pain and neurogenic claudication between spondylolisthesis and non-spondylolisthesis. *p<0.05 **p<0.01

LSS means lumbar spinal stenosis.

Ⅴ. 資料

あなたは大丈夫？L4/5の腰痛借金！

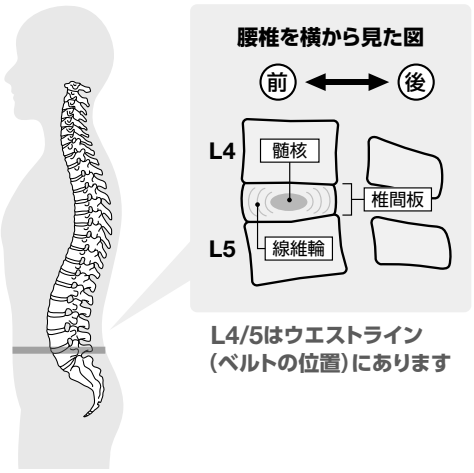
「これだけ体操」ですぐに返済！

一番負担がかかるのは、4番目と5番目の腰骨の間（L4/5椎間板）なのです！

腰痛借金って、なんですか？

●腰痛借金の無い状態

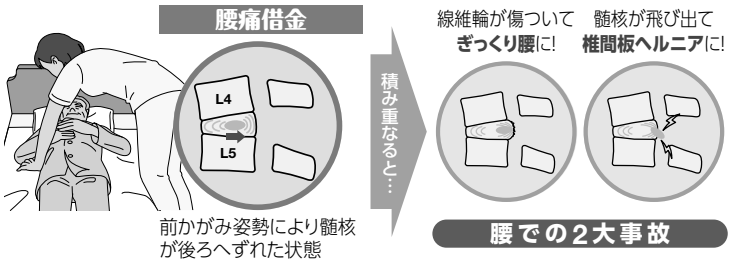
背骨と背骨にはさまれた椎間板の中には、ゼリー状の髄核（ずいかく）という物質があります。
髄核は線維輪（せんいりん）という硬い組織に囲まれており、通常、椎間板の中央に位置しています。
そして、これが腰痛借金の無い状態です。



L4/5はウエストライン（ベルトの位置）にあります

●腰痛借金と、腰痛借金が呼び込む2大事故

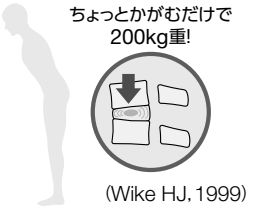
髄核は、通常は椎間板の中央にありますが、前かがみでの仕事を続けていると後ろ（背中側）に移動します。これが腰痛借金のある状態です。
この腰痛借金が積み重なると、髄核が後ろへずれっぱなしとなり、ぎっくり腰やヘルニアといった腰での2大事故が起きる可能性が高くなってしまうのです。



前かがみ姿勢により髄核が後ろへずれた状態

ちょっとした不良姿勢に忍び寄る腰痛借金の魔の手?!

椎間板には、普段の何気ない動作でも思いのほか大きな力が加わっています。
少し前へかがむだけでも、L4/5の椎間板にはなんと200kg重もの力が加わっており、腰痛借金の魔の手はちょっとした不良姿勢にも忍び寄っているのです。



(Wike HJ, 1999)

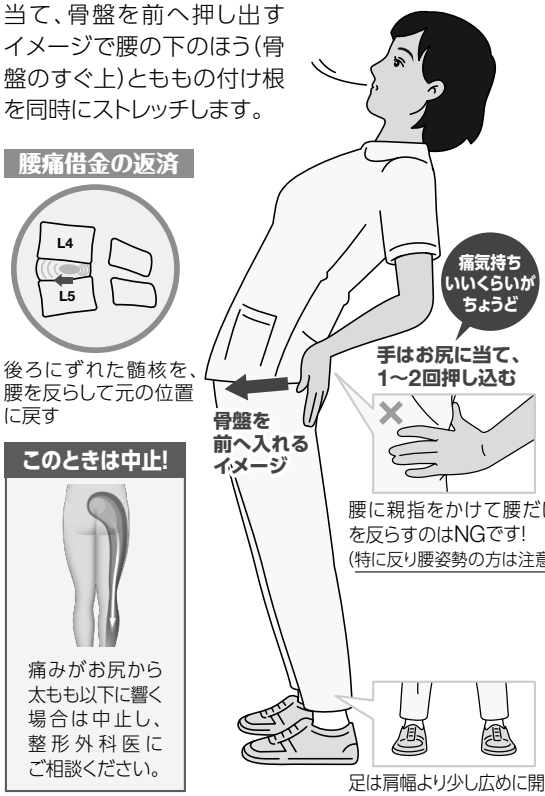
借金はその場で返済！「これだけ体操」で腰痛予防！

●どうやるの？

1〜2回

息を吐きながら、3秒間 腰を反らすだけ

手の指先を下にしてお尻に当て、骨盤を前へ押し出すイメージで腰の下のほう（骨盤のすぐ上）ともものの付け根を同時にストレッチします。



●効果はあるの？

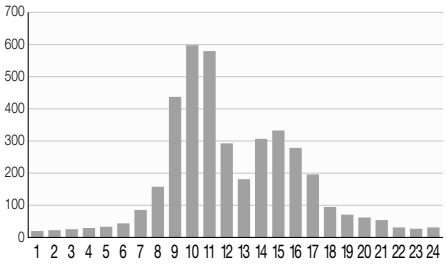
はい、「これだけ体操」を実践し続けた介護施設では、実施しなかった施設に比べ、明らかに「腰痛持ち」が少なくなったという結果が得られています。

(Matsudaira K, 2015)

●いつやればいいの？

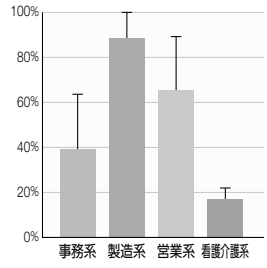
職場でのぎっくり腰は、身体反応の低下している午前中、次に昼休憩後の14〜15時に発生しやすいことがわかっています。一方、職場の始業時体操実施率をみると、他業種に比べ介護・看護系が著しく低いことが報告されています。

職場での「ぎっくり腰」の時間別発生状況（n=4,008）



(厚生労働省, 2013)

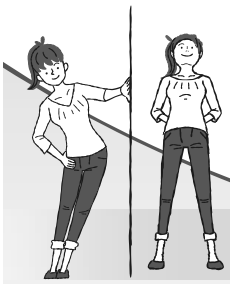
職種別始業時体操実践率（n=1,206）



(高野, 2015)

以上のことから、「これだけ体操」は、次のように行うとよいでしょう。

STEP1 朝の始業時に毎日みんなで実施（朝の貯金）
STEP2 各自、昼休憩時に実施（昼の貯金）
STEP3 作業に応じて、その都度、腰痛借金をチャラにする！



腰痛予防 これだけ体操®

腰が痛いときの安静のしずぎに注意しましょう。

「安静がいちばん」はもう古い!!!
腰痛の研究が飛躍的にすすみ、安静にしすぎないことが、世界的に治療の常識になりました。

東京大学医学部付属病院22世紀医療センター
運動器疼痛メディカルリサーチ&マネジメント講座
特任教授
http://www.lbp4u.com 松平 浩

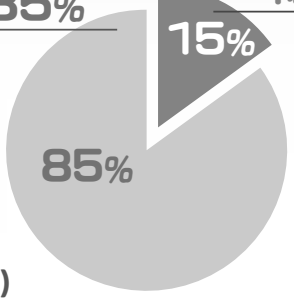
腰痛の種類

腰痛は、「病気が原因の腰痛（特異的腰痛）」と「検査や診断でも痛みの原因がはっきりしない腰痛（非特異的腰痛）」の2種類に別けられます。腰痛で受診する人の85%が原因のはっきりしない非特異的腰痛です。
これには慢性的な腰痛やぎっくり腰なども含まれます。

非特異的腰痛 85%

検査や診断でも痛みの原因がはっきりしない腰痛

腰痛の代表格
- ぎっくり腰
- 慢性的腰痛
- 再発を繰り返す腰痛



特異的腰痛 15%

病気が原因の腰痛

医療機関で検査を受けてください

腰痛の代表格
- 椎間板ヘルニア
- 脊柱管狭窄症
- 骨折

命にかかわる腰痛
- がん
- 大動脈瘤
- 感染

当てはまっていたら要注意！！

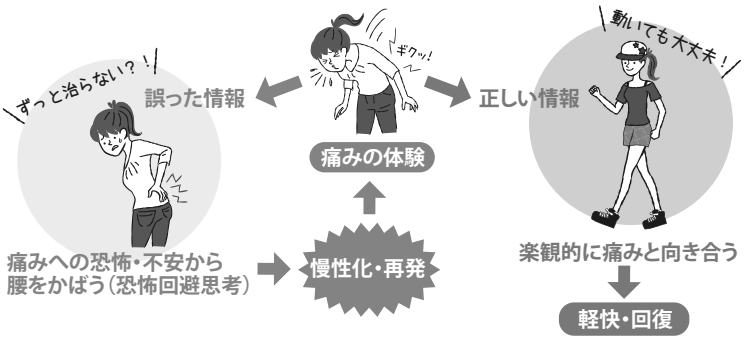
1つでも当てはまれば、
腰痛の原因が病気によるものかもしれませんので、
医療機関で検査を受けてください。

- 転んだりした後に痛みだし、日常生活に支障がでる。
  →骨折の疑い
- 65歳以上で（特に女性）で、朝、布団から起き上がるときに腰や背中に痛みが出た。
  →骨粗しょう症による骨折の疑い
- 横になってもうずく。鎮痛剤を一か月以上使用しても痛みが改善されない。
  →重い病気の疑い
- 痛みやしびれがお尻からひざ下まで広がる。
  →腰部脊柱管狭窄症・椎間板ヘルニアの疑い
- 肛門、性器周辺が熱くなる、しびれがある。尿が出にくい。尿漏れがある。
  →重症な腰部脊柱管狭窄症・椎間板ヘルニアの疑い
- つま先歩き、かかと歩きが難しく、足の脱力がある。
  →重症な腰部脊柱管狭窄症・椎間板ヘルニアの疑い
  脳や脊髄の病気の疑い

これだけ体操で緩和（裏面へ）

痛みを恐れて安静にしすぎると治りにくい

- 適度に体を動かさないと、背骨や周辺の筋肉の柔軟さが失われます。
- 筋肉が酸素不足となり発痛物質が分泌されます。

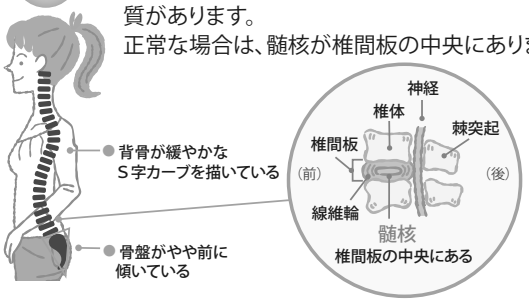


腰痛への不安や恐怖は、痛みを抑制する脳の機能を低下させます。
「怖くない」と自分に言い聞かせましょう。

腰痛の原因

正常

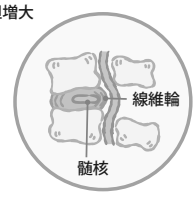
背骨と背骨のあいだの椎間板の中に、硬い組織に囲まれたゼリー状の髄核（ずいかく）という物質があります。
正常な場合は、髄核が椎間板の中央にあります。



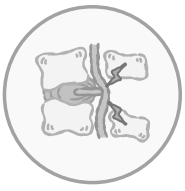
ウエストラインあたりの腰椎を横から見た図

腰痛

前かがみでの作業、猫背の姿勢が続くと髄核が後ろ（背中側）に移動します。この状態が続くと「ぎっくり腰」や椎間板ヘルニアになります。



線維輪が傷ついて
ぎっくり腰に



髄核が飛び出て
椎間板ヘルニアに

〔第1版〕H29.0 00,000

腰痛予防 これだけ体操®

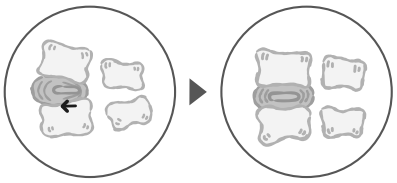
これだけ体操 腰を反らす

最重要！

こんなとき　座り作業で前かがみが続いたとき、重い荷物を持ったあとなど

髄核が後ろにずれるイメージ

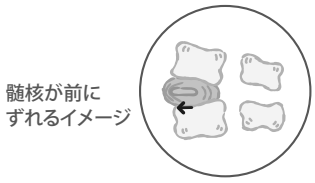
❶足を肩幅よりやや広めに平行に開き、両手を支点に腰をしっかり反らす。
❷息を吐きながら最大限に反らした状態を3秒間保つ(1〜2回)。

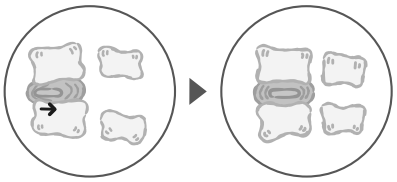

腰をしっかり反らして髄核のずれを戻すイメージ

これだけ体操 腰をかがめる

こんなとき　ヒールで立ち仕事の多い女性や妊婦の方向け

長時間立ったり歩いたりして腰に反りぎみの負荷がかかったとき


髄核が前にずれるイメージ


腰をゆっくりかがめて髄核のずれを戻すイメージ

❶椅子に腰かけ、足を肩幅より広めに開く。
❷息を吐きながら、ゆっくり背中を丸め、床を見ながら3秒間姿勢を保つ(1〜2回)。

これだけ体操 腰を横に曲げる

こんなとき　腰の左右どちらかに違和感があるとき

❶足元が滑らない場所で、安定した壁から離れて立つ。
❷肩の高さで手掌から肘までを壁につき、腰を横に曲げる。左右行う。
❸違和感を感じて曲げにくい側があれば、その方向に、ゆっくりと息を吐きながら徐々に曲げ、きついと感じるところまでしっかり曲げる。
❹左右差がなくなるまで繰り返す(5秒を3〜5回が目安)。

左右で「きつい側」を見つける


髄核の横ずれを戻すイメージ

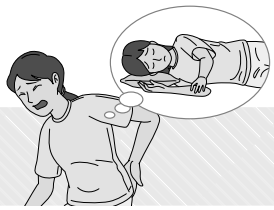
こんなときは中断しましょう

痛みがお尻から太もも以下に響く場合は中止し、整形外科医にご相談ください。

(一社)新潟県労働衛生医学協会　http://www.niwell.or.jp/　〔第1版〕H29.0 00,000

職場の腰痛対策マニュアル

知っていましたか? 腰痛の新常識



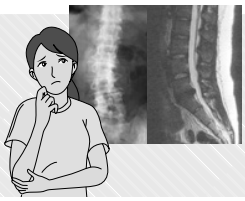
腰痛がある時は「とりあえず安静」と思っていませんか?

新常識

伝統的にはそうですが、明らかな原因疾患のない一般的な腰痛*(言いかえれば心配のない腰痛)に対しては、**今や予防としても治療としても世界的に「安静」は薦められていません!**

＊明らかな原因疾患のある腰痛(**特異的な腰痛**)としては、神経痛を伴う椎間板ヘルニア、腰骨の腫瘍や感染、骨折などが挙げられます(**下の「これは注意!病院で診てもらったほうがよい特異的な病気の潜在を疑う随伴症状」を参照**)。ただその割合は少なく(病院にかかる人のうちでもわずか1%くらい)、腰痛のほとんどが原因疾患のないもので、**非特異的な腰痛**と総称されます。

- 「ぎっくり腰」でさえも、安静を保ち過ぎるとかえってその後の経過がよくないことがわかっており、お仕事を含む普段の活動をできる範囲で維持したほうが望ましいとされています。



レントゲンやMRIの所見をみて、「変形している」「椎間板がつまっている(傷んでいる)」「ずれがある」「ヘルニアがある」「分離症がある」などと言われると、これらが腰痛の原因と思ってしまいませんか?

新常識

画像所見のほとんどは腰痛の原因を説明できません。
また、今後腰痛で困り続けるかどうかの判断材料にもならないことが多いのです。よって腰の画像所見をネガティブイメージで指摘されても、悲観する必要はありません!

- ヘルニア像も含めこのような所見は、腰痛があろうがなかろうが、少なくともどれか一つは多くの人にみられます。逆に腰痛持ちでも画像に全く異常所見がない人もいます。
- 椎間板に負担がかかっている所見は、20代からみられることも珍しくありません。

これは注意! 病院で診てもらったほうがよい特異的な病気の潜在を疑う随伴症状

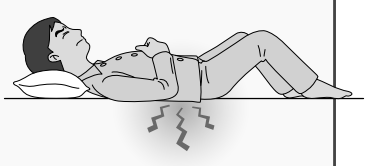
背骨の神経の病気(椎間板ヘルニアや脊柱管狭窄症など)を疑う症状

おしりを含む脚(足)に痛みやしびれ(特に膝を超えて放散するもの)を伴っていると、椎間板ヘルニアや脊柱管狭窄症を疑う必要があります。

非常に稀ではあるが、危険な痛みを疑うサイン

最も重要なことは、**脊椎の感染や腫瘍、大動脈瘤、など命にもかかわりうる危険な病気を疑うサインがある場合**には、「忙しいから」「きっと筋肉痛だから」と言ってほったらかしにしないことです。
以下の兆候があるならば、早急に病院での精密検査を念のため受ける必要があります。

- 腰を動かさなくじっとしていても痛い、夜間等横になっていても痛みが楽にならない
- 微熱でも発熱が続いている
- 体調がすぐれない(冷や汗、動悸、倦怠感など)
- 最近理由もなく体重が減ってきた
- 癌や結核を患ったことがある
- 数値がうまくコントロールされていない糖尿病や高血圧がある
- 鎮痛薬を1ヵ月近く服用しているにもかかわらず頑固な痛みが続いている

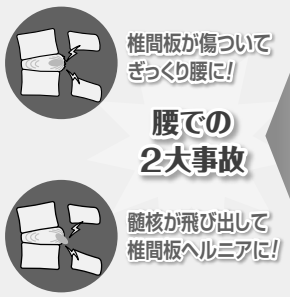


仕事に支障をきたす
腰痛が起こったり長引いたりする 危険因子 は?

1 従来から指摘されている介護作業や運搬といった持ち上げ動作や前屈み動作が頻繁な作業に関わっていることが、重要な危険因子であることは間違いありません。

腰自体の不具合?
それとも脳機能の不具合?
その見極めは?

原因(メカニズム)　　危険因子



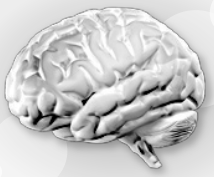
腰自体の不具合
＝
腰痛借金
前屈み姿勢により椎間板内にある髄核が後ろへずれた状態が腰痛借金のある状態です。

積み重なると…

メカニカルな腰へのストレス
持ち上げ、前屈み、
捻り、不良姿勢 など

- 姿勢・動作と腰痛の関係性が明確かつ一貫性がある
- 全く痛くない姿勢が必ずある

不活動、姿勢バランスの乱れ



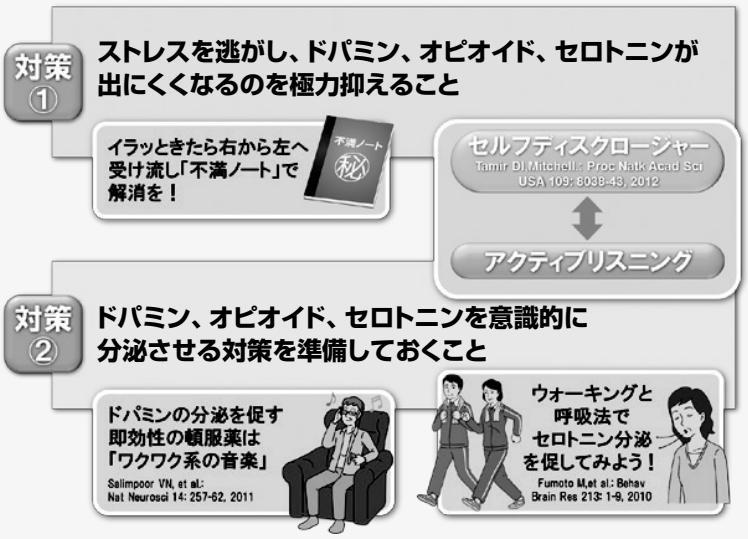
脳機能の不具合

心理的なストレス
周囲のサポート不足、
仕事への不満、人間関係の
ストレス、痛みへの強い不安

- 普通そんな痛くないだろうという刺激で、すごく痛がる
- 睡眠障害、肩こり、胃腸の不調といった身体症状が複数ある、あちこち痛い

2 職場で人間関係のストレスがあることや周囲のサポートが少ないこと、仕事のやりがいが低いこと、腰痛に対する過剰な不安や恐れ感(恐怖回避行動)といった心理・社会面の問題も無視できない重要な因子であることがあきらかになっています。

2 脳機能の不具合への対策



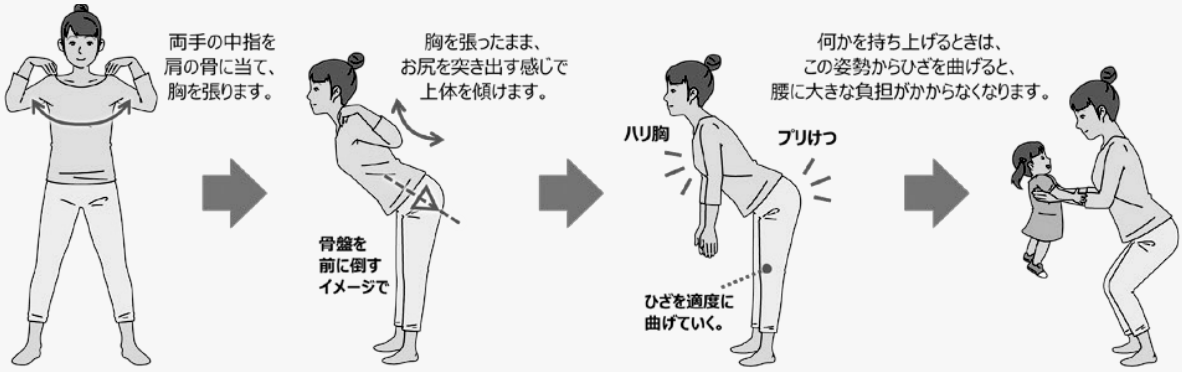
脳機能の不具合に対する対策としては、ストレスの上手な対処、脳の機能を整えることが重要となります。ストレスが強まると内因性のドパミン、オピオイド、セロトニンが分泌されにくくなりますので、①ストレスを逃がし、これらの物質が出にくくなるのを極力抑える、②これらの物質を、意識的に分泌させる対策を準備しておくこと、が肝要です。

①の具体的な方法としては、イラッとしたら「引きずると損だ」とつぶやいてなるべく早く忘れる習慣をつける、不満ノートにイラッとした原因を列挙してみる、相手の悪口を書くといったこと、②については、好きな音楽を聴いて即効性にドパミンの分泌を促す、ウォーキングや呼吸法によってセロトニンの分泌を促すといったことがあげられます。

また、①②双方に関わるものとしては、セルフディスクロージャー(他人と親しく情報や感情を共有する自己を開示する会話)やアクティブリスニング(積極的傾聴)などがあげられます。

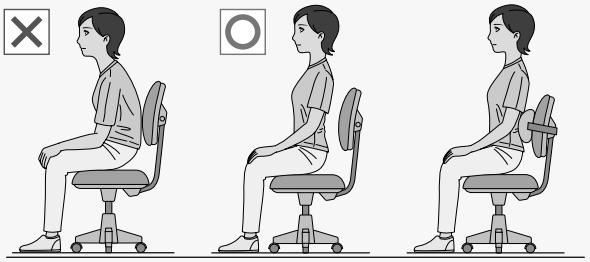
1 腰自体の不具合への対策

前屈みになる時は腰痛借金をつくりにくくする「ハリ胸＆プリけつ」



正しい姿勢

座位の時は猫背になりやすいので、良い姿勢（腰が軽い前弯位：腰をしっかり反らせた後、少し戻したくらいの状態）を保つよう意識しましょう。この姿勢の保持は腹筋、背筋の強化にもなります。
長時間座る必要がある時は、背もたれにクッションや丸めたタオル、あるいはタオルを詰めたウエストポーチなどをベルトの高さに取り付けると良いでしょう。これらは、快適だと感じる高さに微調整してください。

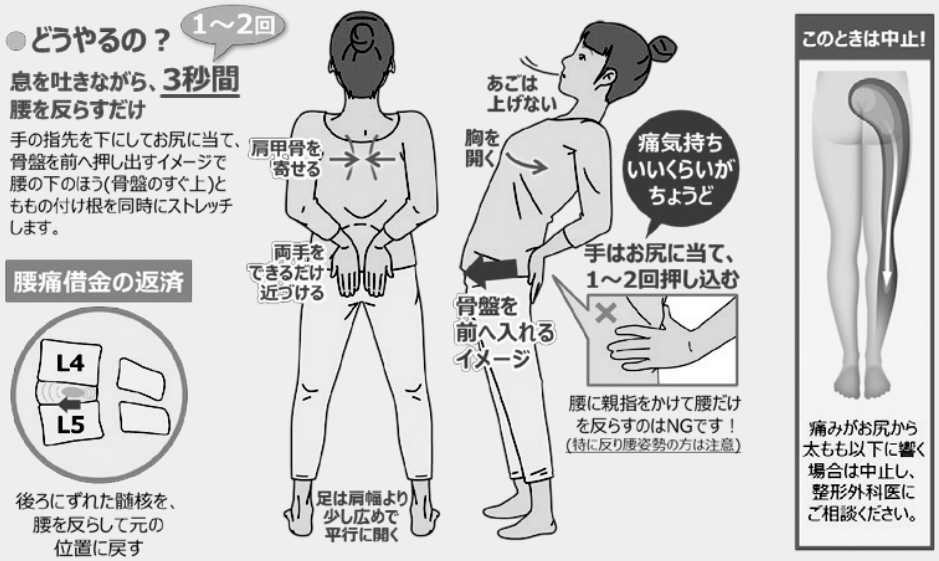


悪い姿勢（猫背）　良い姿勢　長時間座る時の工夫

＊首と肩は緊張させずリラックスした状態を保ちましょう。

腰痛借金はその場で返済！「これだけ体操」

腰を反らす「これだけ体操」



COLUMN 腰痛の重要な危険因子の一つである恐怖回避行動 ―心配し過ぎは要注意！―

「私の腰は、レントゲンで正常でなく傷んでいると言われた、気になってしょうがない」「介護や運送といったいわゆる重労働は、腰にすごく悪いとよく言われる、心配だ」「自分の仕事は重労働過ぎて、このまま続けていると私の腰はとんでもないことに　なってしまうと、ついつい悪い方向に考えてしまう」「今の腰痛が完治するまでは、とにかく無理をせず通常の仕事には戻ら　ないほうがよい」などといった、腰痛に対する強い恐怖感と、それに伴う過剰な活動の制限（専門的には恐怖回避行動と言います）が、かえって腰痛の予防や回復にとって好ましくないことがわかってきました。ここで挙げた事項は、前述しましたとおり医学的根拠はなく事実ではありません。**楽観的に腰痛と上手に付き合い前向きに過ごされることが肝要です。**

よくある質問

ぎっくり腰になったら・・・

腰痛ベルトは着けたほうがいいのですか?

着けた時のほうが痛みが和らぎ、普段の活動を維持することの助けとなるなら装着することは決して悪くはありませんが、**長期にわたり習慣的に使うメリットはほとんどない**とされています。痛みが楽になったら装着を習慣化するかわりに、前述した**「ハリ胸＆プリけつ」**と**「これだけ体操」**を習慣化しましょう!

痛み止めの薬は飲んだほうがいいのですか?

胃潰瘍の経験があるなど胃が弱い、腎臓の機能が悪い、気管支喘息があるなど鎮痛薬（市販のものでもよい）を使用しづらい場合を除いて、**我慢せず数日は定期的に服用するとよいでしょう。**

冷やす? 温める?

足首の捻挫や打撲と違って、冷やすよりも温めるほうが痛みが早くやわらぐ可能性が高いので、腰が冷えないように注意しましょう。

安静期間はどれぐらいみればいいのですか?

もちろん、動けないほどのぎっくり腰を患った場合には、数日程度なら仕事を休んでも構いませんが、**長くても3日以上安静を保つことは避けましょう。**つまり**安静にする期間は、できるだけ最小限にしましょう。**欠勤は最小限とし、数日は軽作業にしてもらうかどうか等は、上司や産業医とよく相談してください。

COLUMN 「くしゃみ」や「せき」をするときの姿勢

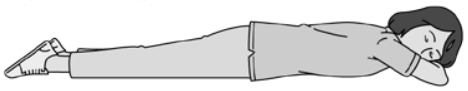
「くしゃみ」や「せき」は、瞬間的に腰に大きな負担がかかり、「ぎっくり腰」や「ヘルニア」を誘発することがあります。上体を後ろに反らせ気味の姿勢とし、可能であれば壁・机、座っていれば机や自分の膝に片手を付き衝撃を和らげましょう。

ぎっくり腰になった場合の対処法

Step 1 うつ伏せになって深呼吸

3分間深呼吸

ゆっくり深呼吸しながら、体をリラックスさせる。



Step 2 胸の下に枕を入れる

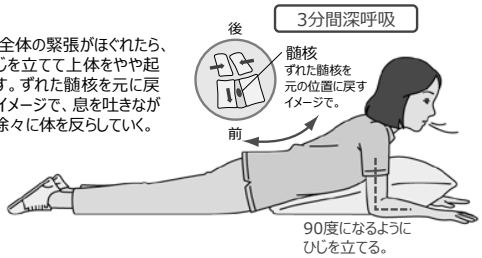
3分間深呼吸

枕が手近にあれば胸の下に入れ、さらに3分間深呼吸を続ける。足は肩幅くらいに開き、下半身（特に尻から太もも）の緊張をほぐす。

Step 3 ひじを立ててゆっくり体を反らす

3分間深呼吸

体全体の緊張がほぐれたら、ひじを立てて上体をやや起こす。ずれた髄核を元に戻すイメージで、息を吐きながら徐々に体を反らしていく。



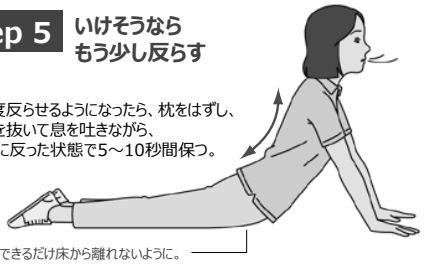
Step 4 腕を立てながらさらに腰を反らしていく

反らして5〜10秒間保つ×10回

腕立て伏せのように腕を立て、息を吐きながら、イタ気持ちいいところまで、さらに反らしていく。おへそが枕から離れないように注意しながら、5〜10秒間保つ。
体をStep2の状態に戻し、ここまでの動きを10回繰り返す。

Step 5 いけそうならもう少し反らす

ある程度反らせるようになったら、枕をはずし、腰の力を抜いて息を吐きながら、最大限に反った状態で5〜10秒間保つ。



この体操は多くの場合、腰を反らすときに痛みが一時的に強まり、うつ伏せに戻れば軽減するというパターンをとります。これは「続ければよくなる可能性が高い」というサインですから、安心して続けましょう。