労災疾病臨床研究事業費補助金

精神疾患により長期療養する労働者が元の職場等、社会復帰するまでの過程及び
その手法に係る医学的知見や文献等に係る調査研究（160102－01）

精神疾患患者の社会復帰指標作成・効果的介入同定の系統的レビュー (16808287)

平成 29 年度 総括・分担研究報告書

研究代表者　　渡辺 範雄

平成 30 年 (2018 年) 3 月

目次

I. 総括研究報告

精神疾患患者の社会復帰指標作成・効果的介入同定の系統的レビュー

京都大学大学院医学研究科　健康増進・行動学分野

渡辺　範雄　-------- 5

II. 分担研究報告

1. 労働者における精神疾患による休業からの復職に関連する
ライフスタイルを含む予測因子についての体系的レビュー

日本大学医学部社会医学系公衆衛生学分野

井谷　修　-------- 19

2. 自閉症スペクトラム障害者の復職、就職のための心理社会的介入

京都大学大学院医学研究科　健康増進・行動学分野

小川 雄右　-------- 35

3. 精神疾患による長期療養者の職場復帰に関する、運動・
睡眠等、ライフスタイルに関する介入

日本大学医学部社会医学系公衆衛生学分野

地家　真紀　-------- 55

4. 精神療法・疾患重症度・レジリエンス等複合化指標による
疾患予後・復職の予測可能性の検討：系統的レビューのオーバービュー

京都大学大学院医学研究科　健康増進・行動学分野

渡辺　範雄　-------- 65

III. 研究成果の刊行に関する一覧表

-------- 93

IV. 研究成果の刊行物・別刷

-------- 99

I. 総括研究報告

労災疾病臨床研究事業費補助金
総括研究報告書

精神疾患患者の社会復帰指標作成･効果的介入同定の系統的レビュー

研究代表者　渡辺　範雄
京都大学大学院医学研究科　健康増進･行動学分野　准教授

研究要旨

精神疾患による就労の長期休業からの復職に関して、これを予測する指標、また復帰を推進する効果的介入の開発・同定が社会的急務である。

本研究では、まず国内外の一次研究のエビデンスを集積して3つの系統的レビューを行い、また既存の系統的レビューを収集してオーバービューを行った。系統的レビュー1では観察研究から復職に関する精神疾患重症度や日常生活因子を探索し、系統的レビュー2では無作為割付対照試験を収集し、既存の系統的レビューになかった自閉症スペクトラム障害に焦点を当て、復職や就職に関する心理社会的介入のメタアナリシスを行った。系統的レビュー3では、運動やリワーク等ライフスタイルに対する介入研究を集積した。オーバービューでは既存の系統的レビューを収集･質評価を行い、異質性等がなければ量的統合･予測モデル作成の方針とした。

結果、系統的レビュー1では復職と有意に関連していた生活習慣として，喫煙・睡眠・飲酒・運動が同定された。しかし、オーバービューでは生活習慣に言及しているものはなかった。本領域では先行研究の多様性や異質性が大きいため、一貫した確固たる結論づけることの困難さを示していると考えられた。系統的レビュー2では自閉症スペクトラム障害に対してコミュニケーションスキルトレーニングによる就労達成者割合や就労インタビュースコア上昇が示された。またSR3では、主に日本国内の復職に関する介入は対照群がない前後比較研究で、運動が介入として選択されているものが多かったが特定の介入が有効であるという結論には至らなかった。オーバービューでは生活指導の介入に関するエビデンスはなく、認知行動療法等の精神療法が有効であると結論した系統的レビューはなかった。

前述のように先行研究には多様性や異質性が大きく、量的統合はもちろん一貫した方向付けを持つ結論が困難であることが示され、今後の研究では研究スコープ・アウトカムの統一化に留意し、大規模データが必要であることが示唆された。

A. 研究目的

近年、精神疾患による労災認定請求・支給決定件数は過去最多を記録した。精神疾患で長期療養に至った労働者では、その後の症状の完全消失・社会復帰が望まれるが、薬物療法だけではこれを達成できない。効果的追加介入と、職場・社会復帰の判断指標が必要となる。しかし研究者らの知る限り、良質エビデンスの社会復帰指標は存在しない。

そこで、本研究では、国内外の質の高いエビデンスを集積して3つの系統的レビュー(SR)を行い、疾患重症度、心理社会的因子、ライフスタイル、精神療法介入等が精神疾患による長期療養者の職場・社会復帰にどのような影響を及ぼすのか、明らかにする。さらに、他の系統的レビューを網羅的収集、質の評価のうえオーバービューを行い、異質性の問題なく統合可能と判断されれば統計学的統合を行い、精神疾患による長期療養者の社会復帰のための予後予測モデルを作成する。具体的には、下記のように各分担研究者の報告として取り扱った。

1. SR1（分担研究者：井谷)では観察研究を集積・統合し、精神疾患の疾患重症度やレジリエンス等の精神状態、睡眠の量・質や活動量等、生活指標の縦断的データを利用した、長期療養者の職場復帰予測指標指標作成を目的とする。

2. SR2(分担研究者：小川)では無作為割り付け対照試験(RCT)を集積して、症状改善や早期復職のための効果的な心理社会的介入を同定する。前年度までに統合失調症やうつ病・適応障害等の精神疾患に関しては、質のよい系統的レビューが存在することが分かったため、本研究ではこれが存在しない自閉症スペクトラム障害患者の復職、就労に焦点を当てた。

3. SR3(分担研究者：地家)では、RCTに限定せずわが国でも広く行われている復職リハビリテーション等に関する観察研究をも集積し、運動・睡眠等、早期復職のための効果的なライフスタイルに関する介入の同定を行う。

4. オーバービュー（分担研究者：渡辺)では、本研究の目的に合致した既存のSRを網羅的に収集・概観することで、既存のエビデンスを包括的に評価し、報告する。可能であれば量的統合・予測モデルを作成して、精神疾患による長期療養者の社会復帰推進のための提言を行う。

これらの知見により、職場・社会復帰促進に加えてその後の再休職予防や、医療資源の効率的利用に資することができると考えている。

B. 研究方法

分担研究ごとに研究方法を述べる。

1. SR1では、前年度に先行研究の網羅的収集を

行い，現時点での知見と限界点をまとめた。今年度はまず先行研究よりもより幅広く高感度に観察研究の文献を収集するための、高感度な文献検索式を作成した。データベースについてはMedline（PubMed）に加えてPsychINFO・EMBASE, CINAHLにおいて横断的・網羅的な検索を行った。4つのデータベースの横断検索結果の統合・重複除去により4,568件の研究が対象となった。その後2人×2チームの独立した研究者により，検索された文献の登録基準該当性確認・データ抽出を行った。

2. SR2では自閉症スペクトラム障害（ASD）者の就労、職場復帰、生産性の回復に関する心理社会的介入の有効性を評価することを目的に、全ての現存の無作為割り付け対照試験（RCT）を収集し、系統的レビュー・メタアナリシスを実施した。文献検索はMEDLINE、CENTRAL、PsycINFO、CINAHL、Scopus、ProQuest Dissertations & Theses、ICTRPを用いて包括的に行ない、独立した２人の研究者によって研究選択、データ抽出、バイアスのリスクの評価を行った。統合が可能と判断したアウトカムに対して、ランダム効果モデルを用いたメタアナリシスを行った。

3. SR3では本課題に対応したPubMedおよび医中誌の検索式を作成し、言語や出版年数を問わず検索を実行した。具体的な用語として労働者、精神疾患に関する用語、休職・復職に関する用語、ライフスタイルに関する用語とした。
検索結果を元に独立した2人の研究者が1次チェックとしてタイトルおよび抄録を、さらに2次チェックとして本文を読み本研究課題に該当するか評価した。1次チェック、2次チェックともに2人の研究者の評価が分かれた文献については議論し該当、非該当の2群に分類した。

4. オーバービューでは、MedlineおよびPsycINFOで文献検索を行い、前向きまたは後ろ向きコホート研究、またはRCTを再現可能な方法によって集積した、系統的レビューを収集した。各系統的レビューにはメタ解析の有無は問わず、質的報告のみ行っているものも許容し、また出版言語は問わなかった。系統的レビューは、精神疾患または精神症状尺度にて一定閾値以上の患者を研究対象とし、かつ患者は休職または失職、または就職していない者に限定した。また、精神疾患の疾患重症度や精神状態、レジリエンス・自己効力感、職場環境、睡眠、運動、食事、または心理社会的介入を扱い、その後の観察で復職・再就職アウトカム情報を収集しているものとした。
データ抽出は事前にデザインされた抽出フォームを使って行ない、また各SRの質の評価はAMSTAR2を利用して評価した。AMSTAR2は2017年に公表された、系統的レビューの質評価を行うツールである。またデータが、系統的レビュー間で異質性が低いと判断され、なおかつ量的データが示されている場合は、各レビューから個々のRCTのデータを抽出して重複を除いた上でメタアナリシスを実施し、量的統合を行うこととした。ただし異質性が高く、量的統合が難しいと判断され

る場合は、各レビュー結果を叙述的・質的にそれを提示することとした。

（倫理面への配慮）

「人を対象とする医学系研究に関する倫理指針」に基づいて行うが、本研究に含まれる全ての系統的レビュー・予測モデル作成は、一次研究を集積・統合した二次研究であり、不利益・危険性の排除や説明同意、プライバシー保護等は一次研究で担保されているものを扱うため、問題は生じないと考えられる。

C. 研究結果

分担研究ごとに研究結果を述べる。

1. 適格性１次チェックの段階で第1グループの2,464件のうち適格性有りと判定されたものは69件であり，第2グループの2,463件のうち適格性有りと判定されたものは62件であり2グループ併せて合計131件について，2次チェックを行った。2次チェックの結果，最終的に14件の研究が本レビューにおける適格研究と判定された。

   最終結果の14件の研究において，休職からの復職のアウトカムに有意に関連していた生活習慣として，喫煙・睡眠・飲酒・運動が同定された。本研究では，抽出対象とする研究のライフスタイルとしてかなり広範囲を対象として検索･抽出を進めたにもかかわらず,最終的に研究の対象となっていたライフスタイルは極めて限定されていた。

2. 文献検索、研究選択の結果、7研究（対象者247名）が同定された。就労者の割合に関しては就労支援の評価をした1研究、コミュニケーションスキルトレーニングの評価をした2研究で報告されていた。販売のためのスキルや、座学、地域の協力者との協力、適応的行動分析技法､社会的コミュニケーションスキルを含む9ヶ月のプログラムであるProject SEARCH及びASD支援 を行った群では1年後に31人中28人が就労を達成したのに対し、行わなかった群では23人中1人のみしか就労していなかった｡コミュニケーションスキルトレーニングの評価をした2研究のメタアナリシスを行った結果､コミュニケーションスキルトレーニングを行った場合は行わなかった場合に比べて就労達成者の割合のリスク比は2.27であったが､有意差はなかった（95%信頼区間0.73-7.07）。就労に関するインタビュースコア については、3研究で統計学的統合を行った。コミュニケーションスキルトレーニングを行った場合は行わなかった場合に比べ､有意にスコアが向上した（標準化平均差 0.56, 95%信頼区間0.05-1.06）。

3. 検索式を用いて抽出された論文数はPubMed 311本、医中誌378本であった。さらに1次チェックを行い該当と評価された論文数はPubMed 14本、医中誌53本であった。そして2次チェックを行い該当と評価された論文数はPubMed3本、医中誌17本であった。

   ライフスタイルに関するプログラムの中でも運動についての介入が行われているものが主であった。スポーツやエクササイズという表記がとられ運動の詳細が不明瞭な文献

のほか有酸素運動療法またはヨガ・卓球、太極拳、ラジオ体操が実施されていた。

4. 精神疾患による休職者の復職に関する要因を探索したものとして5本の系統的レビュー、精神疾患による休職者、または精神疾患によって就職できない者に対する心理社会的介入を扱ったものとして、15本の系統的レビュー(全てRCTを集積したもの)が同定された。AMSTAR-2による質は様々であったが、質の高いものを中心にレビューを行った。
精神疾患による休職者の復職可能性の指標として、各系統的レビューに共通して見られた知見では、高年齢、最近の疾患による休職歴、身体疾患合併等が低い復職可能性と関連し、逆に高い自己効力感や仕事における高い能力または仕事における低い要求水準は復職可能性が上ることを示した。一方、介入に関する系統的レビューのオーバービューでは、IPSモデル等の復職支援プログラムは、統合失調症をはじめとした重症精神疾患者の就職可能性を向上させるが、一般精神疾患では認知行動療法、また運動・食事・飲酒や喫煙・睡眠等生活指導の介入に関する研究は認めなかった。
各系統的レビューで採用した一次研究の研究デザインや方法論のみならず、対象とした疾患、介入・暴露方法、アウトカムの定義にも大きな異質性を認めたため、総括的メタ解析や予測モデル作成を行うことはできなかった。

D. 考察

本研究では、精神疾患による休職者の復職に影響を与える要因・介入の探索を目的に、分担者ごとに複数の包括的な系統的レビュー・オーバービューを行った。
まず復職に影響を与える要因であるが、観察研究の系統的レビュー(SR1) では有意に関連していた生活習慣として，喫煙・睡眠・飲酒・運動が同定された。しかし、オーバービューでは低い復職可能性には高年齢、最近の休職歴、身体疾患合併が、高い復職可能性には高い自己効力感、高い仕事における能力、低い仕事における要求水準が関連することが見出されたものの、SR1で見られたような生活習慣に言及しているものはなかった。収集された研究の中身を見ると、SR1で収集された観察研究とオーバービューに含まれた系統的レビュー内の観察研究に重複はなかったため、研究目的や方法論の差異から生じたものだと考えられる。むしろ、この領域では先行研究の多様性や異質性が大きいため、一貫した確固たる結論づけることの困難さを示している。
また精神疾患による休職者の復職に影響を与える介入に関しては、SR2では自閉症スペクトラム障害に対するコミュニケーションスキルトレーニングによる就労達成者割合や就労インタビュースコア上昇が示された。またSR3では、主に日本国内の復職に関する介入をまとめたが、前後比較研究で対照群がなく、また運動が介入として選択されているものが多かったものの詳細が不明瞭で介入内容・効果の両面から、ある介入が復職に関して有効であるという結論には至らなかった。オーバービューでは、一般精神疾患に関して運動・食事・飲酒や喫煙・睡眠等生活指導の介入に関するものはなく、認知行動療法等の

精神療法が復職アウトカムに関して有効であると結論した系統的レビューはなかった。

以上、前述のように先行研究には多様性や異質性が大きく、今回のSR2のような狭いスコープの研究疑問に絞らないかぎり、量的統合はもちろん一貫した方向付けを持つ結論が困難であることが示された。

これらの知見から、今後の当該領域での研究では、わが国内の職種や労働条件ごとに分けて、異質性を減じた状況での情報収集を行うべきであると言えるだろう。また精神疾患による休職やその後の復職の定義、患者特性や日常生活情報、介入に関しては報告するべき項目や内容・アウトカム尺度において研究間で比較・統合可能な統一フォーマットを整備するべきであると考えられる。

そのためには、個人情報保護などの問題を解決しながら、診療報酬のレセプトデータのように、企業による休職者の登録を制度化するなど、ビッグデータ化していくことも今後計画する必要もあるだろう。

E. 結論

本研究では、先行研究の包括的レビューを行い、精神疾患による休職者の復職に影響を与える要因・介入の探索を行った。先行研究には多様性・異質性とも大きく、量的統合を行って画一的予測モデルを作成するには至らなかった。しかしながら今後の研究として、一定条件ごとに分けた均質なサブグループごとの情報収集が必要であること、精神疾患による休職やその後の復職の定義、収集すべき情報の統一フォーマットを整備するべきであること、個人情報の問題に留意しながら登録等の制度化などビッグデータ化していく可能性があることなどが示唆された。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

1. Itani O, Kaneita Y, Jike M, Furuya M, Uezono C, Oda F, Agematsu R, Tokiya M, Otsuka Y, Ohida T: Sleep-related factors associated with industrial accidents among factory workers and sleep hygiene education intervention. Sleep Biol Rhythms 16:239-251, 2018.
2. Jike M, Itani O, Watanabe N, Buysse DJ, Kaneita Y: Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression. Sleep Med Rev 39:25-36, 2018.
3. Watanabe N, Matsuoka Y, Kumachi M, Hamazaki K, Horikoshi M, Furukawa TA. Omega-3 fatty acids for a better mental state in working populations - Happy Nurse Project: A 52-week randomized controlled trial. *J Psychiatr Res.* 2018;102:72-80.
4. Ogawa Y, Furukawa TA, Takeshima N, Hayasaka Y, Atkinson LZ, Tanaka S, Cipriani A, Salanti G. Efficacy of

antidepressants over placebo is similar in two-armed versus three-armed or more-armed randomized placebo-controlled trials. Int Clin Psychopharmacol. 2018 Mar;33(2):66-72.

5. Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, Ogawa S, Sato H, Kitagawa N, Shinagawa Y, Ikeda Y, Imai H, Tajika A, Ogawa Y, Akechi T, Yamada M, Shimodera S, Watanabe N, Inagaki M, Hasegawa A. Cognitive and Behavioral Skills Exercises Completed by Patients with Major Depression During Smartphone Cognitive Behavioral Therapy: Secondary Analysis of a Randomized Controlled Trial. JMIR Ment Health. 2018 Jan 11;5(1)
6. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JPT, Egger M, Takeshima N, Hayasaka Y, Imai H, Shinohara K, Tajika A, Ioannidis JPA, Geddes JR. Comparative efficacy and acceptability of first- and second-generation antidepressants in the acute treatment of major depressive disorder: a network meta-analysis. Lancet. 2018 Apr 7;391(10128):1357-1366
7. Soares-Weiser K, Rathbone J, Ogawa Y, Shinohara K, Bergman H. Miscellaneous treatments for antipsychotic-induced tardive dyskinesia. Cochrane Database Syst Rev. 2018. CD000208.
8. Itani O, Jike M, Watanabe N, Kaneita Y: Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. Sleep Med 32:246-256, 2017.
9. Kinjo A, Imamoto A, Ikeda M, Itani O, Ohida T, Kaneita Y, Kanda H, Tanihata T, Higuchi S, Osaki Y: The Association Between Alcohol-Flavoured Non-Alcoholic Beverages and Alcohol Use in Japanese Adolescents. Alcohol Alcohol 52:351-357, 2017.
10. Otsuka Y, Kaneita Y, Itani O, Nakagome S, Jike M, Ohida T: Relationship between stress coping and sleep disorders among the general Japanese population: a nationwide representative survey. Sleep Med 37:38-45, 2017.
11. Tokiya M, Kaneita Y, Itani O, Jike M, Ohida T: Predictors of insomnia onset in adolescents in Japan. Sleep Med 38:37-43, 2017.
12. 兼板佳孝, 大井田隆, 今村聡, 羽鳥裕, 地家真紀, 井谷修, 土器屋美貴子, 大塚雄一郎, 中込祥: 2016 年　日本医師会員の喫煙とその関連要因に関する調査. 日本医師会雑誌 146:1409-1422, 2017.
13. Itani O, Kaneita Y, Tokiya M, Jike M, Murata A, Nakagome S, Otsuka Y, Ohida T: Short sleep duration, shift work, and actual days taken off work are predictive life-style risk factors for new-onset metabolic syndrome: a 7-year cohort study of 40,000 male workers. Sleep Med 39:87-94, 2017.
14. Otsuka Y, Kaneita Y, Nakagome S, Jike M, Itani O, Ohida T: Nightmares and sleep paralysis among the general Japanese

population: a nationwide representative survey. Sleep Biol Rhythms 16:187-195, 2017.

15. Psychological and Educational Interventions for Atopic Dermatitis in Adults: A Systematic Review and Meta-Analysis. Hashimoto K, Ogawa Y, Takeshima N, Furukawa TA. Behavior Change. 2017 April; 34(1):48-65.
16. Tateno M, Jovanovic N, Beezhold J, Uehara-Aoyama K, Umene-Nakano W, Nakamae T, Uchida N, Hashimoto N, Kikuchi S, Wake Y, Fujisawa D, Ikari K, Otsuka K, Takahashi K, Okugawa G, Watanabe N, Shirasaka T, Kato TA. Suicidal ideation and burnout among psychiatric trainees in Japan. *Early Interv Psychiatry.* 2017.
17. Kusaka K, Shinohara K, Tada M, Watanabe N, Furukawa TA. Concerns About Selective Outcome Reporting. J Clin Oncol. 2017;35(6):688.

2. 学会発表

（発表誌名巻号・頁・発行年等も記入）

1. Watanabe N, Jike M, Itani O, Buysse D, Kaneita Y: Long Sleep Duration and Health Outcomes: A Systematic Review, Meta-Analysis and Meta-Regression. the 2017 Annual Meeting of the American Psychiatric Association, San Diego, USA, 2017.
2. 井谷修, 兼板佳孝, 土器屋美貴子, 地家真紀, 中込祥, 大塚雄一郎, 日本の中学生・高校生における睡眠随伴症(寝ぼけ・悪夢・金縛り)発症についての縦断的疫学研究. 日本睡眠学会第42回定期学術集会, 横浜, 2017.
3. 大塚雄一郎, 兼板佳孝, 中込祥, 地家真紀, 土器屋美貴子, 井谷修, 高校生に対する睡眠衛生教育の効果に対する検討. 日本睡眠学会第42回定期学術集会, 横浜, 2017.
4. 土器屋美貴子, 兼板佳孝, 井谷修, 地家真紀, 中込祥, 大塚雄一郎, 日本における思春期の不眠発症の予測因子. 日本睡眠学会第42回定期学術集会, 横浜, 2017.
5. Watanabe N, Jike M, Itani O, Buysse D, Kaneita Y: Long Sleep Duration and Health Outcomes: A Systematic Review, Meta-Analysis and Meta-Regression. World Sleep 2017, Prague, Czech Republic, 2017
6. 井谷修, 兼板佳孝, 原野悟, 土器屋美貴子: 労働者の休養の評価尺度質問票の作成. 平成29年度産業保健調査研究発表会, 東京, 2017.
7. 渡辺範雄, 地家真紀, 井谷修, 兼板佳孝：長時間睡眠と健康アウトカム　系統的レビューとメタ回帰分析. 第76回日本公衆衛生学会総会, 鹿児島, 2017.
8. 土器屋美貴子, 兼板佳孝, 井谷修：高校生におけるインターネット依存と睡眠障害との関連性について. 第76回日本公衆衛生学会総会, 鹿児島, 2017.
9. 井谷修: 思春期におけるメディア・ネット依存問題について（シンポジウム: 思春期公衆衛生領域におけるトピックス～研究の最前線から未来への発信～）. 第76回日本公衆衛生学会総会, 鹿児島, 2017.

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

なし

II. 分担研究報告

労災疾病臨床研究事業費補助金
分担研究報告書

労働者における精神疾患による休業からの復職に関連するライフスタイルを含む予測因子についての体系的レビュー

研究分担者　井谷　修
日本大学医学部　社会医学系公衆衛生学分野　准教授

研究要旨

労働者の病気による長期休業の原因として精神疾患が占める割合が増加しており，精神疾患に罹患し休業した労働者に対してReturn To Work（RTW）を促進させる施策を検討する必要がある。更に，施策の検討にあたって，精神疾患の症状改善・治癒といった疾患そのものの回復過程でどのような要因がRTWに関連しているのかを明らかにすることが急務である。そこで本研究では精神疾患による休業からのRTWに関連する要因についての系統的レビューを実施した。

前年度の研究では，先行研究について収集し，現時点での知見と限界点をまとめた。今年度は引き続き研究プロトコルを確立し，系統的レビューの本実施を行った。複数の医学データベースを横断的・網羅的に検索した上で，2段階にわたり複数の独立した研究者が研究を選別した結果，14編の観察研究を抽出した。このうち，精神疾病による休業者の復職に関連する生活習慣要因として有意な関連を認めたものとして，喫煙・睡眠・飲酒・運動が同定された。

しかし本研究では，抽出対象とする研究のライフスタイルとしてかなり広範囲を対象として検索・抽出を進めたにもかかわらず，最終的に研究の対象となっていたライフスタイルは極めて限定されていた。今後の本領域の研究では、より広範囲なライフスタイルとの関連性も検討されるべきである。

A. 研究目的

先進国において労働者の精神疾患罹患による経済的および社会的な損失が大きな社会問題となっている。精神疾患に対応するための費用は，国家総予算の約3～4%に達するという試算もある。これらの費用のうち約3分の1が精神疾患への直接の医療費であるが，残りの3分の2は労働者における生産性の低下や雇用の喪失によるものである。特に，ここ数十年の間で，労働者の病気による長期休業の原因として精神疾患が占める割合は増加してきている。このような状況の改善のため，精神

疾患に罹患し休業した労働者に対してreturn to work（RTW）を促進させる施策を検討する必要がある。施策の検討にあたって，精神疾患の症状改善・治癒といった疾患そのものの回復過程でどのような要因がRTWに関連しているのかを明らかにすることが急務である。

疾病への対応を検討する際に，その要因を科学的に分析することは，対応を効率的に進める上で非常に重要である。一般的に認識されていることとして，病気による休業の要因は多因子的であり，単に原因となっている疾病のみではない。労働者の疾病による休業の先行研究として，現在まで身体疾患や怪我に関するものについては，休業の要因についての研究が数多く行われ，それらのエビデンスを集約した体系的レビューも行われてきている。Brankらは精神疾患とRTWまたは長期休業に関連する要因についての体系的レビューの結果，14編の研究を抽出し， RTWに関連した要因として家族歴やhealth risk behaviorsや社会状況や精神疾患の病状といったものを指摘している[1]。しかし，6ヶ月以上休業を扱った研究は対象から除外されており，対象とする研究デザインとして横断研究を含むすべての研究タイプを対象としているため，同定した関連要因とRTWとの関連について因果関係を述べることが出来ていない。また，Corneliusらは縦断研究のみを対象とした系統的レビューを行っているが，この際の抽出論文は7編のみであり，要因の同定に充分な研究数を確保できていない[2]。その他の研究を含めても，精神疾患による休業の予後を左右する要因ついての先行研究を体系的にレビューしたものは極めて少ない。そこで本研究では精神疾患による休業からのRTWに関連する要因についての系統的レビューの実施を計画した。

前年度の研究では，先行研究について収集し，現時点での知見と限界点をまとめた。今年度は引き続き研究プロトコルを確立し，系統的レビューの本実施を行った。

B. 研究方法

前年度には，本研究に関連する先行研究の網羅的収集を行い，現時点での知見と限界点をまとめた。引き続き，先行研究よりもより幅広く高感度に文献を収集するための、高感度な文献検索式を作成した。データベースについてはMedline(PubMed)に加えてPsycINFO・EMBASE，CINAHLにおいて横断的・網羅的な検索を行った。

その後2人X2チームの独立した研究者により，検索された文献の登録基準該当性確認・データ抽出を行った。

1. 本研究の計画策定

抽出対象とする研究としては，以下の設定とした。

- Type of studies
  All Observational studies, i.e., case-control studies, cohort studies, follow-up studies or longitudinal studies.
- Type of participants
  Workers in the age 16–65 years who are on sick leave due to mental disorders.
  Duration of sick leave or disability is not an in- or exclusion criterion in order to include all durations.
- Types of outcome measures
  Dependent variables: return to work; symptom recovery, improvement of functioning;

reduction of disability; expanding of activities; heightening of social participation.

- Independent variables:
  lifestyles; health service use; adequacy of treatment; coping strategies and social support.

探索する関連要因の種類としては,

- 睡眠時間
- 不眠症者
- 運動習慣非保持者（運動強度など）
- 食事
- メディア（テレビ・スマホ等）の使用
- 休養，レジャー，余暇，趣味
- 業務内容（自己決定権・仕事内容・質・周囲からのサポート・地位職位）
- 家族状況（結婚・子どもなど）
- 経済状況
- 喫煙，飲酒
- 精神状態(治療介入の有無含む）やレジリエンス

とした。

主要アウトカムとしては,

- 復職までの期間（病気休暇日数）

とした。

副次アウトカムとしては

- 最終的な復職の可否
- 復職後の再発の有無
- 罹患疾病の改善

を設定した。

2. 文献検索

文献検索においては，精神疾病について扱った論文を抽出するためのフィルタ[3]および観察研究を抽出するためのフィルタ[4]についてそれぞれ最新のものを使用し，更にRTWについては可能な限りのイディオムを用いた上で網羅的な検索が可能となる検索式を作成した。またデータベースもMedline・PubMedに一本化せず幅広く収集するための検索戦略を検討するため，加えてPsycINFO・EMBASE, CINAHLを用いて横断的に検索を行った。

3. 検索した研究の選別過程

横断的データベース検索により出力された研究について，研究結果の重複を除去した上で統合した上で，第１段階として研究結果を発表年代別に２グループに分けた上で，それぞれ１つのグループについて独立した２名の研究者が，個々の研究のタイトル・要旨のみを閲覧した上で，研究の適格性について判定を行った。第２段階として，２つのグループにおいて的確と判定された研究を統合した上で，２名の独立した研究者が研究の全文を閲覧した上で，最終的な的確研究を判定した。

（倫理面への配慮）

本研究は「人を対象とする医学系研究に関する倫理指針」に基づいて行った。本研究に含まれる全ての系統的レビュー・予測モデル作成は，一次研究を集積・統合した二次研究であり，不利益・危険性の排除や説明同意，プライバシー保護等は一次研究で担保されているものを扱うため，問題は生じないと考えられる。

C. 研究結果

1. 文献検索

PubMed・ PsycINFO・EMBASE・CINAH

Lのデータベースごとに用いた検索式と検索数結果を**表1**に示す。PubMedにおける検索研究結果数は1,817件，PsycINFOでは1,418件，EMBASEでは2,161件，CINAHLでは768件であった。

2. 論文適格性判定結果

論文の適格性の判定結果・過程を**図1**に示す。4つのデータベースの横断検索結果の統合・重複除去により4,568件の研究が対象となった。適格性１次チェックの段階で第1グループの2,464件のうち適格性有りと判定されたものは69件であり，第2グループの2,463件のうち適格性有りと判定されたものは62件であり2グループ併せて合計131件について，２次チェックを行った。２次チェックの結果，最終的に14件の研究が本レビューにおける適格研究と判定された。

3. 最終適格判定研究

２段階にわたる適格性判定の結果，最終的に適格判定となった14件の研究[5~18]について**表2**に示す。Frenchらの研究では，米国の労働者データベースを用いて過去30日の休業の有無や日数と飲酒や喫煙習慣との関連について解析しており，喫煙については有意な関連を認めたことが報告されている[5]。Floderusらの研究では，1,350名の病気休業者を対象として，復職と生活習慣との関連について解析しており，50歳以上の対象者において睡眠との有意な関連が報告されている[6]。Sonnenscheinらの研究では，59名のバーンアウトした従業員を対象に，完全復職の関連要因を解析しており，入眠障害や良好な睡眠と有意な関連を認めたことが報告されている[7]。Akerstedtらの研究では，8,300名の労働者を対象として，病気休業からの復職と睡眠障害との関連について解析している（有意な関連は認めなかった）[8]。Fossらの研究では，8,333名の労働者を対象として，8週間以上の長期病気休業と喫煙・飲酒との関連について解析しており，飲酒問題とは有意な関連を認めたことを報告している[9]。Gliseらの研究では，2,683を対象とし，病気休業と喫煙や運動との関連について解析しており，これらとの有意な関連を報告している[10]。Saloらは56,732名を対象とし，業務不能状態からの復帰と睡眠障害との関連を解析しており，両者に有意な関連が認められたことが報告されている[11]。Lallukkaらは5,986名を対象とし，病気退職と各種睡眠障害（入眠障害や中途覚醒など）との関連を解析しており，有意な関連を認めたことが報告されている[12]。Sandmarkらの研究では233名を対象として，長期病気休業からの業務復帰と自己評価睡眠品質との関連を解析しているが，両者に有意な関連を認めなかった[13]。Holmaらの研究では269名を対象とし，業務不能による年金受給期間と飲酒障害との関連について報告している（有意な関連は認められず）[14]。Quistらの研究では7,401名を対象として長期長期休業の発生と運動や喫煙との関連について解析しており，これらとの有意な関連を報告している[15]。Silva-Juniorらは385名を対象として，長期病気休業と飲酒・喫煙との関連について検討しており，有意な関連を認めたことが報告されている[16]。Moritaらは54名の対象者について病気休業期間と運動習慣について解析しており，両者に有意な関連を認めた[17]。Van den

Bergらは8,364名の対象者について病気休業と喫煙・飲酒習慣との関連について検討しており，喫煙については有意な関連を認めたと報告している[18]。

D. 考察

本研究で抽出された研究数は最終的に14件のみであった。労働者の休職からの復職をアウトカムにした関連するライフスタイルを探索した先行研究は，現時点で非常に限られていることが判明した。また，最終結果の14件の研究において，休職からの復職のアウトカムに有意に関連していた生活習慣として，喫煙・睡眠・飲酒・運動が同定されたが，本研究では，抽出対象とする研究のライフスタイルとしてかなり広範囲を対象として検索・抽出を進めていたにもかかわらず，最終的に研究の対象となっていたライフスタイルは非常に限られたものとなっていた。以上のことより，今後労働者の休職からの復職にどのようなライフスタイルが関わっているのかについての観察研究の精力的な実施によるエビデンスの積み上げが必要であると思われる。この積み上げにより，より広範囲なライフスタイルとの関連性も検討されるべきである。これらの観察研究の積み上げが，より実践的な介入研究（復職に効果的な生活習慣改善）を検討する上での基礎エビデンスとなり得る。

E. 結論

本研究では精神疾患による休業からのRTWに関連する要因についての体系的レビューを実施し，14編の研究を抽出した。このうち，精神疾病による休業者の復職に関連する生活習慣要因として有意な関連を認めたものとして，喫煙・睡眠・飲酒・運動が同定された。

F. 参考文献

1. Blank L, Peters J, Pickvance S, Wilford J, Macdonald E. A systematic review of the factors which predict return to work for people suffering episodes of poor mental health. J Occup Rehabil 18:27-34, 2008.
2. Cornelius LR, van der Klink JJ, Groothoff JW, Brouwer S. Prognostic factors of long term disability due to mental disorders: a systematic review. J Occup Rehabil 21:259-274, 2011.
3. Wilczynski NL, Haynes RB. Optimal search strategies for identifying mental health content in MEDLINE: an analytic survey. Ann Gen Psychiatry 5:4, 2006.
4. Search Filters for Various Databases: PubMed http://libguides.sph.uth.tmc.edu/search_filters/pubmed_filters
5. French MT, Zarkin GA. Mental health, absenteeism and earnings at a large manufacturing worksite. J Ment Health Policy Econ 1:161-172, 1998.
6. Floderus B, Goransson S, Alexanderson K, Aronsson G. Self-estimated life situation in patients on long-term sick leave. J Rehabil Med 37:291-299, 2005.
7. Sonnenschein M, Sorbi MJ, Verbraak MJ, Schaufeli WB, Maas CJ, van Doornen LJ. Influence of sleep on symptom improvement and return to work in clinical burnout. Scand J Work Environ Health 34:23-32, 20

08.
8. Akerstedt T, Kecklund G, Selen J. Disturbed Sleep and Fatigue as Predictors of Return from Long-term Sickness Absence. Industrial Health 48:209-214, 2010.
9. Foss L, Gravseth HM, Kristensen P, Claussen B, Mehlum IS, Skyberg K. Risk Factors for Long-Term Absence Due To Psychiatric Sickness: A Register-Based 5-Year Follow-Up From the Oslo Health Study. Journal of Occupational and Environmental Medicine 52:698-705, 2010.
10. Glise K, Hadzibajramovic E, Jonsdottir IH, Ahlborg G, Jr. Self-reported exhaustion: a possible indicator of reduced work ability and increased risk of sickness absence among human service workers. Int Arch Occup Environ Health 83:511-520, 2010.
11. Salo P, Oksanen T, Sivertsen B, Hall M, Pentti J, Virtanen M, Vahtera J, Kivimaki M. Sleep disturbances as a predictor of cause-specific work disability and delayed return to work. Sleep 33:1323-1331, 2010.
12. Lallukka T, Haaramo P, Lahelma E, Rahkonen O. Sleep problems and disability retirement: a register-based follow-up study. Am J Epidemiol 173:871-881, 2011.
13. Sandmark H. Health, sleep, and professional career in female white-collar workers back to work after long-term sick-listing due to minor mental disorders. Scand J Public Health 39:823-829, 2011.
14. Holma IA, Holma KM, Melartin TK, Rytsala HJ, Isometsa ET. A 5-year prospective study of predictors for disability pension among patients with major depressive disorder. Acta Psychiatr Scand 125:325-334, 2012.
15. Quist HG, Thomsen BL, Christensen U, Clausen T, Holtermann A, Bjorner JB, Andersen LL. Influence of lifestyle factors on long-term sickness absence among female healthcare workers: a prospective cohort study. BMC Public Health 14:1084, 2014.
16. Silva-Junior JS, Fischer FM. Long-term sickness absence due to mental disorders is associated with individual features and psychosocial work conditions. PLoS One 9:e115885, 2014.
17. Morita G, Hori H, Katsuki A, Nishii S, Shibata Y, Kubo T, Suga K, Yoshimura R, Nakamura J, Group SUJ. Decreased Activity at the Time of Return to Work Predicts Repeated Sick Leave in Depressed Japanese Patients. J Occup Environ Med 58:e56-57, 2016.
18. van den Berg S, Burdorf A, Robroek SJW. Associations between common diseases and work ability and sick leave among health care workers. Int Arch Occup Environ Health, 2017.

G. 健康危険情報

なし

H. 研究発表

1. 論文発表

1. Itani O, Jike M, Watanabe N, Kaneita Y:

Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. Sleep Med 32:246-256, 2017.

2. Kinjo A, Imamoto A, Ikeda M, Itani O, Ohida T, Kaneita Y, Kanda H, Tanihata T, Higuchi S, Osaki Y: The Association Between Alcohol-Flavoured Non-Alcoholic Beverages and Alcohol Use in Japanese Adolescents. Alcohol Alcohol 52:351-357, 2017.
3. Otsuka Y, Kaneita Y, Itani O, Nakagome S, Jike M, Ohida T: Relationship between stress coping and sleep disorders among the general Japanese population: a nationwide representative survey. Sleep Med 37:38-45, 2017.
4. Tokiya M, Kaneita Y, Itani O, Jike M, Ohida T: Predictors of insomnia onset in adolescents in Japan. Sleep Med 38:37-43, 2017.
5. 兼板佳孝, 大井田隆, 今村聡, 羽鳥裕, 地家真紀, 井谷修, 土器屋美貴子, 大塚雄一郎, 中込祥: 2016 年　日本医師会員の喫煙とその関連要因に関する調査. 日本医師会雑誌 146:1409-1422, 2017.
6. Itani O, Kaneita Y, Tokiya M, Jike M, Murata A, Nakagome S, Otsuka Y, Ohida T: Short sleep duration, shift work, and actual days taken off work are predictive life-style risk factors for new-onset metabolic syndrome: a 7-year cohort study of 40,000 male workers. Sleep Med 39:87-94, 2017.
7. Otsuka Y, Kaneita Y, Nakagome S, Jike M, Itani O, Ohida T: Nightmares and sleep paralysis among the general Japanese population: a nationwide representative survey. Sleep Biol Rhythms 16:187-195, 2017.
8. Itani O, Kaneita Y, Jike M, Furuya M, Uezono C, Oda F, Agematsu R, Tokiya M, Otsuka Y, Ohida T: Sleep-related factors associated with industrial accidents among factory workers and sleep hygiene education intervention. Sleep Biol Rhythms 16:239-251, 2018.
9. Jike M, Itani O, Watanabe N, Buysse DJ, Kaneita Y: Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression. Sleep Med Rev 39:25-36, 2018.
10. Ohida N, Otsuka Y, Kaneita Y, Nakagome S, Jike M, Itani O, Ohida T: Factors related to alcohol consumption among Japanese physicians. Asia Pac J Public Health, *in press*.
11. Morioka H, Itani O, Jike M, Nakagome S, Otsuka Y, Ohida T: Risk Factors at Birth Predictive of Subsequent Injury among Japanese Pre-School Children: A Nationwide 5-Year Cohort Study. J Dev Behav Pediatr, *in press*.
12. Itani O, Kaneita Y, Doi K, Tokiya M, Jike M, Nakagome S, Ohtuka Y, Ohida T: Longitudinal epidemiologic study of poor mental health status in Japanese adolescents: incidence of predictive lifestyle factors. J Clin Psychiatry, *in press*.
13. Morioka H, Jike M, Kanda H, Osaki Y, Nakagome S, Otuska Y, Kaneita Y, Itani O, Higuchi S, Ohida T: The association between sleep disturbance and second-hand smoke exposure: a large-scale, nationwide, cross-

sectional study of adolescents in Japan. Sleep Med, *in press*.

2. 学会発表

1. Watanabe N, Jike M, Itani O, Buysse D, Kaneita Y: Long Sleep Duration and Health Outcomes: A Systematic Review, Meta-Analysis and Meta-Regression. the 2017 Annual Meeting of the American Psychiatric Association, San Diego, USA, 2017.5
2. 井谷修, 兼板佳孝, 土器屋美貴子, 地家真紀, 中込祥, 大塚雄一郎, 日本の中学生・高校生における睡眠随伴症（寝ぼけ・悪夢・金縛り）発症についての縦断的疫学研究. 日本睡眠学会第42回定期学術集会, 横浜, 2017.6
3. 大塚雄一郎, 兼板佳孝, 中込祥, 地家真紀, 土器屋美貴子, 井谷修, 高校生に対する睡眠衛生教育の効果に対する検討. 日本睡眠学会第42回定期学術集会, 横浜, 2017.6
4. 土器屋美貴子, 兼板佳孝, 井谷修, 地家真紀, 中込祥, 大塚雄一郎, 日本における思春期の不眠発症の予測因子. 日本睡眠学会第42回定期学術集会, 横浜, 2017.6
5. Watanabe N, Jike M, Itani O, Buysse D, Kaneita Y: LONG SLEEP DURATION AND HEALTH Outcomes: A SYSTEMATIC REVIEW, META-ANALYSIS AND META-REGRESSION, World Sleep 2017, Prague, Czech Republic, 2017.10
6. 井谷修, 兼板佳孝, 原野悟, 土器屋美貴子: 労働者の休養の評価尺度質問票の作成. 平成29年度産業保健調査研究発表会, 東京, 2017.10
7. 渡辺範雄, 地家真紀, 井谷修, 兼板佳孝：長時間睡眠と健康アウトカム　系統的レビューとメタ回帰分析. 第76回日本公衆衛生学会総会, 鹿児島, 2017.10
8. 土器屋美貴子, 兼板佳孝, 井谷修：高校生におけるインターネット依存と睡眠障害との関連性について. 第76回日本公衆衛生学会総会, 鹿児島, 2017.10
9. 井谷修: 思春期におけるメディア・ネット依存問題について（シンポジウム: 思春期公衆衛生領域におけるトピックス～研究の最前線から未来への発信～）. 第 76 回日本公衆衛生学会総会, 鹿児島, 2017.10

I. 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
なし

表1-1　Pubmedにおける検索式および検索結果数

	MEDLINE through PubMed (1,817)	
#1	employment [MH] OR employment [TW] OR employee*[TW] OR employable[TW]	151,975
#2	occupations[MH] OR occupation*[TW] OR "Occupational Health" [MH]	304,579
#3	worker[TW] OR labor[TW] OR work[MH]	158,122
#4	vocational[TW] OR "Rehabilitation, Vocational"[MH]	19,150
#5	**#1 OR #2 OR #3 OR #4**	**552,495**
#6	depress*[ALL]	466,949
#7	behav*[ALL]	1,568,435
#8	"mental disorders"[MH]	1,080,291
#9	psych*[ALL]	705,979
#10	**#6 OR #7 OR #8 OR #9**	**2,865,572**
#11	"return to work"[TW] OR "return to work"[MH] OR "return work"[TW] OR "return to job"[TW] OR "return job"[TW]	8,022
#12	absenteeism[TW] OR absenteeism[MH]	10,705
#13	“work disability”[TW]	1,716
#14	(sick*[TW] AND (leave[TW] OR list*[TW] OR absen*[TW])) OR "Sick Leave"[MH]	12,626
#15	"sick time"[TW]	44
#16	retirement[TW]	16,000
#17	reinstate*[TW]	5,380
#18	resumption[TW]	8,760
#19	retread[TW]	6
#20	"work status"[TW]	1,661
#21	"work ability”[TW]	1,263
#22	"comeback"[TW] OR "come back"[TW]	1,110
#23	**#11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #18 OR #19 OR #20 OR #21 OR #22**	**60,893**
#24	"epidemiologic studies"[MH:NOEXP]	7,435
#25	"case control"[TW] OR "case control studies"[MH]	890,627
#26	"cohort study"[TW] OR "cohort studies"[TW] OR "cohort studies"[MH] OR "cohort analy*"[TW]	1,692,964
#27	"follow up study"[TW] OR “follow up studies” [TW]	589,743
#28	"observational study"[TW] OR “observational studies”[TW]	98,506
#29	longitudinal[TW]	238,850
#30	"cross sectional"[TW] OR “cross-sectional studies” [MH:NOEXP]	352,444
#31	**#24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30**	**2,334,705**
#32	**#5 AND #10 AND #23 AND #31**	**1,817**

表1-2 PsycINFOおける検索式および検索結果数

PsycINFO through EBSCO (1,416)

S1	MA employment OR TX employment OR TX employee OR TX employable	132,881
S2	MA occupations OR TX occupation* OR MA “Occupational Health"	144,524
S3	TX worker OR TX labor OR MA work	128,915
S4	TX vocational OR MA "Rehabilitation, Vocational"	66,289
S5	**S1 OR S2 OR S3 OR S4**	**358,808**
S6	TX depress*	323,039
S7	TX behav*	1,466,459
S8	MA "mental disorders"	72,233
S9	**S6 OR S7 OR S8**	**1,697,552**
S10	TX "return to work" OR MA "return to work" OR TX "return work" OR TX "return to job" OR TX "return job"	2,237
S11	TX absenteeism OR MA absenteeism	5,105
S12	TX “work disability”	559
S13	TX "sick leave" OR MA "sick leave"	1,856
S14	TX "sick time"	23
S15	TX retirement	10,091
S16	TX reinstate*	4,294
S17	TX resumption	1,129
S18	TX retread	3
S19	TX "work status"	965
S20	TX "work ability”	862
S21	TX "comeback" OR TX "come back"	495
S22	**S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S18 OR S19 OR S20 OR S21**	**24,157**
S23	(ma “case control studies+”) or (ma “prospective studies+”) or (ma “cross sectional studies”) or (ma “case studies”) or (ma “epidemiological research”)	2
S24	((case or crosssectional or “cross sectional” or epidemiologic* or observational) and (study or studies)) or “case control*” or “case report*” or cohort* or “cross sectional” or followup* or “follow up*” or followed or longitudinal* or prospective* or retrospective* (ma “case control studies+”) or (ma “prospective studies+”) or (ma “cross sectional studies”) or (ma “case studies”) or (ma “epidemiological research”)	721,056
S25	**S23 OR S24**	721,056
S26	**S5 AND S9 AND S22 AND S25**	**1,416**

表1-3 CINAHLおける検索式および検索結果数

CINAHL through EBSCO (759)

S1	MH employment OR TX employment OR TX employee* OR TX employable	85,590
S2	MH occupations OR TX occupation* OR MH “Occupational Health"	148,347
S3	TX worker OR TX labor OR MH work	104,914
S4	TX vocational OR MH "Rehabilitation, Vocational"	9,840
S5	**S1 OR S2 OR S3 OR S4**	**295,173**
S6	TX depress*	94,370
S7	TX behav*	246,379
S8	MH "mental disorders"	365
S9	**S6 OR S7 OR S8**	**318,462**
S10	TX "return to work" OR MH "return to work" OR TX "return work" OR TX "return to job" OR TX "return job"	3,229
S11	TX absenteeism OR MH absenteeism	3,852
S12	TX “work disability”	565
S13	TX "sick leave" OR MH "sick leave"	4,036
S14	TX "sick time"	180
S15	TX retirement	11,426
S16	TX reinstate*	1,387
S17	TX resumption	796
S18	TX retread	2
S19	TX "work status"	648
S20	TX "work ability”	466
S21	TX "comeback" OR TX "come back"	730
S22	**S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S18 OR S19 OR S20 OR S21**	**22,558**
S23	(mh “case control studies+”) or (mh “prospective studies+”) or (mh “cross sectional studies”) or (mh “case studies”) or (mh “epidemiological research”)	4
S24	((case or crosssectional or “cross sectional” or epidemiologic* or observational) and (study or studies)) or “case control*” or “case report*” or cohort* or “cross sectional” or followup* or “follow up*” or followed or longitudinal* or prospective* or retrospective* (mh “case control studies+”) or (mh “prospective studies+”) or (mh “cross sectional studies”) or (mh “case studies”) or (mh “epidemiological research”)	593,220
S25	pt case study	186,195
S26	**S23 OR S24 OR S25**	**732,683**
S27	**S5 AND S9 AND S22 AND S26**	**759**

表1-3　Embaseおける検索式および検索結果数

Embase (2,151)

#1	employment'/exp OR employment OR emploee* OR employable	66,390
#2	occupations'/exp OR occupation* OR 'occupational health'/exp	463,315
#3	worker OR labor OR 'work'/exp	338,753
#4	vocational OR 'rehabilitation, vocational'/exp	21,866
#5	**#1 OR #2 OR #3 OR #4**	**695,492**
#6	depress*	562,928
#7	behav*	1,369,696
#8	mental disorders'/exp	1,506,278
#9	**#6 OR #7 OR #8**	**2,647,018**
#10	return to work' OR 'return to work'/exp OR 'return work' OR 'return to job' OR 'return job'	8,225
#11	absenteeism OR 'absenteeism'/exp	12,483
#12	work disability'	4,866
#13	sick leave' OR 'sick leave'/exp	4,859
#14	sick time'	47
#15	retirement	10,257
#16	reinstate*	5,868
#17	resumption	12,249
#18	retread	3
#19	work status'	1,753
#20	work ability'	1,149
#21	comeback OR 'come back'	1,257
#22	**#10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #18 OR #19 OR #20 OR #21**	**54,986**
#23	epidemiologic studies'/de	142,070
#24	case control' OR "case control studies'/exp	13,090
#25	cohort study' OR 'cohort studies' OR 'cohort studies'/exp OR 'cohort analy*'	316,319
#26	follow up study' OR 'follow up studies'	49,159
#27	observational study' OR 'observational studies'	146,830
#28	longitudinal study' OR 'longitudinal studies'	96,910
#29	cross sectional' OR 'cross-sectional studies'/de	288,780
#30	**#23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29**	**1,062,120**

図1　論文抽出過程のフローチャート

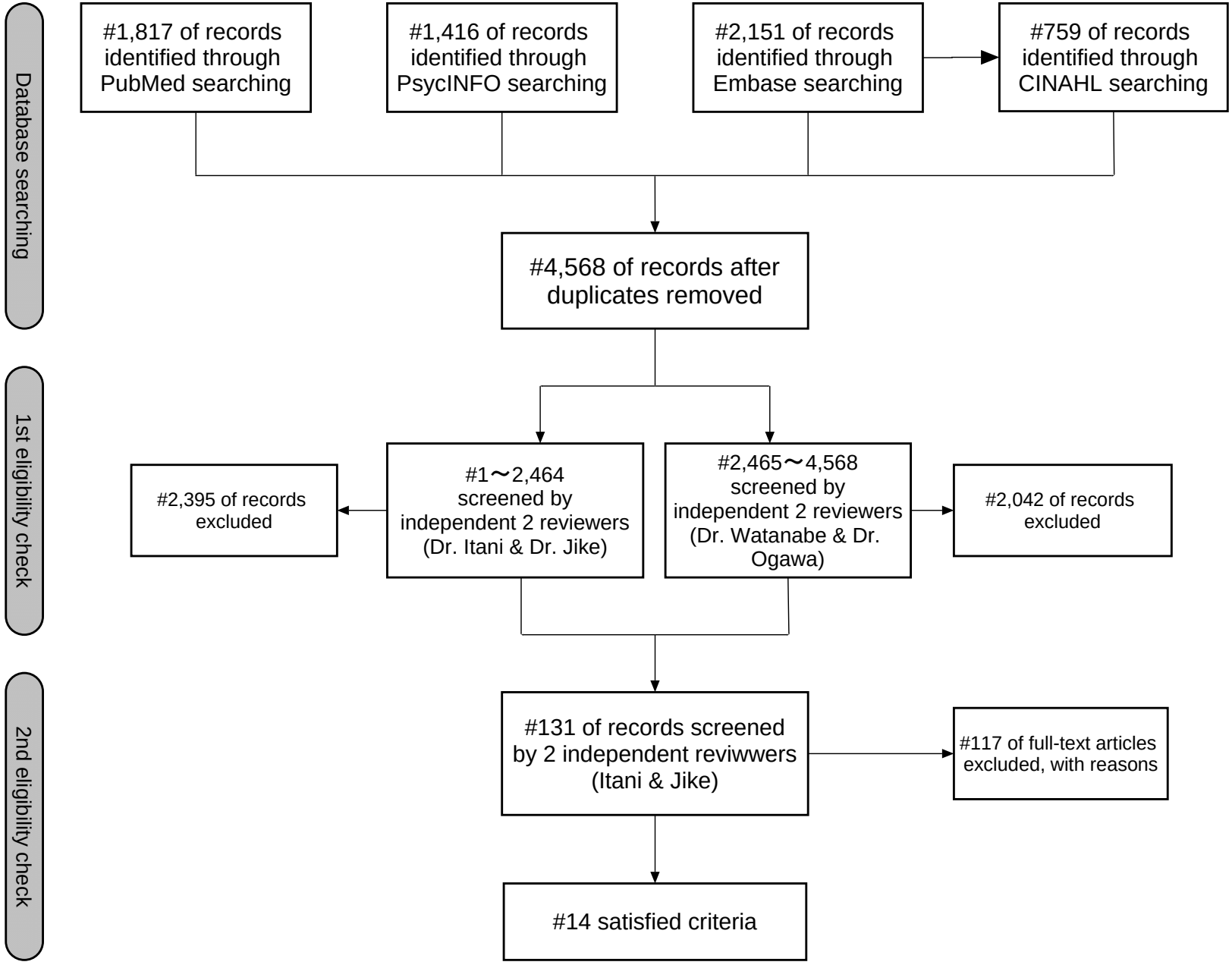


表2　最終抽出研究一覧

研究	報告年	実施国	観察期間	標本数	アウトカム	生活習慣要因	調整因子	主要結果
French et al.	1998	USA	cross-sectional	408	days absent during the past 30 days due to sickness or injury	Daily Drinker **Smoked cigarettes in the past year**	Emotion3	OR 1.023 (NS) **OR 1.704 (*P*<0.05)**
Floderus et al.	2005	Sweden	cross-sectional	1,350	positive and negative consequences of long-term sick leave and return to work	**Sleep** Lifestyle (food habits, physical exercise, reading, etc.) Smoking Use of alcohol	gender and sick leave diagnosis	20-50 years ・Positive consequences sleep OR 0.81(0.38-1.71) lifestyle OR 1.12(0.64-1.95) smoking OR 1.18(0.61-2.31) alcohol OR 0.86(0.54-1.36) ・Negative consequences sleep OR 1.12(0.70-1.80) lifestyle OR 0.94(0.60-1.48) smoking OR 0.58(0.31-1.08) alcohol OR 0.76(0.37-1.60) **>50 years** **・Positive consequences** **sleep OR 3.73(1.38-10.06)** lifestyle OR 0.90(0.40-2.04) smoking OR 1.21(0.42-3.48) alcohol OR 1.22(0.54-2.74) ・Negative consequences sleep OR 0.99(0.45-2.17) lifestyle OR 0.55(0.26-1.12) smoking OR 1.09(0.44-2.71) alcohol OR 2.79(0.69-11.25)
Sonnenschein et al.	2008	Denmark	6 months	99	full return to work	Recovery through sleep **Trouble falling asleep** **Refreshing sleep**	Age, Gender, Education, Follow-up duration Recruited through the Internet, Duration of sick leave, Comorbid depression, Explained variance, Exhaustion, and Partial sick leave	**Trouble falling asleep OR 0.45(0.22-0.93)** **Refresching sleep OR 2.82(1.06-7.46)**
Akerstedt et al.	2010	Sweden	2 years	8,300	return from sickness absence	Disturbed sleep	Age, gender, marital status, children at home, socioeconomic group, heavy physical work, twisted work posture, shift work, overtime work, fulltime/part-time work, work demands, work influence, social support, and fatigue	・14-89 abscent days OR 0.83(0.63-1.10) ・≥90 abscent days OR 0.70(0.41-1.19)
Foss et al.	2010	Norway	4 years	8,333	Long-Term (8 wk) Sickness Absence	Smoke cigarettes **Alcohol problem**	age, education, work-related factors, and self-reported general, mental health, lifestyle and social affiliation	OR 1.3(0.9-1.9) **OR 2.3(1.4-3.7)**
Glise et al.	2010	Sweden	2 years	2,683	・14 days of ongoing sick leave (SA14) ・a period of 60 days of sick leave during the last 12 months (SA60)	**Smoking or using snuff** **Level of physical activity(Moderate/intense v. Light)**	sex, age, marital status, care-giving responsibilities for children and for adult relatives, and length of education	・SA14 **Smoking RR 1.3(1.0-1.8)** **Physical activity RR 1.6(1.1-2.3)** ・SA60 Smoking RR 1.3(0.9-1.3) **Physical activity RR 1.8(1.2-2.6)**

研究	報告年	実施国	観察期間	標本数	アウトカム	生活習慣要因	調整因子	主要結果
Salo et al.	2010	Finland	5 years	56,732	not returning to work after work disability	**Sleep disturbances**	age, socioeconomic status, night/shift work, smoking, alcohol intake, body mass index, physical activity, diagnosed somatic disease, use of pain killers, depression, anxiety, and use of anxiolytics	**・Men** **HR 4.35(1.72-11.11)** ・Women HR 1.04(0.80-1.35)
Lallukka et al.	2011	Finland	8 years	5,986	Disability Retirement	**Any sleep problems** **Difficulties falling asleep** **Nocturnal awakenings** **Difficulties staying asleep** **Nonrestorative sleep**	gender, age, marital status, occupational class, working overtime, shift work, work environmental exposure, physical working conditions, computer work, psychosocial job strain, sleep duration, physiciandiagnosed diseases, smoking, heavy drinking, physical inactivity, and obesity	**HR 3.71(1.29-10.69)** **HR 3.53(1.88-6.61)** **HR 2.15(1.11-4.14)** **HR 2.45(1.34-4.48)** **HR 2.62(1.35-5.09)**
Sandmark	2011	Sweden	34 months	233	back to work after long-term sick-listing	**Self-rated quality of sleep**	none	**No trouble falling asleep OR 2.93(1.53–5.62)** **Sleeping all night without waking up OR 2.90(1.50–5.60)** **Well-rested and not tired during daytime OR 4.06(1.78–9.26)** **Do not wake up too early in the early morning OR 2.24(1.21–4.18)** **Get enough sleep at night (46 hours/night) OR 2.69(1.42–5.08)**
Holma et al.	2012	Finland	5 years	269	time to work disability pension	Alcohol use disorders	gender and age	HR 0.99(0.44-2.26)
Quist et al.	2014	Denmark	12 months	7,401	onset of long-term sickness absence	**Leisure-time physical activity** **Smoking**	age, tenure, physical and psychosocial work factors, occupational type and previous long-term sickness absence	**HR 1.34(1.02-1.77)** **HR 1.23(1.06-1.42)**
Silva-Junior et al.	2014	Brazil	case-control	385	long-term sickness absence	**Tobacco use** **Alcohol intake**	age, sex, Self-reported skin color, Education, Comorbidities, Violence at workplace	**OR 8.08(2.24-29.19)** **OR 9.07(1.05-78.12)**
Morita et al.	2016	Japan	2 years	54	Time to sick leave	**Activity**	age, gender, number of sick leave episodes, number of job changes, number of hospital admissions, marital history, results of the psychiatric symptom assessment, results of the Social Adaptation Self-evaluation Scale score and results of cognitive function tests	**HR 3.38(P=0.018)**
van den Berg et al.	2017	Netherland	cross-sectional	8,364	Sick leave 1–365 days	**Heavy smoking** >10 glasses alcohol/wk	gender, age, and educational level	**OR 2.86(1.11-7.38)** OR 0.44(0.24-0.78)

労災疾病臨床研究事業費補助金
分担研究報告書

自閉症スペクトラム障害者の復職、就職のための心理社会的介入

研究分担者　小川 雄右
京都大学大学院医学研究科　健康増進・行動学分野　助教

研究要旨

自閉症スペクトラム障害（ASD）者の就労、職場復帰のための心理社会的介入の有効性を評価することを目的に、全ての現存する無作為割り付け対照試験を網羅的に収集し、系統的レビュー・メタアナリシスを実施した。

包括的な文献検索を行い、7研究（対象者247名）が同定された。就労者の割合に関しては就労支援の評価をした1研究、コミュニケーションスキルトレーニングの評価をした2研究で報告されていた。販売のためのスキルや、座学、地域の協力者との協力、適応的行動分析技法、社会的コミュニケーションスキルを含む9ヶ月のプログラムであるProject SEARCH及びASD支援を行った群では1年後に31人中28人が就労を達成したのに対し、行わなかった群では23人中1人のみであった。コミュニケーションスキルトレーニングの評価をした2研究の統計学的統合(メタアナリシス)を行った結果、コミュニケーションスキルトレーニングを行った場合は行わなかった場合に比べて就労達成者の割合のリスク比は2.27であったが、有意差はなかった(95%信頼区間0.73-7.07)。就労に関するインタビュースコアについては、3研究でメタアナリシスを行った。コミュニケーションスキルトレーニングを行った場合は行わなかった場合に比べ、有意にスコアが向上した（標準化平均差 0.56, 95%信頼区間0.05-1.06）。

ASD者に就労のための心理社会的介入を行うことにより、就労に必要なコミュニケーションスキルが向上し、就労できる者の割合が増加する可能性が示唆された。今後は就職や復職、復職にまでの期間など、臨床に有用なアウトカムを測定した心理社会的介入を評価する無作為割り付け対照試験の実施が望まれる。

A. 研究目的

精神疾患は長期休職の最大の原因であり(Kessler 2003)、精神疾患による休職は貧困や社会的孤立に関連することが指摘されている(Henderson 2011)。休職期間を短くし、罹患前の生産性を回復することは、患者本人だけではなく

家族、職場、社会にとって重要である。このためには、薬物療法だけではなく、心理社会的介入が重要である。

昨年度は、精神疾患による長期療養者の職場復帰に関する、症状改善・固定や早期復職のための心理社会的介入に関する先行研究の収集を実施し、これまでの知見と限界をまとめた。うつ病、適応障害、統合失調症、自閉症スペクトラム障害（ASD）など、休職、職場復帰に関わることの多い精神障害に関する文献レビューを実施し、それぞれの疾患に関して心理社会的介入の効果を検討されている無作為割り付け対照試験（RCT）や、それらを統合した系統的レビューを同定した。うつ病、適応障害、統合失調症に関しては、様々な心理社会的介入の効果を評価したRCT及び系統的レビュー・メタアナリシスが実施されていたが、ASDについては、文献レビューは存在するものの、包含されたRCTは包括的ではなく、またメタアナリシスも行われていなかった。

そこで、今回、ASD者に対する心理社会的介入のASD者の職場復帰、生産性の回復に関して心理社会的介入の有効性を評価することを目的に、系統的レビュー・メタアナリシスを実施した。これにより、ASD者の職場復帰、生産性の回復、医療資源の効率的利用に資することができると考えられる。

B. 研究方法

本研究では、ASD者の就職、復職、就労のための技能の習得・改善のための心理社会的な介入の効果を検討している全てのRCTを対象とした。準無作為割り付け対照試験は除外した。先行研究よりも幅広く研究を同定するために、感度の高いデータベースの検索式を作成した。検索はMEDLINE、CENTRAL、PsycINFO、CINAHL、Scopus、ProQuest Dissertations & Theses、ICTRPにおいて実施した。検索用語としては、developmental disorders, Asperger等のASDに関する語、return to work, supported employment, absenteeism, sick leaveなどの就職、復職に関する語と、RCTを検索するためのCochrane highly sensitive search strategyを組み合わせて2017年7月6日に検索を行った（参考資料）。言語による制限は行わなかった。更に、本研究のテーマに関連する系統的レビュー、本研究で同定した文献が引用した文献、本研究で同定した文献を引用した文献の調査も行い、関連研究の包括的な調査を行った。

本研究で検討する介入は、就労、復職、仕事能率の改善を目的とした全ての心理社会的介入であり、復職支援プログラム、各種精神療法などを対象とした。個人に対する介入、グループに対する介入いずれも対象とした。アウトカムは復職、就職、仕事能率、就労に関するコミュニケーションスキルである。

検索によって同定された全ての研究のタイトル及び抄録を2人の研究者が独立して採択基準を満たしているかを評価し、満たす可能性があると判断した論文は全文を入手し、更に独立した2人の研究者で採択基準を満たしているかを評価した。意見が一致しない場合は討議して評価を決定した。

データ抽出

採択基準に合致した論文から以下のデータを抽出した。

・就労した人数
・復職した人数
・復職までの期間

・仕事能率
・就労に支援を要した時間
・賃金・就労に関するコミュニケーションスキルの評価尺度。

データ抽出は、2人の評価者が独立して行った。評価の一貫性を保つため、データ抽出用の構造化されたフォームを用いた。意見が一致しない場合は討議を行った。文献より得られなかったアウトカムについて原研究の研究者に問い合わせた。

研究の質の評価

採択した研究のバイアスのリスクはCochrane Collaborationのrisk of bias toolにより評価した。割付け順番の作成がhigh riskで、割付けの隠蔽が明らかになされていない研究は除外した。研究の質の評価は2人の評価者が独立して実施し、意見が一致しない場合は討議して評価を決定した。

統計解析

二値変数については、リスク比（risk ratio; RR）とその95%信頼区間を、連続変数については標準化平均差（standardized mean difference; SMD）とその95%信頼区間を算出した。包含した研究間で、対象者及び介入に臨床的異質性があると考えられたため、重み付け平均にはランダム効果モデルを用いた。統計的異質性の可能性を検討するために、I^2統計量を算出した。有意性の評価には95%信頼区間を用いた。メタアナリシスには、Review Manager 5.3を用いた。

GRADEシステムによる評価

統合されたアウトカムごとにGTADEシステムによる評価を行った。すなわち、研究のバイアス、結果の非一貫性、エビデンスの非直接性，推定値の不精確さ，出版バイアス、効果の大きさの6項目について，統合されたアウトカムごとに評価を行い、それぞれのエビデンスの質及び推奨の強さを判定した。

（倫理面への配慮）

本研究は既存の資料を用いた二次研究であるため、研究対象者に対する人権擁護上の配慮、研究対象者に対する不利益、危険性の排除、説明と同意（インフォームド・コンセント）に関わる状況などに関する問題は生じず、倫理面での問題は生じないと判断した。

C. 研究結果

図1に研究選択のフローダイアグラムを示す。2017年7月に文献検索を実施し、684の重複文献を除き、2207の文献が同定された。タイトルと抄録から採択基準を満たす可能性があると判断された88研究につき本文の評価を行い、8研究が採択基準を満たしていた。1研究は現在実施中であり、今回の研究の対象となったRCTは7研究（対象者247人）であった。

包含された研究の特徴と質

表1に今回包含したRCTの特徴を示す。全てのRCTは2012年以降に出版されていた。対象者は4研究においては10代後半、2研究では雇用されていない成人、1研究では雇用されていないもしくは限定的な雇用下にある成人、1研究では就労中の成人、1研究は学生、非雇用、雇用された成人が含まれていた。7研究のうち、3研究は就労支援プログラムの効果を評価されており、4研究は就労のためのコミュニケーションスキルトレーニングの効果

が評価されていた。3研究の介入は訓練された介入者によって実施され、4研究はコンピュータやスマホ型携帯デバイスによって実施されていた。

介入の詳細

就労支援プログラムの効果を評価した3研究のうち、1研究（Gentry, 2015）は仕事のリマインダー、仕事のリストやイラスト・ビデオによる促し、行動的自己管理適応、ジョブコーチとのコミュニケーションからなるスマートフォン型携帯デバイスを用いたプログラムであり、1研究（Wehman, 2017）は、販売のためのスキルや、座学、地域の協力者との協力、適応的行動分析技法、社会的コミュニケーションスキルを含む9ヶ月のプログラムであるSEARCHプロジェクト、もう1研究（Hagner, 2012）はグループでの訓練・スタッフによる個人を中心とした面接・就労後の調査と計画の実行による追跡援助からなる家族中心の就労移行プログラムであった。

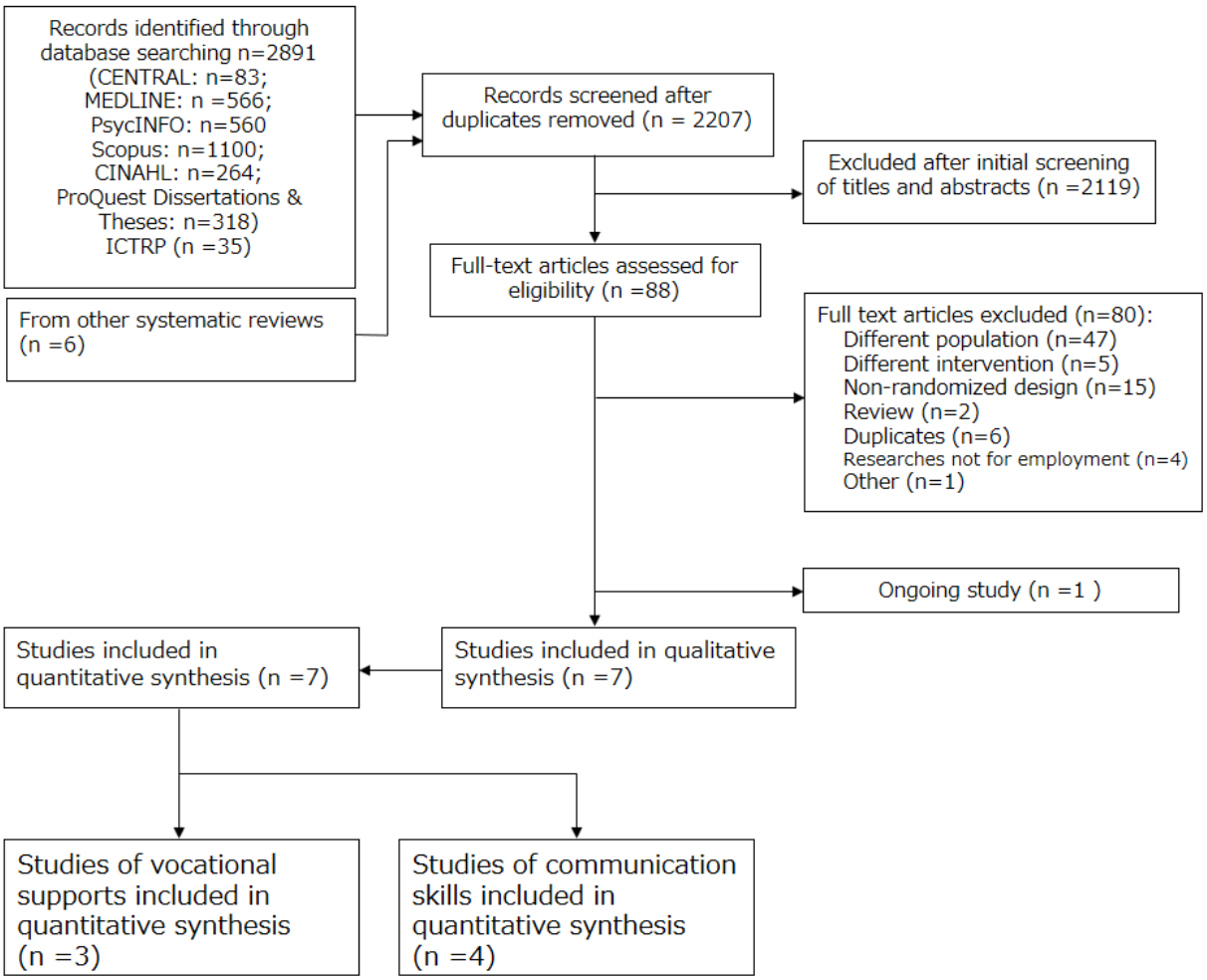


図1. 研究選択のフローダイアグラム

コミュニケーションスキルトレーニングの効果を評価した4研究のうち、コンピュータやタブレット端末用いた介入を行ったのは3研究だった。1研究（Strickland, 2013）は「興味ある仕事の決定」、「仕事を見つける」、「就労する」、「仕事を続ける」、「仕事に関する他のトピック」という5つのセクションからなるコンピュータを用いたマルチメディア就労訓練プログラム、1研究（Smith, 2014）は仕事で必要になる会話と面接でのパフォーマンスに焦点を当てた、会話能力向上のためのコンピュータを用いた仮想現実職業インタビュー訓練、1研究（Hayes, 2015）はモデルのビデオと、参加者自身が応答するビデオを用いたスマートフォン型携帯デバイスによるプログラム、1研究(Morgan, 2014）は討論やロールプレイ、ビデオフィードバック、ピアレビュー、ゲームの要素を含む、面接の成功に必要な実用的な社会スキルの向上を目的とした、グループ介入が評価されていた

包含された研究のバイアスのリスク

バイアスのリスクの評価の結果を図2、3に示す。2研究はsequence generationで低リスクと評価された。allocation concealmentにつき記載されている研究は無かった。今回包含した研究で行われたような心理社会的介入では対象者及び介入者へのブラインド化は困難であるため、全ての研究においてblinding participants and personnelは高リスクであった。blinding assessorに関しては、2研究ではブラインド化がなされていることが報告されていたが、2研究では不明、2研究では高リスクであった。incomplete outcome dataについては、4研究で低リスク、2研究で不明、1研究で高リスクであった。selective outcome reportingは2研究で低リスク、4研究で不明、1研究で高リスクであった。Treatment fidelityについては、3研究で低リスク、4研究で不明であった。他のバイアスのリスクについては、1つの研究で5人の対象者が2回エントリーされており高リスクであったが、他の研究では低リスクであった。

介入の効果

就労者の割合については就労支援の効果を評価をした1研究、コミュニケーションスキルトレーニングの効果を評価した2研究で報告されていた。Project SEARCH及びASD支援を行った場合は31人中28人が1年後に就労していたのに対し、高等学校特別教育プログラムを行った群では23人中1人のみであった。コミュニケーションスキルトレーニングの効果を評価をした2研究のメタアナリシスを行った結果、コミュニケーションスキルトレーニングを行った場合は行わなかった場合に比べて就労者の割合のRRは2.27であったが、有意差はなかった（95% 信頼区間0.73-7.07）（図4）。GRADE評価によるエビデンスの質は、推定値の不正確さと出版バイアスで1ずつのグレードダウンと効果の大きさで1グレードアップで、中等度の質と判断した。

Gentry（2015）らの研究では、スマートフォン型携帯デバイスによる就労支援群では12週の介入直後には62.5±23.1時間就労できていたのに対し、delayed intervention群では週59.4±16.6時間の就労であった。また、スマートフォン型携帯デバイスによる就労支援群では12週の介入直後のSupports Intensity Scale (SIS)が29.7±10.9であったのに対し、delayed intervention群のスコアは33.3±11.4であった。就労に支援を要した時間は、

スマートフォン型携帯デバイスによる就労支援プログラム群では3ヶ月の介入終了後に月に7.6±3.2時間であったのに対し、delayed intervention群では月16.6±7.5時間であった。賃金については1研究（Wehman 2017）で報告されており、Project SEARCH及びASD支援を行った群では1年後には時給8.6±2.9USドルであったのに対し、高等学校特別教育プログラムを行った群では時給0.6±2.4USドルであった。

就労に関するインタビュースコアについては、3研究で評価されており、メタアナリシスを行った結果、コミュニケーションスキルトレーニングを行った場合は行わなかった場合に比べ、有意にスコアが向上していた（SMD 0.56, 95%信頼区間0.05-1.06）（図5）。GRADE評価によるエビデンスの質は、研究のバイアスと推定値の不正確さで1ずつのグレードダウンと効果の大きさで1グレードアップで、中等度の質と判断した。

休職者が復職できた割合、復職までにかかった時間を報告した研究はなかった。

表1 包含した研究の特徴

Study, region	N	Age_ mean women (%)	diagnosis	diagnostic criteria	intellectual level	participants conditions	Format of interventions	Interventions	Control	treatment frequency	intervention period	follow-up from baseline	complete in inter-vention group	main outcomes
1.Studies of vocational supports														
Gentry (2015) USA	Intervention 28 Control 27	24 8 (16%)	ASD	*(school record or medical report)	6% beyond high school, 68% high school diploma, 24% high school certiicate, 2% less than high school	all worked in competitive employment settings	smartphone-like portable device	smartphone-like portable device based vocational support included: (1) task reminders, (2) task lists, (3) picture prompts, (4) video-based task-sequencing prompts, (4)behavioral self-management adaptations, (5) way-finding tools, (6) communication with the job coach	12 weeks delayed intervention	every work day	12 weeks	24 weeks	26 (92.9%)	coaching support hours, support intensity
Wehman (2017) USA	Intervention 31 Control 23	19.8 14 (25.9%)	Autism 71%, PDD-NOS 18%, Asperger 8%	from a previous medical diagnosis by a qualified health care provider and/or the identification on the individual's Individualized Education Plan (IEP))	no information	students	trained staff	Project SEARCH (PS) plus Autism Spectrum Disorder Supports. PS is an intensive 9-month job training program, including internships learning marketable skills, 180hrs of classroom time, collaboration between multiple community partners to support students. This study added structure and intensity of the learning experiences by ensuring the use of applied behavior analytic techniques, in addition to increasing the specific social communication skills. We adjusted instructional and behavioral plans based upon the review and analysis of data.	special education program as determined by their Individualized Education Plan	10–12 week internships (approx. 720 hrs) and classroom time (180hrs)	9 months	21 months	29 (93.5%)	employment, wage, support hours, worked hours, support intensity
Hagner (2012) USA	Intervention 24 Control 23	17.6 2 (4.3%)	ASD	The Autism Diagnostic Observation Schedule	no information	high school students	four fulltime staff who were master's-level professionals	a family-centered transition planning project on the transition Individualized Education Plans consisting of (a) group training sessions for families in the transition process, (b) person-centered planning meetings facilitated by project staff, and (c) follow-up assistance with career exploration and plan implementation	1 year delayed intervention	Once a month for group training sessions for families and 2.5 hrs per week for career exploration activities	9 months	12 months	no data	student and family expectations, self-determination, and career decision-making ability

2.Studies of communication skill training

Hayes (2015) USA	Intervention 8 Control 7	18 2 (13.3%)	ASD	diagnosis in their school records receiving transition services from a public school system	Be able to speak and perform duties related to a variety of job functions	students in high school or graduated high school but still receiving services from the transition program	smartphone-like portable device	smartphone-like portable device application to watch peer model videos as well as record their own videos in response to system-delivered prompts.	no intervention (given the app at the end of the study)	used VidCoach on average 2.02 times per day	1 months	1 months	7 (87.5%)	Mock Job Interviews
Strickland (2013) USA	Intervention 11 Control 11	18 0 (0%)	high functioning autism or Asperger's Disorder	—	but high functioning autism or Asperger's Disorder	between the ages of 16 and 19 years, mean years of school completed 11.2	computer	a multimedia employment training program that offers five sections of "Determining Career Interests," "Finding a Job," "Getting a Job," "Keeping a Job", and "Other Job Topics" like "Leaving a Job."	no intervention	once 30-min virtual practice session	30 minutes	7 days	11 (100%)	responses to interview questions, behaviors related to greetings and farewells
Morgan (2014) USA	Intervention 13 Control 15	24.5 1 (3.6%)	ASD	DSM-IV-TR	above 70	Student 48%, unoccupied 52%, employed 22%	group meeting	low-intensity group-delivered intervention aimed at increasing social–pragmatic skills with an emphasis on those essential to a successful job interview including discussion, role-play, video feedback, peer review, and games.	waitlist	weekly 90-min meetings	12 weeks	12 weeks	12 (92.3%)	Mock Job Interviews
Smith (2014) USA	Intervention 16 Control 10	24.2 6 (23.1%)	ASD	T-score of 60 or higher using parent and self-report versions of the Social Responsiveness Scale, Second Edition (SRS-2)	at least a 6th grade reading level	adult unemployed or underemployed (i.e., working less than half time and looking for additional work)	computer software or via the internet	Virtual Reality Job Interview Training designed to improve job interview performance by targeting job relevant interview content and interviewee performance	treatment as usual	five 10-hrs sessions within a two-week period	2 weeks	6 months	(91.3% session attendance)	Role-Play Performance Total Score, Job interview self-confidence

Abbreviations: ASD = Autism spectrum disorders; PDD-NOS = Pervasive Developmental Disorder - Not Otherwise

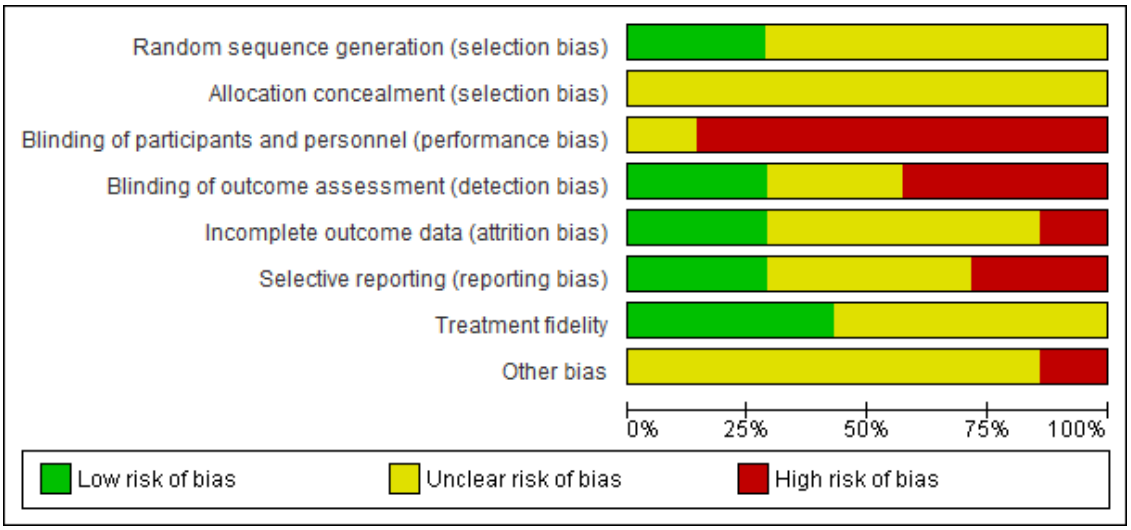


図2 バイアスのリスクの割合

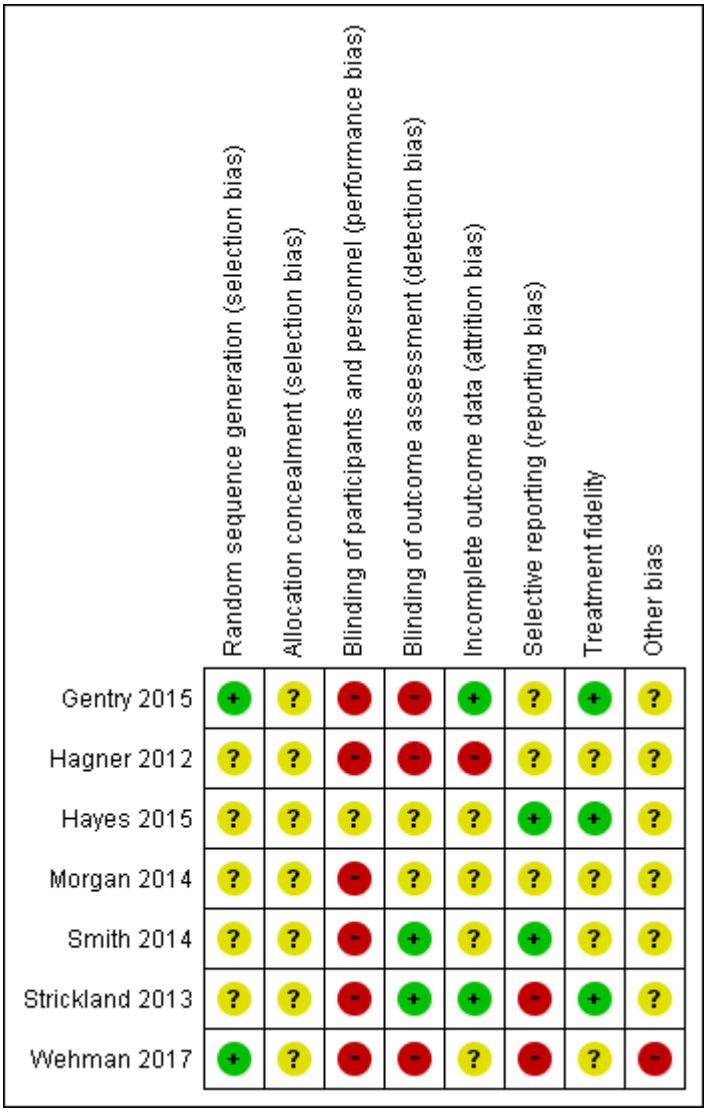


図3. バイアスのリスクのサマリー

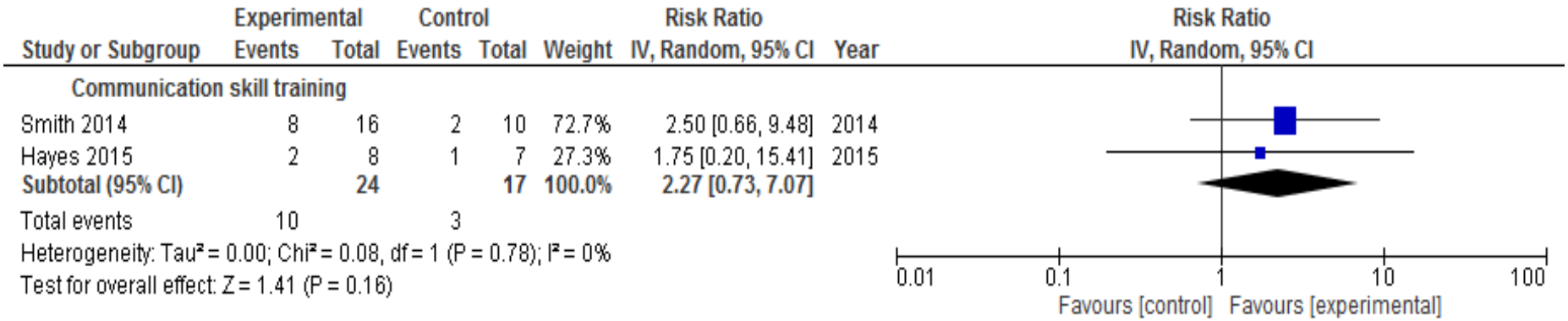


図4. フォレストプロット：就労者の割合

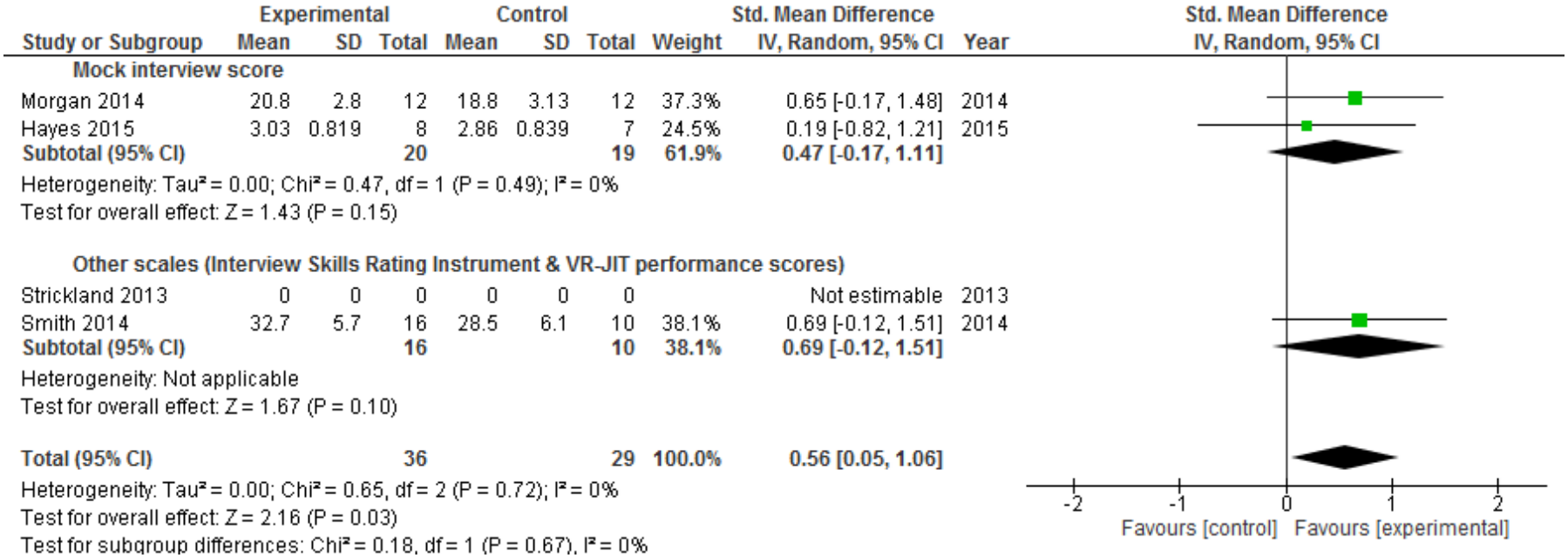


図5. フォレストプロット：コミュニケーション評価尺度のスコア

D. 考察

本研究では、ASD者に対し就労・復職のための心理社会的介入を実施した研究について、包括的な文献検索を行うことにより系統的レビューを実施し、統計学的統合を行った。ASD者に心理社会的介入を行うことにより、就労に必要なコミュニケーションスキルは向上し、就労できる者の割合が増加する可能性が示唆されたが、メタアナリシスの結果では就労できた者の割合について有意差は無かった。しかし、RRの点推定値は2.27であり、ASD者に対する心理社会的介入は就労に役立つ可能性がある。いずれのアウトカムについてもGRADEによるエビデンスの質は中等度であり、コストがかかることや資源が不十分ではあるものの、推奨度は強いと判定した。

就労できる者の割合が介入群と対照群で有意な差を認めなかった要因としては、今回は2つのRCTでしか統計学的統合ができなかった点が挙げられる。また、多くの研究でバイアスのリスクに関する情報は報告されておらず、バイアスのリスクは不明であるドメイン、研究が多かった。また、就職について報告した研究は3研究しかなく、復職した人数や復職までの期間を報告したRCTはなかった。今後は就職や復職、復職にまでの期間など、臨床に有用なアウトカムを測定した心理社会的介入を評価するRCTの実施が望まれる。

先行研究からの知見、限界を検討し、本研究ではデータベースはMEDLINEに限定せず、CENTRALやScopus、PsycINFOに加え、学位論文や臨床試験登録システムなどのデータベースも検索を行い、より広く文献を同定できる検索式を作成し、先行研究よりも網羅的に文献を収集するための戦略を検討した。また、我々の知る限りでは、ASD者の就労、復職のための介入の効果を検討した、初めてのRCTの包括的な系統的レビュー及びメタアナリシスである。

E. 結論

本研究ではASD者に就労のための心理社会的介入を評価したRCTの系統的レビュー、メタアナリシスを実施し、心理社会的介入により就労に必要なコミュニケーションスキルが向上し、就労できる者の割合が増加する可能性が示唆された。今後は就職や復職、復職にまでの期間など、臨床に有用なアウトカムを測定した心理社会的介入を評価するRCTの実施が望まれる。

参考文献

1. Gentry, T., Kriner, R., Sima, A., McDonough, J., & Wehman, P. (2015). Reducing the Need for Personal Supports Among Workers with Autism Using an iPod Touch as an Assistive Technology: Delayed Randomized Control Trial. *Journal of Autism and Developmental Disorders, 45*, 669-684.
2. Hagner, D., Kurtz, A., Cloutier, H., Arakelian, C., Brucker, D. L., & May, J. (2012). Outcomes of a Family-Centered Transition Process for Students With Autism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities, 27*, 42-50.
3. Hayes, G. R., Custodio, V. E., Haimson, O. L., Nguyen, K., Ringland, K. E., Ulgado, R. R., . . . Weiner, R. (2015). Mobile video modeling for employment interviews for individuals with autism. *Journal of Vocational Rehabilitation, 43*, 275-287.

4. Morgan, L., Leatzow, A., Clark, S., & Siller, M. (2014). Interview Skills for Adults with Autism Spectrum Disorder: A Pilot Randomized Controlled Trial. *Journal of Autism and Developmental Disorders, 44*, 2290-2300.
5. Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., . . . Fleming, M. F. (2014). Virtual Reality Job Interview Training in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 44*, 2450-2463.
6. Strickland, D. C., Coles, C. D., & Southern, L. B. (2013). JobTIPS: A Transition to Employment Program for Individuals with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 43*, 2472-2483.
7. Wehman, P., Schall, C. M., McDonough, J., Graham, C., Brooke, V., Riehle, J. E., . . . Avellone, L. (2017). Effects of an employer-based intervention on employment outcomes for youth with significant support needs due to autism. *Autism, 21*, 276-290.

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

1. Hashimoto K, Ogawa Y, Takeshima N, Furukawa TA. Psychological and Educational Interventions for Atopic Dermatitis in Adults: A Systematic Review and Meta-Analysis. Behavior Change. 2017 April; 34(1):48-65.
2. Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, Ogawa S, Sato H, Kitagawa N, Shinagawa Y, Ikeda Y, Imai H, Tajika A, Ogawa Y, Akechi T, Yamada M, Shimodera S, Watanabe N, Inagaki M, Hasegawa A. Cognitive and Behavioral Skills Exercises Completed by Patients with Major Depression During Smartphone Cognitive Behavioral Therapy: Secondary Analysis of a Randomized Controlled Trial. JMIR Ment Health. 2018 Jan 11;5(1)
3. Ogawa Y, Furukawa TA, Takeshima N, Hayasaka Y, Atkinson LZ, Tanaka S, Cipriani A, Salanti G. Efficacy of antidepressants over placebo is similar in two-armed versus three-armed or more-armed randomized placebo-controlled trials. Int Clin Psychopharmacol. 2018 Mar;33(2):66-72.
4. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JPT, Egger M, Takeshima N, Hayasaka Y, Imai H, Shinohara K, Tajika A, Ioannidis JPA, Geddes JR. Comparative efficacy and acceptability of first- and second- generation antidepressants in the acute treatment of major depressive disorder: a network meta-analysis. Lancet. 2018 Apr 7;391(10128):1357-1366
5. Soares-Weiser K, Rathbone J, Ogawa Y, Shinohara K, Bergman H. Miscellaneous

treatments for antipsychotic-induced tardive dyskinesia. Cochrane Database Syst Rev. 2018. CD000208.

2. 学会発表

なし

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

なし

参考資料：検索式

MECLINE (via Ovid)

1 exp child development disorders, pervasive/
2 Developmental Disabilities/
3 pervasive development$ disorder$.tw.
4 (pervasive adj3 child$).tw.
5 (PDD or PDDs or PDD-NOS or ASD or ASDs).tw.
6 autis$.tw.
7 asperger$.tw.
8 kanner$.tw.
9 childhood schizophrenia.tw.
10 or/1-9
11 exp Occupational Therapy/
12 exp Occupational Diseases/
13 exp Occupational Medicine/
14 exp Disability Evaluation/
15 exp WORK/
16 return to work.mp.
17 occupational therap$.mp.
18 occupational intervention$.mp.
19 supported employment.mp.
20 employment.mp.
21 vocational rehabilitation.mp.
22 work capacity evaluation.mp.
23 vocational guidance.mp.
24 absenteeism.mp.
25 occupational health services.mp.
26 occupational health.mp.
27 unemployed.mp.
28 employed.mp.
29 unemployment.mp.
30 re-employment.mp.
31 sick leave.mp.
32 sick$ absence.mp.
33 retirement.mp.
34 disability pension.mp.
35 occupation$.mp.
36 job.mp.
37 vocational.mp.
38 employability.mp.
39 employable.mp.
40 employee*.mp.
41 retirement.mp.
42 work status.mp.
43 work ability.mp.
44 workability.mp.
45 work activity.mp.
46 work retention.mp.
47 job retention.mp.
48 job loss.mp.
49 job performance.mp.
50 work rehabilitation.mp.
51 or/11-50
52 randomized controlled trial.pt.
53 controlled clinical trial.pt.
54 randomized.ab.
55 placebo.ab.
56 drug therapy.fs.
57 randomly.ab.
58 trial.ab.
59 groups.ab.
60 or/52-59
61 exp animals/ not humans.sh.
62 60 not 61
63 10 and 51 and 62

PsycINFO (via Ovid)

1 exp pervasive developmental disorders/
2 Developmental Disabilities/
3 pervasive development$ disorder$.tw.
4 (pervasive adj3 child$).tw.
5 (PDD or PDDs or PDD-NOS or ASD or ASDs).tw.
6 autis$.tw.
7 asperger$.tw.
8 kanner$.tw.
9 childhood schizophrenia.tw.
10 or/1-9
11 exp Occupational Therapy/
12 exp Disability Evaluation/
13 return to work.mp.
14 occupational therap$.mp.
15 occupational intervention$.mp.
16 supported employment.mp.
17 employment.mp.
18 vocational rehabilitation.mp.
19 work capacity evaluation.mp.
20 vocational guidance.mp.
21 absenteeism.mp.
22 occupational health services.mp.
23 occupational health.mp.
24 unemployed.mp.
25 employed.mp.
26 unemployment.mp.
27 re-employment.mp.
28 sick leave.mp.
29 sick$ absence.mp.
30 retirement.mp.
31 disability pension.mp.
32 occupation$.mp.
33 job.mp.
34 vocational.mp.
35 employability.mp.
36 employable.mp.
37 employee*.mp.
38 retirement.mp.
39 work status.mp.
40 work ability.mp.
41 workability.mp.
42 work activity.mp.
43 work retention.mp.
44 job retention.mp.
45 job loss.mp.
46 job performance.mp.
47 work rehabilitation.mp.
48 or/11-47
49 random$.tw.
50 factorial$.tw.
51 crossover$.tw.
52 cross-over$.tw.
53 placebo$.tw.
54 (doubl$ adj blind$).tw.
55 (singl$ adj blind$).tw.
56 assign$.tw.
57 allocat$.tw.
58 volunteer$.tw.
59 control*.tw.
60 "2100".md.
61 or/49-60
62 10 and 48 and 61

CENTRAL (The Cochrane Library)

1 [mh "child development disorders, pervasive"]

2 [mh "Developmental Disabilities"]

3 pervasive development* disorder*

4 (pervasive near/3 child*)

5 (PDD or PDDs or PDD-NOS or ASD or ASDs)

6 autis*

7 asperger*

8 kanner*

9 childhood near/1 schizophreni*

10 {or #1-#9}

11 [mh "Absenteeism"]

12 [mh "Sick Leave"]

13 [mh "Employment"]

14 [mh "Unemployment"]

15 [mh "Occupational Health"]

16 [mh "Occupational Health Services"]

17 [mh "Occupational Medicine"]

18 [mh "Rehabilitation, Vocational"]

19 Occupational Therapy

20 Occupational Diseases

21 occupational intervention$

22 ((sick* NEAR/3 leave) or (sick NEAR/3 list*) or (sick NEAR/3 absen*)):ti,ab,kw

23 (return* NEAR/3 work*):ti,ab,kw

24 "work disability" or employment or "re-employment" or unemployment or unemployed or employability or employable or employee* or "work capacity" or "retirement" or "work status" or "job satisfaction" or "work ability" or workability or "work activity" or "work retention" or "job retention" or "job loss" or "job performance" or "work rehabilitation"

25 ((sick* or absen*) AND (workplace or (work NEAR/2 related) or occupation* or job)):ti,ab,kw

26 disability pension

27 {or #11-#26}

28 #10 and #27

CINAHL

S51 S9 and S39 and S50

S50 S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49

S49 TI ((control* or prospectiv* or volunteer*)) OR AB ((control* or prospectiv* or volunteer*))

S48 (MH "Prospective Studies")

S47 (MH "Evaluation Research+")

S46 TI random* OR AB random*

S45 TI placebo* OR AB placebo*

S44 (MH "Placebos")

S43 TI (((singl* or doubl8 or tripl* or trebl*) N24 (blind* or mask* or dummy*))) OR AB (((singl* or doubl8 or tripl* or trebl*) N24 (blind* or mask* or dummy*))) OR SU (((singl* or doubl8 or tripl* or trebl*) N24 (blind* or mask* or dummy*)))

S42 TI (clin* N24 trial*) OR AB (clin* N24 trial*)

S41 (MH "Clinical Trials+")

S40 PT clinical trial

S39 S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 O

R S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38

S38 TI vocational OR AB vocational OR SU vocational

S37 TI Job OR AB Job OR SU Job

S36 TI Occupation* OR AB Occupation* OR SU Occupation

S35 TI Disability pension OR AB Disability pension OR SU Disability pension

S34 TI retirement OR AB retirement OR SU retirement

S33 TI Sick* absence OR AB sick* absence OR SU sick* absence

S32 TI Sick leave OR AB sick leave OR SU sick leave

S31 TI unemployment OR AB unemployment OR SU unemployment

S30 TI employed OR AB employed OR SU employed

S29 TI unemployed OR AB unemployed OR SU unemployed

S28 TI occupational health OR AB occupational health OR SU occupational health

S27 TI occupational health services OR AB occupational health services OR SU occupational health services

S26 TI absenteeism OR AB absenteeism OR SU absenteeism

S25 TI vocational guidance OR AB vocational guidance OR SU vocational guidance

S24 TI Work capacity evaluation OR AB Work capacity evaluation OR SU Work capacity evaluation

S23 TI vocational rehabilitation OR AB vocational rehabilitation OR SU vocational rehabilitation

S22 TI employment OR AB Employment OR SU Employment

S21 TI Supported employment OR AB Supported employment OR SU Supported employment

S20 TI Occupational intervention* OR AB Occupational intervention* OR SU Occupational intervention*

S19 TI Occupational therap* OR AB Occupational therap* OR SU Occupational therap*

S18 (MH “Occupational Therapy+”)

S17 (MH “Disability Evaluation+”)

S16 (MH “Work”)

S15 (MH “Sick Leave”)

S14 (MH “Rehabilitation, Vocational+”)

S13 (MH “Occupational Health+”)

S12 (MH “Employment+”)

S11 (MH “Job Re-Entry”)

S10 (MH “Job Performance”)

S9 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8

S8 (MH "Developmental Disabilities")

S7 childhood schizophren*

S6 Kanner*

S5 Rett*

S4 Asperger*

S3 (autis* or ASD or ASDs)

S2 (pervasive development* disorder* or PDD or PDDs)

S1 (MH "Child Development Disorders, Pervasive+")

Scopus

(TITLE-ABS-KEY("development* disorder*" OR "development* disabilit*" OR (pervasive W/3 child*) OR (PDD or PDDs or PDD-NOS or ASD or ASDs) OR autis* OR Asperger* OR kanner* OR “childhood schizophrenia”))

AND

(TITLE-ABS-KEY(“Occupational Therapy” OR “Occupational Diseases” OR “Occupational Medicine” OR “Disability Evaluation” OR “ return to work” OR “occupational therap*” OR “occupational intervention*” OR “supported employment” OR employment OR “vocational rehabilitation” OR “work capacity evaluation” OR “vocational guidance” OR absenteeism OR “occupational health service*” OR “occupational health” OR unemployed OR employed OR unemployment OR re-employment OR “sick leave” OR “sick* absence” OR retirement OR “disability pension” OR occupation* OR vocational OR employability OR employable OR employee OR retirement OR “work status” OR “work ability” OR workability OR “work activity” OR “work retention” OR “job retention” OR “job loss” OR “job performance” OR “work rehabilitation”))

AND

((INDEXTERMS ("clinical trials" OR "clinical trials as a topic" OR "randomized controlled trial" OR "Randomized Controlled Trials as Topic" OR "controlled clinical trial" OR "Controlled Clinical Trials" OR "random allocation" OR "Double-Blind Method" OR "Single-Blind Method" OR "Cross-Over Studies" OR "Placebos" OR "multicenter study" OR "double blind procedure" OR "single blind procedure" OR "cross over procedure" OR "clinical trial" OR "controlled study" OR "randomization" OR "placebo")) OR (TITLE-ABS-KEY (("clinical trials" OR " clinical trials as a topic" OR "randomized controlled trial" OR "Randomized Controlled Trials as Topic" OR "controlled clinical trial" OR "Controlled Clinical Trials as Topic" OR "random allocation" OR "randomly allocated" OR "allocated randomly" OR "Double-Blind Method" OR "Single-Blind Method" OR "Cross-Over Studies" OR "Placebos" OR "cross-over trial" OR " single blind" OR "double blind" OR "factorial design" OR "factorial trial"))) OR (TITLE-ABS (clinical trial* OR trial* OR rct* OR random* OR blind*)))

ProQuest Dissertations & Theses A&I

all(("development* disorder*" OR "development* disabilit*" OR (pervasive NEAR/3 child*) OR (PDD OR PDDs OR PDD-NOS OR ASD OR ASDs) OR autis* OR Asperger* OR kanner* OR "childhood schizophrenia"))

AND

all(("Disability Evaluation" OR "return to work" OR "occupational therap*" OR "occupational intervention*" OR "supported employment" OR employment OR "vocational rehabilitation" OR "work capacity evaluation" OR "vocational guidance" OR absenteeism OR "occupational he

alth service*" OR "occupational health" OR un employed OR employed OR unemployment O R re-employment OR "sick leave" OR "sick* a bsence" OR retirement OR "disability pension" OR occupation* OR vocational OR employabili ty OR employable OR employee OR "work sta tus" OR "work ability" OR workability OR "wo rk activity" OR "work retention" OR "job reten tion" OR "job loss" OR "job performance" OR "work rehabilitation"))

AND

(SU("Treatment Effectiveness Evaluation") OR all("Treatment Outcomes" OR "comparative stud*" OR (evaluat* NEAR/3 stud*) OR ((singl* OR doubl* OR trebl* OR tripl*) NEAR/3 (blind* OR mask*))) OR ti(randomized OR randomised OR randomly OR trial OR groups) OR ab(randomized OR randomised OR randomly OR trial OR groups))

労災疾病臨床研究事業費補助金
分担研究報告書

精神疾患による長期療養者の職場復帰に関する、運動・睡眠等、ライフスタイルに関する介入

研究分担者　地家　真紀
日本大学医学部社会医学系公衆衛生学分野　助教

研究要旨

近年の我が国において精神疾患等により休職する労働者が増加傾向にあること、さらには少子高齢化をうけ労働力の確保が大きな課題となっている。これらの休職者が復職できるよう医療機関や企業等においてさまざまな復職支援が行われている。このうち運動・睡眠等ライフスタイルに関するプログラムの介入効果を明らかにするためにPubMedおよび医中誌を用いて検索を実施し、介入研究を探索した。

本研究課題に合致した論文をPubMedでは3本、医中誌では17本存在した。運動の介入研究が多く実施されていたが、複数の内容のひとつとしてプログラム内で実施されていただけでなく、運動強度や実施期間・時間はさまざまであった。従ってライフスタイルに関するプログラム単一の効果を評価することは困難な状況であることが明らかになった。今後単一のプログラムによる介入研究が行われその効果が明らかにされることが重要である。

A. 研究目的

少子高齢化に伴い政府は一億総活躍社会の実現に向けて取り組んでいる一方で精神疾患等により休職する労働者は増加傾向である[1, 2]。多くの者が社会で活躍するためにも休職者の復職支援は行政、事業者、労働者にとって重要な課題となっている。

独立行政法人 労働政策研究・研修機構が実施した「メンタルヘルス、私傷病などの治療と職業生活の両立支援に関する調査」によると通常の年次有給休暇以外で連続して休職できる病気休職制度（慣行含む）をもつ事業場は調査対象の91.1%であること、さらに休暇の期限は事業場により異なり1年以下（「3か月まで」9.6%、「3か月超から6

か月まで」13.3%「6か月超から1年まで」22.3%）と規定する事業場が45.2%であることが報告された[3]。休職に期限が設けられているため休職者は期日までに勤務可能なレベルまで症状の改善をしなければ退職せざるを得ない。従って休職者にとって復職は大きな命題であり、復職支援は今後の人生にかかわる大きなサポートであるといっても過言ではない。

本研究課題において前年度は運動・睡眠等ライフスタイルを含めた復職支援プログラムに着目し日本及び諸外国での先行研究を把握することを目的としてMedline・PubMed及び医中誌を用いて文献の検索を実行した。そして運動・睡眠等ライフスタイルについての介入を実施した研究が少ないこと、さらに我が国ではリワークプログラムという復職支援が医療機関、地域障害者職業センター、企業等で実施されていることを報告した。リワークプログラムでは複数のプログラムが実施されていることが一般的であり本研究課題のライフスタイルに関するプログラムが実施されているかはプログラムごとに異なる。そこで本年度では運動や睡眠等ライフスタイルに関する介入が行われているか我が国の復職支援のプログラムである『リワークプログラム』の内容を確認し、さらには諸外国のライフスタイルに関する復職プログラムについて検索を実施、その効果を把握することを目的とした。

B. 研究方法

本研究課題に対応したPubMedおよび医中誌の検索を作成し、言語や出版年数を問わず検索を実行した。具体的な用語として労働者、精神疾患に関する用語、休職・復職に関する用語、ライフスタイルに関する用語を含めて検索を実行した。さらに以下に示す判定・除外基準を元に本研究課題に合致するか同定を行った。

対象者：精神疾患により休職中労働者であること
　　16歳～65歳
介入内容：ライフスタイルに関するもの
　　―運動、睡眠
　　介入を伴うRCTまたは前後比較・前向きコホート研究（単群・複数群比較も含める）
アウトカム：復職、休職期間の長期化、再休職までの期間の短縮

検索結果を元に独立した2人の研究者がタイトルおよび抄録を吟味し、1次チェックとして研究課題に即した内容か（1. 精神疾患により休職中労働者であること　2. ライフスタイルに関する介入を受けていること）を評価した。2人の研究者の判断に相違があったものは議論して該当、非該当いずれかに評価をした。1次チェックで該当と判断されたものはさらに2次チェックとして独立した2人の研究者が文献を吟味し本研究課題に該当するか（1. 精神疾患により休職中労働者であること　2. ライフスタイルに関する介入を受けていること　3. 復職に関する評価が行われていること）評価した。1次チェック同様、2人の研究者の評価が分かれた文献については議論し該当、非該当の2群に分類した。該当を評価されたものリスト化した（図1，2参照）。

（倫理面への配慮）

「人を対象とする医学系研究に関する倫理指針」に基づいて研究を実施したが、本研究は文献を系統的に検索しレビューを行うため個人の特定は不可能であり、対象者個人への不利益はないと考え

る。

C. 研究結果

検索式を用いて抽出された論文数はPubMed311本、医中誌378本であった。さらに1次チェックを行い該当と評価された論文数はPubMed14本、医中誌53本であった。そして2次チェックを行い該当と評価された論文数はPubMed3本、医中誌17本であった（図1, 2参照）。これらの結果を表1, 2にまとめた。

PubMedで抽出された論文は3論文であり、いずれも運動に関する論文であった。Kroghらは異なる種類の運動による介入を実施しうつ病の患者の症状の感受性に運動は生物学的な介在効果は支持できないが仕事の能力においてstrength trainingは有益な効果があると結論付けた。その一方でJoyceらが報告した精神疾患者の職場での介入についてシステマティックメタレビューでは、身体活動について従業員のメンタルヘルスに影響する可能性はあるが、必要な活動のタイプ、量、強度は不明であると示された[23]。

医中誌において抽出された論文は17論文あり、リワークプログラムという復職のためのプログラムの中でライフスタイルに関する内容が組み込まれていた。リワークプログラムは多岐にわたりそのプログラムの内容は各論文で詳細が示されている。具体的には基礎的作業能力回復プログラムやストレスマネジメントのプログラム、心理教育プログラムなどが確認された。また標準化リワークプログラムの5つの区分①個人プログラム、②特定の心理プログラム、③教育プログラム、④集団プログラム、⑤その他のプログラムにおいて本研究課題のひとつである運動はその他のプログラムに含まれる[24]。本研究課題において抽出された論文は運動に関する介入が多数報告されたが、その内容はヨガや太極拳といったものから有酸素運動まで運動強度はさまざまであった。プログラムスケジュールや実施期間は論文ごとに異なるため復職率やHAM-Dなどの指標を元にどの運動が効果的であるか評価することは不可能であった。

D. 考察

D.-1　PubMedによる論文収集

本研究課題において抽出された論文は3論文であり、いずれも運動に関する論文であった。運動介入は、運動の種類だけでなく介入の実施期間、1回あたりの運動時間等にも左右されることが考えられる。従って現時点でその効果について評価することは困難だと言わざるを得ない。また本研究課題において抽出された文献の中には対照を設定していない論文も存在した。今後質が担保された研究の遂行が期待される。

D.-2　医中誌による論文収集

本研究課題において抽出された論文は運動に関する介入が多数報告されたが、プログラムスケジュールや実施期間は論文ごとに異なるため復職率やHAM-Dなどの指標を元にどの運動が効果的であるか評価することは不可能であった。また、実際に行われている精神医療機関のリワークプログラムにおいてその他のプログラムの実施は減少傾向であることが報告されている[24]。効果的な運動強度や実施期間・時間の尺度がないため、運動による効果が見えづらいことによりその他のプログラムの実施が減少しているのかもしれない。運動単一の研究が進み運動を含めたライフスタイルに関するプログラムの効果が明らかになることを期待する。

E. 結論

運動・睡眠等ライフスタイルについての介入を実施した研究、特に運動を実施する介入が多くみられた。また介入の中で複数のプログラムが実施されていた。従ってライフスタイルに関したプログラム単一の効果を評価することは困難であった。今後単一のプログラムによる介入研究が行われその効果が明らかにされることを期待する。

参考文献

1. 政府広報オンライン　https://www.gov-online.go.jp/tokusyu/ichiokusoukatsuyaku/index.html
2. 業務上疾病発生状況等調査　厚生労働省 http://www.mhlw.go.jp/bunya/roudoukijun/anzeneisei11/h28.html
3. メンタルヘルス、私傷病などの治療と職業生活の両立支援に関する調査 http://www.jil.go.jp/institute/research/2013/documents/0112.pdf
4. Krogh, J., Saltin, B., Gluud, C., Nordentoft, M. The DEMO trial: a randomized, parallel-group, observer-blinded clinical trial of strength versus aerobic versus relaxation training for patients with mild to moderate depression. J Clin Psychiatry. 2009; 70(6): 790-800.
5. Netterstrom, B. and Bech, P. Effect of a multidisciplinary stress treatment programme on the return to work rate for persons with work-related stress. A non-randomized controlled study from a stress clinic. BMC Public Health. 2010 Nov 1; 10: 658.
6. Martin, M. H. and Nielsen, M. B. and Madsen, I. E. and Petersen, S. M. and Lange, T. and Rugulies, R. Effectiveness of a coordinated and tailored return-to-work intervention for sickness absence beneficiaries with mental health problems. J Occup Rehabil. 2013;23(4):621-30.
7. Miyaji S, Yamamoto K, Morita N, Tsubouchi Y, Hoshino S, Yamamoto H, Sakai Y, Tanaka K, Miyaoka H. The relationship between patient characteristics and psychiatric day care outcomes in schizophrenic patients. Psychiatry Clin Neurosci. 2008 Jun;62(3):293-300
8. 泉水 宏臣, 永松 俊哉, 井原 一成, 中川 正俊 回復期にある精神疾患患者を対象とした運動療法の試み。体力研究　107 巻 (2009) p. 15-22
9. 加藤ちえ、岸展江、菅原由衣、北川信樹、賀古勇輝、渡邉紀子、川井朋子。うつ病患者に対する職場復帰支援プログラムの取り組み　北海道作業療法 26(1): 34-43
10. 田島美幸、中村聡美、岡田佳詠、大野裕、秋山剛　うつ病休職者のための集団認知行動療法の効果の検証　産業医学ジャーナル 33(1): 54-59
11. 鈴木瞬、羽岡健史、吉野聡、松崎一葉　うつ病休職者におけるリワークを通した身体活動変化の検討。 スポーツ精神医学 10：95-99
12. 平澤勉、野際陽子。デイケア終了後の復職を予測するものは何か?　うつ病復職デイケア開始後 5 週目までの気分および疲労感の特徴に着目して　作業療法　30(6):707-716
13. 上田知子、深間内文彦、榎本稔。利用者が体験する『休職』と『リワークプログラム』　う

つ病のリワークプログラムに関する一考察　外来精神医療　12(2) :40-47
14. 大木洋子、五十嵐良雄。リワークプログラム利用者の復職後の就労継続性に関する効果研究　産業精神保健　20(4):335-345
15. 田代哲男、吉田周平、田代香織　うつ病復職支援プログラムとしての軽運動と高強度有酸素運動の有用性　スポーツ精神医学　10:62-68
16. 田代哲男、伊藤貴子、吉田周平、田代香織　うつ病寛解期の残遺状態に対する統合的復職支援プログラムの試み　多様な病態に対応するために　精神科治療学　28(6): 805-814
17. 酒井要、西嶋長　当院でのリワークプログラムの概要と利用実績の調査　滋賀県病院薬剤師会誌 36(1): 9-11
18. 林果林、松田由美江、土川怜、端こず恵、藤井悠、山本喜久、黒木宣夫。桂川修一　大学病院における復職支援デイケア利用者の復職にかかわる要因分析　復職要因と復職継続要因の検討　産業ストレス研究　21(2):191-199
19. 原田修一郎、大類真嗣、長谷川淳子、野田承美、森谷郁子、高橋由里、本庄谷奈央、佐々木妙子、伊藤真理子、林みづ穂　うつ病による休職者を対象にした復職デイケアの実践　復職と社会適応能力の関係を中心に　精神科治療学 29(12):1601-1608
20. 山口皓平、兵頭俊次、石束由紀、土居裕見子、山本武、渡部亜矢子、渡部三郎、善家誠、入江寛人　当院のリワークプログラムにおける効果と課題の検討　復職者に対する追跡調査を踏まえて　光会医療研究会誌 12(1): 6-11
21. Hayashi Karin, Taira Yoichi, Maeda Takamitsu, Matsuda Yumie, Kato Yuki, Hashi, Kozue, Kuroki Nobuo, Katsuragawa Shuichi. What inhibits working women with mental disorders from returning to their workplace? A study of systematic re-employment support in a medical institution. Biopsychosocial Medicine 2016,:1-6
22. 油谷元規、眞柄正隆、陣内沙織、堤康彦　森病院リワークプログラムのアウトカム　神奈川県精神医学会誌　0(65):11-18
23. 玉崎愛美、堤輝、松元知美、梅津舞子、手銭宏文、吉村玲児、中村純　研究と報告 うつ状態勤労者の復職支援を目的とした集団精神療法の効果　精神医学　59(2): 139-145
24. Joyce S, Modini M, Christensen H, Mykletun A, Bryant R, Mitchell PB, et al. Workplace interventions for common mental disorders: a systematic meta-review. Psychol Med. 2016; 46: 683-97.
25. 障害者職業総合センター職業センター実践報告書 No.26『精神障害者職場再適応支援プログラム～リワーク昨日を有する医療機関と連携した復職支援～』第 3 章精神科医療機関におけるリワークプログラムの現状 http://www.nivr.jeed.or.jp/download/center/practice26_3.pdf

F. 健康危険情報

なし

G. 研究発表

1. Itani O, Jike M, Watanabe N, Kaneita Y: Short sleep duration and health outcomes: a

systematic review, meta-analysis, and meta-regression. Sleep Med 32:246-256, 2017.
2. Otsuka Y, Kaneita Y, Itani O, Nakagome S, Jike M, Ohida T: Relationship between stress coping and sleep disorders among the general Japanese population: a nationwide representative survey. Sleep Med 37:38-45, 2017.
3. 兼板佳孝, 大井田隆, 今村聡, 羽鳥裕, 地家真紀, 井谷修, 土器屋美貴子, 大塚雄一郎, 中込祥: 2016 年　日本医師会員の喫煙とその関連要因に関する調査. 日本医師会雑誌 146: 1409-1422, 2017.
4. Itani O, Kaneita Y, Tokiya M, Jike M, Murata A, Nakagome S, Otsuka Y, Ohida T: Short sleep duration, shift work, and actual days taken off work are predictive life-style risk factors for new-onset metabolic syndrome: a 7-year cohort study of 40,000 male workers. Sleep Med 39:87-94, 2017.
5. Jike M*, Itani O*, Watanabe N, Buysse DJ, Kaneita Y: Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression. Sleep Med Rev, *in press*. *Jike M and Itani O contributed equally to this work as the first authors.
6. Otsuka Y, Kaneita Y, Nakagome S, Jike M, Itani O, Ohida T: Nightmares and sleep paralysis among the general Japanese population: A nationwide representative survey. Sleep Biol Rhythms, *in press.*
7. Ohida N, Otsuka Y, Kaneita Y, Nakagome S, Jike M, Itani O, Ohida T: Factors related to alcohol consumption among Japanese physicians. Asia Pac J Public Health, *in press*.
8. Itani O, Kaneita Y, Jike M, Furuya M, Uezono C, Oda F, Agematsu R,　Tokiya M, Otsuka Y, Ohida T: Sleep-related factors associated with industrial accidents among factory workers and sleep hygiene education intervention. Sleep Biol Rhythms, *in press.*

2. 学会発表

なし

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

なし

図1

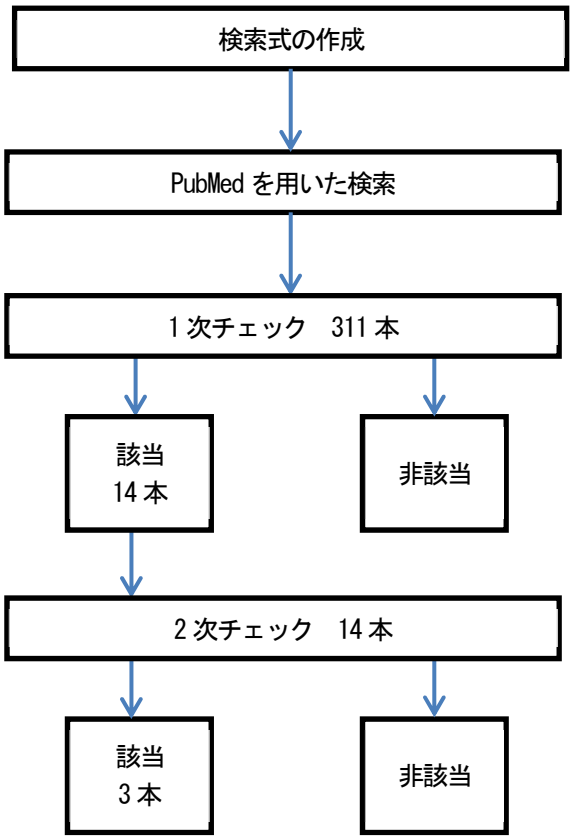
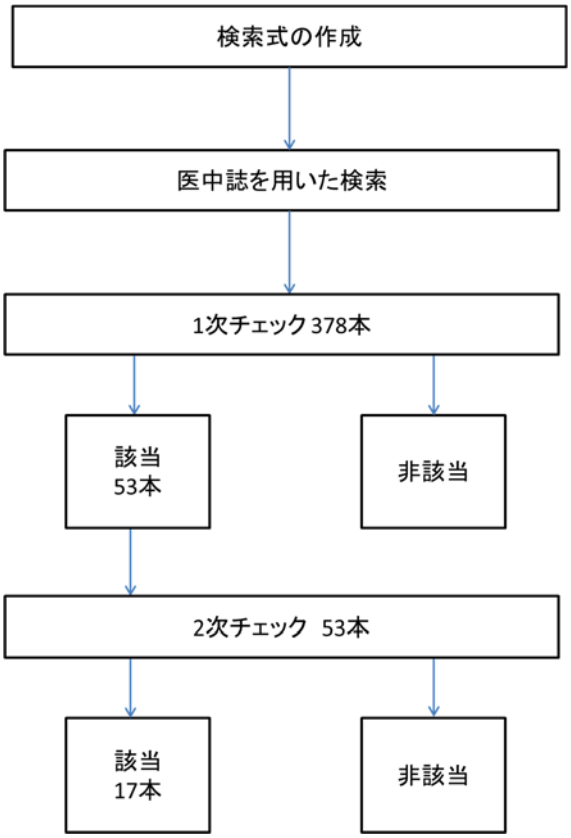

検索式の作成
PubMed を用いた検索
1 次チェック　311 本
該当
14 本
非該当
2 次チェック　14 本
該当
3 本
非該当

図2


検索式の作成
医中誌を用いた検索
1次チェック 378本
該当
53本
非該当
2次チェック 53本
該当
17本
非該当

資料（検索式）

：PubMed （sick[All Fields] AND (leave[All Fields] OR burnout[All Fields]) AND ("work"[MeSH Terms] OR "work"[All Fields]) OR worker[All Fields]) AND (((mental[All Fields]) AND (("disease"[MeSH Terms] OR "disease"[All Fields] OR "disorder"[All Fields]) OR ("disease"[MeSH Terms] OR "disease"[All Fields] OR "disorders"[All Fields]))) OR ((reactive[All Fields] AND (disorder[All Fields] OR disorder'[All Fields] OR disorder"[All Fields] OR disorders[All Fields])) OR "depressed"[All Fields]) OR ("depressive disorder"[MeSH Terms] OR ("depressive"[All Fields] AND "disorder"[All Fields]) OR "depressive disorder"[All Fields] OR "depression"[All Fields] OR "depression"[MeSH Terms]) OR depressive[All Fields] OR ("anxiety"[MeSH Terms] OR "anxiety"[All Fields]) OR ("psychotic disorders"[MeSH Terms] OR ("psychotic"[All Fields] AND "disorders"[All Fields]) OR "psychotic disorders"[All Fields] OR "psychosis"[All Fields]) OR psychotic[All Fields] OR ((("social adjustment"[MeSH Terms] OR ("social"[All Fields] AND "adjustment"[All Fields]) OR "social adjustment"[All Fields] OR "adjustment"[All Fields] OR "adaptation, psychological"[MeSH Terms] OR ("adaptation"[All Fields] AND "psychological"[All Fields]) OR "psychological adaptation"[All Fields])) AND (disorder[All Fields] OR disorders[All Fields])) OR (reactive[All Fields] AND NEXT[All Fields] AND ("depressive disorder"[MeSH Terms] OR ("depressive"[All Fields] AND "disorder"[All Fields]) OR "depressive disorder"[All Fields] OR "depression"[All Fields] OR "depression"[MeSH Terms]))
AND ("Return to Work"[MeSH] OR " absenteeism"[MeSH] OR "Sick time" [All Fields] OR "sickness absence" [All Fields] OR "sickout" [All Fields] OR "reinstatement" [All Fields] OR "resumption" [All Fields] OR "comeback" [All Fields])
AND ("sleep"[All Fields] OR "rest"[All Fields] OR "exercise"[All Fields] OR "sport"[All Fields] OR "stretch "[All Fields] OR "lifestyle" [All Fields])

：医中誌 ((労働/MTH or 職/AL or 労働者/AL) and (メンタル/AL or メンタルヘルス/AL or 精神疾患/AL or 精神病/AL or 精神病性障害/AL or うつ/AL or うつ病/AL or うつ病性障害/AL or 不安/AL or 燃え尽き症候群/AL or バーンアウト/AL) and (休職/AL or 復職/AL or 病欠/AL or 休業/AL) and (PT=原著論文) and (PT=症例報告除く)))。

表 1 PubMed 検索結果

著者名	発行年	対象者	実施国	実施年	介入内容	結果	
						復職率	その他指標
Krogh	2009	165	デンマーク	2005.1～2007.7	1strength 2aerobic traininig 3relaxation training	−	HAM-D 4month Strength 10.0±6.4 Aerobic 12.1±6.4 relaxation10.6±5.6 12month Strength 11.0±7.1 Aerobic 11.9±6.5 relaxation10.0±5.6
Netterstrom	2010	97	デンマーク	2002.11～2004.11	excesise		4months 5.1 (1.2-21.2) 1year 1.5(0.3-7.0) 2year 1.2(0.2-6.7)
Martin	2013	168	デンマーク		physical excesise		HR 0.50(-.34-0.75)

表2 医学中央雑誌 検索結果

著者名	発行年	対象者	実施国	実施年	介入内容(本研究課題に即する内容)		結果	
							復職率	その他指標
Miyajiら	2008	430	日本	1986.5～2004.8	Psychiatric day care	スポーツ、料理	49.3	
泉水	2009	17	日本	2007.10～2008.6	運動教室		-	
加藤ら	2009	22	日本	2006.4～2008.7	職場復帰支援プログラムスケジュール	軽スポーツ、料理	58.3	
田島ら	2010	51	日本	ー	集団認知行動療法		-	
鈴木	2011	28	日本	2008. 9～2011.5	リワークプログラム	ヨガ・卓球	-	**歩数 開始前5959±3011 終了前8229±2762**
平澤	2011		日本	平成20. 4～21.3	複数プログラム	太極拳・軽スポーツ	-	visual analogue scale(身体疲れ):5週目、1週目得点の増減 復職継続群-15 困難群18
上田	2012	11	日本	2009. 6～2011.5	リワークプログラム	スポーツ	45	-
大木ら	2012	556	日本	200.7～2011.6	リワークプログラム	頭と体のストレッチ、卓球	-	Cox比例ハザード プログラムの利用の有無1.869(1.115-3.224)
田代	2013	7	日本		有酸素運動療法		-	**HAM-D_{17} 参加前5.0 参加後2.2** **BDI-Ⅱ プログラム前8.7 終了時3.8**
田代ら	2013	12	日本	2008.10～2011	総合的復職支援プログラム	タウンウォーキング、SUN体操	100	**HAM-D_{17} 参加前5.0 参加後2.2** **BDI-Ⅱ プログラム前8.7 終了時3.8**
酒井	2014	23	日本	2011.11～2013.12	リワークプログラム	ラジオ体操、軽スポーツ	65.2	-
林ら	2014	154	日本	2007.10～2010.10	復職DC	ボディー・ワーク	70.8	
原田ら	2014	48	日本	2010.7～2013.6	復職デイケア	スポーツ	56.3	復職群 **BDI-Ⅱ 開始前21.3±10.2 終了時 10.7±7.7** **SASS 前28.1±6.8 終了時34.±6.5** 休職継続群 BDI-Ⅱ開始前25.4±9.2 終了時15.3±11.1 SASS 開始前26.2±7.1 終了時26.9±11.8 **
山口	2015	33	日本	H25.10～27.5	リワークプログラム	エクササイズ	76	QIDS-J 復職群 **参加前11.64±4.77 参加後6.72±5.13** 中断群 参加前10.63±4.44 参加後10.88±6.09
Hayashi	2016	454	日本	2007.10～2014.12	リワークプログラム	エクササイズ	63	
油谷ら	2016	21	日本	2013.5～2015.1	リワークプログラム	エクササイズ	66.7(リハビリ出社を含めると81)	**BDI-Ⅱ開始時 19.9±6.2 終了時8.4±6.3** **SE 開始時24.7±5.9 終了時32.2±8.6** **状態不安(STAI) 開始時48.7±8.2 終了時42.1±9.9** **特性不安(STAI)開始51.4時 53.8±5.5 終了時45.1±7.4** **CSQ 開始時186.9±49.0 終了時229.4±51.4**
玉崎	2017	48	日本	ー	集団精神療法	心理教育(睡眠に関する話)	70.8	HAM-D17 介入前6.5±2.3 介入後5.0±2.8 SASS-J 介入前28.6±7.7 介入後31.2±6.9

労災疾病臨床研究事業費補助金
平成 29 年度研究報告書

精神療法・疾患重症度・レジリエンス等複合化指標による疾患予後・復職の予測可能性の検討：系統的レビューのオーバービュー

研究分担者 渡辺 範雄
京都大学大学院医学研究科 健康増進・行動学分野 准教授

研究要旨

わが国では近年、精神疾患による労災認定請求・支給決定件数は過去最多を記録した。精神疾患による休職者の復職過程においては、疾患自体に対する効果的介入のみならず、職場・社会復帰のための介入の併用または追加が必要であり、かつ復職可能性を評価する判断指標が望まれる。

本研究では、前年度までに作成した研究計画をもとにこれらを扱ったコホート研究・無作為割り付け対照試験を扱った 20 本の系統的レビューを収集し、研究の質評価を行ったうえでオーバービューを行った。

精神疾患による休職者の復職可能性の指標として、各系統的レビューに共通して見られた知見では、高年齢、最近の疾患による休職歴、身体疾患合併等が低い復職可能性と関連し、逆に高い自己効力感や仕事における高い能力または仕事における低い要求水準は復職可能性が上ることを示した。一方、介入に関する系統的レビューでは、IPS モデル等の復職支援プログラムは、統合失調症をはじめとした重症精神疾患者の就職可能性を向上させるが、一般精神疾患では認知行動療法、また運動・食事・飲酒や喫煙・睡眠等生活指導の介入が有効であるというエビデンスは認めなかった。

今後は、疾患特異的な治療ではなく個々の患者の休職時状況に合わせ、現実的な問題を取り扱う介入方法の開発が必要であろう。

A. 研究目的

精神疾患による労働者に対する負担は大きい。例えば、OECD の 2010 年の報告によると、多くの国では休職の原因のほぼ半数を精神疾患が占める。わが国でも近年、精神疾患による労災認定請求・支給決定件数は過去最多を記録した。精神疾患で長期療養に至った労働者では、症状の完全消失とその後の社会復帰が望まれるが、薬物療法だけではこれを満足には達成できない。例えばうつ病の STAR*D 研究によると薬物療法で寛解まで至るのは 3 割程度[1]であった。この研究ではその後の就労状況に関する報告はないが、疾患自体が寛解しても就労・復職して疾患罹患以前のパフォーマンスを発揮できるようになるには相当の時間を要するし、また復職の段階で疾患が再燃してしまうこともありうる。そのため、精神疾患による休職においては、疾患自体に対する効果的介入のみならず、職場・社会復帰のための介入の併用または追加が必要と考えられる。さらに、どのような特性を持つ患者に高い復職可能性があるのか同定するための判断指標があれば、社会復帰のための治療資源の有効な配分につながる。

しかし、復職可能性の判断指標に関して、良質な系統的レビューの手法でまとめた社会復帰指標は、知られていない。また復職のための追加介入ではうつ病[2]、適応障害[3]等を対象とした系統的レビューはあるものの、患者特性によってどの治療を選択すべきかという指標はなく、また不安障害、自閉症スペクトラム障害等による休職者への復職のための追加介入に関する良質のレビューは、本研究の研究者らの知る限り存在しない。

そこで本研究では、精神疾患の疾患重症度や精神状態・自己効力感、職場環境、睡眠・運動等のライフスタイルデータ等の様々な指標の中から、精神疾患による休職者の復職に関する指標を系統的レビューのオーバービューを行うことで、探索する。また精神療法等の介入が、精神疾患で長期休職にいたった労働者の復職アウトカムやその後の再発予防について与える影響について、同じく系統的レビューのオーバービューを行うことで明らかにする。

平成28年度までに既存の研究を収集して、研究方法論の弱点等を明らかにし、本オーバービューの研究計画書を作成し、平成29年度はそれを実行することで結果のまとめを行った。計画も記すことで報告書として理解が容易になると考え、方法論もここに記載する。

B. 研究方法

本研究目的に合致したテーマを扱った国内外の系統的レビュー(SR)を、網羅的

文献検索を行い、同定した。

SR の登録基準は以下とした。

研究デザイン

前向きまたは後ろ向きコホート研究、または無作為割り付け対象試験(randomized controlled trial: RCT)を文献データベースにて検索して再現可能な方法によって集積した、系統的レビュー。メタ解析の有無は問わず質的報告のみ行っているものも許容し、また出版言語は問わなかった。

対象者

- 性別、国籍、人種を問わないが、研究に包含されている参加者はその 8 割以上が 18 歳から 55 歳までとした。
- ベースライン時点にて、精神疾患または精神症状尺度にて一定閾値以上の患者を研究対象、またはサブグループで独立して上記特性を満たす群について報告しているもの。身体疾患に続発したか否かは問わなかった。
- ベースライン時点にて、休職または失職、または就職していない。復職または再就職または就職をアウトカムとしてみるので、早期退職者を扱ったものは含まなかった。

介入・暴露

下記のいずれか、または複数要素をカテゴリー変数、あるいは連続変数で測定し、アウトカムとの相関を見ているもの扱っているものを選択した。

- 精神疾患の疾患重症度や精神状態
- レジリエンス・自己効力感
- 職場環境
- 睡眠
- 運動
- 食事
- 心理社会的介入。精神療法を含み、対照群があるもの。薬物療法のみを行っているものは、疾患治療が主で復職のための介入とは認められないため、除外した

アウトカム

下記のいずれか、または複数要素を測定し、二値変数であれば群別の実数または母数に対する割合かまたは対照群に対するリスク比・ハザード比・オッズ比等を報告しているもの、連続変数であれば群別の平均値と標準偏差・標準誤差または対照群に対する加重平均差・標準化加重平均差を報告しているものを選択した。

- 復職・再就職の有無
- 復職・再就職までの時間(連続変数)
- 復職後の仕事効率(連続変数)
- 復職後の精神状態(連続変数)
- 復職後の再休職の有無

文献検索

Medline、PsycINFO を利用して網羅的な文献検索により同定した。通常、薬物療法の系統的レビューを行う際には、EMBASE やコクランが管理する CENTRAL 等のデータベース検索も行うが、本研究では非薬物療法または患者要因・暴露要因について復職アウトカムに対する影響を検討するため、EMBASE の検索は行わなかった。コクランレビューに関しては Medline または PsycINFO で検索でき、また CENTRAL は基本的に RCT を集積したデータベースで本研究が扱う系統的レビューを主として集積するわけではないため、利用しなかった。

文献の言語や出版年は問わず包括的収集を意図した。

検索式（参考資料 1・2）

検索式は可能な限り高感度な検索となるよう、精神疾患、休職、復職などのキーワードのそれぞれに対してのバリエーションを考慮して組み合わせて作成した。またこれに、高感度な系統的レビュー同定のための検索フィルターを組み合わせた

Medline では、系統的レビューを同定する検索フィルターとして、 BMJ Clinical Evidence にて開発されたものを採用した。
(http://clinicalevidence.bmj.com/x/set/static/ebm/learn/665076.html)

また PsycINFO では、系統的レビューを同定する検索フィルターとして、School of Public Health, University of Texas にて開発された Search strategy for relevant systematic reviews in PsycINFO for the OVID online search interface を採用した。
(http://libguides.sph.uth.tmc.edu/search_filters/ovid_psycinfo_filters)

登録基準チェック

二人(渡辺、分担研究者：小川)の独立した研究者がタイトルと抄録で一次チェック、次いで登録条件を満たしたものはフルテキストを二次チェックして該当する系統的レビューを同定した。

一次チェックでは主に対象者と介入・暴露をチェックし、二次チェックではアウトカムの有無をチェックした。

データ抽出、研究の質評価

データ抽出は事前にデザインされた抽出フォームを使って行った。各系統的レビューの研究名、研究内の検索で使用された文献データベース、採用された研究デザイン、系統的レビューで収集した一次研究の限定条件(出版年・言語など)、一次研究に含まれた参加者の診断・就労状況、暴露・介入因子、対照因子、アウトカムの種類と結果、一次研究の質評価方法(コクラン バイアスリスク評価ツールな

ど)を抽出した。

またこれらの系統的レビュー間では一次研究を重複して含む可能性がある。同じ一次研究を含む系統的レビューを複数カウントして結果の統合や解釈を行うことは、その一次研究に他のものの数倍の重みづけをするという誤りが起こる可能性があるため、系統的レビューで含んだ一次研究の識別ができるよう、含んだ一次研究の識別番号(著者名＋出版年)の情報も抽出した。

研究の質評価

系統的レビューといえども、その質は様々である。本研究では各系統的レビューの質はAMSTAR2を利用して評価した。AMSTAR2は2017年に公表された、系統的レビューの質評価を行うツールである。もともと第一版であるAMSTARは2007年に同じ目的で発表されたが、観察研究等の非無作為割り付け対照試験を含む系統的レビューには対応していなかったため、AMSTARに改訂を加えて16項目の尺度として開発された。

16項目は、1）研究疑問の定式化、2）研究実施前のレビュー方法の作成と改訂の報告、3）登録基準の明文化、4）包括的文献検索、5）二人以上の独立研究者による研究選択、6）二人以上の独立研究者によるデータ抽出、7）除外研究の理由付きリスト報告、8）登録された研究の十分な報告、9）適切なバイアスリスク評価、10）登録された研究の資金源報告、11）メタ解析が実施された場合の適切な統計方法、12）メタ解析が実施された場合のバイアスリスク結果の適切な影響評価、13）解釈・考察におけるリスクバイアスリスクの適切な利用、14）結果における異質性の解釈や考察、15）量的統合が行われた場合の出版バイアスの探索、16）系統的レビュー自体の実施資金源と利益相反の開示、を評価し、それぞれあり/なし(Yes/No)または項目に応じて部分的あり(Partial yes)またはメタ解析を実施していないため非該当(No meta-analysis conducted)を決定する。

第一版であるAMSTARは合計得点によって8-11点は高品質、4-7点は中程度品質、3点以下は方法論的に低品質とされたが、AMSTAR2は合計得点によって質を決めることはせず重要ドメインによって全体の質を評価するものである。そのため、本研究でも合計点で示すことはしなかった。

統計解析・結果の報告

全体として、結果が系統的レビュー間で異質性が低いと判断され、なおかつ量的データが示されている場合は、各レビューから個々のRCTのデータを抽出して重複を除いた上でメタアナリシスを実施し、量的統合を行うこととした。ただし

異質性が高く、量的統合が難しいと判断される場合は、各レビュー結果を叙述的・質的にそれを提示することとした。

また前者の場合、可能であればエビデンスの質を考慮した上で、結果を統合した予測モデルを作成することとした。

C. 研究結果

系統的レビューの選択

Ovid Medline（最終検索日：2018 年 2 月 5 日）からは 285 文献が同定された。重複を除くと 245 文献となり、二人の研究者が独立して一次チェック・ディスカッションの結果70文献が二次チェックに進んだ。これらのフルテキストを入手して二次チェックを行ったところ、26 文献が残った。

PsychINFO(最終検索日：2018 年 3 月 12 日)からは 108 文献が同定された。Medline との重複を除いた 66 文献が一次チェックへと進んだ。二人の研究者が独立して一次チェックを行い、ディスカッションの結果 7 文献が二次チェックに進んだ。これらのフルテキストを入手して二次チェックを行い、3 文献が残った。

合計 29 文献について、データ抽出と研究の質評価を行い(表 1)、質を評価した(表 2)。しかし、その過程で種々の理由により 9 個の文献が本研究の登録基準に該当しないことが判明し(表 3)、結果として 20 文献についてデータ抽出と質評価を行った。

精神疾患による休職者の復職予測因子

精神疾患による休職者の復職に関する要因を探索したものとして、2008 年から 2017 年に出版された 5 本の系統的レビューが同定された(表 1)。

Lagerveld 2010[4] は 9 本の長期追跡研究を包含していたが、16 本の横断研究のデータも含み、また AMASTAR-2 チェック項目のほとんどが良い質を満たさない(No)結果で全体的に質が低いと認められ(表 2)、また量的データの十分な報告もなかった。

Cornelius 2011[5] は 7 本の前向きコホート研究を含んでいたが、英語論文に限定して集積し、かつ量的データも報告がないなど質が低かったが、復職困難因子として 50 歳以上、また復職までの時間延長因子として過去の休職歴、腰背部痛の存在、スーパーバイザーの態度などを挙げていた。

Ervasti 2017[6] は文献検索や独立した二人以上の研究者による研究遂行などの点で質がよくないものの 13 本のコホート研究を集積して定量的統合を行っており、うつ病を原因として限定した休職者が復職するリスク比として高年齢(リスク比：RR 0.95)、身体疾患合併(RR

0.80)、重症うつ病(RR 0.86)はそれぞれ復職可能性が低くなり、逆に患者特性として律儀であること(RR 1.06)は復職可能性が高くなることを示した。

Nigatu 2017[7] はうつ病も含む一般精神疾患による休職者を扱った 18 本のコホート研究を集積した。これらの一次研究は、E rvasti 2017 に含まれた研究との重複はほとんどなかった。系統的レビューの事前登録データベースである Prospero に登録するなど（登録番号：CRD42016033088）研究の質は他と比較して比較的高く、メタ解析による量的統合も行っており、高年齢(ハザード比：HR 0.77)、性別が女性であること(HR 0.93)、医療機関受診(HR 0.64)、1 年以内の病休(HR 0.84)、同僚との関係性悪さ(HR 0.97)は有意に休職可能性が下がることを示し、逆に高い自己効力感(HR 1.79)、仕事における高い能力または仕事における低い要求水準(HR 1.08)は有意に復職可能性が上ることを示した。医療機関受診で休職可能性が下がっている理由に関しては、精神疾患の重症度が高いこととの関連が考えられる。

Blank 2008[8] は二人以上の独立研究者がレビュープロセスを行ってはいたが、他の項目を満たさず全体的に質が低く、結果も量的データが示されなかった。

精神疾患による休職者への心理社会的介入

精神疾患による休職者、または精神疾患によって就職できない者に対する心理社会的介入を扱ったものとして、15 本の RCT を集積した系統的レビューが同定された。

そのうち、Crowther 2001[9]、Kinoshita 2013[10]、Suijkerbuijk 2017[11] は重症精神疾患、つまり統合失調症に代表される精神病状態を伴う疾患による就職困難者に対する介入を扱っており、後二者は厳格なコクランの方法論に沿って行われたコクランレビューであった。また Suijkerbuijk 2017[11] は Kinoshita 2013[10] が登録する 14 RCT のほとんどを含めた 48 RCT を登録し、さらに洗練された統計手法であるネットワーク・メタ解析(network meta-analysis：NMA)を用いていたため、この研究についてここでは述べる。Individual placement and support (IPS) model をはじめとした定式化された復職支援プログラムは、職業前トレーニングよりも有意に就職を可能なものとして(RR 2.5)、また段階的就職は通常の精神科ケアよりも有意にのちの就職可能性を向上させた(RR 9.0)。ここでいう IPS とは米国で開発された援助付き雇用のモデルの 1 つであり、①働きたい全ての精神障害者が対象、②就労支援と生活支援をセットにしたサービス、③競争的雇用が目標、④社会保障の相談、⑤迅速

な求職活動、⑥継続的な定着支援、⑦系統的な職場開発、⑧全てのサービスは利用者の希望が優先される、等の特徴を持つものである。

他の 12 本の系統的レビューは、いずれもうつや適応障害等の一般精神疾患による休職者の復職介入試験を扱ったものであった。うち、AMSTAR2 による研究の質評価で質が高いと評価され、また過去の研究に含まれた RCT の登録をしていると考えられる Arends 2012[3]、Nieuwenhuijsen 2014[2]、van Vilsteren 2015[12]、Doki 2015[13]、Nigatu 2016[14] について、ここではまとめる。

Arends 2012[3] は適応障害による休職者を扱った 10 RCT を登録し、認知行動療法(cognitive behavioral therapy：CBT)と問題解決療法(problem solving therapy: PST)がこれらを行わない場合と比較した。CBT、PST ともそれぞれ 2 RCT からのエビデンスが得られたが、PST では部分復職までの期間が 17 日短縮したものの、完全復職では有意差がなく、また CBT でも有意差が見られなかった。

Nieuwenhuijsen 2014[2] はうつ病やうつ状態による休職者を扱った 23 RCT を含み、仕事指向的な介入はそれを行わない場合と比較して小程度の治療効果があることを示し(標準化公開サイズ：SMD -0.4)、また暴露を基本とした復職プログラムは小から中程度(SMD 0.45)、電話やオンライン CBT は小程度(SMD 0.2)の効果があることを示した。しかし、2RCT を集積したメタ解析では、有酸素運動では統計学的有意差は見られなかった。

van Vilsteren 2015[12]、Doki 2015[13] は Arends と登録した RCT が重複しているため、ここでは割愛する。

Nigatu 2016[14] はコクランレビューではないものの Prospero に登録するなど（登録番号：CRD42016033092)かなり質はよく、不安障害以外の精神疾患による休職者を集めた 16 RCT を集積した。CBT や PST、コーピング戦略や暴露等の精神療法的介入を一つにまとめてメタ解析を行ったところ、復職率は有意差がなく、完全復職までの期間は介入群 151 日に対して対照群 162 日、加重平均差 13 日と有意に短縮を認めた。

オーバービュー：メタ解析・予測モデル

また総括として、各系統的レビュー結果を集積したメタ解析や予測モデル作成を計画していたが、各系統的レビューで採用した一次研究の研究デザインや方法論のみならず、対象とした疾患、介入・暴露方法、アウトカム定義にも大きな異質性を認め、総括的メタ解析や予測モデル作成を行うことはできなかった。

表１：登録基準に該当した研究の特性

Study name	Included databases	Included study design	Limited selection of studies (if any)	Diagnosis and status required for inclusion among participants	Tool for assessment of methodological quality	Collected variables for exposure or intervention	Comparison group	Outcomes
Investigating factors associated with RTW								
Lagerveld 2010[4]	PubMed, PsycINFO, Scopus	No restrictions to study design, and 9 longitudinal observation (Dewa 2002, Laitinen-Krispijn 2000, Rystala 2007, Simon 2000, Druss 2001, Lerner 2004)Sanderson 2007, Adler 2006, Michon 2008) and 16 cross-sectional studies were included	Studies written in English only	Depression diagnosed by an expert or based on a well-defined cut-off score for depressive symptoms	10-item checklist designed of the study, covering study population, data analysis, data presentation and statistical power	Gender, age, education, marital status, history of sick leave, low self-esteem, hopeless about future, low social functioning, higher neuroticism, more external locus of control, alcoholism, social adjustment, social support, living are, income level	Work participation including absence from work, and work functioning including work limitations	Mixture reports of longitudinal and cross sectional data led to speculate if a factor influenced on work disability and work performance, thus have to say inconclusive.
Cornelius 2011[5]	Medline, PsycINFO, EMBASE, CINAHL, Business Source Premier	7 longitudinal cohort studies with follow-up of 1 year or more (N=7116, Brennninkmeijer 2008, Engstorm 2007, Nieuwenhuijsen 2003, Nieuwenhuijsen 2004, Nieuwenhuijsen 2006, Nieuwenhuijsen 2010, Vaez 2007)	Published only from 1990 to 2009 and limited to papers in English language only	Sick leave or receiving disability benefit due to poor mental health at baseline	According to 16-item quality assessment for prognostic studies recommended by Altman, all studies were deemed high at scores ranging from 13 to 16 points.	Various baseline characteristics		No quantitative data, but older age (>50) was concluded to be associated with longer time to RTW in strong evidence. In limited evidence, personal factors (gender, education, history of previous sickness absence, negative recovery expectation, socioeconomic status), health related (stress-related and shoulder/back pain, depression/ anxiety disorder) and external job-related factors(unemployment, quality and continuity of occupational care, supervisor behavior) were concluded to be associated with

								longer time to RTW.
Ervasti 2017[6]	PubMed, Embase	13 cohort studies (N=36459, Dewa 2003, Nieuwenhuijsen 2004, Koopmans 2008, Kronstrom 2011, Hogelund 2012, Hees 2012, Nordenskjold 2013, Ogawa 2013, Ebrahim 2013, Vemer 2013, Lammers 2015, FPS 2016, FPS 2016)	Not stated	Unable to work due to depression. Sick leave mainly due to common mental disorders other than depression were excluded	Cochrane RoB took, additional 7-item checklist for cohort studies, GRADE	Various baseline characteristics		The pooled estimates were derived from 2 to 5 studies RTW: older age (RR 0.95 (0.84-0.87); somatic comorbidity RR 0.80 (0.77-0.83), psychiatric comorbidity, RR 0.86 (0.83-0.88); more severe depression (RR = 0.96, 95% CI 0.94-0.98) ; personality trait conscientiousness with higher, RR 1.06 (1.02-1.10) While older age and clinical factors predicted slower return, significant heterogeneity was observed between the studies.
Nigatu 2017[7] Prospero registered (CRD42016033088)	Medline, PsycINFO, EMBASE, SocINDEX, Open Grey, Human resource and management databases	18 cohort studies (N=24664, Laaksonen 2015, Lammers 2006, Hees 2002, Brouwers 2009, Prang 216, Gustafsson 2013, de Rijk 2009, Oyeflaten 2014, de Rijk 2008, Landstad 2009, Ekberg 2015, Wahlin 2012, Netterstrom 2010, Netterstrom 2015, Virtanen 2011, D'Amato 2010, van Beurden 2015, Audhoe 2012, Nieuwenhuijsen 2004)	Not stated	Absent from work due to common mental disorders including depressive disorders, anxiety disorders, OCD, PTSD, or adjustment disorders	Newcastle-Ottawa Scale. Full score was 10, and all the studies had the score between 7 and 9.	Various baseline characteristics		Full RTW: age (older), HR 0.77 (0.65, 0.88); sex (women), HR 0.93 (0.74, 1.12); education (high), HR 1.04 (0.96, 1.13); Household persons (>1), HR 0.96, (0.73, 1.19); depression (yes), HR 0.99 (0.84, 1.15); self-rated health (bad), 0.80 (0.54, 1.07); contact with medical specialists (yes), 0.64 (0.49, 0.80); RTW self-efficacy (high), HR 1.79,(1.24, 2.33); prior sickness absence (yes, <365 days), 0.84 (0.42, 1.26); high work ability or low work demands (1.08 (1.06, 1.11); relationship with colleagues (good), 0.97 (0.95, 0.99)
Blank 2008[8]	Medline, CINAHL, PsycINFO, EMBASE, OSH-ROM, Assia, Web	9 cohort studies (Barmby 2001, Stansfeld 1999, Jenkins 1985, Nieuwenhuijsen 2004, Laitinen-Krispijin	Published only from 1985 to 2005, and in English	Not clearly stated, but people suffering episodes of poor mental health and in sick leave	Not stated	Various baseline characteristics		Preventing factors from RTW were discussed, but no quantitative data were provided. Some factors introduced were: work related factors (company sick pay, lower word grade,

	of Science, CINAHL, British Nursing Index, ENB Healthcare, Health Information Management Consortium, Business Source Premier, International Bibliography of Social Sciences	2000, Ginexi 1999, Russell 1995, Young 1995, Parkes 1996), 5 retrospective survey (Salkever 2003, Salkever 2000, Dewa 2003, Dewa 2002, Semmer 1987)						features of benefit plan, risk of unemployment, injury at work, super visor / employer communication, not RTW in first 505 days, high job stressors, not attempted to return to work), high risk behaviors (alcohol intake, drug dependence, overweight, smoker), social status (widowed, divorced, single (male), age (older)), medical factors (psychiatric illness, minor psychiatric disorder, simple phobia, severity of symptoms)
Investigating interventions associated with RTW								
Crowther 2001[9]	Medline, CINAHL, PsycLIT, EMBASE	11 RCT (N=1944, Bond 1995, Drake 1996, Drake 1999, Gervey 1994, McFarlane 2000, Beard 1963, Dincin 1982, Friffiths 1974, Okpaku 1997, Wolkon 1971., Chandler 1996)	Not stated	Severe mental illness	Two items, allocation concealment and independence of evaluators rom those providing the intervention	Supported employment	Prevocational training or supported care	Obtaining competitive employment (supported employment vs prevocational training at 6 months, MA of 3 RCTs, RR 0.74 (0.67, 0.82)
Noordik 2010[15]	Medline, CINAHL, PsycINFO, EMBASE	RCT (N=253Greist 2002, Marks 1988, Cobb 1980), CCT (N=267, Aigner 2004, Foa 1984, Salyards 2005)	Not stated	Anxiety disorders or anxiety complaints, including OCD and PTSD	Cochrane RoB tool and GRADE criteria. 1 OCD and 1 PTSD study had a low RoB, but the others were high or unclear.	Exposure in vivo.	Another intervention including anxiolytics or antidepressants.	Exposure in vivo compared with SSRI (“Work-related effect”, fixed effect of SMD -0.72 (-1.15, -0.28), MA of 2 RCD for OCD) Exposure in vivo compared via a computer or a clinician with self-relaxation ("Work-related effect", SMD 0.35 (0.08, 0.79), 0.72 (0.28, 12

								.17), respectively, 1 RCD for OCD)
Audhoe 2010[16]	Medline, PsycINFO, EMBASE	2 RCT (Vinokur 2000, Vuori 2002), 1 NRCT (Creed 1998), 2 longitudinal studies (Vuori 2005, Vuori 1999)	Published from 1990 to 2008, and in English, German, Dutch, and French	Unemployed and with mental distress complaints or distress	Downs and Black instrument, 27-item checklist	Interventions aiming at RTW, work resumption, job application and/or improved functioning. Work preparation program, enhancing the sense of mastery through acquisition of job search and problem solving skills, job-search training	Booklet, waiting list, literature package	Only p-values were presented, but seemingly RCTs showed JOBS II program (enhancing the sense of mastery through acquisition of job search and problem solving skills) better than the control conditions
Pomaki 2012[17]	Medline, CINAHL, PsycINFO, EMBASE, ISI Web of Science	4 RCT (N=1849, Fleten 2006, Rebergen 2009,Sogaard 2009, Wang 2007), 2 matched control design, (N=185, Grossi 2009, Lander 2009), 1 historical control (N=124, Dewa 2009)	Published only from 2007 to 2009, and in English, French, Dutch or German	Currently absent from work or struggling at work due to common mental health conditions	29-item criteria adapted from Downs and Black, the Effective Public Health Practice Project and Franche et al.	Collaborative mental health care (1 historical cohort), receiving written information about RTW (1 RCT), 3-month group rehabilitation program (1 matched control design), activating intervention (1 RCT), psychological education (1 RCT), assessment and feedback by a psychiatrist (1 RCT), structured telephone-based intervention (1RCT)	Using a control group was declared mandatory.	Collaborative mental health care (RTW at 1 year, 85% vs 63%), receiving written information about RTW (RTW, shorter in the intervention than in the control by 36.6 days), 3-month group rehabilitation program (RTW, no significant difference up to 5 years), activating intervention (RTW, no significant difference), guideline-based care by physicians (partial RTW, 69% vs 54%), assessment and feedback by a psychiatrist to GP and worker (RTW, 61% vs 65%, negative finding), structured telephone-based intervention (job retention at 12 months, 93% vs 88%)
Stergiopoulos 2011[18]	Medline, PsycINFO, EMBASE,	3 RCT (N=86, Gerson 2000, Hogberg 2006, Hogberg 2007), 3 pre-	English or French	Clinical diagnosis of PTSD. PTSD not related to	13-item checklist modifying	Combination of psychodynamic therapy and CBT (1	Mentioned as mandatory quality	Brief eclectic psychotherapy (RTW in 3 months, 86%), EMDR (RTW no usable data),

	ISI Web of Science	post design (N=126, Gunert 1989, Gunert 1992, Weis 1999)		workplace, or combat-related PTSD was excluded.	Lagerveld's checklist. Rated as excellent (2 studies) and good (4)	RCT), EMDR (2 RCT), on-site work evaluations (1 pre-post), graded work exposure (1 pre-post), prolonged imaginal exposure (pre-post)	criterion, but no results from control groups were reported	on-site work evaluations (RTW in 12 months, 87%), graded work exposure (RTW in 6 months, 88%), prolonged imaginal exposure (RTW in 6 months, 83%)
Furlan 2012[19]	Medline, PsycINFO, EMBASE, CINAHL, CENTRAL, Business Source Premier	10 RCTs (Blonk 2006, Smith 2002, Revergen 2009, van der Feltz-Cornelis 2010, Schene 2009, Korgh 2009, Knekt 2008, Lo Sasso 2006, Wang 2007, Schoenbaum 2001), 2 cohort studies (Dewa 2009, Kawakami 1997)	Not stated	Current or remitted depression. Studies that included participants with other mental disorders were included only if 50% or more had depression. Among 12 studies, 4 foucused on those with disability leave or sick leave, and 4 included a mix or those working, on sick leave, or unemployed	Using GRADE, and "very low" for all outcomes identified	Interventions in workplace-based or employer facilitated. RCTs employed CBT, psychodynamic psychotherapy, enhanced care, quality improvement program, occupational therapy, strength training (aerobic training, relaxation training). NRCTs employed collaborative mental health program and worksite stress reduction program.	Not stated	As the GRADE was very low for all outcomes, no synthesis was conducted and no intervention was recommended.
Arends 2012[3] Cochrane review	Medline, PsycINFO, EMBASE, CINAHL, CENTRAL, CCDANCTR, ISI Web of Science, ICTRP, ClicalTrials.gov	10 RCTs (N=1546, Bakker 2007, Blonk labour expert 2006, Blonk psychologists2006, Brouwers 2006, de Vente group 2008, de Vente individual 2008, Rebergen 2009, Stenlund 2009, van der Klink 2003, van	Not stated	Absent from work due to adjustment disorder	Cochrane RoB tool	CBT, problem solving	No treatment	Duration to partial and full RTW (CBT, MA of 2 moderate RoB RCTs, days until partial RTW at one year follow-up, -8.78 (-23.26, 5.71); full RTW, -35.7 (-113.1, 41.7); PST, MA of 2 RCTs, partial RTW, -17.0 (-26.5, -7.5), full RTW -17.7 (-37.4, 1.9)

		Oostrom 2010, Willer 2011)						
Hamberg-van Reenen 2012[20]	Medline, SCOPUS, NHS-EED-HTA, PsycINFO	7 RCT (N=1384, Lo Sasso 2006, Bittman 2003, van Oostrom 2010, Brouwers 2007, Rebergen 2009, Schene 2007, Uegaki 2010), 1 cohort study (N=38, Vogt 2004), 1 simulation study (N=598, Leon 2002), 1 other (N not stated, Wang 2004)	Published only from 2000 to 2011, and in English	Only mentioned as distress, depression, mental health problems	Consensus Health Economic Criteria list (CHEC-list)	Various. Interventions included in RCTs were enhanced treatment, recreational music making, stress management program, RTW-intervention, problem solving, guideline-based care, occupational therapy, minimal intervention by GPs	Usual care	Only studies with economic evaluations were included. Only 6 studies focused on RTW. No evidence favorable cost effectiveness of RTW interventions among 5 studies. Schene 2007 showed the number of days until work resumption was lower at 207 days compared to the control at 299 days.
Kinoshita 2013[10] Cochrane review	Cochrane Schizophrenia Group Trials Register	14 RCTs (N=2265, Bond 2007, Burns 2007, Drake 1996, Drake 1999, Gold 2006, Howard 2010, Killackey 2008, Latimer 2006, Lehman 2002, Macias 2006, Mueser 2004, Tsang 2009, Twanley 2008, Wong 2008)	Not stated	Those with severe mental illness	Cochrane RoB tool	Supported employment including individual placement and support model	No treatment	Days in competitive employment over one year, supported employment (MA of 7 RCTs, RR 3.24 (2.17, 4.82), very low quality of evidence)
Nieuwenhuijsen 2014[2] Cochrane review	Medline, PsycINFO, EMBASE, CINAHL, CENTRAL	23 RCTs (N=5996)	Not stated	Some degree of absent from work due to a major depressive disorder or a high level of depressive symptoms	Cochrane RoB tool	Work directed interventions, antidepressant medications, CBT, telephone or inline CBT, enhanced primary care, exercise	No treatment	reduced sickness absence, work-directed intervention, SMD -0.40 (-0.66 to -0.14; 3 studies, moderate quality); exposure-based return to work program, SMD 0.45; -0.00 to 0.91; 1 study), telephone or online CBT (SMD -0.23; 95% CI -0.45 to -0.01, moderate quality), a structured telephone outreach and care management

								program was more effective in reducing sickness absence than usual care (SMD - 0.21; 95% CI -0.37 to -0.05), aerobic exercise was no more effective in reducing sickness absence than relaxation or stretching (SMD -0.06; 95% CI -0.36 to 0.24, moderate quality, 2 RCTs).
van Vilsteren 2015[12] Cochrane review	Medline, PsycINFO, EMBASE, CENTRAL, Cochrane Work Trials Register	5 RCTs for sick leave due to mental health problems (Blonk 2006, Hees 2012, Noordik 2013, van Oostrom 2010, Vlasveld 2012)	Not stated	Sick leave due to mental health problems	Cochrane RoB tool	Workplace interventions	Usual care	Significant improvement in time until first RTW (HR 2.64, 95% CI 1.41 to 4.95), but no considerable reduction in lasting RTW (HR 0.79, 95% CI 0.54 to 1.17).
Dewa 2015[21]	Medline, PsycINFO, Econlit, Web of Science	6 RCT (N=975, van der Klink 2003, Brouwers 2006, Hees 2010, Nystyuen 2003, Revergen 2009, Vlasvled 2013)	Published only from 2002 to 2014 and limited to papers in English language only	Workers who were certified sickness absences related to mental disorders	Cochrane RoB tool	Problem solving skills	Care as usual, or treatment as usual	RTW (partial or full RTW at 3 months in ven der Klink 2003, 98% vs 87%; full RTW at 3 months in Brouwers 2006, 37% vs 0%; RTW in good health in Hees 2010, 28% vs 24%) Sickness leave duration (RTW in days in van der Klink 2003, 36 vs 53; in Brouwers 2006, 106 vs 121)
Doki 2015[13]	Pubmed, PsycINFO	6 RCT for subgroup 1 (N=846, Brouwers 2007, Rebergen 2009, van der Feltz-¥cornilis 2010, Willert 2011, van Oostrom 2010, Vlasveld 2013)	Published only from 2004 to 2014	Workers with sick leave (subgroup 1) due to mental illness	GRADE	Psychological interventions by occupational health services staff, including problem-solving, Dutch guideline based care, stepwise communication, 8 CBT sessions	Care as usual	Mean difference of RTW (days), MA of 6 RCTs, -10.48 (-22.64, 1.69)
Nigatu 2016[14] Prospero	Medline, PsycINFO, EMBASE, SocINDEX,	16 RCT (N=3345, Bakker 2010, Brouwers 2006, Hees 2013, Martin 2013,	Published only from 1995 to 2016 and	Common Absent from work due to mental disorders (no anxiety	Cochrane RoB tool. High risk of bias was found in	Problem solving, CBT, coping strategies, exposer, occupational	Care as usual in all studies.	RTW, defined as proportion of employees who returned to work after the intervention, was reported 65 and 60% in the

registered (CRD42016033092)	Open GreyHuman resource and management databases	Lagerveld 2012, Netterstrom 2013, Noordik 2013, Nystuen 2006, pdersen 2015, Revergen 2009, Sogaard 2009, ven der Feltz-Cornelis 2010, van der Klink 2003, van Oostrome 2010, Vlasveld 2013, volker 2015))	limited to papers in English language only	disorders)	blinding of participants and personnel. 7 out of 16had unclear risk regarding blinding of the outcomes assessment.	therapy, psychoeducation, diagnosis and consultation and referral		intervention and control groups, respectively (Random effect RR 1.05 (0.97, 1.12), P=0.21, MA of 16 RCTs) Time or duration until full RTW was 151 and 165 days, respectively (Fixed MD -13.38 (-24.07, -2.69), P=0.01, MA of 7 RCTs)
Suijkerbuijk 2017[11] NMA, Cochrane reveiw	Medline, PsycINFO, EMBASE, CINAHL, CENTRAL, CINAHL	48 RCTs (N=8743)	Not stated	Severe mental illness	Cochrane RoB tool	Supported employment, augmented supported employment, prevocational training, transitional employment.	Psychiatric care	Supported employment was more effective than prevocational training (RR 2.52, 95% CI 1.21 to 5.24) and transitional employment (RR 3.49, 95% CI 1.77 to 6.89) and prevocational training was more effective than psychiatric care only (RR 8.96, 95% CI 1.77 to 45.51) in obtaining competitive employment. For the long-term follow-up direct meta-analysis, we could include 22 trials (N = 5233). Augmented supported employment (RR 4.32, 95% CI 1.49 to 12.48), supported employment (RR 1.51, 95% CI 1.36 to 1.68) and prevocational training (RR 2.19, 95% CI 1.07 to 4.46) were more effective than psychiatric care only.

表 2：登録基準に該当した研究の質－AMSTAR-2 を用いて評価

Study name	AMATAR-2															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Investigating factors associated with RTW																
Lagerveld 2010	N	N	N	N	N	N	N	N	N	N	NA	NA	N	N	NA	Y
Cornelius 2011	N	N	Y	N	Y	N	N	PY	Y	N	NA	NA	Y	N	NA	N
Ervasti 2017	N	N	Y	N	N	N	N	N	Y	N	Y	Y	Y	Y	N	Y
Nigatu 2017	N	Y	Y	Y	Y	Y	N	N	Y	N	Y	N	N	Y	N	Y
Blank 2008	N	N	N	N	Y	Y	N	N	N	N	NA	NA	N	N	NA	Y
Investigating interventions associated with RTW																
Crowther 2001	Y	PY	Y	Y	Y	Y	N	Y	PY	N	Y	N	N	Y	N	Y
Noordik 2010	N	Y	Y	Y	Y	Y	N	Y	Y	N	N	N	N	Y	N	Y
Audhoe 2010	Y	N	Y	N	Y	Y	N	Y	PY	N	NA	NA	N	N	NA	N
Pomaki 2012	Y	Y	Y	N	Y	Y	N	Y	Y	N	NA	NA	N	Y	NA	Y
Stergiopoulos 2011	N	N	N	N	Y	Y	N	N	N	N	NA	NA	N	N	NA	Y
Furlan 2012	Y	N	Y	Y	Y	Y	N	Y	Y	N	NA	NA	Y	N	NA	Y
Arends 2012	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Hamberg-van Reenen 2012	N	N	N	N	N	Y	N	Y	PY	N	NA	NA	Y	N	NA	Y
Kinoshita 2013	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Nieuwenhuijsen 2014	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
van Vilsteren 2015	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dewa 2015	Y	Y	Y	N	Y	N	N	PY	PY	N	NA	NA	N	N	N	Y
Doki 2015	Y	N	Y	N	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	N
Nigatu 2016	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	N	Y	N	Y
Suijkerbuijk 2017	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

16 項目は以下のとおり。

1）研究疑問の定式化、2）研究実施前のレビュー方法の作成と改訂の報告、3）登録基準の明文化、4）包括的文献検索、5）二人以上の独立研究者による研究選択、6）二人以上の独立研究者によるデータ抽出、7）除外研究の理由付きリスト報告、8）登録された研究の十分な報告、9）適切なバイアスリスク評価、10）登録された研究の資金源報告、11）メタ解析が実施された場合の適切な統計方法、12）メタ解析が実施された場合のバイアスリスク結果の適切な影響評価、13）解釈・考察におけるリスクバイアスリスクの適切な利用、14）結果における異質性の解釈や考察、15）量的統合が行われた場合の出版バイアスの探索、16）系統的レビュー自体の実施資金源と利益相反の開示

表 3：除外された研究

Study name	Reasons for exclusion
Killackey 2006[22]	Seems not a systematic review because of no description about methodology
van Til 2013[23]	Veteran only and generalizability was in doubt.
Marshall 2014[24]	Including not only cohort studies but also systematic reviews.
Ebrahim 2012[25]	Compared depressive symptomatology change between those with receiving disability benefit and those without
Ebrahim 2014[26]	Overview of systematic reviews.
Dewa 2014[27]	No study design was specified in the article.
Bouvet 2015[28]	Written in French
Joyce 2016[29]	Meta-review of systematic reviews
Matthewson 2015[30]	Study design was not specified in the method and systematic reviews and qualitative studies were include in the results

D. 考察

20 本の系統的レビューを同定し、オーバービューを行った。各系統的レビューの研究の質は、総じて出版年度が最近になるほど改善し、また特に介入を扱った研究ではコクランレビューでは質が高かった。

まず精神疾患による休職者の復職可能性の指標として、各系統的レビューに共通して見られた知見では、高年齢、最近の疾患による休職歴、身体疾患合併等が低い復職可能性と関連し、逆に高い自己効力感や仕事における高い能力または仕事における低い要求水準は復職可能性が上ることを示した。

一方、介入に関する系統的レビューでは、IPS モデル等の復職支援プログラムは、統合失調症をはじめとした重症精神疾患者の就職可能性を向上させ、また段階的就職は通常の精神科ケアよりも有意にのちの就職可能性を向上さる。

しかしながら、一般精神疾患では CBT では復職アウトカムを改善させるという強いエビデンスがなく、PST では部分復職までの期間が 2 週程度短縮したものの、完全復職では有意差がなかった。また限られたエビデンスからは、有酸素運動の効果は見られず、また食事や喫煙・飲酒等について取り扱った系統的レビューはなかった。これらから言えることとして、一般的な生活指導に関するエビデンスはなく、また認知や一般的行動の変容を促す CBT だけでは、復職アウトカムに結びつくことはあまり期待できず、問題の同定や解決方法模索などの仕事に特異的な介入を行う方が良いのかもしれない。

復職を主要アウトカムにするためには、疾患特異的な治療ではなく個々の患者の状況に合わせ、現実的な問題を取り扱う介入方法の開発が必要であろう。

また今回は各系統的レビューに異質性が強かったため、総括的メタ解析や予測モデル作成を行うことはできなかった。おそらく、今後も一次研究を集積してこれらを計画したとしても、その研究の多様性からまとめていくのは難しいかもしれない。DPC による患者の病名・年齢・処方等の基本的データだけではなく、カルテ情報を集積して非薬物療法の実施状況や就労状態、疾患重症度を盛り込んだいわゆるビッグデータとしてまとめることができれば、わが国における精神疾患による休職者の復職指標・介入方法を同定することができる可能性があり、これらの整備が必要であろう。

E. 結論

一般精神疾患では、疾患特異的な治療ではなく個々の患者の休職時状況に合わせ、現実的な問題を取り扱う介入方法の開発が必要であろう。

また研究面では、小さな一次研究を集積するよりも、ビッグデータなどリッチなデータを用いた予測モデル・指標作成が望まれると考えた。

参考文献

1. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry.* 2006;163(1):28-40.
2. Nieuwenhuijsen K, Faber B, Verbeek JH, et al. Interventions to improve return to work in depressed people. *Cochrane Database Syst Rev.* 2014(12):CD006237.
3. Arends I, Bruinvels DJ, Rebergen DS, et al. Interventions to facilitate return to work in adults with adjustment disorders. *Cochrane Database Syst Rev.* 2012;12:CD006389.
4. Lagerveld SE, Bultmann U, Franche RL, et al. Factors associated with work participation and work functioning in depressed workers: a systematic review. *Journal of Occupational Rehabilitation.* 2010;20(3):275-292.
5. Cornelius LR, van der Klink JJL, Groothoff JW, Brouwer S. Prognostic factors of long term disability due to mental disorders: a systematic review. *Journal of Occupational Rehabilitation.* 2011;21(2):259-274.
6. Ervasti J, Joensuu M, Pentti J, et al. Prognostic factors for return to work after depression-related work disability: A systematic review and meta-analysis. *J Psychiatr Res.* 2017;95:28-36.
7. Nigatu YT, Liu Y, Uppal M, et al. Prognostic factors for return to work of employees with common mental disorders: a meta-analysis of cohort studies. *Soc Psychiatry Psychiatr Epidemiol.* 2017;52(10):1205-1215.
8. Blank L, Peters J, Pickvance S, Wilford J, Macdonald E. A systematic review of the factors which predict return to work for people suffering episodes of poor mental health. *Journal of Occupational Rehabilitation.* 2008;18(1):27-34.
9. Crowther R, Marshall M, Bond G, Huxley P. Vocational rehabilitation for people with severe mental illness. *Cochrane Database Syst Rev.* 2001(2):CD003080.
10. Kinoshita Y, Furukawa TA, Kinoshita K, et al. Supported employment for adults with severe mental illness. *Cochrane Database Syst Rev.* 2013(9):CD008297.
11. Suijkerbuijk YB, Schaafsma FG, van Mechelen JC, Ojajarvi A, Corbiere M, Anema JR. Interventions for obtaining and maintaining employment in adults with severe mental illness, a network meta-analysis. *Cochrane Database Syst Rev.* 2017;9:CD011867.
12. van Vilsteren M, van Oostrom SH, de Vet HCW, Franche R-L, Boot CRL, Anema JR. Workplace interventions to prevent work disability in workers on sick leave. *Cochrane Database Syst Rev.*

2015(10):CD006955.

13. Doki S, Sasahara S, Matsuzaki I. Psychological approach of occupational health service to sick leave due to mental problems: a systematic review and meta-analysis. *Int Arch Occup Environ Health.* 2015;88(6):659-667.
14. Nigatu YT, Liu Y, Uppal M, et al. Interventions for enhancing return to work in individuals with a common mental illness: systematic review and meta-analysis of randomized controlled trials. *Psychol Med.* 2016;46(16):3263-3274.
15. Noordik E, van der Klink JJL, Klingen EF, Nieuwenhuijsen K, van Dijk FJH. Exposure-in-vivo containing interventions to improve work functioning of workers with anxiety disorder: a systematic review. *BMC Public Health.* 2010;10:598.
16. Audhoe SS, Hoving JL, Sluiter JK, Frings-Dresen MHW. Vocational interventions for unemployed: effects on work participation and mental distress. A systematic review. *Journal of Occupational Rehabilitation.* 2010;20(1):1-13.
17. Pomaki G, Franche R-L, Murray E, Khushrushahi N, Lampinen TM. Workplace-based work disability prevention interventions for workers with common mental health conditions: a review of the literature. *Journal of Occupational Rehabilitation.* 2012;22(2):182-195.
18. Stergiopoulos E, Cimo A, Cheng C, Bonato S, Dewa CS. Interventions to improve work outcomes in work-related PTSD: a systematic review. *BMC Public Health.* 2011;11:838.
19. Furlan AD, Gnam WH, Carnide N, et al. Systematic review of intervention practices for depression in the workplace. *Journal of Occupational Rehabilitation.* 2012;22(3):312-321.
20. Hamberg-van Reenen HH, Proper KI, van den Berg M. Worksite mental health interventions: a systematic review of economic evaluations. *Occupational & Environmental Medicine.* 2012;69(11):837-845.
21. Dewa CS, Loong D, Bonato S, Joosen MCW. The effectiveness of return-to-work interventions that incorporate work-focused problem-solving skills for workers with sickness absences related to mental disorders: a systematic literature review. *BMJ Open.* 2015;5(6):e007122.
22. Killackey EJ, Jackson HJ, Gleeson J, Hickie IB, McGorry PD. Exciting career opportunity beckons! Early intervention and vocational rehabilitation in first-episode psychosis: employing cautious optimism. *Aust N Z J Psychiatry.* 2006;40(11-12):951-962.

23. Van Til L, Fikretoglu D, Pranger T, et al. Work reintegration for veterans with mental disorders: a systematic literature review to inform research. *Phys Ther.* 2013;93(9):1163-1174.
24. Marshall T, Goldberg RW, Braude L, et al. Supported employment: assessing the evidence. *Psychiatr Serv.* 2014;65(1):16-23.
25. Ebrahim S, Montoya L, Truong W, et al. Effectiveness of cognitive behavioral therapy for depression in patients receiving disability benefits: a systematic review and individual patient data meta-analysis. *PLoS One.* 2012;7(11):e50202.
26. Ebrahim S. Psychotherapy for depression in claimants receiving wage replacement benefits: review of the evidence. *J Insur Med.* 2014;44(1):53-57.
27. Dewa CS, Loong D, Bonato S, Hees H. Incidence rates of sickness absence related to mental disorders: a systematic literature review. *BMC Public Health.* 2014;14:205.
28. Bouvet C, Battin C, Le Roy-Hatala C. [The Clubhouse model for people with severe mental illnesses: Literature review and French experiment]. *Encephale.* 2015;41(6):477-486.
29. Joyce S, Modini M, Christensen H, et al. Workplace interventions for common mental disorders: a systematic meta-review. *Psychol Med.* 2016;46(4):683-697.
30. Matthewson M, Langworthy J, Higgins D. Psychological Predictors of Vocational Success for People with Psychotic Illness. *Australian Journal of Rehabilitation Counselling.* 2015;21(1):29-64.

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

1. Watanabe N, Matsuoka Y, Kumachi M, Hamazaki K, Horikoshi M, Furukawa TA. Omega-3 fatty acids for a better mental state in working populations - Happy Nurse Project: A 52-week randomized controlled trial. *J Psychiatr Res.* 2018;102:72-80.
2. Jike M, Itani O, Watanabe N, Buysse DJ, Kaneita Y. Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression. Sleep Med Rev. 2018;39:25-36
3. Yamazaki H, So R, Matsuoka K, Kobayashi T, Shinzaki S, Matsuura M, Okabayashi S, Kataoka Y, Tsujimoto Y, Furukawa TA, Watanabe N. Certolizumab pegol for

induction of remission in Crohn's disease: protocol. *Cochrane Database Syst Rev.* 2018;2018.

4. Nishizaki D, Ganeko R, Hoshino N, Hida K, Obama K, A Furukawa T, Sakai Y, Watanabe N. Roux-en-Y versus Billroth-I reconstruction after distal gastrectomy for gastric cancer: protocol. *Cochrane Database Syst Rev.* 2018;2018.
5. Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, Ogawa S, Sato H, Kitagawa N, Shinagawa Y, Ikeda Y, Imai H, Tajika A, Ogawa Y, Akechi T, Yamada M, Shimodera S, Watanabe N, Inagaki M, Hasegawa A. Cognitive and Behavioral Skills Exercises Completed by Patients with Major Depression During Smartphone Cognitive Behavioral Therapy: Secondary Analysis of a Randomized Controlled Trial. *JMIR Ment Health.* 2018;5(1):e4.
6. Tateno M, Jovanovic N, Beezhold J, Uehara-Aoyama K, Umene-Nakano W, Nakamae T, Uchida N, Hashimoto N, Kikuchi S, Wake Y, Fujisawa D, Ikari K, Otsuka K, Takahashi K, Okugawa G, Watanabe N, Shirasaka T, Kato TA. Suicidal ideation and burnout among psychiatric trainees in Japan. *Early Interv Psychiatry.* 2017.
7. Kusaka K, Shinohara K, Tada M, Watanabe N, Furukawa TA. Concerns About Selective Outcome Reporting. J Clin Oncol. 2017;35(6):688.
8. Itani O, Jike M, Watanabe N, Kaneita Y. Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. Sleep Med. 2017;32:246-256.

2. 学会発表

1. 渡辺範雄, 地家真紀, 井谷修, 兼板佳孝. 長時間睡眠と健康アウトカム： 系統的レビューとメタ回帰分析. Paper presented at: 第76回日本公衆衛生学会総会; 11月1日, 2017; 鹿児島.
2. Watanabe N, Jike M, Itani O, Buysse D, Kaneita Y: Long Sleep Duration and Health Outcomes: A Systematic Review, Meta-Analysis and Meta-Regression. the 2017 Annual Meeting of the American Psychiatric Association, San Diego, USA, 2017.
3. 渡辺範雄，地家真紀，井谷修，兼板佳孝：長時間睡眠と健康アウトカム　系統的レビューとメタ回帰分析．第76回日本公衆衛生学会総会，鹿児島，2017
4. Watanabe N, Jike M, Itani O, Buysse D, Kaneita Y: Long Sleep Duration and Health Outcomes: A Systematic Review, Meta-Analysis and Meta-Regression. World Sleep 2017, Prague, Czech Republic, 2017

H. 知的財産権の出願・登録状況

なし

1. 特許取得

なし

2. 実用新案登録

3. その他

なし

参考資料 1：Ovid Medline 検索式

1. (review or review,tutorial or review, academic).pt.
2. (medline or medlars or embase or pubmed or cochrane).tw,sh.
3. (scisearch or psychinfo or psycinfo).tw,sh.
4. (psychlit or psyclit).tw,sh.
5. cinahl.tw,sh.
6. ((hand adj2 search$) or (manual$ adj2 search$)).tw,sh.
7. (electronic database$ or bibliographic database$ or computeri?ed database$ or online database$).tw,sh.
8. (pooling or pooled or mantel haenszel).tw,sh.
9. (peto or dersimonian or der simonian or fixed effect).tw,sh.
10. (retraction of publication or retracted publication).pt.
11. or/2-10
12. 1 and 11
13. meta-analysis.pt.
14. meta-analysis.sh.
15. (meta-analys$ or meta analys$ or metaanalys$).tw,sh.
16. (systematic$ adj5 review$).tw,sh.
17. (systematic$ adj5 overview$).tw,sh.
18. (quantitativ$ adj5 review$).tw,sh.
19. (quantitativ$ adj5 overview$).tw,sh.
20. (quantitativ$ adj5 synthesis$).tw,sh.
21. (methodologic$ adj5 review$).tw,sh.
22. (methodologic$ adj5 overview$).tw,sh.
23. (integrative research review$ or research integration).tw.
24. or/13-23
25. 12 or 24
26. depress$.tw
27. anxi$.tw
28. (Mental Disorders).sh
29. psychiatr$.tw
30. (Mental Health).sh
31. mental.tw
32. psychotic.tw
33. psychosis.tw
34. or/26-33
35. (Sick Leave).sh
36. (Return to Work).sh
37. absenteeism.sh
38. (sick time).tw
39. (sickness absence).tw
40. sickout.tw
41. reinstatement.tw
42. resumption.tw
43. retread.tw

44. comeback.tw
45. job.tw
46. (return to work).tw
47. or/35-46
48. 25 and 34 and 47

参考資料 2：Ovid PsycINFO 検索式

(((comprehensive* or integrative or systematic*) adj3 (bibliographic* or review* or literature)) or (meta-analy* or metaanaly* or "research synthesis" or ((information or data) adj3 synthesis) or (data adj2 extract*))).ti,ab,id. or ((review adj5 (rationale or evidence)).ti,ab,id. and "Literature Review".md.) or (cinahl or (cochrane adj3 trial*) or embase or medline or psyclit or pubmed or scopus or "sociological abstracts" or "web of science").ab. or ("systematic review" or "meta analysis").md. and ((depress* or anxi* or psychiatr* or mental or psychotic psychosis).ti,ab,id. or mental disorders/ or mental health/) and ((absenteeism or sick time or sickness absence or sickout or reinstatement or resumption or retread or comeback or return to work).ti,ab,id. or sick leave/ or return to work/)

III. 研究成果の刊行に関する一覧表

研究成果の刊行に関する一覧表

書籍

なし

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
Jike M, Itani O, Watanabe N, Buysse DJ, Kaneita Y	Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression	*Sleep Med Rev*	39	25-36	2018
Itani O, Kaneita Y, Jike M, Furuya M, Uezono C, Oda F, Agematsu R, Tokiya M, Otsuka Y, Ohida T	Sleep-related factors associated with industrial accidents among factory workers and sleep hygiene education intervention	*Sleep Biol Rhythms*	16	239-251	2018
Watanabe N, Matsuoka Y, Kumachi M, Hamazaki K, Horikoshi M, Furukawa TA	Omega-3 fatty acids for a better mental state in working populations - Happy Nurse Project: A 52-week randomized controlled trial.	*J Psychiatr Res*	102	72-80	2018
Ogawa Y, Furukawa TA, Takeshima N, Hayasaka Y, Atkinson LZ, Tanaka S, Cipriani A, Salanti G	Efficacy of antidepressants over placebo is similar in two-armed versus three-armed or more-armed randomized placebo-controlled trials	*Int Clin Psychopharmacol*	33	66-72	2018
Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, Ogawa S, Sato H, Kitagawa N, Shinagawa Y, Ikeda Y, Imai H, Tajika A, Ogawa Y, Akechi T, Yamada M, Shimodera S, Watanabe N, Inagaki M, Hasegawa A	Cognitive and Behavioral Skills Exercises Completed by Patients with Major Depression During Smartphone Cognitive Behavioral Therapy: Secondary Analysis of a Randomized Controlled Trial.	*JMIR Ment Health*	5	e4	2018

Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JPT, Egger M, Takeshima N, Hayasaka Y, Imai H, Shinohara K, Tajika A, Ioannidis JPA, Geddes JR	Comparative efficacy and acceptability of first- and second- generation antidepressants in the acute treatment of major depressive disorder: a network meta-analysis	*Lancet*	391	1357-1366	2018
Soares-Weiser K, Rathbone J, Ogawa Y, Shinohara K, Bergman H. Miscellaneous treatments for antipsychotic-induced tardive dyskinesia	Miscellaneous treatments for antipsychotic-induced tardive dyskinesia	*Cochrane Database Syst Rev*	3	CD000208	2018
Itani O, Jike M, Watanabe N, Kaneita Y	Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression	*Sleep Med*	3(32)	246-256	2017
Kinjo A, Imamoto A, Ikeda M, Itani O, Ohida T, Kaneita Y, Kanda H, Tanihata T, Higuchi S, Osaki Y	The Association Between Alcohol-Flavoured Non-Alcoholic Beverages and Alcohol Use in Japanese Adolescents	*Alcohol Alcohol*	52	351-357	2017
Otsuka Y, Kaneita Y, Itani O, Nakagome S, Jike M, Ohida T	Relationship between stress coping and sleep disorders among the general Japanese population: a nationwide representative survey	*Sleep Med*	37	38-45	2017
Tokiya M, Kaneita Y, Itani O, Jike M, Ohida T	Predictors of insomnia onset in adolescents in Japan	*Sleep Med*	38	37-43	2017
Itani O, Kaneita Y, Tokiya M, Jike M, Murata A, Nakagome S, Otsuka Y, Ohida T	Short sleep duration, shift work, and actual days taken off work are predictive life-style risk factors for new-onset metabolic syndrome: a 7-year cohort study of 40,000 male workers	*Sleep Med*	39	87-94	2017

Otsuka Y, Kaneita Y, Nakagome S, Jike M, Itani O, Ohida T	Nightmares and sleep paralysis among the general Japanese population: a nationwide representative survey	*Sleep Biol Rhythms*	16	187-195	2017
Hashimoto K, Ogawa Y, Takeshima N, Furukawa TA	Psychological and Educational Interventions for Atopic Dermatitis in Adults: A Systematic Review and Meta-Analysis	*Behavior Change.*	34	48-65	2017
Tateno M, Jovanovic N, Beezhold J, Uehara-Aoyama K, Umene-Nakano W, Nakamae T, Uchida N, Hashimoto N, Kikuchi S, Wake Y, Fujisawa D, Ikari K, Otsuka K, Takahashi K, Okugawa G, Watanabe N, Shirasaka T, Kato TA	Suicidal ideation and burnout among psychiatric trainees in Japan	*Early Interv Psychiatry*		1-3	2017
Kusaka K, Shinohara K, Tada M, Watanabe N, Furukawa TA	Concerns About Selective Outcome Reporting	*J Clin Oncol*	35	688	2017

IV. 研究成果の刊行物・別刷

Sleep Medicine Reviews 39 (2018) 25–36

Contents lists available at ScienceDirect

Sleep Medicine Reviews

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

CLINICAL REVIEW

Long sleep duration and health outcomes: A systematic review, meta-analysis and meta-regression

Maki Jike [a, 1], Osamu Itani [b, 1], Norio Watanabe [c, *], Daniel J. Buysse [d], Yoshitaka Kaneita [a]

[a] Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, Tokyo, Japan
[b] Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Oita, Japan
[c] Department of Health Promotion and Human Behavior, Department of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Japan
[d] Sleep Medicine Institute and Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA

ARTICLE INFO

Article history:
Received 31 January 2017
Received in revised form
19 May 2017
Accepted 28 June 2017
Available online 5 July 2017

Keywords:
Diabetes mellitus
Hypertension
Mortality
Meta-regression
Sleep deprivation
Vascular disease

SUMMARY

We examined the dose–response relationship between long sleep duration and health outcomes including mortality and the incidence of diabetes mellitus, hypertension, cardiovascular diseases, stroke, coronary heart diseases, obesity, depression and dyslipidemia. We collected data from 5,134,036 participants from 137 prospective cohort studies. For the independent variable, we categorized participants at baseline as having long sleep duration or normal sleep duration. Risk ratios (RRs) for mortality and incident health conditions during follow-up were calculated through meta-analyses of adjusted data from individual studies. Meta-regression analyses were performed to investigate the association between each outcome and specific thresholds of long sleep. Long sleep was significantly associated with mortality (RR, 1.39; 95% CI, 1.31–1.47), incident diabetes mellitus (1.26, 1.11–1.43), cardiovascular disease (1.25, 1.14–1.37), stroke (1.46, 1.26–1.69), coronary heart disease (1.24, 1.13–1.37), and obesity (1.08, 1.02–1.15). Long sleep was not significantly related to incident hypertension (1.01, 0.95–1.07). Insufficient data were available for depression and dyslipidemia. Meta-regression analyses found statistically significant linear associations between longer sleep duration and increased mortality and incident cardiovascular disease. Future studies should address whether the relationship between long sleep and health outcomes is causal and modifiable.

Introduction

Short and long sleep duration have been associated with a range of adverse health outcomes. The relationship between short sleep duration and health outcomes has been examined more closely. A causal relationship between short sleep duration and adverse health outcomes is supported by meta-analyses of epidemiological and cohort data [1,2], and by laboratory studies of experimental sleep restriction and sleep deprivation [3,4]. Fewer studies have addressed the relationship between long sleep duration and health outcomes, despite the prevalence of long sleep. Long sleep duration, defined more than 9 h of sleep, is prevalent in several developed countries, including Australia (an adjusted proportion of 33% in 2006), Finland (38% in 1999), Germany (40% in 2001), the Netherlands (25.7% in 2005), Sweden (30% in 2000), the U.K. (26% in 2005) and the U.S. (38% in 2007), according to a study using a data from time use survey [5]. A recent survey on behavioral risk factors conducted in the U.S. revealed that approximately 8% of adult respondents reported sleeping nine or longer hours [6].

Several systematic reviews have shown that long sleep duration is associated with important health outcomes including not only mortality [7–10] but also cardiovascular diseases [11], stroke [12,13], and diabetes mellitus [14,15]. However, because these reviews utilized different methodologies, we conducted a systematic review using the same methodology across all health outcomes.

Abbreviations: CER, control event rate; CI, confidence interval; HR, hazard ratio; MOOSE, meta-analysis of observational studies in epidemiology; NOS, Newcastle–Ottawa scale; OR, odds ratio; PRISMA, preferred reporting items for systematic reviews and meta-analyses; RR, risk ratio.

* Corresponding author. Department of Health Promotion and Human Behavior, Department of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.

E-mail addresses: jike.maki@nihon-u.ac.jp (M. Jike), itaniosamu@oita-u.ac.jp (O. Itani), noriowncu@gmail.com (N. Watanabe), BuysseDJ@upmc.edu (D.J. Buysse), kaneita.yoshitaka@gmail.com (Y. Kaneita).

[1] MJ and OI contributed equally to this work as the first authors.

http://dx.doi.org/10.1016/j.smrv.2017.06.011

This approach may allow us to consider the magnitude of long sleep duration effects on different health outcomes.

We have previously examined the association between short sleep duration and important health outcomes, and observed significant associations with outcomes including mortality (risk ratio (RR), 1.12; 95% CI, 1.08–1.16), diabetes mellitus (1.37, 1.22–1.53), hypertension (1.17, 1.09–1.26), cardiovascular diseases (1.16, 1.10–1.23), coronary heart diseases (1.26, 1.15–1.38), and obesity (1.38, 1.25–1.53) [16]. In the present study, we conducted a systematic review, meta-analyses and meta-regression to examine whether long sleep duration is associated with adverse health outcomes using the same methodology. Our method also allowed us to explore the specific sleep duration that is associated with these adverse health outcomes.

Methods

As described in our previous study on short sleep duration [16], we performed the study in accordance with PRISMA (preferred reporting items for systematic reviews and meta-analyses) [17] and MOOSE (meta-analysis of observational studies in epidemiology) [18] guidelines; see Appendices S5 and S6 in Supplementary material for PRISMA and MOOSE checklists.

Two independent researchers (OI and MJ) separately assessed the eligibility, extracted data, and checked the quality of the included studies. Any disagreements were resolved through discussion between these two, with adjudication by a third reviewer (NW) if disagreements persisted.

Data sources and searches

The studies were initially identified on October 17, 2013, through a search of PubMed, PsycINFO, CINAHL and Embase using pre-specified search terms (Appendices S1–S4). The review was not limited to English-language articles. We also hand-searched major medical journals, conference proceedings, and reference lists of included studies and previous systematic reviews for published, unpublished, and ongoing studies. We conducted a search of PubMed using the same search strategy on October 9, 2014 and on May 6, 2016, to identify new studies published during the review process.

Study selection

We included studies that compared individuals with long and "normal" sleep duration on outcomes of mortality and incident health outcomes. All studies included long-term follow-up, used a prospective cohort or randomized controlled trial design, and were conducted in community settings. We limited studies to those with a minimum follow-up duration of 1 y from baseline, and a minimum of 20 participants. Studies were excluded if most participants were aged 20 y or less at baseline, or if participants had been diagnosed with the health outcome at baseline. We also excluded studies that were conducted in inpatient settings and those that involved pharmacological interventions.

The eligibility of each study for inclusion was checked at two stages: 1) review of the title and abstract and 2) review of the full text.

Data extraction and quality assessment

Definition of long duration

Long sleep duration was defined in one of two ways for each paper. For some papers, long sleep duration was defined according to the criteria established by that source paper, given that sleep duration varies among cultures and ethnicities [19,20]. For papers that did not establish a criterion, long sleep was defined as the longest sleep duration range reported for participants in the original article. Durations of sleep in the definition of long sleep were incorporated into subgroup analyses and meta-regression as mediators (see below). When both a self-report (e.g., sleep diary) and objective (e.g., actigraphy or polysomnography) sleep duration were reported, we selected the former as the independent variable. Although self-report may capture a different amount of sleep per night than actigraphy [21] or polysomnography [22], objective measures are less commonly utilized in community settings, making self-report measures more widely applicable. When both sleep duration per night and per 24 h (i.e., including daytime naps) were reported, we selected the former.

The duration of normal sleep was also defined based on each source paper, or defined as the reference range for participants in the original article.

Outcome measures

Outcome measures included mortality and incidence of adverse health outcomes, specifically diabetes mellitus, hypertension, dyslipidemia (hypo or hyperlipidemia), cardiovascular diseases (including events in the heart and brain), coronary heart diseases, stroke, obesity, and depression. When a formal diagnosis was not provided, a surrogate outcome (e.g., coronary artery calcification instead of a diagnosis of coronary artery disease, or a self-report of diabetes mellitus without evidence of formal diagnosis) was included in the primary analyses, but a sensitivity analysis was also conducted (see below).

Assessment of bias

We employed the Newcastle–Ottawa scale (NOS) [23] to assess the studies' quality. The instrument has three broad categories (patient selection, four criteria; comparability of study groups, one criterion; and assessment of the outcome, three criteria). For the comparability criteria, we allotted two stars according to the depth of statistical adjustment for risk factors in the original studies (e.g., one star for age, sex, and race only; two stars for further factors). Therefore, a study could have a maximum quality rating of nine stars. Although previous meta-analyses [24,25] deemed study quality as high with five or more stars on the NOS criteria, we set a threshold of eight or more stars, to focus on very high quality studies.

We defined adequate follow-up for each disorder in terms of duration and % attrition (i.e., 3 y and 10% attrition for all-cause mortality, cardiovascular diseases, and coronary heart disease; and 2 y and 20% attrition for diabetes mellitus, hypertension, dyslipidemia, obesity, and depression).

Data synthesis and analysis

We analyzed data descriptively and conducted a meta-analysis for each health outcome. In the meta-analysis, we calculated risk ratios (RRs) by pooling adjusted RRs between long and normal sleep provided by the original studies, using a random effects model. If a study provided a point estimate and a p-value but not confidence intervals of RRs, the intervals were obtained using a statistical method with log transformations of the estimate [26]. If hazard ratios (HRs) were reported for a study but RRs were not, the HRs were regarded as RRs. Among studies where odds ratios (ORs) were provided but not RRs, we calculated RRs by using the ORs and control event rates (CERs) in normal duration sleepers reported in the original studies. For studies in which neither RRs nor CERs were reported, and only ORs were provided, CERs were borrowed from a study whose characteristics were similar. For studies in which RRs were provided for subgroups separately (e.g., male and female),

data from subgroups were combined using a fixed-effect meta-analysis.

Statistical heterogeneity between studies was investigated using the I^2 statistic [27], assuming an I^2 of 75% or greater to be an important level of inconsistency, as employed by a previous review [28]. To assess publication bias, we used a funnel plot and Egger's test for all primary outcomes [29]. We used the "trim and fill" method to adjust the funnel plot, then recalculated results [30].

Because subgroup analyses should be interpreted with caution [31], we planned a priori to limit our subgroup analyses to a small number of baseline characteristics such as age and sex (i.e., between 20 and 65 y, or 65 y or more; male, or female).

Sensitivity analyses were planned a priori for the primary analyses set by: 1) excluding studies with surrogate outcomes; 2) limiting studies to those in which sleep duration was reported per night; 3) limiting analyses to studies with eight or more stars in the NOS; 4) limiting studies to those with 10 or more y of follow-up; and 5) excluding studies in which CERs from other studies were used to calculate RRs.

In order to explore possible mediation by the definition of long sleep, additional subgroup analyses were conducted by clustering studies according to the definition of long sleep (e.g., more than 9 h or more than 10 h). When studies were clustered into three or more levels of sleep duration for each outcome, meta-regression analyses were also performed. These analyses examined linear associations between sleep duration and the outcome of interest, using a random-effects model and illustrating the regression line and its 95% prediction intervals.

A p-value of less than 0.05 was chosen to test null hypotheses, despite multiple comparisons, in order to avoid type II over type I errors. For all outcomes, 95% confidence intervals (CIs) were calculated. The data were analyzed using the Comprehensive Meta-Analysis Software (Version 3) [32].

Results

Search results

The initial electronic search yielded 3580 articles, and an additional database search identified 182 studies on October 9, 2014 and 388 on May 6, 2016. In total, 2521 studies remained after removing duplicate articles. A hand-search did not identify any studies that had not been included in the electronic search (Fig. 1). At the first and second eligibility check stages, two independent researchers identified 277 articles and 95 articles, respectively.

Characteristics of included studies

All of the 95 included studies were prospective cohort studies. From these, 137 datasets for nine outcomes (N = 5,134,036) were collected. Most studies were conducted in developed countries (see Table 1 and Tables S1–S9 in Supplementary material). The number of participants in each dataset ranged from 276 to 392,164; the duration of follow-up was from 1 to 34 y; and the total NOS scores ranged from five to nine. Although the definition of long sleep varied among studies, most defined long sleep as greater than 8 or 9 h.

We were unable to pool data from six datasets in meta-analyses because no usable data for meta-analyses were provided (Table S10). The number of datasets included in the meta-analyses for each outcome varied from 8 (hypertension) to 36 (mortality). Table 1 shows the characteristics of studies included for the mortality outcome.

Effect estimates of long sleep compared to normal sleep from meta-analyses

Primary analyses

Compared with normal sleep duration, long sleep duration was associated with a statistically significant increase in all-cause mortality, with an RR of 1.39 (95% CI = 1.31–1.47, P < 0.001, I^2 = 83%, N of datasets = 36; Fig. 2). Qualitatively similar significant results were obtained for incident diabetes mellitus (RR = 1.26, 95% CI = 1.11–1.43, P < 0.001, I^2 = 63%, N = 16), cardiovascular disease (RR = 1.25, 95% CI = 1.14–1.36, P < 0.001, I^2 = 81%, N = 25), stroke (RR = 1.46, 95% CI = 1.26–1.69, P < 0.005, I^2 = 71%, N = 14), coronary heart disease (RR = 1.24, 95% CI = 1.13–1.37, P = 0.003, I^2 = 54%, N = 19), and obesity (RR = 1.08, 95% CI = 1.02–1.15, P = 0.010, I^2 = 0%, N = 13) (Fig. 3). Long sleep duration was not associated with a statistically significant increase in incident hypertension compared to normal sleep duration (RR = 1.01, 95% CI = 0.95–1.07, P = 0.309, N = 8). Substantial heterogeneity between datasets was observed in mortality and cardiovascular disease outcomes. Only one study each was identified for depression and dyslipidemia outcomes, and RRs were not provided in the source studies.

Possible publication bias for primary analyses

No significant publication bias was observed for any outcome in the funnel plots or results from Egger's test (see Figs. S2, S11, S22, S32, S42, S52, and S63 in Supplemental material).

Subgroup analyses for age groups

Subgroup analyses were conducted for participants aged ≥65 y or <65 y at baseline (Fig. 3). Compared to normal sleep duration, long sleep duration was associated with a significant increase in the incidence of cardiovascular disease and coronary heart disease among those ≥65 y, but not among those <65 y. On the other hand, long sleep duration was associated with a significant increase in incident obesity only among participants <65 y, and not among those ≥65 y.

Subgroup analyses for sex

In comparison with normal sleep duration, long sleep duration was associated with a significant increase in mortality and incident diabetes, cardiovascular disease, stroke, and coronary heart disease for both men and women. Long sleep duration was associated with a significant increase in incident obesity only among female (Fig. 3).

Sensitivity analyses

Most sensitivity analyses showed qualitatively similar results to those in the primary analyses (Fig. 3). Analyses limited to high quality studies based on the NOS (>8 stars) and to studies with follow-up greater than 10 y did not show statistically significant findings for incident diabetes (High-quality studies: RR = 1.13, 0.94–1.35, P = 0.191; Long follow-up studies: 1.97, 0.96–4.05, P = 0.064), or incident obesity (High-quality studies: RR 0.94, 0.40–2.24, P = 0.896; Long follow-up studies: 1.04, 0.95–1.13, P = 0.441).

Subgroup analyses and meta-regression for specific values of long sleep duration

Subgroup analyses for specific values of long sleep duration were conducted for outcomes other than depression and dyslipidemia (Fig. 4). In comparison with normal sleep duration, long sleep duration defined as >8 h was associated with a significant increase in mortality, stroke, and coronary heart disease; long sleep defined as >9 h was associated with an increase in incident

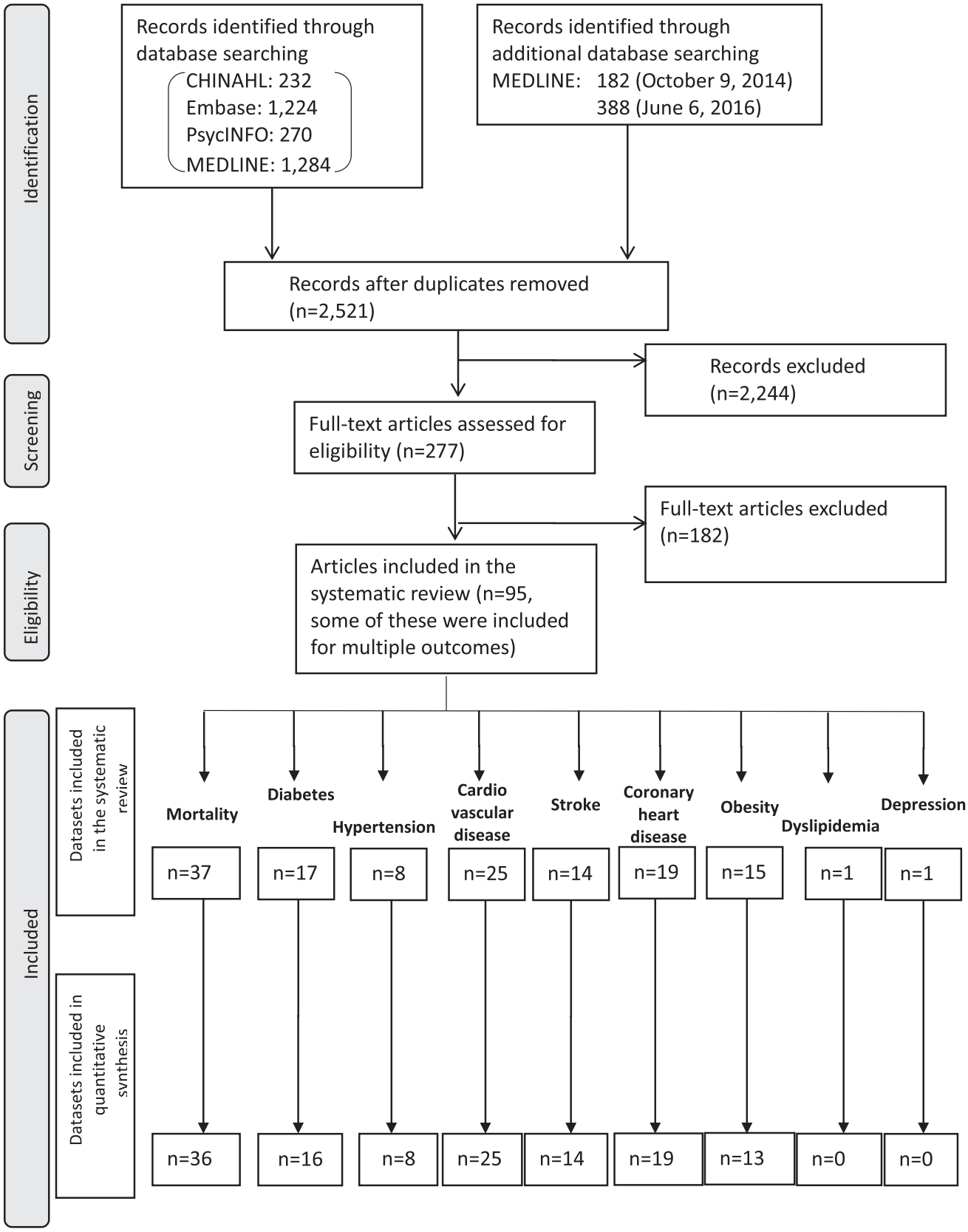


Fig. 1. Flowchart for the included studies. An arrow indicates that the upper limit of the 95% CI is over the scale described here.

Table 1
Characteristics of studies for mortality outcome.

Study	Sample size	Male%	Mean age ± SD (range) in years at baseline	Years of follow up mean ± SD	Definition of normal sleep duration (h)	Definition of long sleep duration (h)	Newcastle–Ottawa Scale: Selection/Comparability/Outcome
Tsubono 1993 [53]	4318	39.8	61.4 (≥40)	4	7–8/night	≥9/night	★★☆★/★★/★★★
Kojima 2000 [54]	5322	45.8	Male: 46.9 (20–67), Female: 47.7 (20–67)	11.9	7.0–8.9/night	≥10/night	☆★☆★/★★/★★★
Seki 2001 [55]	1065	41.3	65.3 ± 3.6 (60–74)	7.5	7/d	≥9/d	★★☆★/★★/★★★
Heslop 2002 [56]	1st screening: 7028 2nd screening: 3030	1st screening: 85.7 2nd screening: 85.4	Male: (≤65), Female; (≤60)	25	7–8/d	>8/d	★★☆★/★★/★★★
Mallon 2002 [57]	1870	48.4	56 (45–65)	12	6–8/night	>8/night	★★☆★/★☆/★★★
Burazeri 2003 [58]	1842	45.7	Male: median 64 (≥50), Female: median 63 (≥50)	9–11	6–8/night	>8/night	★★☆★/★★/★★★
Goto 2003 [59]	724	34.7	Male: median 73 (65–97), Female: median 74 (65–97)	12	6–7/d	>7/d	★★★★/★★/★★★
Amagai 2004 [60]	11,325	39.0	55.1 (19–93)	8.2 ± 1.5	7.0–7.9/night	≥9/night	★★★★/★★/★★★
Patel 2004 [61]	82,969	0.0	53.4 (40–65)	14	7/d	≥9/d	☆★☆★/★★/★★★
Tamakoshi 2004 [62]	104,010	42.2	56.6 (40–79)	9.9	7/d	≥10/d	★★☆☆/★★/★★★
Ferrie 2007 [63]	Phase 1: 9781 Phase3: 7729	NS	(35–55)	Phase 1: 17.1 Phase 3: 11.8	7/night	≥9/night	☆★☆★/★★/★★★
Lan 2007 [64]	3079	56.8	Male: 71.3 (≥64), Female: 71.9 (≥64)	8.4 ± 3.3	7–7.9/night	≥10/night	★★★★/★★/★★★
Gangwisch 2008 [65]	9789	37.2	Male; 45.0 (32–59), Female; 73.0 (60–86)	≤8, ≤10	7/night	≥9/night	★★★★/★★/★★★
Ikehara, 2009 [66]	98,634	42.1	Male; 58.8 (40–79), Female; 60.2 (40–79)	Median 14.3	7/d	≥10/d	★★☆★/★★/★★★
Stone 2009 [67]	8101	0.0	77.0 (≥69)	6.9	6–8/night	>8/night	★★★★/★★/★★★
Suzuki 2009 [68]	12,601	51.1	74.1 ± 5.4 (65–85)	5.3	7/d	≥10/d	☆★☆★/★★/★★★
Chien 2010 [69]	3430	47.3	(≥35)	15.9 (13.1–16.9)	7/d	≥9/d	★★☆★/★★/★★★
Mesas 2010 [70]	3820	43.8	71.8 ± 7.9 (≥60)	6.8	7/d	≥11/d	★★☆★/★★/★★★
Castro-costa 2011 [71]	1512	38.3	68.9 ± 7.1 (63–75)	7.5 Median: 8.9	7–8/night	≥9/night	★★★★/★★/★★★
Kripke 2011 [72]	434	0.0	67.6 ± 7.9 (50–81)	10.5	5–6.5/d	>6.5/d	★★★★/☆☆/★★★
Kronholm 2011 [73]	23,290	48.8	(25–64)	29–34	7–8/night	≥10/night	★★☆★/★★/★★★
Cohen-Mansfield 2012 [74]	1166	55.5	83.4 ± 5.3 (75–94)	20	7–9/night	>9/night	★★☆★/★★/★★★
Chen 2013 [75]	4064	55.8	73.8 ± 5.7 (≥65)	9	7/night	≥9/night	★★★★/★★/★★★
Garde 2013 [76]	4941	100.0	(40–59)	30	6–7/d	≥8/d	☆★★★/★★/★★★
Hale 2013 [77]	3942	0.0	62.1 (50–79)	11–16	7–8/night	≥9/night	★★☆★/★★/★★★
Kakizaki 2013 [78]	49,256	48.2	(40–79)	10.8	7/d	≥10/d	★★☆★/★★/★★★
Kim 2013 [79]	135,685	45.6	(45–75)	12.9	7/d	≥9/d	★★☆★/★★/★★★
Li 2013 [80]	9455	(38.1)	(20–79)	7	7/night	≥9/night	★★☆★/★★/★★★
Magee 2013 [81]	227,815	46.3	(≥45)	2.8	7/d	≥10/d	★★☆★/★★/★☆★
Yeo 2013 [82]	13,164	41.4	(≥20)	9.44	7/d	≥10/d	★★★★/★★/★★★
Bellavia 2014 [83]	70,973	53.3	(45–83)	15	6.6–7.4/d	>8/d	★★☆★/★★/★★★
Lee 2014 [84]	3427	50.9	(≥65)	5.1 ± 0.9	<10/night	≥10/night	☆★☆★/★★/★★★
Rod 2014 [85]	9098	67.2	45 (35–55)	22	7/night	>9/night	☆★☆★/★★/★★★
Xiao 2014 [86]	239,896	56.2	(51–72)	14	7–8/night	≥9/night	★★☆★/★★/★★☆
Zuurbier 2015 [87]	1734	46.6	62.2 ± 9.3 (45–98)	7.3 ± 1.3	6–7.5/night	>7.5/night	☆★★★/★★/★★★
Hall 2015 [88]	3013	48.6	73.6 ± 2.9 (70–79)	8.2 ± 2 0.3	7–8/night	>8/night	★★★★/★★/★★★
Cai 2015 [89]	113,138	60.1	Male: (40–75) Female: (44–79)	Median: Male: 6.07 Female: 7.12	7/d	≥10/d	★★★★/★★/★★★

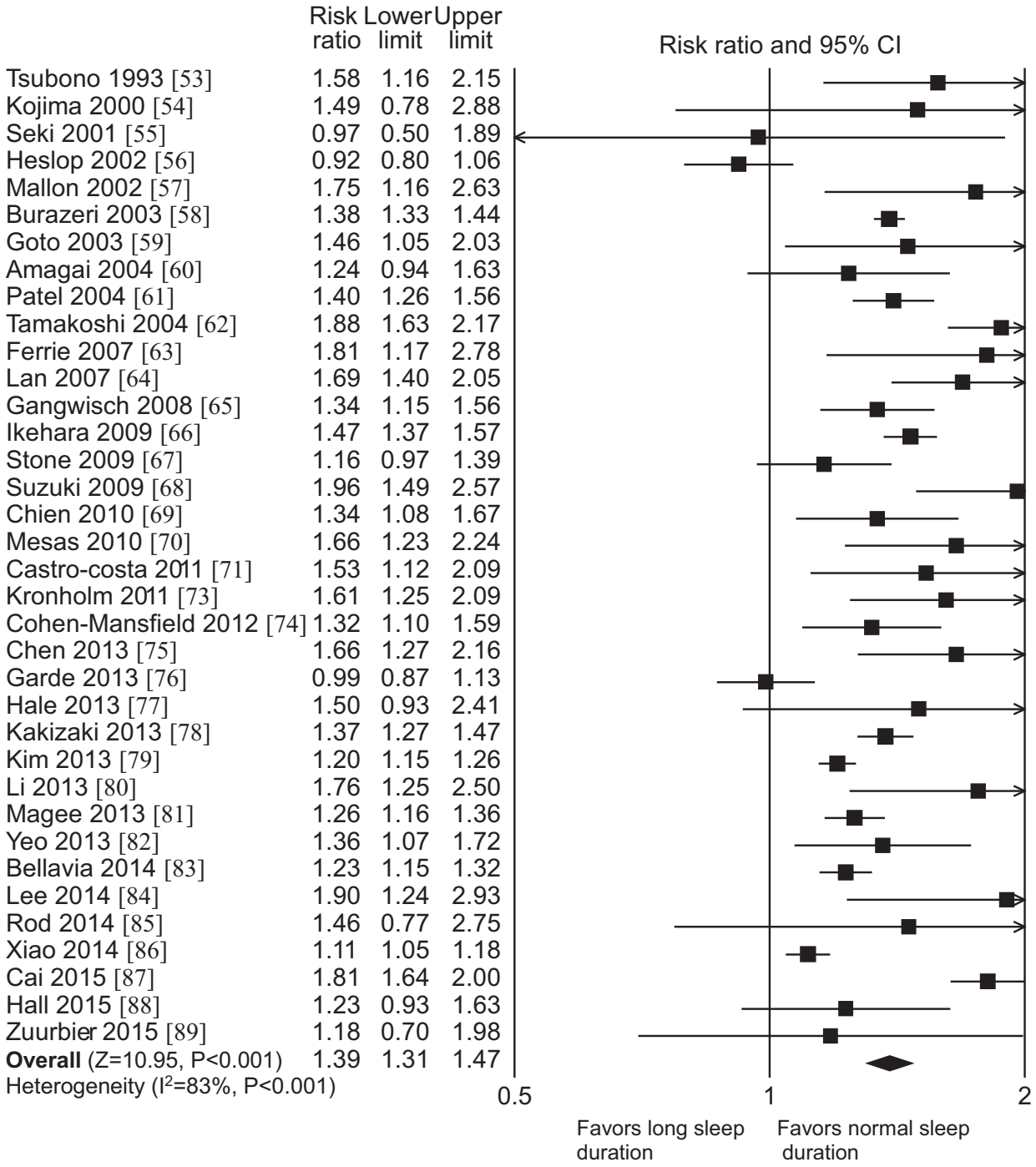


Fig. 2. Mortality outcome comparing long with normal sleepers. An arrow indicates that the upper limit of the 95% CI is over the scale described here.

cardiovascular disease. Similar findings were not observed for diabetes and obesity outcomes.

Meta-regression analyses were performed for mortality, diabetes, cardiovascular disease, stroke, coronary heart disease, and obesity outcomes. Longer sleep duration was significantly associated in a gradient fashion with increased mortality (coefficient = 0.120, P < 0.0001, R^2 analog = 0.03, Fig. 5). The 95% CIs of prediction curves were greater than zero beginning at a sleep duration of >9.5 h. Longer sleep duration was also associated in a gradient fashion with incident cardiovascular disease (coefficient = 0.165, P = 0.002, R^2 analog = 0.05, Fig. S40). No significant gradient associations were observed for the other outcomes.

Discussion

To the best of our knowledge, this is the first systematic review to investigate associations between long sleep duration and multiple health outcomes using the same methodology for each meta-analysis, and to explore the dose–response of long sleep duration on these outcomes using meta-regression analyses. Long sleep duration was associated with greater mortality and increased incidence of multiple health conditions. The strength of association varied across outcomes, with a point estimate RR of 1.46 for stroke, 1.39 for mortality, 1.26 for diabetes mellitus, 1.25 for cardiovascular disease, 1.24 for coronary heart disease and 1.08 for obesity. Long

sleep duration was not significantly associated with increased risk of hypertension, and available evidence was not sufficient to examine depression and dyslipidemia outcomes.

The present results are similar to those from previous systematic reviews. For instance, long sleep duration has been associated with a RR of 1.23 (N of studies = 17) [7] and 1.30 (N = 16) [8] for mortality; with a RR of 1.41 (N = 8) for cardiovascular disease [11]; and with an HR of 1.46 for stroke (N = 11) [12]. Our findings contribute important new information to previous reviews because of our updated comprehensive literature search and our use of the same rigorous methodology for each outcome. However, further epidemiological studies will be needed to investigate the association of long sleep with dyslipidemia and depression.

The lack of a statistically significant association between long sleep duration and hypertension (RR 1.01, 0.95–1.07) is consistent with findings in a previous review (1.02, 0.91–1.14) [33]. On the other hand, our findings differ from previous results for the obesity outcome; we found that long sleep was significantly associated with obesity (RR 1.08, 1.02–1.15, P = 0.010), whereas a previous review did not find a statistically significant association (odds ratio 1.06, 0.98–1.15) [34]. This discrepancy may partly be due to different statistical methodologies, but we also observed narrower confidence intervals, which is likely due to our inclusion of 15 datasets compared to 10 in the previous review.

In subgroup analyses addressing specific thresholds for long sleep duration, values greater than 8 or 9 h were associated with significantly increased mortality and cardiovascular disease in comparison with normal sleep. In meta-regression, longer duration of sleep was linearly associated with increased mortality risk. These findings are again consistent with those in previous studies [9,10]. Although the American Academy of Sleep Medicine and the Sleep Research Society have not published their recommendations for an upper limit of appropriate sleep duration, the National Sleep Foundation issued its recommendations, including 8–10 h for teenagers (aged 14–17 y), 7–9 h for adults (18–64 y) and 7–8 h for older adults (≥65) [35]. Normative total sleep time is reported to decrease with age not only in adolescents but also in adults [36], and most participants included in the present study were aged between 30 and 70 y. The recommendations match well with the results from our analyses.

The findings of the present study complement those of our previous review on short sleep duration [16]. Using identical methodology to that in the current study, we found that sleep duration <6 h is associated with a significant increase in mortality and health outcomes such as diabetes, cardiovascular disease, coronary heart disease, and obesity. In combination, the findings of these two papers are consistent with the "U-shaped" relationship between sleep duration and multiple health outcomes that has been previously described. However, we did not directly test curvilinear relationships in these papers. Moreover, the strength of association with sleep duration varies among the various outcomes, as described above.

Although several theoretical pathways may explain the relationship between long sleep duration and health outcomes [37], our study was unable to examine specific mechanisms. Moreover, we do not have data showing that changing sleep duration modifies health risks. Therefore, while we can confidently state that long sleep duration is a risk factor for adverse health outcomes, we cannot demonstrate that it is a *causal* risk factor [38]. Nevertheless, findings from our studies on the health risks of short and long sleep duration may encourage both mechanistic studies and intervention studies to investigate whether and how sleep duration confers health risk. Regarding education, psychotherapy, and psychosocial interventions, many studies have focused on insomnia and the efficacy of these interventions on the quality of sleep has been reported in recent systematic reviews [39,40]. To the best of our knowledge, there is no evidence base to test these interventions in the community for prevention of the health outcomes. Until such studies are conducted, our findings may be used in education and health literacy efforts to encourage adequate duration of sleep as a key component of overall sleep health [41].

Limitations of the study

Although our findings are internally consistent and consistent with previous studies, we acknowledge several limitations.

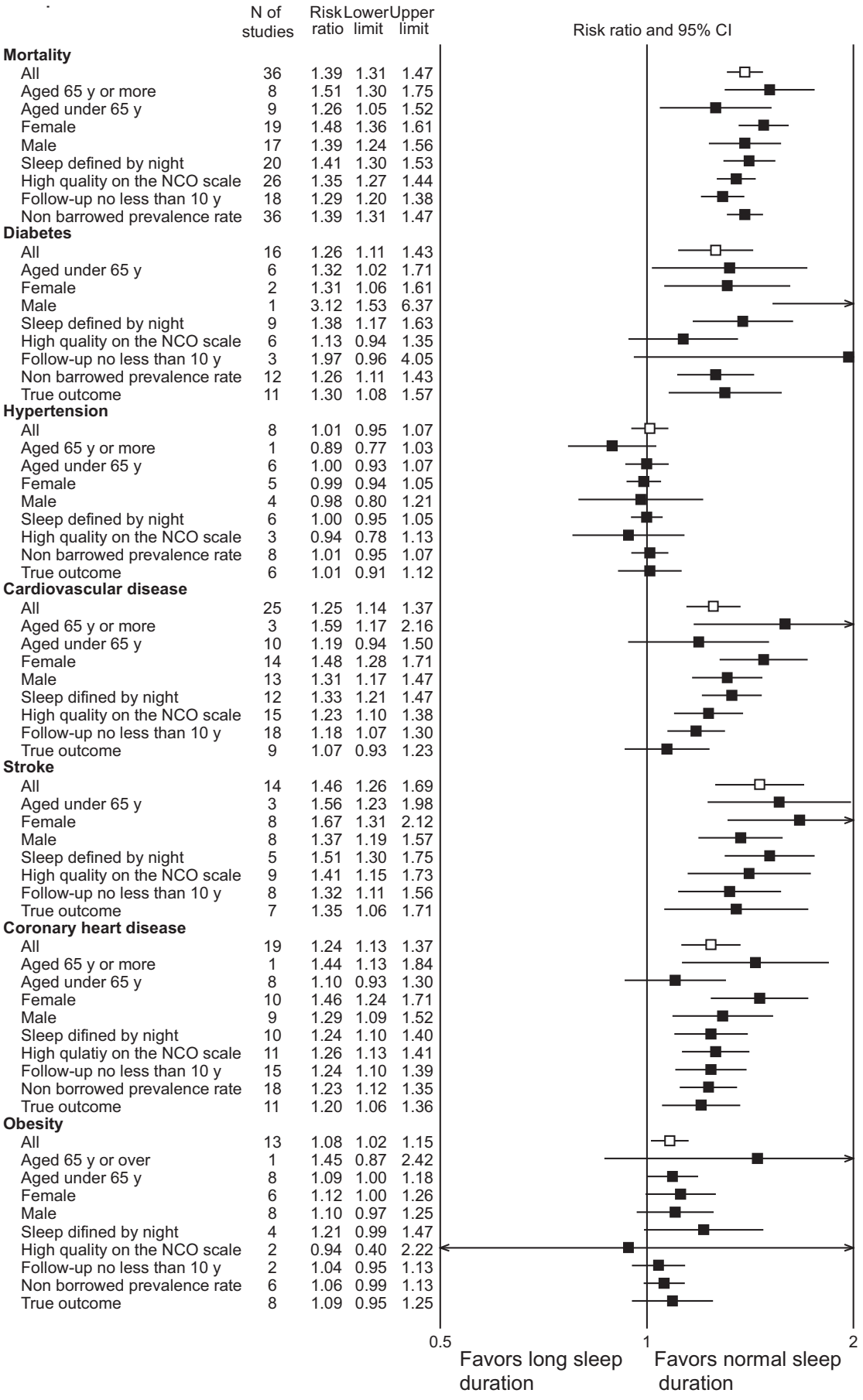
First, our analyses focused on night-time sleep rather than on 24-h sleep. Given the number of primary, subgroup, and sensitivity analyses, we were concerned about inflating the risk of type I errors. A previous systematic review conducted meta-analyses using both night-time and 24-h sleep [9], and reported similar findings with regard to all-cause mortality. Therefore, we believe that our decision to focus on night-time sleep is unlikely to have had a major effect on our conclusions.

Secondly, we collected most of the data on sleep duration based on results from an interview question or a questionnaire, or from a sleep diary, but not from objective measures (e.g., actigraphy or polysomnography), because we believed that self-reported measures were more widely utilized and applicable than objective measures in community settings. However, sleep duration assessment based on a simple question, sleep diary, and objective measures have been reported to be inconsistent [42,43]. In addition, the term "sleep duration" has been sometimes used as the amount of time in bed in previous epidemiological studies [44]. The definition of sleep duration utilized in our systematic review relies on the definition used in original studies and different measurement approaches might lead to discrepancies among the included studies, although we have extracted data on actual sleep duration with the utmost care and attention.

Thirdly, we did not investigate the impact of other dimensions of sleep, such as subjective or objective sleep quality, on health outcomes. Although previous studies have reported an association between sleep problems such as insomnia and mortality [45,46], we intended to focus on the duration of sleep because it is likely that sleep duration is more easily recognized and accurately reported by participants rather than the quality of sleep, in community surveys. However, we believe that interventions to reduce unwanted health outcomes should include information not only about sleep duration but also about the quality of sleep, because the latter may be an important factor in mechanisms linking sleep and subsequent health outcomes.

Fourthly, we have done our meta-analyses by pooling adjusted RRs between long and normal sleep as provided by the original studies. However, we were dependent on the original studies as to whether types of confounders had been adjusted or not, and some important confounders, including employment status, depression and excessive amounts of time in bed, were not adjusted in all studies (Tables S1–S9). This is a limitation of meta-analyses pooling aggregated data.

Fifthly, we recognize that the role of individual differences regarding sleep duration preferences is still uncertain [47], and that the reasons for different durations of sleep vary from person to person. Previous epidemiological studies have shown that short sleep duration is associated with characteristics such as being unmarried [48], more frequent binge drinking [49], lower socioeconomic status [49,50], lower education levels [50], working multiple jobs [51], pre- and post-sleep activities including socializing, self-care and hygiene, and watching TV [51] and African, Asian and


N of studies
Risk ratio
Lower limit
Upper limit
Risk ratio and 95% CI
Mortality
All 36 1.39 1.31 1.47
Aged 65 y or more 8 1.51 1.30 1.75
Aged under 65 y 9 1.26 1.05 1.52
Female 19 1.48 1.36 1.61
Male 17 1.39 1.24 1.56
Sleep defined by night 20 1.41 1.30 1.53
High quality on the NCO scale 26 1.35 1.27 1.44
Follow-up no less than 10 y 18 1.29 1.20 1.38
Non barrowed prevalence rate 36 1.39 1.31 1.47
Diabetes
All 16 1.26 1.11 1.43
Aged under 65 y 6 1.32 1.02 1.71
Female 2 1.31 1.06 1.61
Male 1 3.12 1.53 6.37
Sleep defined by night 9 1.38 1.17 1.63
High quality on the NCO scale 6 1.13 0.94 1.35
Follow-up no less than 10 y 3 1.97 0.96 4.05
Non barrowed prevalence rate 12 1.26 1.11 1.43
True outcome 11 1.30 1.08 1.57
Hypertension
All 8 1.01 0.95 1.07
Aged 65 y or more 1 0.89 0.77 1.03
Aged under 65 y 6 1.00 0.93 1.07
Female 5 0.99 0.94 1.05
Male 4 0.98 0.80 1.21
Sleep defined by night 6 1.00 0.95 1.05
High quality on the NCO scale 3 0.94 0.78 1.13
Non barrowed prevalence rate 8 1.01 0.95 1.07
True outcome 6 1.01 0.91 1.12
Cardiovascular disease
All 25 1.25 1.14 1.37
Aged 65 y or more 3 1.59 1.17 2.16
Aged under 65 y 10 1.19 0.94 1.50
Female 14 1.48 1.28 1.71
Male 13 1.31 1.17 1.47
Sleep difined by night 12 1.33 1.21 1.47
High quality on the NCO scale 15 1.23 1.10 1.38
Follow-up no less than 10 y 18 1.18 1.07 1.30
True outcome 9 1.07 0.93 1.23
Stroke
All 14 1.46 1.26 1.69
Aged under 65 y 3 1.56 1.23 1.98
Female 8 1.67 1.31 2.12
Male 8 1.37 1.19 1.57
Sleep defined by night 5 1.51 1.30 1.75
High quality on the NCO scale 9 1.41 1.15 1.73
Follow-up no less than 10 y 8 1.32 1.11 1.56
True outcome 7 1.35 1.06 1.71
Coronary heart disease
All 19 1.24 1.13 1.37
Aged 65 y or more 1 1.44 1.13 1.84
Aged under 65 y 8 1.10 0.93 1.30
Female 10 1.46 1.24 1.71
Male 9 1.29 1.09 1.52
Sleep difined by night 10 1.24 1.10 1.40
High qulatiy on the NCO scale 11 1.26 1.13 1.41
Follow-up no less than 10 y 15 1.24 1.10 1.39
Non borrowed prevalence rate 18 1.23 1.12 1.35
True outcome 11 1.20 1.06 1.36
Obesity
All 13 1.08 1.02 1.15
Aged 65 y or over 1 1.45 0.87 2.42
Aged under 65 y 8 1.09 1.00 1.18
Female 6 1.12 1.00 1.26
Male 8 1.10 0.97 1.25
Sleep difined by night 4 1.21 0.99 1.47
High quality on the NCO scale 2 0.94 0.40 2.22
Follow-up no less than 10 y 2 1.04 0.95 1.13
Non borrowed prevalence rate 6 1.06 0.99 1.13
True outcome 8 1.09 0.95 1.25
0.5
1
2
Favors long sleep duration
Favors normal sleep duration

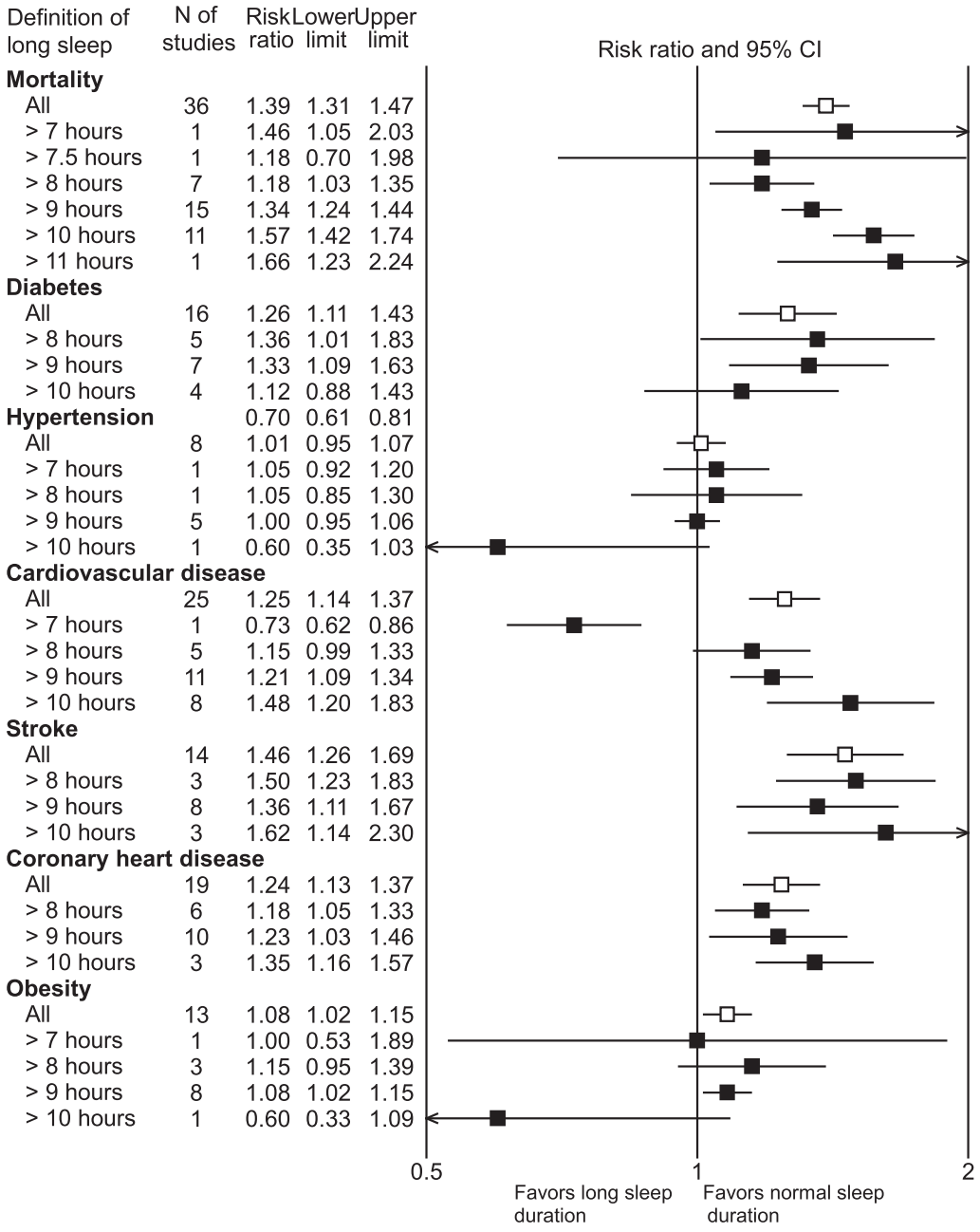


Fig. 4. Subgroup analyses for duration thresholds of long sleep. An arrow indicates that the upper limit of the 95% CI is over the scale described here.

Hispanic ethnicity [50]. On the other hand, characteristics related to long sleep duration include Mexican ethnicity, public liability insurance [50], and less weekly physical activity [52]. Some of these characteristics, such as working multiple jobs, may present barriers for possible interventions aimed at sleep duration because of economic disincentives. However, other characteristics associated with sleep duration may be more amenable to change, including frequent binge drinking, the amount of physical activity, and using smart phones and television viewing prior to sleeping. Future studies focusing on sleep duration must also account for individual preferences.

Finally, shorter or longer periods of sleep duration are unlikely to link directly to sleep disorders such as sleep apnea or insomnia. Therefore, our findings are likely to have limited comparability with findings derived from patients seeking treatment in clinical settings.

Fig. 3. Relative risks of mortality and health outcomes comparing long with normal sleepers. NCO scale, Newcastle–Ottawa Scale. For the outcomes of dyslipidemia and depression, no meta-analyses were performed. An arrow indicates that the upper limit of the 95% CI is over the scale described here.

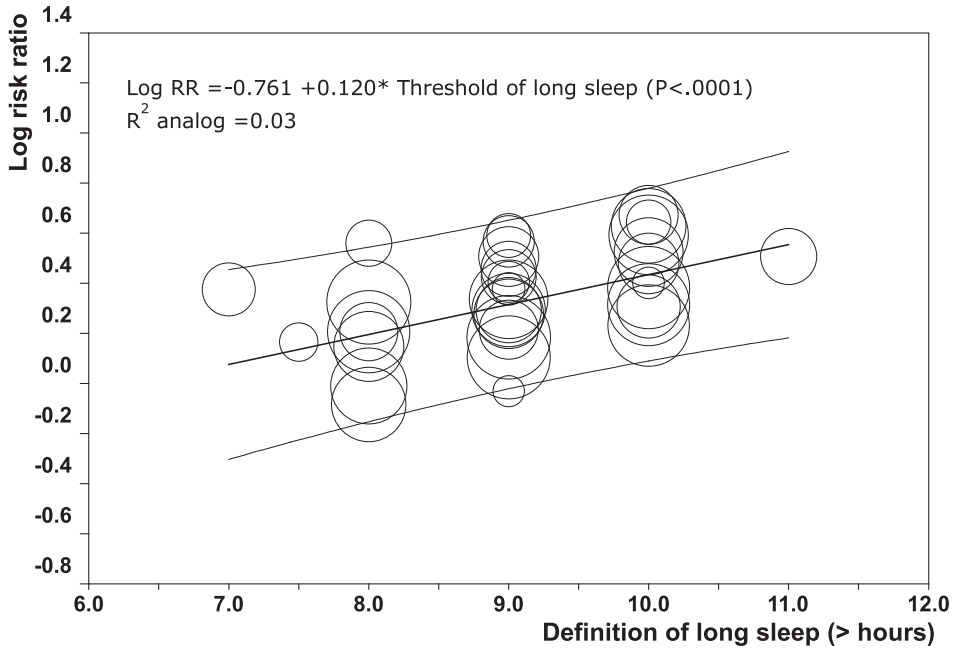


Fig. 5. Meta-regression for specific values of long sleep duration in mortality.

Conclusions

Long sleep duration was associated with a significant increase in risk for mortality and incident diabetes, cardiovascular disease, stroke, coronary heart disease, and obesity. Future studies should address potential mechanisms underlying the relationship between long sleep duration and adverse health outcomes. Whether interventions to reduce long sleep duration also reduce health risk remains an open question.

Practice points

1). Long sleep duration is associated with greater mortality and increased incidence of mortality, diabetes mellitus, cardiovascular disease, stroke, coronary heart disease, and obesity.
2). The strength of association with long sleep duration varies among these outcomes
3). Longer duration of sleep is linearly associated with increased mortality risk.
4). Currently, whether interventions to reduce long sleep duration also reduce health risk remains an open question.

Research agenda

1). Future studies should address potential mechanisms underlying the relationship between long sleep duration and adverse health outcomes.
2). Studies on interventions aimed at sleep duration should be conducted to investigate whether adverse health risks decrease in community settings.

Authors' contributions

All authors contributed to the manuscript as follows:

OI designed the study, developed technical materials, acquired data, and interpreted the data.

MJ designed the study, developed technical materials, acquired data, and interpreted the data.

NW designed the study, interpreted the data, and drafted the manuscript.

DJB interpreted the data, and drafted the manuscript.

YK obtained funding, conceived the study, designed the study, and interpreted the data.

All authors have revised the important intellectual content critically, have read and approved the final manuscript, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

Institution at which the work was performed

Nihon University, Oita University, and Kyoto University, Japan, and University of Pittsburgh, USA.

Disclosure of conflicts of interests

The authors have no conflicts of interests to declare, that may be affected by the publication of the paper. Other conflicts of interests are as follows:

OI has research funds from the Japanese Ministry of Health Labor and Welfare and the Japanese Ministry of Education, Science, and Technology.

MJ has research funds from the Japanese Ministry of Health Labor.

NW has received research funds from the Japanese Ministry of Health Labor and Welfare, the Japanese Ministry of Education, Science, and Technology and National Center of Neurology and Psychiatry, Intramural Research Grant for Neurological and

Psychiatric Disorders. He has also received royalties from Sogensha, Paquet and Akatsuki, and speaking fees from Dai-Nippon Sumitomo, MSD, Otsuka, Eisai, Pfizer, and Takeda during last 5 y.

DJB has served as a paid consultant for the following companies over the past 5 y, each at a level of less than $5000 per 12-mo period: Bayer HealthCare, BeHealth Solutions, Cereve, CMEOutfitters, Emmi Solutions, Medscape, Merck, and Purdue.

YK has research funds from the Japanese Ministry of Health Labor and Welfare and the Japanese Ministry of Education, Science, and Technology.

Acknowledgements

This study was funded by Industrial Disease Clinical Research Grants (160102-01), and by Health Labor Sciences Research Grant (H25-JUNKANKITOU-IPPAN-007) from the Japanese Ministry of Health Labor and Welfare. The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.smrv.2017.06.011.

References

[1] Chaput JP, McNeil J, Despres JP, Bouchard C, Tremblay A. Short sleep duration as a risk factor for the development of the metabolic syndrome in adults. Prev Med 2013;57:872–7.
[2] King CR, Knutson KL, Rathouz PJ, Sidney S, Liu K, Lauderdale DS. Short sleep duration and incident coronary artery calcification. JAMA J Am Med Assoc 2008;300:2859–66.
[3] Spiegel K, Leproult R, Van Cauter E. Impact of sleep debt on metabolic and endocrine function. Lancet 1999;354:1435–9.
[4] Calvin AD, Carter RE, Adachi T, Macedo PG, Albuquerque FN, van der Walt C, et al. Effects of experimental sleep restriction on caloric intake and activity energy expenditure. Chest 2013;144:79–86.
[5] Bin YS, Marshall NS, Glozier N. Sleeping at the limits: the changing prevalence of short and long sleep durations in 10 countries. Am J Epidemiol 2013;177:826–33.
[6] Liu Y, Wheaton AG, Chapman DP, Cunningham TJ, Lu H, Croft JB. Prevalence of healthy sleep duration among adults–United States, 2014. MMWR Morb Mortal Wkly Rep 2016;65:137–41.
[7] Gallicchio L, Kalesan B. Sleep duration and mortality: a systematic review and meta-analysis. J Sleep Res 2009;18:148–58.
*[8] Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep 2010;33:585–92.
*[9] Shen X, Wu Y, Zhang D. Nighttime sleep duration, 24-hour sleep duration and risk of all-cause mortality among adults: a meta-analysis of prospective cohort studies. Sci Rep 2016;6, 21480.
*[10] Liu TZ, Xu C, Rota M, Cai H, Zhang C, Shi MJ, et al. Sleep duration and risk of all-cause mortality: a flexible, non-linear, meta-regression of 40 prospective cohort studies. Sleep Med Rev 2016;32:28–36.
[11] Cappuccio FP, Cooper D, Delia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. Eur Heart J 2011;32:1484–92.
[12] Leng Y, Cappuccio FP, Wainwright NW, Surtees PG, Luben R, Brayne C, et al. Sleep duration and risk of fatal and nonfatal stroke: a prospective study and meta-analysis. Neurology 2015;84:1072–9.
[13] Li W, Wang D, Cao S, Yin X, Gong Y, Gan Y, et al. Sleep duration and risk of stroke events and stroke mortality: a systematic review and meta-analysis of prospective cohort studies. Int J Cardiol 2016;223:870–6.
*[14] Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Quantity and quality of sleep and incidence of type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2010;33:414–20.
[15] Shan Z, Ma H, Xie M, Yan P, Guo Y, Bao W, et al. Sleep duration and risk of type 2 diabetes: a meta-analysis of prospective studies. Diabetes Care 2015;38:529–37.
*[16] Itani O, Jike M, Watanabe N, Kaneita Y. Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. Sleep Med 2017;32:246–56.
[17] Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. PLoS Med 2009;6, e1000097.
[18] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (moose) group. JAMA J Am Med Assoc 2000;283:2008–12.
[19] Zizi F, Pandey A, Murrray-Bachmann R, Vincent M, McFarlane S, Ogedegbe G, et al. Race/ethnicity, sleep duration, and diabetes mellitus: analysis of the national health interview survey. Am J Med 2012;125:162–7.
[20] Adenekan B, Pandey A, McKenzie S, Zizi F, Casimir GJ, Jean-Louis G. Sleep in America: role of racial/ethnic differences. Sleep Med Rev 2013;17:255–62.
[21] Lauderdale DS, Knutson KL, Yan LL, Rathouz PJ, Hulley SB, Sidney S, et al. Objectively measured sleep characteristics among early-middle-aged adults: the cardia study. Am J Epidemiol 2006;164:5–16.
[22] Silva GE, Goodwin JL, Sherrill DL, Arnold JL, Bootzin RR, Smith T, et al. Relationship between reported and measured sleep times: the sleep heart health study (shhs). J Clin Sleep Med JCSM Off Publ Am Acad Sleep Med 2007;3:622–30.
[23] Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa scale (nos) for assessing the quality of nonrandomised studies in meta-analyses. 2008.
[24] Aziz O, Constantinides V, Tekkis PP, Athanasiou T, Purkayastha S, Paraskeva P, et al. Laparoscopic versus open surgery for rectal cancer: a meta-analysis. Ann Surg Oncol 2006;13:413–24.
[25] Millett GA, Flores SA, Marks G, Reed JB, Herbst JH. Circumcision status and risk of HIV and sexually transmitted infections among men who have sex with men: a meta-analysis. JAMA J Am Med Assoc 2008;300:1674–84.
[26] Altman DG, Bland JM. How to obtain the confidence interval from a p value. BMJ 2011;343:d2090.
[27] Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics Med 2002;21:1539–58.
[28] Aburto NJ, Hanson S, Gutierrez H, Hooper L, Elliott P, Cappuccio FP. Effect of increased potassium intake on cardiovascular risk factors and disease: systematic review and meta-analyses. BMJ 2013;346, f1378.
[29] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.
[30] Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455–63.
[31] Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med 1992;116:78–84.
[32] Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive meta analysis. 3.3.070 ed. Englewood, NJ: Biostat.; 2014.
*[33] Guo X, Zheng L, Wang J, Zhang X, Li J, Sun Y. Epidemiological evidence for the link between sleep duration and high blood pressure: a systematic review and meta-analysis. Sleep Med 2013;14:324–32.
[34] Wu Y, Zhai L, Zhang D. Sleep duration and obesity among adults: a meta-analysis of prospective studies. Sleep Med 2014;15:1456–62.
[35] Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National sleep foundation's updated sleep duration recommendations: final report. Sleep Health 2015;1:233–43.
[36] Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep 2004;27:1255–73.
[37] Grandner MA, Drummond SP. Who are the long sleepers? Towards an understanding of the mortality relationship. Sleep Med Rev 2007;11:341–60.
[38] Kraemer HC, Kazdin AE, Offord DR, Kessler RC, Jensen PS, Kupfer DJ. Coming to terms with the terms of risk. Archives General Psychiatry 1997;54:337–43.
*[39] Wu JQ, Appleman ER, Salazar RD, Ong JC. Cognitive behavioral therapy for insomnia comorbid with psychiatric and medical conditions: a meta-analysis. JAMA Intern Med 2015;175:1461–72.
[40] Zachariae R, Lyby MS, Ritterband LM, O'Toole MS. Efficacy of internet-delivered cognitive-behavioral therapy for insomnia – a systematic review and meta-analysis of randomized controlled trials. Sleep Med Rev 2015;30:1–10.
*[41] Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014;37:9–17.
*[42] Cespedes EM, Hu FB, Redline S, Rosner B, Alcantara C, Cai J, et al. Comparison of self-reported sleep duration with actigraphy: results from the hispanic community health study/study of latinos sueno ancillary study. Am J Epidemiol 2016;183:561–73.
[43] Miller CB, Gordon CJ, Toubia L, Bartlett DJ, Grunstein RR, D'Rozario AL, et al. Agreement between simple questions about sleep duration and sleep diaries in a large online survey. Sleep Health 2015;1:133–7.
[44] Sekine M, Tatsuse T, Cable N, Chandola T, Marmot M. U-shaped associations between time in bed and the physical and mental functioning of Japanese civil servants: the roles of work, family, behavioral and sleep quality characteristics. Sleep Med 2014;15:1122–31.
[45] Li Y, Zhang X, Winkelman JW, Redline S, Hu FB, Stampfer M, et al. Association between insomnia symptoms and mortality: a prospective study of U.S. men. Circulation 2014;129:737–46.

* The most important references are denoted by an asterisk.

[46] Parthasarathy S, Vasquez MM, Halonen M, Bootzin R, Quan SF, Martinez FD, et al. Persistent insomnia is associated with mortality risk. Am J Med 2015;128:268–275e2.
[47] Grandner MA, Patel NP. From sleep duration to mortality: implications of meta-analysis and future directions. J Sleep Res 2009;18:145–7.
[48] Stranges S, Dorn JM, Shipley MJ, Kandala NB, Trevisan M, Miller MA, et al. Correlates of short and long sleep duration: a cross-cultural comparison between the United Kingdom and the United States: the Whitehall II study and the western New York Health study. Am J Epidemiol 2008;168:1353–64.
[49] Krueger PM, Friedman EM. Sleep duration in the United States: a cross-sectional population-based study. Am J Epidemiol 2009;169:1052–63.
[50] Whinnery J, Jackson N, Rattanaumpawan P, Grandner MA. Short and long sleep duration associated with race/ethnicity, sociodemographics, and socioeconomic position. Sleep 2014;37:601–11.
[51] Basner M, Spaeth AM, Dinges DF. Sociodemographic characteristics and waking activities and their role in the timing and duration of sleep. Sleep 2014;37:1889–906.
[52] Smagula SF, Koh WP, Wang R, Yuan JM. Chronic disease and lifestyle factors associated with change in sleep duration among older adults in the Singapore Chinese health study. J Sleep Res 2016;25:57–61.
[53] Tsubono Y, Fukao A, Hisamichi S. Health practices and mortality in a rural Japanese population. Tohoku J Exp Med 1993;171:339–48.
[54] Kojima M, Wakai K, Kawamura T, Tamakoshi A, Aoki R, Lin Y, et al. Sleep patterns and total mortality: a 12-year follow-up study in Japan. J Epidemiol 2000;10:87–93.
[55] Seki N. Relationships between walking hours, sleeping hours, meaningfulness of life (ikigai) and mortality in the elderly: prospective cohort study. Nippon eiseigaku zasshi [Japanese] Nihon Eiseigaku Zasshi 2001;56:535–40.
[56] Heslop P, Smith GD, Metcalfe C, Macleod J, Hart C. Sleep duration and mortality: the effect of short or long sleep duration on cardiovascular and all-cause mortality in working men and women. Sleep Med 2002;3:305–14.
[57] Mallon L, Broman JE, Hetta J. Sleep complaints predict coronary artery disease mortality in males: a 12-year follow-up study of a middle-aged Swedish population. J Intern Med 2002;251:207–16.
[58] Burazeri G, Gofin J, Kark JD. Over 8 hours of sleep–marker of increased mortality in mediterranean population: follow-up population study. Croat Med J 2003;44:193–8.
[59] Goto A, Yasumura S, Nishise Y, Sakihara S. Association of health behavior and social role with total mortality among Japanese elders in okinawa, Japan. Aging Clin Exp Res 2003;15:443–50.
[60] Amagai Y, Ishikawa S, Gotoh T, Doi Y, Kayaba K, Nakamura Y, et al. Sleep duration and mortality in Japan: the Jichi medical school cohort study. J Epidemiol 2004;14:124–8.
[61] Patel SR, Ayas NT, Malhotra MR, White DP, Schernhammer ES, Speizer FE, et al. A prospective study of sleep duration and mortality risk in women. Sleep 2004;27:440–4.
[62] Tamakoshi A, Ohno Y, Group JS. Self-reported sleep duration as a predictor of all-cause mortality: results from the JACC study, Japan. Sleep 2004;27:51–4.
[63] Ferrie JE, Shipley MJ, Cappuccio FP, Brunner E, Miller MA, Kumari M, et al. A prospective study of change in sleep duration: associations with mortality in the Whitehall II cohort. Sleep 2007;30:1659–66.
[64] Lan TY, Lan TH, Wen CP, Lin YH, Chuang YL. Nighttime sleep, Chinese afternoon nap, and mortality in the elderly. Sleep 2007;30:1105–10.
[65] Gangwisch JE, Heymsfield SB, Boden-Albala B, Buijs RM, Kreier F, Opler MG, et al. Sleep duration associated with mortality in elderly, but not middle-aged, adults in a large us sample. Sleep 2008;31:1087–96.
[66] Ikehara S, Iso H, Date C, Kikuchi S, Watanabe Y, Wada Y, et al. Association of sleep duration with mortality from cardiovascular disease and other causes for Japanese men and women: the JACC study. Sleep 2009;32:295–301.
[67] Stone KL, Ewing SK, Ancoli-Israel S, Ensrud KE, Redline S, Bauer DC, et al. Self-reported sleep and nap habits and risk of mortality in a large cohort of older women. J Am Geriatrics Soc 2009;57:604–11.
[68] Suzuki E, Yorifuji T, Ueshima K, Takao S, Sugiyama M, Ohta T, et al. Sleep duration, sleep quality and cardiovascular disease mortality among the elderly: a population-based cohort study. Prev Med 2009;49:135–41.
[69] Chien KL, Chen PC, Hsu HC, Su TC, Sung FC, Chen MF, et al. Habitual sleep duration and insomnia and the risk of cardiovascular events and all-cause death: report from a community-based cohort. Sleep 2010;33:177–84.
[70] Mesas AE, Lopez-Garcia E, Leon-Munoz LM, Guallar-Castillon P, Rodriguez-Artalejo F. Sleep duration and mortality according to health status in older adults. J Am Geriatrics Soc 2010;58:1870–7.
[71] Castro-Costa E, Dewey ME, Ferri CP, Uchoa E, Firmo JOA, Rocha FL, et al. Association between sleep duration and all-cause mortality in old age: 9-year follow-up of the bambui cohort study, Brazil. J Sleep Res 2011;20:303–10.
[72] Kripke DF, Langer RD, Elliott JA, Klauber MR, Rex KM. Mortality related to actigraphic long and short sleep. Sleep Med 2011;12:28–33.
[73] Kronholm E, Laatikainen T, Peltonen M, Sippola R, Partonen T. Self-reported sleep duration, all-cause mortality, cardiovascular mortality and morbidity in Finland. Sleep Med 2011;12:215–21.
[74] Cohen-Mansfield J, Perach R. Sleep duration, nap habits, and mortality in older persons. Sleep 2012;35:1003–9.
[75] Chen HC, Su TP, Chou P. A nine-year follow-up study of sleep patterns and mortality in community-dwelling older adults in Taiwan. Sleep 2013;36:1187–98.
[76] Garde AH, Hansen AM, Holtermann A, Gyntelberg F, Suadicani P. Sleep duration and ischemic heart disease and all-cause mortality: prospective cohort study on effects of tranquilizers/hypnotics and perceived stress. Scand J Work Environ Health 2013;39(6):550–8.
[77] Hale L, Parente V, Dowd JB, Sands M, Berger JS, Song Y, et al. Fibrinogen may mediate the association between long sleep duration and coronary heart disease. J Sleep Res 2013;22:305–14.
[78] Kakizaki M, Kuriyama S, Nakaya N, Sone T, Nagai M, Sugawara Y, et al. Long sleep duration and cause-specific mortality according to physical function and self-rated health: the Ohsaki cohort study. J Sleep Res 2013;22:209–16.
[79] Kim Y, Wilkens LR, Schembre SM, Henderson BE, Kolonel LN, Goodman MT. Insufficient and excessive amounts of sleep increase the risk of premature death from cardiovascular and other diseases: the multiethnic cohort study. Prev Med 2013;57:377–85.
[80] Li Y, Sato Y, Yamaguchi N. Potential biochemical pathways for the relationship between sleep duration and mortality. Sleep Med 2013;14:98–104.
[81] Magee CA, Holliday EG, Attia J, Kritharides L, Banks E. Investigation of the relationship between sleep duration, all-cause mortality, and preexisting disease. Sleep Med 2013;14:591–6.
[82] Yeo Y, Ma SH, Park SK, Chang SH, Shin HR, Kang D, et al. A prospective cohort study on the relationship of sleep duration with all-cause and disease-specific mortality in the Korean multi-center cancer cohort study. J Prev Med Public Health Yebang Uihakhoe chi 2013;46:271–81.
[83] Bellavia A, Akerstedt T, Bottai M, Wolk A, Orsini N. Sleep duration and survival percentiles across categories of physical activity. Am J Epidemiol 2014;179:484–91.
[84] Lee JS, Auyeung TW, Leung J, Chan D, Kwok T, Woo J, et al. Long sleep duration is associated with higher mortality in older people independent of frailty: a 5-year cohort study. J Am Med Dir Assoc 2014;15(9):649–54.
[85] Rod NH, Kumari M, Lange T, Kivimaki M, Shipley M, Ferrie J. The joint effect of sleep duration and disturbed sleep on cause-specific mortality: results from the Whitehall II cohort study. PLoS One 2014;9:e91965.
[86] Xiao Q, Keadle SK, Hollenbeck AR, Matthews CE. Sleep duration and total and cause-specific mortality in a large us cohort: interrelationships with physical activity, sedentary behavior, and body mass index. Am J Epidemiol 2014;180:997–1006.
[87] Zuurbier LA, Luik AI, Hofman A, Franco OH, Van Someren EJ, Tiemeier H. Fragmentation and stability of circadian activity rhythms predict mortality: the Rotterdam study. Am J Epidemiol 2015;181:54–63.
[88] Hall MH, Smagula SF, Boudreau RM, Ayonayon HN, Goldman SE, Harris TB, et al. Association between sleep duration and mortality is mediated by markers of inflammation and health in older adults: the health, aging and body composition study. Sleep 2015;38:189–95.
[89] Cai H, Shu XO, Xiang YB, Yang G, Li H, Ji BT, et al. Sleep duration and mortality: a prospective study of 113 138 middle-aged and elderly Chinese men and women. Sleep 2015;38:529–36.

Sleep and Biological Rhythms (2018) 16:239–251
https://doi.org/10.1007/s41105-018-0143-0

ORIGINAL ARTICLE

Sleep-related factors associated with industrial accidents among factory workers and sleep hygiene education intervention

Osamu Itani[1] · Yoshitaka Kaneita[2] · Maki Jike[2] · Misaki Furuya[3] · Chizuru Uezono[4] · Fusako Oda[5] · Ryuji Agematsu[5] · Mikiko Tokiya[1] · Yuichiro Otsuka[2] · Takashi Ohida[2]

Received: 9 April 2017 / Accepted: 9 January 2018 / Published online: 18 January 2018

Abstract

This study was conducted to investigate the association between industrial accidents and sleep-related parameters in factory workers, and to examine the effectiveness of sleep education intervention for improvement of sleep status. 714 factory workers were included in the study. A baseline survey was conducted using a self-administered questionnaire in December 2013. The questionnaire included items for the evaluation of sleep status (containing PSQI and ESS), sleep-related lifestyle habits, and experience of industrial accidents. In January 2014, workers were selected for a sleep education program that included attendance at a lecture and a take-home leaflet containing information on how to improve their sleep habits. All of the workers then participated in a follow-up survey in March 2014 to investigate the effectiveness of the education program. We first analyzed the association between industrial accidents and sleep status at the time of the baseline survey. Then, using data from the follow-up survey, we examined the effectiveness of sleep education by analyzing the differences in the improvement of sleep disorders and sleep habits between the groups who did and did not receive sleep education. We detected a significant association between the occurrence of industrial accidents and PSQI scores from the baseline survey. With regard to the effectiveness of the sleep hygiene education intervention, the percentage of early risers increased significantly in the intervention group among the participants less than 40 years of age. Among the participants aged 40 years or older, the percentage of those who did not drink an alcoholic beverage before going to sleep increased significantly in the intervention group.

Keywords Non-randomized controlled trials · Good sleep habits · Insomnia · Circadian rhythm · Occupational injuries · Epidemiology

Introduction

The diversification of labor and the expansion of the 24-h society in recent years have been increasingly leading to sleep problems [1]. In an epidemiological study of the Japanese general population, 12.2% of adult men and 14.6% of women reported symptoms of insomnia (difficulty initiating sleep, difficulty maintaining sleep with difficulty resuming sleep, and early morning awakening with difficulty resuming sleep), and 3.2% of men and 4.2% of women also suffered from daytime disorders [2]. A further study also found that a higher proportion of workers complained of insomnia compared with the general adult population [3].

Sleep disorders may constitute a risk factor for the lifestyle-related conditions of obesity [4, 5], hypertension [4, 6], glucose intolerance [4, 7], cardiovascular disease [4, 8], and depression [9, 10]. Insomnia not only causes distress during

✉ Yoshitaka Kaneita
nusmpublichealth@gmail.com

1 Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Yufu, Oita, Japan

2 Division of Public Health, Department of Social Medicine, School of Medicine, Nihon University, Itabashi-ward, Tokyo, Japan

3 Ijuin Public Health Center, Hioki, Kagoshima, Japan

4 Department of Health and Welfare, Kagoshima, Kagoshima, Japan

5 Aira Public Health Center, Kirishima, Kagoshima, Japan

the night, but also leads to impaired daytime functioning, with accumulation of fatigue and diminished memory and attention, potentially resulting in a diminished quality of life [11, 12]. Insomnia in workers, in particular, reduces working efficiency and job satisfaction and increases the incidence of work-related injuries and problems such as absenteeism, presenteeism, and early retirement, which incur both direct and indirect financial losses [13]. Many studies have reported a relationship between sleep problems and work injuries. Evidence from these studies recently underwent a systematic review [14]. This review abstracted 27 observational studies and featured a meta-analysis of 54 relative risk estimates from among those studies. The findings of the meta-analysis suggested that workers with sleep problems had a 1.62 times higher risk of being injured than workers without sleep problems. The Pittsburgh Sleep Quality Index (PSQI), developed at University of Pittsburgh School of Medicine, is a questionnaire used worldwide to assess of sleep quality [15]. It consists of 10 questions that assess different aspects of sleep quality. This questionnaire is advantageous because of its widespread use in a variety of settings and translation into many languages, including Japanese [16]. A previous study assessed a healthy (control subjects) and psychiatric disordered group (insomnia, depression, anxiety disorder, and schizophrenia) with the use of the Japanese version of PSQI and reported α reliability coefficient of 0.77, the sensitivity of 80.0–85.7%, and the specificity of 86.6% [17]. Therefore, the Japanese version of the PSQI is convenient and easy-to-use and has sufficient discrimination power.

Sleep disorders cause serious health and lifestyle-related problems, and in particular, excessive daytime sleepiness (EDS) due to a sleep disorder is often a serious problem for workers. According to a previous research, in a survey of the general population in Japan, 2.8% of males and 2.5% of females were aware of EDS [18]. In a survey of the Finnish population, it was reported that 11% of males and 7% of females were aware of daytime sleepiness almost every day [19]. In a similar research conducted in the US population, the Epworth Sleepiness Scale (ESS) [20], which is a questionnaire based on self-awareness for the assessment of EDS, was used [21]. The ESS has been validated clinically in populations of sleep disorder patients, showing 74% sensitivity and 50% specificity relative to the Multiple Sleep Latency Test (MSLT), which is an objective physiological test [22]. Therefore, the ESS has high sensitivity and specificity for the assessment of EDS. In the research that used ESS for the assessment of EDS in the US population, it was reported that 25% of individuals had ESS scores higher than 11 [21]. A reason that EDS is an important problem for workers is that EDS lowers behavioral functioning [23–26] and makes them take poor decisions [27]. Therefore, it has been reported that EDS is a cause of various traffic and occupational accidents [28–31].

Under these circumstances, investigation of measures to improve sleep in workers is an urgent issue. One means of alleviating insomnia is the use of sleeping pills and other drug therapies. Although drug therapy may be effective, it also causes problems such as drug dependency and carry-over effects, and is, therefore, problematic for workers.

Another, non-drug treatment for insomnia is sleep hygiene education directed at the population as a whole using talks and pamphlets. In a previous study, Kakinuma et al. [32] reported the effectiveness of sleep hygiene education intervention among workers. A total of 391 day-shift workers in the Japanese information technology (IT) industry were randomly allocated to either an intervention group or a control group, and the intervention group received 60 min of sleep hygiene education. However, this intervention did not generate a significant improvement in sleep quality (Pittsburgh Sleep Quality Index). There has been a paucity of epidemiological interventional studies on the effects of sleep hygiene education alone on sleep quality. Therefore, we designed a new epidemiological interventional study of sleep hygiene education to overcome the limitations of the study by Kakinuma et al. The present study was characterized by (1) a sufficient sample size to ensure high statistical reliability, (2) the preparation of new sleep hygiene education materials containing information based on evidence incorporating the latest conclusions of sleep science research, and (3) a sufficiently long follow-up period allowing the effect of the intervention to become apparent. In this study, we investigated the effect of intervention by means of sleep hygiene education and the association between sleep quality and experience of occupational accidents in the same group of subjects.

Materials and methods

Study participants and procedures

The study subjects were employees working rotating shifts in a factory in Kagoshima Prefecture, Japan, manufacturing vehicle engine parts on a 24-h basis. The sleep status and lifestyle habits of all the employees working at the factory at that point were surveyed in December 2013 (baseline survey). Study participation was completely voluntary.

The following method was employed in survey: an administrator of this company delivered (1) the instructions, (2) a self-administered questionnaire, and (3) an envelope to each worker. The workers were requested to complete the questionnaire on their own. After completing the questionnaire, each worker put it in the envelope and sealed it. The sealed

envelopes were collected and then opened for the first time at the investigating institution.

The following six common items were included in the questionnaires used in this surveys. (1) Basic attributes: the participants were required to enter their employee number, sex, employment status (regular/non-regular), job category (management/non-management), and departments in which the employees worked. (2) Working conditions: working hours per day, overtime (hours) per month, number of days off per month, and Midnight shift work (22:00–5:00). (3) Sleep status indexes: Japanese version of PSQI [15, 16] and ESS [20, 33]. (4) Practice of 12 lifestyle habits associated with sleep (yes/no): (1) eating snacks after dinner, (2) exercising, (3) taking a bath or a shower, (4) Attempting to read a book, (5) attempting to listen to music, (6) attempting to maintain a healthy lifestyle, (7) attempting to avoid coffee and tea, (8) decreasing the duration and number of times of daytime napping, (9) attempting to go to bed when feeling sleepy, (10) attempting to get up early in the morning, (11) attempting not to have an alcoholic beverage before going to sleep, (12) attempting to get early morning sunlight exposure (5) smoking habit. (6) Occupational accidents or near-miss experience per month (experienced neither event/experienced near-miss experiences only/experienced occupational accidents). Employees were asked if they had ever experienced each of the following accidents: (1) falling from a height, (2) stumbling, (3) collision, (4) being struck by a falling/airborne object, (5) being trapped under/hit by a collapsing structure, (6) being caught or entangled, (7) cuts/abrasions (cut by a box cutter, etc.), (8) directly touching a very hot or cold object (burns, etc.), (9) electric shock/fire, (10) directly touching hazardous material, (11) traffic accident, (12) rebound from bodily movement/excessive bodily movement.

In January 2014, following a baseline survey, specific employees attended a lecture on sleep hygiene education. Attendance was voluntary and limited to employees who were at work on the date of the lecture.

A follow-up survey of all the employees registered as factory workers was conducted in May 2014. In addition to the items covered in the baseline survey, the follow-up survey also included a question (yes/no format) asking (7) whether the respondent had attended the sleep hygiene education lecture in January 2014.

Intervention by sleep hygiene education

In 2013, the Japanese Ministry of Health, Labour and Welfare established a Working Group on Sleep Guidelines Revision, gathering the latest evidence from sleep science research, and published "Sleep Guidelines 2014 for Health Promotion" [34]. The document is written in simple language, aimed at the general public, and includes "12 basic sleep guidelines" designed for use in lifestyle coaching to promote health as part of sleep hygiene education. These 12 sleep guidelines are as follows: (1) good sleep leads to a healthy body and mind, (2) regular meals and routine exercise are important, (3) lack of sleep correlates strongly with lifestyle-related diseases, (4) mental health can be maintained by rest while asleep, (5) it is important to obtain an appropriate amount of sleep for an individual's age and for the time of year, (6) creating an environment conducive to sleep is important, (7) young people should avoid staying up late at night, (8) for working-age people, good sleep allows recovery from fatigue and improves productivity, (9) it is important for older people to make arrangements so that they sleep soundly, (10) go to bed when feeling sleepy, and do not wake up too late, (11) pay attention to any unusual physical changes while sleeping, and (12) consult a specialist if insomnia does not improve. The sleep hygiene education implemented in this study used a four-page leaflet describing the content of these 12 sleep guidelines in accessible, simple language. This was distributed to the audience during a lecture, comprising an hour-long talk explaining the content of the leaflet and the 12 sleep-related lifestyle habits, and providing guidance on how to improve lifestyle habits and put the guidelines into practice.

Statistical analysis

Respondents who reported attending the sleep hygiene education lecture in the follow-up survey were defined as the Intervention Group. Respondents who had not attended the lecture or did not answer that specific question were defined as the Control Group. Data from the baseline survey were then used to calculate the composition of the Intervention and Control Groups according to the following parameters: "basic attributes", "working conditions", "sleep status indices", "lifestyle habits associated with sleep", "smoking", and "occupational accidents or near-miss experiences per month". The two groups were compared using the chi-squared test.

Next, multivariate analysis was used to investigate the association between sleep quality status, level of daytime drowsiness, and experience of occupational accidents, using study group data from the baseline survey. We examined the relationship between two indicators of sleep status and occupational accidents using two different models. In one model, we used the Pittsburgh Sleep Quality Index (PSQI) [15, 16] score as an indicator of insomnia. n the other model, we used the Epworth Sleepiness Scale (ESS) [20, 33] score as an indicator of daytime disorders. This allowed us to separately examine the relationship between daytime drowsiness (daytime disorders) and occupational accidents and the relationship between insomnia (the primary sleep disorder that causes daytime drowsiness) and occupational accidents.

Logistic regression analysis was performed using the forced entry method, with the experience of an occupational accident or near-miss experience in the previous month as the objective variable (respondents who answered "experienced near-miss experiences only" or "experienced occupational accidents" were classified as having had an occupational accident or a near-miss experience, and those who answered "experienced neither event" were classified as not having had one) and PSQI score as the dependent variable (Model 1). Age, sex, employment status, job category, departments in which the employees worked, working hours per day, overtime per month, number of days off per month, midnight shift work, smoking habits, and exercise habits were inputted to the model as adjustment factors. The forced entry method was used because there was no advanced hypothesis regarding the order of importance of the dependent (adjustment) variables. The same analysis was also performed using the ESS score as the dependent variable (Model 2).

We then investigated the effectiveness of sleep hygiene education on sleep-related lifestyle habits. We calculated the proportion of people in the Intervention and Control Groups who were putting the 12 sleep-related lifestyle habits into practice at the time of the baseline survey and of the follow-up survey. We also used McNemar's test [35, 36] to compare the changes between the baseline and follow-up surveys in the Intervention and Control Groups (comparison within group), and used a two-sample McNemar's test [37, 38] to compare the Intervention and Control Groups with each other (comparison between groups). We then performed a sub-group analysis by testing according to age class to examine the difference in effect of sleep hygiene education for each age group. According to a previous study in the field of learning science, comparison of fluid intelligence and crystallized intelligence is a method for characterizing the changes in cognitive and learning functions with age [39, 40]. Fluid intelligence refers to intelligence that reflects items such as calculation ability, concentration, and intelligence quotient, and fluid intelligence decreases after the age of 20 years, and in particular, decreases quickly after the age of 40 years. Crystallized intelligence reflects intelligence that is accumulated with experience, such as empirical knowledge and decision-making, and this type of intelligence improves with age. Based on these previous studies in the field of learning science, we performed our analysis by stratifying participants into two groups, using the age of 40 as the boundary.

Finally, we investigated the effectiveness of sleep hygiene education for improving sleep quality and daytime drowsiness. The proportions of respondents with a PSQI score of ≥ 6 in the baseline and follow-up surveys were calculated separately for the Intervention and Control Groups, and McNemar's test was used to investigate the difference between the scores in the baseline survey and follow-up survey (comparison within group). We also used a two-sample McNemar's test to compare the Intervention and Control Groups with each other (comparison between groups). The same analysis was performed for an ESS score of ≥ 11 or occupational accidents/near-miss experience per month.

We set the level of significance at $P < 0.05$. All analyses were performed using SPSS 22 for Windows (IBM Corp, Armonk, NY, USA).

Informed consent and ethical considerations

The following instructions were provided to the workers: (1) this survey is part of a medical study. What is written in the questionnaire will not be subject to evaluation related to participants' working performance or punishment; (2) subjects' participation in the survey must be voluntary, and subjects who do not participate will not suffer any loss; (3) the completed questionnaires will not be seen by other workers; (4) subjects' privacy will be strictly protected. In this survey, permission to conduct the survey was obtained from the Ethics Committee of Nihon University school of Medicine and the Ethics Committee of Oita University faculty of Medicine.

Results

An overall flow diagram of this survey is shown in Fig. 1. In the baseline survey, responses were received from 798 out of a total of 800 registered employees (response rate 99.8%). The sleep hygiene education program was attended by 287 employees. In the follow-up survey, responses were received from 783 registered employees. The analysis set comprised data from 714 individuals, including participants who had completed both the baseline and follow-up surveys (complete follow-up rate 89.3%). The following individuals ($n = 69$) were excluded due to missing answers, lack of consent to participate in the study, inability to consolidate data (according to lack of employee number), or lack of basic data on sex and/or age. Overall, in the follow-up survey, 104 individuals answered that they had attended the sleep hygiene education lecture (Intervention Group). The Control Group included 217 individuals who answered no participation of the sleep hygiene education lecture, and 393 individuals who did not answered whether or not they had attended the sleep hygiene education lecture.

Table 1 shows the profiles of employees at the time of the baseline survey (Control and Intervention Groups). The analysis set included 596 men (83.5%) and 118 women (16.5%) with an age range of 18–62 years (mean age 38.4 ± 11.6 years). Out of the total, 62.9% were engaged in midnight shift work. The PSQI score was ≥ 6 in 38.0% of the participants, and the ESS score was ≥ 11 in

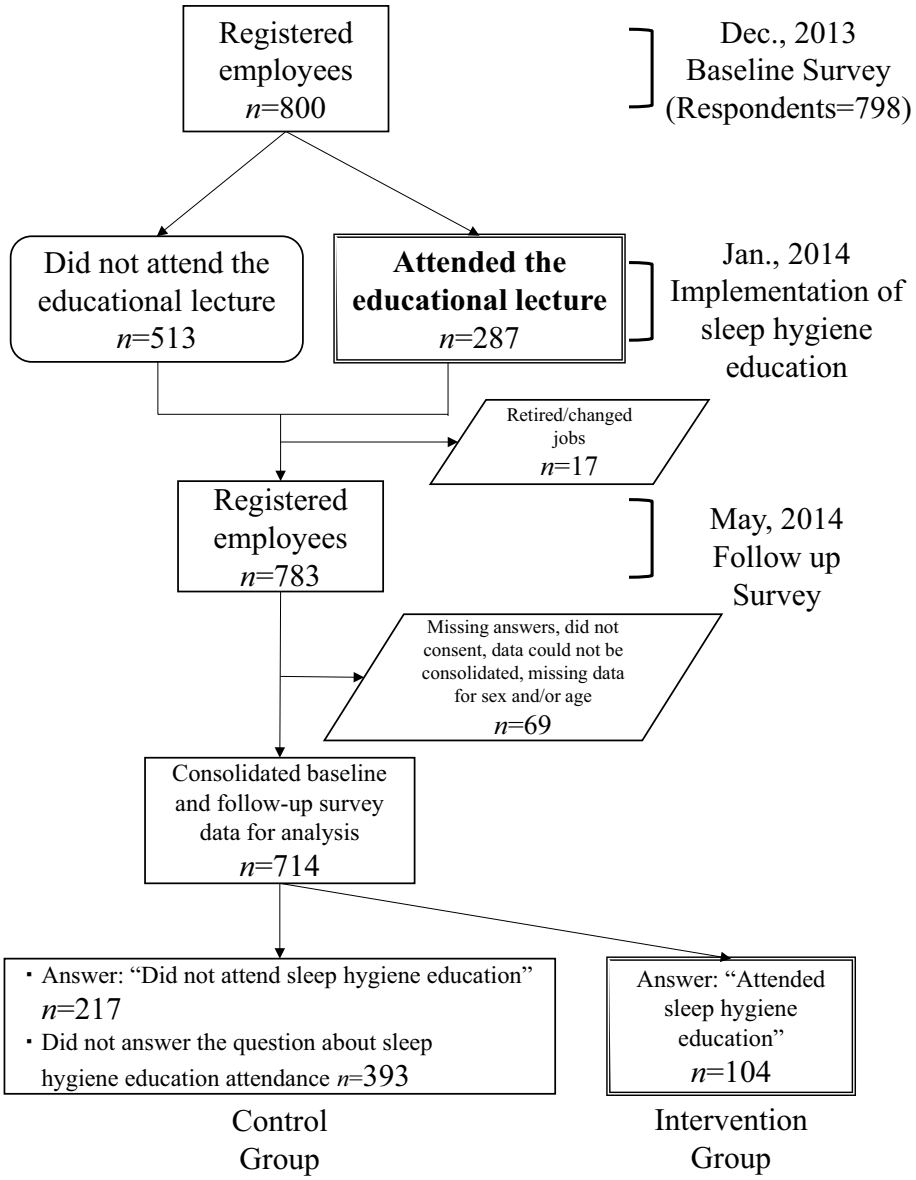


Fig. 1 Overall-flow diagram of this study: baseline and follow-up survey, and sleep hygiene education intervention

26.9%. Comparisons between the Control and Intervention Groups for each parameter using chi-squared test found no significant differences.

Table 2 shows the associations between the occurrence of occupational accidents/near-miss experiences and PSQI/ESS scores found in the baseline cross-sectional study. In Model 1, with the PSQI score input as a dependent variable, the occurrence of occupational accidents/near-miss experiences was significantly associated with exercise habits ($P = 0.034$) or PSQI score ($P = 0.009$). There was no significant association with any other dependent variable. In Model 2, with the ESS score input as a dependent variable, the occurrence of occupational accidents/near-miss experiences was significantly associated with age group ($P = 0.012$) and exercise habits ($P = 0.011$). There was no significant association with any other dependent variable, including the ESS score ($P = 0.204$).

In a comparison of chronological changes in sleep-related lifestyle habits in the Intervention and Control Groups, an investigation of the entire population found no significant difference in lifestyle habits in intragroup (baseline vs. follow-up) or intergroup (Control Group vs. Intervention Group) comparisons. We then performed a sub-group analysis stratified by age group, and found that among employees aged < 40 years (Table 3) in the Intervention Group, significantly more employees were attempting to wake up early in the morning in the follow-up survey compared with the proportion at the baseline ($P = 0.031$). This difference was not observed in the Control Group ($P = 0.162$). There was also a significant

Table 1 Profile of Employees at baseline survey (control and intervention group)

	Control group ($n=610$)	Intervention group (%) ($n=104$)	*P*
Sex			0.422
Male	83.9	80.8	
Female	16.1	19.2	
Age class			0.068
<30 years	26.9	33.7	
30–39 years	22.8	11.5	
40–49 years	26.9	28.8	
≥50 years	23.4	26.0	
Employment status			0.770
Regular	92.5	92.3	
Non-regular	1.8	1.0	
Unknown	5.7	6.7	
Job category			0.739
Management	3.0	1.9	
Non-management	86.9	89.4	
Unknown	10.2	8.7	
Working hours per day			0.502
<8 h/day	26.4	24.0	
≥8 h/day or < 10 h/day	67.5	69.2	
≥10 h/day	4.9	3.8	
Unknown	1.1	2.9	
Overtime per month			0.677
<20 h/month	53.8	48.1	
≥20 h/month or < 45 h/month	34.6	39.4	
≥45 h/month or < 80 h/month	1.1	1.0	
≥80 h/month	0.8	0.0	
Unknown	9.7	11.5	
Number of days off per month			0.352
≤4 days	23.1	22.1	
5–8 days	57.7	62.5	
9–12 days	17.0	11.5	
≥13 days	0.2	1.0	
Unknown	2.0	2.9	
Midnight shift work (22:00–5:00)			0.172
No	63.6	58.7	
Yes	32.6	33.7	
Unknown	3.8	7.7	
Sleep duration			0.334
<5 h/day	5.6	7.7	
≥5 h/day or < 7 h/day	56.7	46.2	
≥7 h/day or < 9 h/day	35.4	42.3	
≥9 h/day	1.0	1.9	
Unknown	1.3	1.9	
PSQI score			0.330
≤5 points	53.1	58.7	
≥6 points	38.2	36.5	
Unknown	8.7	4.8	
ESS score			0.821
≤10 points	71.0	71.2	
≥11 points	27.0	26.0	

Table 1 (continued)

	Control group ($n=610$)	Intervention group (%) ($n=104$)	*P*
Unknown	2.0	2.9	
Nightcap habit			0.813
None	73.9	75.0	
Less than once a week	6.7	5.8	
1–2 times per week	5.1	2.9	
3 or more times per week	13.1	14.4	
Unknown	1.1	1.9	
Smoking habit			0.818
None	53.0	50.0	
Not having smoked for ≥ 1 month	3.3	1.9	
Sometimes	1.6	1.0	
Everyday	39.2	43.3	
Unknown	3.0	3.8	
Exercise Habit			0.907
Very often	4.9	2.9	
Sometimes	24.8	26.0	
Seldom	34.9	33.7	
Rarely	31.1	32.7	
Unknown	4.3	4.8	
Occupational accidents or			0.930
Near-miss experience per month			
Experienced neither event	64.6	65.4	
Experienced near-miss experiences only	23.9	24.0	
Experienced occupational accidents	8.4	6.7	
Unknown	3.1	3.8	

We have combined all the participants in the baseline study ($n=714$)

PSQI Pittsburgh Sleep Quality Index, *ESS* Epworth sleepiness Scale

P was calculated by χ^2 test

difference between the two groups (Control vs. Intervention) ($P=0.005$). Furthermore, among employees aged ≥ 40 years (Table 4) in the Intervention Group, significantly more employees reported attempts not to consume alcohol before going to sleep in the follow-up survey compared with the proportion at the baseline ($P=0.002$). This difference was not observed in the Control group ($P=.883$). There was also a significant difference between the two groups (Control vs. Intervention) ($P=0.005$).

Table 5 shows a comparison of chronological changes in PSQI, ESS scores, and Occupational accidents/Near-miss experience per month in the Intervention and Control groups. No significant difference was evident in the PSQI, ESS scores, or Occupational accidents/Near-miss experience in all intragroup (baseline vs. follow-up) or intergroup (Control vs. Intervention) comparisons.

Discussion

In this study, we investigated whether or not the experience of accidents in the workplace was related to sleep disorder index scores (PSQI or ESS) and whether intervention by means of sleep hygiene education was effective for improving sleep-related lifestyle habits and sleep disorder index scores in employees engaged in rotational shift work. The larger sample size and longer follow-up period for evaluating the effectiveness of intervention, as well as the high response and follow-up rates, make this study more reliable than previous studies.

In this study, we found a significant association between a high PSQI score, indicating low sleep quality, and the experience of accidents. A number of previous studies have also reported that sleep disorders are associated with

Table 2 Associations between the occurrence of occupational accidents/near-miss experiences and PSQI/ESS scores found in the baseline cross-sectional study

	AOR	95% CI	*P*
Model 1			
Age class			0.067
<30 years	1.00		
30–39 years	0.82	0.47–1.44	
40–49 years	0.51	0.28–0.91	
≥50 years	0.48	0.25–0.95	
Exercise Habit			0.034
Very often	1.00		
Sometimes	6.88	1.45–32.69	
Seldom	8.36	1.79–39.15	
Rarely	9.72	2.05–46.17	
PSQI score			0.009
≤5 points	1.00		
≥6 points	1.76	1.15–2.68	
Model 2			
Age class			0.012
<30 years	1.00		
30–39 years	0.77	0.46–1.29	
40–49 years	0.47	0.27–0.80	
≥50 years	0.42	0.22–0.78	
Exercise habit			0.011
Very often	1.00		
Sometimes	4.84	1.32–17.81	
Seldom	6.55	1.80–23.75	
Rarely	7.79	2.13–28.56	
ESS score			0.204
≤10 points	1.00		
≥11 points	1.31	0.87–1.98	

Data at the time of the baseline survey were used

P was calculated by logistic regression analysis (Forced Entry Method)

Missing data were excluded from the statistical analyses (Model 1: $n=490$, Model 2: $n=567$)

PSQI Pittsburgh Sleep Quality Index, *ESS* Epworth sleepiness Scale, *AOR* Adjusted Odds Ratio, *CI* Confidential Interval, *Objective variable* occupational accidents or near-miss experience per month, *Adjusted factors* sex, employment status, job category, departments in which the employees worked, working hours per day, overtime per month, number of days off per month, midnight shift work, smoking habit

accidents. A study of public transport drivers found that the incidence of traffic accidents increased proportionally to the level of subjective drowsiness [41]. Individuals sleeping for < 6 h have been found to fall asleep at the wheel more frequently than those sleeping for a minimum of 7 h [42]. It has also been shown that among drivers causing accidents, those with < 6 h of sleep were more likely to have been involved in a rear-end collision with another vehicle or involved in a single-vehicle accident [43]. Insomnia is associated not only with sleep problems at night, but also physical and mental problems such as fatigue, reduced concentration, drowsiness, and reduced motivation, during the day. A study in the United States found that insomnia accompanied by these daytime symptoms increases the risk of injury [44]. Our results were also consistent with these findings. The logistic regression analysis in this study considered two models. Although the PSQI and ESS are separate indicators that assess insomnia and daytime disorders, respectively, the PSQI includes daytime disorders as one of its assessment indicators, and thus duplicates the primary object of assessment of the ESS. Therefore, entering both indicators as dependent variables in the same model could lead to strong multicollinearity and inappropriate analysis results. We calculated Spearman's rank correlation coefficient to examine the relationship between PSQI and ESS scores, Spearman's rank correlation coefficient between the PSQI and ESS scores was $r=0.261$, $P<0.001$, thus demonstrating a significant correlation between them. Consequently, strong multi-collinearity would be expected when entering both the PSQI score and ESS score as explanatory variables in the same model. Therefore, we deemed it appropriate to enter them into separate models.

In the present study, we did not find a significant correlation between the ESS score (the indicator of daytime disorders) and experience of occupational accidents. Several previous studies had reported a relationship between daytime sleepiness and work injury [45–47]. However, in a systematic review of sleep problems and accidents, Uehli et al. reported that the association between "daytime sleepiness" and work injuries was weak [14] They discussed whether this could be explained by sleepy people being aware of their limitations at the time of risk and adopting coping strategies. A similar reason could also explain the results of the present study. As the ESS is a self-administered questionnaire, it is also possible that the person may not have an accurate grasp of their sleepiness. Going forward, more research needs to be conducted.

In the present study, we found that when sleep hygiene education was provided in the form of a lecture, this resulted in improvements in some lifestyle habits in specific age groups; however, there was no significant improvement in the PSQI or ESS scores, which are indicators of overall sleep status. Previous studies also found no significant improvement in sleep quality resulting from intervention in the form of sleep hygiene education alone [32]. One possible reason for the absence of any effect may have been the education method used. In our study, we used a simple lecture format as the main method of sleep hygiene education. Although the lecture format has the advantage of transmitting knowledge to a large number of people at the same time, in the field of

Table 3 Comparison of chronological changes in sleep-related lifestyle habits in the Control and Intervention groups (among those aged less than 40 years)

	Group	*n*	Baseline (%)	Follow-up (%)	Change from baseline (%)	[a]*P* (within group)	[b]*P* (Control vs. Intervention)
1. Eating snacks after dinner	Control	294	25.5	28.6	3.1	0.306	0.464
	Intervention	45	24.4	22.2	− 2.2	1.000	
2. Exercising	Control	295	16.9	22.0	5.1	0.050	0.832
	Intervention	45	17.8	24.4	6.7	0.581	
3. Taking a bath or a shower	Control	294	34.0	33.3	− 0.7	0.912	0.355
	Intervention	45	20.0	26.7	6.7	0.549	
4. Attempting to read a book	Control	294	13.9	13.3	− 0.7	0.885	0.298
	Intervention	45	4.4	8.9	4.4	0.625	
5. Attempting to listen to music	Control	294	28.6	26.5	− 2.0	0.519	0.688
	Intervention	45	17.8	13.3	− 4.4	0.727	
6. Attempting to maintain a healthy lifestyle	Control	294	33.3	31.0	− 2.4	0.470	0.785
	Intervention	45	26.7	26.7	0.0	1.000	
7. Attempting to avoid coffee and tea	Control	294	18.0	20.1	2.0	0.496	0.848
	Intervention	45	11.1	13.3	2.2	1.000	
8. Decreasing the duration and number of times of daytime napping	Control	293	19.5	14.7	− 4.8	0.066	0.725
	Intervention	45	11.1	8.9	− 2.2	1.000	
9. Attempting to go to bed when feeling sleepy	Control	294	43.2	41.5	− 1.7	0.668	0.378
	Intervention	45	31.1	37.8	6.7	0.629	
10. Attempting to get up early in the morning	Control	294	21.8	17.7	− 4.1	0.162	0.005
	Intervention	45	6.7	20.0	13.3	0.031	
11. Attempting not to have an alcoholic beverage before going to sleep	Control	292	23.3	21.6	− 1.7	0.653	0.514
	Intervention	45	22.2	15.6	− 6.7	0.549	
12. Attempting to get early morning sunlight exposure	Control	294	11.9	12.6	0.7	0.868	0.070
	Intervention	45	4.4	13.3	8.9	0.125	

Missing data were excluded from the statistical analyses

[a]*P* was calculated by McNemar's test

[b]*P* was calculated by two-sample McNemar's test

school education it has long been known that educational effectiveness drops dramatically when students are listening to content delivered unilaterally by a lecturer [48, 49]. Recently, school education has started to incorporate problem-based learning and other forms of learning that involve active participation by students in educational activities, to improve educational effectiveness. These strategies have been found to be effective [50, 51]. Future studies to investigate the effectiveness of sleep hygiene education incorporating active learning are required. A second possible reason for these findings may be problems with the type of subject group for which the interventional approach was used. We utilized the so-called "population approach" [52], involving the entire group in our sleep hygiene education intervention. However, it is possible to adopt a "high-risk approach" [53], where education is focused intensively on the group with the most severe condition, and this strategy may be more effective. Future studies are also required to investigate the effect of different approaches in the field of sleep hygiene education. A third issue is the number and content of educational opportunities. The sleep hygiene education consisted of a single lecture delivered to the study group. It is possible that there were variations in the relative importance of the content to the individuals in the study group at the time the lecture was delivered. It is important for individuals to adopt and persist with healthy lifestyle behavior to prevent and treat lifestyle-related diseases and many other chronic conditions. They must also modify behaviors that have a negative effect on health, and must maintain these modifications. Maintaining both types of behavioral changes would lead to improved health. Health behavior is defined as "any behavior undertaken with the aim of maintaining or promoting health or recovering from illness" [54]. Health behavior can be divided into the following five stages: (1) precontemplation stage, (2) contemplation stage, (3) preparation stage, (4) action stage, and (5) maintenance stage [55]. These five stages are not a one-way progression from (1) to (5); if an individual is unsuccessful halfway through the process, they

Table 4 Comparison of chronological changes in sleep-related lifestyle habits in the Control and Intervention groups (among those aged 40 years or more)

	Group	*n*	Baseline (%)	Follow-up (%)	Change from baseline (%)	*P* (within group)[a]	*P* (Control vs. Intervention)[b]
1. Eating snacks after dinner	Control	282	22.0	19.9	− 2.1	0.504	0.056
	Intervention	55	14.5	25.5	10.9	0.146	
2. Exercising	Control	282	13.8	13.8	0.0	1.000	0.339
	Intervention	55	14.5	10.9	− 3.6	0.625	
3. Taking a bath or a shower	Control	283	30.7	27.6	− 3.2	0.349	0.495
	Intervention	55	25.5	18.2	− 7.3	0.388	
4. Attempting to read a book	Control	281	12.8	13.2	0.4	1.000	0.373
	Intervention	55	16.4	20.0	3.6	0.625	
5. Attempting to listen to music	Control	281	10.3	14.2	3.9	0.109	0.297
	Intervention	55	14.5	12.7	− 1.8	1.000	
6. Attempting to maintain a healthy lifestyle	Control	280	26.8	31.4	4.6	0.118	0.118
	Intervention	54	24.1	18.5	− 5.6	0.508	
7. Attempting to avoid coffee and tea	Control	282	16.7	14.5	− 2.1	0.451	0.474
	Intervention	55	10.9	12.7	1.8	1.000	
8. Decreasing the duration and number of times of daytime napping	Control	277	10.1	11.9	1.8	0.499	0.745
	Intervention	55	16.4	16.4	0.0	1.000	
9. Attempting to go to bed when feeling sleepy	Control	283	43.8	41.7	− 2.1	0.571	1.000
	Intervention	54	46.3	44.4	− 1.9	1.000	
10. Attempting to get up early in the morning	Control	279	20.8	21.5	0.7	0.890	0.929
	Intervention	54	24.1	24.1	0.0	1.000	
11. Attempting not to have an alcoholic beverage before going to sleep	Control	276	15.9	16.7	0.7	0.883	0.005
	Intervention	55	7.3	25.5	18.2	0.002	
12. Attempting to get early morning sunlight exposure	Control	280	12.5	16.8	4.3	0.067	1.000
	Intervention	55	10.9	16.4	5.5	0.508	

Missing data were excluded from the statistical analyses

[a] *P* was calculated by McNemar's test

[b] *P* was calculated by two − sample McNemar's test

Table 5 Comparison of chronological changes in PSQI/ESS scores and experienced occupational accidents or Near-miss in the Control and Intervention groups

	Group	*n*	Baseline (%)	Follow-up (%)	Change from baseline (%)	*P* (within group)[a]	*P* (Control vs. Intervention)[b]
PSQI score ≥ 6 points	Control	544	41.5	39.2	− 2.4	0.259	0.218
	Intervention	96	37.5	39.6	2.1	0.839	
ESS score ≥ 11 points	Control	583	27.8	26.9	− 0.9	0.718	0.376
	Intervention	101	26.7	31.7	5.0	0.359	
Occupational accidents or near-miss experience per month	Control	579	33.2	31.8	− 1.4	0.562	0.669
	Intervention	100	32.0	28.0	− 4.0	0.571	

Missing data were excluded from the statistical analyses

PSQI Pittsburgh Sleep Quality Index, *ESS* Epworth sleepiness Scale

[a] *P* was calculated by McNemar's test

[b] *P* was calculated by two-sample McNemar's test

will revert to the earlier stages. However, continued effective health coaching is believed to result in progress. As the behavior, knowledge, and mindset required at each stage differ, a suitable approach for each stage is needed. Future studies of sleep hygiene education must also address the diverse approaches required for individuals at each different stage.

In the present research, we studied the changes in lifestyles related to sleep that resulted from sleep hygiene education for each age group, and we showed a significant improvement in different lifestyles for the under 40 and the 40-and-over year age groups. There are some previous studies that assess aging influence on sleep-related lifestyles. Changes were noted in the required amount of sleep and aging. Ohayon et al. performed a meta-analysis of 3577 individuals in 65 papers investigating night sleep with the use of polysomnographic data. They reported a decrease in total sleep time (approximately 7 h at 25 years of age compared with <6 h after 65 years); increase in the frequency of wake after sleep onset; decrease in sleep efficiency, which was particularly significant at >40 years of age; and decrease in slow wave sleep with age [56]. A previous study assessed the relationship between insomnia and age. Lichstein et al. [57] reviewed 20 studies and reported that the prevalence and severity of insomnia were associated with age in 60% of studies. These authors found strong evidence for increased "Difficulty Maintaining Sleep" with age, but modest evidence for increased "Difficulty Initiating Sleep" and "Early Morning Awakening" with age. These authors found strong evidence for increased "Difficulty Maintaining Sleep" with age, but modest evidence for increased "Difficulty Initiating Sleep" and "Early Morning Awakening" with age. These changes may be related to the age-related differences in the effect of sleep hygiene education.

Several aspects of our findings limit generalization and, therefore, warrant mention. First, in this study, we were unable to perform a completely random allocation when dividing subjects into the Intervention and Control Groups, as only those individuals who were at work on the day of the lecture and who wished to attend were allocated to the Intervention Group. Therefore, in the process of group allocation, there may have been some bias in terms of factors such as willingness to learn, as this would have depended on whether or not individuals wanted to attend the lecture. Second, as self-administered questionnaires were used, some data obtained would have been subjective. For example, information on working hours or number of days off per month was based on self-reported data, and there may have been some discrepancies between the reported data and the actual status of individuals. Third, we did not screen for sleep problems, such as sleep apnea syndrome, when recruiting participants of this study. Therefore, our study groups may include individuals who were suffering from sleep problems. Fourth, the relationship between sleep-related lifestyles (or sleep hygiene education) and frequency of accidents could not be verified in this study. Of the conceivable causative factors, there might have been an insufficient period of assessment or inadequate assessment to judge whether lifestyle has improved. A future study that includes the above-mentioned parameters will need to be conducted to verify the present study.

Our results suggested that the sleep quality is associated with the frequency of workplace accidents, and that sleep hygiene education may improve some lifestyle habits. In the present study, results of the sub-group analysis of lifestyle changes associated with sleep hygiene education differed by age group. This suggests that sleep hygiene education may yield different changes in lifestyle habits in different age groups. Further studies are required to increase the effectiveness of interventions by addressing the content and methods of sleep hygiene education and the selection of subjects. Further study is also needed to determine factors that would increase motivation to participate in sleep hygiene education.

Acknowledgements This study was funded by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 25460816.

Compliance with ethical standards

Research involving human participants and/or animals None.

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

References

1. Calem M, Bisla J, Begum A, Dewey M, Bebbington PE, Brugha T, Cooper C, Jenkins R, Lindesay J, McManus S, Meltzer H, Spiers N, Weich S, Stewart R. Increased prevalence of insomnia and changes in hypnotics use in England over 15 years: analysis of the 1993, 2000, and 2007 National Psychiatric Morbidity Surveys. Sleep. 2012; 35:377–84.
2. Itani O, Kaneita Y, Munezawa T, Mishima K, Jike M, Nakagome S, Tokiya M, Ohida T. Nationwide epidemiological study of insomnia in Japan. Sleep Med. 2016;25:130–8.
3. Kim K, Uchiyama M, Okawa M, Liu X, Ogihara R. An epidemiological study of insomnia among the Japanese general population. Sleep. 2000; 23:41–7.
4. Itani O, Jike M, Watanabe N, Kaneita Y. Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. Sleep Med. 2017; 32: 246–56.
5. Itani O, Kaneita Y, Murata A, Yokoyama E, Ohida T. Association of onset of obesity with sleep duration and shift work among Japanese adults. Sleep Med. 2011;12:341–5.
6. Gangwisch JE, Heymsfield SB, Boden-Albala B, Buijs RM, Kreier F, Pickering TG, Rundle AG, Zammit GK, Malaspina D. Short sleep duration as a risk factor for hypertension analyses of the first national health and nutrition examination survey. Hypertension. 2006;47:833–9.
7. Chaput JP, Despres JP, Bouchard C, Astrup A, Tremblay A. Sleep duration as a risk factor for the development of type 2

diabetes or impaired glucose tolerance: analyses of the Quebec Family Study. Sleep Med. 2009;10:919–24.
8. Hamazaki Y, Morikawa Y, Nakamura K, Sakurai M, Miura K, Ishizaki M, Kido T, Naruse Y, Suwazono Y, Nakagawa H. The effects of sleep duration on the incidence of cardiovascular events among middle-aged male workers in Japan. Scand J Work Environ Health. 2011;37:411–7.
9. Buysse DJ, Angst J, Gamma A, Ajdacic V, Eich D, Rossler W. Prevalence, course, and comorbidity of insomnia and depression in young adults. Sleep. 2008;31:473–80.
10. Yokoyama E, Kaneita Y, Saito Y, Uchiyama M, Matsuzaki Y, Tamaki T, Munezawa T, Ohida T. Association between depression and insomnia subtypes: a longitudinal study on the elderly in Japan. Sleep. 2010;33:1693–702.
11. Leger D, Scheuermaier K, Philip P, Paillard M, Guilleminault C. SF-36: evaluation of quality of life in severe and mild insomniacs compared with good sleepers. Psychosom Med. 2001;63:49–55.
12. Shekleton JA, Flynn-Evans EE, Miller B, Epstein LJ, Kirsch D, Brogna LA, Burke LM, Bremer E, Murray JM, Gehrman P, Lockley SW, Rajaratnam SM. Neurobehavioral performance impairment in insomnia: relationships with self-reported sleep and daytime functioning. Sleep. 2014;37:107–16.
13. Sarsour K, Kalsekar A, Swindle R, Foley K, Walsh JK. The association between insomnia severity and healthcare and productivity costs in a health plan sample. Sleep. 2011;34:443–50.
14. Uehli K, Mehta AJ, Miedinger D, Hug K, Schindler C, Holsboer-Trachsler E, Leuppi JD, Kunzli N. Sleep problems and work injuries: a systematic review and meta-analysis. Sleep Med Rev. 2014;18:61–73.
15. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiat Res. 1989;28:193–213.
16. Doi Y, Minowa M, Uchiyama M, Okawa M. Development of the Japanese version of the Pittsburgh sleep quality index. Jpn J Psychiatry Treat. 1998;13:755–63.
17. Doi Y, Minowa M, Uchiyama M, Okawa M, Kim K, Shibui K, Kamei Y. Psychometric assessment of subjective sleep quality using the Japanese version of the Pittsburgh Sleep Quality Index (PSQI-J) in psychiatric disordered and control subjects. Psychiat Res. 2000;97:165–72.
18. Kaneita Y, Ohida T, Uchiyama M, Takemura S, Kawahara K, Yokoyama E, Miyake T, Harano S, Suzuki K, Yagi Y, Kaneko A, Tsutsui T, Akashiba T. Excessive daytime sleepiness among the Japanese general population. J Epidemiol. 2005;15:1–8.
19. Hublin C, Kaprio J, Partinen M, Heikkila K, Koskenvuo M. Daytime sleepiness in an adult, Finnish population. J Intern Med. 1996; 239:417–23.
20. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991;14:540–5.
21. Punjabi NM, Bandeen-Roche K, Young T. Predictors of objective sleep tendency in the general population. Sleep. 2003;26:678–83.
22. Johns MW. Sleepiness in different situations measured by the Epworth Sleepiness Scale. Sleep. 1994;17:703–10.
23. Carskadon MA, Dement WC. Effects of total sleep loss on sleep tendency. Percept Mot Skills. 1979;48:495–506.
24. Nicholson AN, Stone BM. Antihistamines: impaired performance and the tendency to sleep. Eur J Clin Pharmacol. 1986;30:27–32.
25. Roehrs T, Kribbs N, Zorick F, Roth T. Hypnotic residual effects of benzodiazepines with repeated administration. Sleep. 1986;9:309–16.
26. Dinges DF, Pack F, Williams K, Gillen KA, Powell JW, Ott GE, Aptowicz C, Pack AI. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. Sleep. 1997;20:267–77.
27. Roehrs T, Greenwald M, Roth T. Risk-taking behavior: effects of ethanol, caffeine, and basal sleepiness. Sleep. 2004;27:887–93.
28. Broughton R, Ghanem Q, Hishikawa Y, Sugita Y, Nevsimalova S, Roth B. Life effects of narcolepsy in 180 patients from North America, Asia and Europe compared to matched controls. The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques. 1981;8:299–304.
29. Findley LJ, Unverzagt ME, Suratt PM. Automobile accidents involving patients with obstructive sleep apnea. Am Rev Respir Dis. 1988;138:337–40.
30. Mitler MM, Carskadon MA, Czeisler CA, Dement WC, Dinges DF, Graeber RC. Catastrophes, sleep, and public policy: consensus report. Sleep. 1988;11:100–9.
31. Dorrian J, Tolley C, Lamond N, van den Heuvel C, Pincombe J, Rogers AE, Drew D. Sleep and errors in a group of Australian hospital nurses at work and during the commute. Appl Ergon. 2008;39:605–13.
32. Kakinuma M, Takahashi M, Kato N, Aratake Y, Watanabe M, Ishikawa Y, Kojima R, Shibaoka M, Tanaka K. Effect of brief sleep hygiene education for workers of an information technology company. Ind Health. 2010;48:758–65.
33. Takegami M, Suzukamo Y, Wakita T, Noguchi H, Chin K, Kadotani H, Inoue Y, Oka Y, Nakamura T, Green J, Johns MW, Fukuhara S. Development of a Japanese version of the Epworth Sleepiness Scale (JESS) based on item response theory. Sleep Med. 2009;10:556–65.
34. Sleep Guidelines 2014 for Health Promotion. Ministry of Health and Welfare, Health Service Bureau. 2014. http://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/0000047221.pdf [**updated 2014; cited 2017 Dec. 4**].
35. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 1947;12:153–7.
36. Berenson ML, Koppel NB. Why McNemar's procedure needs to be included in the business statistics curriculum. Decis Sci J Innov Educ. 2005;3:125–36.
37. Breslow NE, Day NE. Statistical methods in cancer research. Volume I - The analysis of case-control studies. Lyon: IARC Scientific Publications; 1980. pp. 5–338.
38. Adedokun OA, Burgess WD. Analysis of paired dichotomous data: a gentle introduction to the McNemar Test in SPSS. J Multidiscip Eval. 2011;8:125–31.
39. Horn JL, Cattell RB. Age differences in fluid and crystallized intelligence. Acta Psychol. 1967;26:107–29.
40. Baltes PB, Staudinger UM. Wisdom. A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. Am Psychol. 2000;55:122–36.
41. Ozer C, Etcibasi S, Ozturk L. Daytime sleepiness and sleep habits as risk factors of traffic accidents in a group of Turkish public transport drivers. Int J Clin Exp Med. 2014;7:268–73.
42. Maia Q, Grandner MA, Findley J, Gurubhagavatula I. Short and long sleep duration and risk of drowsy driving and the role of subjective sleep insufficiency. Accid Anal Prev. 2013;59:618–22.
43. Abe T, Komada Y, Nishida Y, Hayashida K, Inoue Y. Short sleep duration and long spells of driving are associated with the occurrence of Japanese drivers' rear-end collisions and single-car accidents. J Sleep Res. 2010;19:310–6.
44. Kessler RC, Berglund PA, Coulouvrat C, Fitzgerald T, Hajak G, Roth T, Shahly V, Shillington AC, Stephenson JJ, Walsh JK. Insomnia, comorbidity, and risk of injury among insured Americans: results from the America Insomnia Survey. Sleep. 2012;35:825–34.
45. Akerstedt T. Work hours and sleepiness. Neurophysiologie. 1995;25:367–75.
46. Dinges DF. An overview of sleepiness and accidents. J Sleep Res. 1995;4:4–14.

47. Philip P, Akerstedt T. Transport and industrial safety, how are they affected by sleepiness and sleep restriction?. Sleep Med Rev. 2006;10:347–56.
48. Dale E. The cone of experience. Audio Visual Methods Teach. 1946;1:37–51.
49. Bloom BS. Taxonomy of educational objectives: the classification of educational goals: by a Committee of College and University Examiners: handbook 1. Philadelphia: David McKay; 1969.
50. Hoellwarth C, Moelter MJ. The implications of a robust curriculum in introductory mechanics. Am J Phys. 2011;79:540–5.
51. Olson S, Riordan DG. Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. Executive Office of the President. 2012.
52. Rose G. The strategy of preventive medicine. Oxford: Oxford University Press; 1992.
53. Rose G. Sick individuals and sick populations. Int J Epidemiol. 1985;14:32–8.
54. Gochman DS. Labels, systems and motives: some perspectives for future research and programs. Health Educ Q. 1982;9:263–70.
55. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot. 1997;12:38–48.
56. Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep. 2004;27:1255–73.
57. Lichstein KL, Durrence HH, Riedel BW, Taylor DJ, Bush AJ. Epidemiology of sleep: Age, gender, and ethnicity. Hove: Psychology Press; 2013.

Journal of Psychiatric Research 102 (2018) 72–80

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

Omega-3 fatty acids for a better mental state in working populations - Happy Nurse Project: A 52-week randomized controlled trial

Norio Watanabe[a,*], Yutaka Matsuoka[b], Mie Kumachi[c], Kei Hamazaki[d], Masaru Horikoshi[e], Toshi A. Furukawa[a]

[a] Department of Health Promotion and Human Behavior and of Clinical Epidemiology, Kyoto University Graduate School of Medicine / School of Public Health, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
[b] Division of Health Care Research, Center for Public Health Sciences, National Cancer Center, 5-1-1 Tsukiji, Chuo-ku, 104-0045 Tokyo, Japan
[c] Psychiatric Mental Health Nursing, Japanese Red Cross Akita College of Nursing, 17-3 Azananawashirozawa, Kamikitatesaruta, Akita 010-1493, Japan
[d] Department of Public Health, Faculty of Medicine, University of Toyama, 3190 Gofuku, Toyama 930-8555, Japan
[e] National Center for Cognitive Behavior Therapy and Research, National Center of Neurology and Psychiatry, 4-1-1 Ogawahigashicho, Kodaira, Tokyo 187-0031, Japan

ARTICLE INFO

Keywords:
Anxiety
Burn out
Depression
Fatty acids
Omega-3
Presenteeism

ABSTRACT

The efficacy of omega-3 fatty acids for maintaining a better mental state has not been examined among working populations. We aimed to explore the effectiveness of omega-3 fatty acids for hospital nurses. In a multi-center randomized trial, 80 junior nurses were randomly allocated to either omega-3 fatty acids (1200 mg/day of eicosapentaenoic acid and 600 mg/day of docosahexaenoic acid) or identical placebo pills for 13 weeks. The primary outcome was the total score of the Hospital Anxiety and Depression Scale (HADS), determined by a blinded rater at week 26 from the study enrolment. Secondary outcomes included the total score of the HADS at 13 and 52 weeks; incidence of a major depressive episode; severity of depression, anxiety, insomnia, burnout, and presenteeism; utility scores; and adverse events at 13, 26 and 52 weeks. The mean HADS score at baseline was 7.2. At 26 weeks, adjusted mean scores on the HADS were 6.32 (95% CIs of standard errors: 5.13, 7.52) in the intervention and 6.81 (5.57, 8.05) in the placebo groups, respectively. The coefficient of the group by time interaction was not statistically significant at 0.58 (−1.35, 2.50; P = 0.557). Although the intervention group showed significant superiority on the HADS score at 52 weeks, depression severity at 52 weeks, insomnia severity at 13 weeks, and absolute presenteeism at 26 weeks, no significant superiority or inferiority was observed on the other outcomes. The additive value of omega-3 fatty acids was not confirmed regarding mental state and self-evaluated work efficiency.

Clinical trial registry number and website:

ClinicalTrials.gov: NCT02151162 (registered on May 27, 2014) https://clinicaltrials.gov/ct2/show/NCT02151162.

1. Introduction

Antidepressive effects of omega-3 polyunsaturated fatty acids (PUFAs), in particular eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), have been examined by many previous studies. A recent Cochrane review showed that omega-3 PUFA supplementation resulted in a small or modest benefit for depressive symptomatology compared to placebo for patients with a primary diagnosis of unipolar or major depressive disorder (MDD), with a standardized mean difference of −0.3 (95% CI, −0.10 to −0.50) by pooling results from 25 randomized controlled trials (RCTs) (Appleton et al., 2015).

Some other previous RCTs have examined the effects of omega-3 PUFAs in terms of preventing depression in non depressed population, especially for pregnant women who are either mentally healthy (Blasi et al., 1989; Krauss-Etschmann et al., 2007; Makrides et al., 2010; Mattes et al., 2009; Vaz et al., 2017) or at risk for postpartum depression (Doornbos et al., 2009). However, to the best of our knowledge, trials investigating the efficacy of PUFAs in preventing depression or maintaining healthy mental state among ordinary people, including working populations, are sparse. In a recently published meta-analysis (Grosso et al., 2014), omega-3 PUFAs were effective in patients diagnosed as MDD, but not in preventing depression in healthy subjects. In this analysis, three parallel RCTs investigating preventive effects for

* Corresponding author.
E-mail addresses: watanabe.norio.6x@kyoto-u.ac.jp (N. Watanabe), yumatsuo@ncc.go.jp (Y. Matsuoka), leoaru47@gmail.com (M. Kumachi), keihama@med.u-toyama.ac.jp (K. Hamazaki), mhorikoshi@ncnp.go.jp (M. Horikoshi), furukawa@kuhp.kyoto-u.ac.jp (T.A. Furukawa).

https://doi.org/10.1016/j.jpsychires.2018.03.015
Received 17 January 2018; Received in revised form 27 March 2018; Accepted 29 March 2018

depression in healthy subjects were assessed with a follow-up duration of 3–6 months. The studies employed a daily dose of 1060 mg EPA and 274 mg DHA (Mozurkewich et al., 2013), 3000 mg EPA and 600 mg DHA (DeFina et al., 2011), or 2085 mg EPA and 348 mg DHA (Kiecolt-Glaser et al., 2012). However, a high daily EPA dose between 1000 and 1500 mg/d with a ratio of 2:1 with DHA has been argued as optimal for affective disorders (McNamara, 2009). We believe that another methodologically rigorous trial employing optimal doses of EPA and DHA with long-term follow-up is needed.

Hospital nurses are vulnerable to psychological stress and mental disorders (Calnan et al., 2001), especially depression. The prevalence of depressive symptoms above a clinical cut-off among hospital-employed nurses is 18% in the U.S. (Letvak et al., 2012). Nurses with depression are not only likely to suffer personally, but their illness may also have an impact on the quality of care for patients due to presenteeism. In terms of the monetary burden of patient care, the costs due to increased falls and medication errors caused by presenteeism are estimated at 1346 USD per nurse annually (Letvak et al., 2012).

Hence, the present study aimed to explore the effectiveness of omega-3 PUFAs in maintaining a healthy mental state among hospital nurses. The protocol of the study has already been published elsewhere (Watanabe et al., 2015).

2. Materials and methods

2.1. Trial design

A factorial-design trial was conducted with 1:1:1:1 allocation, with a 52-week follow-up. Participants were randomly assigned to one of the following four intervention arms: a) omega-3 PUFAs plus mindfulness-based stress management program; b) placebo plus mindfulness-based stress management program; c) omega-3 PUFAs plus psychoeducation leaflet; and d) placebo plus psychoeducation leaflet. These interventions were terminated within 13 weeks from the registration of the participant. The present paper focuses on the comparison between combined groups of a) and c), and b) and d).

2.2. Participants

The inclusion criteria for participants were 1) female, aged between 20 and 59 years, because a previous study demonstrated that omega-3 PUFAs were effective in females but not in males (Nishi et al., 2012). We focused on females to maximize the benefit; 2) nurses who worked in inpatient wards at four general hospitals and at one psychiatric hospital in Tama area, Tokyo, Japan; 3) and those mainly engaged in caring for patients but not in administrative responsibilities. Thus, head or senior nurses were excluded. A previous study showed that senior nurses are less burnout than junior nurses (Voltmer et al., 2013).

The exclusion criteria for participants were those 1) planning to resign for any reasons within 26 weeks; 2) engaging in structured psychotherapy; 3) seeing a physician regularly for the treatment of any mood or anxiety disorders; 4) taking antidepressants, mood stabilizers, anticonvulsants or antipsychotics; 5) taking omega-3 PUFAs for four or more weeks within 52 weeks; 6) clinically depressed, judged by satisfying a total score of 11 or more on the Hospital Anxiety and Depression Scale (HADS) Depression Subscale (Kugaya et al., 1998) and that of 15 or more on the Patient Health Questionnaire (PHQ-9) (Spitzer et al., 1999); 7) consuming fish as the main course of a meal four or more times a week; or 8) taking anticoagulant drugs at entry or history of stroke or myocardial infarction.

2.3. Procedures

The participants were randomly assigned to one of the four intervention arms via the EDC using a minimization method, which controlled for place of work; a total score on the HADS of ≥11 (Kugaya et al., 1998) or not; and working as a nurse for ≥1 year or not.

2.4. Interventions

Omega-3 PUFA capsules were formulated to contain 1200 mg EPA and 600 mg DHA according to expert recommendations (McNamara, 2009). The participants took the capsules once a day for 90 days. Placebo capsules contained rapeseed oil (47%), soybean oil (25%), olive oil (25%), and fish oil (3%) and have identical appearance and similar odor to omega-3 PUFA capsules. To assess adherence to the regimen, all the remaining capsules were collected after the 13-week assessment.

Participants who stopped taking the capsules were still asked to complete the assessments. Participants assigned to the placebo group were asked not to take omega-3 PUFA supplements for the first 26 weeks.

2.5. Assessment measures

2.5.1. Primary outcome

The primary outcome was the blindly rated total score of the 14-item HADS (Herrmann, 1997; Zigmond and Snaith, 1983) at week 26, assessed through their mobile phone by a blinded rater located at Kyoto University. All participants were requested not to reveal the assigned treatment to the assessors, in order to keep the assessors' blindness to the groups. After each assessment, an assessor guessed which group the participant has been assigned to, making it possible to examine if the blinding was successful.

The total score of the HADS (HADS-T) ranges from 0 to 42, higher scores indicating more symptoms. The HADS has two sub-scores, each ranging from 0 to 21: HADS-D (depression) and HADS-A (anxiety). The recommended cutoffs of the HADS-T were ≥9 for possible cases and ≥11 for probable cases of anxiety or depressive disorders (Zigmond and Snaith, 1983).

2.5.2. Secondary outcomes

All self-reported measures other than the HADS were gathered through the EDC system, where the participants could input their own data through the Internet at home. Participants were notified at 13, 26, and 52 weeks to fill in the assessment questionnaires within the following 14, 30, and 30 days, respectively.

2.5.2.1. Depression and anxiety symptoms. The HADS was administered at baseline, 13 and 52 weeks as secondary outcomes through their mobile phone by a blinded rater.

2.5.2.2. Major depressive episode. A current major depressive episode, according to DSM-5, were determined using the Primary Care Evaluation of Mental Disorders (PRIME-MD) algorithm in the depression module of the PHQ-9 (Spitzer et al., 1999). The PHQ-9 has been used to assess major depressive disorder, according to DSM-5 (Fried et al., 2013).

2.5.2.3. Anxiety. The GAD-7 (Spitzer et al., 2006) was used to assess the severity of anxiety symptoms in the participants. Total scores from 5 to 9, 10–14, or 15–21 indicate mild, moderate, or severe anxiety symptoms, respectively.

2.5.2.4. Burnout. The Maslach's Burnout Inventory (MBI) (Maslach et al., 1996) was used to assess degree of burnout among nurses. The MBI is a tool for detecting and measuring the severity of burnout syndrome. It is a 22-item questionnaire that assesses the degree of burnout according to the following three subscales: emotional exhaustion (EE), depersonalization (DP) and personal accomplishment (PA).

2.5.2.5. Insomnia. The Insomnia Severity Index (ISI) (Bastien et al.,

2001; Morin and Espie, 2004) was used to assess insomnia. The ISI is now considered a standard global measure for assessing the severity of insomnia and is used in many studies (Morin et al., 2009; Watanabe et al., 2011). Total scores from 8 to 14 and 15–28 indicate subthreshold insomnia and clinical insomnia, respectively.

2.5.2.6. Presenteeism. The World Health Organization Heath and Work Performance Questionnaire (HPQ) was used to assess two types of presenteeism: the absolute presenteeism score, which is calculated as the difference between the score for self over the past 28 days and the score for the average worker in the same job and ranges from 0 (total lack of performance during time on the job) to 100 (no lack of performance during time on the job); and the relative presenteeism score, which is a ratio of actual performance to possible performance (the performance of most workers in the same job) and ranges from 0.25 to 2.0. The validity of the scale has been confirmed in previous studies (Kessler et al., 2004; Kessler and Ustun, 2004).

2.5.2.7. Quality of life (QoL). The EuroQol (EQ-5D) (EuroQol Group, 1990) was used to assess health-related QoL. The five domains include mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. The following are the three levels of severity for each domain: no problems, some or moderate problems, and severe or extreme problems. Each pattern of responses is assigned to an individual utility score, which ranges from 0 (death) to 1 (perfect health).

2.5.2.8. Adverse events and other outcomes. Information about types and severity of an adverse event and degree affecting normal life was collected through the EDC. Information about sick leave and consultations on personal mental state was also collected.

2.6. Sample size

The sample size was based on a power analysis conducted for the HADS scores, basically for the comparison between the stress management program and leaflet groups, because, to the best of our knowledge, there had been no published trial on the efficacy of omega-3 PUFAs in preventing depression among workers. Based on trials that have used nurses as psychotherapists and the HADS as the primary outcome (Romeo et al., 2011), we estimated the mean difference in the HADS scores between the groups to be 4 ± 6 (SD). To detect a significant difference at P = 0.05 (two-tailed) with a power of 0.9, and allowing for a 20% dropout rate, 60 participants would need to be recruited per group, which is a total of 120 participants. We considered that this number was not too small for the comparison between the omega-3 PUFA and placebo groups, because previous trials on the efficacy of omega-3 PUFA for depression in a Cochrane meta-analysis (Appleton et al., 2015) enrolled the median number of participants at around 50 in total.

However, in the interim follow-up during the study, considering extremely low dropout rate for the primary outcome assessment (5% in total), delay in recruitment process, and running out of research grant, we decided to set a new target number of the participants at 80 in total, in July 2015.

2.7. Data management and analysis

All of the participants who were randomized at baseline were included in the primary analyses (intention-to-treat, ITT). First, a descriptive analysis of the variables at baseline was performed. We did not plan any statistical tests to detect a difference at baseline among the trial arms because we aimed to avoid multiple tests; however, when clinically important differences at baseline were noted, we planned to perform our analyses by adjusting for all such possible confounds.

Second, for all the continuous outcomes assessed at 13, 26, and 52 weeks, treatment-by-time interactions were examined using a mixed model repeated measures analysis. The model included participants for a random effect and included treatment, time (categorical), treatment-by-time interaction, and baseline scores for fixed effects. When missing data existed for categorical variables, which were all negative outcomes, an ITT principal was applied by assuming that all dropouts did not satisfy the outcomes.

A P value of < 0.05 was set to test the null hypothesis for all analyses. We did not plan to perform interim analyses to examine the study hypotheses. All the analyses were conducted blindly in terms of the assigned groups, and the results were interpreted before breaking the group assignment. We used IBM SPSS statistics 23 for all analyses.

We, a priori, planned in the protocol to conduct our analyses in the present article by focusing on the efficacy of omega-3 PUFAs and to report that of mindfulness-based intervention in another article. Moreover, in our post-hoc analysis investigating the interaction between the efficacy of omega-3 PUFAs and that of mindfulness-based stress management program, we did not observe any interactions at 13 weeks (P = 0.872), 26 weeks (P = 0.927), and 52 weeks (P = 0.378), respectively. We believe that this may support our decision to report our results from the factorial-design trial in multiple articles.

2.8. Ethical issues

The present study complied with the ethical principles established for research on human beings stipulated in the Declaration of Helsinki and further amendments thereto.

The study protocol was approved by the institutional review boards of the National Center of Neurology and Psychiatry (A2014-017), Kyoto University (C0881) and of Toyama University (26–24). Written informed consent was obtained from all participants. Data for each participant were handled with sequentially assigned numbers to keep participant's confidentiality.

3. Results

3.1. Enrollment and baseline characteristics of the participants

Between June 11, 2014 and August 27, 2015, 83 nurses were assessed for eligibility, of whom 80 were enrolled and randomized to the omega-3 PUFAs (n = 40) or placebo (n = 40) groups. We did not find any clinically important differences between these two groups (Table 1).

3.2. Attrition, adherence and study integrity

We obtained data from 76 participants (95%) for the HADS and 75 (94%) for the other outcomes at 52 weeks (Fig. 1). Reported adverse events were not associated with dropouts.

The remaining capsules were collected from 39 participants (98%) in the omega-3 PUFA group and from 36 (90%) in the placebo group. The mean proportion of taking pills over the total amount provided was 84% (95% CIs: 77%, 91%) in each of the groups, and the median was 91% and 89% for omega-3 PUFAs and placebo groups, respectively. These indicated that the blinding of the participants appeared to be successful.

The agreement rate and kappa value for agreement between the actual assignment and those guessed by blind assessors at 24 weeks were 52.0% and 0.03. These indicated that the blinding of the assessors was successful.

Only one participant in the omega-3 PUFA group reported taking omega-3 PUFA supplements between 13 and 26 weeks, but none did in the placebo group and between 26 and 52 weeks.

Table 1
Characteristics of participants at baseline.

Continuous data: mean ± SD (range)	All (N = 80)	Omega-3 fatty acids (N = 40)	Placebo (N = 40)
Age	30.1 ± 8.4 (21–55)	29.6 ± 9.1 (21–55)	30.5 ± 7.8 (21–49)
Duration of experience as a nurse (year)	4.7 ± 5.6 (0–32)	4.0 ± 6.2 (0–32)	5.4 ± 5.0 (0–20)
Consultation for physical illness	18 (22.5%)	10 (25.0%)	8 (20.0%)
Marital status (Married; Unmarried; Divorced/Widowed)	14; 59; 7	6; 30; 4	8; 29; 3
Education for qualification for a nurse (Vocational school; College; Postgraduate course)	6; 45; 29	3; 21; 16	3; 24; 13
HADS total	7.2 ± 4.6 (0–20)	7.4 ± 4.8 (1–19)	7.1 ± 4.5 (0–20)
HADS depression	3.2 ± 2.7 (0–11)	3.3 ± 2.8 (0–11)	3.1 ± 2.5 (0–9)
HADS anxiety	4.0 ± 2.5 (0–11)	4.0 ± 2.4 (0–9)	4.0 ± 2.7 (0–11)
PHQ-9 total	4.9 ± 3.4 (0–15)	5.1 ± 3.4 (0–12)	4.7 ± 3.4 (0–15)
GAD-7 total	3.1 ± 2.9 (0–14)	3.0 ± 3.0 (0–10)	3.2 ± 2.9 (0–14)
ISI total	5.7 ± 4.1 (0–16)	6.4 ± 4.5 (0–16)	5.0 ± 3.5 (0–13)
MBI Emotional Exhaustion	21.5 ± 11.1 (3–50)	20.9 ± 7.6 (3–50)	19.4 ± 10.0 (5–49)
MBI Depersonalization	7.0 ± 4.9 (0–22)	6.9 ± 4.6 (1–22)	7.1 ± 5.2 (0–20)
MBI Lack of Personal Accomplishment	21.7 ± 8.4 (4–41)	20.9 ± 7.6 (4–40)	22.6 ± 9.1 (6–41)
Absolute presenteeism	50.6 ± 16.4 (10–90)	48.8 ± 16.2 (10–80)	52.5 ± 16.6 (20–90)
Relative presenteeism	0.88 ± 0.29 (0.25–1.80)	0.83 ± 0.28 (0.25–1.50)	0.92 ± 0.30 (0.25–1.80)
Utility score	0.91 ± 0.12 (0.66–1.0)	0.90 ± 0.13 (0.66–1.0)	0.92 ± 0.12 (0.66–1.0)
N of incidents in the last 13 weeks	1.3 ± 1.2 (0–5)	1.4 ± 1.3 (0–5)	1.2 ± 1.2 (0–4)
N of accidents in the last 13 weeks	0.1 ± 0.3 (0–2)	0.1 ± 0.4 (0–2)	0.1 ± 0.3 (0–1)

Abbreviations: HADS, Hospital Anxiety and Depression Scale; ISI, Insomnia Severity Index; MBI, Maslach's Burnout Inventory; PHQ-9, Patient Health Questionnaire-9.

3.3. Primary outcome

At 26 weeks, adjusted mean scores in the HADS-T score were 6.32 (95% CIs of standard errors: 5.13, 7.52) and 6.81 (5.57, 8.05) in the omega-3 PUFA and placebo groups, respectively (Table 2, Fig. 2). The coefficient of the group by time interaction was not statistically significant at 0.58 (−1.35, 2.50; P = 0.557).

3.4. Secondary outcomes

Regarding major depressive episode at 26 weeks, four and three participants satisfied the episode in the omega-3 PUFA and placebo groups, respectively, and there was no statistical difference (P = 0.692) (Table 3). No statistically significant differences were also observed at 13 weeks or at 52 weeks.

Statistically significant superiority was observed in favor of the omega-3 PUFA group in terms of the HADS-T score at 52 weeks (group by time interaction 1.52; 0.46, 2.58; P = 0.005), the HADS-D score at 52 weeks (1.50; 0.26, 2.73; P = 0.018), the Insomnia Severity Index at 13 weeks (2.05; 0.34, 3.76; P = 0.019), and absolute presenteeism at 26 weeks (−6.29; −12.56, 0.02; P = 0.049) (Table 2). No significant superiority or inferiority was observed in the other outcomes, including anxiety, burnout, QoL (Table 2), psychiatrist consultation, or psychotropic medication intake (Table 3).

Adverse events were rare in both groups (Table 3). We did not observe any serious adverse events.

4. Discussion

To the best of our knowledge, this study is the first to investigate the effectiveness of omega-3 PUFAs for depression and anxiety symptom prevention among workers. However, no significant differences between the omega-3 PUFA and placebo groups were observed in terms of our primary outcome, the total score of depression and anxiety at 26 weeks. Although some statistically significant results favorable to the omega-3 PUFA group were observed in the secondary outcomes, including the total score of depression and anxiety at 52 weeks, the insomnia severity score at 13 weeks, and absolute presenteeism at 26 weeks, we may not be able to eliminate the possibility that these plausible results are chance findings due to multiple tests. Moreover, these differences are difficult to be explained biologically, when we consider effects of omega-3 PUFAs on depression which could occur as a result of changes in cell membrane structure and function, leading to impacts on cell communication, inflammatory processes and neurotransmitter activities (Haag, 2003; James et al., 2000). The intervention was terminated at the end of the first 13 weeks, and may not benefit the long-term outcomes without short-term effects.

On the other hand, no significant results of the incidence of adverse events or dropouts were observed between the two intervention arms. We, therefore, concluded that omega-3 PUFAs provide neither benefit or harm in terms of maintaining healthy mental health state and preventing depression among working populations.

A recent systematic review and meta-analysis have shown that omega-3 PUFAs do not prevent depressive symptoms among

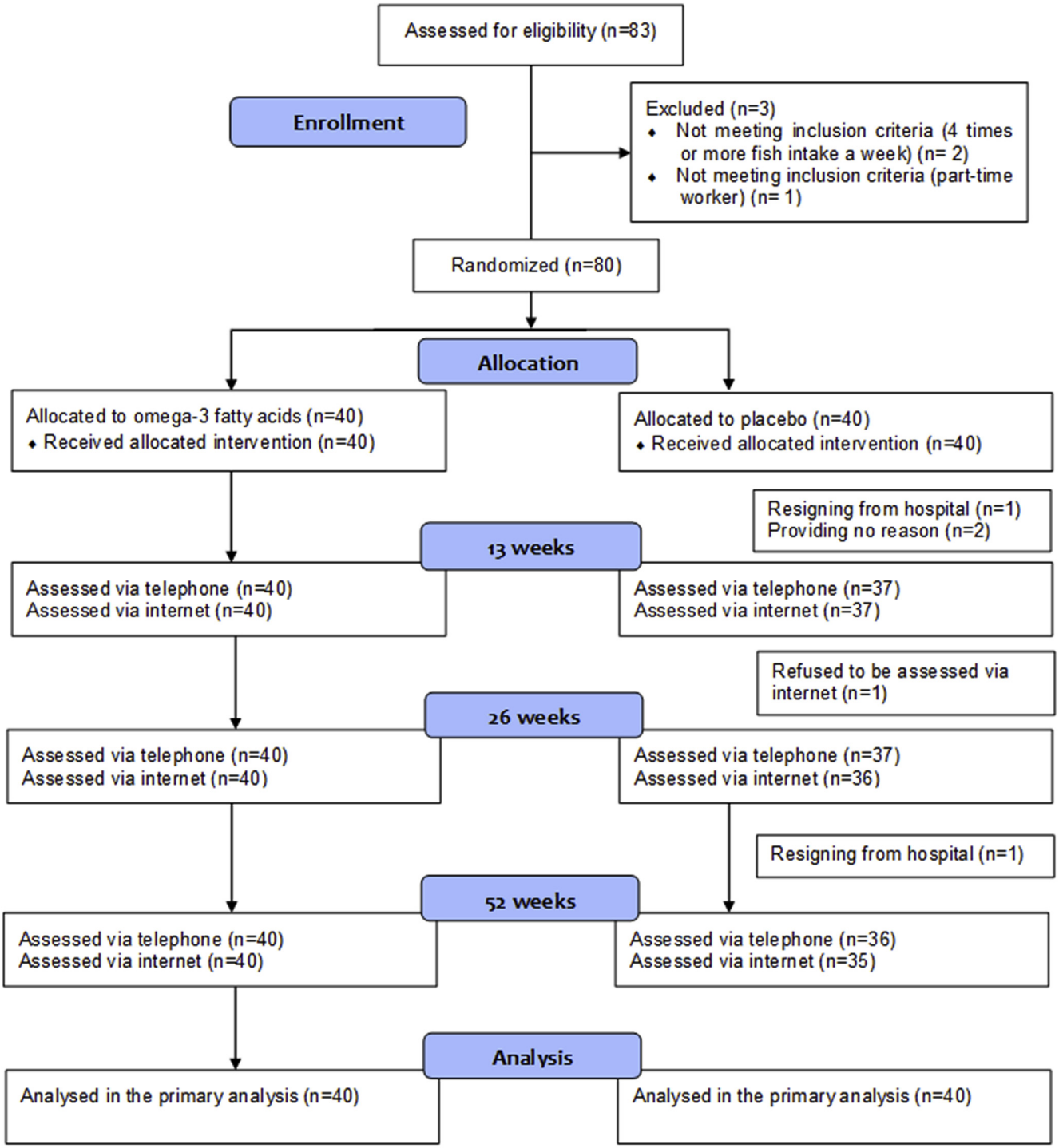


Fig. 1. Participant flow diagram.

populations not diagnosed for depression (Hallahan et al., 2016). Other trials not included in the systematic review and focusing on healthy participants also did not show additional benefit of omega-3 PUFAs in terms of mood (Antypa et al., 2009; van de Rest et al., 2008). Adding the results of optimal dose administration of omega-3 PUFAs from the present study to the current evidence, we would conclude that omega-3 PUFAs were unlikely to be a primary prevention strategy for depressive symptoms in mentally healthy people including working populations.

The present study is not without methodological limitations. First, we did not manage to recruit the number of participants we initially set from a power analysis in the protocol, probably because nurses in the hospitals might think participating in the present study would be burdensome or humiliating. However, our post-hoc analyses showed that the effect size (Cohen's d) (Cohen, 1988) of the primary outcome for the completers was −0.03, with 95% CIs from −0.48 to 0.42. Thus, even from the most optimistic view, the efficacy of PUFAs could only lead to small effect sizes. Hence, we believe that our conclusion that omega-3 PUFAs are unlikely to provide benefit in terms of preventing depression, is accurate.

Secondly, although we reminded participants to take assigned pills

Table 2
Estimated marginal means of Omega-3 and placebo arms for the continuous outcomes.

	Estimated marginal means: mean (95% confidence intervals of standard errors)			
	Baseline	Week 13	Week 26	Week 52
HADS total				
Omega-3 fatty acids	7.20 (6.00, 8.40)	6.40 (5.20, 7.60)	6.32 (5.13, 7.52)	5.85 (4.65, 7.05)
Placebo	7.11 (5.491, 8.31)	7.60 (6.36, 8.83)	6.81 (5.57, 8.05)	8.32 (7.07, 9.58)
Group by time interaction		1.28 (−0.64, 3.21)	0.58 (−1.35, 2.50)	2.56 (0.63, 4.50)
		P = 0.190	P = 0.557	P = 0.010
HADS depression				
Omega-3 fatty acids	3.24 (2.56, 3.92)	2.84 (2.16, 3.52)	2.69 (2.01, 3.37)	2.36 (1.68, 3.04)
Placebo	3.16 (2.50, 3.83)	3.54 (2.84, 4.24)	3.05 (2.35, 3.75)	3.80 (30.9, 4.51)
Group by time interaction		0.78(-0.27, 1.84)	0.45 (−0.61, 1.50)	1.52 (0.46, 2.58)
		P = 0.144	P = 0.404	P = 0.005
HADS anxiety				
Omega-3 fatty acids	3.97 (3.30, 4.64)	3.57 (2.90, 4.24)	3.65 (2.98, 4.32)	3.50 (2.83, 4.17)
Placebo	3.95 (3.28, 4.62)	4.05 (3.36, 4.74)	3.75 (3.06, 4.44)	4.51 (3.82, 5.21)
Group by time interaction		0.50 (−0.54, 1.53)	0.13 (−0.91, 1.16)	1.04 (−0.01, 2.08)
		P = 0.345	P = 0.812	P = 0.050
PHQ-9				
Omega-3 fatty acids	4.94 (4.02, 5.86)	5.42 (4.50, 6.33)	5.39 (4.47, 6.31)	5.42 (4.50, 6.33)
Placebo	4.82 (3.90, 5.74)	5.83 (4.78, 6.88)	5.31 (4.35, 6.26)	4.79 (3.83, 5.76)
Group by time interaction		0.53 (−0.89, 1.95)	0.04 (−1.39, 1.46)	−0.50 (−1.94, 0.93)
		P = 0.460	P = 0.960	P = 0.491
GAD-7				
Omega-3 fatty acids	3.08 (2.33, 3.83)	3.40 (2.65, 4.15)	3.85 (3.11, 4.60)	3.75 (3.01, 4.50)
Placebo	3.13 (2.38, 3.88)	3.45 (2.67, 4.22)	3.30 (2.52, 4.09)	3.73 (2.94, 4.52)
Group by time interaction		−0.01 (−1.18, 1.16)	−0.60 (−1.78, 0.57)	−0.08 (−1.26, 1.10)
		P = 0.986	P = 0.313	P = 0.893
Insomnia Severity Index				
Omega-3 fatty acids	5.94 (4.83, 7.06)	5.07 (3.96, 6.18)	6.09 (4.98, 7.21)	6.07 (4.96, 7.18)
Placebo	5.45 (4.36, 6.56)	6.62 (5.47, 7.77)	5.36 (4.20, 6.21)	5.34 (4.16, 6.51)
Group by time interaction		2.05 (0.34, 3.76)	−0.24 (−1.96, 1.47)	−0.24 (−1.96, 1.49)
		P = 0.019	P = 0.780	P = 0.786
MBI EE				
Omega-3 fatty acids	22.0 (19.3, 24.8)	21.7 (19.0, 24.4)	23.6 (20.9, 26.3)	22.0 (19.3, 24.8)
Placebo	20.8 (18.2, 23.6)	23.0 (20.2, 25.8)	22.3 (19.4, 25.1)	23.5 (20.6, 26.3)
Group by time interaction		2.47 (−2.69, 6.63)	−0.11 (−4.28, 4.07)	2.62 (−1.58, 6.81)
		P = 0.244	P = 0.960	P = 0.221
MBI DP				
Omega-3 fatty acids	6.92 (5.46, 8.38)	8.17 (6.71, 9.63)	8.00 (6.53, 9.46)	9.22 (7.76, 10.68)
Placebo	7.02 (5.60, 8.49)	7.78 (6.27, 9.30)	7.76 (6.23, 9.30)	8.44 (6.89, 10.00)
Group by time interaction		−0.49 (−2.93, 2.00)	−0.33 (−2.79, 2.12)	−0.88 (−3.35, 1.59)
		P = 0.694	P = 0.789	P = 0.484
MBI PA				
Omega-3 fatty acids	21.7 (19.7, 23.8)	20.3 (18.3, 22.4)	23.0 (21.0, 25.0)	22.3 (20.2, 24.3)
Placebo	22.0 (20.0, 24.1)	23.5 (21.4, 25.6)	21.2 (19.1, 23.4)	22.0 (19.9, 24.2)
Group by time interaction		2.88 (−0.28, 6.04)	−2.11 (−5.28, 1.06)	−0.56 (−3.75, 2.63)
		P = 0.00.74	P = 0.192	P = 0.728
Absolute presenteeism				
Omega-3 fatty acids	49.9 (44.4, 55.4)	53.2 (48.7, 57.8)	64.4 (59.9, 68.9)	62.4 (57.9, 66.9)
Placebo	51.7 (47.2, 56.3)	53.0 (48.4, 57.7)	59.9 (55.3, 64.6)	61.8 (57.0, 66.5)
Group by time interaction		−2.05 (−8.31, 4.21)	−6.29 (−12.56, 0.02)	−2.48 (−8.77, 3.82)
		P = 0.520	P = 0.049	P = 0.440
Relative presenteeism				
Omega-3 fatty acids	0.86 (0.78, 0.95)	0.92 (0.83, 1.01)	0.95 (0.86, 1.03)	0.88 (0.79, 0.96)
Placebo	0.90 (0.82, 0.99)	0.85 (0.76, 0.94)	0.86 (0.77, 0.95)	0.82 (0.73, 0.91)
Group by time interaction		−0.12 (−0.25, 0.02)	−0.13 (−0.27, 0.01)	−0.10 (−0.24, 0.03)
		P = 0.088	P = 0.056	P = 0.141
Utility score				
Omega-3 fatty acids	0.90 (0.86, 0.94)	0.84 (0.80, 0.88)	0.86 (0.82, 0.90)	0.87 (0.83, 0.92)
Placebo	0.92 (0.88, 0.97)	0.87 (0.83, 0.92)	0.87 (0.83, 0.92)	0.87 (0.82, 0.91)
Group by time interaction		1.1 (−0.06, 0.08)	−0.01 (−0.08, 0.06)	−0.03 (−0.10, 0.04)
		P = 0.813	P = 0.744	P = 0.376

Group by time interaction is presented as a coefficient, its 95% CIs, and P-value.
Abbreviations: HADS, Hospital Anxiety and Depression Scale; ISI, Insomnia Severity Index; MBI, Maslach's Burnout Inventory; PHQ-9, Patient Health Questionnaire.

regularly by sending monthly e-mails, adherence rates of taking pills, calculated by counting the remains, were the mean of 84% and the median of 90% in both the omega-3 PUFA and the placebo groups, respectively. These seem to be smaller than those in previous trials reporting actual adherence rates among participants in trials on the omega-3 PUFA intervention on mental states among non-clinical depressive participants, including 89% (Poppitt et al., 2009), 93% (Sinn et al., 2012), 96% (van de Rest et al., 2008), and around 97% (Kiecolt-Glaser et al., 2012).

Third, because fish consumption is higher in Japan than in other countries, one may doubt applicability of the findings about omega-3 PUFAs from the present study. A high consumption of fish has been reported to be correlated with a lower countrywide prevalence of major depression, according to a study published in 1998 (Hibbeln, 1998).

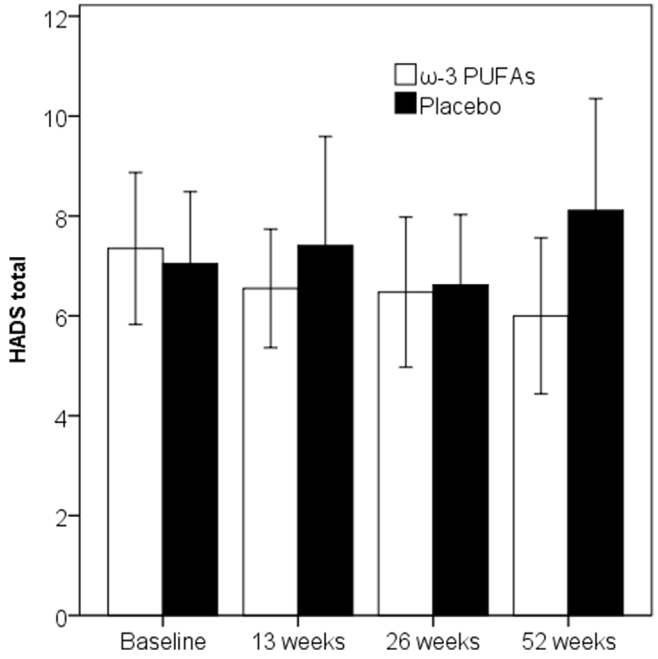


Fig. 2. Hospital Anxiety and Depression Scale scores for the assessment completers (N = 36)
Abbreviations: HADS, Hospital Anxiety and Depression Scale; PUFA, polysaturated
fatty acid.

However, we excluded participants who consumed fish as the main course of meal four or more times a week. We believe that our findings can be applied to the other countries.

Fourth, the overall effects of omega-3 PUFAs might be attenuated by counteracting effects of omega-6 PUFAs (Marventano et al., 2015). We did not collect information about diet including omega-3 PUFAs from the participants during the study, because we aimed a pragmatic trial. Future studies may need to investigate the association between the n-3: n-6 PUFAs ratio and incidence of depression.

5. Conclusion

The additive value of omega-3 PUFAs was not confirmed in terms of mental state and self-evaluated work efficiency in work populations. We do not recommend omega-3 PUFAs for the general population who are willing to maintain their mental health and looking for something beneficial.

Role of the funding source

This study was funded by Intramural Research Grant (25-8) for Neurological and Psychiatric Disorders of National Center of Neurology and Psychiatry, Japan. The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Potential conflicts of interests

The authors have no conflicts of interests to declare, that may be affected by the publication of the manuscript.

Dr Watanabe has received research funds from the Japanese Ministry of Health Labor and Welfare, the Japanese Ministry of Education, Science, and Technology and National Center of Neurology and Psychiatry, Intramural Research Grant for Neurological and Psychiatric Disorders. He has also received royalties from Sogensha and Akatsuki.

Dr Matsuoka has received a Grant-in-Aid for Scientific Research (B) from the Japan Society for the Promotion of Science, an intramural research grant for Neurological and Psychiatric Disorders from the National Center of Neurology and Psychiatry Japan (27-3-2), and a grant from Pfizer Health Research Foundation and SENSHIN Medical Research Foundation.

Dr Kumachi has received research funds from Japan Society for the Promotion of Science.

Dr Hamazaki received Grant-in-Aid for Scientific Research (C) (17K00849), a grant from the First Bank Foundation and intramural research Grant (27-3) for Neurological and Psychiatric Disorders from the National Center of Neurology and Psychiatry, Japan.

Dr Furukawa has received lecture fees from Eli Lilly, Janssen, Meiji, Mitsubishi-Tanabe, MSD and Pfizer. He has received royalties from Igaku-Shoin and Nihon Bunka Kagaku-sha publishers. He has received research support from Mitsubishi-Tanabe and Mochida. He is diplomate of the Academy of Cognitive Therapy.

All the other author reports no conflicts of interest.

Table 3
Numbers of participants satisfying secondary dichotomous outcomes at each assessment.

	Number satisfying the outcome/Number remained in the arm		
	Week 13	Week 26	Week 52
Major depressive episode			
Omega-3 fatty acids	1/39	4/40	4/40
Placebo	3/37	3/36	1/35
P-value for ITT analysis	P = 0.305	P = 0.692	P = 0.166
P-value for completers	P = 0.279	P = 0.802	P = 0.216
Consulting psychiatrists			
Omega-3 fatty acids	0/39	1/40	1/40
Placebo	0/37	0/36	1/35
P-value for ITT analysis		P = 0.340	P = 0.924
P-value for completers		P = 0.314	P = 1.000
Taking psychotropic medication			
Omega-3 fatty acids	0/39	1/40	0/40
Placebo	0/37	1/36	1/35
P-value for ITT analysis		P = 1.000	P = 0.282
P-value for completers		P = 1.000	P = 0.314
Adverse events			
Dropouts due to adverse events			
Omega-3 fatty acids	0/39	0/40	0/40
Placebo	0/37	0/36	0/35
Any adverse events			
Omega-3 fatty acids	3/39	1/40	0/40
Placebo	1/37	1/36	0/35
Headache			
Omega-3 fatty acids	1/39	1/40	0/40
Placebo	0/37	0/36	0/35
Nausea			
Omega-3 fatty acids	2/39	0/40	0/40
Placebo	1/37	1/36	0/35
Depression			
Omega-3 fatty acids	0/39	1/40	0/40
Placebo	0/37	0/36	0/35
Other			
Omega-3 fatty acids	1/39	0/40	0/40
Placebo	1/37	1/36	0/35

Abbreviations: ITT, intention to treat.

Contributors

Norio Watanabe - Participated in the conception and design of the

study, recruited participants, collected data, performed the analysis and wrote the manuscript.

Yutaka Matsuoka - Participated in the conception and design of the study, and recruited participants.

Mie Kumachi - Participated in the conception and design of the study, recruited participants, and collected data.

Kei Hamazaki - Participated in the conception and design of the study, and collected data.

Masaru Horikoshi - Participated in the conception and design of the study.

Toshi A. Furukawa - Participated in the conception and design of the study, collected data, and performed the analysis.

All authors revised the article critically for important intellectual content and approved the final manuscript.

Disclaimer statements

None.

Previous presentation

None.
REFERENCES.

Acknowledgement of assistance

We thank all the collaborators (Tomomi Narisawa, CP, Miyuki Makino, CP, RN, Yuki Oe, CP, Issei Shinmei, CP at National Center of Neurology and Psychiatry, and Masato Usuki, MD, PhD at National Disaster Medical Center), technical assistant staff (Shizuko Takebe at University of Toyama), and the administrative staff (Mayuko Yamano at National Center of Neurology and Psychiatry). We also thank Shinako Ushijima, RN at National Center of Neurology and Psychiatry Hospital, Keiko Osada, RN and Masami Harasawa, RN at National Disaster Medical Center, Miyuki Hatada, RN and Noriko Miyajima, RN at Tokyo Metropolitan Tama Medical Center, Chieko Kosaka, RN and Masako Tanaka, RN at Tama-Hokubu Medical Center, and Jun-ichi Watanabe, RN at Inokashira Hospital for their great assistance in recruiting participants in the study.

References

Antypa, N., Van der Does, A.J., Smelt, A.H., Rogers, R.D., 2009. Omega-3 fatty acids (fish-oil) and depression-related cognition in healthy volunteers. J. Psychopharmacol. 23 (7), 831–840.

Appleton, K.M., Sallis, H.M., Perry, R., Ness, A.R., Churchill, R., 2015. Omega-3 fatty acids for depression in adults. Cochrane Database Syst. Rev.(11) CD004692.

Bastien, C.H., Vallieres, A., Morin, C.M., 2001. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med. 2 (4), 297–307.

Blasi, F., Sommariva, D., Cosentini, R., Cavaiani, B., Fasoli, A., 1989. Bezafibrate inhibits HMG-CoA reductase activity in incubated blood mononuclear cells from normal subjects and patients with heterozygous familial hypercholesterolaemia. Pharmacol. Res. 21 (3), 247–254.

Calnan, M., Wainwright, D., Forsythe, M., Wall, B., Almond, S., 2001. Mental health and stress in the workplace: the case of general practice in the UK. Soc. Sci. Med. 52 (4), 499–507.

Cohen, J., 1988. Statistical Power Analysis in the Behavioral Sciences. Lawrence Erlbaum Associates, Inc, Hillsdale (NJ).

DeFina, L.F., Marcoux, L.G., Devers, S.M., Cleaver, J.P., Willis, B.L., 2011. Effects of omega-3 supplementation in combination with diet and exercise on weight loss and body composition. Am. J. clin. Nutr. 93 (2), 455–462.

Doornbos, B., van Goor, S.A., Dijck-Brouwer, D.A., Schaafsma, A., Korf, J., Muskiet, F.A., 2009. Supplementation of a low dose of DHA or DHA + AA does not prevent peri-partum depressive symptoms in a small population based sample. Prog. Neuro Psychopharmacol. Biol. Psychiatr. 33 (1), 49–52.

EuroQol Group, 1990. EuroQol–a new facility for the measurement of health-related quality of life. Health policy (Amsterdam, Netherlands) 16 (3), 199–208.

Fried, E.I., Nesse, R.M., Zivin, K., Guille, C., Sen, S., 2013. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. Psychol. Med. 1–10.

Grosso, G., Pajak, A., Marventano, S., Castellano, S., Galvano, F., Bucolo, C., Drago, F., Caraci, F., 2014. Role of omega-3 fatty acids in the treatment of depressive disorders: a comprehensive meta-analysis of randomized clinical trials. PLoS One 9 (5), e96905.

Haag, M., 2003. Essential fatty acids and the brain. Can J Psychiatry 48 (3), 195–203.

Hallahan, B., Ryan, T., Hibbeln, J.R., Murray, I.T., Glynn, S., Ramsden, C.E., SanGiovanni, J.P., Davis, J.M., 2016. Efficacy of omega-3 highly unsaturated fatty acids in the treatment of depression. Br. J. Psychiatry 209 (3), 192–201.

Herrmann, C., 1997. International experiences with the Hospital Anxiety and Depression Scale–a review of validation data and clinical results. J. Psychosom. Res. 42 (1), 17–41.

Hibbeln, J.R., 1998. Fish consumption and major depression. Lancet 351 (9110), 1213.

James, M.J., Gibson, R.A., Cleland, L.G., 2000. Dietary polyunsaturated fatty acids and inflammatory mediator production. Am. J. Clin. Nutr. 71 (1 Suppl. l), 343S–348S.

Kessler, R.C., Ames, M., Hymel, P.A., Loeppke, R., McKenas, D.K., Richling, D.E., Stang, P.E., Ustun, T.B., 2004. Using the World health organization health and work performance questionnaire (HPQ) to evaluate the indirect workplace costs of illness. J. Occup. Environ. Med./Am. College Occu. Environ. Med. 46 (6 Suppl. l), S23–S37.

Kessler, R.C., Ustun, T.B., 2004. The World mental health (WMH) survey initiative version of the World health organization (WHO) composite international diagnostic interview (CIDI). Int. J. Meth. Psychiatr. Res. 13 (2), 93–121.

Kiecolt-Glaser, J.K., Belury, M.A., Andridge, R., Malarkey, W.B., Hwang, B.S., Glaser, R., 2012. Omega-3 supplementation lowers inflammation in healthy middle-aged and older adults: a randomized controlled trial. Brain Behav. Immun. 26 (6), 988–995.

Krauss-Etschmann, S., Shadid, R., Campoy, C., Hoster, E., Demmelmair, H., Jimenez, M., Gil, A., Rivero, M., Veszpremi, B., Decsi, T., Koletzko, B.V., Nutrition, Health Lifestyle Study, G., 2007. Effects of fish-oil and folate supplementation of pregnant women on maternal and fetal plasma concentrations of docosahexaenoic acid and eicosa-pentaenoic acid: a European randomized multicenter trial. Am. J. Clin. Nutr. 85 (5), 1392–1400.

Kugaya, A., Akechi, T., Okuyama, T., Okamura, H., Uchitomi, Y., 1998. Screening for psychological distress in Japanese cancer patients. Jpn. J. Clin. Oncol. 28 (5), 333–338.

Letvak, S., Ruhm, C.J., McCoy, T., 2012a. Depression in hospital-employed nurses. Clinical nurse specialist CNS 26 (3), 177–182.

Letvak, S.A., Ruhm, C.J., Gupta, S.N., 2012b. Nurses' presenteeism and its effects on self-reported quality of care and costs. Am. J. Nurs. 112 (2), 30–38 quiz 48, 39.

Makrides, M., Gibson, R.A., McPhee, A.J., Yelland, L., Quinlivan, J., Ryan, P., Team, D.O.I., 2010. Effect of DHA supplementation during pregnancy on maternal depression and neurodevelopment of young children: a randomized controlled trial. JAMA. J. Am. Med. Assoc. 304 (15), 1675–1683.

Marventano, S., Kolacz, P., Castellano, S., Galvano, F., Buscemi, S., Mistretta, A., Grosso, G., 2015. A review of recent evidence in human studies of n-3 and n-6 PUFA intake on cardiovascular disease, cancer, and depressive disorders: does the ratio really matter? Int. J. Food Sci. Nutr. 66 (6), 611–622.

Maslach, C., Jackson, S.E., Leiter, M.P., 1996. Maslach Burnout Inventory Manual 3rd Edition.

Mattes, E., McCarthy, S., Gong, G., van Eekelen, J.A., Dunstan, J., Foster, J., Prescott, S.L., 2009. Maternal mood scores in mid-pregnancy are related to aspects of neonatal immune function. Brain Behav. Immun. 23 (3), 380–388.

McNamara, R.K., 2009. Evaluation of docosahexaenoic acid deficiency as a preventable risk factor for recurrent affective disorders: current status, future directions, and dietary recommendations. Prostaglandins, leukotrienes, and essential fatty acids 81 (2–3), 223–231.

Morin, C.M., Espie, C.A., 2004. Insomnia: a Clinical Guide to Assessment and Treatment. Guilford Press, New York.

Morin, C.M., Vallieres, A., Guay, B., Ivers, H., Savard, J., Merette, C., Bastien, C., Baillargeon, L., 2009. Cognitive behavioral therapy, singly and combined with medication, for persistent insomnia: a randomized controlled trial. JAMA. J. Am. Med. Assoc. 301 (19), 2005–2015.

Mozurkewich, E.L., Clinton, C.M., Chilimigras, J.L., Hamilton, S.E., Allbaugh, L.J., Berman, D.R., Marcus, S.M., Romero, V.C., Treadwell, M.C., Keeton, K.L., Vahratian, A.M., Schrader, R.M., Ren, J., Djuric, Z., 2013. The mothers, Omega-3, and mental health study: a double-blind, randomized controlled trial. Am. J. Obstet. Gynecol. 208 (4), e311–319 313.

Nishi, D., Koido, Y., Nakaya, N., Sone, T., Noguchi, H., Hamazaki, K., Hamazaki, T., Matsuoka, Y., 2012. Fish oil for attenuating posttraumatic stress symptoms among rescue workers after the great east Japan earthquake: a randomized controlled trial. Psychother. Psychosom. 81 (5), 315–317.

Poppitt, S.D., Howe, C.A., Lithander, F.E., Silvers, K.M., Lin, R.B., Croft, J., Ratnasabapathy, Y., Gibson, R.A., Anderson, C.S., 2009. Effects of moderate-dose omega-3 fish oil on cardiovascular risk factors and mood after ischemic stroke: a randomized, controlled trial. Stroke 40 (11), 3485–3492.

Romeo, R., Knapp, M., Banerjee, S., Morris, J., Baldwin, R., Tarrier, N., Pendleton, N., Horan, M., Burns, A., 2011. Treatment and prevention of depression after surgery for hip fracture in older people: cost-effectiveness analysis. J. Affect. Disord. 128 (3), 211–219.

Sinn, N., Milte, C.M., Street, S.J., Buckley, J.D., Coates, A.M., Petkov, J., Howe, P.R., 2012. Effects of n-3 fatty acids, EPA v. DHA, on depressive symptoms, quality of life, memory and executive function in older adults with mild cognitive impairment: a 6-month randomised controlled trial. Br. J. Nutr. 107 (11), 1682–1693.

Spitzer, R.L., Kroenke, K., Williams, J.B., 1999. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA. the journal of the American Medical Association 282 (18), 1737–1744.

Spitzer, R.L., Kroenke, K., Williams, J.B., Lowe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097.

van de Rest, O., Geleijnse, J.M., Kok, F.J., van Staveren, W.A., Hoefnagels, W.H., Beekman, A.T., de Groot, L.C., 2008. Effect of fish-oil supplementation on mental well-being in older subjects: a randomized, double-blind, placebo-controlled trial.

Am. J. Clin. Nutr. 88 (3), 706–713.
Vaz, J.D.S., Farias, D.R., Adegboye, A.R.A., Nardi, A.E., Kac, G., 2017. Omega-3 supplementation from pregnancy to postpartum to prevent depressive symptoms: a randomized placebo-controlled trial. BMC Pregnancy Childbirth 17 (1), 180.
Voltmer, E., Wingenfeld, K., Spahn, C., Driessen, M., Schulz, M., 2013. Work-related behaviour and experience patterns of nurses in different professional stages and settings compared to physicians in Germany. International journal of mental health nursing 22 (2), 180–189.
Watanabe, N., Furukawa, T.A., Horikoshi, M., Katsuki, F., Narisawa, T., Kumachi, M., Oe, Y., Shinmei, I., Noguchi, H., Hamazaki, K., Matsuoka, Y., 2015. A mindfulness-based stress management program and treatment with omega-3 fatty acids to maintain a healthy mental state in hospital nurses (Happy Nurse Project): study protocol for a randomized controlled trial. Trials 16, 36.
Watanabe, N., Furukawa, T.A., Shimodera, S., Morokuma, I., Katsuki, F., Fujita, H., Sasaki, M., Kawamura, C., Perlis, M.L., 2011. Brief behavioral therapy for refractory insomnia in residual depression: an assessor-blind, randomized controlled trial. J. Clin. Psychiatr. 72 (12), 1651–1658.
Zigmond, A.S., Snaith, R.P., 1983. The hospital anxiety and depression scale. Acta Psychiatr. Scand. 67 (6), 361–370.

Efficacy of antidepressants over placebo is similar in two-armed versus three-armed or more-armed randomized placebo-controlled trials

Yusuke Ogawa[a], Toshi A. Furukawa[a,b], Nozomi Takeshima[a], Yu Hayasaka[a], Lauren Z. Atkinson[d,e], Shiro Tanaka[c], Andrea Cipriani[e,f] and Georgia Salanti[g]

Previous studies have reported that effect sizes of antidepressants were larger in two-armed than in three-armed or more-armed (multiarmed) randomized trials, where the probability of being allocated to placebo is lower. However, these studies have not taken into account the publication bias, differences among antidepressants, or covariance in multiarmed studies, or examined sponsorship bias. We searched published and unpublished randomized-controlled trials that compared placebo with 21 antidepressants for the acute treatment of major depression in adults. We calculated the ratio of odds ratios (ROR) of drug response over placebo in two-armed versus multiarmed trials for each antidepressant, and then synthesized RORs across all the included antidepressants using the multivariate meta-analysis. A random-effects model was used throughout. Two hundred and fifty-eight trials (66 two-armed and 192 multiarmed trials; 80 454 patients; 43.0% with unpublished data) were included in the present analyses. The pooled ROR for response of two-armed trials over multiarmed trials was 1.09 (95% confidence interval: 0.96–1.24). The ROR did not materially change between types of antidepressants, publication year, or sponsorship. The differences between two-armed versus multiarmed studies were much smaller than were suggested in previous studies and were not significant. *Int Clin Psychopharmacol* 00:000–000

International Clinical Psychopharmacology 2017, **00**:000–000

Keywords: antidepressants, meta-analysis, number of arms, placebo-controlled trial, randomized-controlled trial, systematic review, trial design

Departments of [a]Health Promotion and Human Behavior, [b]Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, [c]Department of Clinical Biostatistics, Kyoto University Graduate School of Medicine, Kyoto, Japan, [d]Oxford Centre for Human Brain Activity, Wellcome Centre for Integrative Neuroimaging, [e]Department of Psychiatry, University of Oxford, [f]Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK and [g]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

Correspondence to Toshi A. Furukawa, MD, PhD, Department of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
Tel: +81 75 753 9491; fax: +81 75 753 4641;
e-mail: furukawa@kuhp.kyoto-u.ac.jp

Received 1 August 2017 **Accepted** 26 September 2017

Introduction

Pharmacotherapy is the mainstay in today's treatment of major depression, and hundreds of randomized-controlled trials (RCTs) of various antidepressants have been conducted so far to examine their efficacy (Furukawa *et al.*, 2016). Randomized, double-blind, placebo-controlled trials are required by regulatory agencies worldwide to obtain their approval for use with humans, and are considered to be the gold standard for the evaluation of efficacy of antidepressants.

However, overestimation of drug efficacy in traditional placebo-controlled trials has been suggested when effect sizes (ESs) were compared between two-armed and three-armed RCTs. Although the efficacy of the same antidepressant over placebo should not be different whether compared head to head against placebo or compared against another active drug along with placebo, the magnitude of the ES for antidepressants in three-armed RCTs was much smaller than those obtained in previous analyses that included two-armed trials (Greenberg *et al.*, 1992). These authors ascribed this difference to the greater possibility of unblinding in two-armed versus multiarmed studies. Blinding may indeed be difficult to maintain in studies of psychotropic drugs because these drugs have characteristic side effects (Margraf *et al.*, 1991; Even *et al.*, 2000; Moncrieff *et al.*, 2004). When double-blindness is breached, drug efficacy over placebo would probably be overestimated (Leucht *et al.*, 2009).

Some reports have also suggested that antidepressant–placebo difference was associated negatively with the number of treatment arms (Khan *et al.*, 2004; Papakostas and Fava, 2009; Sinyor *et al.*, 2010). These authors implicated the role of expectancy that would lead to greater drug–placebo difference when the expectancy of receiving placebo is high.

All the above studies, however, have several problems. First, previous meta-analyses have unfortunately often been subject to publication bias. Analysis of the trial data submitted to Food and Drug Administration as a requirement of their submission process showed that only half of the phase II or III placebo-controlled trials

 DOI: 10.1097/YIC.0000000000000201

had positive results, and most of the 'negative' trials had not been published (Turner *et al.*, 2008). The reported difference in ESs between two-armed and three-armed trials may be because of greater publication bias among the former as the latter RCTs may be more likely to be published even when there is no significant difference between the antidepressant of interest and placebo because the publication can focus on the comparison between the two active drugs. Second, previous studies have generally assumed that ESs of antidepressants are the same among all antidepressants. However, it has been reported that they may be markedly different (Cipriani *et al.*, 2009). Therefore, intervention effects should be examined and compared for each antidepressant separately. Third, it has been shown that an antidepressant appeared to be more effective when it was the new agent rather than the comparator, suggesting evidence of the so-called 'novelty effect' (Barbui *et al.*, 2004; Salanti *et al.*, 2010). The studies cited above (Greenberg *et al.*, 1992; Khan *et al.*, 2004; Papakostas and Fava, 2009; Sinyor *et al.*, 2010) have not taken this factor into account so that the apparently larger ES reported in two-armed studies might be because of the 'novelty effect' of the agent, which is more likely to be studied in two-armed rather than multiarmed trials when the agent is 'new' and when the trial is sponsored by the manufacturer of the drug.

The aim of the present study is therefore to compare the odds ratios (ORs) of antidepressants over placebo when examined in two-armed versus three-armed or more-armed (heretofore termed multiarmed) trials while taking into account possible differences among different antidepressants on the basis of a dataset compiled with as little publication bias as possible.

Methods

This is a secondary analysis of published and unpublished data from RCTs of antidepressants that were collected for GRISELDA, a multinational project to conduct network meta-analyses of 21 new and old antidepressants for adult major depression. The details of the study methodology have been published (Furukawa *et al.*, 2016) and thus, here, we present its summary as relevant to this secondary analysis.

Criteria for considering studies for this review

All double-blind RCTs that compared placebo with the following selected first-generation and second-generation antidepressants as monotherapy for the acute-phase treatment of depression were included: agomelatine, amitriptyline, bupropion, citalopram, clomipramine, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, nefazodone, paroxetine, reboxetine, sertraline, trazodone, venlafaxine, vilazodone, and vortioxetine. We included RCTs with patients aged 18 years or older, of both sexes, and with a primary diagnosis of unipolar major depression, diagnosed according to any standard operationalized diagnostic criteria.

Search methods for identification of studies

We searched Cochrane CENTRAL, CINAHL, EMBASE, LiLACS, MEDLINE, PSYCINFO, trial databases of the drug-approving agencies, trial registers, and homepages of pharmaceutical companies that market the included drugs up to 8 January 2016. The National Institute for Health and Care Excellence (UK) and the Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Germany) were also contacted. The reference lists of the identified RCTs and recent systematic reviews were checked. No language restriction was applied.

Data collection

Response to the treatment was defined as a reduction of at least 50% from baseline on the total score on the Hamilton Rating Scale for Depression (Hamilton, 1960), the Montgomery–Asberg Depression Rating Scale (Montgomery and Asberg, 1979), or any other validated depression scale at the end of acute-phase treatment. In the present review, acute treatment was defined as an 8-week treatment (Bauer *et al.*, 2002). If 8-week data were not available, we used data ranging between 4 and 12 weeks. When the number of responders was not reported, but the baseline mean and endpoint mean and SD of the depression rating scales were provided, we calculated the number of responding patients by using a validated imputation method (Furukawa *et al.*, 2005).

Two researchers independently examined the titles and abstracts of all reports obtained through the search strategy. Full articles of all the potentially eligible studies were then obtained and inspected by two review authors to identify trials that fulfilled the review criteria. Data from each study were extracted into a structured data abstraction form independently by two researchers. The risk of bias was assessed for each included study using the Cochrane Collaboration 'risk of bias' tool (Higgins and Green, 2011) by two independent researchers. Any disagreement was resolved through discussion or in consultation with a third member of the review team. On the basis of assessments of risks of bias for each domain, we quantified the overall risk of bias for each study as low risk if none of the domains was rated at high risk and three or fewer domains at unclear risk; as moderate risk if one domain was rated at high risk or none rated at high risk but four or more at unclear risk; or as high risk for all other cases.

Statistical analysis

For each antidepressant, we first estimated the overall ORs of response between the antidepressant and placebo by synthesizing ORs from all two-armed or multiarmed comparisons using the random-effects model. We next estimated the ratio of odds ratios (RORs) and their variance of

two-armed versus multiarmed trials for each antidepressant, and finally meta-analytically synthesized RORs across all the included antidepressants using the random-effects model. A random-effects model was used throughout because of possible clinical heterogeneity across the included trials because of differences in clinical populations, drugs, and drug dosages. A summary ROR larger than 1 would mean that two-armed RCTs show larger intervention effects compared with placebo than multiarmed trials do. Because two or more antidepressants were involved in multiarmed studies, the summary RORs were correlated (the placebo arm is the same in two estimates in the same trials in common) and we need to take account of these correlations; for example, the ROR for placebo versus agomelatine and the ROR for placebo versus paroxetine will be dependent because they include data from the same placebo arms in placebo versus agomelatine versus paroxetine trials. The synthesis of these RORs was therefore performed using a multivariate meta-analysis routine in R (rma.mv in the metafor package in R Core Team; R Foundation for Statistical Computing, Vienna, Austria) after specifying the entire variance–covariance matrix (Appendix). We used Review Manager 5.3 (Nordic Cochrane Centre, Cochrane Collaboration, Copenhagen, Denmark), Stata 14 (StataCorp LP, College Station, Texas, USA), and R to carry out the analyses.

We started the assessment of heterogeneity by visual inspection of the forest plots. We also calculated I^2 statistics (Higgins and Green, 2011) and analyzed them on the basis of the Cochrane Handbook's recommendations (I^2 values of 0–40%: might not be important; 30–60%: may represent moderate heterogeneity; 50–90%: may represent substantial heterogeneity; 75–100%: considerable heterogeneity).

Sensitivity analyses

To ascertain the robustness of our findings, we carried out the following sensitivity analyses:

(1) By excluding studies at high risk of bias.
(2) By excluding studies where primary outcomes were imputed rather than reported.
(3) By using the fixed-effect model instead of the random-effects model.

Subgroup analyses

We had a-priori planned to carry out the following subgroup analyses:

(1) Numbers of arms in the multiarmed trials separately (three-armed, four-armed, and five-armed).
(2) Type of antidepressants (tricyclic antidepressants versus new-generation antidepressants).
(3) Publication year [those published until the date of search, until 1990 (Greenberg *et al.*, 1992), and unpublished].
(4) Sponsorship (sponsored drug arms and nonsponsored drug arms in multiarmed trials).

Results

Characteristics of the randomized-controlled trials included

Three hundred and four placebo-controlled trials were identified by the electronic search. However, efficacy data were missing in 35 studies. There were no RCTs comparing milnacipran or clomipramine against placebo providing efficacy data. All placebo-controlled RCTs for fluvoxamine were three-armed or more-armed. Therefore, we could not calculate ROR for these three antidepressants. Altogether, 258 RCTs (80 454 patients) were finally included in the present analyses (Fig. 1). Table 1 presents detailed characteristics for two-armed and multiarmed RCTs. Among the 258 RCTs included in this study, 66 (25.6%) were two-armed and 192 were multiarmed, including, 139 (53.9%) three-armed RCTs, 43 (16.7%) four-armed RCTs, and 10 (3.9%) five-armed RCTs. The median sample size of each active arm was 98.5 (first quartile, 43.5; third quartile, 158) for two-armed trials and 118.5 (first quartile, 66; third quartile, 157) for multiarmed trials. The median number of studies per antidepressant was 13.5 (range: 5–46). Figure 2 summarizes the risk of bias of the studies included. All in all, 46 studies were rated as being at low risk of bias, 214 at moderate risk of bias, and 64 at high risk of bias.

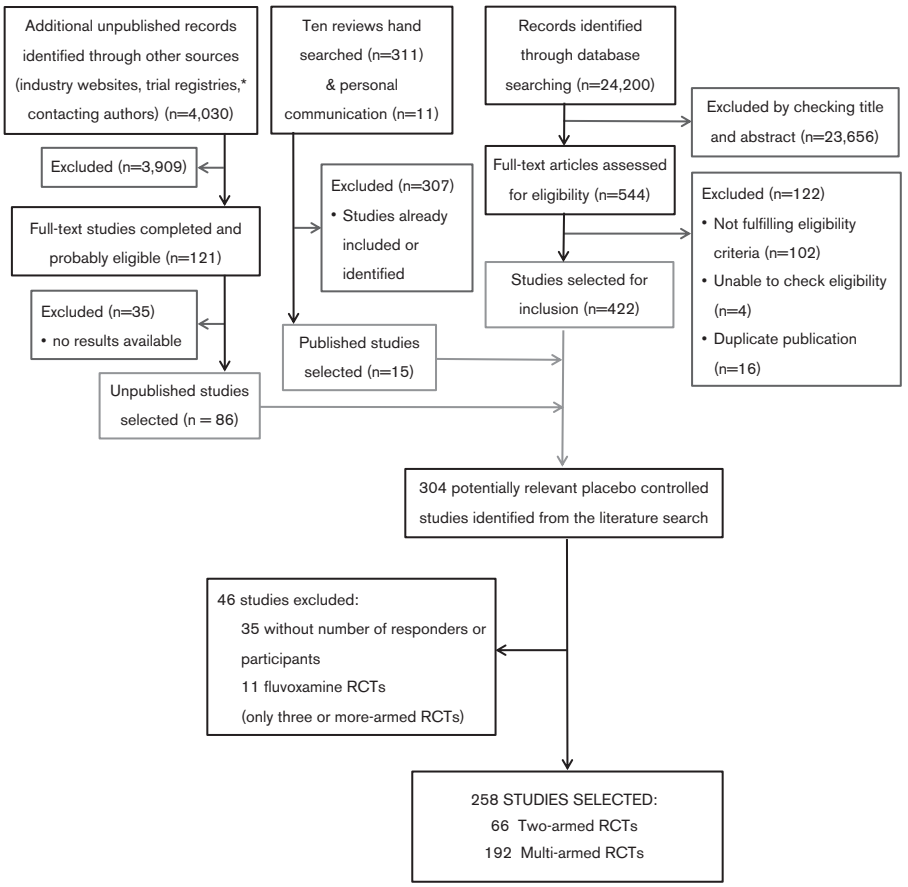
Differences in effect size between two-armed and multiarmed randomized-controlled trials

Pooled response rates for the two treatment groups (antidepressants and placebo) were 45.8 and 31.4% in two-armed RCTs and 49.7 and 37.6% in multiarmed RCTs, respectively (Fig. 3). There was no significant difference between two-armed and multiarmed RCTs in the OR of response between antidepressant and placebo [pooled ROR: 1.09; 95% confidence interval (CI): 0.96–1.24] (Fig. 4). The antidepressants are listed in the order of their approval. There was small to moderate heterogeneity in RORs across antidepressants ($I^2 = 39.6\%$; 95% CI: 0.0–65.6%). Because taking account of the covariance had little influence on the estimated ROR (the simple pooled ROR was 1.09 ($I^2 = 38.6\%$; 95% CI: 0.96–1.24), the following sensitivity and subgroup analyses were carried out without accounting for the covariances because of multiarmed studies.

Sensitivity analyses

After exclusion of studies at high risk of bias, the ROR was 1.06 ($I^2 = 34\%$; 95% CI: 0.92–1.21). After exclusion of studies that imputed the number of responders, ROR was 1.06 ($I^2 = 45\%$; 95% CI: 0.90–1.25). Using the fixed-effect model instead of the random-effects model, ROR was 1.09 ($I^2 = 38.6\%$; 95% CI: 0.99–1.19).

Fig. 1



Flow diagram. RCT, randomized-controlled trial.

Subgroup analyses

The pooled ROR was 1.12 ($I^2 = 22\%$; 95% CI: 0.99–1.26) for two-armed versus three-armed RCTs, 1.03 ($I^2 = 33\%$; 95% CI: 0.87–1.22) for two-armed versus four-armed RCTs, and 1.10 ($I^2 = 36\%$; 95% CI: 0.84–1.43) for two-armed versus five-armed RCTs. ROR of tricyclic antidepressant versus placebo was 2.00 (95% CI: 0.39–10.32) and that of new-generation antidepressants versus placebo was 1.09 ($I^2 = 41\%$; 95% CI: 0.96–1.24). ROR was 1.08 ($I^2 = 40\%$; 95% CI: 0.93–1.25) on the basis of the studies published up to the date of search (i.e. by excluding all unpublished studies), 2.34 ($I^2 = 0\%$; 95% CI: 0.57–9.66) on the basis of the studies up to 1990, and 1.19 ($I^2 = 0\%$; 95% CI: 0.93–1.51) on the basis of the studies that were not published. Similar results were obtained when the drugs in multiarmed studies were marketed by the sponsor of the drug (ROR = 1.09; $I^2 = 32\%$; 95% CI: 0.96–1.25) or when they were not (ROR = 1.07; $I^2 = 37\%$; 95% CI: 0.90–1.28).

Discussion

The differences between the two-armed versus multiarmed studies were much smaller than those found in previous studies, and were not statistically significant. For this study we used the data of the largest systematic review of antidepressants including 66 two-armed RCTs and 192 multiarmed RCTs, corresponding to 80 454 patients. The results of subgroup and sensitivity analyses did not alter this conclusion. RORs appeared larger for tricyclic antidepressants and for studies before 1990, but were not statistically significant either.

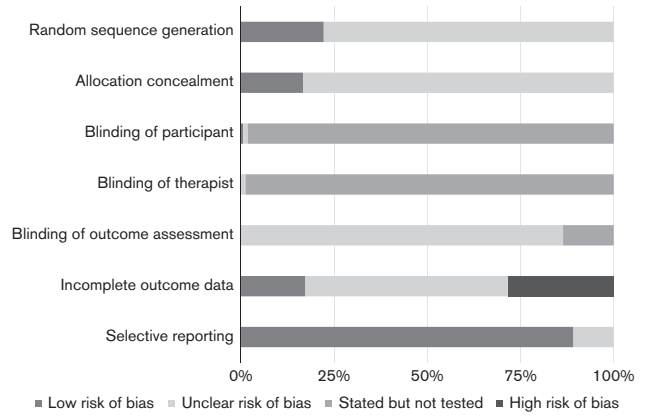
The differences between the previous studies and the present study may be explained as follows: first, the publication bias in our dataset is reduced as we could find unpublished information for 43.0% of the included studies through contacts with pharmaceutical companies and regulatory agencies. We could thus include the largest number of trials to date (258 trials), in comparison with 22 (Greenberg *et al.*, 1992), 52 (Khan *et al.*, 2004), 90 (Sinyor *et al.*, 2010), or 182 (Papakostas and Fava, 2009). Second, we used the random-effects model, which produces wider 95% CI than the fixed-effect model in the presence of heterogeneity. Although the overall the ORs tended to be higher in two-armed studies than multiarmed ones, the differences did not reach statistical

Table 1 Characteristics of two-armed and three-armed or more-armed randomized-controlled trials

	Two-armed RCTs ($n=66$)	Multiarmed RCTs ($n=192$)
Number of RCTs [*n* (%)]	Two-armed: 66 (25.6)	Three-armed: 139 (53.9) Four-armed: 43 (16.7) Five-armed: 10 (3.9)
Sample size per active arm [median (interquartile range)]	98.5 (43.5, 158)	118.5 (66, 157)
Antidepressants examined (*n* of trials, *n* of participants)		
Amitriptyline	2, 176	27, 3112
Trazodone	2, 794	8, 517
Fluoxetine	6, 1018	33, 7431
Bupropion	8, 1531	16, 4144
Sertraline	5, 1374	14, 2775
Paroxetine	8, 734	38, 8899
Venlafaxine	2, 290	20, 4895
Nefazodone	1, 120	8, 1242
Mirtazapine	3, 297	10, 1450
Reboxetine	4, 368	7, 2244
Citalopram	2, 358	11, 3428
Escitalopram	3, 956	16, 5133
Duloxetine	5, 1599	16, 4673
Agomelatine	5, 1112	8, 3061
Desvenlavaxine	2, 876	7, 3503
Vilazodone	4, 1629	4, 1841
Levomilnacipran	3, 1362	2, 1292
Vortioxetine	1, 600	13, 5620
Year of publication [*n* (%)]		
1979–1990	7 (11)	24 (13)
1991–2000	17 (26)	45 (23)
2001–2016	29 (44)	79 (41)
Unpublished	13 (20)	44 (23)

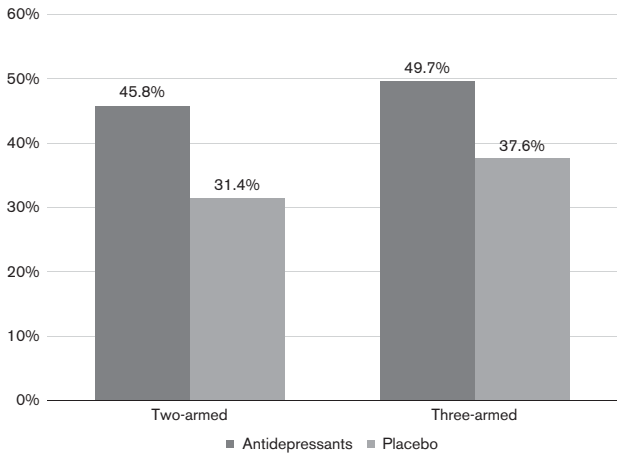
RCT, randomized-controlled trial.

Fig. 2



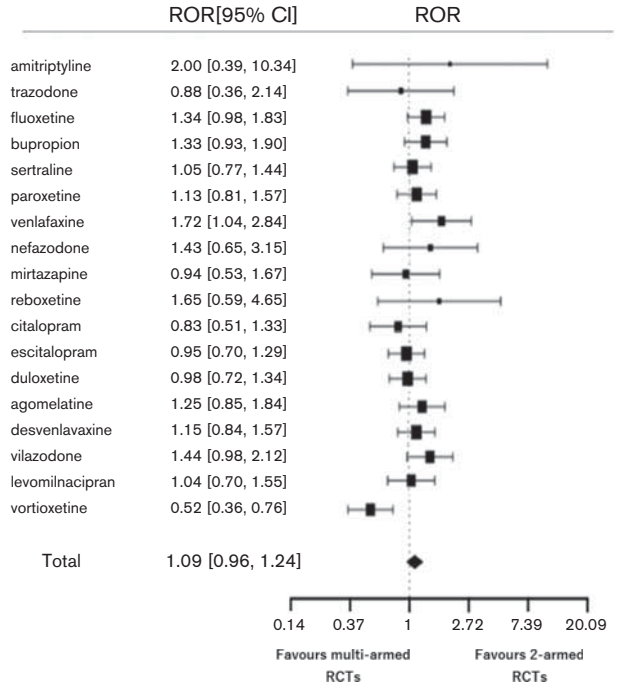
'Risk of bias' graph: review authors' judgments of each risk of bias item presented as percentages across all included studies.

Fig. 3



Antidepressant and placebo response rates.

Fig. 4



Ratio of odds ratios (ROR) between two-armed and multiarmed randomized-controlled trials (RCTs). The antidepressants are listed in the order of their approval. CI, confidence interval.

significance. We believe that our study had performed a more methodologically rigorous synthesis by estimating the ROR for each antidepressant, and then meta-analytically pooling all the RORs of the included antidepressants, instead of assuming a common efficacy for all the antidepressants included. A sensitivity analysis using the fixed-effect model instead of the random-effects model confirmed the primary findings. Third, the novelty effect (Barbui *et al.*, 2004; Salanti *et al.*, 2010) did not appear to be at play to explain the possible differences between two-armed versus multiarmed studies because our subgroup analysis found little difference when the drugs in multiarmed studies were marketed by the sponsor of the drug or when they were not.

Sinyor *et al.* (2010) showed that the response rate for placebo was significantly higher in three-armed studies than in two-armed studies; thus, it is difficult to show the superiority of drugs in studies with more active treatment arms. Although the placebo response rate in multiarmed studies was indeed larger than that in two-armed studies in our dataset, so was the response rate on antidepressant drugs (Fig. 3), resulting in the similar relative efficacy of drugs over placebo in both types of trials (Fig. 4).

Our study has some limitations. We could not consider other trial and patient features that may have an impact on intervention effects, such as the difference in rating scales, countries and cultures, the proportion of melancholic depression, depression severity, and duration of the illness or the number of depressive episodes. Systematic differences in these characteristics between two-armed versus multiarmed studies might have played a role, but we would need individual participant data to examine such effect modifiers. Moreover, given that the field of antidepressant trials in the past has been prone to publication bias, we cannot completely rule out the possibility that some studies are still missing.

In summary, we found that intervention effects were not significantly different between two-armed and multiarmed RCTs. Our original hypotheses that possible breach of the double-blinding in antidepressant clinical trials or the lower expectancy for the active drug in two-armed rather than multiarmed trials would lead to overestimation of antidepressant efficacy was not borne out. Our results were different from those in the previous studies possibly because we appropriately took into account differences among different antidepressants through the random-effects model and also because we could minimize the publication bias.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science to Y.O. (16K09033). A.C. is supported by the NIHR Oxford cognitive health Clinical Research Facility. G.S. is a Marie Skłodowska-Curie fellow.

Conflicts of interest

T.A.F. has received lecture fees from Eli Lilly, Janssen, Meiji, Mitsubishi Tanabe, MSD, and Pfizer, and consultancy fees from Takeda Science Foundation. He has received research support from Mochida and Mitsubishi Tanabe. N.T. has received lecture fees from Otsuka and Meiji. Y.H. has received lecture fees from Yoshitomi. S.T. has received lecture fees from Kobe City, Astra-Zeneca, Taiho Pharmaceutical, and Ono Pharmaceutical. He has received consultation fees from the Pharmaceuticals and Medical Devices Agency, DeNA Life Science, and CanBus. He has received outsourcing fees from Public Health Research Foundation, Japan Breast Cancer Research Group, Satt, and Asahi Kasei Pharma. S.T. has received grants from the Japan Agency for Medical Research and Development, the Japanese Ministry of Health Labor and Welfare, and the Japanese Ministry of Education, Science, and Technology. He engaged in a research project of the Japan Agency for Medical Research and Development. His wife had engaged in a research project of Bayer Yakuhin. A.C. was expert witness for a patent issue about quetiapine extended release. For the remaining authors there are no conflicts of interest.

Appendix: multivariate meta-regression to synthesize RORs

Consider that there are n_A multiarm trials (more than two arms) that involve drug A and n_B multiarm trials that involve drug B. There are also n multi-arm trials that involve both drugs A and B; these studies contribute correlated data to the estimation of $OR^{A}V^{P}$ and $OR^{B}V^{P}$. Consequently, the two ratios of odds ratios

$$\mathrm{ROR}^{\mathrm{A}} = \frac{\mathrm{OR}^{\mathrm{AvP}} \text{ in two-armed studies}}{\mathrm{OR}^{\mathrm{AvP}} \text{in } n_{\mathrm{A}} \text{ multiarmed studies}},$$

$$\mathrm{ROR}^{\mathrm{B}} = \frac{\mathrm{OR}^{\mathrm{BvP}} \text{ in two-armed studies}}{\mathrm{OR}^{\mathrm{BvP}} \text{ in } n_{\mathrm{B}} \text{ multiarmed studies}},$$

are correlated because their denominators are correlated. We need to estimate the covariance $c(\log \mathrm{ROR}^{\mathrm{A}}, \log \mathrm{ROR}^{\mathrm{B}})$. Assuming a fixed-effects model and that the study weights are known and fixed, it is easy to show that:

$$c(\log\mathrm{ROR}^{\mathrm{A}}, \log\mathrm{ROR}^{\mathrm{B}}) =$$

$$c(\log\mathrm{OR}^{\mathrm{A}}\mathrm{V}^{\mathrm{P}} \text{ in multiarm studies;}$$

$$\log\mathrm{OR}^{\mathrm{A}}\mathrm{V}^{\mathrm{P}} \text{ in multiarm studies}).$$

Consequently:

$$c(\log\mathrm{ROR}^{\mathrm{A}}, \log\mathrm{ROR}^{\mathrm{B}}) = \frac{\sum_{\mathrm{i}}^{n} w_{\mathrm{i}}^{\mathrm{A}} w_{\mathrm{i}}^{\mathrm{B}} \left(\frac{1}{S_{\mathrm{i}}} + \frac{1}{F_{\mathrm{i}}}\right)}{\sum_{\mathrm{i}}^{n_{\mathrm{A}}} w_{\mathrm{i}}^{\mathrm{A}} \sum_{\mathrm{i}}^{n_{\mathrm{B}}} w_{\mathrm{i}}^{\mathrm{B}}},$$

where $w_{\mathrm{i}}^{\mathrm{A}}$ is the inverse of the variance of log $\mathrm{OR}^{\mathrm{AvP}}$ in the multiarm study i; $w_{\mathrm{i}}^{\mathrm{B}}$ is the inverse of the variance of log $\mathrm{OR}^{\mathrm{BvP}}$ in the multiarm study i; S_{i} is the number of successes in the placebo arm in the multiarm study i; F_{i} is the number of failures in the placebo arm in the multiarm study i.

The synthesis of the RORs was performed using a multivariate meta-analysis routine in R (rma.mv in the metafor package) after specifying the entire variance–covariance matrix.

References

Barbui C, Cipriani A, Brambilla P, Hotopf M (2004). 'Wish bias' in antidepressant drug trials? *J Clin Psychopharmacol* **24**:126–130.

Bauer M, Whybrow PC, Angst J, Versiani M, Möller HJ, World Federation of Societies Biological Psychiatry Task Force on Treatment Guidelines for Unipolar Depressive Disorders (2002). World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders, part 1: acute and continuation treatment of major depressive disorder. *World J Biol Psychiatry* **3**:5–43.

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, *et al.* (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* **373**:746–758.

Even C, Siobud-Dorocant E, Dardennes RM (2000). Critical approach to antidepressant trials. Blindness protection is necessary, feasible and measurable. *Br J Psychiatry* **177**:47–51.

Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N (2005). Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol* **20**:49–52.

Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, *et al.* (2016). Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ Open* **6**:e010919.

Greenberg RP, Bornstein RF, Greenberg MD, Fisher S (1992). A meta-analysis of antidepressant outcome under 'blinder' conditions. *J Consult Clin Psychol* **60**:664–669.

Hamilton M (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry* **23**:56–62.

Higgins JP, Green S (2011). Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0. Available at: *http://handbook-5-1.cochrane.org/* [Accessed 16 July 2017].

Khan A, Kolts RL, Thase ME, Krishnan KR, Brown W (2004). Research design features and patient characteristics associated with the outcome of antidepressant clinical trials. *Am J Psychiatry* **161**:2045–2049.

Leucht S, Corves C, Arbter D, Engel RR, Li C, Davis JM (2009). Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis. *Lancet* **373**:31–41.

Margraf J, Ehlers A, Roth WT, Clark DB, Sheikh J, Agras WS, *et al.* (1991). How 'blind' are double-blind studies? *J Consult Clin Psychol* **59**:184–187.

Moncrieff J, Wessely S, Hardy R (2004). Active placebos versus antidepressants for depression. *Cochrane Database Syst Rev* **1**:CD003012.

Montgomery SA, Asberg M (1979). A new depression scale designed to be sensitive to change. *Br J Psychiatry* **134**:382–389.

Papakostas GI, Fava M (2009). Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol* **19**:34–40.

Salanti G, Dias S, Welton NJ, Ades AE, Golfinopoulos V, Kyrgiou M, *et al.* (2010). Evaluating novel agent effects in multiple-treatments meta-regression. *Stat Med* **29**:2369–2383.

Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, *et al.* (2010). Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses. *J Clin Psychiatry* **71**:270–279.

Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* **358**:252–260.

Original Paper

Cognitive and Behavioral Skills Exercises Completed by Patients with Major Depression During Smartphone Cognitive Behavioral Therapy: Secondary Analysis of a Randomized Controlled Trial

Toshi A Furukawa[1], MD, PhD; Masaru Horikoshi[2], PhD; Hirokazu Fujita[3], MD, PhD; Naohisa Tsujino[4], MD, PhD; Ran Jinnin[5], MD, PhD; Yuki Kako[6], MD, PhD; Sei Ogawa[1], MD, PhD; Hirotoshi Sato[7], MD; Nobuki Kitagawa[8], MD, PhD; Yoshihiro Shinagawa[9], MD, PhD; Yoshio Ikeda[10], MD, PhD; Hissei Imai[11], MD, PhD; Aran Tajika[11], MD, PhD; Yusuke Ogawa[11], MPH, MD, PhD; Tatsuo Akechi[1], MD, PhD; Mitsuhiko Yamada[12], MD, PhD; Shinji Shimodera[13], MD, PhD; Norio Watanabe[11], MD, PhD; Masatoshi Inagaki[14], MD, PhD; Akio Hasegawa[15], PhD

[1]Department of Psychiatry and Cognitive-Behavioral Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan
[2]Center of Cognitive-Behavior Therapy, National Center of Neurology and Psychiatry, Kodaira, Japan
[3]Center to Promote Creativity in Medical Education, Kochi Medical School, Kochi University, Nankoku, Japan
[4]Department of Neuropsychiatry, Toho University School of Medicine, Tokyo, Japan
[5]Department of Psychiatry and Neurosciences, Hiroshima University Graduate School of Biomedical & Health Sciences, Hiroshima, Japan
[6]Department of Psychiatry, Hokkaido University Graduate School of Medicine, Sapporo, Japan
[7]Harimayabashi Clinic, Kochi, Japan
[8]Hokudai-dori Mental Health Clinic, Sapporo, Japan
[9]Shiki Clinic, Nagoya, Japan
[10]Narumi Himwari Clinic, Nagoya, Japan
[11]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan
[12]National Institute of Mental Health, National Center of Neurology and Psychiatry, Kodaira, Japan
[13]Department of Psychiatry, Kochi Medical School, Kochi University, Nankoku, Japan
[14]Okayama University Hospital, Okayama, Japan
[15]Advanced Telecommunications Research Institute International, Kyoto, Japan

Corresponding Author:
Toshi A Furukawa, MD, PhD
Department of Psychiatry and Cognitive-Behavioral Medicine
Nagoya City University Graduate School of Medical Sciences
Kawasumi 1, Mizuho-cho, Mizuho-ku,
Nagoya, 467-8601
Japan
Phone: 81 52 753 9491
Email: furukawa@kuhp.kyoto-u.ac.jp

Abstract

Background: A strong and growing body of evidence has demonstrated the effectiveness of cognitive behavioral therapy (CBT), either face-to-face, in person, or as self-help via the Internet, for depression. However, CBT is a complex intervention consisting of several putatively effective components, and how each component may or may not contribute to the overall effectiveness of CBT is poorly understood.

Objective: The aim of this study was to investigate how the users of smartphone CBT use and benefit from various components of the program.

Methods: This is a secondary analysis from a 9-week, single-blind, randomized controlled trial that has demonstrated the effectiveness of adjunctive use of smartphone CBT (Kokoro-App) over antidepressant pharmacotherapy alone among patients with drug-resistant major depressive disorder (total n=164, standardized mean difference in depression severity at week 9=0.40, J Med Internet Res). Kokoro-App consists of three cognitive behavioral skills of self-monitoring, behavioral activation, and cognitive restructuring, with corresponding worksheets to fill in. All activities of the participants learning each session of the

program and completing each worksheet were uploaded onto Kokoro-Web, which each patient could use for self-check. We examined what use characteristics differentiated the more successful users of the CBT app from the less successful ones, split at the median of change in depression severity.

Results: A total of 81 patients with major depression were allocated to the smartphone CBT. On average, they completed 7.0 (standard deviation [SD] 1.4) out of 8 sessions of the program; it took them 10.8 (SD 4.2) days to complete one session, during which they spent 62 min (SD 96) on the app. There were no statistically significant differences in the number of sessions completed, time spent for the program, or the number of completed self-monitoring worksheets between the beneficiaries and the nonbeneficiaries. However, the former completed more behavioral activation tasks, engaged in different types of activities, and also filled in more cognitive restructuring worksheets than the latter. Activities such as "test-drive a new car," "go to a coffee shop after lunch," or "call up an old friend" were found to be particularly rewarding. All cognitive restructuring strategies were found to significantly decrease the distress level, with "What would be your advice to a friend who has a similar problem?" found more helpful than some other strategies.

Conclusions: The CBT program offered via smartphone and connected to the remote server is not only effective in alleviating depression but also opens a new avenue in gathering information of what and how each participant may utilize the program. The activities and strategies found useful in this analysis will provide valuable information in brush-ups of the program itself and of mobile health (mHealth) in general.

Trial Registration: Japanese Clinical Trials Registry UMIN CTR 000013693; https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000015984 (Archived by WebCite at http://www.webcitation.org/6u6pxVwik)

(JMIR Ment Health 2018;5(1):e4) doi:10.2196/mental.9092

KEYWORDS

major depressive disorder; smartphone; cognitive therapy; telemedicine

Introduction

Cognitive behavioral therapy (CBT) is the psychotherapy with the strongest evidence base for the treatment of depression [1-3]. CBT is indeed the only psychotherapy that has been shown to beat the pill placebo condition, the gold standard control condition in the evaluation of medical interventions [4]. It has also been demonstrated to show comparable efficacy as antidepressant pharmacotherapy, which is the mainstay of the treatment for major depression today [5].

The broad umbrella term of CBT, however, now subsumes various and different behavioral and cognitive skills such as self-monitoring, behavioral activation, cognitive structuring, assertion training, structured problem solving, mindfulness, and others [6]. The relative contributions of these various components to the overall efficacy of CBT remain uncertain and debated [7-9]. So-called dismantling studies or component studies to disentangle individual constituents of broadly conceived CBT have been largely underpowered and inconclusive, as each study can only examine the value of adding one particular component in question in a relatively limited number of patients [10,11]. Another major issue of such studies is whether the intended components are actually administered by the therapists and received by the patients, although more recent trials attempt to assure their delivery through audio or video recordings.

The development of information and communication technologies, however, is opening new opportunities to monitor the delivery of CBT skills and to study differential contribution of various components of CBT. The CBT itself can be delivered remotely [12], and the patients' progress can be remotely monitored [13,14]. Ecological momentary assessment or experience sampling enables more fine-tuned follow-up of patients' usage of and responses to the program [15,16]. We have developed a smartphone CBT app, named Kokoro-App (*kokoro* means *mind* in Japanese), with the integrated Kokoro-Web secure server to which all the activities of the patients with the app are uploaded. Kokoro-App teaches three distinctive CBT skills, namely self-monitoring, behavioral activation, and cognitive restructuring and provides interactive worksheets that the patients can fill in for each task.

The effectiveness of the system was demonstrated in a randomized controlled trial (RCT) comparing antidepressant medication switch plus Kokoro-App against antidepressant medication switch alone among patients who had been unresponsive to one or more antidepressants: the effect size of the intervention was a standardized mean difference of 0.40 in depression severity as measured by masked assessors (P<.001) [17]. This study aims to examine how the patients used the smartphone CBT app during the trial and to investigate what use characteristics differentiated the more successful users of the CBT app from the less successful ones.

Methods

Study Design

The original study was a 9-week, multicenter, parallel-group RCT comparing antidepressant medication switch plus smartphone CBT against medication switch alone among patients with antidepressant-resistant depression [17] (Japanese Clinical Trials Registry UMIN CTR 000013693). A total of 164 patients who had not responded to one or more antidepressants at adequate dosage for 4 or more weeks [18] were randomized 1:1 to the intervention or the control arm. The RCT has been registered in the Japanese trials registry (UMIN CTR 000013693).

The randomized comparison showed that the adjunctive use of smartphone CBT brought about 2.5 (95% CI 1.2-3.7, *P*<.001) points greater reduction in the Patient Health Questionnaire-9 (PHQ-9) [19] scores and 4.1 (95% CI 1.5-6.6, *P*=.002) points greater reduction in the Beck Depression Inventory-II [20] scores after 9 weeks [17]. This study focuses on the 81 patients who were randomized to the smartphone CBT arm and describes and analyzes the patients' use of Kokoro-App.

Kokoro-App

Kokoro-App is a smartphone CBT app and consists of four parts: sessions, mind maps, actions, and thoughts (Figure 1).

There are eight sessions in which several cartoon characters provide psychoeducation through easy but fun conversations. First, the welcome session explains CBT, as well as how to use iPhone and Siri (voice recognition on iPhone). Sessions 1 and 2 explain how to self-monitor one's reactions to various situations according to the cognitive behavioral model. The sessions introduce mind maps in which the patient can enter details of the situation and his reactions to it in terms of emotion and its degree, automatic thoughts, bodily reactions, and behaviors. The patient chooses between four emotions of sad or depressed, anxious or worried, angry, and happy and rates its strength in five grades between 0 and 5.

Sessions 3 and 4 explain behavioral activation according to two principles of "When the body moves, so does the mind" and "Start small and near." When the patient clicks on actions, lists of candidate activities for behavioral activation personal experiments pop up. The candidates are categorized by the usual time they require to complete into (1) less than 5 seconds, (2) less than 5 min, (3) less than 60 min, and (4) 60 min or more. The patient chooses a candidate and rates his expected mastery and pleasure levels. When the patient completes the personal experiment, he can enter his achieved mastery and pleasure levels. The patient can also enter his own personal experiment task. After his own experiments, the patient can recommend certain activities by clicking on "Nice!" button and the number of "Nice!"s will be shared by all the patients.

Sessions 5 and 6 explain cognitive restructuring. After providing a rationale for cognitive restructuring, the app provides four interactive items to guide the patient to alternative thoughts. The patient first picks up a mind map to work on. The first item, "fact glasses," asks classic questions about evidence for and evidence against the automatic thought, such as "What facts do you have to support this thought?" and "What facts are there that do not support this thought?" Then the item combines the patient's answers automatically and says, "So you believe XXX but YYY. If you think this way, how do you feel now?" and asks the patient to rerate his feeling.

Figure 1. A screenshot from Kokoro-App.

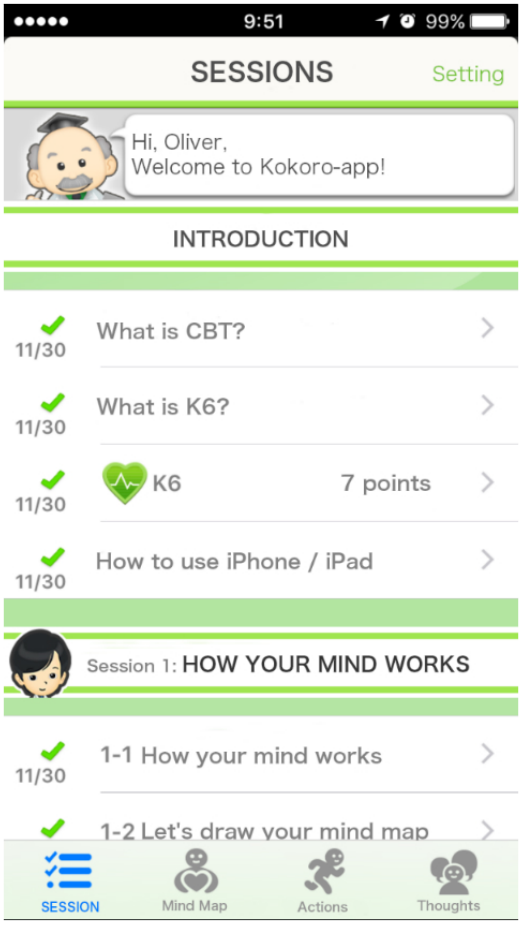
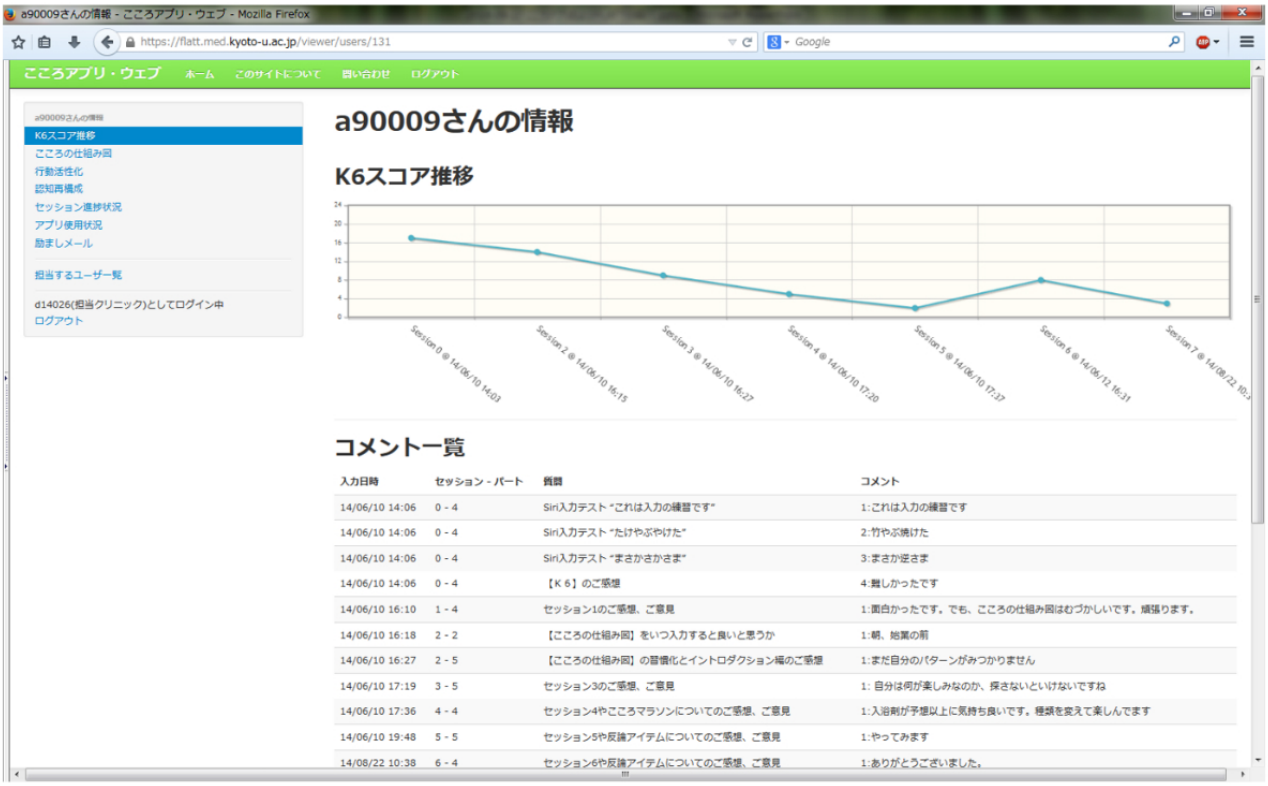


Figure 2. A screenshot from Kokoro-Web.



The second item is called "% Calculator," which does similar things as the fact glasses. It asks, "How confident are you in your thought AAA?" then lets the patient choose between 1% and 99%. Then it asks, "So you think your thought AAA is 99% correct. But then what can there be in your other 1%?" The patient will then answer XXX, and the item will then ask, "So if you think XXX, how would you feel now?" The third item is "friend's call." The item says, "Ring, ring, ring. You have just received a phone call from your best friend, saying AAA. What advice would you give to her?" The rest is the same. The last item is "What-now microphone." The item goes, "Let's just suppose that your thought AAA is true. If so, what can be done now?" The patient will then make an action statement, and the item will then ask, "What if you did do XXX, how would you be feeling?"

The epilogue session summarizes all the previous sessions and also provides tips for relapse prevention.

Each session is expected to take 1 week. The new session can be opened only a week after the last session was started and after one homework has been completed.

Kokoro-Web

All the activities of the patient with Kokoro-App are uploaded to the central server and can be viewed on Kokoro-Web (Figure 2) by the patient, as well as by his treating physician. Kokoro-Web was developed seamlessly and integratively with Kokoro-App. The communication between the app and the server through the Internet was certificated by Secure Sockets Layer.

Statistical Analyses

We first present the descriptive details of how the patients utilized Kokoro-App. The continuous outcomes are summarized by mean and SD and the dichotomous outcomes by number and percentage.

We next subdivide the patients into beneficiaries and nonbeneficiaries from Kokoro-App at the median change score of the PHQ-9 and compare each group's use of the Kokoro-App. Because the same patient contributed a variable number of mind maps, behavioral activations, or cognitive restructurings to account for the within-person clustering effect, we used the mixed effects model where appropriate. Given the observational and hypothesis-generating nature of this study, we set the threshold for statistical significance for each comparison at nominal $P<.05$ throughout. We used STATA (StataCorp) version 14.

Results

Patient Characteristics

Table 1 summarizes the baseline demographic and clinical characteristics of the cohort. Patients were typically around 40 years of age, had some higher education, slightly more than half were in some employment, and slightly less than half were married. They had had several depressive episodes, had been in the current depressive episode for almost 2 years, and were in moderately to severely depression at baseline.

Table 1. Baseline demographic and clinical characteristics of the cohort.

Characteristics	Mean SD[a] or n (%)
Demographic	
Age (years), mean (SD)	40.2 (8.8)
Women, n (%)	46 (57)
Education (years), mean (SD)	14.6 (2.5)
Employment status	
Employed full-time, n (%)	34 (42)
Employed part-time, n (%)	7 (9)
On medical leave, n (%)	21 (26)
Housewife, n (%)	6 (7)
Student, n (%)	0 (0)
Retired, n (%)	0 (0)
Not employed, n (%)	13 (16)
Marital status	
Single, never married, n (%)	34 (42)
Single, divorced, separated or widowed, n (%)	13 (16)
Married, n (%)	34 (42)
Clinical	
Age of onset at first episode (years), mean (SD)	31.8 (10.8)
Number of previous depressive episodes, mean (SD)	3.4 (4.9)
Length of current episode (months), mean (SD)	24.2 (46.3)
PHQ-9[b] at baseline, mean (SD)	13.5 (5.5)
BDI-II[c] at baseline, mean (SD)	28.2 (11.2)

[a]SD: standard deviation.

[b]PHQ-9: Patient Health Questionnaire-9.

[c]BDI-II: Beck Depression Inventory 2nd edition.

Kokoro-App Use Statistics

Table 2 shows the use statistics of Kokoro-App by the 81 patients.

On average, the patients completed 7.0 out of 8 sessions; it took them 10.8 days to complete one session, during which they spent 62 min on the app reading the sessions and also completing their respective homework (self-monitoring, behavioral activation, or cognitive restructuring).

They filled in 10 mind maps, more often for sad or depressed, or anxious or worried feelings but also for angry or happy feelings. They performed 14 behavioral activation personal experiments through which they had anticipated and achieved moderate levels of mastery and pleasure. With respect to cognitive restructuring, they generated an average of six alternative thoughts using fact glasses, % Calculator, friend's call, or what-now microphone almost equally frequently.

Behavioral Activations

We analyzed the behavioral activations completed by the patients according to their frequencies, the level of mastery and pleasure they achieved, and how unexpectedly good they were.

The most frequently chosen behavioral activations were, in the descending order, "Listen to favorite music," "Read books and magazines," "Brew and drink coffee," and so on (Table 3). The levels of mastery or pleasure expected or achieved were typically in the range 3 to 5 on a scale of 0 to 10.

However, when selected for the levels of achieved mastery or pleasure, very different sets of activities emerged (Tables 4 and 5). These tables are limited to such activities that were reported at least three times. Activities that achieved very high levels of mastery or pleasure included "Test-drive a new car," "Go to a coffee shop after lunch," "Call up an old friend," and "Exercise."

Table 2. Use statistics of Kokoro-App.

Use statistics	Mean (SD[a]), range, and median
Overall	
Sessions completed, mean (SD), range	7.0 (1.4), 1-8
Days taken to complete one session, mean (SD), range	10.8 (4.2), 6.3-31
Actual time (min) per session, mean (SD), range, median	62.3 (96.30), 0-677, 39
Self-monitoring	
Mind maps, no. completed per person, mean (SD), range	10.4 (10.5), 0-45
Mind maps, no. completed per person, by emotion	
Sad or depressed, mean (SD), range	3.2 (3.7), 0-16
Anxious or worried, mean (SD), range	3.0 (3.3), 0-18
Angry, mean (SD), range	2.4 (3.4), 0-16
Happy, mean (SD), range	1.9 (3.1), 0-20
Level of emotion recorded (on a scale of 0-5)	
Sad or depressed, mean (SD)	3.4 (1.3)
Anxious or worried, mean (SD)	3.5 (1.3)
Angry, mean (SD)	3.5 (1.3)
Happy, mean (SD)	3.2 (1.3)
Behavioral activation	
Behavioral activations, no. completed per person, mean (SD), range	13.8 (17.3), 0-118
Level of mastery or pleasure by behavioral activation (on a scale of 0-10)	
Mastery expected, mean (SD)	4.5 (2.8)
Mastery achieved, mean (SD)	4.3 (3.0)
Pleasure expected, mean (SD)	4.8 (2.7)
Pleasure achieved, mean (SD)	4.7 (2.9)
Cognitive restructuring	
Cognitive restructuring, no. completed per person, mean (SD), range	6.2 (6.3), 0-31
Cognitive restructuring items used, per person	
Fact glasses, mean (SD), range	2.0 (2.0), 0-11
% Calculator, mean (SD), range	1.5 (1.8), 0-10
Friend's call, mean (SD), range	1.5 (1.6), 0-7
What-now microphone, mean (SD), range	1.4 (1.5), 0-7

[a]SD: standard deviation.

Table 3. Behavioral activations: top 10 activities in terms of frequency and their mastery or pleasure levels.

Activity	Frequency (number of reports)	Mastery		Pleasure	
		Expected Mean (SD[a])	Achieved Mean (SD)	Expected Mean (SD)	Achieved Mean (SD)
Listen to favorite music	95	5.2 (3.0)	5.0 (3.3)	5.9 (2.7)	5.7 (2.9)
Read books and magazines	71	4.9 (3.0)	4.5 (3.3)	5.3 (2.7)	4.8 (2.8)
Brew and drink coffee	50	3.5 (3.2)	3.1 (2.6)	3.8 (2.2)	3.8 (2.5)
Hum a tune	41	2.9 (2.5)	2.3 (1.9)	3.8 (2.1)	3.4 (2.2)
Take a long bath	36	3.6 (2.0)	3.9 (2.5)	4.4 (1.9)	4.5 (2.4)
Throw away something you don't need from the drawer	33	4.3 (1.9)	3.9 (2.4)	3.6 (2.0)	3.5 (2.1)
Go to a coffee shop after lunch	27	8.4 (1.8)	8.3 (2.2)	8.4 (1.8)	8.3 (2.1)
Put some bath powder in the bathtub	24	2.9 (2.0)	3.3 (2.1)	4.1 (2.0)	4.4 (2.0)
Close your eyes for 3 min	21	2.9 (1.7)	2.6 (2.2)	2.7 (1.1)	2.5 (2.3)
Take a different route on the way back home	20	3.0 (1.4)	3.4 (2.3)	3.1 (1.1)	3.3 (2.2)

[a]SD: standard deviation.

Table 4. Behavioral activations: top 10 activities in mastery achieved and their frequencies.

Activity	Frequency (number of reports)	Mastery achieved, mean (SD[a])
Test-drive a new car	3	9.7 (0.6)
Go to a coffee shop after lunch	27	8.3 (2.2)
Exercise	3	8.3 (2.9)
Call up an old friend	4	7.8 (2.9)
Go to yoga with a friend	6	7.2 (1.2)
Go to a hairdresser	3	6.7 (1.5)
Go to a gym	7	6.6 (3)
Walking	8	6.5 (0.8)
Get a haircut	4	6.5 (3)
Go to a meal with a friend	13	6.3 (2.8)

[a]SD: standard deviation.

Table 5. Behavioral activations: top 10 activities in pleasure achieved and their frequencies.

Activity	Frequency (number of reports)	Pleasure achieved, mean (SD[a])
Test-drive a new car	3	9.7 (0.6)
Call up an old friend	4	9.0 (1.4)
Exercise	3	8.7 (2.3)
Go to a coffee shop after lunch	27	8.3 (2.1)
Go to yoga with a friend	6	7.8 (1)
Call up a family and hear their voice	5	7.4 (1.5)
Go to a meal with a friend	13	7.2 (2.3)
Go out for a luxurious lunch	17	6.8 (2.9)
Take a walk	8	6.6 (2)
Nail art	4	6.3 (3.8)

[a]SD: standard deviation.

Some activities brought greater levels of mastery and pleasure than initially expected. Such pleasant surprises included "Exercise," "Do a makeup," "Buy a comic book at a convenience store," "Call up an old friend," or "Go out for a luxurious lunch" (Tables 6 and 7).

Cognitive Restructurings

All the cognitive restructuring items showed statistically significant reductions in sad or depressed, anxious or worried, or angry feelings when the emotion levels were compared pre-post within each situation that the patient worked on (Table 8). Typically, the level of emotion went down from approximately 3.5 to 2.1, showing a reduction greater than 1 point, on a scale of 0 to 5.

When the four tools were compared against each other, again within each situation, friend's call and % Calculator both outperformed fact glasses. The average change in emotion level was −1.6 (SD 1.3), −1.5 (SD 1.3), −1.4 (SD 1.3), and −1.3 (1.3), respectively, for friend's call, what-now microphone, % calculator, and fact glasses.

Contrasts Between Beneficiaries and Nonbeneficiaries of Kokoro-App

The median of the final change score on PHQ-9 was 4. We therefore split the cohort into beneficiaries from Kokoro-App (change greater than 4: n=31) and nonbeneficiaries (change of 4 or less: n=49, including six who showed deterioration from baseline).

Although the beneficiaries tended to complete more sessions, need fewer days to complete one session, and spent more time per session, the group differences were not statistically significant (Multimedia Appendix 1).

Neither did the numbers of mind maps completed, overall and by emotion differ between the two groups, although the beneficiaries tended to report a slightly higher level of happy emotion.

Table 6. Behavioral activations: top 10 activities in unexpected mastery and their frequencies.

Activity	Frequency (number of reports)	Mastery achieved-expected, mean (SD[a])
Exercise	3	1.3 (2.3)
Buy a comic book at a convenience store	4	1.0 (1.4)
Call up an old friend	3	1.0 (1.7)
Do a makeup	5	0.8 (1.3)
Call up a family and hear their voice	5	0.8 (1.3)
Go see a movie	4	0.8 (1.0)
Test-drive a new car	3	0.7 (0.6)
Nail art	3	0.7 (2.1)
Take a long bath	32	0.6 (1.7)
Put some bath powder in the bathtub	21	0.5 (1.0)

[a]SD: standard deviation.

Table 7. Behavioral activations: top 10 activities in unexpected pleasure and their frequencies.

Activity	Frequency (number of reports)	Pleasure achieved-expected, mean (SD[a])
Do a makeup	5	1.0 (1.7)
Call up an old friend	3	1.0 (1.7)
Test-drive a new car	3	0.7 (0.6)
Go out for a luxurious lunch	15	0.6 (1.9)
Say hurray!	9	0.6 (0.7)
Borrow and watch a DVD	8	0.6 (2.7)
Go to a gym	7	0.6 (2.9)
Put some bath powder in the bathtub	21	0.5 (1.0)
Take a long bath	31	0.4 (1.2)
Throw away something you don't need from the drawer	25	0.4 (1.7)

[a]SD: standard deviation.

Table 8. Changes in emotion levels by cognitive restructuring items. The statistical test was done with within-situation paired *t* test.

Item and emotion	Before, mean (SD[a])	After, mean (SD)	Change, mean (SD)	*P* value
Fact glasses				
Sad or depressed (n=68)	3.5 (1.3)	2.4 (1.3)	−1.5 (1.3)	<.001
Anxious or worried (n=61)	3.7 (1.1)	2.2 (1.4)	−1.5 (1.2)	<.001
Angry (n=44)	3.8 (1.2)	2.3 (1.3)	−1.1 (1.3)	<.001
% Calculator				
Sad or depressed (n=48)	3.5 (1.3)	2.2 (1.2)	−1.3 (1.3)	<.001
Anxious or worried (n=49)	3.7 (1.1)	2.1 (1.4)	−1.6 (1.2)	<.001
Angry (n=33)	3.8 (1.3)	2.2 (1.1)	−1.5 (1.3)	<.001
Friend's call				
Sad or depressed (n=50)	3.6 (1.3)	2.1 (1.2)	−1.5 (1.3)	<.001
Anxious or worried (n=46)	3.7 (1.2)	2.1 (1.2)	−1.7 (1.3)	<.001
Angry (n=30)	3.4 (1.3)	2.0 (1.1)	−1.4 (1.3)	<.001
What-now microphone				
Sad or depressed (n=41)	3.4 (1.3)	2.0 (1.2)	−1.4 (1.3)	<.001
Anxious or worried (n=45)	3.9 (1.1)	2.1 (1.2)	−1.8 (1.2)	<.001
Angry (n=30)	3.6 (1.3)	2.4 (1.2)	−1.2 (1.3)	<.001

[a]SD: standard deviation.

The use of behavioral activation differed significantly in many aspects between the successful users and the less successful ones. The former conducted almost twice as many behavioral activations and expected and achieved greater levels of mastery or pleasure. The kinds of behavioral activation tasks chosen were significantly different: differences by more than 3% were found for activities such as "Listen to favorite music," "Read books and magazines," "Go to a coffee shop after lunch" (all more frequent among the beneficiaries), and "Take a long bath" (more frequent among nonbeneficiaries). The time categories of activities chosen were also different: the beneficiaries chose activities likely to require 60 min, whereas the nonbeneficiaries chose activities requiring 5 min or less.

The successful users of Kokoro-App also conducted more cognitive restructuring than the less successful ones, especially those using fact glasses and % Calculator. The decrease in emotion levels, however, was significantly different between the two groups only when using % Calculator.

Discussion

Summary of Findings

Kokoro-App was well accepted among the patients who had been unresponsive to one or more antidepressants and were moderately to severely depressed at baseline. The patients proceeded with the sessions in Kokoro-App at their own pace, spending approximately 60 min across 10 days for each session. Over the course of 9 weeks, on average, they completed 10 mind maps for self-monitoring own emotions and thoughts, conducted 14 behavioral activation personal experiments, and filled in six cognitive restructuring worksheets.

To the best of our knowledge, this study is the first study to examine specific details of behavioral activation or cognitive restructuring tasks conducted by the patients undergoing remote CBT or face-to-face CBT among a sizable number of clinical patients. Very interesting pictures emerged. Although patients often conducted activities with expected and achieved mastery or pleasure levels, around 4 to 5, a number of candidate activities emerged that achieved higher than expected mastery or pleasure. Such activities included, among others, "Test-drive a new car," "Go to a coffee shop after lunch," "Exercise," "Call up an old friend," "Do a makeup," and "Go out for a luxurious lunch." All the cognitive restructuring items were able to reduce the emotion levels significantly.

The study is also the first to compare details of the behavioral and cognitive skills practiced by the patients with regard to the outcome. The successful users of Kokoro-App conducted twice as many behavioral experiments of different kinds and of different time requirements than those who were less successful. The former also conducted significantly more cognitive restructuring tasks, especially using % Calculator.

Limitations of the Study

These are several caveats in the interpretation of the current findings. First, although the original study was an RCT examining the value of adjunctive use of Kokoro-App, this study is by nature an observational study of the users of Kokoro-App. The current results therefore indicate association but not necessarily causation. The amount and nature of behavioral activation tasks or cognitive restructuring worksheets are potential mediators in the causative process from using the smartphone CBT to reduction in depressive symptomatology. It is possible that the beneficiaries of Kokoro-App got better

because they engaged in more behavioral activations, or it is also possible that they conducted more behavioral activations because they had already felt better and had more energy. Second, in accordance with the observational nature of the study, we did not correct for multiple statistical testing and the analyses remain hypothesis-generating rather than confirmatory. The insights gained need be examined in future confirmatory experiments to provide ultimate guidance on how to conduct CBT. Third, strictly speaking, the findings only apply to Kokoro-App and the Japanese patients with moderate to severe depression on an antidepressant treatment. Whether they would apply to other smartphone or Internet CBT (iCBT), or whether they apply to Japanese nonclinical populations or to non-Japanese patients when they use Kokoro-App cannot be taken for granted. For example, "Take a long bath" or "Put some bath powder in the bathtub" may be particularly comforting for the Japanese people who traditionally take great pleasure in taking baths and may not necessarily apply to people living in different cultural traditions. The candidate activities list must certainly be culturally adapted when Kokoro-App is transferred to different countries. It must also be emphasized that app contents need be contextualized for each user's age, sex, personal relationships, disabilities, and so on to make them more specific.

Implications of the Study Findings

Nonetheless, our findings have important implications at three levels. First, they suggest how Kokoro-App can be improved in the next upgrade. Currently Kokoro-App lists the candidate activities according to the number of "Nice!"s that the patients have voted. This is probably a good feature of the app, creating an atmosphere of a therapeutic community. The next version of Kokoro-App can probably add another dimension to the recommendation by highlighting such activities that may not have been experimented by many but which turned out to produce great mastery or pleasure. The next version of Kokoro-App may also choose to emphasize % Calculator, and possibly do away with fact glasses as the latter has been found to be less effective than the other items. % Calculator and fact glasses aim to derive the same kinds of information, namely evidence for and evidence against the automatic thoughts but through different Socratic questions. % Calculator may be easier to understand for the users.

Second, they provide some insight on how iCBT and CBT in general can be better practiced. Our results suggest that behavioral activation best distinguishes the more from the less successful users of smartphone CBT. This finding is in line with a growing number of RCTs showing similar effectiveness of behavioral activation in comparison with the full CBT package, including cognitive restructuring [7-9]. However, these studies compared the different versions of CBT in the face-to-face settings. Whether the iCBT may as well consist only of behavioral activation or need to include cognitive restructuring is an empirical question warranting a direct randomized comparison.

Third, they also provide suggestions for the next generation of mobile health (mHealth). Providing the mHealth intervention on the Web or via a smartphone increases accessibility of the intervention but is only taking advantage of one aspect of the technology. The program can be used to collect valuable information of what the users of the program do or feel. It may also be combined with habit formation activities. Development of such an e-monitoring system has its own difficulties and complexities, including privacy, integration, and customization [21] but our Kokoro-Web presents one successful example and our study an example of how such a system enables collection of important and fertile information.

Acknowledgments

The study was funded by the Ministry of Health, Labor, and Welfare, Japan (H-22-Seishin-Ippan-008) from April 2013 to March 2014 and thereafter by the Japan Foundation for Neuroscience and Mental Health (JFNMH). The JFNMH received donations from Asahi Kasei, Eli-Lilly, GSK, Janssen, MSD, Meiji Seika, Mitsubishi-Tanabe, Mochida, Otsuka, Pfizer, Shionogi, and Taisho. The study was supported in part by JSPS KAKENHI (Grant-in-Aid for Scientific Research) grant number 17K19808 to TAF. A complete list of participating centers and investigators in the Fun to Learn to Act and Think through Technology (FLATT) trial appears in Multimedia Appendix 2.

Authors' Contributions

TAF conceived the study. TAF, MH, and NY designed the study. KF, NT, RJ, YK, SO, HS, NK, YS, YI, HI, AT, YO, NT, TA, MY, SS, NW, MI, and AH carried out the study. TAF conducted the analyses and wrote the first draft manuscript. All authors contributed to the critical revision of the draft and read and approved the final manuscript.

Conflicts of Interest

TAF has received lecture fees from Eli Lilly, Janssen, Meiji, MSD, Otsuka, Pfizer, and Tanabe-Mitsubishi and consultancy fees from Takeda Science Foundation. He has received royalties from Igaku-Shoin and Nihon Bunka Kagaku-sha publishers. He has received research support from Mochida and Tanabe-Mitsubishi. He is diplomate of the Academy of Cognitive Therapy. MH has received royalties from Igaku Shoin, Shogakukan Shuei-sha Production, Shindan-to-Chiryo-sha, Sogen-sha, Kango-Kyokai, Kitaoji-Shobo, and Kongo-Shuppan publishers. HF has received lecture fees from Meiji and Mochida. NT has received speaking fees from Astellas, Shionogi, Novartis, FUJIFILM RI Pharma, Meiji, Mochida, Janssen, Eli Lilly, and Dainippon-Sumitomo. He has received royalties from Igaku-Shoin, Nanzando, Medical View, and Kanehara publishers. YK has received speaking fee from Otsuka, Yoshitomi, Tanabe-Mitsubishi, Dainippon-Sumitomo, and Eli Lilly. SO has received speaking fee from Eli Lilly and

Mochida. He has received royalties from Igaku-Shoin and Nihon-Hyoron-sha publishers. NK has received lecture fees from Eli Lilly, Janssen, Dainippon-Sumitomo, and Otsuka and consultancy fees from Meiji and Otsuka. He has received royalties from Igaku-Shoin, Nakayama-Shoten, Seronjihou-sha, and Iwasakigakujutu-Shuppan publishers. He has received research funds from Shionogi, Pfizer, and Meiji-Seika. YO has received honoraria for speaking at meetings sponsored by Eli Lilly. NT has received lecture fees from Otsuka and Meiji. AT has received honoraria for speaking at a meeting sponsored by Eli Lilly and Tanabe-Mitsubishi. He has received royalties from Kagaku-Hyoron-sha. TA has received speaking fees or research funds from Daiichi-Sankyo, Eizai, Hisamitsu, Lilly, MSD, Meiji, Mochida, Otsuka, Pfizer, Novartis, and Terumo. He has received royalties from Igaku-Shoin, Nanzando, and Nankodo publishers. MY has received speaking fees from Meiji and has contracted research with Nippon Chemiphar. MI has received lecture fees from Pfizer, Mochida, Shionogi, Daiichi-Sankyo, Meiji, Takeda, and Sumitomo Dainippon Pharma. He has received royalties from Nippon-Hyoron-Sha, Nanzando, Seiwa-Shoten, Igaku-Shoin, and Technomics. SS has received lecture fees from Otsuka, MSD, Meiji, Eli Lilly, Mochida, Pfizer, and Tanabe-Mitsubishi. He has received royalties from Seiwa-Shoten. He has received royalties from Sentan-Igaku-sha, Chuohoki, and Medical Review publishers. NW has received royalties from Sogen-sha, Paquet and Akatsuki publishers.

Multimedia Appendix 1

Beneficiaries and nonbeneficiaries of Kokoro-App.

[PDF File (Adobe PDF File), 63KB - mental_v5i1e4_app1.pdf]

Multimedia Appendix 2

FLATT investigators and committee members.

[PDF File (Adobe PDF File), 29KB - mental_v5i1e4_app2.pdf]

References

1. Cuijpers P, van Straten A, Andersson G, van Oppen P. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol 2008 Dec;76(6):909-922. [doi: 10.1037/a0013075] [Medline: 19045960]
2. Baardseth TP, Goldberg SB, Pace BT, Wislocki AP, Frost ND, Siddiqui JR, et al. Cognitive-behavioral therapy versus other therapies: redux. Clin Psychol Rev 2013 Apr;33(3):395-405. [doi: 10.1016/j.cpr.2013.01.004] [Medline: 23416876]
3. Barth J, Munder T, Gerger H, Nüesch E, Trelle S, Znoj H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med 2013;10(5):e1001454 [FREE Full text] [doi: 10.1371/journal.pmed.1001454] [Medline: 23723742]
4. Furukawa TA, Weitz ES, Tanaka S, Hollon SD, Hofmann SG, Andersson G, et al. Initial severity of depression and efficacy of cognitive-behavioural therapy: individual-participant data meta-analysis of pill-placebo-controlled trials. Br J Psychiatry 2017 Mar;210(3):190-196. [doi: 10.1192/bjp.bp.116.187773] [Medline: 28104735]
5. Weitz ES, Hollon SD, Twisk J, van Straten A, Huibers MJH, David D, et al. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy: an individual patient data meta-analysis. JAMA Psychiatry 2015 Nov;72(11):1102-1109. [doi: 10.1001/jamapsychiatry.2015.1516] [Medline: 26397232]
6. Churchill R, Moore T, Caldwell D, Davies P, Jones H, Furukawa T, et al. Cognitive behavioural therapies versus other psychological therapies for depression. Cochrane Database Syst Rev 2010(9) [FREE Full text] [doi: 10.1002/14651858.CD008698] [Medline: 25411559]
7. Jacobson NS, Dobson KS, Truax PA, Addis ME, Koerner K, Gollan JK, et al. A component analysis of cognitive-behavioral treatment for depression. J Consult Clin Psychol 1996 Apr;64(2):295-304. [Medline: 8871414]
8. Dimidjian S, Hollon SD, Dobson KS, Schmaling KB, Kohlenberg RJ, Addis ME, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. J Consult Clin Psychol 2006 Aug;74(4):658-670. [doi: 10.1037/0022-006X.74.4.658] [Medline: 16881773]
9. Richards DA, Ekers D, McMillan D, Taylor RS, Byford S, Warren FC, et al. Cost and outcome of behavioural activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. Lancet 2016 Aug 27;388(10047):871-880 [FREE Full text] [doi: 10.1016/S0140-6736(16)31140-0] [Medline: 27461440]
10. Cuijpers P. Are all psychotherapies equally effective in the treatment of adult depression? The lack of statistical power of comparative outcome studies. Evid Based Ment Health 2016 Dec;19(2):39-42. [doi: 10.1136/eb-2016-102341] [Medline: 26984413]
11. Bell EC, Marcus DK, Goodlad JK. Are the parts as good as the whole? A meta-analysis of component treatment studies. J Consult Clin Psychol 2013 Aug;81(4):722-736. [doi: 10.1037/a0033004] [Medline: 23688145]
12. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based Cognitive Behavioral Therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 01;74(4):351-359. [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]

13. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
14. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
15. Armey MF, Schatten HT, Haradhvala N, Miller IW. Ecological momentary assessment (EMA) of depression-related phenomena. Curr Opin Psychol 2015 Aug 1;4:21-25. [doi: 10.1016/j.copsyc.2015.01.002] [Medline: 25664334]
16. aan het Rot M, Hogenelst K, Schoevers RA. Mood disorders in everyday life: a systematic review of experience sampling and ecological momentary assessment studies. Clin Psychol Rev 2012 Aug;32(6):510-523. [doi: 10.1016/j.cpr.2012.05.007] [Medline: 22721999]
17. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone cognitive-behavior therapy as adjunct to pharmacotherapy for refractory depression: a randomized controlled trial. J Med Internet Res 2017 Nov 03;19(11):e373 [FREE Full text] [doi: 10.2196/jmir.8602] [Medline: 29101095]
18. Thase ME, Rush AJ. When at first you don't succeed: sequential strategies for antidepressant nonresponders. J Clin Psychiatry 1997;58(Suppl 13):23-29 [FREE Full text] [Medline: 9402916]
19. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. J Am Med Assoc 1999 Nov 10;282(18):1737-1744. [Medline: 10568646]
20. Beck A, Steer R, Brown G. BDI-II: Beck Depression Inventory, Second Edition. San Antonia: The Psychological Corporation; 1996.
21. Sundram F, Hawken SJ, Stasiak K, Lucassen MF, Fleming T, Shepherd M, et al. Tips and traps: lessons from codesigning a clinician e-monitoring tool for computerized cognitive behavioral therapy. JMIR Ment Health 2017 Jan 11;4(1):e3 [FREE Full text] [doi: 10.2196/mental.5878] [Medline: 28077345]

Abbreviations

CBT: cognitive behavioral therapy
iCBT: Internet cognitive behavioral therapy
mHealth: mobile health
SD: standard deviation

Edited by G Eysenbach; submitted 03.10.17; peer-reviewed by N Bragazzi, S Aggarwal; comments to author 08.11.17; revised version received 16.11.17; accepted 17.11.17; published 11.01.18

Please cite as:
Furukawa TA, Horikoshi M, Fujita H, Tsujino N, Jinnin R, Kako Y, Ogawa S, Sato H, Kitagawa N, Shinagawa Y, Ikeda Y, Imai H, Tajika A, Ogawa Y, Akechi T, Yamada M, Shimodera S, Watanabe N, Inagaki M, Hasegawa A
Cognitive and Behavioral Skills Exercises Completed by Patients with Major Depression During Smartphone Cognitive Behavioral Therapy: Secondary Analysis of a Randomized Controlled Trial
JMIR Ment Health 2018;5(1):e4
URL: http://mental.jmir.org/2018/1/e4/
doi:10.2196/mental.9092
PMID:29326098

Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis

Andrea Cipriani, Toshi A Furukawa, Georgia Salanti*, Anna Chaimani, Lauren Z Atkinson, Yusuke Ogawa, Stefan Leucht, Henricus G Ruhe, Erick H Turner, Julian P T Higgins, Matthias Egger, Nozomi Takeshima, Yu Hayasaka, Hissei Imai, Kiyomi Shinohara, Aran Tajika, John P A Ioannidis, John R Geddes*

Summary

Background Major depressive disorder is one of the most common, burdensome, and costly psychiatric disorders worldwide in adults. Pharmacological and non-pharmacological treatments are available; however, because of inadequate resources, antidepressants are used more frequently than psychological interventions. Prescription of these agents should be informed by the best available evidence. Therefore, we aimed to update and expand our previous work to compare and rank antidepressants for the acute treatment of adults with unipolar major depressive disorder.

Methods We did a systematic review and network meta-analysis. We searched Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Process, PsycINFO, the websites of regulatory agencies, and international registers for published and unpublished, double-blind, randomised controlled trials from their inception to Jan 8, 2016. We included placebo-controlled and head-to-head trials of 21 antidepressants used for the acute treatment of adults (≥18 years old and of both sexes) with major depressive disorder diagnosed according to standard operationalised criteria. We excluded quasi-randomised trials and trials that were incomplete or included 20% or more of participants with bipolar disorder, psychotic depression, or treatment-resistant depression; or patients with a serious concomitant medical illness. We extracted data following a predefined hierarchy. In network meta-analysis, we used group-level data. We assessed the studies' risk of bias in accordance to the Cochrane Handbook for Systematic Reviews of Interventions, and certainty of evidence using the Grading of Recommendations Assessment, Development and Evaluation framework. Primary outcomes were efficacy (response rate) and acceptability (treatment discontinuations due to any cause). We estimated summary odds ratios (ORs) using pairwise and network meta-analysis with random effects. This study is registered with PROSPERO, number CRD42012002291.

Findings We identified 28 552 citations and of these included 522 trials comprising 116 477 participants. In terms of efficacy, all antidepressants were more effective than placebo, with ORs ranging between 2·13 (95% credible interval [CrI] 1·89–2·41) for amitriptyline and 1·37 (1·16–1·63) for reboxetine. For acceptability, only agomelatine (OR 0·84, 95% CrI 0·72–0·97) and fluoxetine (0·88, 0·80–0·96) were associated with fewer dropouts than placebo, whereas clomipramine was worse than placebo (1·30, 1·01–1·68). When all trials were considered, differences in ORs between antidepressants ranged from 1·15 to 1·55 for efficacy and from 0·64 to 0·83 for acceptability, with wide CrIs on most of the comparative analyses. In head-to-head studies, agomelatine, amitriptyline, escitalopram, mirtazapine, paroxetine, venlafaxine, and vortioxetine were more effective than other antidepressants (range of ORs 1·19–1·96), whereas fluoxetine, fluvoxamine, reboxetine, and trazodone were the least efficacious drugs (0·51–0·84). For acceptability, agomelatine, citalopram, escitalopram, fluoxetine, sertraline, and vortioxetine were more tolerable than other antidepressants (range of ORs 0·43–0·77), whereas amitriptyline, clomipramine, duloxetine, fluvoxamine, reboxetine, trazodone, and venlafaxine had the highest dropout rates (1·30–2·32). 46 (9%) of 522 trials were rated as high risk of bias, 380 (73%) trials as moderate, and 96 (18%) as low; and the certainty of evidence was moderate to very low.

Interpretation All antidepressants were more efficacious than placebo in adults with major depressive disorder. Smaller differences between active drugs were found when placebo-controlled trials were included in the analysis, whereas there was more variability in efficacy and acceptability in head-to-head trials. These results should serve evidence-based practice and inform patients, physicians, guideline developers, and policy makers on the relative merits of the different antidepressants.

Funding National Institute for Health Research Oxford Health Biomedical Research Centre and the Japan Society for the Promotion of Science.

Published Online
February 21, 2018
http://dx.doi.org/10.1016/S0140-6736(17)32802-7

See Online/Comment
http://dx.doi.org/10.1016/S0140-6736(18)30421-5

*Joint first authors

Department of Psychiatry, University of Oxford, Oxford, UK (A Cipriani MD, L Z Atkinson MSc, H G Ruhe PhD, Prof J R Geddes MD); **Oxford Health NHS Foundation Trust** (A Cipriani, Prof J R Geddes) **and Oxford Centre for Human Brain Activity, Department of Psychiatry** (L Z Atkinson), **Warneford Hospital, Oxford, UK; Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine and School of Public Health, Kyoto, Japan** (Prof T A Furukawa MD, Y Ogawa MD, N Takeshima MD, Y Hayasaka MD, H Imai MD, K Shinohara MD, A Tajika MD); **Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland** (G Salanti PhD, Prof M Egger PhD); **School of Medicine, Paris Descartes University, Paris, France** (A Chaimani PhD); **INSERM, UMR1153 Epidemiology and Statistics, Sorbonne Paris Cité Research Center, METHODS Team, Paris, France** (A Chaimani); **Cochrane France, Paris, France** (A Chaimani); **Department of Psychiatry and Psychotherapy, Technische Universität München, Munich, Germany** (Prof S Leucht MD); **Department of Psychiatry, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands** (H G Ruhe); **Behavioral Health and Neurosciences Division, VA Portland Health Care System, Portland, OR, USA** (E H Turner MD); **Department of**

Psychiatry and Department Pharmacology, Oregon Health & Science University, Portland, OR, USA (E H Turner)**; School of Social and Community Medicine, University of Bristol, Bristol, UK** (Prof J P T Higgins PhD)**; Department of Medicine, Department of Health Research and Policy, Department of Biomedical Data Science, and Department of Statistics, Stanford University, Stanford, CA, USA** (Prof J P A Ioannidis MD)**; and Meta-Research Innovation Center at Stanford, Stanford University, Stanford, CA, USA** (J P A Ioannidis)

Correspondence to:
Dr Andrea Cipriani, Department of Psychiatry, University of Oxford, Oxford, OX3 7JX, UK
andrea.cipriani@psych.ox.ac.uk

Research in context

Evidence before this study

Antidepressants are routinely used worldwide for the treatment of major depressive disorder, which is one of the most important global health challenges; however, in the scientific literature, there remains considerable debate about both their effectiveness as a group, and the potential differences in effectiveness and tolerability between individual drugs. With the marketing of new antidepressants and increasing numbers of trials published every year, an updated systematic review and network meta-analysis was required to synthesise the evidence in this important clinical area.

Added value of this study

This network meta-analysis represents a major update and extension of our previous study, which addressed 12 antidepressants with data for head-to-head comparisons only, and provides the best currently available evidence base to guide the choice about pharmacological treatment for adults with acute major depressive disorder. We now include a more comprehensive list of 21 antidepressants and placebo, consider three new clinical outcome measures and many potential effect modifiers, and use the most advanced statistical methodology for network meta-analysis to date.

Implications of all the available evidence

Our findings should inform clinical guidelines and assist the shared decision making process between patients, carers, and clinicians in routine practice on selecting the most appropriate antidepressant for adults with acute major depressive disorder. Future research should seek to extend network meta-analysis to combine aggregate and individual-patient data from trials in a so-called individual-patient data network meta-analysis. This analysis will allow the prediction of personalised clinical outcomes, such as early response or specific side-effects, and the estimate of comparative efficacy at multiple timepoints.

Introduction

Psychiatric disorders account for 22·8% of the global burden of diseases.[1] The leading cause of this disability is depression, which has substantially increased since 1990, largely driven by population growth and ageing.[2] With an estimated 350 million people affected globally, the economic burden of depressive disorders in the USA alone has been estimated to be more than US$210 billion, with approximately 45% attributable to direct costs, 5% to suicide-related costs, and 50% to workplace costs.[3] This trend poses a substantial challenge for health systems in both developed and developing countries, with the need to treat patients, optimise resources, and improve overall health care in mental health.

Grouped into various classes of drugs with slightly different mechanisms of action, antidepressants are widely used treatments for major depressive disorder, which are available worldwide. However, there is a long-lasting debate and concern about their efficacy and effectiveness, because short-term benefits are, on average, modest; and because long-term balance of benefits and harms is often understudied.[4] Therefore, innovation in psychopharmacology is of crucial importance, but the identification of new molecular targets is difficult, primarily because of the paucity of knowledge about how antidepressants work.[5] In routine practice, clinicians have a wide choice of individual drugs and they need good evidence to make the best choice for each individual patient. Network meta-analyses of existing datasets make it possible to estimate comparative efficacy, summarise and interpret the wider picture of the evidence base, and to understand the relative merits of the multiple interventions.[6] Therefore, in this study, we aimed to do a systematic review and network meta-analysis to inform clinical practice by comparing different antidepressants for the acute treatment of adults with unipolar major depressive disorder.

Methods

Search strategy and selection criteria

We did a systematic review and network meta-analysis. We searched the Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Process, PsycINFO, AMED, the UK National Research Register, and PSYNDEX from the date of their inception to Jan 8, 2016, with no language restrictions. We used the search terms "depress*" OR "dysthymi*" OR "adjustment disorder*" OR "mood disorder*" OR "affective disorder" OR "affective symptoms" combined with a list of all included antidepressants.

We included double-blind, randomised controlled trials (RCTs) comparing antidepressants with placebo or another active antidepressant as oral monotherapy for the acute treatment of adults (≥18 years old and of both sexes) with a primary diagnosis of major depressive disorder according to standard operationalised diagnostic criteria (Feighner criteria, Research Diagnostic Criteria, DSM-III, DSM-III-R, DSM-IV, DSM-5, and ICD-10). We considered only double-blind trials because we included placebo in the network meta-analysis, and because this study design increases methodological rigour by minimising performance and ascertainment biases.[7] Additionally, we included all second-generation antidepressants approved by the regulatory agencies in the USA, Europe, or Japan: agomelatine, bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, levomilnacipran, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, venlafaxine, vilazodone, and vortioxetine. To inform clinical practice

globally, we selected the two tricyclics (amitriptyline and clomipramine) included in the WHO Model List of Essential Medicines). We also included trazodone and nefazodone, because of their distinct effect and tolerability profiles. Additionally, we included trials that allowed rescue medications so long as they were equally provided among the randomised groups. We included data only for drugs within the therapeutic range (appendix pp 133, 134). Finally, we excluded quasi-randomised trials and trials that were incomplete or included 20% or more of participants with bipolar disorder, psychotic depression, or treatment-resistant depression; or patients with a serious concomitant medical illness.

The electronic database searches were supplemented with manual searches for published, unpublished, and ongoing RCTs in international trial registers, websites of drug approval agencies, and key scientific journals in the field.[8] For example, we searched ClinicalTrials.gov using the search term "major depressive disorder" combined with a list of all included antidepressants. We contacted all the pharmaceutical companies marketing antidepressants and asked for supplemental unpublished information about both premarketing and post-marketing studies, with a specific focus on second-generation antidepressants. We also contacted study authors and drug manufacturers to supplement incomplete reports of the original papers or provide data for unpublished studies.

Six pairs of investigators (ACi, TAF, LZA, SL, HGR, YO, NT, YH, EHT, HI, KS, and AT) independently selected the studies, reviewed the main reports and supplementary materials, extracted the relevant information from the included trials, and assessed the risk of bias. Any discrepancies were resolved by consensus and arbitration by a panel of investigators within the review team (ACi, TAF, LZA, EHT, and JRG).

The full protocol of this network meta-analysis has been published.[8]

Outcomes

Our primary outcomes were efficacy (response rate measured by the total number of patients who had a reduction of ≥50% of the total score on a standardised observer-rating scale for depression) and acceptability (treatment discontinuation measured by the proportion of patients who withdrew for any reason).[8] All-cause discontinuation was used as a measure for the acceptability of treatments, because it encompasses efficacy and tolerability.[9] Secondary outcomes were endpoint depression score, remission rate, and the proportion of patients who dropped out early because of adverse events. When depressive symptoms had been measured with more than one standardised rating scale, we used a predefined hierarchy, based on psychometric properties and consistency of use across included trials.[8] In the absence of information or supplemental data from the authors, response rate was calculated according to a validated imputation method.[10] We recorded the outcomes as close to 8 weeks as possible for all analyses.[9] If information at 8 weeks was not available, we used data ranging between 4 and 12 weeks (we gave preference to the timepoint closest to 8 weeks; if equidistant, we took the longer outcome). We checked trial protocols where available and compared published with unpublished data. We extracted data following a predefined hierarchy described in our protocol and gave priority to unpublished information in case of disagreement.[8]

For the **20th WHO Model List of Essential Medicines** see http://www.who.int/medicines/publications/essentialmedicines/20th_EML2017.pdf?ua=1

See **Online** for appendix

Data analysis

For studies published more than once (ie, duplicates), we included only the report with the most informative and complete data. Full details of the applied statistical approaches are provided in the protocol.[8] We estimated summary odds ratios (ORs) for dichotomous outcomes and standardised mean differences (SMD, Cohen's *d*) for continuous outcomes using pairwise and network meta-analysis. In network meta-analysis, we used group-level data; the binomial likelihood was used for dichotomous outcomes and the normal likelihood for continuous outcomes. The study effect sizes were then synthesised using a random-effects network meta-analysis model. We accounted for the correlations induced by multi-group studies by using multivariate distributions. The variance in the random-effects distribution (heterogeneity variance) was considered to measure the extent of across-study and within-comparison variability on treatment effects. Additionally, in network meta-analysis, we assumed that the amount of heterogeneity was the same for all treatment comparisons. To assess the amount of heterogeneity, we compared the posterior distribution of the estimated heterogeneity variance with its predictive distribution.[11] To rank the treatments for each outcome, we used the surface under the cumulative ranking curve (SUCRA) and the mean ranks.[12] The transitivity assumption underlying network meta-analysis was evaluated by comparing the distribution of clinical and methodological variables that could act as effect modifiers across treatment comparisons.[8] We did a statistical evaluation of consistency (ie, the agreement between direct and indirect evidence) using the design-by-treatment test[13] and by separating direct evidence from indirect evidence.[14]

We assessed the studies' risk of bias in accordance to the Cochrane Handbook for Systematic Reviews of Interventions. Additionally, we assessed the certainty of evidence contributing to network estimates of the main outcomes with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework.[15]

We evaluated whether treatment effects for the two primary outcomes were robust in subgroup analyses and network meta-regression using study year, sponsorship, depressive severity at baseline, dosing schedule, study precision (ie, small study effect), and novelty effect.[16] The appendix (pp 133–36) summarises

For the **Cochrane Handbook for Systematic Reviews of Interventions** see http://handbook-5-1.cochrane.org/

the definition of covariates. The sensitivity of our conclusions was evaluated by analysing the dataset with the following restrictions: studies with reported response rate, studies using accepted doses in all groups, studies with unpublished data, multi-centre studies, and head-to-head studies. We used comparison-adjusted funnel plots to investigate whether results in imprecise trials differ from those in more precise trials.[17]

We fitted all models in OpenBUGS (version 3.2.2)[18] using the binomial likelihood for dichotomous outcomes, uninformative prior distributions for the treatment effects, and a minimally informative prior distribution for the common heterogeneity SD. We assumed uninformative priors—ie, N(0,1000)—for all meta-regression coefficients. Convergence of models was ensured by visual inspection of three chains and after considering the Brooks–Gelman–Rubin diagnostic. The codes of analyses, statistical details of the meta-analysis, and meta-regression models are presented in the appendix (pp 182, 183). Statistical evaluation of inconsistency and production of network graphs and result figures were done using the network and network graphs packages in Stata (version 14.2).[19] Network meta-analyses of the primary outcomes were duplicated using the netmeta 0.9-6 package in R (version 3.4.0).[20] The appendix (p 289) lists the changes to the original protocol. The study was done from March 12, 2012, to June 4, 2016, and data analysis was done from June 5, 2016, to Sept 18, 2017.

This study is registered with PROSPERO, number CRD42012002291.

Data sharing

With the publication of this Article, the full dataset will be freely available online in Mendeley Data, a secure online repository for research data, which allows archiving of any file type and assigns a permanent and unique digital object identifier (DOI) so that the files can be easily referenced (DOI:10.17632/83rthbp8ys.2).

Role of the funding source

The funder of this study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or in the decision to submit for publication. ACi, TAF, GS, ACh, LZA, and YO had full access to all the data, and ACi was responsible for the decision to submit for publication.

Results

28 552 citations were identified by the search and 680 potentially eligible articles were retrieved in full text (figure 1). We included 421 trials from the database search, 86 unpublished studies from trial registries and pharmaceutical company websites, and 15 from personal communication or hand-searching other review articles. Overall, 522 double-blind, parallel, RCTs (comprising 116 477 patients) done between 1979 and 2016, and comparing 21 antidepressants or placebo were included in the analysis (appendix pp 6–64). The appendix (pp 65–114) summarises the characteristics of included studies. The mean study sample size was 224 participants (SD 186). In total, 87 052 participants were randomly assigned to an active drug and 29 425 were randomly assigned to placebo. The mean age was 44 years (SD 9) for both men and women; 38 404 (62·3%) of 61 681 of the sample population were women. The median duration of the acute treatment was 8 weeks (IQR 6–8). 243 (47%) of 522 studies randomly assigned participants to three or more groups, and 304 (58%) of 522 were placebo-controlled trials. 391 (83%) of 472 were multi-centre studies and 335 (77%) of 437 studies recruited outpatients only. 252 (48%) of 522 trials recruited patients from North America, 37 (7%) from Asia, and 140 (27%) from Europe (59 [11%] trials were cross-continental and the remaining 34 [7%] were either from other regions or did not specify). The great majority of patients had moderate-to-severe major depressive disorder, with a mean reported baseline severity score on the Hamilton Depression Rating Scale 17-item of 25·7 (SD 3·97) among 464 (89%) of 522 studies. Response rate was imputed in 20 608 (17·7%) of 116 447 cases. Rescue medications (typically benzodiazepines or other sedative hypnotics) were allowed in 187 (36%) of 522 studies. 409 (78%) of 522 studies were funded by pharmaceutical companies. We retrieved unpublished information for 274 (52%) of the included trials. Consistent with the study protocol, the primary analysis was based on the 474 studies (comprising 106 966 patients) that used drugs within the licensed dose range (ie, the dosage approved by the regulatory agencies in the USA and Europe; appendix pp 133, 134).

Figure 2 shows the network of eligible comparisons for efficacy and acceptability. All antidepressant drugs, except milnacipran, had at least one placebo-controlled trial. Only levomilnacipran was not directly compared with at least another active drug in any of the networks. The appendix (pp 139–44) provides detailed results of pairwise meta-analyses. Figure 3 shows the network meta-analysis' results for the primary outcomes. In terms of efficacy (432 RCTs, comprising 102 443 patients), all antidepressants were more effective than placebo, with ORs ranging between 2·13 (95% credible interval [CrI] 1·89–2·41) for amitriptyline and 1·37 (1·16–1·63) for reboxetine. In terms of acceptability (422 RCTs, comprising 99 787 patients), agomelatine (OR 0·84, 95% CrI 0·72–0·97) and fluoxetine (0·88, 0·80–0·96) were associated with fewer dropouts than placebo; by contrast, clomipramine was worse than placebo (1·30, 1·01–1·68).

The relative efficacy of antidepressants compared with placebo is also shown for remission (appendix pp 152, 153). The random-effects summary SMD for all antidepressants was 0·30 (95% CrI 0·26–0·34; p<0·0001; appendix pp 150, 151). In terms of dropouts due to adverse events, all active drugs were associated with higher withdrawal rates than placebo with ORs ranging

between 1·64 and 4·44, and 95% CrI excluding the null, except agomelatine (OR 1·21, 95% CrI 0·94–1·56; appendix pp 154–55). For the full results of the secondary outcomes see the appendix (pp 150–55).

In the analysis of response rate, 8% of the loops were inconsistent (17 of 219 loops; p value of the design by treatment test was 0·063), and also 8% of the loops were inconsistent for dropouts (16 of 210 loops; p=0·219). The median heterogeneity variances were estimated at 0·044 (95% CrI 0·028–0·063) for response and 0·040 (0·023–0·062) for dropout, suggesting moderate-to-low heterogeneity. Subgroup meta-regression analyses revealed that the use of placebo in trials was the strongest explanation of heterogeneity and inconsistency in those evaluated. Exclusion of placebo-controlled trials resulted in a 24% relative reduction in heterogeneity variance for response and 45% for dropout. Additionally, we found that smaller and older studies presented larger effects of the active interventions versus placebo (in particular for amitriptyline, bupropion, fluoxetine, and reboxetine; appendix pp 182–96). The year of randomisation or study precision did not materially impact on the relative treatment effects between active interventions (appendix p 228). Overall, 46 (9%) of 522 trials were rated as high risk of bias, 380 (73%) trials as moderate, and 96 (18%) as low (appendix pp 115–32).

We also synthesised head-to-head studies separately to assess the differences between drugs. Figure 4 presents these data for the primary outcomes (194 studies with at least two active groups at licensed dose and

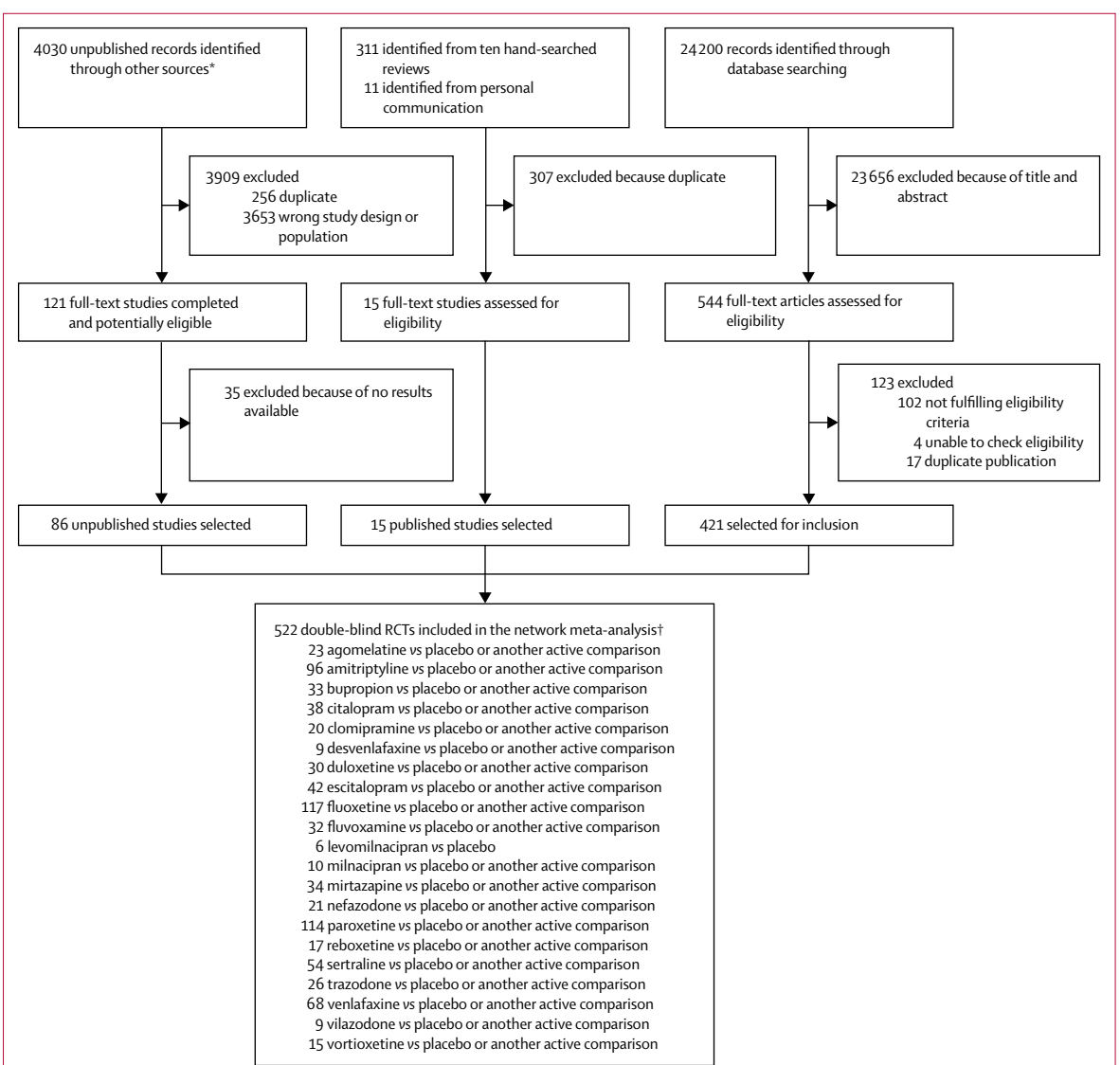


Figure 1: **Study selection process**
RCTs=randomised controlled trials. *Industry websites, contact with authors, and trial registries. The total number of unpublished records is the total number of results for each drug and on each unpublished database source. †522 RCTs corresponded to 814 treatment groups.

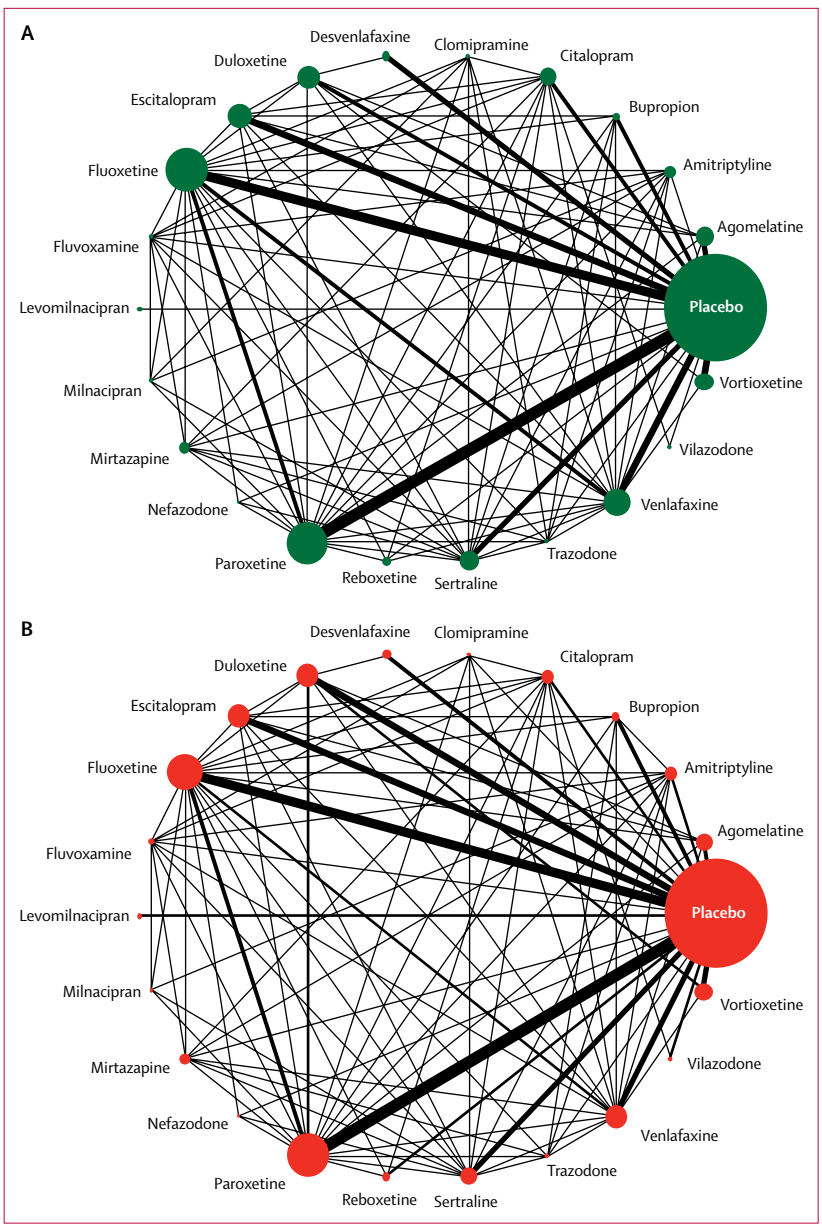


Figure 2: **Network meta-analysis of eligible comparisons for efficacy (A) and acceptability (B)**
Width of the lines is proportional to the number of trials comparing every pair of treatments. Size of every circle is proportional to the number of randomly assigned participants (ie, sample size).

comprised 34 196 patients). Agomelatine, amitriptyline, escitalopram, mirtazapine, paroxetine, venlafaxine, and vortioxetine were more effective than other antidepressants (ORs ranging between 1·19 and 1·96), whereas fluoxetine, fluvoxamine, reboxetine, and trazodone were among the least efficacious drugs (ORs ranging between 0·51 and 0·84). In terms of acceptability, agomelatine, citalopram, escitalopram, fluoxetine, sertraline, and vortioxetine were more tolerable than other antidepressants (ORs ranging between 0·43 and 0·77), whereas amitriptyline, clomipramine, duloxetine, fluvoxamine, reboxetine, trazodone, and venlafaxine were the antidepressants associated with the highest dropout rates (ORs ranging between 1·30 and 2·32). When all trials were considered, differences in ORs between antidepressants ranged from 1·15 to 1·55 for efficacy (appendix p 147) and from 0·64 to 0·83 for acceptability (appendix p 149), with wide CrIs on most of the comparative analyses. Figure 5 reports the two-dimensional graphs about efficacy and acceptability in all studies and head-to-head studies. Results for the secondary outcomes were in line with the findings for the primary outcomes (appendix pp 197–230). Within the head-to-head comparisons, when a treatment was the novel or experimental drug of comparison, it appeared to be significantly more effective than when that same treatment was the older or control drug of comparison (difference 1·18-times, 95% CrI 1·09–1·27). Adjusting for this novelty effect diminished the differences between antidepressants.

We incorporated the GRADE judgments in figure 4. The certainty of evidence for the relative treatment effects of efficacy and acceptability varied; it was moderate for most of the comparisons involving agomelatine, escitalopram, citalopram, and mirtazapine, and low to very low for most comparisons involving vortioxetine, nefazadone, clomipramine, bupropion, and amitriptyline (appendix pp 231–65). The appendix (pp 266–85) presents the ranking of treatments based on cumulative probability plots and SUCRAs.

In accordance with the review protocol, we also did a sensitivity analysis, including all the studies that used the drugs within the accepted doses (ie, doses recommended in some international clinical guidelines; appendix pp 133, 134) and the results did not change substantially (appendix p 187).

Discussion

This study is based on 522 double-blind studies, which included 116 477 patients randomly assigned to 21 individual first-generation and second-generation antidepressant drugs or placebo. The project extends our previous work that had addressed 12 antidepressants with data for head-to-head comparisons.[9] The present analysis is substantially more comprehensive because it includes 21 active treatments and placebo. The much larger evidence base (about 117 000 *vs* 26 000 patients), obtained through exhaustive search for published and unpublished information, allowed us to investigate additional important outcomes, such as remission, change in mood symptoms and dropouts due to side-effects, and a number of methodological issues, such as sponsorship, dosing schedule, study precision, and novelty effect.[16]

We found that all antidepressants included in the meta-analysis were more efficacious than placebo in adults with major depressive disorder and the summary effect sizes were mostly modest. Some antidepressants, such as

escitalopram, mirtazapine, paroxetine, agomelatine, and sertraline had a relatively higher response and lower dropout rate than the other antidepressants. By contrast, reboxetine, trazodone, and fluvoxamine were associated with generally inferior efficacy and acceptability profiles compared with the other antidepressants, making them less favourable options. To make our results as relevant and robust as possible to inform clinical practice, we decided to focus on head-to-head studies and at the same time emphasise the certainty of the retrieved evidence. Our assessment overall found few differences between antidepressants when all data were considered, while there was more diversity in the range of efficacy and dropout patterns seen across the head-to-head comparisons than the meta-analysis of antidepressants versus placebo.

The present findings in adults contrast with the efficacy of antidepressants in children and adolescents, for which fluoxetine is probably the only antidepressant that might reduce depressive symptoms.[21] This differential efficacy across age groups might reflect heterogeneous mechanisms and causes of depression,[22] smaller number of studies in young people, or different methodological issues affecting adult and paediatric trials.[23] The effect sizes were also smaller in more recent and larger placebo-controlled trials than in older and smaller ones, which might be an indicator of bias.

Estimated differences between drugs were smaller in placebo-controlled trials than in head-to-head studies. There are several potential explanations, as many factors have been associated with higher placebo response rates, such as randomisation ratio and the expectation of receiving an active treatment, the therapeutic setting, or the frequency of study visits.[24] In our dataset, we found that response to the same antidepressant was on average smaller and dropouts more likely to occur in placebo controlled trials than in head-to-head studies. Moreover, for the same drug and the same probability of receiving placebo, larger all-cause dropout rates were associated with a lower response to treatment. The use of the last observation carried forward (LOCF) approach for imputing missing outcome data might have affected the estimates of treatment effect.[25] Depressive symptoms tend to spontaneously improve over time and this phenomenon contributes to the high percentage of placebo responders in antidepressant trials.[26] Patients randomly assigned to the active drug in a double-blind, placebo-controlled trial might leave studies earlier than in head-to-head studies because they might suspect they have been allocated to the placebo group than to the intervention group. Antidepressants usually take full effect only after weeks of treatment; therefore, participants who dropped out earlier tend to have poorer responses than those who remain on treatment, which are carried forward to the end of the trial by the LOCF analysis. The final result can be an underestimate of the true efficacy of the active drug.

Another possible explanation could be a bias in conduct, analysis, or reporting of head-to-head trials, driven by

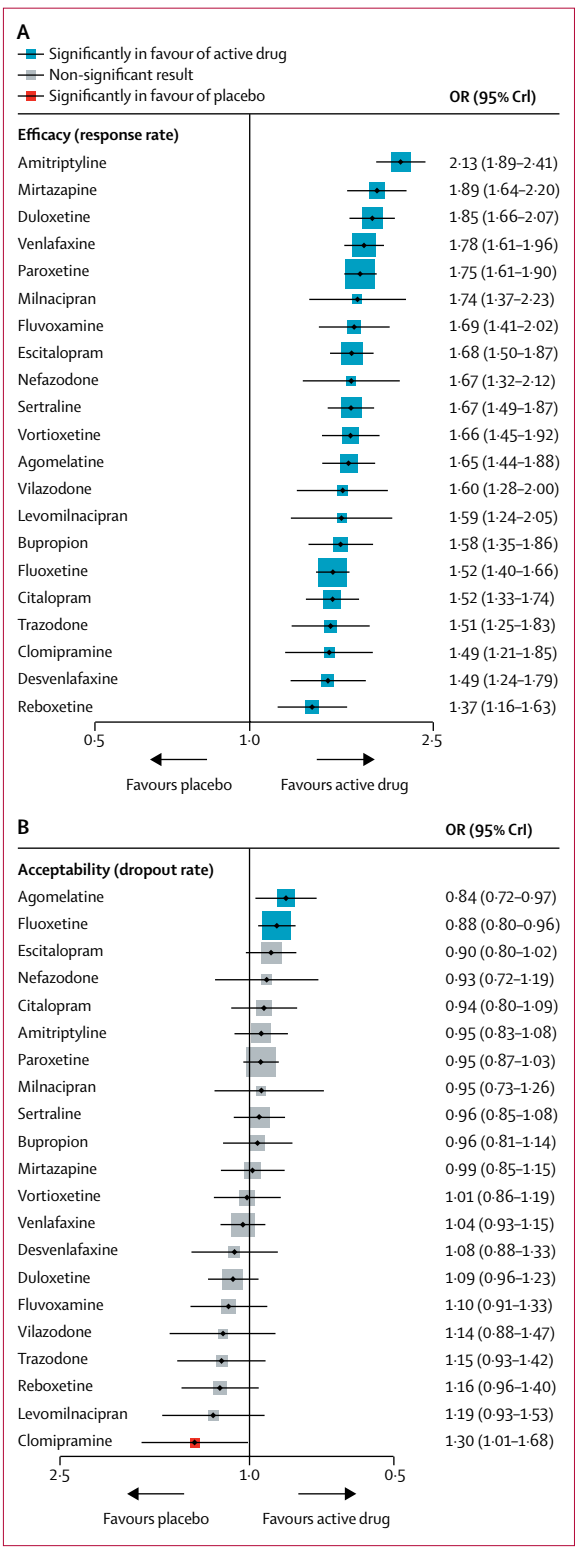


Figure 3: **Forest plots of network meta-analysis of all trials for efficacy (A) and acceptability (B)**
Antidepressants were compared with placebo, which was the reference compound. OR=odds ratio. CrI=credible interval.

Efficacy (response rate) | Comparison | Acceptability (dropout rate)

Agom	**0·72* (0·55–0·92)**	0·80* (0·54–1·15)	0·89* (0·66–1·19)	**0·57* (0·42–0·77)**	**0·62† (0·47–0·82)**	0·97* (0·74–1·27)	0·85† (0·68–1·05)	**0·69† (0·51–0·97)**	0·79* (0·58–1·09)	0·81* (0·61–1·05)	0·70* (0·44–1·14)	0·81* (0·65–1·00)	**0·53* (0·36–0·80)**	0·86* (0·66–1·13)	**0·69* (0·48–0·98)**	**0·74† (0·58–0·92)**	1·24† (0·71–2·19)
0·96* (0·76–1·24)	Amit	1·10‡ (0·78–1·58)	1·23* (0·94–1·64)	0·79† (0·60–1·05)	0·87† (0·66–1·15)	**1·35* (1·05–1·74)**	1·18† (0·99–1·42)	0·97† (0·74–1·24)	1·10† (0·84–1·45)	1·12* (0·89–1·42)	0·98‡ (0·62–1·55)	1·12† (0·95–1·34)	0·74† (0·51–1·10)	1·20* (0·97–1·47)	0·96‡ (0·70–1·31)	1·02† (0·83–1·26)	**1·72† (1·00–3·05)**
0·87† (0·59–1·30)	0·91‡ (0·62–1·31)	Bupr	1·11‡ (0·76–1·67)	0·71† (0·49–1·07)	0·78† (0·53–1·18)	1·23* (0·84–1·80)	1·07‡ (0·76–1·50)	0·87‡ (0·59–1·30)	1·00‡ (0·66–1·49)	1·01† (0·70–1·47)	0·89‡ (0·51–1·54)	1·02‡ (0·73–1·43)	0·67† (0·42–1·08)	1·08‡ (0·75–1·56)	0·87‡ (0·57–1·30)	0·92‡ (0·66–1·30)	1·55† (0·85–2·94)
1·13* (0·88–1·47)	1·18* (0·93–1·49)	1·30† (0·88–1·93)	Cita	**0·64† (0·47–0·87)**	**0·70* (0·51–0·95)**	1·09* (0·85–1·42)	0·96* (0·76–1·21)	0·78* (0·57–1·06)	0·89* (0·64–1·23)	0·91† (0·68–1·21)	0·79‡ (0·49–1·32)	0·91* (0·71–1·17)	**0·60† (0·41–0·87)**	0·97‡ (0·74–1·25)	0·77* (0·53–1·13)	0·83† (0·64–1·07)	1·40† (0·78–2·48)
1·20* (0·91–1·59)	1·24† (0·98–1·58)	1·37† (0·93–2·04)	1·06* (0·82–1·38)	Clom	1·10† (0·80–1·51)	**1·71* (1·27–2·29)**	**1·49† (1·16–1·90)**	1·22† (0·88–1·67)	**1·40† (1·00–1·92)**	**1·41* (1·05–1·91)**	1·24‡ (0·76–2·00)	**1·42† (1·12–1·79)**	0·94‡ (0·62–1·41)	**1·51† (1·15–1·96)**	1·21† (0·83–1·73)	1·29† (0·99–1·67)	**2·20† (1·22–3·90)**
1·06* (0·82–1·37)	1·10† (0·84–1·42)	1·21† (0·81–1·81)	0·93* (0·71–1·22)	0·88† (0·66–1·18)	Dulo	**1·56* (1·19–2·01)**	**1·37* (1·06–1·73)**	1·12* (0·80–1·53)	1·28† (0·91–1·75)	1·30* (0·96–1·72)	1·13‡ (0·69–1·83)	**1·30* (1·02–1·63)**	0·86‡ (0·57–1·29)	**1·38† (1·04–1·80)**	1·10† (0·76–1·59)	1·18‡ (0·92–1·49)	**1·99† (1·13–3·52)**
0·90* (0·71–1·14)	0·93* (0·74–1·17)	1·03† (0·70–1·51)	**0·79* (0·65–0·97)**	**0·75* (0·58–0·97)**	0·85* (0·67–1·08)	Esci	0·87* (0·70–1·09)	**0·71* (0·53–0·96)**	0·81* (0·60–1·11)	0·83* (0·63–1·08)	0·72† (0·45–1·18)	0·83* (0·67–1·03)	**0·55* (0·37–0·81)**	0·88* (0·69–1·12)	0·70* (0·49–1·00)	**0·75* (0·60–0·94)**	1·27‡ (0·73–2·25)
1·20* (0·99–1·48)	**1·25† (1·06–1·48)**	1·38† (0·97–1·97)	1·06* (0·87–1·29)	1·00‡ (0·81–1·24)	1·14* (0·91–1·44)	**1·34* (1·11–1·61)**	Fluo	0·82* (0·64–1·04)	0·94* (0·72–1·20)	0·95* (0·77–1·16)	0·83† (0·54–1·30)	0·95* (0·83–1·09)	**0·63† (0·44–0·90)**	1·01† (0·84–1·21)	0·81* (0·60–1·09)	0·87† (0·74–1·01)	1·46† (0·85–2·53)
1·20* (0·91–1·61)	1·25† (0·99–1·59)	1·38† (0·93–2·07)	1·06* (0·82–1·39)	1·01‡ (0·76–1·32)	1·14† (0·85–1·54)	**1·34* (1·03–1·75)**	1·00* (0·80–1·25)	Fluv	1·14† (0·84–1·56)	1·16* (0·89–1·52)	1·01‡ (0·62–1·71)	1·16* (0·90–1·49)	0·77† (0·51–1·17)	1·23* (0·94–1·63)	0·99‡ (0·69–1·42)	1·06* (0·80–1·38)	**1·78‡ (1·00–3·24)**
1·07* (0·80–1·44)	1·11† (0·86–1·43)	1·23† (0·81–1·85)	0·94† (0·71–1·26)	0·89† (0·67–1·19)	1·01‡ (0·74–1·38)	1·19* (0·90–1·58)	0·89* (0·70–1·13)	0·89† (0·67–1·17)	Miln	1·02† (0·75–1·37)	0·88‡ (0·54–1·44)	1·02‡ (0·80–1·31)	0·67† (0·45–1·03)	1·08* (0·82–1·44)	0·86* (0·60–1·25)	0·93* (0·71–1·22)	1·56† (0·89–2·84)
0·93* (0·72–1·21)	0·97* (0·77–1·21)	1·07† (0·73–1·57)	0·82* (0·65–1·05)	0·78* (0·60–1·01)	0·88* (0·67–1·16)	1·04* (0·82–1·32)	**0·78* (0·64–0·94)**	**0·78* (0·60–0·99)**	0·87* (0·66–1·15)	Mirt	0·87† (0·55–1·41)	1·00* (0·82–1·23)	**0·66* (0·45–0·99)**	1·06* (0·84–1·35)	0·85* (0·62–1·18)	0·91* (0·73–1·13)	1·53† (0·89–2·72)
1·15† (0·76–1·76)	1·19† (0·80–1·78)	1·32‡ (0·80–2·20)	1·01‡ (0·67–1·54)	0·96‡ (0·63–1·45)	1·09‡ (0·71–1·68)	1·28* (0·86–1·94)	0·96‡ (0·66–1·40)	0·95‡ (0·63–1·46)	1·07‡ (0·70–1·67)	1·23* (0·82–1·86)	Nefa	1·15‡ (0·74–1·78)	0·75‡ (0·43–1·32)	1·23† (0·77–1·90)	0·98‡ (0·57–1·64)	1·04‡ (0·66–1·65)	1·76† (0·90–3·56)
1·01* (0·82–1·24)	1·05† (0·89–1·23)	1·16† (0·81–1·64)	0·89* (0·72–1·09)	0·84† (0·68–1·03)	0·95† (0·76–1·19)	1·12* (0·93–1·35)	**0·84* (0·73–0·95)**	0·84* (0·67–1·04)	0·94† (0·75–1·18)	1·08* (0·89–1·30)	0·88‡ (0·60–1·27)	Paro	**0·66† (0·46–0·94)**	1·06* (0·88–1·28)	0·85† (0·63–1·15)	0·91* (0·77–1·07)	1·53† (0·90–2·66)
1·44* (1·02–2·04)	**1·50† (1·07–2·07)**	**1·65† (1·05–2·60)**	1·27† (0·92–1·75)	1·20† (0·84–1·70)	1·36† (0·95–1·95)	**1·60* (1·14–2·23)**	1·20† (0·88–1·62)	1·20† (0·83–1·71)	1·35† (0·92–1·95)	**1·54* (1·09–2·17)**	1·25‡ (0·77–2·01)	**1·43† (1·05–1·94)**	Rebo	**1·61† (1·09–2·34)**	1·29† (0·81–2·01)	1·38† (0·94–1·99)	**2·32† (1·24–4·41)**
1·07* (0·85–1·37)	1·11* (0·92–1·35)	1·23† (0·85–1·79)	0·95† (0·76–1·18)	0·90† (0·71–1·13)	1·02‡ (0·79–1·32)	1·20* (0·97–1·48)	0·89‡ (0·76–1·05)	0·89† (0·70–1·13)	1·00† (0·77–1·30)	1·15* (0·93–1·43)	0·93‡ (0·63–1·37)	1·07* (0·90–1·26)	0·75† (0·54–1·04)	Sert	0·80* (0·58–1·11)	0·86* (0·70–1·05)	1·45† (0·84–2·54)
1·36* (0·99–1·87)	**1·41† (1·06–1·86)**	**1·56† (1·04–2·31)**	1·20* (0·88–1·63)	1·13† (0·83–1·54)	1·28† (0·92–1·79)	**1·51* (1·12–2·04)**	1·13† (0·87–1·46)	1·13† (0·82–1·55)	1·27* (0·91–1·76)	**1·45* (1·09–1·94)**	1·18‡ (0·75–1·84)	**1·35* (1·04–1·75)**	0·94‡ (0·64–1·39)	1·26† (0·95–1·67)	Traz	1·07‡ (0·77–1·47)	1·80† (0·98–3·38)
1·01* (0·82–1·26)	1·05† (0·87–1·27)	1·16† (0·82–1·65)	0·90† (0·72–1·10)	0·85† (0·67–1·06)	0·96† (0·77–1·21)	1·13* (0·93–1·37)	**0·84† (0·73–0·97)**	0·84* (0·66–1·07)	0·95* (0·73–1·23)	1·09* (0·89–1·33)	0·88‡ (0·59–1·30)	1·01† (0·86–1·17)	**0·70† (0·51–0·97)**	0·94* (0·78–1·13)	**0·75† (0·57–0·98)**	Venl	**1·69† (1·01–2·86)**
0·73‡ (0·42–1·26)	0·76‡ (0·44–1·29)	0·83‡ (0·45–1·54)	0·64† (0·37–1·11)	0·61† (0·35–1·05)	0·69† (0·40–1·20)	0·81‡ (0·47–1·39)	0·60† (0·36–1·02)	0·60† (0·34–1·05)	0·68† (0·39–1·20)	0·78‡ (0·45–1·34)	0·63† (0·33–1·19)	0·72† (0·43–1·22)	**0·51† (0·28–0·92)**	0·68† (0·39–1·16)	**0·54† (0·30–0·95)**	0·72† (0·43–1·19)	Vort

Figure 4: **Head-to-head comparisons for efficacy and acceptability of the 21 antidepressants**

Drugs are reported in alphabetical order. Data are ORs (95% CrI) in the column-defining treatment compared with the row-defining treatment. For efficacy, ORs higher than 1 favour the column-defining treatment (ie, the first in alphabetical order). For acceptability, ORs lower than 1 favour the first drug in alphabetical order. To obtain ORs for comparisons in the opposite direction, reciprocals should be taken. Significant results are in bold and underscored. The certainty of the evidence (according to GRADE) was incorporated in this figure (appendix pp 231–65). OR=odds ratio. CrI=credible interval. Agom=agomelatine. Amit=amitriptyline. Bupr=bupropion. Cita=citalopram. Clom=clomipramine. Dulo=duloxetine. Esci=escitalopram. Fluo=fluoxetine. Fluv=fluvoxamine. Miln=milnacipran. Mirt=mirtazapine. Nefa=nefazodone. Paro=paroxetine. Rebo=reboxetine. Sert=sertraline. Traz=trazodone. Venl=venlafaxine. Vort=vortioxetine. *Moderate quality of evidence. †Low quality of evidence. ‡Very low quality of evidence.

commercial interests.[27] In our analyses, funding by industry was not associated with substantial differences in terms of response or dropout rates. However, non-industry funded trials were few and many trials did not report or disclose any funding. We also observed that drugs tended to show a better efficacy profile when they were novel and used as experimental treatments than when they had become old. This novelty effect might arise where a novel agent is perceived to be more effective and better tolerated; alternatively, selective analyses and outcome reporting bias might be more prominent when a treatment is first launched.[16]

Our literature search was as comprehensive as possible, including the largest amount of unpublished data to date, which are associated with less favourable effect sizes for antidepressants.[28] Although it is possible that a certain amount of unpublished data could not be retrieved, our comparison-adjusted funnel plots did not suggest that small studies gave different results from larger studies either among placebo-controlled trials or head-to-head comparison trials (appendix pp 179–81, 225–27). The estimates of treatment effect from our study are in line with previous reviews on the same matter,[28] but they are considerably more precise because of our larger quantity of data and resulting statistical power.

Our review has some limitations. According to the GRADE framework, the quality of many comparisons was assessed as low or very low for amitriptyline, bupropion, and venlafaxine, whereas it was often rated as moderate for agomelatine, escitalopram, and mirtazapine. We incorporated the certainty of evidence in the main results of our analysis to highlight the most robust findings for further use in clinical judgment. However, many trials did not report adequate information about randomisation and allocation concealment, which restricts the interpretation of these results. To increase the methodological rigour of the contributing evidence, we included only double-blind trials, which were generally very similar in design and

conduct. The poor information in terms of risk of bias assessment might be a matter of reporting; however, we presented full details about the risk of bias of all included studies in the appendix (pp 115–32). We did not do a formal cost-effectiveness analysis. All of the most effective antidepressants are now off patent and available in generic form. Some of the antidepressants are included in the WHO Model List of Essential Medicines, which makes them available worldwide and ready to use also in developing countries.

We analysed only average treatment effects and were not able to investigate potentially important clinical and demographical modifiers of treatment response at the individual patient level (eg, age, sex, severity of symptoms, or duration of illness). Patients recruited in randomised trials tend to be highly selected and we also excluded patients with psychotic or treatment-resistant depression, which might limit the applicability of the results to these clinical subgroups, but it was intended as a methodological strength to assure transitivity in the network. We did not cover important clinical issues that might inform treatment decision making in routine clinical practice (eg, specific adverse events, withdrawal symptoms, or combination with non-pharmacological treatments). Additionally, because of the paucity of information reported in the original studies, we were not able to quantify some outcomes, such as global functioning. It should also be noted that some of the adverse effects of antidepressants occur over a prolonged period, meaning that positive results need to be taken with great caution, because the trials in this network meta-analysis were of short duration. The current report summarises evidence of differences between antidepressants when prescribed as an initial treatment. Given the modest effect sizes, non-response to antidepressants will occur. Our information unfortunately cannot guide next-step choices after failure of such a first step (ie, they do not apply to treatment-resistant depression), for which well performed trials are scarce.[29]

Using the data made available on the websites of the US Food and Drug Administration and European Medicines Agency, on the international trial registries, and from contacting study authors and pharmaceutical companies, we managed to incorporate in the analysis a considerable amount of unpublished data for some drugs—namely, agomelatine, escitalopram, paroxetine, reboxetine, sertraline, venlafaxine, vilazodone, and vortioxetine—but not for all the antidepressants included in the network meta-analysis. This limitation in the primary trials might affect the validity of the findings for some antidepressants, but the incorporation of both direct and indirect comparisons might have contributed to reduce the potential risk of bias.[30] We did our best to retrieve all unpublished data and contacted study authors for supplemental material, but we are aware that a substantial amount of information is still not available to the public. There are online archives where trials are prospectively registered; however, they collect reliable information only about the most recent studies and we cannot rule out the possibility that some studies are absent or the same study has been counted twice in our analyses. It is not uncommon for the same study to go by different names in different publications, which complicates the process of data synthesis.[31] By making the dataset fully and freely available, we welcome any

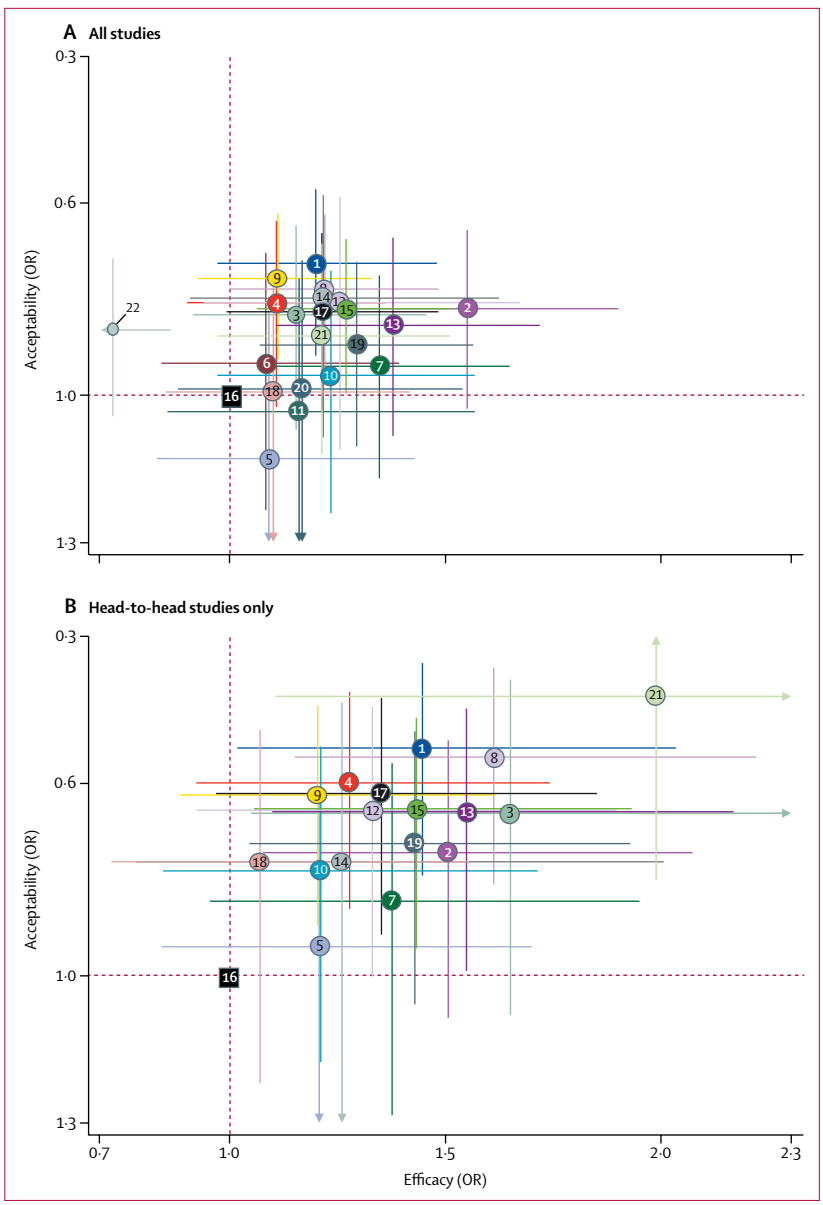


Figure 5: **Two-dimensional graphs about efficacy and acceptability in all studies (A) and head-to-head (B) studies only**
Data are reported as ORs in comparison with reboxetine, which is the reference drug. Error bars are 95% CrIs. Individual drugs are represented by different coloured nodes. Desvenlafaxine, levomilnacipran, and vilazodone were not included in the head-to-head analysis because these three antidepressants had only placebo-controlled trials. ORs=odds ratios. 1=agomelatine. 2=amitriptyline. 3=bupropion. 4=citalopram. 5=clomipramine. 6=desvenlafaxine. 7=duloxetine. 8=escitalopram. 9=fluoxetine. 10=fluvoxamine. 11=levomilnacipran. 12=milnacipran. 13=mirtazapine. 14=nefazodone. 15=paroxetine. 16=reboxetine. 17=sertraline. 18=trazodone. 19=venlafaxine. 20=vilazodone. 21=vortioxetine. 22=placebo.

information that might help clarify any mistakes in our dataset.

Notwithstanding these limitations, the findings from this network meta-analysis represent the most comprehensive currently available evidence base to guide the initial choice about pharmacological treatment for acute major depressive disorder in adults. All statements comparing the merits of one antidepressant with another must be tempered by the potential limitations of the methodology,[32] the complexity of specific patient populations, and the uncertainties that might result from choice of dose or treatment setting. We hope that these results will assist in shared decision making between patients, carers, and their clinicians.

Contributors

ACi, TAF, GS, and JRG conceived and designed the study. ACi, TAF, LZA, SL, HGR, YO, NT, YH, EHT, HI, KS, and AT selected the articles and extracted the data. GS, Ach, JPTH, and ME analysed the data. ACi, TAF, GS, and JRG wrote the first draft of the manuscript. ACh, LZA, YO, SL, HGR, EHT, JPTH, ME, and JPAI interpreted the data and contributed to the writing of the final version of the manuscript. All authors agreed with the results and conclusions of this Article.

Declaration of interests

ACi is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility. TAF has received lecture fees from Eli Lilly, Janssen, Meiji, Mitsubishi-Tanabe, Merck Sharp & Dohme, and Pfizer; consultancy fees from Takeda Science Foundation; and research support from Mochida and Mitsubishi-Tanabe. SL has received honoraria for consulting from LB Pharma, Lundbeck, Otsuka, TEVA, Geodon Richter, Recordati, LTS Lohmann, and Boehringer Ingelheim; and for lectures from Janssen, Lilly, Lundbeck, Otsuka, SanofiAventis, and Servier. NT has received lecture fees from Otsuka and Meiji. YH has received lecture fees from Yoshitomi. JRG is an NIHR Senior Investigator. All other authors declare no competing interests.

Acknowledgments

This study was funded by the National Institute for Health Research (NIHR) Oxford Health Biomedical Research Centre (BRC-1215-20005) and the Japan Society for the Promotion of Science (17K19808). The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR, or the UK Department of Health.

References

1 GBD 2013 DALYs and HALE Collaborators, Murray CJ, Barber RM, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. *Lancet* 2015; **386:** 2145–91.

2 GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016; **388:** 1603–58.

3 WHO. Depression: fact sheet. Geneva: World Health Organisation, 2017. http://www.who.int/mediacentre/factsheets/fs369/en/ (accessed Sept 21, 2017).

4 Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? *Philos Ethics Humanit Med* 2008; **3:** 14.

5 Harmer CJ, Duman RS, Cowen PJ. How do antidepressants work? New perspectives for refining future treatment approaches. *Lancet Psychiatry* 2017; **4:** 409–18.

6 Higgins JP, Welton NJ. Network meta-analysis: a norm for comparative effectiveness? *Lancet* 2015; **386:** 628–30.

7 Hróbjartsson A, Thomsen AS, Emanuelsson F, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *CMAJ* 2013; **185:** E201–11.

8 Furukawa TA, Salanti G, Atkinson LZ, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. *BMJ Open* 2016; **6:** e010919.

9 Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatment meta-analysis. *Lancet* 2009; **373:** 746–58.

10 Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol* 2005; **20:** 49–52.

11 Rhodes KM, Turner RM, Higgins JP. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *J Clin Epidemiol* 2015; **68:** 52–60.

12 Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011; **64:** 163–71.

13 Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* 2012; **3:** 98–110.

14 Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; **29:** 932–44.

15 Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014; **9:** e99682.

16 Salanti G, Dias S, Welton NJ, et al. Evaluating novel agent effects in multiple-treatments meta-regression. *Stat Med* 2010; **29:** 2369–83.

17 Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013; **8:** e76654.

18 Lunn D, Spiegelhalter D, Thomas A, et al. The BUGS project: evolution, critique and future directions. *Stat Med* 2009; **28:** 3049–67.

19 Chaimani A, Salanti G. Visualizing assumptions and results in network meta-analysis: the network graphs package. *Stata J* 2015; **15:** 905–50.

20 Schwarzer G. Network meta-analysis. In: Schwarzer G, Carpenter JR, Rücker G, eds. Meta-analysis with R. Berlin: Springer, 2015: 187–216.

21 Cipriani A, Zhou X, Del Giovane C, et al. Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis. *Lancet* 2016; **388:** 881–90.

22 Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. *Lancet* 2012; **379:** 1056–67.

23 Walkup JT. Antidepressant efficacy for depression in children and adolescents: industry- and NIMH-funded studies. *Am J Psychiatry* 2017; **174:** 430–37.

24 Rutherford BR, Roose SP. A model of placebo response in antidepressant clinical trials. *Am J Psychiatry* 2013; **170:** 723–33.

25 Cook RJ, Zeng L, Yi GY. Marginal analysis of incomplete longitudinal binary data: a cautionary note on LOCF imputation. *Biometrics* 2004; **60:** 820–28.

26 Furukawa TA, Cipriani A, Atkinson LZ, et al. Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. *Lancet Psychiatry* 2016; **3:** 1059–66.

27 Perlis RH, Perlis CS, Wu Y, Hwang C, Joseph M, Nierenberg AA. Industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry. *Am J Psychiatry* 2005; **162:** 1957–60.

28 Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008; **358:** 252–60.

29 Furukawa TA, Akechi T, Shimodera S, et al. Strategic use of new generation antidepressants for depression: SUN(^_^)D study protocol. *Trials* 2011; **12:** 116.

30 Furukawa TA, Miura T, Chaimani A, et al. Using the contribution matrix to evaluate complex study limitations in a network meta-analysis: a case study of bipolar maintenance pharmacotherapy review. *BMC Res Notes* 2016; **9:** 218.

31 Wager E. The need for trial identifiers. *Curr Med Res Opin* 2004; **20:** 203–06.

32 Ioannidis JP. Meta-analyses can be credible and useful: a new standard. *JAMA Psychiatry* 2017; **74:** 311–12.

Cochrane Library

Cochrane Database of Systematic Reviews

Miscellaneous treatments for antipsychotic-induced tardive dyskinesia (Review)

Soares-Weiser K, Rathbone J, Ogawa Y, Shinohara K, Bergman H

Soares-Weiser K, Rathbone J, Ogawa Y, Shinohara K, Bergman H.
Miscellaneous treatments for antipsychotic-induced tardive dyskinesia.
Cochrane Database of Systematic Reviews 2018, Issue 3. Art. No.: CD000208.
DOI: 10.1002/14651858.CD000208.pub2.

www.cochranelibrary.com

WILEY

[Intervention Review]

Miscellaneous treatments for antipsychotic-induced tardive dyskinesia

Karla Soares-Weiser[1], John Rathbone[2], Yusuke Ogawa[3], Kiyomi Shinohara[4], Hanna Bergman[5]

[1]Editorial & Methods Department, Cochrane, London, UK. [2]Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Australia. [3]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan. [4]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto, Japan. [5]Cochrane Response, Cochrane, London, UK

Contact address: Hanna Bergman, Cochrane Response, Cochrane, St Albans House, 57-59 Haymarket, London, SW1Y 4QX, UK. hbergman@cochrane.org, behanna@gmail.com.

Editorial group: Cochrane Schizophrenia Group.
Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 3, 2018.

Citation: Soares-Weiser K, Rathbone J, Ogawa Y, Shinohara K, Bergman H. Miscellaneous treatments for antipsychotic-induced tardive dyskinesia. *Cochrane Database of Systematic Reviews* 2018, Issue 3. Art. No.: CD000208. DOI: 10.1002/14651858.CD000208.pub2.

ABSTRACT

Background

Antipsychotic (neuroleptic) medication is used extensively to treat people with chronic mental illnesses. Its use, however, is associated with adverse effects, including movement disorders such as tardive dyskinesia (TD) - a problem often seen as repetitive involuntary movements around the mouth and face. This review, one in a series examining the treatment of TD, covers miscellaneous treatments not covered elsewhere.

Objectives

To determine whether drugs, hormone-, dietary-, or herb-supplements not covered in other Cochrane reviews on TD treatments, surgical interventions, electroconvulsive therapy, and mind-body therapies were effective and safe for people with antipsychotic-induced TD.

Search methods

We searched the Cochrane Schizophrenia Group's Study-Based Register of Trials including trial registers (16 July 2015 and 26 April 2017), inspected references of all identified studies for further trials and contacted authors of trials for additional information.

Selection criteria

We included reports if they were randomised controlled trials (RCTs) dealing with people with antipsychotic-induced TD and schizophrenia or other chronic mental illnesses who remained on their antipsychotic medication and had been randomly allocated to the interventions listed above versus placebo, no intervention, or any other intervention.

Data collection and analysis

We independently extracted data from these trials and we estimated risk ratios (RR) or mean differences (MD), with 95% confidence intervals (CIs). We assumed that people who left early had no improvement. We assessed risk of bias and created 'Summary of findings' tables using GRADE.

Main results

We included 31 RCTs of 24 interventions with 1278 participants; 22 of these trials were newly included in this 2017 update. Five trials are awaiting classification and seven trials are ongoing. All participants were adults with chronic psychiatric disorders, mostly schizophrenia, and antipsychotic-induced TD. Studies were primarily of short (three to six6 weeks) duration with small samples size (10 to 157 participants), and most (61%) were published more than 20 years ago. The overall risk of bias in these studies was unclear, mainly due to poor reporting of allocation concealment, generation of the sequence, and blinding.

Nineteen of the 31 included studies reported on the primary outcome 'No clinically important improvement in TD symptoms'. Two studies found moderate-quality evidence of a benefit of the intervention compared with placebo: valbenazine (RR 0.63, 95% CI 0.46 to 0.86, 1 RCT, n = 92) and extract of *Ginkgo biloba* (RR 0.88, 95% CI 0.81 to 0.96, 1 RCT, n = 157), respectively. However, due to small sample sizes we cannot be certain of these effects.

We consider the results for the remaining interventions to be inconclusive: Low- to very low-quality evidence of a benefit was found for buspirone (RR 0.53, 95% CI 0.33 to 0.84, 1 RCT, n = 42), dihydrogenated ergot alkaloids (RR 0.45, 95% CI 0.21 to 0.97, 1 RCT, n = 28), hypnosis or relaxation, (RR 0.45, 95% CI 0.21 to 0.94, 1 study, n = 15), pemoline (RR 0.48, 95% CI 0.29 to 0.77, 1 RCT, n = 46), promethazine (RR 0.24, 95% CI 0.11 to 0.55, 1 RCT, n = 34), insulin (RR 0.52, 95% CI 0.29 to 0.96, 1 RCT, n = 20), branched chain amino acids (RR 0.79, 95% CI 0.63 to 1.00, 1 RCT, n = 52), and isocarboxazid (RR 0.24, 95% CI 0.08 to 0.71, 1 RCT, n = 20). There was low- to very low-certainty evidence of no difference between intervention and placebo or no treatment for the following interventions: melatonin (RR 0.89, 95% CI 0.71 to 1.12, 2 RCTs, n = 32), lithium (RR 1.59, 95% CI 0.79 to 3.23, 1 RCT, n = 11), ritanserin (RR 1.00, 95% CI 0.70 to 1.43, 1 RCT, n = 10), selegiline (RR 1.37, 95% CI 0.96 to 1.94, 1 RCT, n = 33), oestrogen (RR 1.18, 95% CI 0.76 to 1.83, 1 RCT, n = 12), and gamma-linolenic acid (RR 1.00, 95% CI 0.69 to 1.45, 1 RCT, n = 16).

None of the included studies reported on the other primary outcome, 'no clinically significant extrapyramidal adverse effects'.

Authors' conclusions

This review has found that the use of valbenazine or extract of *Ginkgo biloba* may be effective in relieving the symptoms of tardive dyskinesia. However, since only one RCT has investigated each one of these compounds, we are awaiting results from ongoing trials to confirm these results. Results for the remaining interventions covered in this review must be considered inconclusive and these compounds probably should only be used within the context of a well-designed evaluative study.

PLAIN LANGUAGE SUMMARY

Miscellaneous treatments for antipsychotic-induced tardive dyskinesia

What is the aim of this review?

The aim of this Cochrane Review was to find out if drugs, supplements, surgical interventions, electroconvulsive therapy, or mind-body therapies not covered in other Cochrane reviews of tardive dyskinesia can improve tardive dyskinesia. We collected and analysed all relevant randomised controlled trials to answer this question.

Key messages

The drug valbenazine and extract of the herb *Ginkgo biloba* probably improves symptoms of tardive dyskinesia. But we still need more high-quality studies to confirm these findings that were taken from only one study per intervention.

What was studied in the review?

Antipsychotic drugs are used to treat chronic mental illnesses such as schizophrenia by controlling, for instance, abnormal perceptions (hallucinations), disordered thinking and fixed false beliefs (delusions). Tardive dyskinesia is a disfiguring and disabling disorder of abnormal, repetitive and involuntary movements, and it is often caused by antipsychotic drugs. More than 20% of people who rely on antipsychotic drugs to control their mental illness have developed tardive dyskinesia. Many different interventions have been studied for easing the symptoms of tardive dyskinesia. Several Cochrane reviews have summarised the effects of the many treatments used to manage these involuntary movements. This review focusses on 'miscellaneous', a group of other non-connected, interventions not covered in the other Cochrane reviews on tardive dyskinesia.

What are the main results of the review?

We found 31 studies that reported on 24 different interventions to improve tardive dyskinesia in 1278 people who take antipsychotic medication for their chronic mental illnesses. Unfortunately most studies followed up on participants for a short time (most were three to six weeks) and included few participants (the average number of participants was 41 per study).

• Valbenazine probably reduces symptoms of tardive dyskinesia to a clinically important extent compared with placebo (moderate-certainty evidence). However, this evidence is based on only one study in the USA with 92 participants; we are awaiting results from recently completed and ongoing trials to confirm these results.

• Extract of *Ginkgo biloba* probably reduces symptoms of tardive dyskinesia to a clinically important extent compared with placebo (moderate-certainty evidence). However, this evidence is based on only one study in China with 157 participants; we are awaiting results from recently completed and ongoing trials to confirm these results.

• Evidence for the remaining interventions was of low- to very low-certainty evidence and we consider the results for these other interventions to be inconclusive.

How up-to-date is this review?

We searched for studies that had been published up to 26 April 2017.

Sleep Medicine 32 (2017) 246–256

Contents lists available at ScienceDirect

Sleep Medicine

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

Original Article

Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression

Osamu Itani [a, 1], Maki Jike [b, 1], Norio Watanabe [c, *], Yoshitaka Kaneita [a]

[a] Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Oita, Japan
[b] Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, Tokyo, Japan
[c] Department of Health Promotion and Human Behavior, Department of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan

ARTICLE INFO

Article history:
Received 21 April 2016
Received in revised form
1 August 2016
Accepted 15 August 2016
Available online 26 August 2016

Keywords:
Diabetes mellitus
Hypertension
Mortality
Meta-regression
Sleep deprivation
Vascular disease

ABSTRACT

Objective: The dose–response of short sleep duration in mortality has been studied, in addition to the incidences of notable health complications and diseases such as diabetes mellitus, hypertension, cardiovascular diseases, stroke, coronary heart diseases, obesity, depression, and dyslipidemia.
Methods: We collected data from prospective cohort studies with follow-ups of one year or more on associations between short sleep duration and the outcomes. For the independent variable, we divided participants at baseline into short sleepers and normal sleepers. The primary outcomes were defined as mortality and an incident of each health outcome in the long-term follow-up. Risk ratios (RRs) for each outcome were calculated through meta-analyses of adjusted data from individual studies. Sub-group and meta-regression analyses were performed to investigate the association between each outcome and the duration of short sleep.
Results: Data from a cumulative total of 5,172,710 participants were collected from 153 studies. Short sleep was significantly associated with the mortality outcome (RR, 1.12; 95% CI, 1.08–1.16). Similar significant results were observed in diabetes mellitus (1.37, 1.22–1.53), hypertension (1.17, 1.09–1.26), cardiovascular diseases (1.16, 1.10–1.23), coronary heart diseases (1.26, 1.15–1.38), and obesity (1.38, 1.25–1.53). There was no sufficient usable evidence for meta-analyses in depression and dyslipidemia. Meta-regression analyses found a linear association between a statistically significant increase in mortality and sleep duration at less than six hours. No dose–response was identified in the other outcomes.
Conclusions: Based on our findings, future studies should examine the effectiveness of psychosocial interventions to improve sleep on reducing these health outcomes in general community settings.

1. Introduction

Short sleepers are prevalent throughout the world. In the U.S., the age-adjusted mean sleep duration was 7.18 hours and the prevalence of sleepers reporting less than six hours of sleep was 29.2% in 2012 [1]. In several developed countries, the prevalence rate is not higher than in the U.S., but 11.3% in Canada and 9.8% in the U.K. [2].

Several systematic reviews have shown that short sleep duration is associated with important health outcomes including not only mortality [3–6] but also hypertension [5,7], cardiovascular diseases [8], stroke [9], diabetes mellitus [10,11], and obesity [12]. These have been regarded as phenotypes of metabolic abnormalities [13] or arteriosclerosis promotion [14,15] associated with short sleep duration. However, because these reviews investigated associations between short sleep duration and these health outcomes utilized various methodologies in conducting reviews, another systematic review may be needed where the same methodology is used across health outcomes. In this review, associations between short sleep and incidents of some important health outcomes, including dyslipidemia and depression, which have not yet been examined in previous reviews, should also be investigated. Although sleep duration less than six hours is reported to be

Abbreviations: CER, control event rate; CI, confidence interval; HR, hazard ratio; MOOSE, meta-analysis of observational studies in epidemiology; NOS, Newcastle–Ottawa scale; OR, odds ratio; PRISMA, preferred reporting items for systematic reviews and meta-analyses; RR, risk ratio.

Previous presentation: Poster presented at the 73rd Annual Meeting of Japanese Society of Public Health, Tochigi, Japan, November 5–7, 2014.

* Corresponding author. Department of Health Promotion and Human Behavior, Department of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. Fax: +81-75-753-4641.
E-mail address: noriowncu@gmail.com (N. Watanabe).

[1] OI and MI contributed equally to this work as the first authors.

http://dx.doi.org/10.1016/j.sleep.2016.08.006
1389-9457/

associated with higher risk than that of seven to eight hours (especially in terms of mortality outcome in previous cohort studies [16–19]), to the best of our knowledge, this has not been systematically examined in meta-analyses and meta-regression, which can contribute to publication bias and to let researchers speculate on mediator effects of sleep duration on health outcomes.

We therefore conducted a systematic review, meta-analyses, and meta-regression to examine if short sleep duration is associated with a higher prevalence of health outcomes using the same.

2. Methods

We performed the study in accordance with the PRISMA (preferred reporting items for systematic reviews and meta-analyses) [20] and the MOOSE (meta-analysis of observational studies in epidemiology) [21] guidelines, with these checklists (see Appendices S5 and S6 in the Supplementary material).

Two independent researchers (OI and MJ) separately assessed the eligibility, extracted data, and checked the quality of the included studies. Any disagreements were resolved through discussion between these two, and with a third reviewer (NW) if disagreements persisted.

2.1. Data sources and searches

The studies were initially identified on October 17, 2013, through a search of PubMed, PsycINFO, CINAHL, and Embase using pre-specified search terms (Appendices S1–S4). Major medical journals, conference proceedings, and reference lists of included studies and previous systematic reviews were also hand-searched for published, unpublished, and ongoing studies. To identify new studies published during the review process, we conducted a search of PubMed using the same search strategy on October 9, 2014 and on May 6, 2016.

2.2. Study selection

We included studies with a prospective cohort or randomized controlled trial design, conducted in community settings, which compared short with normal sleepers for mortality and incidence of health outcomes in a long-term follow-up. We limited studies to those with a minimum follow-up duration of one year from baseline, and a minimum of 20 participants. Studies were excluded if most participants were aged 20 years or younger at baseline, or if participants had been diagnosed with the health outcome at baseline. We also excluded studies that were conducted in inpatient settings and those that involved pharmacological interventions.

The eligibility of each study for inclusion was checked at two stages: (1) looking through the title and abstract and (2) checking the full text.

2.3. Data extraction and quality assessment

2.3.1. Definition of sleep duration

The definition of short sleep was based on the original paper because common sleep duration varies among cultures and ethnicities [22,23]. Durations of short sleep were incorporated into subgroup analyses and meta-regression as mediators (see below). When both a subjective (eg, sleep diary) and objective sleep duration (eg, actigraphy or polysomnography) were reported, we selected the former as the independent variable. Although a self-report survey may be unlikely to capture the actual amount of sleep per night in comparison with actigraphy [24] or polysomnography [25], objective measures may not always be utilized in general community settings and subjective measures might be preferable because of their applicability. When both sleep durations per day (possibly including a daytime nap) and per night were reported, we selected the latter.

The duration of normal sleep was also defined based on the original paper.

2.3.2. Outcome measures

The outcome was defined as mortality and incidence of health outcomes, which were diabetes mellitus, hypertension, dyslipidemia (hypo or hyperlipidemia), cardiovascular diseases (including events in the heart and brain), coronary heart diseases, stroke, obesity, and depression. When a formal diagnosis was not provided, a surrogate outcome (eg, coronary artery calcification instead of diagnosis of coronary artery diseases, a self-report of diabetes mellitus without evidence of formal diagnosis) was included in the primary analyses, but a sensitivity analysis was planned (see below).

2.3.3. Assessment of bias

We employed the Newcastle–Ottawa scale (NOS) [26] to assess the studies' quality. The instrument has three broad categories (patient selection, four criteria, comparability of study groups, one criterion, and assessment of the outcome, three criteria). For the comparability criteria, we allotted two stars according to the depth of statistical adjustment for risk factors in the original studies (eg, one star for age, sex, and race only, two stars for beyond these). Therefore, a study could reach a full mark with nine stars. For the second and third items of the outcome criteria, we defined, a priori, follow-up durations as reasonably long enough, and adequate follow-up of cohorts in terms of the percent lost to follow-up that was allowed for each disorder (ie, three years and 10% for any cause of mortality, two years and 20% for diabetes mellitus, two years and 20% hypertension, two years and 20% for dyslipidemia, three years and 10% for cardiovascular diseases, three years and 10% for coronary heart diseases, two years and 20% for obesity, and two years and 20% for depression, respectively).

Although previous meta-analyses [27,28] deemed quality of a study as high when it had five or more stars on the NOS criteria, we (a priori) set eight or more stars as high in order to focus on very high quality studies.

2.4. Data synthesis and analysis

We analyzed data y and conducted a meta-analysis for each dependent outcome. In the meta-analysis, we calculated risk ratios (RRs) by pooling adjusted RRs between short and normal sleep provided by the original studies with a random effects model. If hazard ratios (HRs) were reported in a study but RRs were not, the HRs were regarded as RRs. Among studies where odds ratios (ORs) were provided but not RRs, we calculated RRs by using the ORs and control event rates (CERs) in normal sleepers reported in the original studies. Regarding studies where both RRs and CERs were not reported, and only ORs were provided, CERs were borrowed from a study whose characteristics were similar. In the primary analyses, regarding studies where RRs were provided for subgroups separately (eg, male and female), data from these subgroups were combined using a fixed-effect meta-analysis.

Statistical heterogeneity between studies was investigated using the I^2 statistic [29], assuming an I^2 of 75% or greater to be an important level of inconsistency, as a previous review employed [30]. To assess publication bias, we used a funnel plot and Egger's test for all primary outcomes [31]. We used the "trim and fill" method to adjust the funnel plot and recalculated the results [32].

Although subgroup analyses should be interpreted with caution [33], we planned, a priori, to perform analyses for several types of baseline characteristics (ie, between 20 and 65 years of age, or aged 65 years or more; male or female).

Sensitivity analyses were planned for the primary outcome by: (1) excluding studies with surrogate outcomes, and limiting studies to those (2) where sleep duration was reported per night, (3) with eight or more stars in the NOS, (4) those following up participants for no less than ten years, and (5) excluding studies where CERs were borrowed from the others to calculate RRs. Moreover, for the purpose of exploring mediator effects of the duration of definition of short sleep, subgroup analyses were performed for clustering included studies according to the duration (eg, less than five hours or six hours) for each outcome. When ten or more studies were included [34] and studies were clustered into three or more levels in each outcome, meta-regression analyses were also performed to examine a linear association between sleep duration and increase in the outcome frequency, using a random-effects model and illustrating, if any, regression line and its 95% prediction intervals.

A P-value of less than 0.05 was chosen to test the null hypothesis despite multiple comparisons in order to avoid type II over type I errors. For all outcomes, 95% confidence intervals (CIs) were calculated. The data were analyzed using the Comprehensive Meta-Analysis Software (Version 3) [35], with double data entry to avoid input errors.

3. Results

3.1. Search results

The initial electronic search yielded 3580 articles, and an additional database search identified 182 studies on October 9, 2014 and 388 on May 6, 2016. In total, 2521 studies remained after removing duplicate articles. A hand-search did not identify any studies that had not been included in the electronic search (Fig. 1).

At the first and second eligibility check stages, two independent researchers identified 277 articles with an inter-rater reliability of

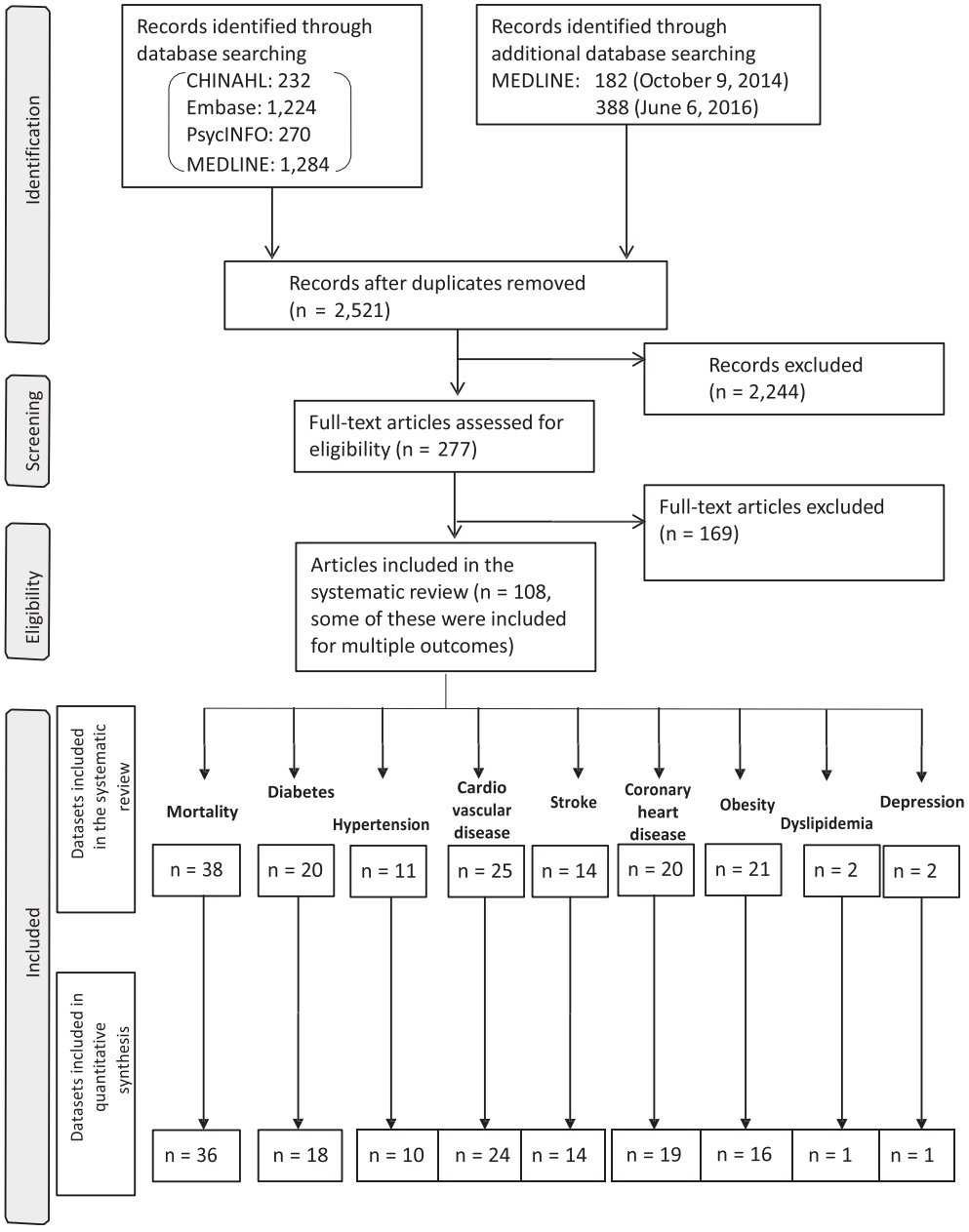


Fig. 1. Flowchart for the included studies.

kappa statistic at 0.79 (95% CI: 0.75–0.82) and 108 articles at 0.64 (0.54–0.73), respectively.

3.2. Characteristics of included studies

All of the 108 included studies were prospective cohort studies. Forty-five studies reported data for two or more health outcomes, and 153 datasets for nine outcomes (N = 5,172,718) were collected from the 108 studies. Most studies were conducted in developed countries (see Table 1 and Tables S1–S8 in the Supplementary material). The number of participants in each dataset ranged from 276 to 392,164, the duration of follow-up was from one to 30 years, and the total NOS scores were from five to nine. The studies' definitions of short sleep varied, but most defined it as less than five or six hours.

Of the included studies for the systematic review, we could not pool data from 14 datasets in the meta-analyses, mainly because no usable data for meta-analyses were provided (Table S9). The number of datasets included in the meta-analyses on each outcome varied from ten (hypertension) to 36 (mortality). Table 1 shows the characteristics of the studies included for the mortality outcome.

3.3. Effect estimates of short sleep compared to normal sleep from meta-analyses

3.3.1. Primary analyses

Compared with normal sleep, short sleep showed a statistically significant increase in mortality due to all causes at a RR of 1.12 (95% CI 1.08–1.16, $P < 0.005$, $I^2 = 25\%$, N of datasets = 36; Figs 2 and 3). Qualitatively similar significant results were obtained for diabetes mellitus, hypertension, cardiovascular disease, coronary heart disease, and obesity, at RR of 1.37 (1.22–1.53, $P < 0.005$, $I^2 = 53\%$, N = 18), 1.17 (1.09–1.26, $P < 0.005$, $I^2 = 48\%$, N = 10), 1.16 (1.10–1.23, $P < 0.005$, $I^2 = 49\%$, N = 24), 1.26 (1.15–1.37, $P < 0.005$, $I^2 = 51\%$, N = 19), and 1.38 (1.25–1.53, $P < 0.005$, $I^2 = 60\%$, N = 16), respectively (Fig. 2). In terms of stroke, short sleep showed no statistically significant increase compared to normal sleep at a RR of 1.08 (0.98–1.19, P = 0.10, N = 14). No important level of heterogeneity between datasets was observed for all outcomes. For the depression and the dyslipidemia outcomes, only one study was included in each and meta-analyses were therefore not performed.

3.3.2. Possible publication bias for primary analyses

In the funnel plots and results from Egger's test (see Figs S1–S3 in the Supplementary material), significant publication bias was observed for coronary heart disease (P = 0.005) and obesity (P = 0.037). However, all the imputed RRs using the trim and fill method still indicated a statistically significant increase in incident of these outcomes among short sleepers.

3.3.3. Subgroup analyses for age groups

Subgroup analyses for those aged 65 years or more and those under 65 at baseline were conducted (Fig. 2). For outcomes where both age groups were analyzed, short sleep was associated with a significant increase in incidents of outcome among those aged under 65 years but did not among those aged over 65 years for hypertension, cardiovascular disease, and obesity, in comparison with normal sleep. Short sleep was not associated with a significant increase in either age group for the mortality outcome; however, short sleep was associated significantly in both age groups for the coronary heart disease outcome. For outcomes where only subgroup analyses were conducted for those aged less than 65 years, short sleep was associated with a significant increase in incidents of diabetes mellitus, but not for stroke.

3.3.4. Subgroup analyses for sex

Subgroup analyses for sex were performed for outcomes other than depression and dyslipidemia. In comparison with normal sleep, short sleep was associated with a significant increase in both sexes in mortality, cardiovascular disease, coronary heart disease, and obesity. Short sleep was associated with a significant increase among males but not females for diabetes mellitus, but vice versa for hypertension and stroke (Fig. 2).

3.3.5. Sensitivity analyses

A majority of sensitivity analyses showed qualitatively similar results to those in the primary analyses (Fig. 2), although the result was not statistically significant in an analysis limiting studies judged to be high quality using the NOS in hypertension outcome (RR = 0.93, 95% CI 0.47–1.84, N = 2) (P = 0.835).

3.3.6. Subgroup analyses and meta-regression for the duration of definition of short sleep

Subgroup analyses for the duration were conducted for the outcomes other than depression and dyslipidemia (Fig. 4). In comparison with normal sleep, short sleep defined as the duration less than six hours was associated with a significant increase in mortality, diabetes, cardiovascular disease, and obesity. In hypertension, stroke, and coronary heart disease outcomes, this trend was not observed probably due to a small number of studies in a cluster.

Meta-regression analyses were performed for mortality, diabetes, hypertension, cardiovascular, and obesity outcomes. In mortality outcome, shorter duration of definition of short sleep was significantly associated with increase in the outcome (coefficient = −0.056, SE = 0.021, P = 0.008, R^2 analog = 0.84, Fig. 5), and the 95% CIs of prediction curves were above zero at the definition of sleep less than six hours. In obesity outcomes, however, shorter sleep duration was associated with decrease in the outcome (coefficient = 0.181, SE = 0.089, P = 0.042, R^2 analog = 0.33, Fig. S4), and the 95% CIs of prediction curves were above zero at the definition of sleep more than five hours. No significant association was observed in diabetes (P = 0.237), hypertension (P = 0.655), or cardiovascular disease (P = 0.819) outcomes.

4. Discussion

To the best of our knowledge, this is the first systematic review where the association between short sleep duration and multiple important health outcomes was investigated with the same methodology used in meta-analyses, and where dose–response of short sleep duration on these outcomes was explored in meta-regression analyses. The present review revealed that short sleepers were likely to be associated with greater mortality than normal sleepers with a RR of around 1.12, in other words around a 12% absolute increase. For the other outcomes, short sleepers were likely to have a point estimate of an absolute increase of 37% for diabetes mellitus, 17% for hypertension, 16% for cardiovascular disease, 26% for coronary heart disease and 38% for obesity. In terms of depression and dyslipidemia, no sufficient evidence from meta-analyses existed to conclude whether short sleep was associated with an increase in the incidents in the meta-analyses.

The present results are similar to those of previous systematic reviews, such as a RR of 1.12 (N of studies = 15) for the mortality outcome [4] and of 1.23 (N = 6) for the hypertension outcome [5]. We believe that the findings of our review add to those in these previous reviews because of our updated comprehensive search and rigorous methodology. However, further epidemiological studies are needed to investigate the association of short sleep with dyslipidemia and depression.

Table 1
Characteristics of studies for the mortality outcome.

Author, Year	Sample size	Male %	Mean age ± SD (range) in years at baseline	Years of follow up Mean ± SD	Definition of Short sleep duration (h)	Definition of normal sleep duration (h)	Newcastle–Ottawa Scale: Selection/Comparability/Outcome
Tsubono et al. 1993 [36]	4,318	39.8	61.4 (40+)	4	≤6/night	7–8/night	★★☆★/★★/★★★
Kojima et al. 2000 [37]	5,322	45.8	Male: 46.9 (20–67) Female: 47.7 (20–67)	11.9	−6.9/night	7.0–8.9/night	☆★☆★/★★/★★★
Seki et al. 2001 [38]	1,065	41.3	65.3 ± 3.6 (60–74)	7.5	<6/day	7/day	★★☆★/★★/★★★
Heslop et al. 2002 [39]	Base line: 7,028 2nd screening: 3,030	Baseline: 85.7 2nd screening: 85.4	Male: (−65) Female: (−60)	25	<7/day	7–8/day	★★☆★/★★/★★★
Mallon et al. 2002 [40]	1,870	48.4	56 (45–65)	12	<6/night	6–8/night	★★☆★/★☆/★★★
Goto et al. 2003 [41]	724	34.7	Male: median 73 (65–97) Female: median 74 (65–97)	12	<6/day	6–7/day	★★★★/★★/★★★
Amagai et al. 2004 [18]	11,325	39.0	55.1 (19–93)	8.2 ± 1.5	≤5.9/night	7.0–7.9/night	★★★★/★★/★★★
Patel et al. 2004 [42]	82,969	0.0	53.4 (40–65)	14	≤5/day	7/day	☆★☆★/★★/★★★
Tamakoshi et al. 2004 [19]	104,010	42.2	56.6 (40–79)	9.9	≤4/day	7/day	★★☆☆/★★/★★★
Ferrie et al. 2007 [43]	Phase1: 9,781 Phase3: 7,729		(35–55)	Phase1: 17.1 Phase3: 11.8	≤5/night	7/night	☆★☆★/★★/★★★
Lan et al. 2007 [44]	3,079	56.8	Male: 71.3 (64+) Female: 71.9 (64+)	8.4 ± 3.3	<7/night	7–7.9/night	★★★★/★★/★★★
Gangwisch et al. 2008 [45]	9,789	37.2	Male: 45.0 (32–59) Female: 73.0 (60–86)	−8, −10	≤5/night	7/night	★★★★/★★/★★★
Ikehara et al. 2009 [46]	98,634	42.1	Male: 58.8 (40–79) Female: 60.2 (40–79)	Median 14.3	≤4/day	7/day	★★☆★/★★/★★★
Mallon et al. 2009 [47]	3,523	49.7	40 ± 10 (30–65)	20	<6/night	≥6/night	★★☆★/★★/★★★
Stone et al. 2009 [48]	8,101	0.0	77.0 (69+)	6.9	<6/night	6–8/night	★★★★/★★/★★★
Suzuki et al. 2009 [49]	12,601	51.1	74.1 ± 5.4 (65–85)	5.3	≤5/day	7/day	☆★☆★/★★/★★★
Chien et al. 2010 [50]	3,430	47.3	(35+)	15.9 (13.1–16.9)	≤5/day	7/day	★★☆★/★★/★★★
Mesas et al. 2010 [51]	3,820	43.8	71.8 ± 7.9 (60+)	6.8	≤5/day	7/day	★★☆★/★★/★★★
Vgontzas et al. 2010 [52]	1,741	42.6	Male: 50.1 ± 14.5 (20+) Female: 47.4 ± 12.6 (20+)	Male: 14 Female: 10	<6/night without insomnia, <6/night with insomnia	≥6/night without insomnia	★★★★/★★/★★★
Castro-Costa et al. 2011 [53]	1,512	38.3	68.9 ± 7.1 (63–75)	7.5 Median: 8.9	<6/night	7 to <8/night	★★★★/★★/★★★
Kripke et al. 2011 [54]	434	0.0	67.6 ± 7.9 (50–81)	10.5	<5/day	6–6.5/day	★★★★/☆☆/★★★
Kronholm et al. 2011 [55]	23,290	48.8	(25–64)	29–34	≤5/night	7–8/night	★★☆★/★★/★★★
Cohen-Mansfield et al. 2012 [56]	1,166	55.5	83.4 ± 5.3 (75–94)	20	<7/night	7–9/night	★★☆★/★★/★★★
Ensrud et al. 2012 [57]	2,505	100.0	75.7 ± 5.2 (67+)	3.4 ± 0.5	≤5/day	>5/day	☆★★★/★★/★★★
Chen et al. 2013 [58]	4,064	55.8	73.8 ± 5.7 (65+)	9	≤4/night	7/night	★★★★/★★/★★★
Garade et al. 2013 [59]	4,941	100.0	(40–59)	30	<6/day	6–7/day	☆★★★/★★/★★★
Hale et al. 2013 [60]	3,942	0.0	62.1 (50–79)	11–16	≤5/night	7–8/night	★★☆★/★★/★★★
Kakizaki et al. 2013 [61]	49,256	48.2	(40–79)	10.8	≤6/day	7/day	★★☆★/★★/★★★
Kim et al. 2013 [62]	135,685	45.6	(45–75)	12.9	≤5/day	7/day	★★☆★/★★/★★★
Li et al. 2013 [63]	9,455	(38.1)	(20–79)	7	≤5/night	7/night	★★☆★/★★/★★★
Magee et al. 2013 [64]	227,815	46.3	(45+)	2.8	<6/day	7/day	★★☆★/★★/★☆★
Yeo et al. 2013 [65]	13,164	41.4	(20+)	9.44	≤5/day	7/day	★★★★/★★/★★★
Bellavia et al. 2014 [66]	70,973	53.3	(45–83)	15	<6/day	6.6–7.4/day	★★☆★/★★/★★★
Rod et al. 2014 [67]	9,098	67.2	45 (35–55)	22	≤5/night	7/night	☆★☆★/★★/★★★
Xiao 2014 [68]	239,896	56.2	(51–72)	14	<5/night	7–8/night	★★☆★/★★/★★☆
Zawisza 2015 [69]	2,449	35.0	Male: 72 ± 5.8 Female: 2.5 ± 5.7	22	≤5/day		★★★★/★★/★★★
Zuurbier 2015 [70]	1,734	46.6	62.2 ± 9.3 (45–98)	7.3 ± 1.3	<6/night	6–7.5/night	☆★★★/★★/★★★
Hall 2015 [71]	3,013	48.6	73.6 ± 2.9 (70–79)	8.2 ± 2.3	<6/night	7–8/night	★★★★/★★/★★★
Cai 2015 [72]	113,138	60.1	Male: (40–75) Female: (44–79)	Median: Male: 6.07 Female: 7.12	4–5/day	7/day	★★★★/★★/★★★

Filled stars indicate applicable for the item. The empty stars indicate not applicable for the item.

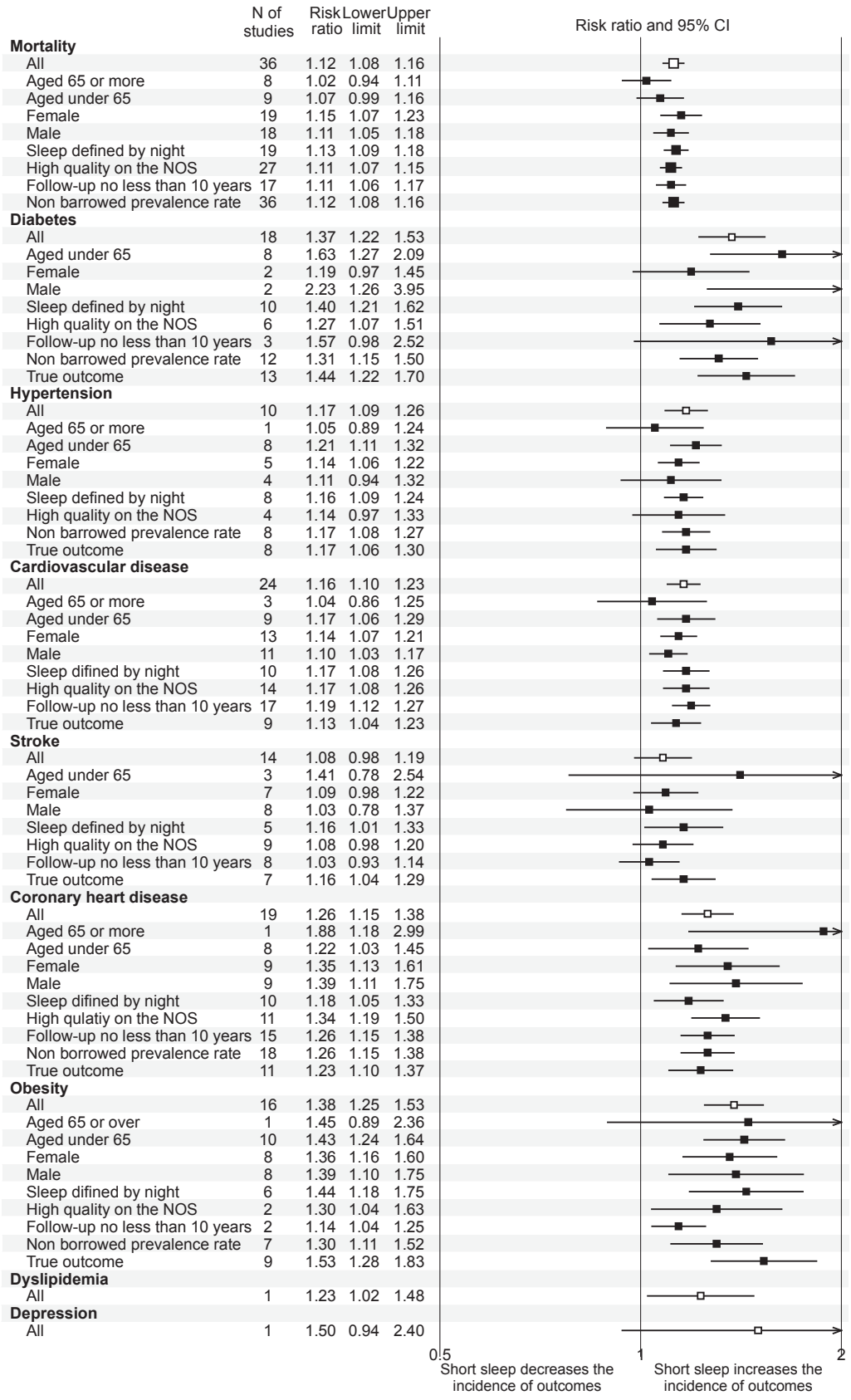


Fig. 2. Relative risks of mortality and health outcomes comparing short with normal sleepers. NOS, Newcastle–Ottawa Scale. For the outcomes of dyslipidemia and depression, no meta-analyses were performed.

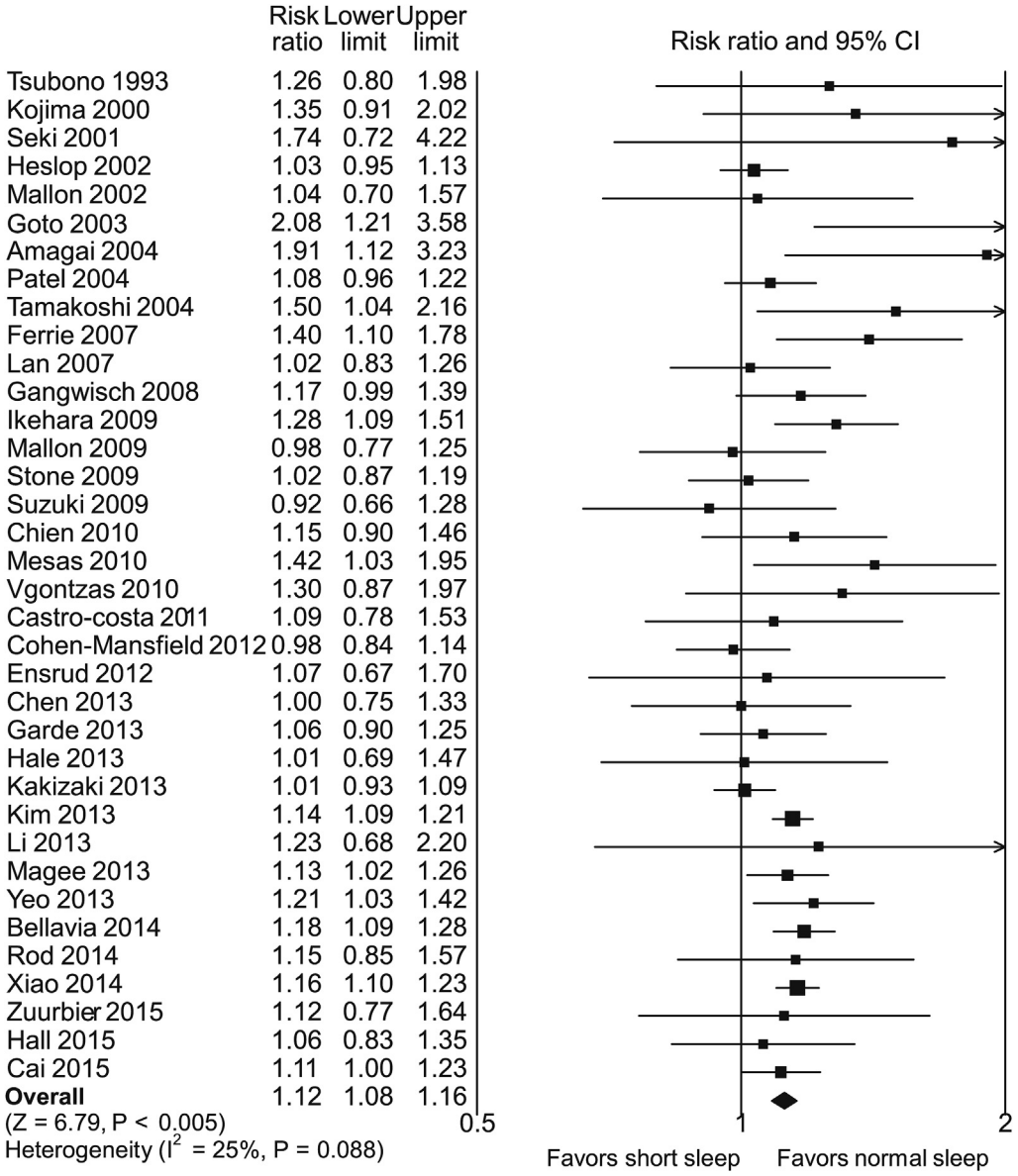


Fig. 3. All-cause mortality outcome comparing short with normal sleepers.

Our study (as in a previous review [5]) reported that in subgroup analyses, a gender difference was found in associations between short sleep duration and health outcomes, such as diabetes mellitus and hypertension. However, each gender may have one illness but not the other in association with short sleep, as observed in the present review.

Subgroup analyses also showed that short sleep defined as a duration less than six hours was associated with a significant increase in mortality, diabetes, cardiovascular disease, and obesity. In meta-regression, shorter duration of sleep definition was linearly associated with an increase in the mortality outcome. A significant association in the opposite trend was observed in the obesity outcome, and no significantly linear association was observed for the other outcomes; however, we were not able to conclude whether the association existed or not because of the small numbers of studies included in each outcome. An important limitation of prior epidemiological studies is that definitions of short sleep across studies can lead to prevention of adequate comparisons [73]. We believe this issue can be approached by the methodology we used in the present review.

Although the study confirmed that short sleep duration was associated with an increase in the important health outcomes, the mechanisms do not seem straightforward. Previous studies have showed that sleep deprivation reduces energy expenditure in healthy men [74], an insulin-resistant state in human adipocytes [75], and inadequate pancreatic insulin secretion and increased plasma glucose concentrations after a meal [76]. Thus, chronic sleep curtailment may lead to consequent outcomes related to metabolic systems. In addition, short sleep duration has been reported to be associated with increased ghrelin and decreased leptin, which is likely to increase appetite [77,78], possibly explaining increase in obesity and diabetes mellitus [79]. As well as changes in these hormones, sympathetic nervous system overactivity and changes in circadian rhythm may play an important role in development of hypertension [80,81]. Cardiovascular diseases can be caused by complex combinations of these metabolic and circulatory abnormalities. Although these endocrine and cardiovascular illnesses had been adjusted in statistical model in most of the studies investigating the mortality outcome, these illnesses might have occurred or worsened during the course in participants and caused

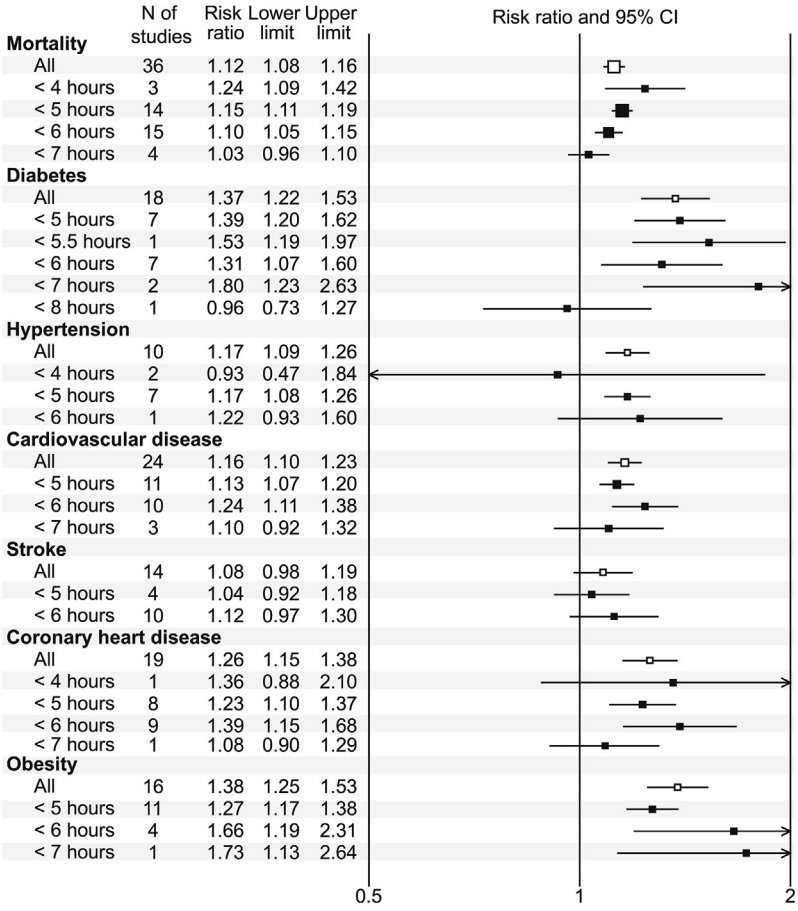


Fig. 4. Subgroup analyses for the duration of the definition of short sleep.

subsequent death. Although our findings appear to be very convincing, one may need to practice caution in terms of considering implications of the results in the community. First, although short sleep is associated with increased mortality and other health outcomes, there is no rigorous evidence that lengthening sleep duration can lead to smaller frequency of these outcomes. We do not intend to recommend hypnotics for people whose sleep is less than six hours a night. With regard to psychoeducation, psychotherapy, and psychosocial interventions, many studies have focused on insomnia and the efficacy of these interventions for the quality of sleep, which was reported in recent systematic reviews [82,83]. To the best of our knowledge, there has been no evidence to recommend these interventions for the community for the purpose of decreasing the frequency of mortality and other health outcomes. It should also be noted that the role of individual differences regarding sleep duration preferences is still uncertain [84]. Unmarried persons [85], those with lower socioeconomic status, or those who engage in frequent binge drinking [86] have been reported to be more likely to be short sleepers. It seems likely that psychoeducation or psychosocial intervention about the present review's results of the relationship between short sleep and health outcomes can encourage people to modify their behaviors; yet, the effects of this intervention may not lead to subsequent decrease in the frequency of health outcomes, and their preferences in terms of sleep should be taken into consideration. A pragmatic effectiveness comparative study in general community settings is needed to determine if psychoeducation, psychosocial, or lifestyle interventions can lead to prevention of premature death and the other important health outcomes.

4.1. Limitations of the study

First, we did not investigate the relationship between long sleep duration and mortality or health outcomes. Long sleep duration has been reported to be more widespread than short sleep [2]. Although a critical review argues that it is premature to conclude a U-shaped association between sleep duration and mortality risks [87], previous systematic reviews have repeatedly suggested that long sleep can be associated with an increase in mortality [3,4,6,73,88,89], hypertension [5,90,91], diabetes mellitus [91,92], and cardiovascular diseases [8,92]. We decided not to report outcomes associated with long sleep duration because of the journal's limited space, but we are planning (July 2016) a systematic review of long sleep duration using the same methodology.

Impacts of the quality of sleep in either subjective or objective measures on the prevalence of the health outcomes were also not investigated in the present review. Although previous studies have reported the association between insufficient sleep such as insomnia and mortality [63,93], we intended to focus on the duration of sleep because it is likely that sleep duration is more easily recognized and accurately reported by participants rather than the quality of sleep in community surveys. However, interventions to reduce mortality and the incidents of the other health outcomes may include information about not only sleep duration but also sleep quality. Previous systematic reviews on cognitive behavioral therapy for insomnia (CBT-I) delivered through the Internet [83] and self-help interventions [94] have revealed that CBT-I is likely to improve the sleep quality, thus components of these low-intensity interventions might be

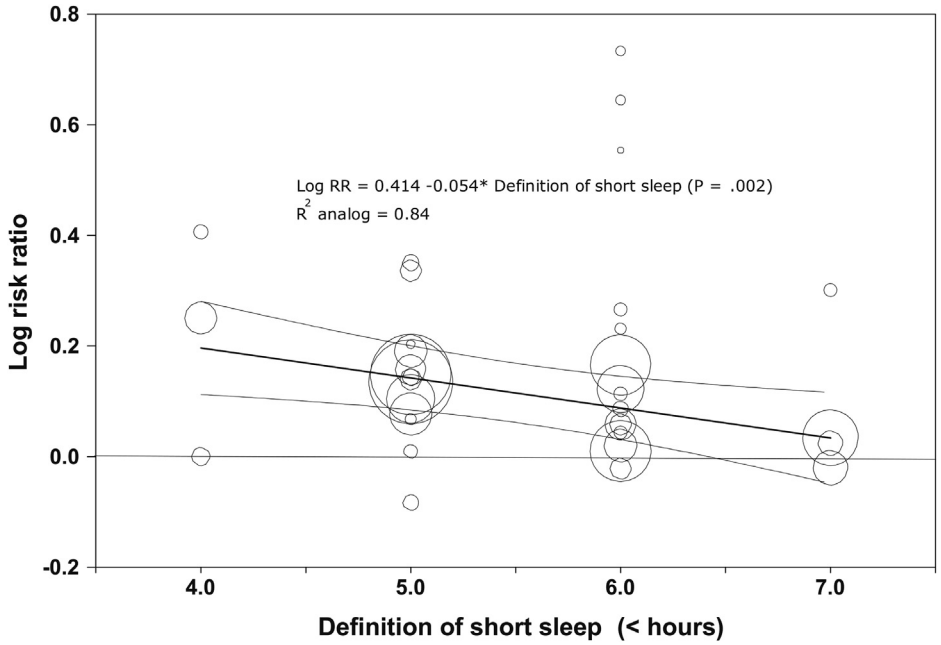


Fig. 5. Meta-regression for the duration of the definition of short sleep in mortality. Regression line and its prediction intervals (95% CIs) were presented in the figure.

considered to be incorporated to those conducted in the community settings.

Third, we did not include previous important surveys [15,69,95] that seemed eligible for the present systematic review, because we focused on studies that defined short and normal sleep duration and compared the incidence of health outcomes between those within these two conditions. These studies reported odds ratios of outcomes for each hour increase in sleep duration, and we regretted that we were not able to include the results from the study. However, we intended to include as many studies as we could by using several methodologies, such as calculating a RR from an OR using a CER. We believe that this is among merits of our systematic review.

5. Conclusions

Short sleep defined as the duration less than six hours was associated with a significant increase in mortality, diabetes, cardiovascular disease, coronary heart disease, and obesity. For clinical practice, the effectiveness of psychoeducation or psychosocial intervention to improve sleep in short sleepers on reducing mortality and the incidence of these health outcomes should be examined in general community settings.

Authors' contributions

All authors contributed to the manuscript as follows:

OI designed the study, developed technical materials, acquired data, and interpreted the data. MJ designed the study, developed technical materials, acquired data, and interpreted the data. NW designed the study, interpreted the data, and drafted the manuscript.

YK obtained funding, conceived the study, designed the study, and interpreted the data.

All authors have revised the important intellectual content critically, have read and approved the final manuscript, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

Funding sources

This study was funded by Health Labor Sciences Research Grant (H25-JUNKANKITOU-IPPAN-007) from the Japanese Ministry of Health Labor and Welfare.

The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Conflict of interest

The authors have no conflicts of interests to declare that may be affected by the publication of the paper. Other conflicts of interests are as follows:

OI has research funds from the Japanese Ministry of Health Labor and Welfare and the Japanese Ministry of Education, Science, and Technology.

MJ has research funds from the Japanese Ministry of Health Labor.

NW has received research funds from the Japanese Ministry of Health Labor and Welfare, the Japanese Ministry of Education, Science, and Technology and National Center of Neurology and Psychiatry, Intramural Research Grant for Neurological and Psychiatric Disorders. He has also received royalties from Sogensha, Paquet, and Akatsuki.

YK has research funds from the Japanese Ministry of Health Labor and Welfare and the Japanese Ministry of Education, Science, and Technology. He has also received royalties from Dai-Nippon Sumitomo, MSD, Otsuka, Eisai, Pfizer, and Takeda during the last 5 years.

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: http://dx.doi.org/10.1016/j.sleep.2016.08.006.

Acknowledgments

This work was performed at Oita University, Nihon University, and Kyoto University, Japan.

Appendix A. Supplementary material

Supplementary data to this article can be found online at doi:10.1016/j.sleep.2016.08.006.

References

[1] Ford ES, Cunningham TJ, Croft JB. Trends in self-reported sleep duration among US adults from 1985 to 2012. Sleep 2015;38:829–32.
[2] Bin YS, Marshall NS, Glozier N. Sleeping at the limits: the changing prevalence of short and long sleep durations in 10 countries. Am J Epidemiol 2013;177:826–33.
[3] Gallicchio L, Kalesan B. Sleep duration and mortality: a systematic review and meta-analysis. J Sleep Res 2009;18:148–58.
[4] Cappuccio FP, D'Elia L, Strazzullo P, et al. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep 2010;33:585–92.
[5] Guo X, Zheng L, Wang J, et al. Epidemiological evidence for the link between sleep duration and high blood pressure: a systematic review and meta-analysis. Sleep Med 2013;14:324–32.
[6] Shen X, Wu Y, Zhang D. Nighttime sleep duration, 24-hour sleep duration and risk of all-cause mortality among adults: a meta-analysis of prospective cohort studies. Sci Rep 2016;6:21480.
[7] Wang Y, Mei H, Jiang YR, et al. Relationship between duration of sleep and hypertension in adults: a meta-analysis. J Clin Sleep Med 2015;11:1047–56.
[8] Cappuccio FP, Cooper D, Delia L, et al. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. Eur Heart J 2011;32:1484–92.
[9] Leng Y, Cappuccio FP, Wainwright NW, et al. Sleep duration and risk of fatal and nonfatal stroke: a prospective study and meta-analysis. Neurology 2015;84:1072–9.
[10] Cappuccio FP, D'Elia L, Strazzullo P, et al. Quantity and quality of sleep and incidence of type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2010;33:414–20.
[11] Shan Z, Ma H, Xie M, et al. Sleep duration and risk of type 2 diabetes: a meta-analysis of prospective studies. Diabetes Care 2015;38:529–37.
[12] Wu Y, Zhai L, Zhang D. Sleep duration and obesity among adults: a meta-analysis of prospective studies. Sleep Med 2014;15:1456–62.
[13] Chaput JP, McNeil J, Despres JP, et al. Short sleep duration as a risk factor for the development of the metabolic syndrome in adults. Prev Med 2013;57:872–7.
[14] Sands MR, Lauderdale DS, Liu K, et al. Short sleep duration is associated with carotid intima-media thickness among men in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Stroke 2012;43:2858–64.
[15] King CR, Knutson KL, Rathouz PJ, et al. Short sleep duration and incident coronary artery calcification. JAMA 2008;300:2859–66.
[16] Hammond EC. Some preliminary findings on physical complaints from a prospective study of 1,064,004 men and women. Am J Public Health Nations Health 1964;54:11–23.
[17] Kaplan GA, Seeman TE, Cohen RD, et al. Mortality among the elderly in the Alameda County Study: behavioral and demographic risk factors. Am J Public Health 1987;77:307–12.
[18] Amagai Y, Ishikawa S, Gotoh T, et al. Sleep duration and mortality in Japan: the Jichi Medical School Cohort Study. J Epidemiol 2004;14:124–8.
[19] Tamakoshi A, Ohno Y, Group JS. Self-reported sleep duration as a predictor of all-cause mortality: results from the JACC study, Japan. Sleep 2004;27:51–4.
[20] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
[21] Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008–12.
[22] Zizi F, Pandey A, Murrray-Bachmann R, et al. Race/ethnicity, sleep duration, and diabetes mellitus: analysis of the National Health Interview Survey. Am J Med 2012;125:162–7.
[23] Adenekan B, Pandey A, McKenzie S, et al. Sleep in America: role of racial/ethnic differences. Sleep Med Rev 2013;17:255–62.
[24] Lauderdale DS, Knutson KL, Yan LL, et al. Objectively measured sleep characteristics among early-middle-aged adults: the CARDIA study. Am J Epidemiol 2006;164:5–16.
[25] Silva GE, Goodwin JL, Sherrill DL, et al. Relationship between reported and measured sleep times: the sleep heart health study (SHHS). J Clin Sleep Med 2007;3:622–30.
[26] Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2008.
[27] Aziz O, Constantinides V, Tekkis PP, et al. Laparoscopic versus open surgery for rectal cancer: a meta-analysis. Ann Surg Oncol 2006;13:413–24.
[28] Millett GA, Flores SA, Marks G, et al. Circumcision status and risk of HIV and sexually transmitted infections among men who have sex with men: a meta-analysis. JAMA 2008;300:1674–84.
[29] Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58.
[30] Aburto NJ, Hanson S, Gutierrez H, et al. Effect of increased potassium intake on cardiovascular risk factors and disease: systematic review and meta-analyses. BMJ 2013;346:f1378.
[31] Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.
[32] Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455–63.
[33] Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med 1992;116:78–84.
[34] Borenstein M, Hedges LV, Higgins JPT, et al. PART 4: meta-regression is observational. In: Borenstein M, Hedges LV, Higgins JPT, et al., editors. Regression in meta-analysis. Englewood (NJ): Biostat, Inc.; 2015. p. 19–20.
[35] Borenstein M, Hedges L, Higgins J, et al. Comprehensive meta analysis, Version 3. 3.3.070 ed. Englewood (NJ): Biostat, Inc.; 2014.
[36] Tsubono Y, Fukao A, Hisamichi S. Health practices and mortality in a rural Japanese population. Tohoku J Exp Med 1993;171:339–48.
[37] Kojima M, Wakai K, Kawamura T, et al. Sleep patterns and total mortality: a 12-year follow-up study in Japan. J Epidemiol 2000;10:87–93.
[38] Seki N. Relationships between walking hours, sleeping hours, meaningfulness of life (ikigai) and mortality in the elderly: prospective cohort study. Nihon Eiseigaku Zasshi 2001;56:535–40. in Japanese.
[39] Heslop P, Smith GD, Metcalfe C, et al. Sleep duration and mortality: the effect of short or long sleep duration on cardiovascular and all-cause mortality in working men and women. Sleep Med 2002;3:305–14.
[40] Mallon L, Broman JE, Hetta J. Sleep complaints predict coronary artery disease mortality in males: a 12-year follow-up study of a middle-aged Swedish population. J Intern Med 2002;251:207–16.
[41] Goto A, Yasumura S, Nishise Y, et al. Association of health behavior and social role with total mortality among Japanese elders in Okinawa, Japan. Aging Clin Exp Res 2003;15:443–50.
[42] Patel SR, Ayas NT, Malhotra MR, et al. A prospective study of sleep duration and mortality risk in women. Sleep 2004;27:440–4.
[43] Ferrie JE, Shipley MJ, Cappuccio FP, et al. A prospective study of change in sleep duration: associations with mortality in the Whitehall II cohort. Sleep 2007;30:1659–66.
[44] Lan TY, Lan TH, Wen CP, et al. Nighttime sleep, Chinese afternoon nap, and mortality in the elderly. Sleep 2007;30:1105–10.
[45] Gangwisch JE, Heymsfield SB, Boden-Albala B, et al. Sleep duration associated with mortality in elderly, but not middle-aged, adults in a large US sample. Sleep 2008;31:1087–96.
[46] Ikehara S, Iso H, Date C, et al. Association of sleep duration with mortality from cardiovascular disease and other causes for Japanese men and women: the JACC study. Sleep 2009;32:295–301.
[47] Mallon L, Broman JE, Hetta J. Is usage of hypnotics associated with mortality? Sleep Med 2009;10:279–86.
[48] Stone KL, Ewing SK, Ancoli-Israel S, et al. Self-reported sleep and nap habits and risk of mortality in a large cohort of older women. J Am Geriatr Soc 2009;57:604–11.
[49] Suzuki E, Yorifuji T, Ueshima K, et al. Sleep duration, sleep quality and cardiovascular disease mortality among the elderly: a population-based cohort study. Prev Med 2009;49:135–41.
[50] Chien KL, Chen PC, Hsu HC, et al. Habitual sleep duration and insomnia and the risk of cardiovascular events and all-cause death: report from a community-based cohort. Sleep 2010;33:177–84.
[51] Mesas AE, Lopez-Garcia E, Leon-Munoz LM, et al. Sleep duration and mortality according to health status in older adults. J Am Geriatr Soc 2010;58:1870–7.
[52] Vgontzas AN, Liao D, Pejovic S, et al. Insomnia with short sleep duration and mortality: the Penn State cohort. Sleep 2010;33:1159–64.
[53] Castro-Costa E, Dewey ME, Ferri CP, et al. Association between sleep duration and all-cause mortality in old age: 9-year follow-up of the Bambui Cohort Study, Brazil. J Sleep Res 2011;20:303–10.
[54] Kripke DF, Langer RD, Elliott JA, et al. Mortality related to actigraphic long and short sleep. Sleep Med 2011;12:28–33.
[55] Kronholm E, Laatikainen T, Peltonen M, et al. Self-reported sleep duration, all-cause mortality, cardiovascular mortality and morbidity in Finland. Sleep Med 2011;12:215–21.
[56] Cohen-Mansfield J, Perach R. Sleep duration, nap habits, and mortality in older persons. Sleep 2012;35:1003–9.
[57] Ensrud KE, Blackwell TL, Ancoli-Israel S, et al. Sleep disturbances and risk of frailty and mortality in older men. Sleep Med 2012;13:1217–25.
[58] Chen HC, Su TP, Chou P. A nine-year follow-up study of sleep patterns and mortality in community-dwelling older adults in Taiwan. Sleep 2013;36:1187–98.
[59] Garde AH, Hansen AM, Holtermann A, et al. Sleep duration and ischemic heart disease and all-cause mortality: prospective cohort study on effects of tranquilizers/hypnotics and perceived stress. Scand J Work Environ Health 2013;39:550–8.
[60] Hale L, Parente V, Dowd JB, et al. Fibrinogen may mediate the association between long sleep duration and coronary heart disease. J Sleep Res 2013;22:305–14.
[61] Kakizaki M, Kuriyama S, Nakaya N, et al. Long sleep duration and cause-specific mortality according to physical function and self-rated health: the Ohsaki Cohort Study. J Sleep Res 2013;22:209–16.

[62] Kim Y, Wilkens LR, Schembre SM, et al. Insufficient and excessive amounts of sleep increase the risk of premature death from cardiovascular and other diseases: the Multiethnic Cohort Study. Prev Med 2013;57:377–85.
[63] Li Y, Zhang X, Winkelman JW, et al. Association between insomnia symptoms and mortality: a prospective study of U.S. men. Circulation 2014;129:737–46.
[64] Magee CA, Holliday EG, Attia J, et al. Investigation of the relationship between sleep duration, all-cause mortality, and preexisting disease. Sleep Med 2013;14:591–6.
[65] Yeo Y, Ma SH, Park SK, et al. A prospective cohort study on the relationship of sleep duration with all-cause and disease-specific mortality in the Korean Multi-center Cancer Cohort study. J Prev Med Public Health 2013;46:271–81.
[66] Bellavia A, Akerstedt T, Bottai M, et al. Sleep duration and survival percentiles across categories of physical activity. Am J Epidemiol 2014;179:484–91.
[67] Rod NH, Kumari M, Lange T, et al. The joint effect of sleep duration and disturbed sleep on cause-specific mortality: results from the Whitehall II cohort study. PLoS ONE 2014;9. e91965.
[68] Xiao Q, Keadle SK, Hollenbeck AR, et al. Sleep duration and total and cause-specific mortality in a large US cohort: interrelationships with physical activity, sedentary behavior, and body mass index. Am J Epidemiol 2014;180:997–1006.
[69] Zawisza K, Tobiasz-Adamczyk B, Galas A, et al. Sleep duration and mortality among older adults in a 22-year follow-up study: an analysis of possible effect modifiers. Eur J Ageing 2015;12:119–29.
[70] Zuurbier LA, Luik AI, Hofman A, et al. Fragmentation and stability of circadian activity rhythms predict mortality: the Rotterdam study. Am J Epidemiol 2015;181:54–63.
[71] Hall MH, Smagula SF, Boudreau RM, et al. Association between sleep duration and mortality is mediated by markers of inflammation and health in older adults: the Health, Aging and Body Composition Study. Sleep 2015;38:189–95.
[72] Cai H, Shu XO, Xiang YB, et al. Sleep duration and mortality: a prospective study of 113,138 middle-aged and elderly Chinese men and women. Sleep 2015;38:529–36.
[73] Grandner MA, Hale L, Moore M, et al. Mortality associated with short sleep duration: the evidence, the possible mechanisms, and the future. Sleep Med Rev 2010;14:191–203.
[74] Benedict C, Hallschmid M, Lassen A, et al. Acute sleep deprivation reduces energy expenditure in healthy men. Am J Clin Nutr 2011;93:1229–36.
[75] Broussard JL, Ehrmann DA, Van Cauter E, et al. Impaired insulin signaling in human adipocytes after experimental sleep restriction: a randomized, cross-over study. Ann Intern Med 2012;157:549–57.
[76] Buxton OM, Cain SW, O'Connor SP, et al. Adverse metabolic consequences in humans of prolonged sleep restriction combined with circadian disruption. Sci Transl Med 2012;4:129ra43.
[77] Taheri S, Lin L, Austin D, et al. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. PLoS Med 2004;1:e62.
[78] Spiegel K, Tasali E, Penev P, et al. Brief communication: sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. Ann Intern Med 2004;141:846–50.
[79] Van Cauter E, Holmback U, Knutson K, et al. Impact of sleep and sleep loss on neuroendocrine and metabolic function. Horm Res 2007;67(Suppl. 1):2–9.
[80] Ogawa Y, Kanbayashi T, Saito Y, et al. Total sleep deprivation elevates blood pressure through arterial baroreflex resetting: a study with microneurographic technique. Sleep 2003;26:986–9.
[81] Gangwisch JE. Epidemiological evidence for the links between sleep, circadian rhythms and metabolism. Obes Rev 2009;10:37–45.
[82] Wu JQ, Appleman ER, Salazar RD, et al. Cognitive behavioral therapy for insomnia comorbid with psychiatric and medical conditions: a meta-analysis. JAMA Intern Med 2015;175:1461–72.
[83] Zachariae R, Lyby MS, Ritterband LM, et al. Efficacy of internet-delivered cognitive-behavioral therapy for insomnia – a systematic review and meta-analysis of randomized controlled trials. Sleep Med Rev 2015;30:1–10.
[84] Grandner MA, Patel NP. From sleep duration to mortality: implications of meta-analysis and future directions. J Sleep Res 2009;18:145–7.
[85] Stranges S, Dorn JM, Shipley MJ, et al. Correlates of short and long sleep duration: a cross-cultural comparison between the United Kingdom and the United States: the Whitehall II Study and the Western New York Health Study. Am J Epidemiol 2008;168:1353–64.
[86] Krueger PM, Friedman EM. Sleep duration in the United States: a cross-sectional population-based study. Am J Epidemiol 2009;169:1052–63.
[87] Kurina LM, McClintock MK, Chen JH, et al. Sleep duration and all-cause mortality: a critical review of measurement and associations. Ann Epidemiol 2013;23:361–70.
[88] Grandner MA, Drummond SPA. Who are the long sleepers? Towards an understanding of the mortality relationship. Sleep Med Rev 2007;11:341–60.
[89] Liu TZ, Xu C, Rota M, et al. Sleep duration and risk of all-cause mortality: a flexible, non-linear, meta-regression of 40 prospective cohort studies. Sleep Med Rev 2016. http://dx.doi.org/10.1016/j.smrv.2016.02.005.
[90] Wang Q, Xi B, Liu M, et al. Short sleep duration is associated with hypertension risk among adults: a systematic review and meta-analysis. Hypertens Res 2012;35:1012–8.
[91] Araghi MH, Thomas GN, Taheri S. The potential impact of sleep duration on lipid biomarkers of cardiovascular disease. Clin Lipidol 2012;7:443–53.
[92] Alvarez GG, Ayas NT. The impact of daily sleep duration on health: a review of the literature. Prog Cardiovasc Nurs 2004;19:56–9.
[93] Parthasarathy S, Vasquez MM, Halonen M, et al. Persistent insomnia is associated with mortality risk. Am J Med 2015;128:268–75. e2.
[94] Ho FY, Chung KF, Yeung WF, et al. Self-help cognitive-behavioral therapy for insomnia: a meta-analysis of randomized controlled trials. Sleep Med Rev 2015;19:17–28.
[95] Petrov ME, Kim Y, Lauderdale D, et al. Longitudinal associations between objective sleep and lipids: the CARDIA study. Sleep 2013;36:1587–95.

Alcohol and Alcoholism, 2017, 52(3) 351–357
doi: 10.1093/alcalc/agw081
Advance Access Publication Date: 25 October 2016
Article

Article

The Association Between Alcohol-Flavoured Non-Alcoholic Beverages and Alcohol Use in Japanese Adolescents

Aya Kinjo[1,*], Aya Imamoto[1], Maki Ikeda[2], Osamu Itani[3], Takashi Ohida[2], Yoshitaka Kaneita[3], Hideyuki Kanda[4], Takeo Tanihata[5], Susumu Higuchi[6], and Yoneatsu Osaki[1]

[1]Division of Environmental and Preventive Medicine, Department of Social Medicine, Faculty of Medicine, Tottori University, Nishi-machi 86, Yonago-shi, Tottori 683-8503, Japan, [2]Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, Itabashi-ku, Tokyo 173-8610, Japan, [3]Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Yufu-shi, Oita 879-5593, Japan, [4]Department of Public Health, Faculty of Medicine, Shimane University, Izumo-shi, Shimane 693-8501, Japan, [5]Department of Health, Bureau of Public Health and Welfare, Kobe City, 6 Chome-5-1 Kano-cho, Chuo Ward, Kobe, Hyogo 650-0001, Japan, and [6]National Hospital Organization Kurihama Medical and Addiction Center, Yokosuka-shi, Kanagawa 239-0841, Japan

*Corresponding author: Division of Environmental and Preventive Medicine, Department of Social Medicine, Faculty of Medicine, Tottori University, Nishi-machi 86, Yonago-shi, Tottori 683-8503, Japan. Tel.: +81-859-38-6103; Fax: +81-859-38-6100; E-mail: kinjo_aya@med.tottori-u.ac.jp

Received 12 July 2016; Revised 4 October 2016; Editorial Decision 6 October 2016; Accepted 11 October 2016

Abstract

Aims: There are no legal regulations in Japan governing minors' consumption of alcohol-flavoured non-alcoholic beverages (AFNAB); therefore, we examined if their consumption could lead to increased alcohol use among adolescents in Japan.

Methods: This cross-sectional study used a nonclinical, nationally representative sample of 38,494 junior (19,662 boys) and 61,556 senior (31,925 boys) high school students recruited in 2012. We measured AFNAB consumption rates and the order that adolescents first consumed AFNAB and alcohol.

Results: The AFNAB consumption was strongly associated with alcohol use in high school students. Among all age groups, alcohol was more commonly consumed before AFNAB for both males and females.

Conclusions: Consumption of AFNAB is more prevalent among minors than alcohol consumption and it has a strong association with alcohol consumption. However, concerns that AFNAB use would lead to increased alcohol use were not supported because AFNAB consumption usually started after adolescents began consuming alcohol.

Short Summary: Consumption of AFNAB is more prevalent among high school students than alcohol consumption and it has a strong association with alcohol consumption. However, concerns that AFNAB use would lead to increased alcohol use were not supported because AFNAB consumption usually started after adolescents began consuming alcohol.

INTRODUCTION

Relative to adults, adolescents experience greater physical and social harm with alcohol use. Adolescent alcohol use elevates the risk of unintentional injury (Hingson *et al.*, 2000) and unsafe sex (Hingson *et al.*, 2003) and leads to alcohol dependence (Hingson *et al.*, 2006). Furthermore, early alcohol consumption is a strong predictor of

illicit drug use (Kandel and Yamaguchi, 1993). Young people who have an earlier drinking onset are at more risk of alcohol dependence/abuse and of unintentionally injuring themselves after drinking (Hingson and Zha, 2009). Therefore, adolescent alcohol use is prohibited in numerous countries. However, in a Japanese study conducted in 2010, the prevalence of alcohol use in 12th grade students within the 30 days preceding the study was 21.0% in boys and 18.5% in girls (Central Research Services, Inc., 2010). In a similar study conducted in the United States in 2011, prevalence rates were 42.1 and 37.5% in boys and girls, respectively (Patrick *et al.*, 2013). Furthermore, while the prevalence of adolescent alcohol use tends to decrease annually, it remains high (Osaki *et al.*, 2009; Patrick and Schulenberg, 2010).

In Japan, the sale of alcohol-flavoured non-alcoholic beverages (AFNAB) has increased in recent years (SUNTORY, 2015), and there is some concern that consumption of these beverages could lead to an increase in adolescents' alcohol use. The average price of AFNAB (1.40 US dollars) is 1.4 times higher than that of soft drinks (1.00 US dollar) and is the same price as alcohol. The demand for these beverages has increased in Japan since 2007 due to the implementation of a campaign to reduce drinking and driving. Alcohol-free beer was launched in Japan in 2009, and beverage companies subsequently released various non-alcoholic beverages, such as those that taste like wine, cocktails, and alcopops, increasing sales of AFNAB. The number of adults who consume AFNAB is increasing because people are permitted to drive after drinking AFNAB and women can drink AFNAB when they are pregnant and breastfeeding. In addition, some AFNAB are authorized as 'food for specified health uses', which means that they are deemed positive for health by the Consumer Affairs Agency under the Japanese Government (Nikkei Asian Review, 2015). Moreover, beverages with an alcohol content of <1% are classified as soft drinks in Japan. Therefore, the consumption and purchase of AFNAB by minors is permitted in Japanese legislation. In addition to alcohol, beverage companies and retailers self-regulate AFNAB advertisements towards minors (Liaison Committee about Alcohol Advertising, 2012).

A survey of Japanese parents (aged in their 30s and 40s) revealed that more than 60% permitted their minor children to consume AFNAB (Miyaki, 2012). Reasons for this included that AFNAB are permitted by law, there is no reason to prohibit AFNAB, and AFNAB prevent minors from drinking alcohol. AFNAB are not considered detrimental to one's health even though they appear and taste remarkably similar to alcoholic drinks. On the other hand, the reason why some parents did not allow their children to consume AFNAB was that it creates a habit of drinking alcohol and strengthens minors' interest in alcohol. Consequently, there is some concern that consumption of these beverages could lead to an increase in adolescents' alcohol use.

A study conducted in 2011 that examined AFNAB consumption rates among 9775 Japanese senior high school students (Kubo *et al.*, 2015) revealed that 25.8% of senior high school boys and 26.1% of senior high school girls had consumed AFNAB at least once. In the current study, we investigated AFNAB consumption in a nationally representative sample of Japanese adolescents containing both junior and senior high school students in order to examine whether AFNAB consumption occurs at an earlier age. We hypothesized that AFNAB would be consumed at an earlier age than alcohol.

MATERIALS AND METHODS

Participants

A single-stage cluster sampling method was employed, with 10,018 junior and 4603 senior high schools (14,621 in total) in Japan registered for the study as of May 2012. Of these, 140 junior (selection rate: 1.4%) and 124 senior high schools (selection rate: 2.7%) were randomly selected. Probability proportional to size sampling was employed to ensure that the probability of selection was proportional to the number of students enrolled. The sample size was determined according to school response rates and confidence interval ranges based on the variance in the results of five preceding nationwide surveys examining alcohol consumption behaviour (1996, 2000, 2004, 2008 and 2010; Suzuki *et al.*, 2000; Ohida *et al.*, 2004; Kaneita *et al.*, 2006; Munezawa *et al.*, 2011; Morioka *et al.*, 2013). The sample size was similar to those of the previous surveys (Morioka *et al.*, 2013).

Japanese children enter junior high school at the age of 13 years and study for 3 years; they then attend senior high school for a further 3 years. In this report, years 1–3 of junior and senior high school are referred to as grades 7–9 and 10–12, respectively.

Survey Procedure

Letters of invitation to participate in the survey were sent to the principals of the selected schools, with sufficient questionnaires and envelopes for all students included. Homeroom teachers distributed the questionnaires to the students. To protect respondents' privacy and obtain frank responses, teachers were asked to comply with the following guidelines: (a) please do not make positive or negative remarks about alcohol use prior to the survey, (b) please do not read the students' responses while they are completing the questionnaires and (c) envelopes should remain sealed to protect students' privacy. Teachers were also asked to inform the students that the questionnaire envelopes would remain sealed. To ensure compliance with these guidelines, teachers were provided with survey implementation guides. In addition, the questionnaire included a statement confirming that the teachers would not see the completed questionnaires. Students were instructed to insert the completed questionnaire into the envelope provided and seal the adhesive flap. The teachers sent the sealed envelopes to the Nihon University School of Medicine. Participants gave their informed consent on the first page of the questionnaire and they did not receive any credit or money for participation. The survey was conducted between October 2012 and March 2013. The study was approved by the ethics committee of the School of Medicine, Nihon University, and conducted in accordance with the Declaration of Helsinki.

Response Rates

In total, 101,134 responses to the 2012 survey were obtained from 94 junior (response rate 67.1%) and 85 senior high schools (68.5%), which accounted for 59.8% and 61.3% of the junior and senior school students, respectively. The number of valid responses was 100,050 (51,587 boys and 48,463 girls) following exclusion of invalid responses with missing sex or grade data. We combined unweighted data for all grades and compared subtotal prevalence for the junior and senior high schools. School characteristics, such as type (private, vocational or general), were chosen to represent the study population. Average numbers of enrolled students did not differ between responding and non-responding schools.

Measures

Two questionnaire items were used to measure the consumption of AFNAB. The first item asked participants, 'Do you drink beverages without alcohol (e.g. non-alcohol, alcohol-free, alcohol zero) such as

non-alcoholic beer, non-alcoholic cocktails and non-alcoholic alcopops?' Participants chose between five answers: 'I have never; I have before, but not now; I do sometimes; I usually do; and I do not know'. If participants had consumed both alcohol and AFNAB, they were asked the second item, 'What did you drink first: beverages without alcohol (e.g. non-alcohol, alcohol-free, alcohol zero) such as non-alcoholic beer, non-alcoholic cocktails and non-alcoholic alcopops or alcohol?' Again, participants chose between five answers: 'beverages without alcohol, alcohol, both almost at the same time, I have drunk either only beverages without alcohol or alcohol, and I do not know'.

The other questionnaire about alcohol use was based on those of the preceding five surveys (1996, 2000, 2004, 2008 and 2010; Suzuki *et al.*, 2000; Ohida *et al.*, 2004; Kaneita *et al.*, 2006; Munezawa *et al.*, 2011; Morioka *et al.*, 2013). Accordingly, the questionnaire concerning alcohol use focused on participants' experiences, frequency of alcohol use, amount of alcohol consumed, and binge drinking frequency. Binge drinking was defined as consumption of more than six glasses of alcohol per occasion.

Statistical analysis

Participants' alcohol and AFNAB use behaviour were calculated according to sex, grade (7–12), and school stage (junior and senior). The odds ratio of grades in relation to the 7th grade, and of girls in relation to boys regarding each AFNAB consumption behaviour ('consume at least once' or not, 'sometimes consume' or not and 'usually consume' or not) were calculated using the odds ratio formula. The odds ratio of 'AFNAB consumed at least once' in relation to 'never consume AFNAB' in terms of alcohol drinking behaviour (at least one instance of alcohol consumption or not, and alcohol consumption within the last 30 days or not) were also calculated using the odds ratio formula, according to sex and school stage (junior and senior). Then, a binominal logistic regression analysis was performed to calculate the odds ratios of AFNAB consumption frequency in relation to 'never consume AFNAB' regarding alcohol drinking behaviour. Alcohol drinking behaviour (at least one instance of alcohol consumption or not, and alcohol consumption within the last 30 days or not) was treated as the dependent variable, and AFNAB consumption status ('never consume', 'consumed before, but not now', 'consume sometimes' or 'usually consume') was treated as an explanatory variable. With regard to which type of drink the students consumed first (i.e. AFNAB or alcohol), we used the data from participants who had consumed both alcohol and AFNAB, and then calculated the odds ratio according to sex, grade (7–12) and school stage (junior and senior). All analyses were performed using the SPSS 22.0 software programme (IBM Corp., Armonk, New York).

RESULTS

Table 1 shows the participants' characteristics concerning alcohol use behaviour by sex and grade. As the grade increased, the frequency of alcohol drinking and binge drinking increased. Drinking 'one time or more in the lifetime' and 'more than once a month' was higher in girls; however, drinking 'more than once a week' was higher in boys.

Table 2 shows the consumption rate of AFNAB by sex and grade. Compared to the data shown in Table 1, the 'consumed at least once' rates of AFNAB were higher than those of alcohol consumption in 7th to 10th grade boys and girls. The percentages of students who had consumed AFNAB were lowest in seventh grade boys and girls. The grade odds ratios increased until grade 10, and then decreased or did not change in all AFNAB consumption categories. The sex odds ratios of consuming AFNAB were higher in girls than boys, except for 'usually consume' in grades 7–9 and 11.

Table 3 shows the percentages and odds ratio of alcohol drinking behaviour according to AFNAB consumption. The odds ratio for both 'at least one instance of consuming alcohol' and 'drank alcohol within 30 days' increased as AFNAB consumption increased. These relationships were shown regardless of the school stage or gender.

Table 4 shows the proportion of students for whom AFNAB consumption preceded and followed alcohol use according to alcohol use behaviour among consumers of both AFNAB and alcohol.

Table 1. Participants' characteristics concerning alcohol use behaviour by sex and grade

	Grade	Total *n*	Alcohol drinking frequency (%)			Alcohol binge drinking frequency[c] (%)
			One time or more in the lifetime	More than once a month[a]	More than once a week[b]	One time or more in the lifetime
Boys						
Junior high school	7th	6880	16.2	3.3	0.9	10.2
	8th	6527	18.0	4.6	1.1	11.4
	9th	6161	20.3	6.9	2.3	14.1
Senior high school	10th	11,207	26.3	11.0	3.0	16.3
	11th	10,448	28.5	13.3	3.9	18.2
	12th	10,178	28.8	14.7	4.9	20.8
Girls						
Junior high school	7th	6457	14.8	3.4	0.9	10.1
	8th	6302	18.0	5.7	1.1	11.2
	9th	5994	19.2	6.8	1.6	12.0
Senior high school	10th	10,220	26.9	12.2	2.8	15.8
	11th	9514	29.7	13.6	3.0	17.4
	12th	9793	31.7	15.2	3.0	19.2

[a]This contains 'once or twice a month', 'weekly', 'several times per week' and 'every day'.
[b]This contains 'weekly', 'several times per week' and 'every day'.
[c]'Alcohol binge drinking' means drinking more than six glasses of alcohol per occasion.

Table 2. The rate of alcohol-flavoured non-alcoholic beverages consumption behaviour and the odds ratios of grade and sex in terms of alcohol-flavoured non-alcoholic beverages consumption behaviour

	Consumed at least once[a]					Consume sometimes					Usually consume				
	Boys		Girls		Girls' odds ratio (95% CI)[c]	Boys		Girls		Girls' odds ratio (95% CI)[c]	Boys		Girls		Girls odds ratio (95% CI)[c]
	%	Grade odds ratio (95% CI)[b]	%	Grade odds ratio (95% CI)[b]		%	Grade odds ratio (95% CI)[b]	%	Grade odds ratio (95% CI)[b]		%	Grade odds ratio (95% CI)[b]	%	Grade odds ratio (95% CI)[b]	
7th	22.6	1.0 (reference)	26.7	1.0 (reference)	1.3 (1.2–1.4)	5.0	1.0 (reference)	6.3	1.0 (reference)	1.3 (1.1–1.5)	1.0	1.0 (reference)	0.8	1.0 (reference)	0.7 (0.5–1.0)
8th	24.8	1.1 (1.1–1.2)	29.4	1.1 (1.1–1.2)	1.3 (1.2–1.4)	6.3	1.3 (1.1–1.5)	7.8	1.3 (1.1–1.5)	1.3 (1.1–1.4)	1.1	1.0 (0.7–1.4)	1.3	1.7 (1.2–2.5)	1.2 (0.9–1.7)
9th	26.3	1.2 (1.1–1.3)	30.7	1.2 (1.1–1.3)	1.2 (1.1–1.3)	6.6	1.4 (1.2–1.6)	7.6	1.2 (1.1–1.4)	1.2 (1.0–1.3)	1.4	1.4 (1.0–1.9)	1.7	2.3 (1.6–3.3)	1.2 (0.9–1.6)
10th	29.0	1.4 (1.3–1.5)	36.2	1.6 (1.5–1.7)	1.4 (1.3–1.5)	7.6	1.6 (1.4–1.8)	11.4	1.9 (1.7–2.2)	1.6 (1.4–1.7)	1.4	1.4 (1.0–1.8)	2.0	2.6 (1.9–3.6)	1.4 (1.1–1.7)
11th	27.7	1.3 (1.2–1.4)	35.1	1.5 (1.4–1.6)	1.4 (1.3–1.5)	7.1	1.5 (1.3–1.7)	11.5	1.9 (1.7–2.2)	1.7 (1.6–1.9)	1.4	1.3 (1.0–1.8)	1.7	2.3 (1.7–3.1)	1.2 (1.0–1.5)
12th	27.0	1.3 (1.2–1.4)	34.8	1.5 (1.4–1.6)	1.5 (1.4–1.5)	6.6	1.4 (1.2–1.6)	10.4	1.7 (1.6–2.0)	1.7 (1.5–1.8)	1.2	1.2 (0.9–1.6)	1.7	2.3 (1.7–3.2)	1.4 (1.1–1.8)
Junior high school	24.5	1.0 (reference)	28.9	1.0 (reference)	1.3 (1.2–1.3)[d]	5.9	1.0 (reference)	7.2	1.0 (reference)	1.2 (1.1–1.3)[d]	1.2	1.0 (reference)	1.3	1.0 (reference)	1.1 (0.9–1.3)[d]
Senior high school	27.9	1.2 (1.2–1.2)[e]	35.4	1.4 (1.3–1.4)[e]	1.4 (1.4–1.5)[d]	7.1	1.2 (1.1–1.3)[e]	11.1	1.6 (1.5–1.7)[e]	1.6 (1.5–1.7)[d]	1.3	1.1 (1.0–1.3)[e]	1.8	1.5 (1.2–1.7)[e]	1.3 (1.2–1.5)[d]

Note: AFNAB = alcohol-flavoured non-alcoholic beverages; CI = confidence interval.

[a]This contains 'consumed before, but not now' (not shown in this table), 'consume sometimes' and 'usually consume'.

[b]The 8th to 12th grades' odds ratios in relation to the seventh grade for each AFNAB consumption behaviour ('consume at least once' or not, 'sometimes consume' or not and 'usually consume' or not) were all calculated by sex.

[c]The girls' odds ratios in relation to boys for each AFNAB consumption behaviour ('consume at least once' or not, 'sometimes consume' or not and 'usually consume' or not) were all calculated by grade.

[d]The girls' odds ratios in relation to boys for each AFNAB consumption behaviour ('consume at least once' or not, 'sometimes consume' or not and 'usually consume' or not) were calculated by each school stage.

[e]The senior high school students' odds ratios in relation to junior high school students for each AFNAB consumption behaviour ('consume at least once' or not, 'sometimes consume' or not and 'usually consume' or not) were calculated by sex.

Table 3. The percentages of students who drink alcohol for each alcohol-flavoured non-alcoholic beverages consumption behaviour and the odds ratios of alcohol-flavoured non-alcoholic beverages consumption behaviours for 'never consume' in relation to alcohol drinking behaviour by each school stage and sex

	Consumption behaviour of AFNAB							
	Among never consume		Among consumed at least once[a]		Among consume sometimes		Among usually consume	
	%[b]	Odds ratio (95% CI)	%[b]	Odds ratio[c] (95% CI)	%[b]	Odds ratio[d] (95% CI)	%[b]	Odds ratio[d] (95% CI)
Boys								
At least one instance of drinking alcohol								
Junior high school total	10.1	1.00 (reference)	45.6	7.5 (6.9–8.1)	61.7	14.3 (12.6–16.3)	67.4	18.4 (13.9–24.4)
Senior high school total	17.5	1.00 (reference)	57.9	6.5 (6.1–6.8)	74.1	13.5 (12.2–14.9)	74.6	13.8 (11.1–17.3)
Drank alcohol within the last 30 days								
Junior high school total	3.2	1.00 (reference)	20.6	8.0 (7.1–9.0)	34.9	16.5 (14.1–19.2)	46.5	26.7 (20.3–35.1)
Senior high school total	7.8	1.00 (reference)	31.3	5.4 (5.0–5.8)	49.7	11.6 (10.6–12.8)	56.6	15.4 (12.6–18.7)
Girls								
At least one instance of drinking alcohol								
Junior high school total	8.4	1.00 (reference)	41.6	7.8 (7.2–8.4)	60.4	16.6 (14.6–18.8)	72.6	28.9 (21.6–38.8)
Senior high school total	16.8	1.00 (reference)	54.9	6.0 (5.7–6.4)	70.5	11.8 (10.8–12.8)	79.5	19.2 (15.5–23.8)
Drank alcohol within the last 30 days								
Junior high school total	2.7	1.00 (reference)	20.0	9.0 (7.9–10.2)	34.8	19.2 (16.5–22.4)	56.4	46.5 (35.2–61.4)
Senior high school total	7.6	1.00 (reference)	29.5	5.1 (4.7–5.4)	43.6	9.4 (8.6–10.2)	57.7	16.5 (13.8–19.8)

Note: AFNAB = alcohol-flavoured non-alcoholic beverages; CI = confidence interval.

[a]This contains 'consumed before, but not now' (not shown in this table), 'consume sometimes' and 'usually consume'.

[b]% indicates the proportion of 'at least one instance of alcohol drinking' or 'alcohol drinking within the last 30 days' among each AFNAB consumption behaviour ('never consume', 'consume at least once', 'sometimes consume' and 'usually consume').

[c]The odds ratio of 'AFNAB consumed at least once' in relation to 'never consume AFNAB' in terms of alcohol drinking behaviour (at least one instance of alcohol consumption or not, and alcohol consumption within the last 30 days or not) were calculated by each sex and school stage (junior and senior).

[d]A binominal logistic regression analysis was performed to calculate the odds ratios of AFNAB consumption frequency in relation to 'never consume AFNAB' in terms of alcohol drinking behaviour. Alcohol drinking behaviour (at least one instance of alcohol consumption or not, and alcohol consumption within the last 30 days or not) was treated as the dependent variable, and AFNAB consumption status ('never consume', 'consumed before, but not now' (not shown in this table), 'consume sometimes' or 'usually consume') was treated as an explanatory variable. The values represent the odds ratio (95% Cl) for students who responded 'never consume'.

The proportion of students for whom alcohol consumption preceded AFNAB use was similar in boys and girls, and was higher in senior high school students relative to junior high school students. The proportion of students for whom AFNAB consumption preceded alcohol use was higher in junior high school students relative to senior high school students. The percentages of students who answered 'I do not know' ranged from 16.7 to 26.0% (not shown in Table 4).

DISCUSSION

This study confirmed that more than 20% of seventh grade students in Japan have consumed AFNAB, and that the percentage of students who have consumed AFNAB is higher than the percentage of students who have consumed alcohol, with the exception of grade 11 and 12 boys. Therefore, there is the possibility that AFNAB are easier for minors to obtain. Many adults do not see AFNAB consumption as a problem (Miyaki, 2012), and AFNAB are easier to consume than alcohol.

Moreover, the percentage of students who have consumed AFNAB increased as students aged, reaching a peak in grade 10. Only in senior high school boys was the percentage of students who consumed AFNAB lower than the percentage who consumed alcohol. A previous study reported that 12th grade students had high motivation for drinking to get drunk, which cannot be achieved via AFNAB consumption (Patrick and Schulenberg, 2013). Therefore, senior high school students may be more likely to demand alcohol rather than AFNAB relative to junior high school students.

On the other hand, this study revealed that the percentage of female students who had consumed AFNAB was higher than the percentage who had consumed alcohol and was higher than in boys, respectively. One reason for the higher experimentation rates in girls may be the sweet taste of AFNAB (SUNTORY, 2015). A recent report found that the sweet taste of alcopops was preferred by girls (Metzner and Kraus, 2008). Additionally, the packaging of AFNAB is colourful and cute, and actors whom girls prefer appear in commercials to encourage girls' attraction to AFNAB. Moreover, AFNAB being good for beauty and health is emphasized in advertisements. Another reason is that the Japanese culture has typically been intolerant toward women drinking alcohol; therefore, girls might be more encouraged to consume AFNAB.

It was found that AFNAB consumption was strongly associated with alcohol use among junior and high school students. The odds ratio for alcohol use increased as the frequency of consuming AFNAB increased. The students who consumed AFNAB more often were the students who consumed alcohol more frequently. This is consistent with previous results (Kubo *et al.*, 2015). Research addressing junior high school students' drinking behaviour and girls' drinking behaviour is lacking (Osaki *et al.*, 2009; Patrick and Schulenberg, 2010) and examining the link with AFNAB may shed light on this issue.

We determined the order in which AFNAB and alcohol were first consumed among students who have experience with both AFNAB and alcohol. In all grades and sexes, alcohol use preceded AFNAB consumption for the majority of students. The percentages were not

Table 4. Proportions of participants for whom alcohol-flavoured non-alcoholic beverages consumption preceded/followed alcohol use among students who have consumed both alcohol-flavoured non-alcoholic beverages and alcohol[a]

	Drank AFNAB first		Drank alcohol first		Drank both at the same time[c]	
	Total *n*	%[b]	Total *n*	%[b]	Total *n*	%[b]
Boys						
7th	164	25.0	294	44.7	28	4.3
8th	152	20.9	364	50.1	49	6.7
9th	127	16.3	432	55.4	54	6.9
10th	284	15.7	1099	60.8	105	5.8
11th	234	14.0	1043	62.3	95	5.7
12th	203	12.5	1048	64.3	106	6.5
Junior high school total	443	20.5	1090	50.4	131	6.1
Senior high school total	721	14.1	3190	62.4	306	6.0
Girls						
7th	161	24.8	287	44.2	35	5.4
8th	162	20.4	391	49.2	54	6.8
9th	136	17.1	457	57.4	62	7.8
10th	299	15.2	1165	59.3	122	6.2
11th	272	15.1	1122	62.3	98	5.4
12th	233	12.0	1252	64.4	111	5.7
Junior high school total	459	20.5	1135	50.7	151	6.7
Senior high school total	804	14.1	3539	62.0	331	5.8

Note: AFNAB = alcohol-flavoured non-alcoholic beverages.

[a]The participants were the students who had experienced both AFNAB consumption and alcohol use.

[b]The participants who answered 'I do not know' are not shown in the table; thus, the sums of the percentages do not equal 100%.

[c]'At the same time' indicates the students who had experienced both AFNAB and alcohol.

different between boys and girls. Considering the order, it seems that AFNAB was not a trigger for using alcohol for most students. Current alcohol drinkers might consume AFNAB as an alternative to alcohol. The difference between AFNAB consumption preceding alcohol use and alcohol use preceding AFNAB consumption was smaller in junior high school students and larger in high school students. More than 20% of junior high school students consumed AFNAB first; therefore, drinking AFNAB first might be a trigger for drinking alcohol for some students, especially junior high school students. Because AFNAB are rapidly spreading beverages and the proportion of participants who consumed AFNAB before alcohol was higher in the younger high school students than in the older students in this study, there is the possibility that AFNAB could be a gateway drug in future. Further studies are required to reveal the characteristics of students who are triggered by AFNAB to consume alcohol.

This study has some limitations. First, this was a cross-sectional study; therefore, the analysis concerning drinking order only contained students who had consumed both AFNAB and alcohol, and did not contain students who will drink in the future. To reveal any real order effects, a cohort study is necessary. Second, the participants in this study were all Japanese, there are many types of AFNAB products and advertisements, and it is easy to acquire AFNAB in Japan. Therefore, the amount of AFNAB consumption in Japan might be higher than it is in other countries. However, when beverage industries expand the sale of AFNAB to other countries these results might be useful. Third, the measuring method of AFNAB consumption and alcohol use in this study was completed by self-report. It is presumed that the amount of consumption of AFNAB or alcohol was underestimated because minors drinking alcohol is illegal. To offset this limitation, teachers did not look at students' answer sheets, the questionnaires were anonymous, and they were placed in an envelope and sealed when submitted. Additionally, there was the probability of recall bias concerning the consumption of AFNAB and alcohol. A survey of 18–89-year-old individuals showed that recall bias regarding the onset age of alcohol consumption was underestimated by 0.31 years and it was smaller in younger individuals (Liang *et al.*, 2012). The recall bias concerning the consumption of AFNAB and alcohol seems small because high school students first drink AFNAB and alcohol within several years.

In conclusion, AFNAB consumption was strongly associated with alcohol use, but AFNAB consumption generally did not precede alcohol use. Our results did not illustrate that AFNAB consumption leads to increased alcohol use among adolescents in Japan. A longitudinal study is needed to reveal the actual order. In Japan, AFNAB are easier to consume than alcohol; therefore, we have to determine the influence of AFNAB on minors in order to prevent minors from drinking alcohol.

ACKNOWLEDGEMENTS

The data collection was conducted by Central Research Services, Inc. We would like to thank Editage (www.editage.jp) for English language editing.

FUNDING

This work was supported by the Health and Labour Sciences Research Grants (Comprehensive Research on Life-Style Related Diseases including Cardiovascular Diseases and Diabetes Mellitus) of the Ministry of Health, Labour and Welfare [2010–2012 Cardiovascular, etc. (Lifestyle)–Designation 020].

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

Central Research Services, Inc. (2010) *Adolescents' smoking status and the effect of increase in tobacco price in 2010. Central Investigation Report (No. 649)* [Central Research Services' website]. http://www.crs.or.jp/backno/No649/6491.htm (19 September 2015, date last accessed).

Hingson RW, Heeren T, Jamanka A, *et al.* (2000) Age of drinking onset and unintentional injury involvement after drinking. *J Am Med Assoc* **284**: 1527–33. doi:10.1001/jama.284.12.1527.

Hingson RW, Heeren T, Winter MR. (2006) Age at drinking onset and alcohol dependence: age at onset, duration, and severity. *Arch Pediatr Adolesc Med* **160**:739–46. doi:10.1001/archpedi.160.7.739.

Hingson RW, Heeren T, Winter MR, *et al.* (2003) Early age of first drunkenness as a factor in college students' unplanned and unprotected sex attributable to drinking. *Pediatrics* **111**:34–41. doi:10.1542/peds.111.1.34.

Hingson RW, Zha W. (2009) Age of drinking onset, alcohol use disorders, frequent heavy drinking, and unintentionally injuring oneself and others after drinking. *Pediatrics* **123**:1477–84.

Kandel D, Yamaguchi K. (1993) From beer to crack: developmental patterns of drug involvement. *Am J Public Health* **83**:851–55. doi:10.2105/AJPH.83.6.851.

Kaneita Y, Ohida T, Osaki Y, *et al.* (2006) Insomnia among Japanese adolescents: a nationwide representative survey. *Sleep* **29**:1543–50.

Kubo M, Nozu Y, Kataoka C, *et al.* (2015) Correlation between non-alcoholic beverage consumption and alcohol drinking behavior among Japanese youths. *Open J Prev Med* **5**:31–7. doi:10.4236/ojpm.2015.52004.

Liaison Committee About Alcohol Advertising. (2012) *Self-standard about the indication of the advertisement of alcoholic beverages and alcoholic beverages container*. http://www.rcaa.jp/standard/pdf/jishukijun.pdf (28 March 2016, date last accessed).

Liang W, Chikritzhs T, Lee A. (2012) Modest time-dependent misclassification in recalling age at first use of alcohol and tobacco. *Contemp Drug Prob* **39**:147–54. doi:10.1177/009145091203900107.

Metzner C, Kraus L. (2008) The impact of alcopops on adolescent drinking: a literature review. *Alcohol Alcohol* **43**:230–9. doi:10.1093/alcalc/agm148.

Miyaki Y. (2012) Pros and cons: consumption of non-alcoholic beverages. *Life Design Report* (28 March 2016, date last accessed) **204**:38–40 http://group.dai-ichi-life.co.jp/dlri/ldi/watching/wt1208.pdf

Morioka H, Itani O, Kaneita Y, *et al*. (2013) Associations between sleep disturbance and alcohol drinking: a large-scale epidemiological study of adolescents in Japan. *Alcohol* **47**:619–28. doi:10.1016/j.alcohol.2013.09.041.

Munezawa T, Kaneita Y, Osaki Y, *et al*. (2011) Nightmare and sleep paralysis among Japanese adolescents: a nationwide representative survey. *Sleep Med* **12**:56–64.

Nikkei Asian Review. (2015) *Suntory, Asahi tussle over nonalcoholic beer in patent suit*. http://asia.nikkei.com/Business/Trends/Suntory-Asahi-tussle-over-nonalcoholic-beer-in-patent-suit (28 March 2016, date last accessed).

Ohida T, Osaki Y, Doi Y, *et al*. (2004) An epidemiologic study of self-reported sleep problems among Japanese adolescents. *Sleep* **27**:978–85.

Osaki Y, Tanihata T, Ohida T, *et al*. (2009) Decrease in the prevalence of adolescent alcohol use and its possible causes in Japan: periodical nationwide cross-sectional surveys. *Alcohol Clin Exp Res* **33**:247–54.

Patrick ME, Schulenberg JE. (2010) Alcohol use and heavy episodic drinking prevalence and predictors among national samples of American eighth- and tenth-grade students. *J Stud Alcohol Drugs* **71**:41–5.

Patrick ME, Schulenberg JE. (2013) Prevalence and predictors of adolescent alcohol use and binge drinking in the United States. *Alcohol Res* **35**:193–200.

Patrick ME, Schulenberg JE, Martz ME, *et al*. (2013) Extreme binge drinking among 12th-grade students in the United States: prevalence and predictors. *JAMA Pediatr* **167**:1019–25. doi:10.1001/jamapediatrics.2013.2392.

SUNTORY. (2015) *Factual report on nonalcoholic beverages*. http://www.suntory.co.jp/news/2015/12378.html (28 March 2016, date last accessed).

Suzuki K, Minowa M, Osaki Y. (2000) Japanese national survey of adolescent drinking behavior in 1996. *Alcohol Clin Exp Res* **24**:377–81.

Sleep Medicine 37 (2017) 38–45

Contents lists available at ScienceDirect

Sleep Medicine

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

Original Article

Relationship between stress coping and sleep disorders among the general Japanese population: a nationwide representative survey

Yuichiro Otsuka [a], Yoshitaka Kaneita [b,*], Osamu Itani [b], Sachi Nakagome [a], Maki Jike [a], Takashi Ohida [a]

[a] Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, Tokyo, Japan
[b] Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Oita, Japan

ARTICLE INFO

Article history:
Received 21 January 2017
Received in revised form
18 May 2017
Accepted 1 June 2017
Available online 24 June 2017

Keywords:
Stress
Stress coping
Sleep disorder
Epidemiology
Prevalence
Japanese

ABSTRACT

Objective: To clarify the prevalence of stress, and examine the relationship between sleep disorders and stress coping strategies among highly stressed individuals in the general Japanese population.
Methods: A cross-sectional nationwide survey was undertaken in November 2007. Men and women were randomly selected from 300 districts throughout Japan. Data from 7671 (3532 men (average age 53.5 ± 17.0 years) and 4139 women (average age 53.9 ± 17.7 years)) were analyzed. Participants completed a self-reported questionnaire on stress, sleep disorders, and stress coping strategies in the previous month.
Results: Highly stressed individuals comprised 16.6% (95% confidence interval 15.8–17.5%) of the total sample, and most were aged 20–49 years. In multiple logistic regression, symptoms of insomnia (ie, difficulty initiating sleep, difficulty maintaining sleep, and early morning awakening), excessive daytime sleepiness, nightmares, daytime malfunction, and lack of rest due to sleep deprivation were more prone to occur in highly stressed individuals. In addition, logistic regression analysis controlling for other adjustment factors revealed that stress coping strategies such as 'giving up on problem-solving', 'enduring problems patiently', 'smoking' and 'drinking alcohol' were positively associated with the above-mentioned sleep disorders. On the other hand, stress coping strategies such as 'exercising', 'enjoying hobbies', and 'sharing worries' were inversely associated with the above-mentioned sleep disorders.
Conclusions: Distraction-based stress coping (eg, hobbies, exercise, and optimistic thinking) was found to be preferable to problem-based stress coping in a highly stressed Japanese general population.

1. Introduction

Sleep disorders are important conditions in modern society because they have been found to increase the risk of developing various diseases such as obesity, impaired glucose tolerance, cardiovascular diseases, and metabolic syndrome [1–4]. A meta-analysis reported an association between sleep and total deaths [5]. Sleep has even been linked to an increase in mistakes during exertion and a greater risk of accidents [6]. Moreover, symptoms of insomnia are often seen in depression [7], while insomnia itself is known as a potential risk factor for depression [8,9].

Like sleep disorders, stress is also considered an important social issue in modern society. Chronic exposure to psychological stress has been associated with multi-system physiological dysregulation [10,11]. For example, several meta-analyses have reported an association between mental stress and various diseases such as coronary artery disease [12,13], metabolic syndrome [14], and diabetes [15]. Stress is also widely known to induce sleep disorders [16–18]. For example, when subjective stress is high, people sleep about 10% less and feel the lack of sleep more than they do when they are not stressed [19]. Furthermore, life events such as divorce [20] or the death of a spouse [21] are stressors that have been linked to potential insomnia. In addition, stressful times can have negative effects on adolescents' sleep fragmentation [22].

In very stressful situations, implementing an appropriate method of coping with stress (stress coping) can prevent stress-related diseases [23]. Lazarus and Folkman [23] defined stress

* Corresponding author. Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Hasama-machi, Yufu-shi, Oita 879-5593, Japan. Fax: +81 97 586 5739.
E-mail address: nusmpublichealth@gmail.com (Y. Kaneita).

http://dx.doi.org/10.1016/j.sleep.2017.06.007
1389-9457/

coping as a person's constantly changing cognitive and behavioral efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the person's resources. In general, stress coping that aggressively resolves problems by facing health and behavioral problems is effective in stressful situations that can be controlled under one's own power. On the other hand, stress coping through diversion rather than directly solving the problem itself is known to be more effective at reducing stress in situations that are impossible to control [24]. Workers who use problem-solving-based or diversion-based stress coping have better mental health, and several studies have reported an association between avoidance-based stress coping and mental disorders [25,26]. One meta-analysis reported that illness perceptions and coping played an important role in the explanation of distress outcomes [27]. Therefore, the continuous accumulation of epidemiological knowledge on stress coping is essential for examining the problem of mental health disorders in modern society.

To date, there have been few studies on sleep disorders and stress coping [28–30]. Among these limited previous studies is one that reports that people who pessimistically implement stress coping have poorer quality of sleep than those who implement positive stress coping [28]. In addition, some maladaptive coping strategies are mechanisms by which stress exposure precipitates insomnia [31]. Furthermore, Abe et al. [29] reported that stress coping to resolve problems is associated with an improvement in sleep disorders. It is believed that the present study is the only one to examine this topic in Japanese men and women [29].

There are many reports regarding the relationship between sleep disorders and stress, and the influence of stress on sleep is known; however, it is not sufficiently clear which specific sleep disorders are strongly associated with high stress. Moreover, no previous studies could be found that examined which stress coping strategies contribute to the alleviation or exacerbation of sleep disorders in highly stressed people.

Therefore, the aims of the present study were to: 1) collect epidemiological data on the nature of the relationship between stressful situations and sleep disorders, and 2) examine which stress coping strategies alleviate or exacerbate sleep disorders in highly stressed people if a relationship between stress and sleep disorders does indeed exist. This study could help identify methods of preventing the exacerbation of sleep disorders, even in situations where stress is difficult to alleviate.

2. Methods

2.1. Data source

The present study was performed using data collected by the National Health and Nutrition Survey (NHNS) conducted in November 2007 by the Japanese Ministry of Health, Labour and Welfare. The NHNS is a nationally representative sample, cross-sectional survey of the general Japanese population. The study comprised approximately 18,000 members aged ≥1 year from 300 survey district units randomly selected from the national census unit districts. The study had three components: physical examination of individuals; dietary assessment of households and individuals; and a questionnaire on health-related behaviors, habits, and lifestyle factors. The examination protocol and data collection design were completely documented in the NHNS annual reports [32,33]. Quality assurance in data processing has been described elsewhere [34]. All survey respondents provided verbal informed consent to participate. This study conformed to the ethics guidelines of the Ministry of Health, Labour and Welfare, and the Ministry approved this study.

2.2. Participants

Of the 9611 men and women enrolled in NHNS 2007, those aged ≥20 years, excluding pregnant or lactating women, with available data on lifestyle factors were selected as study participants.

2.3. Measures and definitions

The main questions of this study pertained to the following seven issues: (1) basic attributes (sex and age), (2) bedtime habits (use of hypnotics), (3) sleep situation (sleep duration, sufficiency level of rest obtained through sleep, difficulty initiating sleep (DIS), difficulty maintaining sleep (DMS), early morning awakening (EMA) and excessive daytime sleepiness (EDS)); (4) smoking habit, (5) alcohol drinking habit, (6) stress, and (7) stress coping strategies.

The question for bedtime habits was: Did you take a sedative or sleeping aid at bedtime in the previous month? (1. Never, 2. Seldom, 3. Sometimes, 4. Often, 5. Always). 'Often', and 'Always' were taken as affirmative answers to the question.

The six questions for sleep situation over the previous month were as follows:

1. How many hours did you sleep at night on average? (1. <5, 2. ≥5 but <6, 3. ≥6 but <7, 4. ≥7 but <8, 5. ≥8).
2. How would you rate your amount of sleep? (1. Very sufficient, 2. Sufficient, 3. Insufficient, 4. Very insufficient). 'Insufficient' and 'Very insufficient' were taken to indicate lack of rest due to sleep deprivation [35].
3. Did you have difficulty falling asleep at night? (1. Never, 2. Seldom, 3. Sometimes, 4. Often, 5. Always). 'Often' and 'Always' were taken to indicate DIS [36,37].
4. Did you wake up during the night after you had gone to sleep? (1. Never, 2. Seldom, 3. Sometimes, 4. Often, 5. Always). 'Often' and 'Always' were taken to indicate DMS [36,37].
5. Did you wake up too early in the morning and have difficulty getting back to sleep? (1. Never, 2. Seldom, 3. Sometimes, 4. Often, 5. Always). 'Often' and 'Always' were taken to indicate EMA [36,37].
6. Did you feel excessively sleepy during the daytime? (1. Never, 2. Seldom, 3. Sometimes, 4. Often, 5. Always). 'Often' and 'Always' were taken to indicate EDS [36,37].

Insomnia was defined as being present when an affirmative answer was obtained for any of DIS, DMS or EMA [34,38].

The questions for smoking and alcohol habits were as described below:

Have you smoked in the previous month? (1. Every day, 2. Sometimes, 3. No). 'Every day' and 'Sometimes' were taken as affirmative answers to the question.

How many days did you consume alcoholic beverages in the previous month? (1. Every day, 2. 5–6 days/week, 3. 3–4 days/week, 4. 1–2 days/week, 5. 1–3 days/month, 6. None). 'Every day', '5–6 days/per week' and '3–4 days/per week' were taken as affirmative answers to the question.

The question related to stress was: Did you feel dissatisfied or distressed, or have difficulties and stress in the past month? (1. Frequently, 2. Sometimes, 3. Rarely, 4. Never). 'Frequently' was taken to indicate high stress in accordance with previous surveys [29,34]. For stress coping strategies, men and women were asked to choose how they cope with stress from the following list (multiple answers permitted):

1. Aggressively solving problems, 2. Being physically active and performing exercise, 3. Enjoying hobbies or relaxing alone,

4. Watching television or listening to the radio, 5. Sharing worries and concerns with family and friends, 6. Giving up on problem-solving, 7. Enduring problems patiently, 8. Thinking positively and working through problems, 9. Looking for stimulation and excitement, 10. Drinking alcoholic beverages, 11. Smoking, 12. Eating, and 13. Nothing'.

2.4. Statistical analysis

First, the percentage of highly stressed individuals and the 95% confidence interval (CI) were calculated. Second, the associations between sleep disorders during the previous month and stress were examined by multiple logistic regression analyses. Third, the percentage of stress coping strategies by age group among highly stressed individuals was calculated using the Chi-squared test. Finally, for highly stressed individuals, multiple logistic regression analyses were performed to examine the stress coping strategies associated with each sleep disorder during the previous month. The study adjusted for confounding variables by using each of the basic attributes (sex and age), use of hypnotics, alcohol drinking, smoking, and stress as co-variables. The level of significance was set at $p < 0.05$. All analyses were performed using SPSS version 17.0 for Windows (SPSS, Inc., Chicago, IL).

3. Results

3.1. Prevalence of highly stressed individuals among the general Japanese population

Table 1 shows the prevalence of highly stressed individuals among the general Japanese population. The sample was limited to adults aged ≥20 years, and further excluded those with any missing data for the variables included in the analysis. The final sample for analysis comprised 7671 adults: 3532 men (average age 53.5 ± 17.0 years) and 4139 women (average age 53.9 ± 17.7 years), with a mean age of 53.6 ± 17.4 years (age range 20–99). Among both men and women, most highly stressed individuals were aged 20–49. Individuals aged ≥50 were found to be less prone to stress.

Table 1
Prevalence of highly stressed individuals in the general Japanese population.

Population	*N*	Highly stressed individuals (*n*)	Percentage	95% CI		*p*
Total						
Age group (years)						
20–29	720	156	21.7	18.7	24.7	<0.001
30–39	1262	292	23.1	20.8	25.5	
40–49	1200	276	23.0	20.6	25.4	
50–59	1399	263	18.8	16.7	20.8	
60–69	1494	165	11.0	9.5	12.6	
≥70	1596	125	7.8	6.5	9.2	
Total	7671	1277	16.6	15.8	17.5	
Men						
Age group (years)						
20–29	324	68	21.0	16.5	25.4	<0.001
30–39	570	137	24.0	20.5	27.6	
40–49	575	132	23.0	19.5	26.4	
50–59	655	109	16.6	13.8	19.5	
60–69	698	64	9.2	7.0	11.3	
≥70	710	49	6.9	5.0	8.8	
Total	3532	559	15.8	14.6	17.0	
Women						
Age group (years)						
20–29	396	88	22.2	18.1	26.3	<0.001
30–39	692	155	22.4	19.3	25.5	
40–49	625	144	23.0	19.7	26.4	
50–59	744	154	20.7	17.8	23.6	
60–69	796	101	12.7	10.4	15.0	
≥70	886	76	8.6	6.7	10.4	
Total	4139	718	17.3	16.2	18.5	

p-value was calculated by the Chi-squared test.
Subjects with missing data were excluded from the analysis.
CI: confidence interval.

3.2. Association between each sleep disorder and high stress among the general Japanese population

Table 2 shows the association between each sleep disorder and high stress among the general Japanese population. Multiple logistic regression showed that all sleep disorders were significantly positively associated with high stress.

3.3. Percentage of stress coping strategies among highly stressed individuals by age group

Table 3 shows the percentage of stress coping strategies among highly stressed individuals by age group. Significant associations were observed for almost all stress coping strategies except for 'aggressively solving problems', 'being physically active and performing exercise', 'enduring problems patiently' and 'nothing'. The frequency of using many of the stress coping strategies tended to be lower as age increased. However, 'being physically active and performing exercise', 'watching television or listening to the radio', 'sharing worries and concerns with family and friends' and 'enduring problems patiently' were highly frequent stress coping strategies among respondents aged ≥70 years. 'Enjoying hobbies or relaxing alone', 'watching television or listening to the radio', 'sharing worries and concerns with family and friends', 'enduring problems patiently', and 'thinking positively and working through problems' were practiced by >30% of the respondents.

3.4. Association between sleep disorders and stress coping strategies among highly stressed individuals

Highly stressed individuals were only selected, then the association between sleep disorders and stress coping strategies was analyzed using multiple logistic regression. The results are summarized in Tables 4 and 5. Table 4 shows the association between insomnia, EDS, and stress coping strategies among highly stressed individuals. Multiple logistic regression analysis with insomnia (ie, DIS, DMS, and EMA) and EDS during the previous month as the dependent variable showed significant associations with sex and age group. Multivariate logistic analysis showed the two stress coping strategies of 'being physically active and performing exercise' and 'thinking positively and working through problems' had a significant negative relationship with DIS. Conversely, stress coping strategies with a significant positive relationship with DIS included 'enduring problems patiently', 'looking for stimulation and excitement', and 'smoking'. None of the stress coping strategies had a significant negative relationship with DMS. On the other hand, the three stress coping strategies of 'enduring problems patiently', 'giving up on problem-solving', and 'drinking alcoholic beverages' had a significant positive relationship with DMS. None of the stress coping strategies had a significant positive relationship with EMA. In contrast, 'being physically active and performing exercise' and 'drinking alcoholic beverages' had a significant negative relationship with DMS. None of the stress coping strategies had a significant negative relationship with EDS. Only 'smoking' had a significant positive relationship with EDS.

Table 5 shows the associations between nightmares, daytime malfunction, and lack of rest due to sleep deprivation and stress coping strategies among highly stressed individuals. Multiple logistic regression analysis with each of these sleep disorders during the previous month as the dependent variable showed

Table 2
Association between each sleep disorder and high stress in the general Japanese population.

Sleep disorder	DIS	DMS	EMA	EDS	Nightmares	Daytime malfunction	Lack of rest due to sleep deprivation
Crude OR	3.01	1.99	1.44	3.04	3.45	5.57	4.70
95% CI	2.59–3.51	1.73–2.30	1.25–1.66	2.29–4.03	2.63–4.53	4.48–6.92	4.14–5.34
Adjusted OR	2.58	2.10	1.59	2.46	3.49	3.67	3.31
95% CI	2.17–3.06	1.80–2.46	1.36–1.86	1.81–3.35	2.59–4.71	2.90–4.64	2.87–3.83

The presence of EDS was defined as an Epworth Sleepiness Scale score of ≥P11.
Subjects with missing data were excluded from the analysis.
Adjusted for sex, age, sleep duration, use of hypnotics, alcohol drinking, smoking and stress by multiple logistic regression.
p-value was calculated by multiple logistic regression analysis.
AOR: adjusted odds ratio; CI: confidence interval; DIS: difficulty initiating sleep; DMS: difficulty maintaining sleep; EMA: early morning awakening; EDS: excessive daytime sleepiness.

Table 3
Stress coping strategy use among highly stressed individuals by age group.

Variable								
Age group (years)	20–29	30–39	40–49	50–59	60–69	≥70	Total	*p*
n	156	292	276	263	165	125	1277	
Stress coping strategy								
Aggressively solving problems	20.5	19.9	22.5	24.0	22.4	14.4	21.1	0.366
Being physically active and performing exercise	10.9	15.8	15.6	9.9	14.5	16.0	13.8	0.239
Enjoying hobbies or relaxing alone	52.6	31.2	34.4	27.0	27.3	18.4	31.9	<0.001
Watching television or listening to the radio	33.3	27.7	37.7	34.6	36.4	43.2	34.6	0.040
Sharing worries and concerns with family and friends	46.8	39.7	42.4	42.2	29.1	28.8	39.2	0.002
Giving up on problem-solving	12.8	8.6	5.1	4.9	8.5	7.2	7.4	0.035
Enduring problems patiently	39.7	40.4	44.9	35.7	41.8	36.0	40.1	0.321
Thinking positively and working through problems	34.6	35.6	40.9	36.9	37.6	19.2	35.6	0.002
Looking for stimulation and excitement	4.5	3.4	1.1	1.1	1.2	0.0	2.0	0.019
Drinking alcoholic beverages	16.7	21.6	28.3	20.9	12.1	6.4	19.6	<0.001
Smoking	23.1	24.3	19.6	16.0	9.7	7.2	17.9	<0.001
Eating	25.6	23.6	21.4	16.3	9.1	6.4	18.3	<0.001
Nothing	2.6	4.8	1.4	2.3	0.6	2.4	2.5	0.075

p-value was calculated by the Chi-squared test.

Table 4
Association between insomnia, EDS, and stress coping strategies among highly stressed individuals.

Stress coping strategy	DIS				DMS				EMA				EDS			
	AOR	95% CI		*p*	AOR	95% CI		*p*	AOR	95% CI		*p*	AOR	95% CI		*p*
Aggressively solving problems	0.95	0.67	1.34	0.777	1.01	0.72	1.42	0.947	1.11	0.79	1.56	0.550	0.96	0.51	1.80	0.901
Being physically active and performing exercise	0.58	0.37	0.90	0.016	0.83	0.55	1.24	0.361	0.64	0.44	0.93	0.018	1.03	0.50	2.11	0.944
Enjoying hobbies or relaxing alone	0.79	0.58	1.07	0.133	0.93	0.69	1.26	0.638	1.05	0.77	1.42	0.775	1.01	0.59	1.72	0.986
Watching television or listening to the radio	0.88	0.66	1.18	0.391	0.81	0.61	1.07	0.142	1.27	0.95	1.70	0.101	0.78	0.46	1.32	0.363
Sharing worries and concerns with family and friends	0.91	0.68	1.23	0.543	0.96	0.71	1.28	0.765	1.03	0.76	1.39	0.840	0.69	0.41	1.17	0.170
Giving up on problem-solving	1.14	0.70	1.86	0.594	2.01	1.25	3.24	0.004	0.94	0.56	1.57	0.804	1.90	0.89	4.04	0.096
Enduring problems patiently	1.44	1.10	1.89	0.008	1.35	1.03	1.77	0.030	0.79	0.60	1.04	0.098	0.77	0.47	1.27	0.308
Thinking positively and working through problems	0.73	0.54	0.97	0.031	0.80	0.60	1.07	0.128	1.07	0.80	1.42	0.664	0.68	0.40	1.17	0.165
Looking for stimulation and excitement	2.81	1.20	6.60	0.018	0.28	0.06	1.25	0.094	1.03	0.36	2.97	0.955	1.65	0.45	6.08	0.455
Drinking alcoholic beverages	0.73	0.50	1.05	0.089	1.66	1.18	2.33	0.004	0.70	0.50	0.99	0.043	0.98	0.53	1.80	0.943
Smoking	1.53	1.07	2.17	0.018	0.71	0.48	1.04	0.079	1.31	0.89	1.94	0.173	2.70	1.54	4.74	0.001
Eating	0.98	0.68	1.40	0.902	1.00	0.70	1.43	1.000	1.08	0.74	1.56	0.701	1.67	0.96	2.92	0.070
Nothing	0.53	0.19	1.44	0.210	1.97	0.89	4.40	0.096	1.54	0.55	4.29	0.410	1.59	0.43	5.90	0.487

Subjects with missing data were excluded from the analysis.
Adjusted for sex and age by multiple logistic regression analysis.
p-value was calculated by multiple logistic regression analysis.
AOR: adjusted odds ratio; CI: confidence interval; DIS: difficulty initiating sleep; DMS: difficulty maintaining sleep; EDS: excessive daytime sleepiness; EMA: early morning awakening.

significant associations with sex and age group. 'Sharing worries and concerns with family and friends' had a significant negative relationship with nightmares. Conversely, 'drinking alcoholic beverages' had a significant positive relationship with nightmares. The four stress coping strategies of 'being physically active and performing exercise', 'enjoying hobbies or relaxing alone', 'sharing worries and concerns with family and friends', and 'thinking positively and working through problems' had a significant negative relationship with daytime malfunction. 'Enduring problems patiently' had a significant positive relationship with daytime malfunction. 'Enjoying hobbies or relaxing alone' and 'watching television or listening to the radio' had a significant negative relationship with sleep deprivation; enduring problems patiently had a significant positive relationship with lack of rest due to sleep deprivation.

4. Discussion

This study used data from a survey with a high level of national representation to examine the relationship between stress coping

Table 5
Association between nightmares, daytime malfunction and lack of rest due to sleep deprivation and stress coping strategies among highly stressed individuals.

Sleep disorder												
Stress coping strategy	Nightmares				Daytime malfunction				Lack of rest due to sleep deprivation			
	AOR	95% CI		*p*	AOR	95% CI		*p*	AOR	95% CI		*p*
Aggressively solving problems	0.88	0.48	1.61	0.675	0.83	0.52	1.32	0.434	0.88	0.66	1.18	0.385
Being physically active and performing exercise	0.55	0.24	1.23	0.144	0.29	0.13	0.63	0.002	0.88	0.62	1.24	0.464
Enjoying hobbies or relaxing alone	0.61	0.35	1.05	0.075	0.60	0.40	0.90	0.014	0.64	0.49	0.83	0.001
Watching television or listening to the radio	1.03	0.64	1.66	0.894	0.87	0.60	1.25	0.437	0.72	0.56	0.92	0.010
Sharing worries and concerns with family and friends	0.53	0.31	0.91	0.020	0.62	0.43	0.90	0.012	0.94	0.73	1.21	0.624
Giving up on problem-solving	1.28	0.61	2.68	0.517	0.63	0.32	1.25	0.188	0.97	0.62	1.52	0.893
Enduring problems patiently	1.14	0.72	1.79	0.578	1.84	1.31	2.59	<0.001	1.31	1.03	1.66	0.029
Thinking positively and working through problems	0.79	0.48	1.29	0.342	0.62	0.43	0.91	0.014	0.82	0.64	1.05	0.120
Looking for stimulation and excitement	2.37	0.72	7.83	0.157	1.76	0.55	5.63	0.339	2.19	0.93	5.19	0.074
Drinking alcoholic beverages	2.31	1.37	3.87	0.002	0.95	0.61	1.47	0.813	0.87	0.64	1.17	0.345
Smoking	1.01	0.56	1.81	0.978	1.33	0.86	2.05	0.204	1.06	0.77	1.46	0.710
Eating	0.78	0.42	1.48	0.451	0.94	0.61	1.44	0.760	1.15	0.85	1.56	0.359
Nothing	0.73	0.16	3.34	0.682	0.74	0.26	2.08	0.564	1.55	0.70	3.41	0.278

Subjects with missing data were excluded from the analysis.
Adjusted for sex and age by multiple logistic regression analysis.
p-value was calculated by multiple logistic regression analysis.
AOR: adjusted odds ratio; CI: confidence interval.

strategies and sleep disorders among highly stressed people. The results showed that: (1) highly stressed individuals may be more prone to have symptoms of insomnia, nightmares, daytime malfunction, and lack of rest due to sleep deprivation, and (2) appropriate stress coping strategies can improve sleep disorders even in the most highly stressed individuals. Conversely, this study suggested that certain stress coping strategies could also exacerbate sleep disorders. To date, there have been very few reports of epidemiological studies examining the association between stress coping strategies and sleep disorders in the general Japanese population. The results of the present study could serve as an important reference in devising countermeasures for stress and approaches for good sleep hygiene habits.

In a 2008 Cabinet Office survey of stress in the general Japanese population [39], approximately 50% of the total respondents answered that they felt stressed; 14.9%, many of which were in their 20s–50s, answered that they felt highly stressed. The present study found a similar prevalence of highly stressed individuals in the general Japanese population at 16.6%. This study showed that developing high stress due to dissatisfaction, worries, hardships and other factors makes people more conscious of symptoms of insomnia, EDS, nightmares, daytime malfunction, and lack of rest due to sleep deprivation. Previous surveys of the general Japanese public also reported a facilitatory association between mental stress and symptoms of insomnia or EDS [40,41]. Furthermore, a survey of Japanese workers found that occupational stress is significantly associated with symptoms of insomnia and poor sleep quality [17,42]. Many epidemiological studies conducted in other countries have also demonstrated a strong correlation between stress and insomnia [43–45]. Morin et al. wrote that stress is the most significant cause of insomnia [46]. Stress is therefore a very important factor that hinders sleep in humans.

The present study showed that those who exercised, enjoyed hobbies, watched television or listened to radio, shared worries and concerns with family or friends, and thought positively and worked through problems as strategies to cope with stress had a significantly lower odds ratio of sleep disorders than those who practiced none of these strategies. These strategies were therefore considered means of coping with stress that are beneficial to sleep. Conversely, those who patiently endured problems, looked for stimulation and excitement, and smoked as strategies to cope with stress had a significantly higher odds ratio of sleep disorders than those who did not practice these strategies, suggesting that they were means of coping with stress that are disadvantageous to sleep. To date, very few studies have investigated the relationship between sleep disorders and stress coping strategies. Among this small number of studies is one by Kales et al., who reported that insomniacs were less satisfied with their interpersonal relationships and less able to engage in proper stress coping than the control group [47]. Other reports have described how insomniacs are more likely to choose emotion-focused stress coping strategies [46,48]. Yet another report described how patients with breast cancer and prostate cancer who implement avoidance-based stress coping also had poor sleep quality [49]. People with hypertension and insomnia symptoms showed less effective coping strategies than non-insomniacs [50]. The samples in these previous studies had poor representation, were performed in people with specific diseases, and thus had many flaws that prevented the results from being generalizable. However, emotion-focused stress coping, especially avoidance-based stress coping, was generally assumed to be bad for sleep. The present study also categorized 'enduring problems patiently' and 'looking for stimulation and excitement' as avoidance-based or emotion-focused stress coping, which was consistent with previous studies. The results of the present study have strong academic value because a sample with high national representation was used. Many reports correlated stress coping by means of suppressing emotions or avoidance with poor mental health [51–54], and these means of stress coping themselves may potentially increase stress.

In the present study, 'thinking positively and working through problems' emerged as a good stress coping strategy. Thinking positively is considered to improve sleep disorders by developing the ability to explore new means of adaptation through re-examination of cognitive process, suppressing feelings of anxiety, and seeking diversion. In their study, Abe et al. [29] suggested that stress coping by 'taking it easy' is a good coping strategy for insomniacs with depression, which was supported by the present results.

This research also suggested that exercising, as a means of stress coping, is beneficial to DIS, EMA, and daytime malfunction. An epidemiological study in the general Japanese public identified a relationship between a lack of routine exercise and insomnia [40]. Furthermore, a longitudinal study in healthy elderly Japanese people found that physical activity at least five days a week inhibits the onset of insomnia [55]. Another study reported that using

exercise as a stress coping strategy could potentially reduce overall stress levels [56,57], and that the mechanism behind exercise could also improve sleep disorders. Exercising as a means of stress coping therefore has a positive effect on sleep disorders. However, the present study did not investigate intensity, frequency or duration of exercise. A report has also suggested that intensive exercise before going to bed prevents the onset of sleep [58]. An epidemiological study taking these matters into consideration is needed to clarify these issues going forward.

The results of the present study indicated that 'enjoying hobbies or relaxing alone' are beneficial stress coping strategies for daytime malfunction and lack of rest, 'watching television or listening to the radio' are beneficial for daytime malfunction, and 'sharing worries and concerns with family and friends' is beneficial for nightmares and daytime malfunction. Morin et al. [59] reported that instead of stress coping, cognitive behavioral therapy, such as listening to music before going to bed and relaxing by reading a book, improves insomnia, while Abe et al. [29] reported that Japanese insomniacs who watched television or listened to the radio experienced exacerbated insomnia. The results of the present study revealed no significant relationship between coping with stress by watching television or listening to the radio and symptoms of insomnia. Instead, the results of the present study suggested that coping with stress by watching television and listening to the radio has minimal effect on stress-induced symptoms of insomnia. Sharing worries and concerns with family and friends has a positive effect on mental health, as this allows a person to reveal emotions and organize feelings. Moreover, this behavior has been reported to prevent the onset of stress-related diseases [51]. Improving sleep disorders by using stress coping strategies imparts mental relaxation and reduces stress, which could allow for rest during sleep and lower sensitivity to feeling daytime malfunction.

The present study indicated that smoking, as a means of coping with stress, was disadvantageous for DIS and EDS. Previous studies have reported a close relationship between smoking and stress [60–66]. Smoking has also been reported to alleviate stress [62–64,66]; however, other reports have claimed that smoking has harmful effects on stress [66–68], indicating a lack of consensus. Among smokers, those who do not practice any other particular stress coping strategy besides smoking are known to be especially prone to feeling the urge to smoke in stressful situations [69]. People with a lot of stress and fewer other stress coping strategies have a greater tendency to smoke and may be more susceptible to exacerbation of sleep. While much remains unclear about the association between smoking and sleep, the results of the present study suggest that quitting smoking is important to improving sleep disorders in highly stressed people.

Drinking alcohol was also found to be a stress coping strategy that was disadvantageous to DMS and nightmares. Drinking alcohol has the unique ability to both relieve stress and be the cause of it [70,71]. Drinking immediately after experiencing a stress-inducing event is reported to suppress cortisol and markedly prolong feelings of tension [72]. Furthermore, people who use drinking as a means of coping with stress often use avoidance-based coping strategies or practice few coping strategies [73]. Drinking as a means of coping with stress is therefore considered to have a negligible effect on improving stress. Stress has also been cited as a cause of nightmares [74]. When stress coping strategies fail to work, stress accumulates, which may explain the results of the present study.

This study revealed no clear association between aggressive problem-solving and sleep disorders. The study by Abe et al. [29], which examined stress coping strategies in Japanese people with insomnia, revealed that only problem-solving-based coping was reported to improve insomnia. Folkman and Lazarus [24] reported that the stress response is reduced when problem-focused coping is used when coping is deemed possible, whereas emotion-focused coping is used when coping is deemed impossible. It is therefore likely that the use of emotion-focused coping reduces stress and helps deal with sleep disorders in highly stressed people, such as those in the present study, who are already in situations of inevitable stress where solutions to the problem cannot be actively sought.

The present study had several limitations. First, this was a cross-sectional study in which the relationship between sleep disorders, stress and stress coping strategies could not be determined in terms of trends over time and causal relationships. A future longitudinal study is needed to examine how sleep disorders change and how stress coping strategies are learned in highly stressed people. Secondly, this study used self-administered questionnaires. Measurements of objective indices, such as polysomnography to examine sleep and measurement of salivary cortisol to examine stress, may provide stronger evidence to support the results of this study. Thirdly, no data on work, educational background or income of the participants could be obtained in this study. Given that these factors have a high likelihood of influencing stress coping, they should be examined in more detail in future studies. Fourthly, stress coping itself might directly influence sleep, independently of stress. This study only asked about coping to alleviate stress, and not about coping to improve sleep; however, while minimal, there remains a possibility that coping to improve sleep might directly influence sleep.

Despite the above limitations, one of this study's strengths was that it involved a national survey performed by the Ministry of Health, Labour and Welfare that focused on highly stressed people in the general Japanese population. This study did not evaluate stress coping strategies by category, but rather examined the individual effects of such strategies on sleep disorders. Very few prior studies have conducted such an analysis. Stress is strongly associated with sleep disorders and this study demonstrated that certain stress coping strategies might improve or exacerbate sleep disorders even in highly stressed people. Stress has a wide range of causes and while it is important to alleviate stress itself, it is not easy. The stress coping strategies presented in this study are relatively easy to implement in daily life, and while they may not completely eliminate stress, they could be of great significance if they were to somewhat improve sleep. To date, few epidemiological studies have examined the relationship between stress coping strategies and sleep disorders, leaving a demand for future studies that expand on this topic. Many treatments for sleep disorders use drug therapy as their main component, while few reports have described treatments that rectify stress. It is essential for doctors who treat sleep disorders to increase patient awareness of the strong association between stress and sleep disorders, as well as their understanding of favorable stress coping during routine visits.

5. Conclusion

This study revealed that the prevalence of stressed individuals in the general Japanese population is as high as 16.6%. Some stress coping strategies used by highly stressed individuals, such as 'enduring problems patiently', and 'smoking', may act as precipitating or perpetuating factors for sleep disorders. On the other hand, it was found that several acceptable stress coping strategies were practiced by participants in this study, such as 'being physically active and performing exercise' and 'enjoying hobbies or relaxing'. These findings may offer critical insights for developing effective sleep education and preventative programs for sleep disorders.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

The authors report no other financial affiliation or relationship relevant to the subject of the article.

Conflict of interest

None.

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: http://dx.doi.org/10.1016/j.sleep.2017.06.007.

References

[1] Itani O, Kaneita Y, Murata A, et al. Association of onset of obesity with sleep duration and shift work among Japanese adults. Sleep Med 2011;12:341–5.
[2] Beihl DA, Liese AD, Haffner SM. Sleep duration as a risk factor for incident type 2 diabetes in a multiethnic cohort. Ann Epidemiol 2009;19:351–7.
[3] Amagai Y, Ishikawa S, Gotoh T, et al. Sleep duration and incidence of cardiovascular events in a Japanese population: the Jichi Medical School cohort study. J Epidemiol 2010;20:106–10.
[4] Troxel WM, Buysse DJ, Matthews KA, et al. Sleep symptoms predict the development of the metabolic syndrome. Sleep 2010;33:1633–40.
[5] Cappuccio FP, D'Elia L, Strazzullo P, et al. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep 2010;33:585–92.
[6] Katz DA, McHorney CA. The relationship between insomnia and health-related quality of life in patients with chronic illness. J Fam Pract 2002;51:229–35.
[7] Kaneita Y, Ohida T, Uchiyama M, et al. The relationship between depression and sleep disturbances: a Japanese nationwide general population survey. J Clin Psychiatry 2006;67:196–203.
[8] Chang PP, Ford DE, Mead LA, et al. Insomnia in young men and subsequent depression the Johns Hopkins Precursors study. Am J Epidemiol 1997;146:105–14.
[9] Li L, Wu C, Gan Y, et al. Insomnia and the risk of depression: a meta-analysis of prospective cohort studies. BMC Psychiatry 2016;16:375.
[10] Gallo LC, Jimenez JA, Shivpuri S, et al. Domains of chronic stress, lifestyle factors, and allostatic load in middle-aged Mexican-American women. Ann Behav Med 2011;41:21–31.
[11] Glei DA, Goldman N, Chuang YL, et al. Do chronic stressors lead to physiological dysregulation? Testing the theory of allostatic load. Psychosom Med 2007;69:769–76.
[12] Wei J, Rooks C, Ramadan R, et al. Meta-analysis of mental stress-induced Myocardial ischemia and subsequent cardiac events in patients with coronary artery disease. Am J Cardiol 2014;114:187–92.
[13] Backé E-M, Seidler A, Latza U, et al. The role of psychosocial stress at work for the development of cardiovascular diseases: a systematic review. Int Arch Occup Environ Health 2012;85:67–79.
[14] Bergmann N, Gyntelberg F, Faber J. The appraisal of chronic stress and the development of the metabolic syndrome: a systematic review of prospective cohort studies. Endocr Connect 2014;3:R55–80.
[15] Cosgrove MP, Sargeant LA, Caleyachetty R, et al. Work-related stress and Type 2 diabetes: systematic review and meta-analysis. Occup Med 2012;62:167–73.
[16] Deguchi Y, Iwasaki S, Ishimoto H, et al. Relationships between temperaments, occupational stress, and insomnia among Japanese workers. PLoS One 2017;12:e0175346.
[17] Doi Y, Minowa M, Tango T. Impact and correlates of poor sleep quality in Japanese white-collar employees. Sleep 2003;26:467–71.
[18] Tachibana H, Izumi T, Honda S, et al. The prevalence and pattern of insomnia in Japanese industrial workers: relationship between psychosocial stress and type of insomnia. Psychiatry Clin Neurosci 1998;52:397–402.
[19] Hicks RA, Garcia ER. Level of stress and sleep duration. Percept Mot Ski 1987;64:44–6.
[20] Cartwright RD, Wood E. Adjustment disorders of sleep: the sleep effects of a major stressful event and its resolution. Psychiatry Res 1991;39:199–209.
[21] Pasternak RE, Reynolds 3rd CF, Hoch CC, et al. Sleep in spousally bereaved elders with subsyndromal depressive symptoms. Psychiatry Res 1992;43:43–53.
[22] Dewald JF, Meijer AM, Oort FJ, et al. Adolescents' sleep in low-stress and high-stress (exam) times: a prospective quasi-experiment. Behav Sleep Med 2014;12:493–506.
[23] Lazarus RS, Folkman S. Stress, coping and adaptation. New York: Springer; 1984.
[24] Folkman S, Lazarus RS. An analysis of coping in a middle-aged community sample. J Health Soc Behav 1980;21:219–39.
[25] Nagase Y, Uchiyama M, Kaneita Y, et al. Coping strategies and their correlates with depression in the Japanese general population. Psychiatry Res 2009;168:57–66.
[26] Antonovsky A. Unraveling the mystery of health: how people manage stress and stay well. San Francisco, CA, US: Jossey-Bass; 1987.
[27] Dempster M, Howell D, McCorry NK. Illness perceptions and coping in physical health conditions: a meta-analysis. J Psychosomatic Res 2015;79:506–13.
[28] Fortunato VJ, Harsh J. Stress and sleep quality: the moderating role of negative affectivity. Personality Individ Differ 2006;41:825–36.
[29] Abe Y, Mishima K, Kaneita Y, et al. Stress coping behaviors and sleep hygiene practices in a sample of Japanese adults with insomnia. Sleep Biol Rhythms 2011;9:35–45.
[30] Palagini L, Mauri M, Faraguna U, et al. Insomnia symptoms, perceived stress and coping strategies in patients with systemic lupus erythematosus. Lupus 2016;25:988–96.
[31] Pillai V, Roth T, Mullins HM, et al. Moderators and mediators of the relationship between stress and insomnia: stressor chronicity, cognitive intrusion, and coping. Sleep 2014;37:1199–208.
[32] Ministry of Health, Labour and Welfare, Japan. Annual report of the National health and nutrition survey in 2007. Japan: Ministry of Health, Labour and Welfare; 2016 [Online] Available from: http://www.mhlw.go.jp/bunya/kenkou/eiyou09/dl/01-04.pdf [Accessed 16 October 15].
[33] Ikeda N, Takimoto H, Imai S, et al. Data resource profile: the Japan National Health and nutrition survey (NHNS). Int J Epidemiol 2015;44:1842–9.
[34] Furihata R, Kaneita Y, Jike M, et al. Napping and associated factors: a Japanese nationwide general population survey. Sleep Med 2016;20:72–9.
[35] Kaneita Y, Ohida T, Osaki Y, et al. Insomnia among Japanese adolescents: a nationwide representative survey. Sleep 2006;29:1543–50.
[36] Foley DJ, Monjan AA, Brown SL, et al. Sleep complaints among elderly persons: an epidemiologic study of three communities. Sleep 1995;18:425–32.
[37] Liu X, Uchiyama M, Kim K, et al. Sleep loss and daytime sleepiness in the general adult population of Japan. Psychiatry Res 2000;93:1–11.
[38] Munezawa T, Kaneita Y, Osaki Y, et al. The association between use of mobile phones after lights out and sleep disturbances among Japanese adolescents: a nationwide cross-sectional survey. Sleep 2011;34:1013–20.
[39] Cabinet Office of Japan. White paper on the national lifestyle. 2008. Available from: http://www5.cao.go.jp/seikatsu/whitepaper/h20/06_eng/index.html [Accessed 16 October 15].
[40] Kim K, Uchiyama M, Okawa M, et al. An epidemiological study of insomnia among the Japanese general population. Sleep 2000;23:41–7.
[41] Kaneita Y, Ohida T, Uchiyama M, et al. Excessive daytime sleepiness among the Japanese general population. J Epidemiol 2005;15:1–8.
[42] Utsugi M, Saijo Y, Yoshioka E, et al. Relationships of occupational stress to insomnia and short sleep in Japanese workers. Sleep 2005;28:728–35.
[43] Urponen H, Vuori I, Hasan J, et al. Self-evaluations of factors promoting and disturbing sleep: an epidemiological survey in Finland. Soc Sci Med 1988;26:443–50.
[44] Ancoli-Israel S, Roth T. Characteristics of insomnia in the United States: results of the 1991 National sleep foundation survey. Sleep 1999;22:S347–53.
[45] Ribet C, Derriennic F. Age, working conditions, and sleep disorders: a longitudinal analysis in the French cohort E.S.T.E.V. Sleep 1999;22:491–504.
[46] Morin CM, Rodrigue S, Ivers H. Role of stress, arousal, and coping skills in primary insomnia. Psychosom Med 2003;65:259–67.
[47] Kales A, Kales JD. Evaluation and treatment of insomnia. USA: Oxford University Press; 1984.
[48] Hoyt MA, Thomas KS, Epstein DR, et al. Coping style and sleep quality in men with cancer. Ann Behav Med 2009;37:88–93.
[49] Thomas KS, Bower J, Hoyt MA, et al. Disrupted sleep in breast and prostate cancer patients undergoing radiation therapy: the role of coping processes. Psycho Oncol 2010;19:767–76.
[50] Palagini L, Bruno RM, Cheng P, et al. Relationship between insomnia symptoms, perceived stress and coping strategies in subjects with arterial hypertension: psychological factors may play a modulating role. Sleep Med 2016;19:108–15.
[51] Gross JJ, John OP. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. J Personal Soc Psychol 2003;85:348.
[52] Spira AP, Zvolensky MJ, Eifert GH, et al. Avoidance-oriented coping as a predictor of panic-related distress: a test using biological challenge. J Anxiety Disord 2004;18:309–23.
[53] Kashdan TB, Barrios V, Forsyth JP, et al. Experiential avoidance as a generalized psychological vulnerability: comparisons with coping and emotion regulation strategies. Behav Res Ther 2006;44:1301–20.
[54] Kaholokula JK, Antonio MC, Ing CK, et al. The effects of perceived racism on psychological distress mediated by venting and disengagement coping in Native Hawaiians. BMC Psychol 2017;5:2.
[55] Inoue S, Yorifuji T, Sugiyama M, et al. Does habitual physical activity prevent insomnia? A cross-sectional and longitudinal study of elderly Japanese. J Aging Phys Act 2013;21:119–39.
[56] Cairney J, Kwan MY, Veldhuizen S, et al. Who uses exercise as a coping strategy for stress? Results from a national survey of Canadians. J Phys Act Health 2014;11:908–16.
[57] Bahrke MS, Morgan WP. Anxiety reduction following exercise and meditation. Cognitive Ther Res 1978;2:323–33.
[58] Driver HS, Taylor SR. Exercise and sleep. Sleep Med Rev 2000;4:387–402.
[59] Morin CM, LeBlanc M, Daley M, et al. Epidemiology of insomnia: prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. Sleep Med 2006;7:123–30.

[60] Todd M. Daily processes in stress and smoking: effects of negative events, nicotine dependence, and gender. Psychol Addict Behav 2004;18:31–9.
[61] Liu X. Cigarette smoking, life stress, and behavioral problems in Chinese adolescents. J Adolesc Health 2003;33:189–92.
[62] Revell AD, Warburton DM, Wesnes K. Smoking as a coping strategy. Addict Behav 1985;10:209–24.
[63] Hadaway P, Beyerstein B, Kimball M. Addiction as an adaptive response: is smoking a functional behavior? J Drug Issues 1986;16:371–90.
[64] Choi D, Ota S, Watanuki S. Does cigarette smoking relieve stress? Evidence from the event-related potential (ERP). Int J Psychophysiol 2015;98:470–6.
[65] Hajek P, Taylor T, McRobbie H. The effect of stopping smoking on perceived stress levels. Addiction 2010;105:1466–71.
[66] Taylor G, McNeill A, Girling A, et al. Change in mental health after smoking cessation: systematic review and meta-analysis. BMJ (Clinical research ed) 2014;348:g1151.
[67] Epstein LH, Perkins KA. Smoking, stress, and coronary heart disease. J Consult Clin Psychol 1988;56:342.
[68] Gilbert DG, Robinson JH, Chamberlin CL, et al. Effects of smoking/nicotine on anxiety, heart rate, and lateralization of EEG during a stressful movie. Psychophysiology 1989;26:311–20.
[69] Shadel WG, Mermelstein RJ. Cigarette smoking under stress: the role of coping expectancies among smokers in a clinic-based smoking cessation program. Health Psychol 1993;12:443.
[70] Anthenelli RM. Overview: stress and alcohol use disorders revisited. Alcohol Res Curr Rev 2012;34:386–90.
[71] Sudhinaraset M, Wigglesworth C, Takeuchi DT. Social and cultural contexts of alcohol use: influences in a social-ecological framework. Alcohol Res Curr Rev 2016;38:35–45.
[72] Bernardy NC, King AC, Parsons OA, et al. Altered cortisol response in sober alcoholics: an examination of contributing factors. Alcohol 1996;13:493–8.
[73] Cooper ML, Russell M, Skinner JB, et al. Stress and alcohol use: moderating effects of gender, coping, and alcohol expectancies. J Abnorm Psychol 1992;101:139–52.
[74] Spoormaker VI, Schredl M, van den Bout J. Nightmares: from anxiety symptom to sleep disorder. Sleep Med Rev 2006;10:19–31.

Sleep Medicine 38 (2017) 37–43

Contents lists available at ScienceDirect

Sleep Medicine

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

Original Article

Predictors of insomnia onset in adolescents in Japan

Mikiko Tokiya [a], Yoshitaka Kaneita [b, *], Osamu Itani [a], Maki Jike [b], Takashi Ohida [b]

[a] Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Oita, Japan
[b] Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, Tokyo, Japan

ARTICLE INFO

Article history:
Received 28 December 2016
Received in revised form
28 June 2017
Accepted 29 June 2017
Available online 24 July 2017

Keywords:
Cohort studies
Lifestyle
Parasomnias
Risk factors
Incidence

ABSTRACT

Objective: The objective of this study was to clarify the incidence rate and predictive factors of insomnia in Japanese junior and senior high school students.
Methods: We conducted a baseline survey on first year junior and senior high school students (seventh and 10th graders) throughout the nation. A follow-up survey was then conducted two years later. For both surveys, we used self-administered questionnaires inquiring about sleep, mental health status, lifestyle, participation in club activities, and study hours.
Results: A total of 3473 students (776 junior high and 2697 senior high) were suitable for analysis. During the two years leading to the follow-up study, the incidence rate of newly developed insomnia was 7.8% among junior high and 9.2% among senior high school students. Multiple logistic regression analyses revealed that factors associated with new insomnia onset were 'sleep paralysis experience' and 'poor mental health status' in junior high school students, and 'being woken by a nightmare', 'poor mental health status', '≥2 h of extracurricular learning per day' and 'mobile phone use for ≥2 h per day' in senior high school students.
Conclusions: In junior and senior high school students, parasomnias such as nightmares and sleep paralysis, and mental health status can be predictors of insomnia onset. For senior high school students, longer use of mobile phones can be a predictor of insomnia onset. The present findings suggest that sleep health must be promoted among junior and senior high school students in the future.

1. Introduction

Previous studies have demonstrated that sleep plays a crucial role in the healthy development of adolescents, especially the development of important psychosomatic functions such as behavior, emotion, and attention [1–7]. However, qualitative or quantitative disturbance of sleep is common among adolescents, and is now considered a serious school health problem. Studies in various countries have indicated that 4–39% of adolescents have insomnia symptoms [8–25], although the definitions of insomnia used in these studies vary. A few longitudinal studies have clarified the factors associated with the onset of insomnia symptoms in adolescents. Our cohort study of 698 Japanese junior high school students revealed that sleep disorders are associated with future development of poor mental health. In addition, poor mental health increases the risk of onset of sleep disorders [26]. A longitudinal study that was announced later, confirmed the bidirectional association between sleep disorders and onset of a number of neurological symptoms [27,28]. In the United States, Robert et al. reported a similar finding in a longitudinal study of 3134 youths aged 11–17 years conducted between 2000 and 2001 [29]. It is also known that chronic insomnia is a predictor of future problems related to somatic health and interpersonal and psychological function, including poor mental health, lower subjective life satisfaction, and depressive mood [30]. Thus, insomnia is a predictor of poor mental and somatic health, and of alcohol and drug abuse among adolescents. However, the factors that affect insomnia onset among adolescents have not yet been fully elucidated due to the limited number of longitudinal studies completed to date [31,32]. As described above, limited predictors for insomnia onset have been identified to date. We postulated that predictors of insomnia onset can be identified in the lifestyle of junior high and high school students and that these may be useful for developing preventive measures for insomnia.

* Corresponding author. Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, 30-1 Ohyaguchikami-machi, Itabashi-ku, Tokyo, 173-8610, Japan. Fax: +81 3 3972 5878.
E-mail address: nusmpublichealth@gmail.com (Y. Kaneita).

http://dx.doi.org/10.1016/j.sleep.2017.06.028
1389-9457/

Therefore, we conducted an epidemiological, self-administered questionnaire-based survey of junior and senior high school students sampled nationwide to examine the factors affecting the sleep of adolescents. The objectives of this study were (1) to clarify the incidence rates of insomnia among junior and senior high school students in Japan, and (2) to elucidate the factors predicting the onset of insomnia in adolescents.

2. Material and methods

2.1. Study population and design

In 2008, we conducted a cross-sectional nationwide survey evaluating the lifestyle habits of 170 Japanese junior and senior high school students who were selected by random sampling [33]. In January 2010, we invited the head teachers of 170 schools to participate in the present longitudinal-epidemiological survey. The following documents were sent to the schools: (1) a document seeking cooperation in the study, (2) a research plan describing the study's purpose and methodology, and (3) a sample of the questionnaire. With regard to the survey process, we explained that a baseline survey involving students in the seventh and 10th grades would be conducted, and that a follow-up survey would be conducted two years later, when these students had advanced to the ninth and 12th grades, respectively. Furthermore, we explained that the data collected from the self-administered questionnaires would not be used or made available for any purposes other than for this study and that the privacy of the respondents would be safeguarded. Out of the 170 schools contacted, a total of 10 and 14 junior and senior high schools, respectively, agreed to participate in this study.

The 24 schools consisted of 5687 students (1304 and 4383 students in the seventh and 10th grades, respectively) who participated in the study. The baseline survey was conducted between October and November 2010, and the follow-up survey of the same cohort of students was conducted two years later over the same months in 2012. Each survey adopted the same methodology with the homeroom teacher delivering the following information to each student during class: (1) the instructions, (2) the self-administered questionnaire, and (3) an envelope. After completing the questionnaire, each student placed and sealed the forms in the envelopes. The sealed envelopes were collected and opened for the first time at the investigating institution. Permission to conduct the study was obtained from the Ethics Committee of the institution to which the authors are affiliated (Nihon University School of Medicine Ethics Committee, Faculty of Medicine, Oita University Ethics Committee).

2.2. Measurement

The following five categories of data were included in the questionnaires used for each survey. (1) Personal data, including, school name, class name, student name, sex, and birth date. (2) Sleep status, including sleep duration, time of going to bed, time of getting out of bed, difficulty initiating sleep (DIS), difficulty maintaining sleep (DMS), early morning awakening (EMA), subjective sleep assessment, disorders of arousal, nightmares, sleep paralysis, self-treatment to aid sleep onset, and the Japanese version of the Epworth Sleepiness Scale (JESS) [34,35], comprising eight questions designed to measure subjective daytime sleepiness. The total JESS score (0–24) is obtained by summing the scores (0–4) for each question. Individuals with a JESS score of $\geq$11 points were considered to have excessive daytime sleepiness. This cutoff point was also adopted in the present study. (3) Lifestyle factors, including, eating habits, exercise habits, club activities, study time, commuting time, attendance at cram schools or after-school lessons, coffee or tea intake, time spent watching television, playing electronic games, and using mobile phones, respectively, incidences of bullying, reasons for bullying, personal worries, presence of an advisor, and days spent in the school nurse's office. (4) Physical status such as height and weight. (5) Mental health status, including, contentment with daily life, and the Japanese version of the 12-item General Health Questionnaire (GHQ-12) consisting of 12 questions [36]. The total GHQ-12 score (0–12) was obtained by summing the scores (0–1) for each question. Higher scores were considered to indicate poorer mental health status. Although the GHQ was developed for surveys targeting adults, it is known that valid results can also be obtained when administered to adolescents [26,37]. In previous studies conducted by our research group, the GHQ-12 was used for the measurement of poor mental health, and reasonable results could be obtained even at puberty [26,33,37]. Therefore, the GHQ 12 was also used in this study.

The students were provided with the following instructions: (1) this survey is part of a medical study, and their answers for the questionnaire would not affect the participant's academic performance or result in penalties, (2) Participation in the survey must be voluntary, and subjects not opting to participate would not be penalized, (3) The completed questionnaires would not be seen by the participants' teachers, (4) The participants' privacy would be strictly protected.

Participation in the study was voluntary and written informed consent was obtained from all of the participating students.

2.3. Definition

We questioned the participants about the following three issues: (1) difficulty initiating sleep, (2) difficulty maintaining sleep, and (3) early morning awakening, for which the respective questions were phrased as follows: “Over the past month, have you had difficulty falling asleep at night?”, “Over the past month, have you woken up during the night after going to sleep?”, and “Over the past month, have you woken up too early in the morning and had difficulty getting back to sleep?” Each question had five possible replies: “never,” “seldom,” “sometimes,” “often,” or “always.” “Often” and “always” were interpreted as affirmative answers to each question, and such responses were considered to indicate that students had DIS, DMS, or EMA. Insomnia was defined as the presence of one or more of these symptoms.

We also questioned participants about nightmares and sleep paralysis. Regarding nightmares, the question was phrased as “Over the past month, have you been woken by a nightmare?” There were five possible replies: “never,” “seldom,” “sometimes,” “often,” and “always.” “Often” and “always” were taken as affirmative answers, and these responses were considered to indicate that the student suffered from nightmares. Regarding sleep paralysis (*kanashibari* in Japanese), the question was phrased as: “Over the past month, have you suffered a sensation characterized by not being able to move your hands, feet and body when waking up or falling asleep?” There were two possible replies: “yes” or “no.” Students with GHQ-12 scores of $\geq$4 were defined as having poor mental health status.

2.4. Statistical analyses

All analyses in this study were performed separately for the junior and senior high school students. First, the prevalence of insomnia at baseline was calculated. Second, individuals did not demonstrate baseline insomnia, but those who had insomnia at the time of the follow-up survey were defined as subjects developing insomnia during the survey period. The incidence rate of insomnia onset was then calculated. Third, the factors associated with

insomnia onset were examined using the chi-square test and multiple logistic regression analysis. Only participants who did not have insomnia at baseline were included in the multiple logistic regression analysis. The presence of insomnia at the time of the follow-up survey was used as a target variable, and the following items at baseline were used as covariates: sex, sleep duration, extracurricular learning (studying at home or at a cram school, a specialized private school that trains students to pass entrance examinations, after regular school hours), hours of mobile phone use per day, and presence or absence of nightmares, sleep paralysis, poor mental health status, breakfast, coffee or tea intake, exercise habits, and availability of an advisor. The forced entry method was used for multiple logistic regression analyses. SPSS for Windows version 22 (IBM Corp., Armonk, NY, USA) was used for statistical analyses.

3. Results

At the time of the baseline survey, 5687 students (1304 juniors and 4383 seniors) agreed to participate in the study, however, only 3473 students (776 juniors and 2697 seniors) agreed to participate in the follow-up survey. The overall response rate was 61.1% (59.5% and 61.5% for junior and for senior high school students, respectively). Subject characteristics at baseline are shown in Table 1. Individuals with a sleep duration <6 h per night accounted for 8.4% and 25.8% of junior and senior high school students, respectively. The proportion of junior high school students who undertook ≥2 h per day of extracurricular learning was higher than that for the senior cohort. In contrast, nearly half of the senior cohort used their mobile phones for ≥2 h per day. The proportions of individuals with poor mental health, those who skipped breakfast, and who habitually consumed caffeine were higher in the senior cohort than in the junior cohort. The proportions of students who suffered from nightmares and sleep paralysis, and who had access to an advisor were comparable between junior and senior high school students.

The prevalence rate of insomnia at baseline, shown in Table 2, was 9.3% (72 out of 776) and 10.2% (275 out of 2697) among junior and senior high school students.

The incidence rate of insomnia, shown in Table 3, over the two years between the baseline and follow-up surveys was 7.8% and 9.2% in the junior and among senior high school students, respectively.

The results of the chi-squared analysis to examine the factors associated with insomnia onset are shown in Table 4. Sleep paralysis, poor mental health status, and the presence/absence of an advisor were significantly associated with insomnia onset among junior high school students. Among senior high school students, hours of extracurricular learning, hours of mobile phone use, nightmares, poor mental health status, and the presence/absence of an advisor, were significantly associated with insomnia onset.

The results of multiple logistic regression analyses of the factors associated with insomnia onset among each cohort of junior and senior high school students are shown in Table 5. The presence of sleep paralysis experiences and poor mental health status at baseline facilitated insomnia onset among junior high school students. Four factors facilitated insomnia onset among senior high school students: the presence of nightmares, poor mental health status, extracurricular learning for ≥2 h per day, and mobile phone use for ≥2 h per day.

Table 1
Subject characteristics at baseline.

	Junior high school students		Senior high school students	
	N	%	*N*	%
Sex				
Male	372	47.9	1556	57.7
Female	404	52.1	1141	42.3
Sleep duration (h/day)				
<6	65	8.4	697	25.8
≧6	708	91.2	1994	73.9
Unknown	3	0.4	6	0.2
Extracurricular learning (h/day)				
<2	592	76.3	2531	93.8
≧2	172	22.2	153	5.7
Unknown	12	1.5	13	0.5
Mobile phone use (h/day)				
<2	660	85.1	1520	56.4
≧2	106	13.7	1165	43.2
Unknown	10	1.3	12	0.4
Nightmares				
No	747	96.3	2614	96.9
Yes	26	3.4	81	3.0
Unknown	3	0.4	2	0.1
Sleep paralysis				
No	678	87.4	2247	83.3
Yes	30	3.9	156	5.8
Unknown	68	8.8	294	10.9
Poor mental health (GHQ score ≥4)				
No	557	71.8	1595	59.1
Yes	215	27.7	1093	40.5
Unknown	4	0.5	9	0.3
Skipped breakfast				
No	727	93.7	2325	86.2
Yes	44	5.7	367	13.6
Unknown	5	0.6	5	0.2
Habitually consumed caffeine				
No	669	86.2	2101	77.9
Yes	103	13.3	590	21.9
Unknown	4	0.5	6	0.2
Exercise habits				
No	108	13.9	981	36.4
Yes	666	85.8	1703	63.1
Unknown	2	0.3	13	0.5
Advisor				
No	99	12.8	407	15.1
Yes	672	86.6	2250	83.4
Unknown	5	0.6	40	1.5

4. Discussion

In this study, we performed longitudinal epidemiological surveys of junior and senior high school students, and examined predictors of insomnia onset in each cohort. This study had a number of strengths. First, the sample size was sufficient. Second, participants were sampled from schools across Japan. Third, predictive factors for insomnia onset were examined by performing two surveys of the same cohort using a longitudinal design. This was also the first study to have evaluated predictors of insomnia onset in junior and senior high school students in Japan. To date, predictors of insomnia in puberty have not been sufficiently elucidated. The novel finding of this study is the identification of a predictor of insomnia in the normal lifestyle of junior and high school students selected from across Japan. The predictor identified in this study will contribute to the development of preventive

Table 2
The prevalence rate of insomnia at the baseline.

Insomnia at the baseline	Junior high school students		Senior high school students		Total	
	N	%	*N*	%	*N*	%
No	695	89.6	2407	89.2	3102	89.3
Yes	72	9.3	275	10.2	347	10.0
Unknown	9	1.2	15	0.6	24	0.7

Table 3
The incidence rate of insomnia during the 2 years between the baseline and follow-up surveys.

Insomnia at follow up	Junior high school students		Senior high school students		Total	
	N	%	*N*	%	*N*	%
No	638	91.8	2174	90.3	2812	90.7
Yes	54	7.8	221	9.2	275	8.9
Unknown	3	0.4	12	0.5	15	0.5

Table 4
The incidence rate of insomnia and the factors associated with insomnia onset.

	Junior high school students			Senior high school students		
	%	N	*p*	%	N	*p*
Overall	7.8	692		9.2	2395	
Sex			0.54			0.76
Male	8.5	331		9.1	1378	
Female	7.2	361		9.4	1017	
Sleep duration (h/day)			0.62			0.94
<6	6.0	50		9.2	600	
≧6	8.0	639		9.3	1791	
Extracurricular learning (h/day)			0.14			<0.01
<2	8.8	525		8.8	2252	
≧2	5.1	156		16.5	133	
Mobile phone use (h/day)			0.64			<0.01
<2	7.4	596		7.8	1381	
≧2	8.8	91		11.2	1005	
Nightmares			0.81			<0.01
No	7.9	673		8.7	2342	
Yes	6.3	16		34.6	52	
Sleep paralysis			<0.01			0.07
No	7.3	606		9.1	2011	
Yes	24.0	25		14.0	121	
Poor mental health (GHQ score ≥ 4)			<0.01			<0.01
No	5.4	520		7.1	1473	
Yes	15.4	169		12.8	915	
Skipped breakfast			0.66			0.12
No	8.0	653		8.9	2093	
Yes	5.9	34		11.7	299	
Habitually consumed caffeine			0.72			0.44
No	8.0	601		9.0	1887	
Yes	6.9	87		10.1	504	
Exercise habits			0.94			0.92
No	7.6	92		9.3	845	
Yes	7.8	599		9.2	1539	
Advisor			0.01			0.02
No	15.0	80		12.7	347	
Yes	6.9	608		8.7	2017	

Subjects with missing data were excluded from the analysis. *p* was calculated by χ^2 test.

Table 5
Multiple logistic regression analyses of the factors associated with insomnia onset among each cohort (junior high school students, senior high school students).

	Junior high school students			Senior high school students		
	AOR	95% CI	*p*	AOR	95% CI	*p*
Sex			0.44			0.54
Male	1.00			1.00		
Female	0.78	0.41–1.48		0.90	0.65–1.25	
Sleep duration (h/day)			0.44			0.33
<6	1.00			1.00		
≧6	0.60	0.16–2.20		0.84	0.59–1.19	
Extracurricular learning (h/day)			0.24			0.01
<2	1.00			1.00		
≧2	0.61	0.27–1.39		2.10	1.23–3.60	
Mobile phone use (h/day)			0.58			0.03
<2	1.00			1.00		
≧2	1.27	0.55–2.96		1.39	1.03–1.89	
Nightmares			0.41			<0.01
No	1.00			1.00		
Yes	0.39	0.04–3.72		4.46	2.36–8.42	
Sleep paralysis			0.02			0.31
No	1.00			1.00		
Yes	3.59	1.19–10.83		1.34	0.76–2.36	
Poor mental health (GHQ score ≥ 4)			<0.01			<0.01
No	1.00			1.00		
Yes	2.69	1.41–5.15		1.62	1.18–2.21	
Skipped breakfast			0.54			0.19
No	1.00			1.00		
Yes	0.63	0.14–2.82		1.32	0.87–1.99	
Habitually consumed caffeine			0.64			0.98
No	1.00			1.00		
Yes	0.79	0.29–2.12		1.00	0.69–1.43	
Exercise habits			0.91			0.97
No	0.95	0.38–2.36		1.01	0.73–1.39	
Yes	1.00			1.00		
Advisor			0.05			0.32
No	2.17	0.98–4.78		1.22	0.82–1.82	
Yes	1.00			1.00		

Subjects with missing data were excluded from the analysis. *p* was calculated by the multiple logistic regression analysis. AOR, adjusted odds ratio; CI, confidence interval; GHQ, general health questionnaire.

measures targeting individuals going through puberty. Specifically, it is necessary for students with sleep paralysis, nightmares, or poor mental health status to be aware that there is a risk of developing insomnia in the following years. We suggest that it is important to recognize this predictor and provide health counseling and health guidance for students. For students with extensive mobile phone use time, it will be necessary to draw their attention to this habit. In our view, the results of this research will be valuable for the future development of health guidance aimed at students.

The incidence rate of insomnia among Japanese adolescents was 7.8% and 9.2% in the junior and senior high school students, respectively. In the United States, Robert et al. reported an annual incidence rate of insomnia of 14.0% among adolescents aged 11–17 years ($N = 3134$), based on the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [32]. In China, Luo et al. reported an annual incidence rate of insomnia of 16.0% among adolescents aged 11–18 years ($N = 2787$), as defined by the Insomnia Severity Index [31]. While a direct comparison may not be appropriate because the definitions of insomnia used in these studies differed, the incidence rates of insomnia among the Japanese adolescents were lower than that reported for the Chinese and North American adolescents.

The multiple logistic regression analyses conducted in this study revealed a number of predictors of insomnia onset. First, poor mental health status at baseline appeared to be a predictor of insomnia onset for both high school cohorts. This result was consistent with those of previous studies investigating the association between sleep and mental health among adolescents [26,29]. Previously, Robert et al. suggested that symptoms of depression (a depressed mood, anhedonia, and irritability) could be predictive factors for insomnia [29]. Furthermore, in a two-year follow-up study of 516 Japanese junior high school students, Kaneita et al. indicated that poor mental health status could be a predictor of sleep disorders [26]. Similarly, other studies using a range of cohorts have also reported poor mental health status as a predictor of insomnia onset, suggesting that this factor may be a universal phenomenon [27,28,38,39]. Two explanations for this should be considered. First, insomnia symptoms are widely reported in the majority of patients diagnosed with depression, while improvement in depressive symptoms was reportedly correlated with an improvement in insomnia status. This supports the possibility that poor mental health status can cause insomnia. Second, despite the absence of a causal association between poor mental health status and insomnia, the presence of a confounding factor that is correlated with both of these factors may support the association.

Moreover, students incidentally recognized poor mental health status at an earlier stage than insomnia. While we adjusted for the effects of confounding factors by entering assumed models in our multiple logistic regression analyses, we could not rule out the possibility that some confounding factors were not adjusted for in our analysis.

This study revealed that sleep paralysis was a predictor of insomnia among Japanese junior high school students, i.e. students who did not have insomnia symptoms, but who had experienced sleep paralysis at baseline tended to become aware of having insomnia over the following two years. Sleep paralysis is a state in which a person is unable to move their limbs, body, and head, although he/she is subjectively awake. It is known to be one of the symptoms of narcolepsy [40]. Takeuchi et al. conducted a physiological experiment using polysomnography on 1314 Japanese university students who had neither cataplexy nor narcolepsy. They investigated the sleep phase during which sleep paralysis occurred by intentionally interrupting the sleep of these participants. They found that once sleep had been interrupted, REM sleep occurred upon resuming sleep without a non-REM sleep phase, and that sleep paralysis was induced during this REM sleep after sleep resumption [41]. This study indicated that sleep paralysis and DMS (one of the symptoms of insomnia) were closely associated. The results of our study suggested that participants had first become aware of their sleep paralysis before becoming aware of insomnia symptoms by the time of the follow-up survey. Therefore, sleep paralysis may be recognized earlier than symptoms of insomnia.

This study also revealed that having nightmares could be a predictor of insomnia onset among senior high school students. Nightmares are defined as awakening from a frightening dream and rapidly becoming oriented and alert, according to the DSM-IV [42], and this is known to be a phenomenon incidental to DMS. Our previous study showed that experiences of sleep paralysis were significantly more frequent among adolescents in whom the frequency of nightmares was higher, and an association between these factors was identified [43]. The key finding in the present study is that the experience of nightmares tended to be recognized before insomnia, as with sleep paralysis. Sleep paralysis and nightmares are both classified as 'REM-related parasomnia' in the third edition of the International Classification of Sleep Disorders (ICSD-3), and are considered to be linked physiologically [44]. The present study revealed that sleep paralysis and nightmares could be potential predictors of DMS. Accordingly, they are important signs that may appear before individuals become aware of insomnia. DMS tends to occur under conditions of physical and mental stress, in individuals exhibiting unhealthy lifestyles, or who experience disturbances in their normal sleep rhythm [45]. Treatment of sleep paralysis or nightmares as soon as such disorders are recognized may help to prevent DMS onset. Therefore, such disorders need to be taken into consideration in the context of student health.

This study revealed that extracurricular learning for $\geq$2 h per day could be a predictor of insomnia onset in senior high school students. This result was somewhat unexpected, as initially, it would appear more likely that reduced sleep duration resulting from longer extracurricular learning would increase sleepiness and, therefore, promote sleep. This result needs to be interpreted carefully. One possible explanation is that another factor promoted both longer extracurricular learning and insomnia, and that longer extracurricular learning was the first 'symptom' to develop. For instance, it could be postulated that strong psychological stress due to the preparation for entrance exams may increase the motivation for longer extracurricular learning, leading to insomnia. Another possible mechanism is that studying until just before bedtime may disturb the mechanism for sleep regulation. Around a few hours before the time of sleep onset, the level of arousal (state of excitement) becomes maximal, therefore overriding any increased need to sleep. Sleepiness then increases rapidly, and changes appear in different sleep-promoting physiological parameters. The appropriate phase associations among these parameters are sustained, allowing sleep to be initiated and maintained [46]. However, it can be inferred that studying during this time zone may maintain the high level of arousal until immediately before sleep [47,48]. Previous studies have indicated that a higher arousal level before sleep onset prevents sleep, and that allowing at least 1 h of relaxation before bedtime effectively provides good sleep [47,48]. However, considering that in contemporary Japanese society, many junior and senior high school students are pressured with heavy homework loads, it is quite common for these students to continue studying until bedtime. It is highly probable that the resulting brain activation during this period may induce insomnia. To mitigate the risk of insomnia onset, it is essential for students to consider these results, and adjust their study schedule accordingly, after returning home from school.

This study revealed that $\geq$2 h of mobile phone use per day by senior high school students could be a predictor of insomnia onset. In recent years, studies on the association between the use of electronic devices, such as mobile phones, and insomnia among adolescents have been conducted from different perspectives [33,49,50]. A cross-sectional study of 95,680 junior and senior high school students in Japan by Munezawa et al. in 2008 indicated that the number of students displaying insomnia symptoms increased significantly with the frequency of mobile phone usage after switching off their bedroom light [33]. In 2010, Aora et al. conducted a cross-sectional study of 738 youths aged 11–13 years in the United Kingdom to investigate the association between the frequent use of technology, including, mobile phones, and the quality/quantity of sleep. The authors reported that the odds ratios (ORs) of insomnia symptoms (DIS, DMS, EMA) for students who used their devices every night were significantly higher [49]. Hysing et al. performed a cross-sectional study in 2010 involving 9846 Norwegian youths aged 16–19 years to investigate the association between the use of electronic devices with a display (including, personal computers, mobile phones, tablets, game consoles, televisions) and sleep. The results indicated that students who used the devices for $\geq$2 h during the day or before the time of sleep onset had a sleep onset latency of at least 60 min [50]. However, all of these studies were cross-sectional and a causal association between electronic device usage and insomnia in adolescents could not be determined. Our study was longitudinal and the results indicated that longer mobile phone usage increased the risk of insomnia onset after a two-year period.

A number of studies have investigated the association between use of electronic devices, such as mobile phones, and melatonin [51–53]. The blue light emitted from the display of electronic devices has been reported to have the strongest suppressive effects on melatonin secretion, compared to lights in other parts of the visual spectrum [51,52]. Furthermore, the blue light was reported to prematurely advance the plasma melatonin rhythm phase. Even at a low illuminance level (8 lux [28 $\mu W/cm^2$]), blue light exerts the same degree of phase advancement as 12,000 lux of white light (4300 $\mu W/cm^2$) [53]. Meanwhile, a longitudinal study by Fossum et al., reported that using a mobile phone in bed, or before going to bed was a predictor of insomnia onset and evening chronotype onset [54]. Light exposure by the device displays may delay the sleep phase. It is important to conduct a comprehensive examination of the associations among factors such as the use of mobile phones, insomnia, DSPD, and chronotype. Furthermore, similarly to

the effects of extracurricular learning, mobile phone use before bedtime may activate the user's brain and induce insomnia. In 2008, Foley et al. conducted a nationwide survey of 2107 children aged 5–18 years in New Zealand to investigate the association between their behaviors during the 90 min before sleep and after sleep onset. The authors reported that sleep onset was delayed in the screen-based sedentary behavior group [55]. Furthermore, adolescence is a period when sleep onset shifts towards more nocturnal patterns [2,4,56]. Specifically, sleep/wake patterns are reported to change among adolescents of senior high school age [1]. Therefore, adolescents need to be educated on sleep-related health issues so that they have a better idea of how to spend time before sleep onset.

This study had some limitations. First, the definitions of the insomnia symptoms we used were not the diagnostic criteria defined in the second edition of the ICSD (ICSD-2). The ICSD-2 stipulates that the following two features are essential for the diagnosis of insomnia: (1) complaints of insomnia symptoms during nocturnal sleep, and (2) daytime sleep-related mental and/or physical impairment [57]. Previous self-administered questionnaire surveys included questions regarding the presence of DIS, DMS, and EMA in order to identify the presence of insomnia symptoms [17,24,33,58]. The present study was consistent with previous studies in this aspect. Second, because of ethical considerations and space limitations in the questionnaire, questions regarding a number of items could not be included. Examples are socioeconomic status (family financial conditions, academic performance), problematic behavior (drinking alcohol and smoking), sleep environment (bedroom temperature and humidity), and reliability of anamnesis. We cannot exclude the possibility that these factors may affect the predictors found in the present study as well as insomnia onset, and should be considered in future studies. Third, this study used self-reported data. Fourth, we did not measure the Delayed Sleep Phase Disorder (DSPD) or chronotype. DSPD or chronotype may be a confounding factor in relation to lifestyle and insomnia. It is crucial to gather information on DSPD and chronotype for future research. Finally, the number of junior high school students who participated in this study was smaller than that of the senior students. This may have decreased the statistical power of the study. Future studies will need to include a sufficient number of schools and participants.

5. Conclusions

This study has obtained useful epidemiological data on predictors of insomnia onset among Japanese junior and senior high school students. We found that parasomnia, such as nightmares and sleep paralysis, as well as mental health status were associated with insomnia. We also found that the use of mobile phones for $\geq$2 h per day and extracurricular learning for $\geq$2 h per day affected the sleep of senior high school students. These results provide scientific evidence that will be useful for the development of public health strategies targeting adolescents. The present findings suggest that sleep health should be promoted among junior and senior high school students in the future.

Acknowledgements

The authors are deeply grateful to the students and teachers of the junior and senior high schools for their cooperation. We wish to express our thanks to Mrs Yukiko Abe and Ms Eri Hirano for their help with this study. This study was supported by a Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (No. 21390208).

Conflict of interest

All authors declare no competing financial interest.

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: http://dx.doi.org/10.1016/j.sleep.2017.06.028.

References

[1] Wolfson AR, Carskadon MA. Sleep schedules and daytime functioning in adolescents. Child Dev 1998;69:875–87.
[2] Dahl RE, Lewin DS. Pathways to adolescent health sleep regulation and behavior. J Adolesc Health 2002;31:175–84.
[3] Fallone G, Owens JA, Deane J. Sleepiness in children and adolescents: clinical implications. Sleep Med Rev 2002;6:287–306.
[4] Carskadon MA, Acebo C, Jenni OG. Regulation of adolescent sleep: implications for behavior. Ann N Y Acad Sci 2004;1021:276–91.
[5] Beebe DW, Rose D, Amin R. Attention, learning, and arousal of experimentally sleep-restricted adolescents in a simulated classroom. J Adolesc Health 2010;47:523–5.
[6] Brand S, Kirov R. Sleep and its importance in adolescence and in common adolescent somatic and psychiatric conditions. Int J Gen Med 2011;4:425–42.
[7] Casement MD, Keenan KE, Hipwell AE, et al. Neural reward processing mediates the relationship between insomnia symptoms and depression in adolescence. Sleep 2016;39:439–47.
[8] Price VA, Coates TJ, Thoresen CE, et al. Prevalence and correlates of poor sleep among adolescents. Am J Dis Child 1978;132:583–6.
[9] Kirmil-Gray K, Eagleston JR, Gibson E, et al. Sleep disturbance in adolescents: sleep quality, sleep habits, beliefs about sleep, and daytime functioning. J Youth Adolesc 1984;13:375–84.
[10] Yang L, Zuo C, Eaton LF. Research note: sleep problems of normal Chinese adolescents. J Child Psychol Psychiatry 1987;28:167–72.
[11] Morrison DN, McGee R, Stanton WR. Sleep problems in adolescence. J Am Acad Child Adolesc Psychiatry 1992;31:94–9.
[12] Huang YS, Wang CH, Guilleminault C. An epidemiologic study of sleep problems among adolescents in North Taiwan. Sleep Med 2010;11:1035–42.
[13] Blader JC, Koplewicz HS, Abikoff H, et al. Sleep problems of elementary school children. A community survey. Arch Pediatr Adolesc Med 1997;151:473–80.
[14] Johnson EO, Chilcoat HD, Breslau N. Trouble sleeping and anxiety/depression in childhood. Psychiatry Res 2000;94:93–102.
[15] Liu X, Uchiyama M, Okawa M, et al. Prevalence and correlates of self-reported sleep problems among Chinese adolescents. Sleep 2000;23:27–34.
[16] Patten CA, Choi WS, Gillin JC, et al. Depressive symptoms and cigarette smoking predict development and persistence of sleep problems in US adolescents. Pediatrics 2000;106:E23.
[17] Roberts RE, Roberts CR, Chen IG. Ethnocultural differences in sleep complaints among adolescents. J Nerv Ment Dis 2000;188:222–9.
[18] Liu X, Zhou H. Sleep duration, insomnia and behavioral problems among Chinese adolescents. Psychiatry Res 2002;111:75–85.
[19] Roberts RE, Roberts CR, Chen IG. Impact of insomnia on future functioning of adolescents. J Psychosom Res 2002;53:561–9.
[20] Johnson EO, Roth T, Breslau N. The association of insomnia with anxiety disorders and depression: exploration of the direction of risk. J Psychiatr Res 2006;40:700–8.
[21] Hysing M, Pallesen S, Stormark KM, et al. Sleep patterns and insomnia among adolescents: a population-based study. J Sleep Res 2013;22:549–56.
[22] Amaral O, Garrido A, Pereira C, et al. Sleep patterns and insomnia among Portuguese adolescents: a cross-sectional study. Aten Primaria 2014;46(Suppl 5):191–4.
[23] Kronholm E, Puusniekka R, Jokela J, et al. Trends in self-reported sleep problems, tiredness and related school performance among Finnish adolescents from 1984 to 2011. J Sleep Res 2015;24:3–10.
[24] Kaneita Y, Ohida T, Osaki Y, et al. Insomnia among Japanese adolescents: a nationwide representative survey. Sleep 2006;29:1543–50.
[25] Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev 2002;6:97–111.
[26] Kaneita Y, Yokoyama E, Harano S, et al. Associations between sleep disturbance and mental health status: a longitudinal study of Japanese junior high school students. Sleep Med 2009;10:780–6.
[27] Alvaro PK, Roberts RM, Harris JK, et al. The direction of the relationship between symptoms of insomnia and psychiatric disorders in adolescents. J Affect Disord 2017;207:167–74.
[28] Shanahan L, Copeland WE, Angold A, et al. Sleep problems predict and are predicted by generalized anxiety/depression and oppositional defiant disorder. J Am Acad Child Adolesc Psychiatry 2014;53:550–8.
[29] Roberts RE, Duong HT. Depression and insomnia among adolescents: a prospective perspective. J Affect Disord 2013;148:66–71.
[30] Roberts RE, Roberts CR, Duong HT. Chronic insomnia and its negative consequences for health and functioning of adolescents: a 12-month prospective study. J Adolesc Health 2008;42:294–302.
[31] Luo C, Zhang J, Pan J. One-year course and effects of insomnia in rural Chinese adolescents. Sleep 2013;36:377–84.

[32] Roberts RE, Roberts CR, Chan W. Persistence and change in symptoms of insomnia among adolescents. Sleep 2008;31:177–84.
[33] Munezawa T, Kaneita Y, Osaki Y, et al. The association between use of mobile phones after lights out and sleep disturbances among Japanese adolescents: a nationwide cross-sectional survey. Sleep 2011;34:1013–20.
[34] Takegami M, Suzukamo Y, Wakita T, et al. Development of a Japanese version of the Epworth Sleepiness Scale (JESS) based on item response theory. Sleep Med 2009;10:556–65.
[35] Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 1991;14:540–5.
[36] Doi Y, Minowa M. Factor structure of the 12-item general health questionnaire in the Japanese general adult population. Psychiatry Clin Neurosci 2003;57:379–83.
[37] Kaneita Y, Ohida T, Osaki Y, et al. Association between mental health status and sleep status among adolescents in Japan: a nationwide cross-sectional survey. J Clin Psychiatry 2007;68:1426–35.
[38] Alvaro PK, Roberts RM, Harris JK. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. Sleep 2013;36:1059–68.
[39] Urrila AS, Paunio T, Palomaki E, et al. Sleep in adolescent depression: physiological perspectives. Acta Physiol Oxf 2015;213:758–77.
[40] Vaughn BV, D'Cruz ONF. Cardinal manifestations of sleep disorders. In: Kryger MH, Roth T, editors. Principles and practice of sleep medicine. 6th ed. Elsevier; 2017. p. 582.
[41] Takeuchi T, Miyasita A, Sasaki Y, et al. Isolated sleep paralysis elicited by sleep interruption. Sleep 1992;15:217–25.
[42] Nielsen T, Carr M. Nightmare and nightmare function. In: Kryger MH, Roth T, editors. Principles and practice of sleep medicine. 6th ed. Elsevier; 2017. p. 547.
[43] Munezawa T, Kaneita Y, Osaki Y, et al. Nightmare and sleep paralysis among Japanese adolescents: a nationwide representative survey. Sleep Med 2011;12:56–64.
[44] Mizuno S, Horiguchi J. ICSD-3 to nihon no suiminiryou (ICSD-3 and Japanese sleep medicine) Suiminjizuihanshyougun (Sleep parasomnia). suimin iryou (Japanese Journal of Sleep Medicine) 2015;9:185–93 [in Japanese].
[45] Ogawa K. Fumin (Insomnia) Nemuri no huchyou to sono chiryou akumu to kanashibari (Bad sleep and its treatment nightmare and sleep paralysis its psychophysiological background). Kokoro no Kagaku (Human Mind Science). 2015. p. 29–34 [in Japanese].
[46] Lavie P. Ultrashort sleep-waking schedule. III. 'Gates' and 'forbidden zones' for sleep. Electroencephalogr Clin Neurophysiol 1986;63:414–25.
[47] Morin CM. Psycological and behavioral treatments for insomnia: approaches and efficacy. In: Kryger MH, Roth T, Dement WC, editors. Principles and practice of sleep medicine. 5th ed. St. Louis, Missouri: Elsevier/Saunders; 2011. p. 866–83.
[48] Turner PL, Van Someren EJ, Mainster MA. The role of environmental light in sleep and health: effects of ocular aging and cataract surgery. Sleep Med Rev 2010;14:269–80.
[49] Arora T, Broglia E, Thomas GN, et al. Associations between specific technologies and adolescent sleep quantity, sleep quality, and parasomnias. Sleep Med 2014;15:240–7.
[50] Hysing M, Pallesen S, Stormark KM, et al. Sleep and use of electronic devices in adolescence: results from a large population-based study. BMJ Open 2015;5, e006748.
[51] Brainard GC, Hanifin JP, Greeson JM, et al. Action spectrum for melatonin regulation in humans: evidence for a novel circadian photoreceptor. J Neurosci 2001;21:6405–12.
[52] Thapan K, Arendt J, Skene DJ. An action spectrum for melatonin suppression: evidence for a novel non-rod, non-cone photoreceptor system in humans. J Physiol 2001;535:261–7.
[53] Warman VL, Dijk DJ, Warman GR, et al. Phase advancing human circadian rhythms with short wavelength light. Neurosci Lett 2003;342:37–40.
[54] Fossum IN, Nordnes LT, Storemark SS, et al. The association between use of electronic media in bed before going to sleep and insomnia symptoms, daytime sleepiness, morningness, and chronotype. Behav Sleep Med 2014;12:343–57.
[55] Foley LS, Maddison R, Jiang Y, et al. Presleep activities and time of sleep onset in children. Pediatrics 2013;131:276–82.
[56] Roenneberg T, Kuehnle T, Juda M, et al. Epidemiology of the human circadian clock. Sleep Med Rev 2007;11:429–38.
[57] American Academy of Sleep Medicine. The international classification of sleep disorders. 2nd ed. Igakusyoin: Diagnostic &Cording Manual; 2010 [in Japanese].
[58] Ohayon MM, Roberts RE, Zulley J, et al. Prevalence and patterns of problematic sleep among older adolescents. J Am Acad Child Adolesc Psychiatry 2000;39:1549–56.

Sleep Medicine 39 (2017) 87–94

Contents lists available at ScienceDirect

Sleep Medicine

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

Original Article

Short sleep duration, shift work, and actual days taken off work are predictive life-style risk factors for new-onset metabolic syndrome: a seven-year cohort study of 40,000 male workers

Osamu Itani [a], Yoshitaka Kaneita [b, *], Mikiko Tokiya [a], Maki Jike [b], Atsushi Murata [b], Sachi Nakagome [b], Yuichiro Otsuka [b], Takashi Ohida [b]

[a] Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Oita, Japan
[b] Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, Tokyo, Japan

ARTICLE INFO

Article history:
Received 17 April 2017
Received in revised form
13 July 2017
Accepted 14 July 2017
Available online 9 September 2017

Keywords:
Diabetes mellitus
Dyslipidemias
Holidays
Hypertension
Longitudinal studies
Obesity

ABSTRACT

Background: This longitudinal study investigated the effects of various lifestyle-related factors – including sleep duration, shift work, and actual days taken off work – on new-onset metabolic syndrome (MetS).

Methods and results: A total of 39,182 male employees (mean age 42.4 ± 9.8 years) of a local government organization in Japan were followed up for a maximum of seven years, between 1999 and 2006. Multivariate analysis (Cox proportional hazard method) identified seven high-risk lifestyle factors that were significantly associated with new-onset MetS or a range of metabolic factors (obesity, hypertension, hyperglycemia, dyslipidemia): (1) short sleep duration (<5 h/day), (2) shift work, (3) insufficient number of days off work, (4) always eating until satiety, (5) not trying to take every opportunity to walk, (6) alcohol intake ≥60 g/day, and (7) smoking. In addition, a higher number of these high-risk lifestyle factors significantly promoted the onset of MetS. The hazard ratio for MetS associated with 0–1 high-risk lifestyle parameters per subject at the baseline was set at 1.00. Hazard ratios associated with the following numbers of high-risk lifestyle parameters were: 1.22 (95% CI 1.15–1.29) for 2–3 of these parameters; and 1.43 (1.33–1.54) for 4–7.

Conclusion: An increase in the number of high-risk lifestyle factors – such as short sleep duration, shift work, and an insufficient number of days off work – increased the risk of MetS onset. Comprehensive strategies to improve a range of lifestyle factors for workers, such as sleep duration and days off work, could reduce the risk of MetS onset.

1. Introduction

According to the World Health Organization (WHO), fatalities resulting from non-communicable diseases (NCDs) accounted for 36 million (63.2%) of 57 million deaths from all causes, occurring worldwide in 2008 [1]. The fatalities according to disease were as follows: 17.3 million from cardiovascular diseases, 7.6 million from cancer, 4.2 million from chronic respiratory disease, and 1.3 million from diabetes mellitus [1]. The total direct medical costs for treating these four diseases were estimated to be 2.88 trillion dollars in 2010 [1]. NCDs are currently the leading cause of death in several countries, impoverishing individuals and families but also limiting the social and economic development of countries. The United Nations and WHO have issued warnings that NCDs and their socioeconomic effects are the largest health problem of the 21st century; however, it is anticipated that implementation of national-level comprehensive action plans would decrease NCD-related morbidity and adverse socioeconomic effects [1,2].

In Japan, the National Health Promotion Movement for the 21st Century has been strongly promoted as a national project to combat NCDs since 2002 [3]. Implementation of Specific Health Checkups and Specific Health Guidance [4], focusing specifically on metabolic syndrome (MetS), was endorsed for this project [3]. MetS is a cluster of different diseases affecting a single individual, including abdominal obesity, insulin resistance, dyslipidemia, and

* Corresponding author. Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, 30-1 Ohyaguchikami-machi, Itabashi-ku, Tokyo 173-8610, Japan. Fax: +81 3 3972 5878.
E-mail address: nusmpublichealth@gmail.com (Y. Kaneita).

http://dx.doi.org/10.1016/j.sleep.2017.07.027
1389-9457/

elevated blood pressure, and is known to significantly increase the risk of cardiovascular disease and type 2 diabetes mellitus [5]. MetS is currently attracting attention as an important target in strategies for the prevention of arteriosclerotic diseases.

Improvement of lifestyle is regarded as the most important prevention strategy against MetS [6]. Previous studies have identified a range of lifestyle factors that pose a high risk for MetS onset, including exercise [7], diet [8], alcohol intake [9,10], and smoking [11]. More recently, further risk factors have been identified, such as short sleep duration and insufficient days off work [12]. Itani et al. reported that insufficient days off work per week was a risk factor for new-onset MetS in a cohort study involving approximately 30,000 male public servants who worked for local governments in Japan [13]. However, these studies had considerable limitations. First, no causal association could be determined, due to the cross-sectional design of the studies [12]. Second, the diagnosis of MetS may have been potentially inaccurate, due to insufficient data on the treatment of the individual diseases constituting MetS [13]. In addition, evaluation of the cumulative incidence over time was insufficient because the data collected during the observation period were not included in data analysis [13].

A longitudinal epidemiological study was planned to examine the association between MetS and lifestyle factors, avoiding the above-mentioned limitations. The present study met the following criteria: (1) identification of multiple individual risk factors for MetS onset among the key lifestyle-related parameters affecting workers, including sleep, work conditions (shift work), and days off; (2) a sample size of several tens of thousands of subjects; (3) a longitudinal study design; (4) inclusion of all data obtained during the observation period in the analysis; and (5) use of a database including treatment of individual diseases constituting MetS. Furthermore, a novel parameter was assessed: the effects of an increase in the number of high-risk lifestyle parameters on the onset of MetS, or the individual diseases constituting MetS. The effects of individual lifestyle factors on MetS onset have been investigated in previous epidemiological studies. To date, however, limited research on the effects of an increase in the number of lifestyle-related factors, including sleep and days off work, on MetS onset has been conducted in Japan and worldwide. The present study included sleep, shift work, and days off work as lifestyle-related factors for analysis. This resulted in a comprehensive analysis of the key lifestyle parameters affecting workers. It was considered that identifying the effects of an increase in the number of lifestyle habits on the onset of MetS might contribute to the development of recommendations for prevention of MetS.

2. Methods

2.1. Study subjects and data collection

This cohort study retrospectively analyzed the medical checkup data for employees of a local government organization in Japan. Medical checkups of all employees were conducted annually, and the data included in this study were from medical checkups between 1999 and 2006. The number of female employees (3286) was markedly smaller than that of male employees (42,136), and the types of work and work environment for males and females were substantially different in this organization. The study therefore excluded data for female employees from analyses, in order to ensure sample homogeneity. A total of 39,182 male employees underwent medical checkups in 1999 (consultation rate: 93.0%) and were followed up for a maximum of 7 years between 1999 and 2006. Most employees reaching the mandatory retirement age (60 years of age in this organization) retired, and were precluded from further annual medical checkups. Therefore, these employees were treated as dropouts. Finally, a total of 22,423 male employees receiving annual medical checkups between 1999 and 2006 were included in this study (complete follow-up rate: 57.2%).

2.2. Study variables and measurements

The checkups included the following: (1) body height/weight, blood pressure measurement; (2) blood test; (3) urinalysis; (4) a self-administered questionnaire; (5) electrocardiography; and (6) chest X-ray. Physical assessments included measurements of height, body weight, and blood pressure (systolic/diastolic). Blood pressure was measured twice in the right upper arm with the subject seated on a chair, using the auscultatory method.

As a condition of blood collection, the subjects were instructed not to take any food or beverage, except for water and tea, within 10 h before the checkup. The parameters included in the blood test were high-density lipoprotein (HDL) cholesterol, triglyceride, and fasting plasma glucose.

The questionnaire sought information on previous or present diseases (including current medication and outpatient care), physical and mental complaints, sleep duration, actual availability of weekly rest days, presence/absence of shift work, eating habits, exercise, alcohol consumption, and smoking. With regard to sleep hours, the subjects were requested to select their sleep hours per day as <5 h, ≥5 h but <7 h, or ≥7 h. The work regime followed in this organization was either fixed daytime work or shift work; the shift work cycles were: (1) a night shift every 3 days, (2) a night shift every 4 days, (3) a night shift every 5 days, (4) a night shift every 6 days, and (5) a night shift every 8 days. In this study, all five of these patterns were classified as “shift work”. The organization for which the subjects worked stipulated that 1 or 2 rest days per week were periodically provided in principle. It was also asked whether the subjects had been able to take weekly rest days over the past 1–2 months, with the following answer options: “able to take most of the available weekly rest days off” and “unable to often take weekly rest days off”.

Regarding eating habits, subjects were asked whether they ate in moderation, and were requested to select one of the following answers: always, sometimes, or never. Regarding exercise habits, subjects were asked whether they tried to take every opportunity to walk, for instance, by using the stairs, where possible. A two-point scale (yes or no) was used for the evaluation of exercise habits.

With regard to drinking habits, the subjects answered a question on whether they consumed alcoholic beverages. Subjects were asked to report their alcohol consumption per drinking session on a Japanese sake (rice wine)-converted basis. The pure alcohol intake was then calculated on the basis that the pure alcohol content per 180 mL of Japanese sake corresponds to 20 g. With regard to smoking habits, the subjects answered a question on whether they smoked, by selecting an option from among “yes”, “no”, and “quit smoking”. Smokers were asked to report the number of cigarettes smoked per day. With regard to mental complaints, the subjects answered questions such as whether they experienced irritability, reduced concentration, and lethargy.

2.3. Definitions of various metabolic diseases

2.3.1. Obesity

Obesity was defined as a body mass index (BMI) of ≥25 kg/m^2 according to the standards of the Japan Society for the Study of Obesity [14].

2.3.2. Hypertension

Hypertension was defined as a mean value of two systolic blood pressure measurements of ≥140 mmHg or a mean value of two

diastolic blood pressure measurements of ≥90 mmHg, in accordance with criteria determined by WHO [15] and the Japanese Society of Hypertension [16]. Patients prescribed antihypertensive medication were also defined as hypertensive.

2.3.3. Hyperglycemia

Furthermore, in accordance with the criteria stipulated by the Japan Diabetes Society [17], a fasting plasma glucose level of ≥126 mg/dL was considered to represent hyperglycemia. Patients prescribed hypoglycemic medication were also defined as hyperglycemic.

2.3.4. Dyslipidemia

In accordance with the criteria stipulated by the Japan Atherosclerosis Society [18], a serum triglyceride level of ≥150 mg/dL was considered to be high, and an HDL cholesterol level of <40 mg/dL was defined as low. Patients prescribed antilipidemic medication were also defined as dyslipidemia.

2.4. Definition of metabolic syndrome

The definition of MetS in this study was adapted from the diagnostic criteria set by the Japanese Committee to Evaluate Diagnostic Standards for Metabolic Syndrome [19]. Specifically, the present study replaced the endpoint of visceral fat accumulation from the abdominal circumference criteria in the original criteria with the BMI criteria. The diagnostic criteria for MetS used in the present study were as follows: BMI ≥25 kg/m^2 and two or more of the following conditions being met: (1) systolic blood pressure measurements of ≥130 mmHg, or diastolic blood pressure measurements of 85 mmHg, or use of antihypertensive medication; (2) triglyceride ≥150 mg/dL, or HDL cholesterol <40 mg/dL, or use of antilipidemic medication; and (3) glucose ≥110 mg/dL, or use of hypoglycemic medication.

2.5. Definition of clustering of lifestyle factors

The following seven lifestyle-related parameters were defined as conferring a significantly high risk for new-onset obesity, hypertension, hyperglycemia, dyslipidemia, or MetS: (1) <5 h/day of sleep; (2) shift work; (3) insufficient days off work; (4) always eating until satiety; (5) not attempting to take every opportunity to walk; (6) consuming ≥60 g/day of pure alcohol (≥540 mL of Japanese sake); and (7) smoking.

2.6. Statistical analyses

The prevalence of obesity, hypertension, hyperglycemia, dyslipidemia, and MetS was calculated by age class (<30, 30–39, 40–49, >50 years) at the baseline (1999). A Mantel–Haenszel test for trend was performed to examine the age-class-based trends in the prevalence of each disease. The level of significance was set at $p < 0.05$.

After selecting subjects who were obese at the baseline, the Cox proportional hazards model (forced entry method) was applied, using new-onset obesity during the observation period as the endpoint. The following seven lifestyle-related parameters were used as explanatory variables in this test: sleep duration (≥5/<5 h/day); shift work (no/yes); actual use of weekly rest days (most rest days taken/most rest days not taken); eating in moderation (always/sometimes/never); trying to take every opportunity to walk (yes/no); alcohol intake (no current drinking habit/<60/≥60 g/day pure alcohol intake); smoking (non-smoker or quit smoking/<40/≥40 cigarettes/day). The analysis was adjusted for the following factors: age class; hypertension; hyperglycemia; dyslipidemia; and mental health complaints (irritability, reduced concentration, and lethargy). Subjects who did not undergo a medical checkup during the observation period were withdrawn from follow-up at that point. For hypertension, hyperglycemia, dyslipidemia, and MetS, the tests were performed in the same way as for obesity.

Finally, any lifestyle-related parameters found to be significantly associated with new onset of any of the following diseases was defined as high-risk: obesity, hypertension, hyperglycemia, dyslipidemia, or MetS. The Cox proportional hazards model was then applied, using the onset of each of the above-mentioned diseases as the objective variable, and the number of high-risk lifestyle parameters present at the baseline as the explanatory variable. The model was adjusted for the following variables: age class, hypertension, hyperglycemia, dyslipidemia, and mental health complaints. Subjects who did not undergo a medical checkup during the observation period were withdrawn from follow-up at that point. For hypertension, hyperglycemia, dyslipidemia, and MetS, the tests were performed in the same way as for obesity. All analyses were performed using SPSS 22 for Windows (IBM Corp, Armonk, NY, USA).

2.7. Ethical considerations

For the present study, the following measures were exercised to safeguard the privacy of the subjects: (1) only one researcher had access to the personal data of the subjects, and (2) the files containing the subjects' personal data were managed separately from those used for statistical analyses. This study was conducted in accordance with the tenets of the Personal Information Protection Act enforced in Japan and the Ethical Guidelines for Epidemiological Studies jointly announced by the Ministry of Health, Labor and Welfare and the Ministry of Education, Culture, Sports, Science and Technology of Japan.

3. Results

The characteristics of the subjects at the baseline (1999) are shown in Table 1. They ranged in age from 18 to 65 years (mean ± standard deviation (SD), 42.4 ± 9.8 years). Overall, 44.6% of the subjects worked shifts.

The prevalence of obesity, hypertension, hyperglycemia, dyslipidemia, and MetS by age class at the baseline (1999) are shown in Table S1. The prevalence (95% CI) of obesity, hypertension, hyperglycemia, dyslipidemia, and MetS were 44.7% (44.2–45.2%), 20.6% (20.2–21.0%), 7.9% (7.6–8.2%), 40.7% (40.2–41.2%), and 16.9% (16.5–17.3%), respectively. The Mantel–Haenszel test for trend indicated that the prevalence of each disease increased with age, and this correlation was statistically significant (all $p < 0.001$).

The associations between lifestyle-related parameters (sleep duration, shift work, actual use of weekly rest days, eating habits (eating in moderation), exercise habits (trying to take every opportunity to walk), alcohol intake, and smoking) at the baseline and new-onset obesity, hypertension, hyperglycemia, dyslipidemia, and MetS, during the seven-year follow-up period are shown in Table 2. The associations between the following lifestyle-related parameters and diseases were significant: sleep duration and new-onset obesity, hypertension, and MetS (all $p < 0.001$); shift work and new-onset hypertension ($p = 0.018$), dyslipidemia ($p = 0.034$), and MetS ($p = 0.029$); actual use of weekly rest days and new-onset obesity ($p = 0.039$), dyslipidemia ($p = 0.016$), and MetS ($p = 0.001$); eating in moderation and new-onset obesity ($p < 0.001$), hypertension ($p = 0.040$), dyslipidemia ($p = 0.051$), and MetS ($p < 0.001$); trying to take every opportunity to walk and new-onset hypertension ($p = 0.020$) and MetS ($p = 0.010$); alcohol intake and new-onset obesity ($p = 0.026$), hypertension ($p < 0.001$),

Table 1
Baseline characteristics of the study population.

	n	%
Age, years		
18–19	32	0.1
20–29	5572	14.2
30–39	8886	22.7
40–49	13,786	35.2
50–59	10,692	27.3
60–65	214	0.5
Sleep duration, hours/day		
<5	15,518	38.5
≥5 and <7	24,478	60.8
≥7	269	0.7
Shift work		
Day work only	16,835	45.4
A night shift every 3 days	889	2.4
A night shift every 4 days	14,504	39.1
A night shift every 5 days	2507	6.8
A night shift every 6 days	2236	6.0
A night shift every 8 days	43	0.1
Others	78	0.2
Actual use of weekly rest days		
Most taken	32,390	83.5
Most not taken	6396	16.5
Eating in moderation		
Always	23,084	59.0
Sometimes	11,216	28.7
Never	4834	12.4
Taking every opportunity to walk		
Yes	25,659	65.6
No	13,438	34.4
Drinking habits		
No drinking	6800	17.4
Intake of pure alcohol <40 g/day	14,463	37.0
Intake of pure alcohol 40–60 g/day	11,315	29.0
Intake of pure alcohol 60–100 g/day	5707	14.6
Intake of pure alcohol ≥100 g/day	773	2.0
Smoking habits		
No or quit smoking	16,609	42.5
Number of cigarettes smoked <20/day	4138	10.6
Number of cigarettes smoked ≥20/day and <40/day	15,480	39.6
Number of cigarettes smoked ≥40/day and <60/day	2675	6.8
Number of cigarettes smoked ≥60/day	204	0.5
Mental complaints		
Irritability		
No	37,876	96.7
Yes	1306	3.3
Reduced concentration		
No	38,710	98.8
Yes	472	1.2
Lethargy		
No	38,452	98.1
Yes	730	1.9

Of 39,182 male employees who underwent a medical checkup at the baseline (1999), those with missing data were excluded before the calculation was performed.

dyslipidemia ($p = 0.009$), and MetS ($p < 0.001$); and smoking and new-onset obesity ($p = 0.010$), hypertension ($p = 0.002$), hyperglycemia ($p < 0.001$), and dyslipidemia ($p < 0.001$).

The associations between the number of high-risk lifestyle parameters present at baseline and new-onset obesity, hypertension, hyperglycemia, dyslipidemia, and MetS during the observation period are shown in Table 3. Obesity ($p = 0.002$), hyperglycemia ($p < 0.001$), dyslipidemia ($p < 0.001$), and MetS ($p < 0.001$) were significantly associated with the number of high-risk lifestyle parameters present at baseline. The hazard ratio for first onset of MetS associated with 0–1 high-risk lifestyle parameters per subject at the baseline was set at 1.00. The hazard ratios associated with the following numbers of high-risk lifestyle parameters were: 1.22 (95% CI 1.15–1.29) for 2–3 of these parameters; 1.43 (1.33–1.54) for 4–7. It was noteworthy that an increase in the number of high-risk lifestyle parameters was associated with an increase in the hazard ratio (Table 3 and Fig. 1).

4. Discussion

The present longitudinal epidemiological study identified the high-risk lifestyle parameters associated with new-onset MetS, and examined the effect of the number of such parameters on the onset of MetS. This study had several advantages over previous investigations. First, it was a longitudinal study that included a sufficiently large sample of subjects, in the order of several tens of thousands. This ensured high epidemiological reliability and allowed conclusions to be drawn about causal associations between variables. Furthermore, the quality of the research was high, as the sampled population consisted of employees of a single organization, thus achieving high sample homogeneity and a high survey participation rate. Second, lifestyle-related parameters – such as sleep duration, shift work, and actual use of days off work – were included among the risk factors for first onset of MetS. The association between MetS onset and these factors has not been widely examined, and hours of sleep per day, and days off work, represent two of the key parameters related to lifestyle outside working hours. Therefore, the present study allowed a comprehensive evaluation of factors related to the daily lifestyles of workers that impacted on MetS. Third, the present study examined the effects of the number of high-risk lifestyle parameters on new-onset MetS. Although the effects of individual lifestyle habits on the onset of MetS have been reported previously, this is the first study to have examined the effects of the number of such lifestyle-related factors (including sleep duration, shift work, and actual use of days off).

4.1. Effect of the number of high-risk lifestyle parameters present on incidence of MetS

The present findings suggested that the number of high-risk lifestyle parameters present in any given individual is able to predict the onset of MetS. Previous studies have examined the effects of the number of lifestyle habits on a range of diseases: Breslow et al. examined the association between seven healthy habits and mortality [20]; Morimoto et al. examined the association between eight healthy habits and the onset of cardiovascular diseases [21]; and Ikeda examined the association between six healthy habits and the prevention of lifestyle-related diseases [22]. Wada et al. conducted a seven-year cohort study of 9554 individuals to analyze the association between the three sets of healthy habits proposed by Breslow [20], Morimoto [21], and Ikeda [22], in the studies referred to above, and the onset of MetS. The authors concluded that Ikeda's six healthy habits (smoking, food intake, alcohol intake, exercise, rest, and enjoyable pursuits) [22] had the greatest impact on the risk of MetS, and that there was a significant association between the number of lifestyle habits engaged in and the onset of MetS. In addition to Ikeda's six health-related parameters, the present study revealed that short sleep duration (<5 h) and shift work were additional risk factors for new-onset MetS [23]. These findings give valuable insight into the lifestyle-related parameters affecting the health of workers. Furthermore, the reliability of the current research is high, due to the larger sample size that was employed, in comparison to the study by Wada et al.

From a public health perspective, the current results are significant for a number of reasons. First, it was found that lifestyle-related parameters, which have not conventionally been regarded as important factors in health guidance strategies, such as sleep and actual use of days off work, could prevent the onset of MetS. The authors believe that health guidance on how to improve daily lifestyle should be provided after identifying the factors that can

Table 2
The associations between lifestyle habits at baseline and new-onset obesity, hypertension, hyperglycemia, dyslipidemia, or MetS (1999–2006).

Lifestyle habits at baseline	Obesity[a]			Hypertension[b]			Hyperglycemia[c]			Dyslipidemia[d]			Metabolic syndrome[e]		
	HR	95% CI	p	HR	95% CI	p	HR	95% CI	p	HR	95% CI	p	HR	95% CI	p
Sleep duration, hours/day			<0.001			<0.001			0.444			0.991			0.001
≥5	1.00			1.00			1.00			1.00			1.00		
<5	1.13	1.06–1.20		0.92	0.88–0.96		0.98	0.91–1.04		1.00	0.95–1.05		1.08	1.03–1.14	
Shift work			0.115			0.018			0.883			0.034			0.029
No shift work	1.00			1.00			1.00			1.00			1.00		
Shift work	1.05	0.99–1.12		1.06	1.01–1.10		1.01	0.94–1.07		1.05	1.00–1.10		1.06	1.01–1.11	
Actual use of weekly rest days			0.039			0.602			0.222			0.016			0.001
Most taken	1.00			1.00			1.00			1.00			1.00		
Most not taken	1.09	1.00–1.19		0.99	0.93–1.04		1.05	0.97–1.14		1.08	1.01–1.14		1.12	1.05–1.19	
Eating in moderation			<0.001			0.040			0.913			0.015			<0.001
Always	1.00			1.00			1.00			1.00			1.00		
Sometimes	1.24	1.15–1.33		1.07	1.01–1.12		1.02	0.95–1.09		1.05	1.00–1.10		1.40	1.33–1.47	
Never	1.30	1.18–1.44		1.03	0.96–1.10		1.00	0.91–1.10		1.10	1.03–1.19		1.53	1.42–1.64	
Taking every opportunity to walk			0.600			0.020			0.061			0.051			0.010
Yes	1.00			1.00			1.00			1.00			1.00		
No	1.02	0.95–1.09		0.95	0.91–0.99		0.94	0.88–1.00		1.05	1.00–1.10		1.07	1.02–1.12	
Drinking			0.026			<0.001			0.338			0.009			<0.001
No drinking	1.00			1.00			1.00			1.00			1.00		
Intake of pure alcohol <60 g/day	0.90	0.83–0.97		1.34	1.26–1.42		0.95	0.88–1.04		0.95	0.89–1.01		0.95	0.89–1.01	
Intake of pure alcohol ≥60 g/day	0.94	0.85–1.05		1.65	1.53–1.78		1.00	0.90–1.11		1.04	0.96–1.12		1.13	1.04–1.22	
Smoking			0.010			0.002			<0.001			<0.001			0.133
No or quit smoking	1.00			1.00			1.00			1.00			1.00		
Number of cigarettes smoked <40/day	0.91	0.85–0.97		0.92	0.88–0.96		1.28	1.20–1.37		1.28	1.22–1.34		0.98	0.94–1.03	
Number of cigarettes smoked ≥40/day	0.92	0.81–1.04		0.94	0.87–1.02		1.43	1.28–1.60		1.34	1.22–1.46		1.08	0.98–1.18	

The analysis was performed after excluding subjects with missing data or subjects with a follow-up period <1 year.
Only male employees were subjected to analysis.
Data for subjects who did not fulfill the criteria for each disease at the baseline were analyzed, and subjects were monitored for new-onset disease over a maximum period of 7 years.
Adjusted factors; obesity: age, hypertension, hyperglycemia, dyslipidemia, mental complaints (irritability, reduced concentration, lethargy).
Hypertension: age, obesity, hyperglycemia, dyslipidemia, mental complaints (irritability, reduced concentration, lethargy).
Hyperglycemia: age, obesity, hypertension, dyslipidemia, mental complaints (irritability, reduced concentration, lethargy).
Dyslipidemia: age, obesity, hypertension, hyperglycemia, mental complaints (irritability, reduced concentration, lethargy).
Metabolic syndrome: age, mental complaints (irritability, reduced concentration, lethargy).
P-value was calculated by Cox proportional hazards model (forced entry method).
Abbreviations: CI, Confidence Interval; HR, Hazard Ratio.
[a] Obesity: BMI ≥25 kg/m^2.
[b] Hypertension: ≥140/90 mmHg, or use of antihypertensive medication.
[c] Hyperglycemia: ≥126 mg/dL, or use of hypoglycemic medication, hypertriglyceridemia: ≥150 mg/dL, Low HDL cholesterol: <40 mg/dL.
[d] Dyslipidemia: hypertriglyceridemia and/or low HDL cholesterol, or use of antilipidemic medication.
[e] Metabolic syndrome: BMI ≥25 kg/m^2 and two or more of the following conditions being met: 1) blood pressure ≥130/85 mmHg, or use of antihypertensive medication, 2) triglyceride ≥150 mg/dL, or HDL cholesterol <40 mg/dL, or use of antilipidemic medication; and 3) glucose ≥110 mg/dL, or use of hypoglycemic medication.

Table 3
The associations between the number of high-risk lifestyle parameters present at baseline and new-onset obesity, hypertension, hyperglycemia, dyslipidemia, or MetS (1999–2006).

Number of high-risk lifestyle parameters at the baseline	Obesity[a]			Hypertension[b]			Hyperglycemia[c]			Dyslipidemia[d]			Metabolic syndrome[e]		
	HR	95% CI	p	HR	95% CI	p	HR	95% CI	p	HR	95% CI	p	HR	95% CI	p
0–1	1.00		0.002	1.00		0.853	1.00		<0.001	1.00		<0.001	1.00		<0.001
2–3	1.11	1.03–1.18		0.99	0.94–1.04		1.18	1.09–1.26		1.26	1.20–1.32		1.22	1.15–1.29	
4–7	1.18	1.07–1.30		0.98	0.92–1.05		1.20	1.09–1.32		1.42	1.32–1.52		1.43	1.33–1.54	

The analysis was performed after excluding subjects with missing data or subjects with a follow-up period of <1 year.
Only male employees were subjected to analysis.
Data for subjects who did not fulfill the criteria for each disease at baseline were analyzed, and subjects were monitored for new-onset disease over a maximum period of 7 years.
Adjusted factors; obesity: age class, hypertension, hyperglycemia, dyslipidemia, mental complaints (irritability, reduced concentration, lethargy).
Hypertension: age class, obesity, hyperglycemia, dyslipidemia, mental complaints (irritability, reduced concentration, lethargy).
Hyperglycemia: age class, obesity, hypertension, dyslipidemia, mental complaints (irritability, reduced concentration, lethargy).
Dyslipidemia: age class, obesity, hypertension, hyperglycemia, mental complaints (irritability, reduced concentration, lethargy).
Metabolic syndrome: age class, mental complaints (irritability, reduced concentration, lethargy).
P-value was calculated by Cox proportional hazards model (forced entry method).
Abbreviations: CI, Confidence Interval; HR, Hazard Ratio.
[a] Obesity: BMI ≥25 kg/m^2.
[b] Hypertension: ≥140/90 mmHg, or use of antihypertensive medication.
[c] Hyperglycemia: ≥126 mg/dL, or use of hypoglycemic medication, hypertriglyceridemia: ≥150 mg/dL, low HDL cholesterol: <40 mg/dL.
[d] Dyslipidemia: hypertriglyceridemia and/or low HDL cholesterol or use of antilipidemic medication.
[e] Metabolic syndrome: BMI ≥25 kg/m^{2,} and two or more of the following conditions being met: 1) blood pressure ≥130/85 mmHg, or use of antihypertensive medication, 2) triglyceride ≥150 mg/dL, or HDL cholesterol <40 mg/dL, or use of antilipidemic medication; and 3) glucose ≥110 mg/dL, or use of hypoglycemic medication.

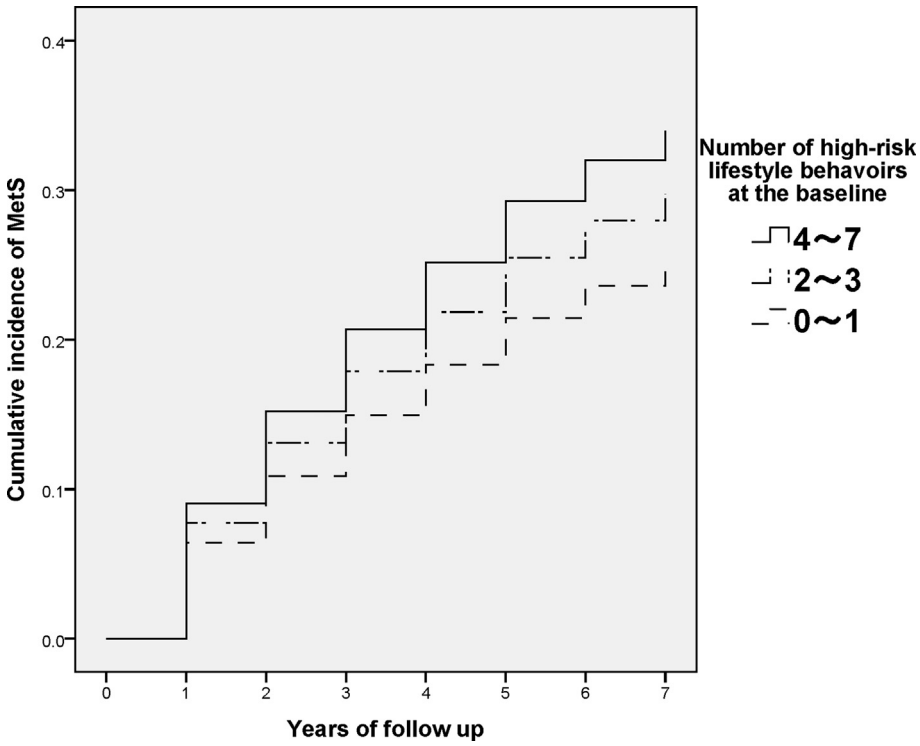


Fig. 1. Kaplan–Meier curves for onset of MetS based on the number of high-risk lifestyle parameters present at the baseline. MetS, metabolic syndrome.

affect workers' health both during and outside working hours, including sleep and days off. Second, the study found that an increase in the number of high-risk lifestyle parameters significantly increased the incidence of MetS onset. Therefore, a comprehensive evaluation of the daily lifestyles of workers may be warranted, with encouragement to decrease the number of risk factors to which they are exposed, and thus helping to reduce the rate of MetS onset. For example, shift work is a risk factor for MetS, and it is important to establish whether shift workers also have other associated risk factors, such as those related to diet or exercise. In addition to the provision of appropriate education and information on the risk of MetS, shift workers should be offered guidance on reducing the number of their risk factors as far as possible, in order to decrease their risk of developing MetS. Currently, Specific Health Checkups and Specific Health Guidance [4] to MetS are widely available in Japan [3], and particular support is provided for patients diagnosed with borderline MetS, in order to reduce or eliminate their risk factors. As part of this health guidance aimed at high-risk patients, it may be effective for counsellors to focus on reducing the number of high-risk lifestyle parameters per patient. Specifically, patients may be able to better understand how they can improve their lifestyle if they are made aware of the number of these high-risk parameters to which they are exposed. Furthermore, improvements in public health require wide dissemination of the relevance of a number of high-risk lifestyle-related factors. It is anticipated that the current findings will contribute to the development of public health strategies (high risk or population approach) aimed at reducing risk factors for MetS in the Japanese population.

4.2. Short sleep duration and MetS

The findings of the present study indicated that short sleep duration was significantly associated with the onset of MetS. Xi et al. conducted a systematic review and meta-analysis of the association between sleep duration and MetS, and identified 12 related studies (10 cross-sectional and two longitudinal) [24]. They then performed a meta-analysis, and found a significant association between short sleep duration and MetS (Odds Ratio 1.27, $p = 0.002$), but no significant association between long sleep duration and MetS. The results of the present study do not contradict the findings of this meta-analysis. However, in comparison with previous studies, the level of epidemiological evidence in the present study was high, due to the large sample size and longitudinal design. There are a number of possible biological mechanisms that might underlie the association between short sleep duration and MetS. It is suggested that short sleep duration increases body weight and changes glucose metabolism [25]. In experimental studies, inadequate sleep significantly changes the main components of energy homeostasis, including glucose tolerance, food craving, and hormones critical to appetite regulation [26]. For instance, sleep restriction could reduce leptin and elevate ghrelin, which regulate satiety and hunger, respectively [27], thereby increasing the cravings for calorie-dense and carbohydrate-rich food.

4.3. Shift work and MetS

The present study indicated that shift work significantly promoted the onset of MetS. In a systematic review on the association between shift work and MetS, Canuto et al. selected 10 studies (three longitudinal, six cross-sectional, and one case–control study) and found a positive association between shift work and MetS in eight of them [28]. However, only three of the studies included sleep duration as a confounding variable, and the findings of the studies were contradictory. Canuto et al. concluded that there was insufficient epidemiological evidence for a correlation between shift work and onset of MetS, and identified three desirable prerequisites for epidemiological studies investigating this association: (1) a sufficiently large sample size; (2) a clear definition of shift work (night work only or rotating shifts); and (3) inclusion of sleep duration as a confounder.

The present longitudinal study satisfied all of these conditions and is one of the highest-quality epidemiological studies to have investigated the association between shift work and MetS onset. The effect of shift work on cardiovascular risk factors may be attributable to a disturbance of circadian rhythm. It is well established that mice with mutation in the Clock gene, which regulates circadian rhythm, tend to develop dyslipidemia and hyperglycemia [29]. Other studies have also suggested that circadian rhythm disturbance may induce MetS [30–32]. Steals has noted that there is interplay between the metabolic pathways associated with lipogenesis and catabolism, involving the glucocorticoid receptor, peroxisome proliferator-activated receptors α and γ, and the Clock gene. Disruption of these pathways leads to onset of MetS [33].

4.4. Actual use of weekly rest days and MetS

The present study demonstrated a significant association between actual use of weekly rest days and new-onset MetS. A number of epidemiological studies have reported an association between working hours and MetS [34–37]. Although resting status is, effectively, the opposite of working status, few studies have investigated the effects of actual use of weekly days off. It has previously been reported that insufficient use of days off significantly promoted MetS onset [13], and these findings were corroborated by the present study. Importantly, the present study had higher data reliability, and resolved the limitations of earlier research, suggesting that weekly days off are not simply days off work but have a direct effect on the health status of workers. However, the present study did not identify the physiological mechanism whereby insufficient use of weekly days off increases the risk of MetS onset.

In addition to weekly days off, generally corresponding to 1 or 2 days of rest from work per week, workers can take longer periods off work as vacations. Gump et al. studied the effects of vacations on health, and followed up 12,338 middle-aged men at high risk of cardiovascular disease in the United States for 9 years [38]. In comparison with men who did not take annual vacations, the relative risk of all-cause mortality for subjects who did take annual vacations was 0.83, and the relative risk of mortality due to cardiovascular disease and non-cardiovascular disease was 0.71 and 0.98, respectively. Gump et al. argued that vacationing could release subjects from continuous or potential stress that might trigger or exacerbate coronary heart disease, and concluded that vacationing may offer opportunities for social contact with family members and friends and for physical activities, thus also allowing physiological recovery. Weekly days off may confer benefits similar to those of vacationing. However, the underlying mechanisms of any such beneficial effects have not yet been clarified, and further epidemiological and physiological data are required.

4.5. Limitations

The present study had a number of limitations. First, as it was retrospective and utilized previously collected data, not all the necessary data were included in the questionnaire. For example, it did not consider data on how leisure time was spent on weekly days off. Therefore, the effects of different types of leisure time could not be analyzed. Furthermore, the data relating to weekly days taken off work were only available for a limited period of time – 1–2 months before the subjects were asked about it. Therefore, it is possible that any health problems resulting from difficulties in taking weekly days off in the mid-term to long-term may not have been evaluated. Second, some of the data, including those pertaining to taking weekly days off work, sleep duration, and smoking habits, were obtained by a self-administered questionnaire and could not be independently verified. A prospective study including these parameters would therefore be useful for corroborating the findings.

5. Conclusions

In conclusion, the present study demonstrated that several lifestyle-related parameters, including sleep duration, shift work, and days off, are risk factors for new-onset MetS, and that an increase in the number of these high-risk lifestyle parameters in any given individual further increases the risk of developing MetS. Therefore, it is important to comprehensively evaluate the daily living habits of workers, and also consider lifestyle-related parameters that have not traditionally been targeted by health guidance, such as sleep duration and days off, in order to devise strategies for the prevention of MetS using multiple approaches.

Acknowledgements

This work was supported by Health Labor Sciences Research Grant from the Japanese Ministry of Health Labor and Welfare [grant number: H25-JUNKANKITOU-IPPAN-007]. None of this funding source had a role in the design, conduct, analysis, or reporting of the study.

Conflict of interest

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: http://dx.doi.org/10.1016/j.sleep.2017.07.027.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.sleep.2017.07.027.

References

[1] World Health Organization. Global status report on noncommunicable diseases 2010. Available from: http://www.who.int/nmh/publications/ncd_report2010/en. [Accessed 13 April, 2017].
[2] Marrero SL, Bloom DE, Adashi EY. Noncommunicable diseases: a global health crisis in a new world order. JAMA 2012;307:2037–8.
[3] Ministry of Health, Labour and Welfare. The second term of National Health Promotion Movement in the twenty first century. Available from: http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryou/kenkou/kenkounippon21/en. [Accessed 13 April, 2017].
[4] Ministry of Health, Labour and Welfare. Specific health checkups and specific health guidance. Available from:: http://www.mhlw.go.jp/english/wp/wp-hw3/dl/2-007.pdf. [Accessed 13 April, 2017].
[5] Cornier MA, Dabelea D, Hernandez TL, et al. The metabolic syndrome. Endocr Rev 2008;29:777–822.
[6] Esposito K, Giugliano D. Healthy lifestyle for metabolic health: no more excuse! Endocrine 2014;46:176–8.
[7] Carroll S, Dudfield M. What is the relationship between exercise and metabolic abnormalities? A review of the metabolic syndrome. Sports Med (Auckland, NZ) 2004;34:371–418.
[8] Djousse L, Padilla H, Nelson TL, et al. Diet and metabolic syndrome. Endocr Metab Immune Disord Drug Targets 2010;10:124–37.
[9] Sun K, Ren M, Liu D, et al. Alcohol consumption and risk of metabolic syndrome: a meta-analysis of prospective studies. Clin Nutr (Edinburgh, Scotland) 2014;33:596–602.
[10] Alkerwi A, Boutsen M, Vaillant M, et al. Alcohol consumption and the prevalence of metabolic syndrome: a meta-analysis of observational studies. Atherosclerosis 2009;204:624–35.
[11] Sun K, Liu J, Ning G. Active smoking and risk of metabolic syndrome: a meta-analysis of prospective studies. PLoS One 2012;7, e47791.
[12] Ikeda M, Kaneita Y, Uchiyama M, et al. Epidemiological study of the associations between sleep complaints and metabolic syndrome in Japan. Sleep Biol Rhythms 2014;12:269–78.
[13] Itani O, Kaneita Y, Ikeda M, et al. Associations of work hours and actual availability of weekly rest days with cardiovascular risk factors. J Occup Health 2013;55:11–20.

[14] New criteria for 'obesity disease' in Japan. Circ J 2002;66:987–92.
[15] Whitworth JA. 2003 World Health Organization (WHO)/International Society of Hypertension (ISH) statement on management of hypertension. J Hypertens 2003;21:1983–92.
[16] Shimamoto K, Ando K, Fujita T, et al. The Japanese Society of Hypertension Guidelines for the Management of Hypertension (JSH 2014). Hypertens Res 2014;37:253–390.
[17] Seino Y, Nanjo K, Tajima N, et al. Report of the committee on the classification and diagnostic criteria of diabetes mellitus. J Diabetes Investig 2010;1:212–28.
[18] Teramoto T, Sasaki J, Ueshima H, et al. Diagnostic criteria for dyslipidemia. Executive summary of Japan Atherosclerosis Society (JAS) guideline for diagnosis and prevention of atherosclerotic cardiovascular diseases for Japanese. J Atheroscler Thromb 2007;14:155–8.
[19] Matsuzawa Y, Ikeda Y, Katayama S, et al. Definition and the diagnostic standard for metabolic syndrome—Committee to Evaluate Diagnostic Standards for Metabolic Syndrome. Nippon Naika Gakkai Zasshi 2005;4:794–809 [in Japanese].
[20] Belloc NB, Breslow L. Relationship of physical health status and health practices. Prev Med 1972;1:409–21.
[21] Morimoto K. Lifestyle and health (1) Physical and mental health. Kosyu-Eisei 1987;51:135–43 [In Japanese].
[22] Ikeda Y. Self control. Tokyo: Practice suppl Ishiyaku Publishers; 1991. p. 119–28 [In Japanese].
[23] Wada T, Fukumoto T, Ito K, et al. Of the three classifications of healthy lifestyle habits, which one is the most closely associated with the prevention of metabolic syndrome in Japanese? Intern Med (Tokyo, Japan) 2009;48:647–55.
[24] Xi B, He D, Zhang M, et al. Short sleep duration predicts risk of metabolic syndrome: a systematic review and meta-analysis. Sleep Med Rev 2014;18:293–7.
[25] Hall MH, Muldoon MF, Jennings JR, et al. Self-reported sleep duration is associated with the metabolic syndrome in midlife adults. Sleep 2008;31:635–43.
[26] Spiegel K, Knutson K, Leproult R, et al. Sleep loss: a novel risk factor for insulin resistance and Type 2 diabetes. J Appl Physiol 2005;99:2008–19.
[27] Taheri S, Lin L, Austin D, et al. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. PLoS Med 2004;1:e62.
[28] Canuto R, Garcez AS, Olinto MT. Metabolic syndrome and shift work: a systematic review. Sleep Med Rev 2013;17:425–31.
[29] Turek FW, Joshu C, Kohsaka A, et al. Obesity and metabolic syndrome in circadian Clock mutant mice. Science 2005;308:1043–5.
[30] Broberger C. Brain regulation of food intake and appetite: molecules and networks. J Intern Med 2005;258:301–27.
[31] Buijs RM, Kreier F. The metabolic syndrome: a brain disease? J Neuroendocrinol 2006;18:715–6.
[32] Wolk R, Somers VK. Sleep and the metabolic syndrome. Exp Physiol 2007;92:67–78.
[33] Staels B. When the Clock stops ticking, metabolic syndrome explodes. Nat Med 2006;12:54–5. discussion 5.
[34] Jeong JU, Jeon MJ, Sakong J. The association between long working hours and the metabolic syndrome: evidences from the 5th Korean National Health and Nutrition Examination Survey of 2010 and 2012. Ann Occup Environ Med 2014;26:53.
[35] Kobayashi T, Suzuki E, Takao S, et al. Long working hours and metabolic syndrome among Japanese men: a cross-sectional study. BMC Public Health 2012;12:395.
[36] Violanti JM, Burchfiel CM, Hartley TA, et al. Atypical work hours and metabolic syndrome among police officers. Arch Environ Occup Health 2009;64:194–201.
[37] Pimenta AM, Bes-Rastrollo M, Sayon-Orea C, et al. Working hours and incidence of metabolic syndrome and its components in a Mediterranean cohort: the SUN project. Eur J Public Health 2015;25:683–8.
[38] Gump BB, Matthews KA. Are vacations good for your health? The 9-year mortality experience after the multiple risk factor intervention trial. Psychosom Med 2000;62:608–12.

Sleep and Biological Rhythms (2018) 16:187–195
https://doi.org/10.1007/s41105-017-0138-2

ORIGINAL ARTICLE

Nightmares and sleep paralysis among the general Japanese population: a nationwide representative survey

Yuichiro Otsuka[1] · Yoshitaka Kaneita[1] · Sachi Nakagome[1] · Maki Jike[1] · Osamu Itani[2] · Takashi Ohida[1]

Received: 25 June 2017 / Accepted: 12 December 2017 / Published online: 27 December 2017

Abstract

The objective of this study was to determine the prevalence of nightmares and sleep paralysis and their associated factors in the general population in Japan. This study was designed as a cross-sectional sampling survey conducted in November 2007. Subjects were selected randomly from among 300 districts throughout Japan. Data from 8099 people (3748 men and 4351 women) were used. Participants completed a self-reported questionnaire on the frequency of nightmares and sleep paralysis in the previous month. In the total sample, the prevalence (95% confidence interval) of nightmares and sleep paralysis was 3.0% (2.6–3.4%) and 2.4% (2.1–2.8%). As for nightmares, multiple logistic analyses revealed that long sleep duration, use of hypnotic medication, difficulty initiating sleep, difficulty maintaining sleep, early morning awakening, excessive daytime sleepiness, and distress had higher odds ratios. As for sleep paralysis, younger age, smoking, difficulty maintaining sleep, and distress had higher odds ratios. This study revealed the prevalence of nightmares and sleep paralysis among the general Japanese population. The results indicate that nightmares and sleep paralysis are induced by high levels of stress and insomnia symptoms in Japanese adults. Therefore, adequate stress-coping skills, sleep education to raise awareness of the importance of maintaining an appropriate circadian rhythm, and effective treatments for insomnia are required.

Keywords Parasomnia · Nightmares · Sleep paralysis · Epidemiology · Prevalence · Japan

Introduction

Sleep disorders are highly prevalent in developed countries [1]. They are known to increase susceptibility to various physical problems [2–4]. The association between sleep and all-cause mortality [5] was reported in a meta-analysis. Sleep disorders include many clinical conditions, such as insomnia, hypersomnia, and sleep apnea [6]. However, most studies examining sleep problems have focused on sleep deprivation, insomnia, and sleep apnea; other types of sleep disturbances have not been adequately investigated.

The term parasomnia encompasses any sleep disorder characterized by abnormal behavioral or physiological activity during sleep [7]. The phases of sleep can be broadly divided into rapid eye movement and non-rapid eye movement. Nightmares and sleep paralysis are considered to be rapid eye movement sleep parasomnias. Usually, dreams occur during rapid eye movement sleep [8]. Occasionally, we remember the content of our dreams. Sometimes, we have nightmares, which we are more inclined to remember. Hence, we feel that we only have nightmares [9]. Recurrent nightmares may cause a lack of recovery from fatigue, as well as sleepiness, depression, and fear of sleep, which could hamper performance in normal daily activities [10]. Sleep paralysis is a phenomenon in which individuals experience inability to move as if someone/something is holding them down, and this causes a heightened fear response. In Japanese, this phenomenon is called *kanashibari* [11]. Sleep paralysis tends to be considered a spiritual phenomenon among different cultures and religions [11, 12]. It is easily imaginable that people experiencing sleep paralysis feel strong fear or anxiety.

According to the previous studies, the onset of nightmares is mostly observed between the ages of 3 and 6 years [13]. It is characterized by repeated nightmares during rapid

✉ Yoshitaka Kaneita
nusmpublichealth@gmail.com

[1] Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, 30-1 Oyaguchi-kamimachi, Itabasi-ku, Tokyo 173-8610, Japan

[2] Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, 1-1 Hasama-machi Idaigaoka, Yufu, Oita 879-5593, Japan

eye movement sleep followed by awakening with intense fear and anxiety [14]. Large-scale epidemiological studies worldwide have reported that the prevalence of nightmares once a week or more is 0.9–6.8% in the general population [14, 15]. However, the prevalence of nightmares in patients with post-traumatic stress disorder (PTSD) or borderline personality disorder is increased to between 49 and 71% [16, 17]. Nightmares have been reported to be associated with mental stress [17], physical problems [18], and suicide risk [19]. Hence, the presence nightmares is a clinically important symptom and may be a treatment target. A large systematic review [20] estimated the lifetime prevalence of sleep paralysis to be 7.6% in the general population, 28.3% in college students, and 31.9% in patients with psychiatric disorders. Sleep paralysis is often experienced by patients with narcolepsy, along with other symptoms, such as hypnagogic/hypnopompic hallucinations, hypersomnolence, and cataplexy [21]. Sleep paralysis can also occur in the absence of narcolepsy; for instance, it may occasionally occur in patients with obstructive sleep apnea, sleep deprivation, and circadian misalignment [22]. Therefore, sleep paralysis is a clinically important symptom and requires early detection.

Nightmares and sleep paralysis have not been routinely studied, because they frequently diminish with age and they involve marked individual differences in severity [10, 11]. Sleep paralysis is more frequently encountered in children and adolescents than in adults [20]. This may be attributable to physical and mental factor affecting the occurrence of nightmares and sleep paralysis. Unlike studies on other sleep disorders, most of the physical and mental effects of nightmares and sleep paralysis are yet to be clarified despite their clinical importance. The only reliable Japanese study on nightmares and sleep paralysis published to date had an exclusively adolescent sample [23]. Another report [24], with young Japanese adults, had a small sample that was not representative of the general population. Therefore, no large-scale epidemiological studies on this topic have been conducted on the general adult population in Japan.

Thus, the prevalence of nightmares and sleep paralysis and their associated factors in Japanese adults are unknown. Elucidating the factors associated with nightmares and sleep paralysis may facilitate the promotion of health and the improvement of the quality of life of those afflicted, as was accomplished with studies on other sleep disorders. In the present study, we used the data collected in the National Health and Nutrition Survey (NHNS) conducted in 2007 by the Japanese Ministry of Health, Labour, and Welfare as data of a nationally representative sample. The objectives of our study were as follows: (1) to determine the prevalence of nightmares and sleep paralysis in Japanese adults and (2) to investigate the factors associated with nightmares and sleep paralysis and compare them to those reported by previous studies.

Methods

Data source

This study was performed using data collected by the NHNS, which was conducted in November 2007 by the Japanese Ministry of Health, Labour, and Welfare. The NHNS was a cross-sectional survey of a nationally representative sample of the general Japanese population. The survey intended to develop measures for national health promotion and included an interview and data pertaining to the physical examination of individuals, nutritional intake and diet, and lifestyle. The survey targeted approximately 18,000 individuals aged 1 year or older in 300 survey district units, randomly selected from the national census unit districts. The number of individuals that finally participated in the survey was 9611. The response rate was 53.4%. The examination protocol and data collection design are fully documented in the NHNS annual reports [25]. Quality assurance in data processing has been described elsewhere [26]. This study conformed to the ethical guidelines of the Ministry of Health, Labour, and Welfare, and the Ministry approved this study. The staff of the public health centers where the physical examinations were conducted obtained oral informed consent from the participants.

Participants

Of the 9611 subjects that participated in the 2007 NHNS, those aged 20 years or older, excluding pregnant or lactating women, with available data for lifestyle factors were selected as study subjects. The number of participants was 3748 men and 4351 women with an average age of 53.5 ± 17.0 and 53.9 ± 17.7 years, respectively.

Measures and definitions

The following questions about nightmares and sleep paralysis were included.

"Have you been awakened by a nightmare during the previous month?" with response choices of "never," "seldom," "sometimes," "often," and "always." Responses of "often," and "always" were considered to be affirmative answers regarding the occurrence of nightmares.

"Have you experienced 'kanashibari' in which you could not move your hands, feet, or body when waking up or falling asleep during the previous month?" with response choices of "yes," "no," and "I don't know." The response "yes" was considered to be an affirmative answer regarding the occurrence of sleep paralysis.

The other questions pertained to the following six issues: (1) basic personal attributes (sex, age); (2) use of hypnotic medication; (3) sleep situation [sleep duration, sufficiency level of rest obtained through sleep, difficulty initiating sleep (DIS), difficulty maintaining sleep (DMS), early morning awakening (EMA), and excessive daytime sleepiness (EDS)]; (4) smoking; (5) alcohol drinking; and (6) distress.

The question for the use of hypnotic medication was phrased as: "Have you taken a sleeping pill at bedtime in the previous month?" with response choices of "never," "seldom," "sometimes," "often," and "always." Responses of "often" and "always" were considered to be affirmative answers to the question.

The six questions for sleep-related factors over the previous month were phrased as follows:

1. How many hours did you sleep at night on average? With response choices of "< 5 h," "≥ 5 h but < 6 h", "≥ 6 h but < 7 h", "≥ 7 h but < 8 h", and "≥ 8 h".
2. How would you rate your amount of sleep? With response choices of "very good", "good", "bad", and "very bad". Responses of "bad" and "very bad" were considered to signify subjective insufficient sleep (SIS).
3. Did you have difficulty falling asleep at night? With response choices of "never", "seldom", "sometimes", "often", and "always". Responses of "often" and "always" were considered to signify DIS [27].
4. Did you wake up during the night after you went to sleep? With response choices of "never", "seldom", "sometimes", "often", and "always". Responses of "often" and "always" were considered to signify DMS [27].
5. Did you wake up too early in the morning and had difficulty getting back to sleep? With response choices of "never", "seldom", "sometimes", "often", and "always". Responses of "often" and "always" were considered to signify EMA [27].
6. Did you feel excessively sleepy during the daytime? With response choices of "never", "seldom", "sometimes", "often", and "always". Responses of "often" and "always" were considered to signify EDS [27].

The questions for smoking and alcohol drinking were phrased as follows:

1. Have you smoked in the previous month? With response choices of "every day", "sometimes", and "no". Responses of "every day" and "sometimes" were considered to be affirmative answers to the question.
2. How many days did you consume alcoholic beverages in the previous week? With response choices of "every day", "5–6 days/per week", "3–4 days/per week", "1–2 days/per week", and "1–3 days/per week", and "none". Responses of "every day", "5–6 days/per week" and "3–4 days/per week" were considered to be affirmative answers to the question.

The question related to distress was, "Did you feel dissatisfied or distressed, or experienced difficulties and stress in the past month?" with response choices of "frequently," "sometimes," "rarely," and "never." The response of "frequently" was considered to signify distress [26].

Statistical analysis

First, the prevalence of nightmares and sleep paralysis and the 95% confidence intervals (CIs) were calculated. Second, the prevalence of both relative to sex and age group was calculated. Third, the associations among nightmares and sleep paralysis during the previous month and factors pertaining to personal attributes, lifestyle, sleep status, and distress were examined. The significance of categorical data, such as the prevalence of nightmares and sleep paralysis, was analyzed using the chi-squared test. Finally, multiple logistic regression analyses were performed to separately examine the factors associated with nightmares or sleep paralysis during the previous month.

Regarding the responses to the questions pertaining to nightmares, 162 questionnaires were excluded, because sex or age was not specified or the answers were inconsistent. The data of the remaining 7657 questionnaires were analyzed. Regarding the responses to the questions pertaining to sleep paralysis, 150 were excluded, because sex or age was not specified or the answers were inconsistent. The data of the remaining 7669 questionnaires were analyzed. We set the level of significance at $P < 0.01$. All analyses were performed using SPSS version 17.0 for Windows (SPSS Inc., Chicago, IL, USA).

Results

Prevalence of nightmares and sleep paralysis in the general Japanese population

Table 1 shows the prevalence of nightmares and sleep paralysis in the general Japanese population. The prevalence (95% CI) of nightmares was 3.0% (2.6–3.4%) in the total sample, 3.1% (2.6–3.7%) among men, and 2.9% (2.4–3.4%) among women. A significant difference was observed between age groups ($P < 0.001$); however, no significant difference was observed between the sexes ($P = 0.489$). Among men, the prevalence of nightmares gradually increased from the sixth to the eighth decade of life. Among women, the prevalence of nightmares was highest in those aged 20–29 years, followed by those aged 70 years or older. The prevalence (95% CI) of

Table 1 Prevalence of nightmares and sleep paralysis in the Japanese general population

Population	Men				Women				
	N	Prevalence (%)	95% CI	*P* value[a]	*N*	Prevalence (%)	95% CI	*P* value[a]	*P* value[b]
Nightmares age group, years									
20–29	323	1.2	0.0–2.5	< 0.001	20–29	394	5.1	2.9–7.3	0.039	0.004
30–39	567	2.3	1.1–3.5		30–39	692	2.9	1.6–4.1		0.509
40–49	574	2.8	1.4–4.1		40–49	624	2.1	1.0–3.2		0.428
50–59	651	1.8	0.8–2.9		50–59	742	2.0	1.0–3.0		0.810
60–69	699	3.7	2.3–5.1		60–69	793	2.5	1.4–3.6		0.182
70+	710	5.6	3.9–7.3		70+	888	3.5	2.3–4.7		0.039
Total	3524	3.1	2.6–3.7		Total	4133	2.9	2.4–3.4		0.489
Sleep paralysis age group, years									
20–29	324	5.2	2.8–7.7	0.002	20–29	396	6.6	4.1–9.0	< 0.001	0.486
30–39	570	1.4	0.4–2.4		30–39	692	3.0	1.8–4.3		0.057
40–49	574	2.1	0.9–3.3		40–49	624	1.9	0.8–3.0		0.836
50–59	655	2.2	1.0–3.2		50–59	744	2.4	1.3–3.5		0.763
60–69	698	1.3	0.5–2.1		60–69	795	2.1	1.1–3.1		0.215
70+	710	2.7	1.5–3.9		70+	887	1.6	0.8–2.4		0.120
Total	3531	2.2	1.7–2.7		Total	4138	2.6	2.1–3.1		0.317

Participants with missing date were excluded from the analysis

CI confidence interval

[a]The *P* value was calculated with the chi-squared test among age groups

[b]The *P* value was calculated with the chi-squared test by sex for each age group

sleep paralysis was 2.4% (2.1–2.8%) in the total sample, 2.2% (1.7–2.7%) among men, and 2.6% (2.1–3.1%) among women. No statistically significant difference was observed between the sexes ($P = 0.292$). However, a statistically significant difference was observed between the age groups ($P < 0.001$). For men and women, sleep paralysis was most prevalent among those in their 20 s.

Association of the prevalence of nightmares and sleep paralysis with personal, lifestyle, sleep status, and distress factors in the Japanese general population

Table 2 shows the association of the prevalence of nightmares and sleep paralysis with personal, lifestyle, sleep status, and distress factors in the Japanese general population. With regard to nightmares, significant associations were observed in almost all variables except for sex, smoking, and alcohol drinking. With regard to sleep paralysis, significant associations were observed in almost all variables except for sex, alcohol drinking, and use of hypnotic medication.

Association of nightmares with personal, lifestyle, sleep status, and distress factors in the Japanese general population

Table 3 shows the adjusted odds ratios of nightmares with personal, lifestyle, sleep status, and distress factors in the Japanese general population. When nightmares were used as the dependent variable in multiple logistic regression analysis, significant associations were observed with sleep duration, use of hypnotic medication, DIS, DMS, EMA, EDS, and distress. Particularly, a substantially higher odds ratio was observed for the use of hypnotic medication, DMS, and distress. In addition, no significant associations were observed with the frequency of sleep paralysis.

Association of sleep paralysis with personal, lifestyle, sleep status, and distress factors in the Japanese general population

Table 4 shows the adjusted odds ratios of sleep paralysis with personal, lifestyle, sleep status, and distress factors in

Table 2 Association of the prevalence of nightmares and sleep paralysis with personal, lifestyle, sleep status, and distress factors in the Japanese general population

Variables	Nightmares				Sleep paralysis			
	N	%	95% CI	*P* value	*N*	%	95% CI	*P* value
Sex								
Men	3524	3.1	2.6–3.7	0.489	3531	2.2	1.7–2.7	0.292
Women	4133	2.9	2.4–3.4		4138	2.6	2.1–3.1	
Age groupy								
20–29	717	3.3	2.0–4.7	0.002	720	6.0	4.2–7.7	< 0.001
30–39	1259	2.6	1.7–3.5		1262	2.3	1.5–3.1	
40–49	1198	2.4	1.5–3.3		1198	2.0	1.2–2.8	
50–59	1393	1.9	1.2–2.7		1399	2.3	1.5–3.1	
60–69	1492	3.1	2.2–4.0		1493	1.7	1.1–2.4	
70+	1598	4.9	3.4–5.5		1597	2.1	1.4–2.8	
Smoking								
No	5714	3.0	2.6–3.5	0.858	5724	2.1	1.7–2.5	0.002
Yes	1934	2.9	2.2–3.7		1936	3.4	2.6–4.2	
Drinking alcohol								
No	5260	2.8	2.4–3.3	0.146	5269	2.4	2.0–2.8	0.807
Yes	2394	3.4	2.7–4.2		2397	2.5	1.9–3.1	
Sleep duration (h)								
< 5	484	4.5	2.7–6.4	< 0.001	483	5.2	3.2–7.2	< 0.001
≥ 5 but < 6	1690	2.9	2.1–3.7		1695	2.8	2.0–3.6	
≥ 6 but < 7	2896	2.1	1.6–2.6		2901	2.1	1.6–2.6	
≥ 7 but < 8	1606	3.7	2.8–4.7		1607	1.7	1.1–2.3	
≥ 8	979	4.0	2.8–5.2		981	2.7	1.6–3.7	
Hypnotic								
No	7152	2.4	2.0–2.7	< 0.001	7163	2.4	2.0–2.7	0.271
Yes	504	11.7	8.9–14.5		505	3.2	1.6–4.7	
Difficulty initiating sleep								
No	6725	2.2	1.9–2.6	< 0.001	6737	2.2	1.8–2.5	< 0.001
Yes	931	8.8	7.0–10.6		931	4.3	3.0–5.6	
Difficulty maintaining sleep								
No	6327	1.8	1.5–2.1	< 0.001	6340	2.1	1.7–2.4	< 0.001
Yes	1327	8.6	7.1–10.1		1326	4.2	3.1–5.3	
Early morning awakening								
No	6123	2.1	1.8–2.5	< 0.001	6134	2.2	1.9–2.6	0.026
Yes	1524	6.4	5.2–7.7		1524	3.2	2.3–4.1	
Excessive daytime sleepiness								
No	7441	2.8	2.4–3.1	< 0.001	7452	2.4	2.0–2.7	0.010
Yes	214	11.2	7.0–15.5		215	5.1	2.1–8.1	
Subjective insufficient sleep								
No	5972	2.4	2.0–2.8	< 0.001	5984	1.9	1.6–2.3	< 0.001
Yes	1683	5.1	4.0–6.1		1683	4.3	3.3–5.2	
Distress								
No	6380	2.2	1.8–2.5	< 0.001	6390	1.9	1.5–2.2	< 0.001
Yes	1274	7.1	5.7–8.6		1276	5.3	4.0–6.5	
Sleep paralysis/nightmares								
No	7467	2.9	2.5–3.2	< 0.001	7423	2.3	1.9–2.6	< 0.001
Yes	186	8.6	4.5–12.7		230	7.0	3.6–10.3	

The presence of excessive daytime sleepiness was defined as an Epworth Sleepiness Scale score of P11

Participants with missing data were excluded from the analysis

The *P* value was calculated by the chi-squared test

CI confidence interval

Table 3 Association of nightmares with personal, lifestyle, sleep status, and distress factors in the Japanese general population

Variables	Odds ratios			
	N	AOR	95% CI	*P* value
Sleep duration (h)				
< 5	482	1.18	0.68–2.05	< 0.001
≥ 5 but < 6	1684	2.55	1.43–4.55	
≥ 6 but < 7	2885	1.00		
≥ 7 but < 8	1600	1.93	1.04–3.56	
≥ 8	970	1.21	0.69–2.10	
Hypnotic				
No	7120	1.00		< 0.001
Yes	501	2.78	1.91–4.05	
Difficulty initiating sleep				
No	6695	1.00		0.005
Yes	926	1.64	1.16–2.32	
Difficulty maintaining sleep				
No	6301	1.00		< 0.001
Yes	1320	2.63	1.91–3.62	
Early morning awakening				
No	6102	1.00		< 0.001
Yes	1519	1.83	1.34–2.51	
Excessive daytime sleepiness				
No	7408	1.00		0.003
Yes	213	2.12	1.29–3.49	
Distress				
No	6351	1.00		< 0.001
Yes	1270	2.40	1.74–3.30	

The presence of excessive daytime sleepiness was defined as an Epworth Sleepiness Scale score of P11

Participants with missing data were excluded from the analysis

The *P* value was calculated with multiple logistic regression analysis

Adjusted for sex, age group, smoking, drinking alcohol, subjective insufficient sleep, and sleep paralysis

CI confidence interval, *AOR* adjusted odds ratio

Table 4 Association of sleep paralysis with personal, lifestyle, sleep status, and distress factors in the Japanese general population

Variables	Odds ratios			
	N	AOR	95% CI	*P* value
Age group, years				
20–29	715	3.68	2.15–6.30	< 0.001
30–39	1258	1.23	0.70–2.15	
40–49	1194	1.00		
50–59	1385	1.35	0.78–2.34	
60–69	1481	1.17	0.65–2.10	
70+	1588	1.36	0.75–2.46	
Smoking				
No	5691	1.00		0.004
Yes	1930	1.66	1.17–2.35	
Difficulty maintaining sleep				
Never/seldom/sometimes	6301	1.00		0.006
Often/always	1320	1.71	1.17–2.51	
Distress				
No	6351	1.00		< 0.001
Yes	1270	2.02	1.43–2.86	

CI confidence interval, *AOR* adjusted odds ratio

The presence of excessive daytime sleepiness was defined as an Epworth Sleepiness Scale score of P11

Participants with missing data were excluded from the analysis

The *P* value was calculated with multiple logistic regression analysis

Adjusted for sex, drinking alcohol, sleep duration, hypnotic, difficulty maintaining sleep, early morning awakening, excessive daytime sleepiness, subjective insufficient sleep, and nightmares

the Japanese general population. When sleep paralysis during the previous month was used as the dependent variable in the multiple logistic regression analysis, significant associations were observed with sex, age group, smoking, DMS, and distress. Particularly, a higher odds ratio was observed for distress. In addition, a higher odds ratio was observed for being in the third compared to being in the fifth decade of life. No significant associations were observed in the frequency of nightmares.

Discussion

This study is the first to reveal the prevalence of nightmares and sleep paralysis in the general Japanese population. This survey has two main strengths. First, it included a nationwide sample. Second, the survey sample comprised of randomly selected individuals from the general Japanese population without regard to age, sex, place of residence, or occupation. Therefore, the participants were fairly representative of the general Japanese population. In addition, our results revealed that distress was associated with nightmares and sleep paralysis. This result has not been highlighted in previous sleep surveys conducted anywhere in the world.

The prevalence of nightmares in the previous month in our study was approximately 3%, which was similar to that reported by studies conducted in other countries [15, 28]. The prevalence of sleep paralysis in both men and women was slightly higher than 2%. Although it was slightly higher among women than men, no significant difference was observed. Notably, these values appear to be lower than those reported by studies conducted in other countries [20]. However, as the survey methods used in this and other studies differed, simple comparisons should be avoided. For example, in the present study, the frequency of nightmares or sleep paralysis during the previous month was investigated. In contrast, other studies investigated the lifetime prevalence or the prevalence of nightmares or sleep paralysis during the

previous week. According to a study on Japanese junior and high school students that utilized a survey method similar to that employed in this study, the prevalence of nightmares and sleep paralysis was 35.2 and 8.3% [23]. Furthermore, the prevalence of both disorders was distinctly higher in junior and high school students than in adults. It was suggested that the prevalence of nightmares and sleep paralysis was significantly affected by age.

The present study indicated that nightmares and sleep paralysis were correlated with distress. According to the previous studies on the association between nightmares and stress, as stress increased, the frequency of nightmares [29] increased concomitantly. The findings of these reports suggested that there is continuity between nightmares and waking thoughts [30]. In addition, nightmares have been associated with stress-related diseases such as anxiety [31], depression [32], suicide [33], and with neurotic characteristics [34]. Furthermore, recurrent nightmares can lead to inadequate daytime activity and functioning, which in turn increase stress, resulting in a vicious cycle [35]. It has been reported that the prevalence of nightmares is very high in patients with PTSD [16]. Picchioni et al. [36] reported that nightmares were positively associated with adequate social support and stress-coping strategies, and proposed the hypothesis that nightmares themselves serve as coping mechanisms to alleviate stress.

Similarly, several studies have also reported associations between sleep paralysis and stress [37–39]. The cortex of the frontal lobe, the cingulate convolution, hypothalamus, hippocampus, and amygdala in the limbic system has been reported to be associated with the evaluation and cognition of stress [40]. According to Phan et al. [41], fear processing implicates amygdala hyperactivity. Maquet et al. [42] found functional interactions between the amygdala and the brainstem during rapid eye movement sleep. Owing to the fact that sleep paralysis is characterized by augmented rapid eye movement sleep activity and significant fear, it is possible that patients with sleep paralysis have an overactive amygdala. Therefore, stress and sleep paralysis may be closely associated.

DMS was a factor associated with both nightmares and sleep paralysis in the present study. Some previous studies have argued for a strong association between nightmares and insomnia [43, 44]. Because nightmares frequently appears in the latter half of the sleep cycle, DMS and EMA could result; subsequently, resuming sleep could be difficult. Among the few studies on the association between sleep paralysis and insomnia, Miyashita et al. [45] conducted a study using a forced awakening method; the sleep of healthy subjects was interrupted for certain periods of time during the non-rapid eye movement phase of nocturnal sleep. The results indicated that sleep paralysis was experimentally induced during the sleep-onset rapid eye movement period at a high rate. This suggests that frequent awakening during sleep may cause sleep-onset rapid eye movement sleep and induce sleep paralysis even among healthy individuals. In addition, sleep-onset rapid eye movement sleep is reportedly prone to occur, while the amplitude of the circadian rhythm is reduced, such as during shifted sleep periods [46], sleep in inversed sleep-waking cycles [47], and in patients with depression [48]. Frequent awakening during sleep decreases the amplitude of the circadian rhythm, which induces sleep-onset rapid eye movement sleep. Consequently, this can induce sleep paralysis. In the present study, nightmares were associated with both DMS and EMA and with DIS and EDS. That is, it was suggested that nightmares have stronger associations with insomnia symptoms than with sleep paralysis.

This study indicated some differences in the factors associated with nightmares and sleep paralysis. First, the associations with age differed. There was no recognized association between nightmares and age. However, this study suggested that sleep paralysis develops in adolescence and that its appearance rate decreases in people in their 30 s and onwards. In contrast, nightmares were frequently observed in young children, according to a previous study [34]. Fukuda et al. [11] found that many subjects experienced the first episode of sleep paralysis in the latter half of adolescence, and inferred that their circadian rhythm was more predisposed to be disrupted because of preparation for the university entrance examinations in Japan. This was considered to be one of the background factors for sleep paralysis. The age of new-onset sleep paralysis is adolescence [11]. These ages corresponded to the later phases of development of secondary sexual characteristics. Second, the associations with sex differed. In our study, nightmares had no association with sex, but the odds ratio of sleep paralysis for women was higher than that for men. Specifically, Ohayon et al. [49] reported that the risk of nightmares was higher for women than for men, which did not coincide with the findings of the present study. However, Munezawa et al. [23] reported a higher odds ratio for boys than for girls. In this regard, further epidemiological studies are required. Third, the associations with smoking differed. In the present study, while nightmares were not associated with smoking, sleep paralysis was. Nicotine contained in tobacco stimulates the central nervous system, promoting wakefulness, resulting in prolonged sleep latency and reduced total sleep duration and rapid eye movement sleep [50]. Zhang et al. [51] reported that the sleep of habitual smokers tends to become shallower. Degraded sleep quality caused by smoking may induce sleep paralysis. Since the action of nicotine absorbed by smoking can last approximately 1 h, smoking before sleeping may induce DIS. A future investigation of the presence/absence of smoking soon before sleep may clarify the association between nightmares and smoking. Fourth, associations with the use of hypnotic medications

were different. In our study, hypnotics were associated with nightmares, but not with sleep paralysis. Hypnotics, along with beta-adrenergic antagonists and alpha-adrenergic agonists, were reported to increase the prevalence of nightmares [52]. Nevertheless, their pharmacological mechanism has not been elucidated. Insomnia treatment is frequently practiced with poor medication adherence and compliance [53]. Such inappropriate use of hypnotic medication may worsen the sleep rhythm, which in turn, could easily induce nightmares.

This study had some limitations. First, the study was a cross-sectional survey, and this limits any consideration of causality. Second, as the data were collected through a self-administered questionnaire, there could have been information bias present. Third, the definitions of nightmares and sleep paralysis were not provided and could have been interpreted differently by each participant. The limited space in the questionnaires did not allow us to expand on the definitions of nightmares and sleep paralysis; therefore, the participants could have understood the meaning of nightmares and sleep paralysis only on the basis of the information in the questions related to these phenomena. Particularly, sleep paralysis may have been confused with other phenomena such as sleep inactivity or difficulty in waking up. Sleep paralysis would be difficult to distinguish from such phenomena on the basis of the explanation provided in the questionnaire. In addition, sleep paralysis is often observed in patients with narcolepsy [22]. Nightmares may also have been confused with bad dreams. Therefore, this study may not necessarily represent the actual prevalence of pathological nightmares and sleep paralysis. Patients with pathological nightmares and sleep paralysis are difficult to select based only on epidemiological data. However, a more accurate diagnosis may be possible by providing highly detailed information about the symptoms to the participants. In future studies, we would like to resolve these problems to more accurately determine the prevalence of nightmares and sleep paralysis. Fourth, sociodemographic data, such as those relating to occupation, academic achievement, and income, were not collected in the present study. However, these factors may affect sleep habits, and their effects should be surveyed to improve the validity of future studies. Fifth, the influence of drugs other than hypnotics was not evaluated. Therefore, a future study should be conducted to evaluate the effects of these drugs on nightmares.

Conclusion

This study was a nationwide survey of nightmares and sleep paralysis in the general Japanese adult population. Nightmares and sleep paralysis were associated with high levels of stress and insomnia symptoms in Japanese adults. Therefore, adequate stress-coping skills, sleep education to raise awareness about the importance of maintaining an appropriate circadian rhythm, and effective treatments for insomnia are required.

Compliance with ethical standards

Conflict of interest The authors of this study received no financial support and declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

Ethical approval This study conformed to the ethical guidelines of the Ministry of Health, Labour, and Welfare and the Ministry approved this study.

Informed consent The staff of the public health centers where the physical examinations were conducted obtained oral informed consent from the participants.

References

1. Kim K, Uchiyama M, Okawa M, Liu X, Ogihara R. An epidemiological study of insomnia among the Japanese general population. Sleep. 2000;23:41–7.
2. Gangwisch JE, Heymsfield SB, Boden-Albala B, Buijs RM, Kreier F, Pickering TG, et al. Short sleep duration as a risk factor for hypertension: analyses of the first National Health and Nutrition Examination Survey. Hypertension. 2006;47:833–9.
3. Mallon L, Broman JE, Hetta J. High incidence of diabetes in men with sleep complaints or short sleep duration: a 12-year follow-up study of a middle-aged population. Diabetes Care. 2005;28:2762–7.
4. Troxel WM, Buysse DJ, Matthews KA, Kip KE, Strollo PJ, Hall M, et al. Sleep symptoms predict the development of the metabolic syndrome. Sleep. 2010;33:1633–40.
5. Cappuccio FP, D'Elia L, Strazzullo P, Miller MA. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep. 2010;33:585–92.
6. Tanigawa T, Tachibana N, Yamagishi K, Muraki I, Kudo M, Ohira T, et al. Relationship between sleep-disordered breathing and blood pressure levels in community-based samples of Japanese men. Hypertens Res. 2004;27:479–84.
7. Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. Chest. 2014;146:1387–94.
8. Aserinsky E, Kleitman N. Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. Science (New York, NY). 1953;118:273–4.
9. Foulkes D. Home and laboratory dreams: four empirical studies and a conceptual reevaluation. Sleep. 1979;2:233–51.
10. Leung AK, Robson WL. Nightmares. J Natl Med Assoc. 1993;85:233–5.
11. Fukuda K, Miyasita A, Inugami M, Ishihara K. High prevalence of isolated sleep paralysis: kanashibari phenomenon in Japan. Sleep. 1987;10:279–86.
12. Dahlitz M, Parkes JD. Sleep paralysis. Lancet. 1993;341:406–7.
13. Liddon SC. Sleep paralysis and hypnagogic hallucinations. Their relationship to the nightmare. Arch Gen Psychiatry. 1967;17:88–96.

14. Nielsen T, Levin R. Nightmares: a new neurocognitive model. Sleep Med Rev. 2007;11:295–310.
15. Li SX, Zhang B, Li AM, Wing YK. Prevalence and correlates of frequent nightmares: a community-based 2-phase study. Sleep. 2010;33:774–80.
16. Leskin GA, Woodward SH, Young HE, Sheikh JI. Effects of comorbid diagnoses on sleep disturbance in PTSD. J Psychiatric Res. 2002;36:449–52.
17. Semiz UB, Basoglu C, Ebrinc S, Cetin M. Nightmare disorder, dream anxiety, and subjective sleep quality in patients with borderline personality disorder. Psychiatry Clin Neurosci. 2008;62:48–55.
18. Krakow B. Nightmare complaints in treatment-seeking patients in clinical sleep medicine settings: diagnostic and treatment implications. Sleep. 2006;29:1313–9.
19. Pigeon WR, Pinquart M, Conner K. Meta-analysis of sleep disturbance and suicidal thoughts and behaviors. The Journal of clinical psychiatry. 2012;73:e1160–e7.
20. Sharpless BA, Barber JP. Lifetime prevalence rates of sleep paralysis: a systematic review. Sleep Med Rev. 2011;15:311–5.
21. American Academy of Sleep M. International classification of sleep disorders: diagnostic and coding manual. 2005. Westchester, IL.2.
22. Snyder S. Isolated sleep paralysis after rapid time-zone change ('jet-lag') syndrome. Chronobiologia. 1983;10:377–9.
23. Munezawa T, Kaneita Y, Osaki Y, Kanda H, Ohtsu T, Suzuki H, et al. Nightmare and sleep paralysis among Japanese adolescents: a nationwide representative survey. Sleep Med. 2010;12:56–64.
24. Fukuda K, Inamatsu N, Kuroiwa M, Miyasita A. Personality of healthy young adults with sleep paralysis. Perceptual Motor Skills. 1991;73:955–62.
25. Nakamura M, Kiyama M, Kitamura A, Ishikawa Y, Sato S, Noda H, et al. Revised system to evaluate measurement of blood chemistry data from the Japanese National Health and Nutrition Survey and Prefectural Health and Nutrition Surveys. J Epidemiol Jpn Epidemiol Assoc. 2013;23:28–34.
26. Furihata R, Kaneita Y, Jike M, Ohida T, Uchiyama M. Napping and associated factors: a Japanese nationwide general population survey. Sleep Med. 2016;20:72–9.
27. Foley DJ, Monjan AA, Brown SL, Simonsick EM, Wallace RB, Blazer DG. Sleep complaints among elderly persons: an epidemiologic study of three communities. Sleep. 1995;18:425–32.
28. Tanskanen A, Tuomilehto J, Viinamaki H, Vartiainen E, Lehtonen J, Puska P. Nightmares as predictors of suicide. Sleep. 2001;24:844–7.
29. Köthe M, Pietrowsky R. Behavioral effects of nightmares and their correlations to personality patterns. Dreaming. 2001;11:43.
30. Domhoff GW. A new neurocognitive theory of dreams. Dreaming. 2001;11:13–33.
31. Nielsen TA, Laberge L, Paquet J, Tremblay RE, Vitaro F, Montplaisir J. Development of disturbing dreams during adolescence and their relation to anxiety symptoms. Sleep. 2000;23:727–36.
32. Levin R, Fireman G. Nightmare prevalence, nightmare distress, and self-reported psychological disturbance. Sleep. 2002;25:205–12.
33. Kales A, Soldatos CR, Caldwell AB, Charney DS, Kales JD, Markel D, et al. Nightmares: clinical characteristics and personality patterns. Am J Psychiatry. 1980;137:1197–201.
34. Blagrove M, Farmer L, Williams E. The relationship of nightmare frequency and nightmare distress to well-being. J Sleep Res. 2004;13:129–36.
35. Sjostrom N, Waern M, Hetta J. Nightmares and sleep disturbances in relation to suicidality in suicide attempters. Sleep. 2007;30:91–5.
36. Picchioni D, Goeltzenleucher B, Green DN, Convento MJ, Crittenden R, Hallgren M, et al. Nightmares as a coping mechanism for stress. Dreaming. 2002;12:155–69.
37. Bell CC, Shakoor B, Thompson B, Dew D, Hughley E, Mays R, et al. Prevalence of isolated sleep paralysis in black subjects. J Natl Med Assoc. 1984;76:501–8.
38. Ohayon MM, Zulley J, Guilleminault C, Smirne S. Prevalence and pathologic associations of sleep paralysis in the general population. Neurology. 1999;52:1194–200.
39. Simard V, Nielsen TA. Sleep paralysis-associated sensed presence as a possible manifestation of social anxiety. Dreaming. 2005;15:245.
40. Davidson RJ, Putnam KM, Larson CL. Dysfunction in the neural circuitry of emotion regulation—a possible prelude to violence. Science. 2000;289:591–4.
41. Phan KL, Wager T, Taylor SF, Liberzon I. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. NeuroImage. 2002;16:331–48.
42. Maquet P, Peters J, Aerts J, Delfiore G, Degueldre C, Luxen A, et al. Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. Nature. 1996;383:163–6.
43. Sandman N, Valli K, Kronholm E, Revonsuo A, Laatikainen T, Paunio T. Nightmares: risk factors among the Finnish general adult population. Sleep. 2014;38:507–14.
44. Nakajima S, Inoue Y, Sasai T, Okajima I, Komada Y, Nomura T, et al. Impact of frequency of nightmares comorbid with insomnia on depression in Japanese rural community residents: a cross-sectional study. Sleep Med. 2014;15:371–4.
45. Miyasita A, Fukuda K, Inugami M. Effects of sleep interruption on REM-NREM cycle in nocturnal human sleep. Electroencephalogr Clin Neurophysiol. 1989;73:107–16.
46. Decoster F, Foret J. Sleep onset and first cycle of sleep in human subjects: change with time of day. Electroencephalogr Clin Neurophysiol. 1979;46:531–7.
47. Weitzman ED, Kripke DF, Goldmacher D, McGregor P, Nogeire C. Acute reversal of the sleep-waking cycle in man. Effect on sleep stage patterns. Arch Neurol. 1970;22:483–9.
48. Schulz H, Lund R, Cording C, Dirlich G. Bimodal distribution of REM sleep latencies in depression. Biol Psychiatry. 1979;14:595–600.
49. Ohayon MM, Morselli PL, Guilleminault C. Prevalence of nightmares and their relationship to psychopathology and daytime functioning in insomnia subjects. Sleep. 1997;20:340–8.
50. Phillips BA, Danner FJ. Cigarette smoking and sleep disturbance. Arch Internal Med. 1995;155:734–7.
51. Zhang L, Samet J, Caffo B, Punjabi NM. Cigarette smoking and nocturnal sleep architecture. Am J Epidemiol. 2006;164:529–37.
52. Thompson DF, Pierce DR. Drug-induced nightmares. Ann Pharmacother. 1999;33:93–8.
53. Wagner J, Wagner ML. Non-benzodiazepines for the treatment of insomnia. Sleep Med Rev. 2003;4:551–81.

Psychological and Educational Interventions for Atopic Dermatitis in Adults: A Systematic Review and Meta-Analysis

Kayoko Hashimoto, Yusuke Ogawa, Nozomi Takeshima and Toshi A. Furukawa

Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine, School of Public Health, Kyoto, Japan

Psychological and educational interventions are valuable adjuncts in the management of adult atopic dermatitis. We performed a systematic review and meta-analysis of randomised controlled trials (RCTs) of the efficacy of these interventions. Twelve articles published between 1986 and 2013 were identified through electronic searches. The methodological quality was assessed according to the *Cochrane Handbook for Systematic Reviews of Interventions*, version 5.1.1 (Higgins & Green, 2015). A random-effects model was used to estimate the standardised mean difference (SMD). No significant difference was found in eczema severity determined in three RCTs (124 participants; SMD, -0.29; 95% CI [-0.64, 0.07]) and dropout rate in five RCTs (198 participants; relative risk, 0.66; 95% CI [0.20, 2.17]). Education via online video was significantly superior to handouts in ameliorating eczema severity in one RCT (80 participants). We conclude that, rather than a combination of these interventions with conventional therapy being of no value, the data did not have sufficient power to provide evidence-based conclusions.

■ *Keywords:* atopic dermatitis, meta-analysis, psychological and educational interventions, systematic review

Atopic dermatitis (AD) is a chronically relapsing skin disorder with a significant social and financial burden (Arkwright, Stafford, & Sharma, 2014; Bieber, 2008; Leung & Guttman-Yassky, 2014; Weidinger & Novak, 2015). AD commonly presents before the age of 5 years, with achievement of remission before adolescence in 70% of patients. However, severe cases may persist until adulthood. In addition, adult-onset AD differs in morphology and distribution of eruptions from childhood AD (Bannister & Freeman, 2000; Garmhausen et al., 2013; Ozkaya, 2005). It is currently estimated that 10–20% of children and 1–3% of adults in developed countries are affected by the disorder (Asher et al., 2006; Silverberg & Hanifin, 2013). The overall prevalence of AD has increased by two- to threefold over the last 3 decades in industrialised countries (Leung & Bieber, 2003; Odhiambo, Williams, Clayton, Robertson, & Asher, 2009).

Intractable itching, a hallmark of the disease, causes sleep disturbance and associated daytime tiredness, a lack of concentration, anxiety, depression, feelings of

Address for correspondence: Kayoko Hashimoto, MD, PhD, MPH, Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine, School of Public Health, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606–8501, Japan. E-mail: hashykay@gmail.com

Behaviour Change | Volume **34** | Number **1** | 2017 | pp. 48–65 | © The Author(s) 2017
doi 10.1017/bec.2017.4

unattractiveness, and financial strain (Carroll et al., 2005; Garmhausen et al., 2013; Linnet & Jemec, 2001; Slattery et al., 2011). Stressful life events can aggravate skin symptoms through several possible mechanisms, such as alteration of pruritus perception and modulation of the immune response (Peters et al., 2014). Management of AD is thus complex and often requires a well-planned, multidisciplinary approach for optimal care to break the vicious cycle of itching and scratching (Arkwright et al., 2013; Schneider et al., 2012). In general, the choice of therapy depends on the severity of the eczema and the age of the patient, including identification and avoidance of exacerbating factors, skin hydration, and topical and systemic remedies. While topical corticosteroids and emollients are the first-line treatment for AD, these are not always effective. Furthermore, an unpleasant sensation from greasy ointment, steroid phobia related to fear of adverse effects, and repeated hospital visits lead to frustration, disappointment, and poor adherence to treatment (Charrman, Morris, & Williams, 2000). Thus, AD represents a major public health problem because of the significant morbidity, impaired quality of life (QOL), and healthcare costs.

Two previous meta-analyses (Chida, Steptoe, Hirakawa, Sudo, & Kubo, 2007; Ersser et al., 2014) suggested that psychological and educational interventions represented a critical option for successful management of childhood and adult AD. Because of differences in the course of disease and cognition between children and adults (Arkwright et al., 2014; Bieber, 2008; Friedman, Nessler, Cycowicz, & Horton, 2009; Leung & Guttman-Yassky, 2014; Weidinger & Novak, 2015), we assumed that the efficacy of these interventions in the management of AD should be verified separately. In this analysis, we extended our studies and aimed to systematically review randomised control trials (RCTs) to determine the effectiveness of psychological and educational interventions in addition to conventional treatment modalities in the management of adult AD.

Methods

We conducted a review of RCTs in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009). All of the objective, inclusion and exclusion criteria, primary and secondary outcomes, and methods of synthesis were prespecified. Approval from the Institutional Review Board was withheld as the data review had no direct subject involvement.

Inclusion Criteria

We included RCTs that compared psychological and educational interventions in addition to conventional treatment (intervention group) with conventional treatment only (control group) in AD patients aged 16 years or older irrespective of publication status, date, country, or language. In this review, conventional treatment was primarily pharmacotherapy. Quasi-randomised studies were excluded.

Interventions

The psychological interventions included supportive, cognitive, behavioural, and psychodynamic psychotherapies, cognitive-behavioural therapy, and physical training such as progressive muscle relaxation. Lectures, audiotapes, books, videotapes, and question-and-answer sessions for the educational interventions contained information on the disease, treatment options, and strategies for management and prevention. We

considered simple provision of handouts as a component of conventional treatment. Both individual and group formats of the interventions were conducted. A trial in which psychological intervention followed educational intervention was classified as psychological, as the components of psychological interventions are often educational (Chida et al., 2007; Ersser et al., 2014). We included interventions of any duration and those remotely administered through telephone calls or video.

Outcome Measures

Primary outcomes were the severity of AD based on previously published named scoring systems, and the dropout rate defined as the proportion of patients who could not complete the required program for any reason. Secondary outcomes included QOL, anxiety, depression, sleep disturbance, use of medical resources such as topical corticosteroids and emollients, and financial cost.

Identification of Trials

We electronically searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, PsycINFO, Scopus, and CINAHL up to May 2014 with appropriate variations of the following terms: [atopic dermatitis; eczema; neurodermatitis; Besnier prurigo] (diagnosis or keyword) and [psychotherapy; behavior; behavior, cognitive; training; counseling; relaxation; program; education] (intervention). We also searched clinical trial databases such as the International Standard Randomized Controlled Trial Number, the U.S. National Institutes of Health ClinicalTrials.gov, and the World Health Organization International Clinical Trial Registry platform. A manual search of major dermatology and medical journals from reference lists of the articles was also conducted. No language restriction was imposed on the search. When necessary, the relevant authors were contacted to supplement any incomplete data in the original manuscript. Two authors (UO and KH) independently reviewed the titles and abstracts of the references retrieved from the searches and selected all potentially relevant trials. Disagreement in the trial selection was discussed with another author (NT) to reach consensus.

Data Extraction and Management

Using data extraction forms, as exemplified in Furukawa, McGuire, and Barbui (2002), two authors (UO and KH) extracted the following data from the original reports: the first author's last name, publication year, characteristics of the participants (sample size, age, sex, and baseline disease severity), characteristics of the interventions (type of intervention, number and length of session, and duration of follow-up), and outcome measurements.

Assessment of the Risk of Bias

The same authors (UO and KH) independently assessed the methodological quality of the selected trials based on the quality checklist recommended by the *Cochrane Handbook for Systematic Reviews of Interventions*, version 5.1.1 (Higgins & Green, 2015). The checklist contains six main sections, that is, sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other potential sources of bias. Any discrepancies were resolved by discussion among all authors.

Measures of Treatment Effect, Data Synthesis, and Assessment of Heterogeneity

For dichotomous variables, the relative risk (RR) and 95% confidence interval (CI) were calculated using a random-effects model. For continuous variables, the standardised weighted mean difference (SMD) and 95% CI were calculated using a random-effects model (Furukawa, Guyatt, & Griffith, 2002). Data were entered into Review Manager 5.3 (http://www.cc-ims.net/RevMan) and double-checked for accuracy (Furukawa, Guyatt, & Griffith, 2002). Statistical heterogeneity among the trials was evaluated using the I^2 statistic and Q statistic (χ^2 test) for heterogeneity among the results of each outcome analysed. The significance level of the χ^2 test was set at 0.10.

Results

Study Selection

Figure 1 shows a flowchart of the search process. The electronic search conducted on May 19, 2014 identified 394 trials from CENTRAL, 231 from MEDLINE, 128 from PsycINFO, 2,027 from Scopus, and four from CINAHL. After looking over their titles and abstracts, we identified 67 articles as potentially relevant for inclusion in the review. Ten additional studies were identified through the clinical trial databases. Finally, a total of 12 trials (Armstrong et al., 2011; Bae et al., 2012; Bostoen et al., 2012; Guerra-Tapia et al., 2007; Habib & Morrissey, 1999; Jaspers et al., 2000; Linnet & Jemec, 2001; Melin et al., 1986; Norén & Melin, 1989; Schut et al., 2013; Shi et al., 2013; Van Os-Medendorp et al., 2012) satisfied the eligibility criteria (Table 1).

Study Characteristics

The included RCTs were published between 1986 and 2013 with sample sizes of 18–109 in studies with six educational interventions (Armstrong, Kim, Idriss, Larsen, & Lio, 2011; Bostoen, Bracke, De Keyser, & Lambert, 2012; Guerra-Tapia, Lleonart, & Balañá, 2007; Jaspers, Span, Molier, & Coenraads, 2000; Shi, Nanda, Lee, Armstrong, & Lio, 2013; Van Os-Medendorp et al., 2012) and sample sizes of 16–45 in studies with six psychological interventions (Bae et al., 2012; Habib & Morrissey, 1999; Linnet & Jemec, 2001; Melin, Frederiksen, Norén, & Swebilius, 1986; Norén & Melin, 1989; Schut et al., 2013). The age of participants ranged from 30 to 40 years, and severity of symptoms varied from mild to severe. All interventions were delivered in primary- and secondary-care settings except two RCTs with either population-based (Habib & Morrissey, 1999; Schut et al., 2013) or unclear (Bae et al., 2012; Guerra-Tapia et al., 2007) settings. The intervention formats were individual except three trials where interventions were given in a group (Jaspers et al., 2000; Bostoen et al., 2012; Shut et al., 2013). The length of the intervention varied from 3–5 minutes to 52 weeks, with two to six sessions in most RCTs. The follow-up duration after the interventions ranged from 1 day to 1 year, with a corresponding period in the control groups. The severity of AD was scored based on the SCORAD index (Bostoen et al., 2012; Jaspers et al., 2000; Linnet & Jemec, 2001; Schut et al., 2013), EASI (Bae et al., 2012; Bostoen et al., 2012), IGA (Guerra-Tapia et al., 2007; Shi et al., 2013), ADAM (Habib & Morrissey, 1999), POEM (Armstrong et al., 2011), the extent plus severity part of ISDL (Van Os-Medendorp et al., 2012), and an unnamed scoring system (Melin et al., 1986; Norén & Melin, 1989). Psychological interventions consisted of habit reversal

TABLE 1

Characteristics of the Trials

Source (No. of participants)	Setting	Sample size: Intervention/ control	Intervention	Duration	No. of sessions	Follow-up	Outcome measure: Severity	QOL
Melin et al., 1986 (28)	Clinic based	7/9	Behavioural therapy. Individual. Habit reversal (behavioural habit-breaking method): patients were instructed to use the competing responses such as grasping objects when they felt an urge to scratch.	4 weeks	2	4 weeks	Physician's own evaluation	
Norén and Melin, 1989 (29)	Clinic based	23/22	Behavioural therapy. Individual. Habit reversal (behavioural habit-breaking method): patients were instructed to use the competing responses such as clenching their fists when they felt an urge to scratch.	4 weeks	2	4 weeks	Physician's own evaluation	

TABLE 1

Continued

Source (No. of participants)	Setting	Sample size: Intervention/ control	Intervention	Duration	No. of sessions	Follow-up	Outcome measure: Severity	QOL	
Habib and Morrissey, 1999 (30)	Population based	9/8	Cognitive behavioural therapy. Individual. Program includes cognitive restructuring with reference to the itch-scratch cycle, self-consciousness, anxiety, and anger management.	6 weeks	6	14 weeks	ADAM		PANAS PANAS
Jaspers et al., 2000 (31)	Hospital based	31/32	Educational therapy. Group. 6-hour sessions to inform patients about influencing factors, appropriate skin care, and to promote coping with AD, including stress management and self-control.	2 weeks	16	42 weeks	SCORAD	SF-36	

TABLE 1

Continued

Source (No. of participants)	Setting	Sample size: Intervention/ control	Intervention	Duration	No. of sessions	Follow-up	Outcome measure: Severity	QOL	
Linnet and Jemec., 2001 (12)	Clinic based	16/16	Psychodynamic therapy. Individual. Method focused on illness-related conflicts involving perception of AD, disfigurement, anxiety and aggression related to itch-scratch patterns, and depressive feelings about living with a chronic illness.	26 weeks	15.5 (11–18)	52 weeks	SCORAD		STAI
Guerra-Tapia et al., 2007 (32)	Unclear	56/53	Educational therapy. Individual. Educational material contained information about everyday patient-oriented aspects of AD and a diary for recording itch and redness intensity using VAS.	26 weeks	3	26 weeks	IGA		STAI

TABLE 1

Continued

Source (No. of participants)	Setting	Sample size: Intervention/ control	Intervention	Duration	No. of sessions	Follow-up	Outcome measure: Severity	QOL			
Armstrong et al., 2011 (33)	Clinic based	40/40	Educational therapy. Individual. Online video education on the clinical manifestations of AD, contributing environmental factors, and abating techniques.	12 weeks	1*	12 weeks	POEM	VAS			
Bae et al., 2012 (34)	Unclear	15/10	Relaxation therapy. Individual. Progressive muscle relaxation using video and audio programs at home twice a day after procedures being supervised by a psychologist.	4 weeks	56	4 weeks	EASI		STAI	BDI	VAS

TABLE 1

Continued

Source (No. of participants)	Setting	Sample size: Intervention/ control	Intervention	Duration	No. of sessions	Follow-up	Outcome measure: Severity	QOL	
Bostoen et al., 2012 (35)	Hospital based	10/11	Educational therapy. Group. 2-hour sessions containing education on AD, a healthy lifestyle such as diet, physical training, sleep hygiene and smoking.	13 weeks	24	39 weeks	SCORAD, EASI	DLQI, Skindex-29	BDI
Van Os-Medendorp et al., 2012 (36)	Clinic based	Unclear	Educational therapy. Individual Eczema portal consists of internet-guided monitoring, self-management training with general information about AD. Nurses could provide advices by e-consultation.	52 weeks		52 weeks	ISDL	DLQI	

TABLE 1

Continued

Source (No. of participants)	Setting	Sample size: Intervention/ control	Intervention	Duration	No. of sessions	Follow-up	Outcome measure: Severity	QOL		
Schut et al., 2013 (37)	Population based	17/14	Cognitive behavioural therapy. Group. 3-hour session focused on cognitive restructuring and enhancing problem-solving strategies.	5 weeks	5	8 weeks	SCORAD		HADS	HADS
Shi et al., 2013 (38)	Clinic based	8/10	Educational therapy. Individual. Eczema Action Plan contained handout information on appropriate skin care method with an emphasis on controlling skin inflammation, moisturising and relieving itching.	3 to 5 minutes**	1	1 day	IGA		HADS	

Note: ADAM = Atopic Dermatitis Assessment Measure; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; HADS = Hospital Anxiety and Depression Scale; IGA = Investigator's Global Assessment; ISDL = Impact of chronic skin disease on daily life; PANAS = Positive Affect, Negative Affect Scale; POEM = Patient-Oriented Eczema Measure; SCORAD = Severity scoring of atopic dermatitis; SF-36, Short form (36) health survey; STA = Spielberger's State-Trait Anxiety; VAS = Visual Analog Scale

*Online video available to participants to view more than once.; **Time for answering the questionnaire.

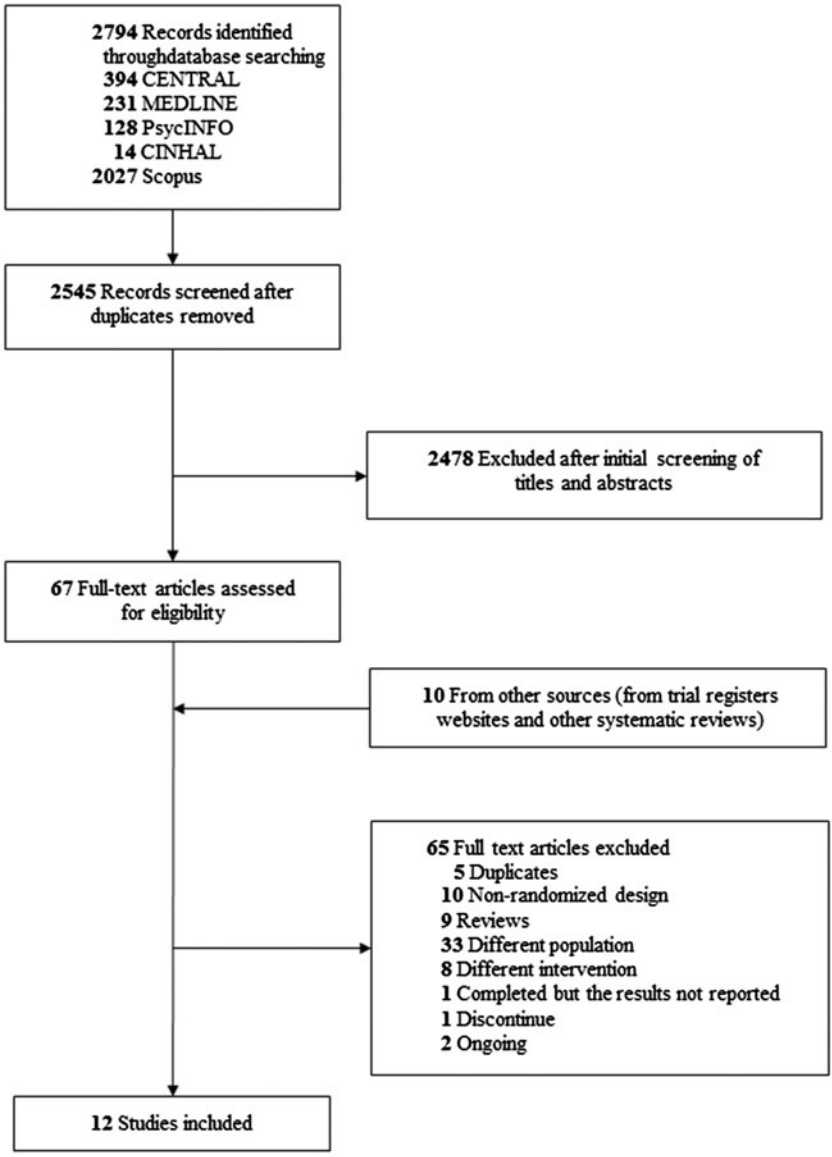


FIGURE 1

Flowchart of the literature search and study selection process.

(Melin et al., 1986; Norén & Melin, 1989), cognitive behaviour therapy (Habib & Morrissey, 1999; Schut et al., 2013), psychodynamic therapy (Linnet & Jemec, 2001), and progressive muscle relaxation (Bae et al., 2012). The interventions in six other studies (Armstrong et al., 2011; Bostoen et al., 2012; Guerra-Tapia et al., 2007; Jaspers et al., 2000; Shi et al., 2013; Van Os-Medendorp et al., 2012) were categorised as educational.

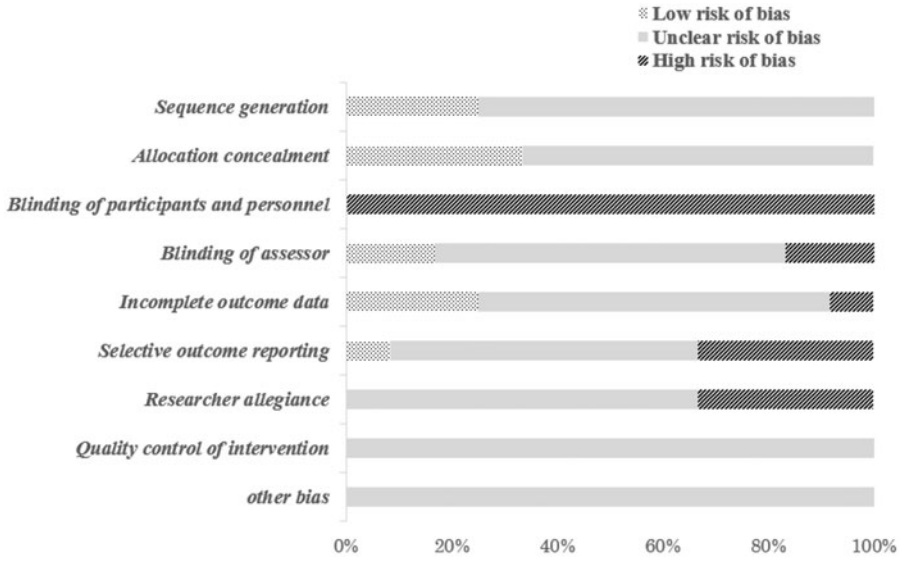


FIGURE 2

Risk of bias. Review authors' judgment of each trial is presented as percentage across all included trials.

Quality Assessment

Figure 2 shows the risk of bias, which indicates the quality of the included trials. Three of the trials (Armstrong et al., 2011; Bostoen et al., 2012; Van Os-Medendorp et al., 2012) had adequate sequence generation, and four (Armstrong et al., 2011; Bostoen et al., 2012; Van Os-Medendorp et al., 2012; Schut et al., 2013) concealed treatment allocation appropriately. Blinding of the assessor was adequate in two trials (Bostoen et al., 2012; Schut et al., 2013). Risk of bias was low for incomplete outcome data in three trials (Armstrong, 2011; Bae et al., 2012; Norén & Melin, 1989) and for selective outcome reporting in one trial (Van Os-Medendorp et al., 2012). Some RCTs had methodological weaknesses. Risk of bias was high for researchers' allegiance in four trials (Bostoen et al., 2012; Melin et al., 1986; Norén & Melin, 1989; Van Os-Medendorp et al., 2012). Blinding of participants and personnel was not possible in all 12 RCTs because of the nature of the psychological interventions. No studies reported quality control of psychotherapy regarding, for example, qualifications of the therapists and sessions.

Outcome Measurements

The follow-up duration was arbitrarily divided into <6 months and 6 months to 1 year, designated as short term and long term respectively. Of 11 RCTs that assessed the severity at short-term follow-up (Armstrong et al., 2011; Bae et al., 2012; Bostoen et al., 2012; Guerra-Tapia et al., 2007; Habib & Morrissey, 1999; Linnet & Jemec, 2001; Melin et al., 1986; Norén & Melin, 1989; Schut et al., 2013; Shi et al., 2013; Van Os-Medendorp et al., 2012), the intervention group

showed better results compared with the control group in four trials (Armstrong et al., 2011; Habib & Morrissey, 1999; Melin et al., 1986; Norén & Melin, 1989). However, the eczema severity, when measured at long-term follow-up, did not significantly differ between intervention and control groups in five RCTs (Armstrong et al., 2011; Guerra-Tapia et al., 2007; Jaspers et al., 2000; Linnet & Jemec, 2001; Van Os-Medendorp et al., 2012). Five trials (Armstrong et al., 2011; Jaspers et al., 2000; Melin et al., 1986; Schut et al., 2013; Shi et al., 2013) disclosed the dropout rates, four at short-term follow-up and one at long-term follow-up, and ranged from 0 to 17.6% in the intervention group, and 0 to 12.5% in the control group.

Results of the secondary outcome measurements were as follows. QOL was comparable between two groups at short-term follow-up in four trials (Armstrong et al., 2011; Bostoen et al., 2012; Jaspers et al., 2000; Van Os-Medendorp et al., 2012). Anxiety was significantly reduced in the intervention group compared with the control group as assessed at short-term follow-up in two studies (Habib & Morrissey, 1999; Bae et al., 2012) and at long-term follow-up in three studies (Bae et al., 2012; Habib & Morrissey, 1999; Linnet & Jemec, 2001) of the six RCTs (Bae et al., 2012; Guerra-Tapia et al., 2007; Habib & Morrissey, 1999; Linnet & Jemec, 2001; Schut et al., 2013; Shi et al., 2013). Four trials (Bae et al., 2012; Bostoen et al., 2012; Habib & Morrissey, 1999; Schut et al., 2013) examined depression at short-term follow-up, with one trial (Bae et al., 2012) reporting significant alleviation only in the intervention group. Loss of sleep was significantly ameliorated in the intervention group at short-term follow-up in one trial (Bae et al., 2012). The intervention group used more emollients and less topical corticosteroids than the control group when examined at short- and long-term follow-up in one study (Jaspers et al., 2000). Another study reported a reduction in long-term costs in the intervention group because of fewer days of absenteeism from work compared with the control group (Van Os-Medendorp et al., 2012). However, cost-effectiveness was not apparent 6 months after the interventions when the program costs and use of medical resources were taken into account (Bostoen et al., 2012). Clinical heterogeneity could not be examined because of the paucity of the pooled data.

Meta-Analyses

When the data from three trials (Armstrong et al., 2011; Bae et al., 2012; Schut et al., 2013) were examined in a meta-analysis, the severity of AD at short-term follow-up did not differ between 65 participants in the intervention group and 59 participants in the control group (SMD = -0.29, 95% CI [-0.64, 0.07], I^2 = 0%; Figure 3). However, Armstrong et al. (2011) showed that education through online video (intervention), including clinical manifestations, contributing environmental factors, and bathing and handwashing techniques, was significantly superior to handouts (control) in improving the severity of AD. The dropout rate obtained from the data of five studies (Armstrong et al., 2011; Jaspers et al., 2000; Melin et al., 1986; Schut et al., 2013; Shi et al., 2013) was 5.82% in 103 participants in the intervention group and 8.42% in 95 participants in the control group, with no significant difference between the groups (RR = 0.66, 95% CI [0.20, 2.17], I^2 = 10%; Figure 4). These data suggested that there was no difference in the acceptability of the treatment between both groups. A meta-analysis of secondary outcomes was not possible because of limitations in the available data.

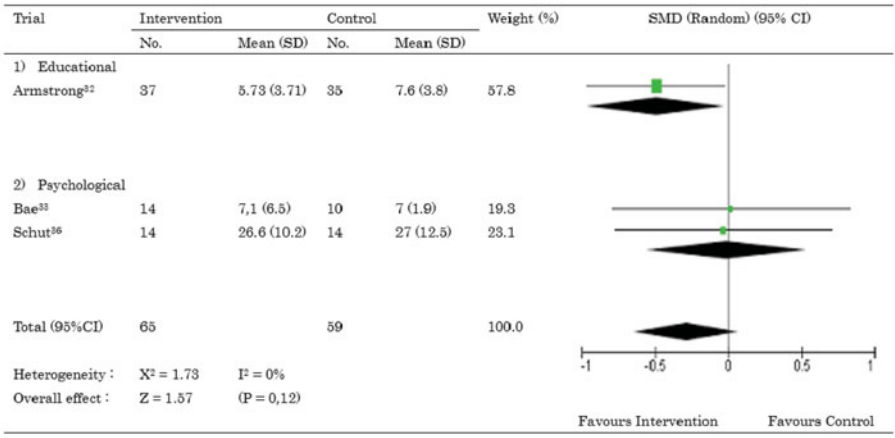


FIGURE 3

Forest plot of the severity scores for the intervention group versus control group. The square boxes indicate effect size and sample size (the larger the box, the larger the sample size) with the line indicating the 95% CI in each trial. The diamond shape shows the pooled effect size and the width the 95% CI.

Note: CI = confidence interval, *SD* = standard deviation, SMD = standardised mean difference.

Discussion

The present study systematically and comprehensively reviewed RCTs to assess the effect of psychological and educational interventions as an adjunct to conventional therapy on skin and psychological symptoms, and financial burden in adult AD. To this end, an attempt was made wherever possible to adhere to the Cochrane methodology of conducting systematic reviews. We not only searched the electronic databases, international conference abstracts, and references for relevant trials, but also contacted the authors of relevant studies to retrieve detailed information essential for the analyses. Thus, a large number of studies on psychological and educational interventions as the effect modifier were included. Furthermore, a meta-analysis of the relevant trials was performed to answer physicians' concerns, with statistically greater power and better quality of analyses.

This review revealed that there is currently only limited research evidence to provide a definite conclusion about the efficacy of educational and psychological approaches when used alongside conventional treatment of adult AD. A strong correlation between clinical manifestations of AD and mental burden has been well documented (Carroll et al., 2005; Linnet & Jemec, 2001). Dermatitis severity is positively associated with symptoms of anxiety and depression, and sleep deprivation, with resultant impairment of daytime functioning. Educational interventions in adolescent and adult patients improve self-care and disease control, and increase knowledge of treatments (Barbarot & Stalder, 2014; Staab et al., 2006; Williams, 2006). In confirmation of the previous studies (Chida et al., 2007; Ersser et al., 2014), addition of psychological and educational interventions to conventional treatment provided better therapeutic results regarding alleviation of eczema severity, and psychological symptoms. One study (Van Os-Medendorp et al., 2012) showed that health care prac-

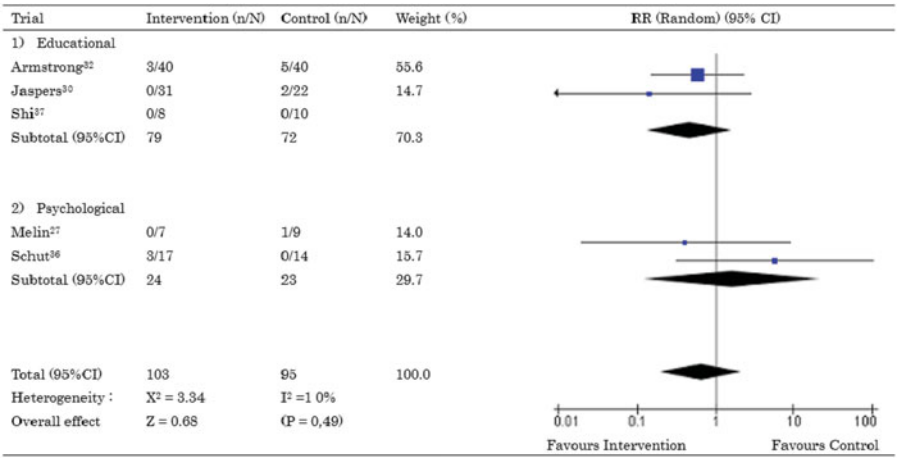


FIGURE 4

Forest plot of the dropout rate for the intervention group versus control group. The square boxes indicate effect size and sample size with the line indicating the 95% CI in each trial. The diamond shape shows the pooled effect size and the width the 95% CI.
Note: CI = confidence interval, RR = relative risk.

tice supported by electronic communication was as effective as usual face-to-face care in improving disease severity and QOL, with substantial cost savings. These results strongly suggest that psychological support and education are a promising adjunct to pharmacotherapy in AD. However, the meta-analysis showed no clear difference in the therapeutic effect of either intervention on eczema severity and the acceptability of the intervention as exemplified by the dropout rate in the two groups. Of note is the trial by Armstrong et al. (2011) showing that attainment of optimal treatment results depended on the strategy of disseminating information on AD. Good treatment results are obtained through appreciation by both patients and family members of what AD is and how to treat the disease. At the same time, clear verbal and written outlines should also be provided. In this context, an online video for patient education is an effective and appealing tool for improving clinical outcomes compared with pamphlet-based education. To maintain a satisfactory skin condition for an extended period of time, educational interventions should be regularly delivered and reinforced.

This study has several limitations. First, although we contacted authors to supplement missing data if necessary, the included trials were too small in number to perform some of the preplanned analyses. The fact that the present meta-analysis could not identify psychological and educational interventions as the effect modifier could be attributed to the low statistical power. Second, the included studies provided a range of interventions, from a single 15-minute consultation to a comprehensive series of sessions delivered to either individuals or groups. Third, the diagnostic criteria and outcome measures were heterogeneous in the included trials. The Hanifin and Rajka diagnostic criteria were employed in five trials (Armstrong et al., 2011; Bae et al., 2012; Linnet & Jemec, 2001; Schut et al., 2013; Shi et al., 2013). EASI, SCORAD, and POEM were used for the severity score. Tools for secondary outcome measures were varied and included, for example, DLQI, SF-36, Skindex-29, and VAS for QOL.

Furthermore, a point in time for outcome measurement varied from trial to trial, suggesting that the efficacy was diverse even under the same interventions. Finally, the risk of bias was unclear or high in many trials, resulting in a meta-analysis with a limited number of the trials and only for the primary outcome measures.

Poor adherence to therapy is a significant issue in the management of AD (Bass, Anderson, & Feldman, 2015). Self-management and behavioural changes to improve adherence have been increasingly emphasised in the treatment of chronic diseases. Psychological and educational interventions may be an important option in this regard. However, this review and others (Chida et al., 2007; Ersser et al., 2014) have confirmed that the evidence base remains limited regarding the efficacy of these interventions in the management of adult AD. We conclude that, rather than the combination of these interventions with conventional therapy being of no value, the data reported to date are not sufficiently powerful to detect evidence-based efficacy.

Acknowledgments

We are deeply indebted to Dr Masahiro Takigawa at the Tobi-kai Incorporated Medical Institution for critical review of the manuscript.

Conflict of Interest

The authors declare no conflicts of interest.

References

Arkwright, P.D., Motala, C., Subramanian, H., Spergel, J., Schneider, L.C., & Wollenberg, A. (2013). Atopic Dermatitis Working Group of the Allergic Skin Diseases Committee of the AAAAI. Management of difficult-to-treat atopic dermatitis. *Journal of Allergy and Clinical Immunology: In Practice, 1*, 142–151.

Arkwright, P.D., Stafford, J.C., & Sharma, V. (2014). Atopic dermatitis in children. *Journal of Allergy and Clinical Immunology: In Practice*, *2*, 388–395.

Armstrong, A.W., Kim, R.H., Idriss, N.Z., Larsen, L.N., & Lio, P.A. (2011). Online video improves clinical outcomes in adults with atopic dermatitis: A randomized controlled trial. *Journal of the American Academy of Dermatology, 64*, 502–507.

Asher, M.I., Montefort, S., Björkstén, B., Lai, C.K.W., Strachan, D.P., Weiland, S.K., ... ISAAC Phase Three Study Group. (2006). Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. *Lancet, 368*, 733–743.

Bae, B.G., Oh, S.H., Park, C.O., Noh, S., Noh, J.Y., Kim, K.R., & Lee, K.H. (2012). Progressive muscle relaxation therapy for atopic dermatitis: Objective assessment of efficacy. *Acta Dermato-Venereologica, 92*, 57–61.

Bannister, M.J., & Freeman, S. (2000). Adult-onset atopic dermatitis. *Australasian Journal of Dermatology, 41*, 225–228.

Barbarot, S., & Stalder, J.F. (2014). Therapeutic patient education in atopic eczema. *British Journal of Dermatology, 170*(Suppl. 1), 44–48.

Bass, A.M., Anderson, K.L., & Feldman, S.R. (2015). Interventions to increase treatment adherence in pediatric atopic dermatitis: A systematic review. *Journal of Clinical Medicine, 4*, 231–242.

Bieber, T. (2008). Atopic dermatitis. *New England Journal of Medicine, 358*, 1483–1494.

Bostoen, J., Bracke, S., De Keyser, S., & Lambert, J. (2012). An educational programme for patients with psoriasis and atopic dermatitis: A prospective randomized controlled trial. *British Journal of Dermatology*, *167*, 1025–1031.

Charrman, C.R., Morris, A., & Williams, H.C. (2000). Topical corticosteroids phobia in patients with atopic eczema. *British Journal of Dermatology*, *142*, 931–936.

Carroll, C.L., Balkrishnan, R., Feldman, S.R., Fleischer, A.B. Jr., & Manuel, J.C. (2005). The burden of atopic dermatitis: Impact on the patient, family, and society. *Pediatric Dermatology, 22*, 192–199.

Chida, Y., Steptoe, A., Hirakawa, N., Sudo, N., & Kubo, C. (2007). The effects of psychological intervention on atopic dermatitis. A systematic review and meta-analysis. *International Archives of Allergy and Immunology, 144*, 1–9.

Ersser, S.T., Cowdell, F., Latter, S., Gardiner, E., Flohr, C., Thompson, A.R., … Drury, A. (2014). Psychological and educational interventions for atopic eczema in children. *Cochrane Database of Systematic Reviews 2014*, Issue 1: CD004054.

Friedman, D., Nessler, D., Cycowicz, Y.M., & Horton, C. (2009). Development of and change in cognitive control: A comparison of children, young and older adults. *Cognitive, Affective, & Behavioral Neuroscience, 9*, 91–102.

Furukawa, T.A., Guyatt, G.H., & Griffith, L.E. (2002). Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. *International Journal of Epidemiology, 31*, 72–76.

Furukawa, T.A., McGuire, H., & Barbui, C. (2002). Meta-analysis of effects and side effects of low dosage tricyclic antidepressants in depression: Systematic review. *BMJ, 325*, 1–9.

Garmhausen, D., Hagemann, T., Bieber, T., Dimitriou, I., Fimmers, R., Diepgen, T., & Novak, N. (2013). Characterization of different courses of atopic dermatitis in adolescent and adult patients. *Allergy, 68*, 498–506.

Guerra-Tapia, A., Lleonart, M., & Balañá, M. (2007). Observational study to evaluate the impact of an educational/informative intervention in the emotional status (anxiety) of patients with atopic dermatitis (CUIDA-DEL). *Actas Dermo-Sifiliográficas, 98*, 250–258.

Habib, S., & Morrissey, S. (1999). Stress management for atopic dermatitis. *Behaviour Change, 16*, 226–236.

Higgins, J.P., & Green, S. (Eds.). (2015). *Cochrane handbook for systematic reviews of interventions version 5.1.0* [updated March 2011]. The Cochrane Collection. Retrieved from http://handbook.cochrane.org/

Jaspers, J.P.C., Span, L., Molier, L., & Coenraads, P.J. (2000). A multimodal education and treatment program for young adults with atopic dermatitis: A randomized controlled trial. *Dermatology and Psychosomatics, 1*, 148–153.

Leung, D.Y., & Bieber, T. (2003). Atopic dermatitis. *Lancet, 361*, 151–160.

Leung, D.Y., & Guttman-Yassky, E. (2014). Deciphering the complexities of atopic dermatitis: Shifting paradigms in treatment approaches. *Journal of Allergy and Clinical Immunology*, *134*, 769–779.

Linnet, J., & Jemec, G.B. (2001). Anxiety level and severity of skin condition predicts outcome of psychotherapy in atopic dermatitis patients. *International Journal of Dermatology, 40*, 632–636.

Melin, L., Frederiksen, T., Norén, P., & Swebilius, B.G. (1986). Behavioural treatment of scratching in patients with atopic dermatitis. *British Journal of Dermatology*, *115*, 467–474.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ, 339*, b2535.

Norén, P., & Melin, L. (1989). The effect of combined topical steroids and habit-reversal treatment in patients with atopic dermatitis. *British Journal of Dermatology*, *121*, 359–366.

Odhiambo, J.A., Williams, H.C., Clayton, T.O., Robertson, C.F., & Asher, M.I. (2009). Global variations in prevalence of eczema symptoms in children from ISAAC Phase Three. *Journal of Allergy and Clinical Immunology*, *124*, 1251–1258.

Ozkaya, E. (2005). Adult-onset atopic dermatitis. *Journal of the American Academy of Dermatology, 52*, 579–582.

Peters, E.M., Michenko, A., Kupfer, J., Kummer, W., Wiegand, S., Niemeier, V., … Gieler, U. (2014). Mental stress in atopic dermatitis-neuronal plasticity and the cholinergic system are

affected in atopic dermatitis and in response to acute experimental mental stress in a randomized controlled pilot study. *PLoS One, 9*, e113552.

Schneider, L., Tilles, S., Lio, P., Boguniewicz, M., Beck, L., LeBovidge, J., & Novak, N. (2013). Atopic dermatitis: A practice parameter update. *Journal of Allergy and Clinical Immunology, 131*, 295–299.

Schut, C., Weik, U., Tews, N., Gieler, U., Deinzer, R., & Kupfer, J. (2013). Psychophysiological effects of stress management in patients with atopic dermatitis: A randomized controlled trial. *Acta Dermato-Venereologica, 93*, 57–61.

Shi, V.Y., Nanda, S., Lee, K., Armstrong, A.W., & Lio, P.A. (2013). Improving patient education with an eczema action plan: A randomized controlled trial. *JAMA Dermatology, 149*, 481–483.

Silverberg, J.I., & Hanifin, J.M. (2013). Adults eczema prevalence and associations with asthma and other health and dermographic factors: A US population-based study. *Journal of Allergy and Clinical Immunology*, *132*, 1132–1138.

Slattery, M.J., Essex, M.J., Paletz, E.M., Vanness, E.R., Infante, M., Rogers, G.M., & Gern, J.E. (2011). Depression, anxiety, and dermatologic quality of life in adolescents with atopic dermatitis. *Journal of Allergy and Clinical Immunology*, *128*, 668.

Staab, D., Diepgen, T.L., Fartasch, M., Kupfer, J., Lob-Corzilius, T., Ring, J., … Gieler, U. (2006). Age related, structured educational programmes for the management of atopic dermatitis in children and adolescents: Multicentre, randomized controlled *trail. BMJ, 332*, 933–938.

Van Os-Medendorp, H., Koffijberg, H., Eland-de Kok, P.C.M., Van der Zalm, A., De Bruin-Weller, M.S., Pasmans, S.G.M.A., … Bruijnzeel-Koomen, C.A.F.M. (2012). E-health in caring for patients with atopic dermatitis: A randomized controlled cost-effectiveness study of internet-guided monitoring and online self-management training. *British Journal of Dermatology*, *166*, 1060–1068.

Weidinger, S., & Novak, N. (2015). Atopic dermatitis. *Lancet, 387*, 1109–1122.

Williams, H.C. (2006). Education programmes for young people with eczema. *BMJ, 332*, 923–924.

Received: 16 September 2016 | Revised: 24 May 2017 | Accepted: 17 June 2017

DOI: 10.1111/eip.12466

WILEY

BRIEF REPORT

Suicidal ideation and burnout among psychiatric trainees in Japan

Masaru Tateno[1,2,3] | Nikolina Jovanović[4] | Julian Beezhold[5,6] | Kumi Uehara-Aoyama[1] | Wakako Umene-Nakano[1] | Takashi Nakamae[1] | Naoki Uchida[1] | Naoki Hashimoto[1] | Saya Kikuchi[1] | Yosuke Wake[1] | Daisuke Fujisawa[1] | Keisuke Ikari[1] | Kotaro Otsuka[1] | Katsuyoshi Takahashi[1] | Gaku Okugawa[1] | Norio Watanabe[1] | Tomohiro Shirasaka[1] | Takahiro A. Kato[1]

[1]Burnout Syndrome Study (BoSS) Japan Group, Tokyo, Japan

[2]Tokiwa Child Development Center, Tokiwa Hospital, Sapporo, Japan

[3]Department of Neuropsychiatry, School of Medicine, Sapporo Medical University, Sapporo, Japan

[4]Centre for Psychiatry, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

[5]Norfolk and Suffolk NHS Foundation Trust, Norwich, UK

[6]Norfolk and Suffolk NHS Foundation Trust, Norwich, UK, University of East Anglia, Norwich, UK

Correspondence
Masaru Tateno, Tokiwa Child Development Center, Tokiwa Hospital, Sapporo, Japan.
Email: tatema@sapmed.ac.jp

Aim: Burnout is a psychological condition that may occur in all workers after being exposed to excessive work-related stresses. We investigated suicidal ideation and burnout among Japanese psychiatric trainees as a part of the Burnout Syndrome Study (BoSS) International.

Methods: In the Japanese branch, 91 trainees fully completed suicide ideation and behaviour questionnaire (SIBQ) and Maslach Burnout Inventory—General Survey (MBI-GS).

Results: Passive suicidal ideation was reported by 38.5% of Japanese trainees and 22.0% of them had experienced active suicidal ideation. The burnout rate among Japanese subjects was 40.0%. These results were worse compared to the all 1980 trainees who fully completed the main outcome measure in BoSS International, 25.9%, 20.4% and 36.7%, respectively.

Conclusions: Our results suggest a higher risk of suicide among Japanese residents. Japan has a higher suicide rate than other countries. Early detection of, and appropriate intervention for, suicidal ideation is important in preventing suicide in psychiatry residents.

KEYWORDS

suicide, suicidal ideation, burnout, depression, psychiatric trainee

1 | BACKGROUND

Psychological well-being is an important factor for residents to complete a specialty training programme. Mental health problems in the medical workforce can have negative effects on clinical performance and patient care (Demerouti, Bakker, & Leiter, 2014).

Burnout is a psychological condition that may occur in all workers after being exposed to excessive and prolonged work-related stresses, and is characterized by three dimensions: emotional exhaustion (EE), cynicism (CY) and reduced personal efficacy (PE) (Maslach, Schaufeli, & Leiter, 2001). Although burnout is closely related to depression, these two conditions can be differentiated from each other because burnout is limited to work-related situations, while depression is more general.

Suicidal ideation is a significant symptom of depression that may result in a catastrophic outcome such as completed suicide. Previous studies have demonstrated that burnout is independently associated with suicidal ideation in medical professionals (Shanafelt, Balch, Dyrbye, et al., 2011). Early detection of, and appropriate intervention for, suicidal ideation is important in preventing suicide in medical workers.

In this paper, we report the results of our survey which investigated burnout and suicidal ideation among psychiatric trainees in Japan.

2 | METHODS

This survey on suicidal ideation and burnout among psychiatric trainees in Japan was conducted as a part of the Burnout Syndrome Study (BoSS) International (Jovanovic, Podlesek, Volpe, et al., 2016). BoSS International was initiated by members of the European

Federation of Psychiatric Trainees (EFPT) and the European Psychiatric Association–European Early Career Psychiatrists (EPA–EECP). The aim was to assess burnout rates among psychiatric trainees and explore factors that are associated with burnout. The subjects were psychiatric trainees in 22 countries/regions and they were requested to answer online questionnaires. A national coordinator in each country was recruited using the EFPT, EPA-EECP and World Psychiatric Association (WPA) fellowship programme networks. For the Japanese branch of the study, a national coordinator (the first author of this paper) recruited study collaborators all over Japan using the Japan Young Psychiatric Organization (JYPO) mailing list, and psychiatric trainees working at their medical institutes were invited to participate in BoSS International by email. In Japan, all trainees from 15 university hospitals and their affiliated teaching institutions were enrolled. Out of 7468 psychiatric trainees internationally who received an email invitation, 1980 trainees (26.0%) fully completed the main outcome measure. Of them, 95 were from Japan (response rate: 41.5%). In the Japanese branch, 91 out of 95 trainees fully answered questionnaires to assess suicidal ideation and burnout.

Suicidal ideation and burnout were assessed by using the suicide ideation and behaviour questionnaire (SIBQ) (Marusic, Roskar, & Zorko, 2007) and the Maslach Burnout Inventory—General Survey (MBI-GS) (Maslach, Jackson, & Leiter, 1996), respectively. SIBQ is a short instrument to assess suicidality as a trait that arises from passive and active suicidal ideation through to suicidal behaviour and suicide attempts (Jovanović, Podlesek, Medved, et al., 2013). The criteria for severe burnout in this survey were defined as EE scores of 2.20 and higher, and CY of 2.00 and higher. Statistical analysis was performed by using IBM SPSS Statistics 22.

The study protocol was approved by the ethics committee of Sapporo Medical University, and this study was conducted in compliance with the Helsinki Declaration. The aim of this study was clearly stated on the online survey's first page. Answering the questionnaire was deemed to constitute consent, with an initial question asking for the subject's agreement to participate. All respondents participated in this study without any incentive provided by the study investigators.

TABLE 1 A two-group comparison of work environment and burnout between absence and presence of passive suicidal ideation assessed by using suicide ideation and behaviour questionnaire

	Absence (n = 56)	Presence (n = 35)	P-value
Age (years old)	32.1 ± 5.1	31.3 ± 4.4	0.4180
Clinical experience (years)	4.0 ± 2.6	3.0 ± 1.9	0.0628
Weekly working hours	64.8 ± 21.0	72.3 ± 24.1	0.1219
Weekly supervision	5.5 ± 6.1	4.2 ± 3.9	0.1380
Maslach Burnout Inventory—General Survey subscore			
Emotional exhaustion	2.86 ± 1.4	3.26 ± 1.5	0.1961
Cynicism	2.06 ± 1.4	2.62 ± 1.5	0.0704
Personal efficacy	3.86 ± 1.2	3.44 ± 1.0	0.0831

3 | RESULTS

The mean age of the 91 Japanese psychiatric trainees (67.0% male) enrolled in BoSS International was 31.8 ± 4.8. Their mean clinical experience was 3.6 ± 2.4 years. The mean weekly working hours in the workplace were 67.7 ± 22.4.

Regarding experienced suicidal ideation measured by SIBQ, passive suicidal ideation (*Have you ever thought it would be better if you die?*) was reported by 38.5% of Japanese trainees and 22.0% of them had experienced active suicide ideation (*Have you ever thought about committing suicide?*). These rates were higher than the means of the whole study, 25.9% and 20.4%, respectively. Statistical analysis demonstrated a significant difference between the Japanese sample and the whole data set with regard to passive suicidal ideation (P = .0089), while the difference was not significant on active suicide ideation (P = .7964). With regard to the period of suicidal ideation, 77.1% of psychiatric trainees with passive suicidal ideation and 70.0% of those with active suicidal ideation experienced it before entering psychiatry training programmes.

Mean scores on MBI-GS subscales for Japanese subjects were EE 3.0 ± 1.5, CY 2.3 ± 1.4 and PE 3.7 ± 1.1. All three scores were worse compared to the means for all 1980 subjects: 2.6 ± 1.4, 2.0 ± 1.4 and 4.5 ± 1.1, respectively. Two group comparisons between presence and absence of suicidal ideation demonstrated no statistically significant difference. However, the trainees with current or previous suicidal ideation had higher EE and CY, and lower PE scores on MBI-GS, that is, they had more severe burnout compared to those without suicidal ideation (Table 1).

4 | DISCUSSION

Japan has a higher suicide rate than most other countries. Organisation for Economic Co-operation and Development (OECD) data show that the annual suicide rate in Japan is 20.9 per 100 000 population, fourth highest among 34 member countries (OECD, 2014). This is much higher than the mean for all OECD countries, 12.4 per 100 000.

Previous studies have revealed that the suicide rate among physicians is much higher than in the general population (Schernhammer & Colditz, 2004). This fact may be explained, at least partially, by the stigma against psychiatric services among non-psychiatry specialists and the fear that receiving psychiatric treatment could lead to the loss of their medical license although they are aware of the necessity of psychiatric intervention (Rubin, 2014).

To the best of our knowledge, this study is the first to investigate suicidal ideation among psychiatric trainees in Japan by using SIBQ. Our results demonstrated that, among 91 Japanese psychiatric trainees who replied to our invitation, 38.5% had experienced passive suicidal ideation and 22.0% of them had experienced active suicidal ideation. Both of the rates were higher than those among psychiatry trainees in other countries, suggesting a higher risk of suicide.

The burnout rate among Japanese psychiatric trainees was 40.0% according to the results of the MBI-GS subscales, while the rate of burnout was 36.7% in the whole study. A tendency for correlation was observed between current/previous suicidal ideation and

burnout. Burnout may be linked to drop out from speciality training programmes and malpractice in clinical settings (Thomas, 2004).

Statistical analyses of the whole data set from BoSS International (n = 1980) showed that younger age and longer working hours were associated with burnout (Jovanović et al., 2016). Further analyses revealed that reduced contact hours with their supervisors could be related to burnout among psychiatric trainees. Clinical mentors should be aware of the higher risk of burnout in young residents and the importance of regular and sufficient supervision to prevent burnout and to diminish suicidal ideation. Furthermore, programme directors should pay careful attention to trainees with current/previous suicidal ideation because they may be more susceptible to burnout.

This study has several limitations. Sample size was too small to draw definitive conclusions because the total number of trainees in Japan at the time of study was estimated as 1922 (representative rate: 4.9%). However, this is the largest study on suicidal ideation and burnout among psychiatric trainees in Japan. Previous history of suicidal ideation may affect motivation for entering psychiatry training programmes and this fact may bias the results. There may be differences in the time when thoughts and behaviours assessed by SIBQ and MBI-GS were experienced.

The Japanese Government revised the Industrial Safety and Health Act in order to improve the mental health of workers in every field, and obliged employers to regularly conduct "psychological stress checks" of all workers since December 2015. The results of BoSS International reveal that burnout was associated with long working hours, lack of supervision and not having regular rest. More attention should be paid to the work conditions of psychiatric trainees because some of them might have a risk of suicide.

REFERENCES

Demerouti, E., Bakker, A. B., & Leiter, M. (2014). Burnout and job performance: The moderating role of selection, optimization, and compensation strategies. *Journal of Occupational Health Psychology*, *19*, 96–107.

Jovanović, N., Podlesek, A., Medved, V., Grubišin, J., Mihaljevic-Peleš, A., Goran, T., & Lovretić, V. (2013). Association between psychopathology and suicidal behavior in schizophrenia. A cross-sectional study of 509 participants. *Crisis*, *34*, 374–381.

Jovanovic, N., Podlesek, A., Volpe, U., Barrett, E., Ferrari, S., Rojnic Kuzman, M., … Beezhold, J. (2016). Burnout syndrome among psychiatric trainees in 22 countries: Risk increased by long working hours, lack of supervision, and psychiatry not being first career choice. *European Psychiatry*, *32*, 34–41.

Marusic, A., Roskar, S., & Zorko, M. (2007). Questionnaire on suicide ideation and behaviour. In A. Shrivastava (Ed.), *Suicide prevention in developing countries* (pp. 201–209). London, England: Gaskell.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *Maslach burnout inventory manual* (). Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*, 397–422.

OECD. (2014). *Making mental health count – The social and economic costs of neglecting mental health care (Report)* (). Paris, France: OECD.

Rubin, R. (2014). Recent suicides highlight need to address depression in medical students and residents. *Journal of the American Medical Association*, *312*, 1725–1727.

Schernhammer, E. S., & Colditz, G. A. (2004). Suicide rates among physicians: A quantitative and gender assessment (meta-analysis). *The American Journal of Psychiatry*, *161*, 2295–2302.

Shanafelt, T. D., Balch, C. M., Dyrbye, L., Bechamps, G., Russell, T., Satele, D., … Oreskovich, M. R. (2011). Special report: Suicidal ideation among American surgeons. *Archives of Surgery*, *146*, 54–62.

Thomas, N. K. (2004). Resident burnout. *Journal of the American Medical Association*, *292*, 2880–2889.

How to cite this article: Tateno M, Jovanović N, Beezhold J, et al. Suicidal ideation and burnout among psychiatric trainees in Japan. *Early Intervention in Psychiatry*. 2017;1–3. https://doi.org/10.1111/eip.12466

VOLUME 35 · NUMBER 6 · FEBRUARY 20, 2017

JOURNAL OF CLINICAL ONCOLOGY

CORRESPONDENCE

Concerns About Selective Outcome Reporting

TO THE EDITOR: In their recent article in *Journal of Clinical Oncology* titled "Randomized Controlled Trial of Family Therapy in Advanced Cancer Continued Into Bereavement," Kissane et al[1] reported on results from their trial of family therapy for at-risk families when one of their members had advanced cancer. In their abstract as well as in their main article, the authors showed a significant effect of family-focused grief therapy (FFGT) based on the results of the Complicated Grief Inventory (CGI). We are concerned that their manuscript is an example of selective outcome reporting[2] and of abstract spin.[3]

First, in their protocol (published online simultaneously with the main article), the primary outcomes were clearly defined as the Brief Symptom Inventory (BSI) and the Beck Depression Inventory (BDI; pp. 15) and used as the basis for sample size calculation (pp. 32), whereas the article reported the results of the CGI and BDI as if they were the primary outcomes. Moreover, the article did not report on the BSI at all. Thus, it seems as if the CGI became the main focus of the published article because of its significant results. If there were any justifiable revisions to the original protocol, the authors should have explained those changes in the article. Compliance with the CONSORT 2010 guidelines should be a prerequisite for publication.

Second, in the abstract, Kissane et al concluded that FFGT was efficacious by focusing on the positive results derived from the analysis of the CGI. We are concerned that such a focus on the positive findings of the CGI analysis in the abstract, with a passing note on only one of their two predefined primary outcomes, will overemphasize the beneficial effects of intervention and therefore subsequently mislead readers.[4-6]

We agree with the authors' statement that family-centered cancer care is needed. However, reporting all primary results transparently, in both the text and the abstract, is essential to promote the appropriate and informed application of FFGT and to lead future research in the right direction.

Keiko Kusaka, Kiyomi Shinohara, and Masafumi Tada
Kyoto University Graduate School of Medicine, Kyoto, Japan

Norio Watanabe and Toshi A. Furukawa
Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at ascopubs.org/journal/jco.

REFERENCES

1. Kissane DW, Zaider TI, Li Y, et al: Randomized controlled trial of family therapy in advanced cancer continued into bereavement. J Clin Oncol 34: 1921-1927, 2016
2. Chan AW, Hróbjartsson A, Haahr MT, et al: Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. JAMA 291:2457-2465, 2004
3. Boutron I, Dutton S, Ravaud P, et al: Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA 303:2058-2064, 2010
4. Shinohara K, Tajika A, Imai H, et al: Protocol registration and selective outcome reporting in recent psychiatry trials: New antidepressants and cognitive behavioural therapies. Acta Psychiatr Scand 132:489-498, 2015
5. Furukawa TA, Watanabe N, Omori IM, et al: Association between unreported outcomes and effect size estimates in Cochrane meta-analyses. JAMA 297:468-470, 2007
6. Boutron I, Altman DG, Hopewell S, et al: Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: The SPIIN randomized controlled trial. J Clin Oncol 32:4120-4126, 2014

DOI: 10.1200/JCO.2016.69.7557; published at ascopubs.org/journal/jco on January 3, 2017.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Concerns About Selective Outcome Reporting

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Keiko Kusaka
No relationship to disclose

Kiyomi Shinohara
Research Funding: Japan Primary Care Association

Masafumi Tada
No relationship to disclose

Norio Watanabe
Patents, Royalties, Other Intellectual Property: Sogensha, Paquet and Akatsuki

Toshi A. Furukawa
Honoraria: Eli Lilly, Janssen Pharmaceuticals, Meiji Seika Kaisha, Merck Sharp & Dohme, Otsuka, Pfizer, Mitsubishi Tanabe Pharma
Consulting or Advisory Role: Sekisui Chemical, Takeda Science Foundation
Research Funding: Mochida Pharmaceutical (Inst), Mitsubishi Tanabe Pharma (Inst), Japan Foundation for Neuroscience and Mental Health (Inst)
Patents, Royalties, Other Intellectual Property: Igaku-Shoin, Nihon Bunka Kagaku-sha