

|                             |     |
|-----------------------------|-----|
| 平成28年3月16日                  | 資料2 |
| 第29回レセプト情報等の<br>提供に関する有識者会議 |     |

# 基本データセットの提供について

平成28年3月16日

厚生労働省保険局医療介護連携政策課  
保険システム高度化推進室

# 現在提供されているNDBデータの種類

|                           | 特別抽出   | サンプリング<br>データセット                                     | 基本データセット   | 集計表情報                              |
|---------------------------|--|--|--|------------------------------------|
| 基本的な<br>イメージ              | 申出者の要望に応じ、データベースにある全データのなかから、該当する個票の情報を抽出し、提供する          | 探索的研究へのニーズに対応し、抽出、匿名化などを施して安全性に十分配慮した、単月分のデータセット     | 入院、外来、疾患別など目的に合わせて年度ごとに紐付けが可能で、簡易に分析することが可能なデータセット | 申出者の要望に応じ、データを加工して作成した集計表を提供する     |
| 提供データ                     | 個票   | 一部匿名化等を行った個票   | 大幅に加工した個票  | 集計表                                |
| 含まれている<br>データ項目例          | レセプト情報、特定健診等情報に含まれている、ほぼすべての項目                           | 希少な情報があらかじめ匿名化・削除されたレセプトデータ                          | 患者の基本属性情報以外は、主傷病名、診療識別情報、要望に応じたコードなど               | 集計表                                |
| 利用にあたり<br>具備すべき<br>セキュリティ | データ利用時に、情報セキュリティマネジメントシステムを確実に運用できる利用環境を整える              | 特別抽出で求められるセキュリティ水準と比較してある程度具備しやすいセキュリティ水準での利用が可能     |  |                                    |
| 想定される<br>利用者像             | レセプト研究に一定の知見があり、申出内容や抽出条件を吟味し、大量のデータを高速に処理することを想定している利用者 | レセプト研究に関心はあるが経験がまだ十分でなく、データの特徴や各項目の概要を把握したいと考えている利用者 | レセプトの構造を踏まえながら研究するよりも、基本的項目について簡単に分析を試みたいと考えている利用者 | 集計された結果を必要とし、データ処理を行うことを想定していない利用者 |
| 提供実績<br>(計66件)            | 42件  | 15件  | 2件   | 7件                                 |

# 第三者提供の承諾状況：平成25年9月～平成26年3月

| 審査時期   | 申出者      | 所属機関                             | 研究名称  | 提供情報                           | 該当要件         |
|--------|----------|----------------------------------|---|--------------------------------|--------------|
| H25.09 | 大洞 清香    | 文部科学省科学技術・学術政策局企画評価課             | 糖尿病を中心とした生活習慣病の、全国的な合併症費用構造および、患者分布や医療費などの試算        | サンプリングデータセット                   | ①            |
|        | 堀口 裕正    | 国立病院機構総合研究センター                   | エビデンスに基づいた診療報酬改定を行うためのレセプトデータ利活用の手法についての研究          | レセプト情報                         | ③            |
|        | 森由希子     | 京都大学医学部付属病院                      | 高齢者胃がん患者における治療実態調査に関する研究                            | レセプト情報                         | ④            |
|        | 三宅 康史    | 昭和大学                             | レセプト情報を用いた夏季熱中症例の発生実態調査                             | レセプト情報(集計表情報)                  | ④            |
|        | 上嶋 健治    | 京都大学医学部付属病院                      | メタボリック症候群・肥満症の臨床疫学研究－ナショナルデータベースを用いた横断・縦断解析         | 特定健診等情報                        | ④            |
|        | 加藤 元嗣    | 北海道大学病院                          | National databaseからみたピロリ菌除菌と胃癌診療の実態調査               | レセプト情報                         | ④            |
|        | 梶尾 雅宏    | 厚生労働省医政局指導課                      | 医療計画作成支援データベースの構築及び医療計画を踏まえた医療の連携体制構築に関する評価の検討について  | レセプト情報                         | ①            |
|        | 豊川 智之    | 東京大学                             | 脳性麻痺児の原因別発生頻度及び脳性麻痺患者の医療費の推計                        | レセプト情報                         | ④            |
|        | H26.03   | 飯原なおみ                            | 徳島文理大学  | ふらつき等の危険を有する薬の使用と骨折との関連性に関する研究 | サンプリングデータセット |
| 宮下 光令  |          | 東北大学                             | 厚生労働省が提供するレセプト情報等を用いた終末期がん医療の実態調査及び終末期がん医療の質評価方法の開発 | サンプリングデータセット                   | ④            |
| 今中 雄一  |          | 京都大学                             | 超高齢化社会における認知症及び脳卒中の地域別・病床機能別医療提供状況と関連要因の分析          | ④                              | ④            |
| 西 信雄   |          | 国立健康・栄養研究所                       | 特定健診・特定保健指導における生活習慣病予防のための運動と果に関する研究                | ③                              | ③            |
| 頭金 正博  |          | 名古屋市立大学                          | 特種集団における医薬品処方実態調査                                   | ④                              | ④            |
| 小野 順子  |          | 福岡県立大学                           | 我が国における重複受診、頻回受診、重複処方の現状                            | ④                              | ④            |
| 吉田 愛   |          | 全日本病院協会総合研究所                     | 亜急性期入院医療管理料を算定した患者の診療内容に関する研究                       | ⑥                              | ⑥            |
| 今野 俊範  |          | 神奈川県保健福祉局保健医療部                   | 在宅医療実態状況調査  | ②                              | ②            |
| 恒石美登里  |          | 日本歯科総合研究機構                       | 歯科治療状況と医療受療動向との関連調査                                 | ⑥                              | ⑥            |
| 東 尚弘   |          | 国立がん研究センター                       | 我が国のがん医療におけるがん診療連携拠点病院の役割及び連携の実態に関する研究              | ③                              | ③            |
| 大江 和彦  |          | 東京大学医学部附属病院                      | ナショナルデータベースを用いた脳血管疾患の実態に関する研究                       | 基本データセット                       | ④            |
| 野口 晴子  |          | 早稲田大学                            | 急性期心疾患患者の医療費と治療効果に関する実証的研究                          | 基本データセット                       | ④            |
| 武藤 正樹  | 国際医療福祉大学 | 後発医薬品普及促進のためのレセプトナショナルデータベース活用研究 | ④   | ④                              |              |

平成26年3月審査にて、  
提供の承諾あり

# 基本データセットの特徴

●厚生労働科学研究「汎用性の高いレセプト基本データセット作成に関する研究」等にて開発を進めた(注)

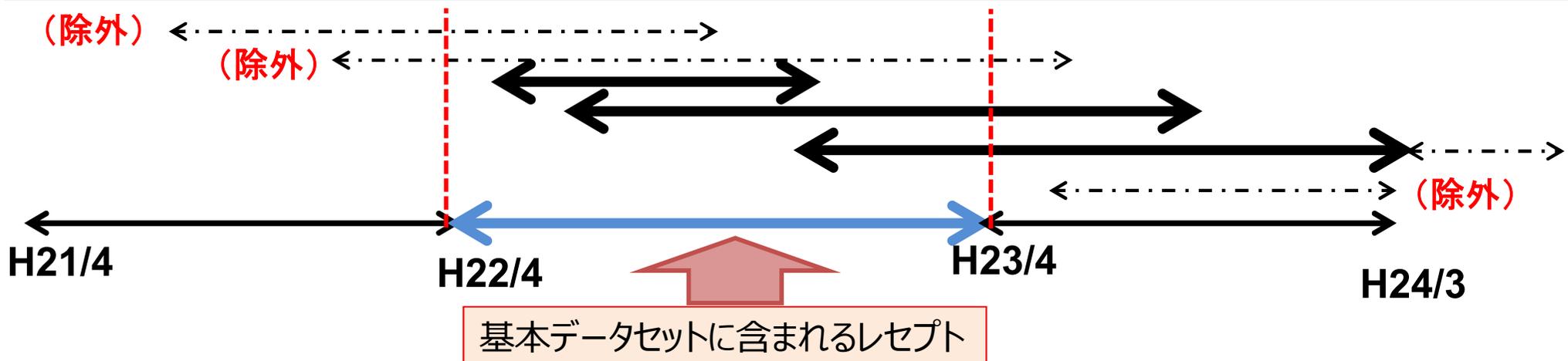
●抽出条件

- ・対象レセプト：平成22年度 医科入院レセプトおよびDPCレセプト
- ・入院日：平成22年度内のレセプト(提供するレセプトは平成24年3月診療分まで)  
入院日と診療実日数はデータ項目に含まれる
- ・対象疾患：ICD-10 中分類:121分類より申出者が選択し抽出
- ・都道府県情報及び保険者コードは削除済み

●H26年3月承諾案件への提供が完了(申出者:大江和彦(東京大学))

(注)

- ・平成27年度～平成29年度厚生労働科学研究「レセプト情報・特定健診等情報データベースの利活用の推進に関する研究」(研究代表者:大江 和彦(東京大学医学部附属病院))
- ・平成26年度～平成27年度厚生労働科学研究「レセプト情報・特定健診等情報データベースを利用した医療需要の把握・整理・予測分析および超高速レセプトビッグデータ解析基盤の整備(戦略研究)」(研究代表者:満武 巨裕(医療経済研究機構))
- ・平成24年度～平成25年度厚生労働科学研究「汎用性の高いレセプト基本データセット作成に関する研究」(研究代表者:満武 巨裕(医療経済研究機構))



# 基本データセット：医科入院・DPCデータセット

## ●レコードフォーマット：データセットC（記述統計等に活用）

| 項番 | データ項目名（日本語）    | 条件等記入欄 |
|----|----------------|--------|
| 1  | ID1            |        |
| 2  | ID2            |        |
| 3  | ID3(研究班試作)     |        |
| 4  | 男女区分           |        |
| 5  | 年齢階級コード(H22年次) |        |
| 6  | 入院回数           |        |
| 7  | 合計点数           |        |
| 8  | 合計診療実日数(入院)    |        |
| 9  | 医療機関数          |        |
| 10 | 入院回数シーケンス番号    |        |
| 11 | 入院-DPCフラグ      |        |
| 12 | 合計点数           |        |
| 13 | 診療実日数(入院)      |        |

## ●レコードフォーマット：データセットD（診療行為・処方情報を追加）

| 項番  | データ項目名（日本語）        | 条件等記入欄                          |
|-----|--------------------|---------------------------------|
| 1   | ID1                | C1と同じ                           |
| 2   | ID2                |                                 |
| 3   | ID3(研究班試作)         |                                 |
| 4   | 男女区分               |                                 |
| 5   | 年齢階級コード(H22年次)     |                                 |
| 6   | 入院回数               |                                 |
| 7   | 合計点数               |                                 |
| 8   | 合計診療実日数(入院)        |                                 |
| 9   | 医療機関数              |                                 |
| 10  | 入院回数シーケンス番号        |                                 |
| 11  | 入院-DPCフラグ          |                                 |
| 12  | 合計点数               |                                 |
| 13  | 診療実日数(入院)          |                                 |
| 14  | 主病名1               | SYの主傷病(10)にフラグが付いていた傷病名コード(5)   |
| 15  | 診療開始日1             | SYの主傷病(10)にフラグが付いていた診療開始日コード(6) |
| 16  | 転帰区分1              | SYの主傷病(10)にフラグが付いていた転帰区分(7)     |
| 17  | 主病名1(ICD10_大分類)    |                                 |
| 18  | 主病名1(ICD10_中分類)    |                                 |
| ... | 傷病名コード             | オプション 傷病名コードが存在するかのフラグ等が設定可能    |
| ... | 診療行為コード・フィールド      | オプション 診療行為レコード(SI)の情報格納：データセットE |
| ... | 医薬品コード・フィールド       | オプション 医薬品レコード(SI)の情報格納：データセットE  |
| ... | 特定器材コード・フィールド      | オプション 特定器材レコード(SI)の情報格納：データセットE |
| ... | ...                | オプション ...                       |
| 256 | その他要望されたコード、コード群など | オプション 行数は最大256行とする              |

申出者の研究テーマに合わせた条件の設定が可能

申出者の希望でフラグ  
を設定・指定が可能

# 入院データセット C, Dについて

|                 | 入院データセットC  | 入院データセットD   |
|-----------------|--|---|
| レコード数<br>(=患者数) | <ul style="list-style-type: none"> <li>保険者単位で抽出した患者レコードから、さらに研究者用に抽出</li> </ul>  | <ul style="list-style-type: none"> <li>左記患者の入院エピソード全て</li> </ul>  |
| 抽出患者            | <p>【保険者の抽出方法】</p> <ul style="list-style-type: none"> <li>都道府県単位の保険者（協会けんぽ等）は、全国8ブロックから1つずつ、<u>人口25%分の保険者をランダムに抽出</u></li> <li>全国にまたがる保険者（組合健保、共済等）は、<u>全保険者から20%をランダムに抽出</u></li> </ul> <p>【抽出された保険者間での、患者抽出方法】</p> <ul style="list-style-type: none"> <li>抽出された保険者の全患者から<u>20%分に該当する患者をランダムに抽出</u></li> </ul> |   |
| 対象期間            | <ul style="list-style-type: none"> <li>入院イベントごと</li> <li>年間の入院イベントを把握できる。</li> </ul>   | <ul style="list-style-type: none"> <li>入院イベントごと</li> <li>データセットCと併用することで、年間の入院イベントの詳細を把握できる。</li> </ul> |
| 患者属性情報          | <ul style="list-style-type: none"> <li>研究用作成IDおよび性、5歳刻み年齢階級</li> </ul>   |   |
| 項目              | <ul style="list-style-type: none"> <li>入院／外来合計点数</li> <li>入院／外来合計診療実日数</li> <li>入院／外来受診医療機関数</li> <li>主傷病名</li> <li>診療識別情報</li> </ul>  | <p>【入院データセットCに加え】</p> <ul style="list-style-type: none"> <li>申出者の要望により、診療行為・医薬品情報等のデータ項目も提供</li> </ul>  |
| データ構造           | <ul style="list-style-type: none"> <li>データ項目数は固定</li> </ul>  | <ul style="list-style-type: none"> <li>データ項目数は一定範囲内に確定する</li> </ul>                                     |
| 研究の例            | 患者別の診療行為別医療費分析等<br>(社会医療診療行為別調査の個票のイメージ)   | 患者別の診療行為・医薬品・特定器材の利用別分析等<br>(研究に必要なデータ項目を含む個票)  |



- 抽出方法から、個人の特定性は非常に低い
- 提供依頼申出者は、各ファイルのデータ項目を確認し、あらかじめ用意された上記データセットを選択し、さらに研究目的に応じて必要となるデータ項目の指定を行う

# 提供するデータの特徴とセキュリティの確保について

個人の特定期間

高リスク

**匿名化等の  
処理なし**  
匿名化等の処理は行  
わず、申出に沿って  
データ提供を行う

**匿名化等**  
個人などの特定可能  
性を下げる、一定の  
処理を行っている

**集計表化**  
依頼に基づいて  
集計表を作成し  
提供する

**定型表化**  
ごく基本的な項目を  
毎年集計し、HP等  
で公表のうえ、ダウ  
ンロード可能とする

NDBオープン  
データ

**サンプリング  
データセット**  
・抽出条件の提供は不要。  
・一部匿名化等を行った個票

**基本データセット**  
・傷病名などの基本的な項目  
から、要望に応じた一部(最大  
256項目)の診療行為情報を  
含み、年度単位で患者をまと  
めたデータセット

**集計表提供**  
・申出者の要望に応じ、データを  
加工して作成した集計表を提供  
する

**特別抽出**  
・申出者の提供依頼に応じ、  
有識者会議の審査を経て  
データを提供  
・抽出条件、およびセキュリ  
ティの確保も審査

低リスク

**オープン化**  
例: ウェブサイトで公開するなど、  
誰でも利用できるデータとする

**セキュリティ要求**  
例: 申出者に一定の水準のセキュ  
リティ環境を要求し、利用環境の  
実地監査を行う

緩い

厳しい

利用者に求めるセキュリティ要件

# 基本データセットに関する今後の予定

- ① 承諾された申出者へのデータ提供を完了
- ② 平成22年度基本データセットの仕様の確定と提供
- ③ 平成23－26年度基本データセットの作成及び提供の開始  
( \* 提供を通して得られた知見を反映し、基本データセットを適宜バージョンアップする )



平成28年度内に「基本データセット」本格運用を目指す