

厚生労働行政推進調査事業費補助金  
(政策科学総合研究事業(政策科学推進研究事業))

レセプト情報・特定健診等データベースの利活用の推進に関する研究

平成 28 年度 総括・分担研究報告書

研究代表者 大江 和彦

平成 29 (2017) 年 3 月

# 目 次

## I. 総括研究報告

レセプト情報・特定健診等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 ----- 1

## II. 分担研究報告書

レセプトNDB（ナショナルデータベース）利活用推進のための課題と対策

研究分担者：今 中 雄 一 ----- 7

レセプト情報・特定健診等情報データベースの情報提供機能の充実と利用

研究分担者：満 武 巨 裕 ----- 37

III. 研究成果の刊行に関する一覧表 ----- 47

厚生労働行政推進調査事業費補助金  
(政策科学総合研究事業(政策科学推進研究事業))  
総括研究報告書

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企画情報運営部 教授

**研究要旨**

レセプト情報・特定健診等情報データベース(NDB)について、平成 28 年度から提供が開始された NDB オープンデータとともに、その情報提供機能の充実と利用に関して潜在的に存在していると考えられる課題も含めてとりあげ、その解決方法についても検討する。具体的に検討すべき課題として 28 年度は、①ID による同一患者の名寄せ手法、②信頼できる傷病名情報の取得方法、③NDB 特別抽出データの取得と処理方法の効率化、④基本データセットの作成方法、⑤NDB オープンデータの役割および利用例としての国際比較データの作成、の 5 点について検討した。レセプトデータが患者ごとに連結されず、医療機関 x 保険者 x 個人 x 受診年月、に分断されたデータ単位として匿名化されているためにこれらの分断データ単位を再結合することが非常に困難になっている。この問題を解決するには保険者が厚労省にデータを匿名化して出す場合に、医療機関 x 個人 x 受診年月についてはその同一個人ごとに 1 匿名化 ID となるように処理するとともに、その匿名化 ID が保険者間での結合できるように、全レセプトデータを管理する機関が正確な被保険者台帳を管理運用することであろう。一方、データ提供にかかる時間の短縮、受け取ったデータの再 DB 化が効率的に実現できるように、研究者が利用するデータベースエンジンを想定したデータ提供も必要で、現状のテキストデータでの提供には限度がある。また傷病名の精度確保に関する研究をさらにすすめる必要もある。さらに今後も、NDB からのデータ提供としての特別抽出、サンプリングデータセット、基本データセット、NDB オープンデータをはじめとする情報提供機能の充実と利用が必要である。

以上のように課題は残っているが、患者 ID の取扱い、オープンデータの利用、特別抽出データの利用を含む NDB データ活用につき、具体的課題と今後の対応のあり方を示した。今後の NDB データの提供側およびデータの利用側にとっても活用推進に役立つと思われる。また NDB オープンデータから生成できるデータを使用した国際統計報告は、諸外国の行政関係者や研究者への波及効果も生み出し得るものと考えられた。

### <研究分担者>

今中 雄一 京都大学大学院医学研究科医療経済学分野 教授  
満武 巨裕 一般財団法人 医療経済研究・社会保険福祉協会 医療経済研究機構  
副部長

### <研究協力者>

國澤 進 京都大学大学院医学研究科医療経済学分野 講師  
大坪 徹也 京都大学大学院医学研究科医療経済学分野 助教  
佐藤 大介 東京大学医学部附属病院企画情報運営部 助教

## A. 研究目的

日本におけるレセプト情報・特定健診等情報データベース（以下、NDB）の情報提供機能の充実と利用についての課題および解決方法について検討する。日本のNDBは、提供開始（2011年11月）から本年度までの利用承諾件数が100件を超えた。引き続き、研究利用として提供している特別抽出、サンプリングデータ、基本データセット、NDBオープンデータをはじめとする情報提供機能の充実と利用が必要であるが、潜在的に存在していると考えられる課題をとりあげ、その検討を行う。

## B. 研究方法

具体的に検討すべき課題として今年度は、①IDによる同一患者の名寄せ手法、②信頼できる傷病名情報の取得方法、③NDB 特別抽出データの取得と処理方法の効率化、④基本データセットの作成方法、⑤NDB オープンデータの役割および利用例としての国際比較データの作成、の5点について取り上げた。①③⑤については分担研究者の今中が、④および⑤の国際比較データ作成については分担研究者の満武は、②について研究代表者の大江がそれぞれ分担して検討した。

## C. 研究結果と考察

①IDによる同一患者の名寄せ手法に関する検討

「全患者ID1数」に占める「患者ID2を一つもつ患者ID1の数」は、63%であり、「全患者ID2数」に占める「患者ID1を一つもつ患者ID2の数」は、68%であった。これにより、氏名の入力ゆれや改名の発生頻度は、保険者番号・記号番号の入力ゆれや保険の変更の発生頻度よりも低いものの、小さいものではないといえる。開発した統合ID生成アルゴリズムを適用した結果、患者ID1と患者ID2の組合せ数に対する統合ID数は42%となった。

こうしたハッシュ化患者IDの問題に対する根本的な解決は、全レセプトデータを管理する機関が、正確な被保険者台帳を管理運用することであろう。

②信頼できる傷病名情報の取得方法の検討

2012年のある4ヶ月間での入院歴のないレセプトをランダムに40名分調査したところ、レセプト1件あたりの傷病名数は平均15.2個であり、レセプトに登録された診療行為と傷病名の適用関係データベースを利用し

た診療行為から傷病名にウエイトをつけて評価した先行研究では、ほぼ疑いなく患者に存在する傷病名の推定は93%の正確さであったとの報告があり、連続する数ヶ月のレセプトを処理することで性能が上がる可能性が指摘されている。

機械学習と上記手法とを組み合わせた傷病名推定を行うことにより、精度の高い傷病名推定ができる可能性があり、入院患者レセプトについて同一患者で連結した最長6ヶ月のレセプトの傷病名、診療行為、診療科、時間軸で出現情報などを入力とし、正解データとしてDPCの最も医療資源を消費した病名を用いて機械学習することが考えられた。

機械学習により傷病名推定を行うには、大量の正解データが必要で、入院レセプトではDPCの主たる医療資源を消費した病名以外には、退院時サマリの病名を電子的に収集する手法が考えられる。外来レセプトにおいて正解データである傷病名をどのように入手するかについてはさらに検討が必要であるが、特定の傷病名で実施される診療行為と、そうでない診療行為のリストを作成して分析することが必要であろう。

### ③NDB 特別抽出データの取得と処理方法の効率化に関する検討

ボトルネックになっている行程が、大きく3つあると考えられた。

i) 有識者会議の承諾から承諾通知発出の間

ii) データ返却、或いは承諾通知発出からデータ抽出作業開始の間

iii) データ抽出作業完了から提供データ格納用外付けハードディスク到着の間

これらに対し、i) は厚生労働省の事務処理、ii) は厚生労働省およびデータ抽出作業担当も含めた全体的なスケジュール調整・管理、iii) はデータ抽出作業担当からの連絡を改善することによって、それぞれ解消が図られると考えられた。

詳細は今中の分担研究報告書を参照されたい。

### ④基本データセットの作成方法に関する検討

基本データセットは、試行として2利用者にデータが提供されたが、課題となったのが、ユーザーが独自に指定する診療行為コードや医薬品コードの設定である。指定数は、厚生労働省・保険局との協議により256項目としていた。利用者は、診療行為、医薬品コードの指定にあたり、診療行為マスタは約6,700種類、医薬品マスタは

約20,000種類の中から設定しなければならなかった。このように、利用者が分析上条件設定として必要とする診療行為や医薬品は256程度よりはるかに多いことがわかる。これを解決するには、各診療行為や医薬品をカテゴリーに抽象化(指定粒度を意味的に粗くする)する必要がある。医薬品の場合、一般名化、WHO-ATC分類などの粒度に抽象化する、などの手法が必要である。

### ⑤NDB オープンデータの役割、および利用例としての国際比較データの作成に関する検討

都道府県ごとに、高齢者人口を医療需要総量を反映するものとみなして、どれだけの提供量があるかを示した。これらの指標は、治療へのアクセスの指標ともなりうると考えられる。一方で、都道府県内でも、地域間の格差が大きく、これらの指標は、各県内の格差の大きい地域を足し合わせた平均値である点に、特に留意すべきである。指標の解釈の例もコメントとして付記しているが、解釈も一例であることに留意が必要である。

NDB オープンデータを利用による国際比較データを作成は、CT の対 1000 人当たりの施行件数、MRI の対 1000 人当たりの施行件数について、国際比較を事例として OECD（経済協力開発機構）が公表している諸外国の医療の質データに関連した日本の新しい指標作成について検討した。OECD 加盟国間で比較可能なデータを作成できた。

現在、先進国は限りある財政状況の中、効率性と公平性を考慮しながら、医療および介護保険システムを運営している。我が国は、これまで、質の高い医療サービスを比較的少ない医療費で提供していると WHO(世界保健機関)等から評価されてきた。だが、増加傾向が続く医療費・介護費の対 GDP 比率は OECD(経済協力開発機構)の加盟国中で上位（第 3 位）となり、今後は一層、現在の医療・介護の質を確保しつつ、費用を適正化しなければならない課題に直面している。こうした状況は多くの先進国およびアジア諸国においても同様であり、このため OECD、WHO、国連等の国際機関は、医療及び介護分野における政策立案に資する国際統計報告として様々な HQ(ヘルスインディケータ:保健医

療指標)の迅速な提供を各国に求めているが、日本が国際機関に提出している厚生労働統計分野の項目数は少ない。

例えば、2015 年に OECD が提出を求めている 129 項目の HQ において、日本は 53 項目を提出しているが、加盟 35 カ国中 32 位と低い。ちなみに OECD 加盟国の平均提出件数は、89 項目である。そのため、日本はデータ提出状況を改善することが望ましい。NDB オープンデータから今回新しいエビデンスを作成することができたが、日本の全人口の保険医療記録から作成されたデータは、国内だけではなく諸外国に向けた情報としても価値のあるものと考えられる。

## E.結論

レセプトデータが患者ごとに連結されず、医療機関 x 保険者 x 個人 x 受診年月、に分断されたデータ単位として匿名化されているためにこれらの分断データ単位を再結合することが非常に困難になっている。この問題を解決するには保険者が厚労省にデータを匿名化して出す場合に、医療機関 x 個人 x 受診年月についてはその同一個人ごとに 1 匿名化 ID となるように処理するとともに、その匿名化 ID が保険者間での結合できるようにインフラを整備すべきである。

データ提供にかかる時間の短縮、受け取ったデータの再 DB 化が効率的に実現できるように研究者が利用するデータベースエンジンを想定したデータ提供も必要で、テキストデータでの提供には限度がある。また傷病名の精度確保に関

する研究をさらにすすめる必要もある。

さらに今後も、NDB からのデータ提供としての特別抽出、サンプリングデータセット、基本データセット、NDB オープンデータをはじめとする情報提供機能の充実と利用が必要である。

以上のように課題はまだ多いが、患者 ID の取扱い、オープンデータの利用、特別抽出データの利用を含む NDB データ活用につき、具体的課題と対応を明確にした。当研究成果が公に広く提供されることにより、今後の NDB データの提供側、およびデータの利用側にとっても、活用推進に役立つと思われる。また NDB オープンデータから生成できるデータを使用した国際統計報告は、諸外国の行政関係者や研究者への波及効果も生み出し得るものであると考える。

厚生労働省が 2009 年から収集を開始した NDB は、ヘルスケア分野における最大規模のリアルワールド・データベースであり、そこから作成された二次データである NDB オープンデータとともに、国内だけではなく諸外国に向けた情報としても価値のあるものと考えられる。

#### F.研究発表

- 1) 満武巨裕、大江和彦、今中雄一：NDB オープンデータを研究利用に活用す

る：医療技術(CT,MRI,PET)の利用に関する国際比較の試み、社会保険旬報、第 2661 巻:12-16,2016 年

- 2) 満武巨裕（「諸外国の医療ビッグデータ」、第 2 回データヘルス時代の質の高い医療の実現に向けた有識者検討会（平成 28 年 5 月 23 日）
- 3) 大江和彦.医療情報データベースの基盤整備,情報管理.2016,vol.59,no.5, p.277-283.
- 4) 大江和彦,医療ビッグデータとこれからの医療,日本臨床検査自動化学会会誌(0286-1607)41 巻 4 号,371 (2016.08).
- 5) 松居宏樹,佐藤大介,大江和彦,レセプト情報等オンサイトリサーチセンターにおける、今後の第三者提供の方向性についてレセプト情報等オンサイトリサーチセンターにおける NDB データの利用システム環境と NDB の特性に関する報告,医療情報学連合大会論文集 36 回 1 号,138-140 (2016.11).
- 6) 佐藤大介,大江和彦,医療データベース利活用の国内基盤の最新状況レセプト情報等オンサイトリサーチセンターの試行的利用について,日本薬剤疫学会学術総会抄録集,巻 22,p51(2016.11).

#### G.知的所有権の取得状況

該当なし

厚生労働行政推進調査事業費補助金  
政策科学総合研究事業(政策科学推進研究事業))

「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」  
分担研究報告書

**レセプト NDB (ナショナルデータベース) 利活用推進のための課題と対応策**

研究分担者：今中雄一（京都大学大学院医学研究科医療経済学分野 教授）

研究要旨

目的：

- 1) NDBにおける ID：同一個人の複数患者 ID 問題を定量的に評価し、患者 ID1 と患者 ID2 を統合した新たな患者 ID の生成（換言すれば、両 ID を用いた名寄せ）するアルゴリズムを開発する。また、統合患者 ID の長所と短所について整理を行なう。
- 2) NDB オープンデータの活用： NDB オープンデータが公表されており、この利活用を行う。
- 3) NDB データ活用の課題：レセプト情報等の特別抽出データについて、申出書を提出してからレセプトデータが本研究室へ提供されるまでの事務的やりとりをまとめ、レセプト情報等の速やかな提供につながるよう問題点を検討した。

方法：

- 1) 当分野が特別抽出請求により入手した、一部の疾患の医科、DPC、調剤に関する 4 年分の NDB レセプトデータを用いた。同じ患者 ID1 をもつ患者 ID2 を名寄せし、名寄せされた同じ患者 ID2 をもつ患者 ID1 を名寄せすることを繰り返し行なうアルゴリズムを開発し、統合患者 ID を生成した。  
また、入院レセプトのみ、外来レセプトのみで統合患者 ID を生成した場合と入院外来調剤レセプトを用いた場合の患者 ID の個数を比較した。
- 2) NDB オープンデータから以下の指標に関連する項目を選定し、算定回数の合計を都道府県の高齢者すなわち 65 歳以上の人口で除し（一部、総人口で除すものを追加）、都道府県ごとの指標値を算出し、グラフに表した。年齢性別についても同様に算出とグラフ化を行った。
- 3) 本研究室とレセプト情報等の提供に係る支援業務を担当する民間企業および厚生労働省との間で主にメールで行われたやりとりを整理し、ボトルネックになる行程を洗い出した。

結果・考察：

- 1) 「全患者 ID1 数」に占める「患者 ID2 を一つもつ患者 ID1 の数」は、63%であり、「全患者 ID2 数」に占める「患者 ID1 を一つもつ患者 ID2 の数」は、68%であった。



これにより、氏名の入力ゆれや改名の発生頻度は、保険者番号・記号番号の入力ゆれや保険の変更の発生頻度よりも低いものの、小さいものではないといえる。

開発した統合 ID 生成アルゴリズムを適用した結果、患者 ID1 と患者 ID2 の組合せ数に対する統合 ID 数は 42%となった。ただし、同一の氏名・生年月日・性をもつ者が複数いた場合、または、同一の保険者番号・記号番号・生年月日・性である場合（双子など）は、いずれの場合も本アルゴリズムでは同一人物とみなしてしまう点に留意が必要である。この問題に対し、統一個人 ID で同月同一医療機関から複数レセプトが発生していないかなど、より多角的に個人 ID を評価することで、個人集計の精度が一層高まることが期待される。

こうしたハッシュ化患者 ID の問題に対する根本的な解決は、全レセプトデータを管理する機関が、正確な被保険者台帳を管理運用することである。

2) 都道府県ごとに、高齢者人口を医療需要総量を反映するものとみなして、どれだけの提供量があるかを示した。これらの指標は、治療へのアクセスの指標ともなりうると思われる。一方で、都道府県内でも、地域間の格差が大きく、これらの指標は、各県内の格差の大きい地域を足し合わせた平均値である点に、特に留意すべきである。指標の解釈の例もコメントとして付記しているが、解釈も一例であることに留意が必要である。

3) ボトルネックになっている行程が、大きく 3 つあると考えられた。

①有識者会議の承諾から承諾通知発出の間

②データ返却、或いは承諾通知発出からデータ抽出作業開始の間

③データ抽出作業完了から提供データ格納用外付けハードディスク到着の間

これらに対し、①は厚生労働省の事務処理、②は厚生労働省およびデータ抽出作業担当も含めた全体的なスケジュール調整・管理、③はデータ抽出作業担当からの連絡を改善することによって、それぞれ解消が図られると考える。

研究協力者：

國澤 進 (京都大学大学院医学研究科医療経済学分野 講師)

大坪 徹也 (京都大学大学院医学研究科医療経済学分野 助教)

## A. 目的

### 1) NDBにおける ID

レセプトデータの活用に関して想定される長所のひとつとして、個人単位での受診歴の追跡可能が挙げられる。個人情報保護の観点に配慮したデータ管理運用を実施するため、各々複数の異なる被保険者情報を用いてハッシュ化することに

より生成される、患者 ID1 (使用項目:保険者番号・記号番号・生年月日・性別)

と患者 ID2 (使用項目:氏名・生年月日・性別) が用いられる。

一方、ハッシュ化対象項目において、同一人物であってもデータの入力状況が異なる可能性が指摘されている。このため、同一人物であっても異なる患者 ID が複数

にわたって生成され、個人単位での集計が困難となっている。

そこで、同一個人の複数患者 ID 問題を定量的に評価し、患者 ID1 と患者 ID2 を統合した新たな患者 ID の生成(換言すれば、両 ID を用いた名寄せ)するアルゴリズムを開発する。また、統合患者 ID の長所と短所について整理を行なう。

## 2) NDB オープンデータの活用

NDB レセプトは全国民が受療した診療内容に関する豊富な情報を有している(生活保護等公費を除く)。NDB レセプトは個票データであるが、このたび集計データとして都道府県ごとの診療行為件数がオープンデータとして公表されることになった。

医療のニーズとの関連が強いと考えられる高齢者(65歳以上)人口で、公表された診療行為件数を調整し、都道府県地域間で比較することによって、必要な医療へのアクセスの状況、新しい技術の普及状況、施設の整備状況など医療提供体制の把握の参考となる情報を得うる。

そこで、NDB オープンデータにおいて各都道府県別に公表された診療行為件数を高齢者人口で除し、医療提供量の指標とした。

超高齢社会でますます重要になると思われる脳卒中、心筋梗塞、骨折、認知症、在宅医療、リハビリテーション、がん治療(特に高度な放射線治療と高額な分子標的薬)の実施状況を算出の対象とし、地域(都道府県)および年齢性別の違いによる差違について検討する。

## 3) NDB データ活用の課題

レセプト情報等の特別抽出データについて、申出書を提出してからレセプトデータが本研究室へ提供されるまでの事務的やりとりをまとめ、レセプト情報等の速やかな提供につながるよう問題点を検討した。

## B. 方法

### 1) NDB における ID

当分野が特別抽出請求により入手した、一部の疾患の医科、DPC、調剤に関する4年分のNDBレセプトデータを用いた。同じ患者 ID1 をもつ患者 ID2 を名寄せし、名寄せされた同じ患者 ID2 をもつ患者 ID1 を名寄せすることを繰り返し行なうことで、統合患者 ID を生成した。統合患者 ID に対する患者 ID の個数を集計した。患者 ID 生成に用いる項目の入力ゆれを評価するため、入院レセプト、外来レセプト、調剤レセプトについて、統合患者 ID ごとの受診機関数と患者 ID の数の相関分析を行なった。

また、入院レセプトのみ、外来レセプトのみで統合患者 ID を生成した場合と入院外来調剤レセプトを用いた場合の患者 ID の個数を比較した。

統合患者 ID 生成の手順を以下に示す。

名寄せ準備：

患者 ID1 と患者 ID2 のユニークなペアに対してペア ID を振る。このペア ID を初期の個人番号とする。(図 1)

名寄せ手順 1：

同一の患者 ID1 を持っている個人番号の

グループに対して、すべてにグループ内の最小の個人番号を割り振る。(図 2)

名寄せ手順 2 :

同一の患者 ID2 を持っている個人番号のグループに対して、すべてにグループ内の最小の個人番号を割り振る。(図 3)

以後、名寄せ手順 1 および 2 を繰り返す。停止条件は、同一の患者 ID1 や患者 ID2 に対して複数の個人番号が存在しないこととした。(図 4 完成イメージ)

次に、入院レセプトのみまたは外来レセプトのみで統合患者 ID を生成した場合について、医科 DPC 調剤レセプトを用いた場合と比べて患者 ID の統合状況がどの程度異なるのかを比較した。

## 2) NDB オープンデータの活用

NDB オープンデータから以下の指標に関連する項目を選定し、算定回数の合計を都道府県の高齢者すなわち 65 歳以上の人口で除し(一部、総人口で除すものを追加)、都道府県ごとの指標値を算出し、グラフに表した。年齢性別についても同様に算出とグラフ化を行った。高齢者人口を使ったのは総人口よりも医療需要を反映すると考えられたためである。

ただし、オープンデータの項目は全ての診療行為コード、医薬品コードを網羅していないので、指標を表す診療行為コード、医薬品コードが不十分な場合がある。結果を見るときは、十分な注意が必

要である。

指標に利用した診療行為コード、医薬品コードは、別添 (1) に示す。

指標一覧(高齢者人口あたり)

	指標	グラフ
1	tPA	A
2	虚血性心疾患(急性心筋梗塞、狭心症含む)に対する PCI [指標 3 を含む]	B
3	安定狭心症に対する経皮的冠動脈ステント留置術(その他)	C
4	人工股関節置換術	D
5	認知症入院	E
6	重度認知症患者デイ・ケア料(1日につき)	F
7	在宅患者訪問診察料	G
	医総管(在宅時、特定施設、在宅がん)	H
	[医総管なし]在宅患者訪問診察料	I
8	訪問看護・指導料	J
6	訪問看護指示料	K
10	リハビリテーション 運動器 心大血管疾患 認知症 脳血管疾患等 がん	L M N O P
11	放射線治療管理料(IMRTによる体外照射)	Q
	体外照射(IMRT)	R
12	高額分子標的薬	S
13	外来化学療法(抗がん剤)	T

## 3) NDB データ活用の課題

当研究室は、2015 年 1 月にレセプト情報等に関する申出書(様式 1)を提出した。提供を申し出たレセプト情報等の内容は、傷病名を限定した医科、DPC、調剤の全国 4 年分のレセプトデータの特別抽出であ

る。

当研究室の申し出は第24回有識者会議(2015年3月)において承諾され、同年5月に厚生労働省より承諾通知が発出された。

なお、この申出書提出は、レセプト情報等利用期間中に行っている(※)。この場合、データ提供の時期については、厚生労働省担当者に拠れば、前回利用していたレセプト情報等を返却後、次のレセプトデータの抽出作業にかかる段取りであるということであった。

本報告書は、2015年1月にレセプト情報等に関する申出書を提出した後、レセプトデータが本研究室に到着する迄の間、本研究室とレセプト情報等の提供に係る支援業務を担当する民間企業(以下、支援業務担当と記す)および厚生労働省との主にメールでのやりとりを時系列に示す。

その後、ボトルネックになっている可能性が高い過程を洗い出し、改善につながると思われる点を記述する。

※「レセプト情報・特定健診等情報の提供に関するガイドライン」平成26年10月改訂において、「現にレセプト情報等の利用を承諾された提供依頼申出者が、レセプト情報等の利用が終了していない場合に、新たな申出を行うことは原則認めない。」ことが追記されている。(ガイドライン第6の4(3)参照)。

## C. 結果

### 1) NDBにおけるID

当該データに含まれた患者IDの種類(ユニーク)数は、患者ID1が96,819,611、患者ID2が91,746,700であった。「全患者ID1数」に占める「患者ID2を一つもつ患者ID1の数」は、63%であり、「全患者ID2数」に占める「患者ID1を一つもつ患者ID2の数」は、68%であった。

開発した名寄せアルゴリズムは、5往復で停止条件を満たした。初期状態でのペアID数は、137,903,651であったのに対し、開発した名寄せアルゴリズムを実行した結果、統合患者ID数は、57,382,498となった。また、統合個人IDに対して、42%が複数の患者ID1をもっていた。患者ID1, ID2, 統合IDの関係について、事例を示す。

患者IDの統合事例として、後期高齢者医療制度への保険の変更があった、患者ID1が2種あり、患者ID2が2種ある個人を対象として、名寄せアルゴリズムを適用した例を示す。事例1の個人は、13の医療機関のいずれかから、4年間いずれの月もレセプトが発生していた。表1の通り、事例対象者は、13のうち12の医療機関で同一の患者ID1, ID2のレセプトが発行されていたが、A医療機関のみで2種の患者ID1および異なる1種の患者ID2でレセプトが発行されていた。なお、2種の患者ID1は、ある時点を境に切替わっていた。以上より、医療機関により被保険者情報の入力規則が異なるため、同一個人に対して複数の患者IDが生じることが確認された。(表1)

次に、入院レセプトのみ、外来レセプトのみまたは調剤レセプトのみで統合患者IDを生成した場合について比較した結果を示す。

まず、入院・外来・調剤のレセプトを用いた場合の結果を表2に示す。

まず、入院・外来・調剤のレセプトを用いた場合の結果を表2に示す。

入院レセプトのみでの名寄せに対して、外来や調剤も含めて名寄せを行った効果、統合個人 ID 数の減少は、80,649 であり、その減少率は 0.49%であった。

次に、外来レセプトのみから統合患者 ID を作成した場合の結果を表 3 に示す。

外来レセプトのみでの名寄せに対して、入院や調剤も含めて名寄せを行った効果、統合個人 ID 数の減少は、96,807 であり、その減少率は 0.17%であった。

最後に、調剤レセプトのみから統合患者 ID を作成した場合の結果を表 4 に示す。

調剤レセプトのみでの名寄せに対して、入院や外来も含めて名寄せを行った効果、統合個人 ID 数の減少は、515,778 であり、その減少率は 0.96%であった。

## 2) NDB オープンデータの活用

各指標を算出しグラフ化した結果は別添 2 および別添 3 に示す。

都道府県ごとに、高齢者人口を医療需要総量を反映するものとみなして、どれだけの提供量があるかを示すものである。これらの指標は、治療へのアクセスの指標ともなりうると考えられる。

一方で、都道府県内には、地域間の格差が大きく、モザイク状態になっていることが、従来の研究より知られている。これらの指標は、各県内の格差の大きい地域を足し合わせた平均値である点に、特に留意すべきである。

以下に各指標に対するコメント例を記載する。これらのデータで現時点で十分に深く詳細な考察は困難であるので、あくまで試行的に参考例としてコメント等を作成したものである。

また、当報告で『高齢者人口当たり』とあるのは、全人口の件数を高齢者人口で除した指標を指す。その際、分子には、高齢者でない場合の件数も含まれている。

### [指標 1 グラフ A]

tPA -高齢者人口あたり

指標の概要 年齢 (65 歳以上) に対する tPA 投与について算定された回数  
高齢者人口当たりの tPA 使用状況について算定された回数

指標をみることの意義 医療資源消費の傾向や、心筋梗塞、脳梗塞治療で主に利用される薬剤の使用状況をみることができる。

コメント例 65 歳から 75 歳まででは男性の方に多く使われるが、以降は女性に増えている。年齢人口 (寿命) に影響するため、人口当たりなどの別指標も考慮できる。

地域により大きく異なる。ただし、罹患率の違い、罹患患者あたりの違い、また治療法の違い (tPA は血栓溶解療法であるが、手術等による血栓除去など) を併せて評価する必要がある。

### [指標 2 グラフ B]

虚血性心疾患 (急性心筋梗塞、狭心症含む) に対する PCI 算定回数 -総人口及び高齢者人口あたり

指標の概要 虚血性心疾患に対する治療法の 1 つである経皮的冠動脈形成術について算定された回数

指標をみることの意義 経皮的冠動脈形成術の適応となる虚血性心疾患患者の分布を間接的に把握することができる。

コメント例 PCIは男性に多く算定されており、特に65歳以下での性差が大きい。また大規模PCI実施可能施設が多い東京・大阪・神奈川・愛知のPCI算定回数が少ない印象がある。150318310、150260350、150284310はステント治療と併用して行われることに注意

#### [指標3 グラフC]

安定狭心症に対するステント留置術算定回数 -総人口及び高齢者人口あたり  
指標の概要 急性心筋梗塞・不安定狭心症以外の虚血性心疾患に対するステント留置術について算定された回数  
指標をみることの意義 急性心筋梗塞・不安定狭心症以外の虚血性心疾患の分布を把握することができる。

コメント例 ステント治療もPCIと同様に男性に多く算定されており、特に特に65歳以下の差が大きい。またPCI治療数が多い県では、ステント治療算定数も多い。

#### [指標4 グラフD]

人工股関節置換術 -高齢者人口あたり  
指標の概要 人工股関節置換手術(一回目と二回目の計)について算定された回数  
指標をみることの意義 人工股関節置換手術の実施の分布を把握することができる。

コメント例 西高東低の傾向が見える。東北及び岐阜県、静岡県、愛知県、三重県で算定回数が少ない。全年齢で女性が多い。

#### [指標5 グラフE]

認知症入院 -高齢者人口あたり  
指標の概要 精神病棟に入院している認知症患者について算定された回数  
指標をみることの意義 認知症の周辺症状(BPSD)など入院が必要となった場合の、認知症患者の受け皿の充実度がわかる。

コメント例 ・全国規模で見ると西高東低だが、より多くのニーズが予測される大都市圏・都市部の方が、低い傾向を示している。

・どの年齢層でも1.5~2倍女性が多い。政策として、都市部の受け皿の充実を促す際、根拠データとなりうる。熱心な地域もわかるので、実際の状況を調べて妥当であれば、それらの地域をbest practiceとして呈示できる可能性がある。

#### [指標6 グラフF]

重度認知症患者デイ・ケア料(1日につき) -高齢者人口あたり

指標の概要 重度認知症患者デイ・ケア料(1日につき)について算定された回数

指標をみることの意義 重度認知症患者デイ・ケアを行う施設があることがわかり、認知症患者の受け皿のみならず、介護者のレスパイトケアの充実度がわかる。

コメント例 全国的に少なく、西高東低である。より多くのニーズが予測される大都市圏、都市部の低さが際立つ。特に、東京都で圧倒的に少ないことが一目

瞭然でわかる。

政策として、都市部の受け皿の充実を促す際、根拠データとなりうる。

熱心な地域（特に沖縄、山形、大分、広島等）について、実際の状況を調べて妥当であれば、それらの地域を best practice として呈示できる可能性がある。

[指標 7 グラフ G, H, I]

在宅患者訪問診察料 - 高齢者人口あたり  
指標の概要 定期訪問時について算定された回数

但し、在宅患者訪問診察料は訪問する毎に算定、医総管は施設要件あり、月 2 回以上訪問した際に月 1 回だけ算定可能である

指標をみることの意義 医総管（在宅時、特定施設、在宅がん） [医総管なし]在宅患者訪問診察料 においては、医療機関数とあわせて評価することで、在宅医療提供体制のストラクチャー指標としての参考値になると考えられる。

コメント例 大都市圏・一部地域に算定数が多く、医療機関数・規模との関連性が推察される。医総管 vs [医総管なし]在宅患者訪問診察料の比率を評価することで、当該地域の提供体制の課題をより詳細に評価できると考えられる。

患者ごとの重複算定が生じる指標であるため、在宅医療のアウトカム指標としては看取り関連加算（看取り加算、死亡診断加算）をあわせて評価する必要がある。

在宅医療の需要は高齢人口との関連が特に深いため、人口動態による医療需要とあわせて評価が重要である。その観点からも、医療圏別・市区町村別の評価が重

要であると考えられる。

[指標 8 グラフ J]

訪問看護・指導料 - 全年齢及び高齢者人口あたり

指標の概要 医療保険による訪問看護において、訪問毎に算定された回数  
指標をみることの意義 訪問看護提供体制のストラクチャー指標の参考値として評価できる。

コメント例 訪問診療、「訪問看護指示料」との関連が少なく、指標の高低には複合的な要因が関与している推察される。

訪問看護においては、介護保険が大部分を占めるため、介護保険とあわせた評価が必須である。

[指標 9 グラフ K]

訪問看護指示料 - 全年齢及び高齢者人口あたり

指標の概要 訪問看護指示書（訪問看護内容を医師が規定するために発行される文書）を発行する際に、算定された回数。但し、医療機関 1-6 ヶ月毎に更新する必要があるが、更新間隔は不定（医療保険・介護保険双方）である

指標をみることの意義 訪問看護実施のアウトカム指標としての参考値となる。

コメント例 在宅医療指標と相関が疑われ、提供機関数と関連するものと推察される。

発行間隔が不定（任意）であることが、指標としての限界である。

[指標 10 グラフ L, M, N, O, P]

リハビリテーション -高齢者人口あたり

指標の概要 高齢者人口あたりリハビリテーションについて算定された回数  
指標をみることの意義 運動器、心大血管疾患、認知症、脳血管疾患等、がんに関するリハビリテーションの実施状況を見ることができる。

コメント例 認知症患者リハビリテーションを除いて、西高東低、とくに九州地方に多い傾向にある。

九州の中では宮崎県はいずれのリハビリテーション指標も低かった。

脳血管疾患等リハビリテーションにおいて、都道府県で3倍以上の差がみられた。当施設数や医師数などの資源量との関連が想定される。

該指標群は、プロセス指標に該当する。アウトカム指標との関係性をみることで、地域差解消のあり方について検討することが出来よう。

[指標 11 グラフ Q, R]

IMRTによる体外照射 -高齢者人口あたり

指標の概要 高齢者人口あたり放射線治療管理料（IMRTによる体外照射）について算定された回数

高齢者人口あたり体外照射（IMRT）について算定された回数

指標をみることの意義 先進的ながん治療の実施状況を把握できる。

高度な放射線治療であるIMRTを行える体制の整った医療施設の充実度がわかる。特に算定には放射線治療の常勤医2名が必要であり、放射線治療専門医などの人

的資源を確保できているかの評価に有用と期待される。

コメント例 都道府県格差が著しく、東北、北関東、九州四国に実施が少ない県が見られる一方、石川県、沖縄県が突出して多い。

全国的には西高東低を示すが、中では北陸や沖縄県で高い値を示している。

専門人材と設備整備に地域差があることがうかがえる。

必ずしも都会だけにIMRTの行える体制が集中しているわけではないことが示唆され、さらに現地の状況を調査することで、各地方でのIMRT実施体制の充実を図る上での有用な検討材料となると期待される。IMRTの対象患者は主に前立腺癌と頭頸部癌であり、これらの腫瘍の有病割合や当該診療科の体制の充実度も、別途数値に影響しうることに留意が必要であろう。

[指標 12 グラフ S]

高額分子標的薬 -高齢者人口あたり

指標の概要 高額分子標的薬であるリツキサンおよびベクティビックスの65歳以上人口千人あたりの算定回数を指標としている

年齢（65歳以上）に対する高額な分子標的薬の使用状況について算定された回数  
高齢者人口当たりの高額分子標的薬使用状況について算定された回数

指標をみることの意義 リツキサンは悪性リンパ腫において、ベクティビックスは大腸がんにおいて用いられている高額分子標的薬である。いずれも発売から一定の期間が経過し、高度専門機関の医



師のみならず、化学療法を行う一般病院の医師も広く治療に成熟しており、多くの患者が全国の医療機関で投与を受けている薬剤である。

従ってこの指標は、市販後数年が経過し、全国的に広く流通した分子標的薬の使用傾向を反映した指標と考える。

医療資源消費の傾向をみることができる。

コメント例 都道府県により算定回数には大きくバラツキがあり、最も多い県と少ない県では約二倍の差がある。しかし、高齢者人口あたり外来抗がん剤算定回数のグラフではそこまでの地域差は見られない。また、高齢者人口あたりの病床数および施設数の地域分布とも一致しない。従って、この地域差を生じる原因は、化学療法を行う施設数の違いのみならず、地域による分子標的薬の選好の違いも大きく影響しているのではないかと推測する。

どの年代でも男性が多く消費している。人口当たり、罹患者あたりなど別指標も考慮できるが、この分野において男性による需要の多さが際立つ。

地域により大きく異なる。ただし、罹患率の違い、罹患者あたりの違い、また治療法の違い（レジメンの違い等）を併せて評価する必要がある。

#### [指標 13 グラフ T]

外来抗がん剤治療 - 高齢者人口あたり  
指標の概要 外来抗がん剤の高齢者人口千人あたりの算定回数を指標としている

高齢者に対する外来抗がん剤治療実施状況について算定された回数

高齢者人口あたりの外来抗がん剤治療実施状況について算定された回数

指標をみることの意義 我が国の悪性腫瘍（がん腫特定なし）に対する外来化学療法の地域や年齢による施行状況を見る事ができる。

医療資源消費の傾向をみることができる。

コメント例 算定回数はほとんどの都道府県で65歳以上人口千人あたり6～7回となっており、地域差は比較的少なくなっている。男女で算定回数に占める65歳以上の割合に差があるのは、乳がんや婦人科がんなど、罹患する癌腫の性差によるものと考えられる。

どの年代でも男性が多く消費している。人口当たり、罹患者あたりなど別指標も考慮できるが、この分野において男性による需要の多さが際立つ。

地域により大きく異なる。ただし、罹患率の違い、罹患者あたりの違い、また治療法の違い（入院治療、外来治療の違い等）を併せて評価する必要がある。

各指標を算出しグラフ化した結果は別添2および別添3に示す。

都道府県ごとに、高齢者人口を医療需要総量を反映するものとみなして、どれだけの提供量があるかを示すものである。これらの指標は、治療へのアクセスの指標ともなりうると考えられる。

一方で、都道府県内には、地域間の格差が大きく、モザイク状態になっていることが、従来の研究より知られている。これらの指標は、各県内の格差の大きい地域を足し合わせた平均値である点に、特に留意すべきである。

表5として別に各指標に対するコメント例を記載する。これらのデータで現時点で十分に深く詳細な考察は困難であるので、あくまで試行的に参考例としてコメント等を作成したものである。

また、当報告で『高齢者人口当たり』とあるのは、全人口の件数を高齢者人口で除した指標を指す。その際、分子には、高齢者でない場合の件数も含まれている。

### 3) NDB データ活用の課題

支援業務担当および厚生労働省との主にメールでのやりとりを、全申し出者に共通する部分と当研究室の個別事例に分けてまとめた。(図5)(①②③についてはD.考察にて後述)

## D. 考察

### 1) NDBにおけるID

特別抽出データを用いて患者ID1と患者ID2のペア数を調査すると共に、患者ID1と患者ID2を用いた名寄せアルゴリズムを開発した。

患者ID1とID2のペア数より、氏名の入力ゆれの発生頻度は、保険者番号・記号番号の入力ゆれや保険の変更の発生頻度よりも低いものの、小さいものではないといえる。

開発した統合ID生成アルゴリズムを適用した結果、患者ID1と患者ID2の組合せ数に対する統合ID数は42%となった。ただし、同一の氏名・生年月日・性をもつ者が複数いた場合、または、同一の保険者番号・記号番号・生年月日・性である場合(双子など)は、いずれの場合も本アルゴリズムでは同一人物とみなしてし

まう点に留意が必要である。この問題に対し、統一個人IDで同月同一医療機関から複数レセプトが発生していないかどうかなど、より多角的に個人IDを評価することで、個人集計の精度が一層高まることが期待される。

入院レセプトのみまたは外来レセプトのみで統合患者IDを生成した場合について、医科DPC調剤レセプトを用いた場合と比べて患者IDの統合状況がどの程度異なるのかを比較した結果、入院のみ、外来のみ、調剤のみのレセプトから名寄せを行っても、すべてのレセプトを用いて名寄せを行ったものと比較して、それぞれのレセプトに現れる患者IDのペアに限れば、名寄せ結果に大きな違いはない。これは、レセプトを増やすことによって新たに患者ID1を名寄せできるパターンは図6枠内のようなブロックのみであるが、このような複雑な結合が少ないためと思われる。

一方で、入院レセプトを持つ個人番号に対し、外来・調剤レセプトも含めて紐づけられる患者ID1を広げると、その数は18,159,947から26,241,390に増える。(45%増加する)

これは、同一の個人が入院と外来・調剤で違う患者ID1が使われている可能性を示唆する。

こうしたハッシュ化患者IDの問題に対する根本的な解決は、全レセプトデータを管理する機関が、正確な被保険者台帳を管理運用することである。

### 2) NDB オープンデータの活用

前述のごとく、本解析は都道府県ごとに、高齢者人口を医療需要総量を反映す

るものとみなして、どれだけの提供量があるかを示すものである。これらの指標は、治療へのアクセスの指標ともなりうると考えられる。ただし、都道府県内には、地域間の格差が大きく、モザイク状態になっているため、これらの指標は、各県内の格差の大きい地域を足し合わせた平均値である点に、特に留意が必要である。

このように解析には限界を含めた解釈が併記されることが望ましい。本報告ではその解釈の例としてコメントも作成した。解釈もあくまで試行的なものであることも留意が必要である。

### 3) NDB データ活用の課題

今回の申出に関してデータ抽出作業のボトルネックと考えられる出来事は、下記の3つと考える。

①有識者会議の承諾から承諾通知発出の間

②データ返却、或いは承諾通知発出からデータ抽出作業開始の間

③データ抽出作業完了から提供データ格納用外付けハードディスク到着の間

これらに関わるやりとりについて、問題点を簡潔に挙げていく。

① 有識者会議の承諾から承諾通知発出の間

厚生労働省から承諾通知が発出されるまで有識者会議から7週間経っている。

現時点で既に改善されている可能性は大いにあるが、迅速な事務処理を希望したい。

② データ返却、或いは承諾通知発出

からデータ抽出作業開始の間

データ返却は、申出書提出時に NDB データを利用していない申出者にとっては、厚生労働省から承諾通知が発出される時点と同じ扱いになる。

データ抽出に必要な書類は全て支援業務担当へ提出済みであり、当研究室からのデータ返却を以て特別抽出データの抽出作業にかかることとされていたが、データ抽出作業へ着手した旨の連絡がなく、支援業務担当からマスタについての質問がくるまで、どのような進捗になっているのか不明だった。

今のところ、データ抽出の順番 (queuing) のルールや順番そのものが明らかになっていないため、何ヶ月後にデータを入手出来るのかの目処すら分からないことが問題である。大幅なデータ提供の遅れは研究スケジュールに混乱を生じさせ、さらに年度をまたぐ場合、科研費や研究に携わる者の人事異動にも影響を来す。

よって、NBD データ提供以外の業務(新規レセプトのNDB データベース登録作業、厚生労働省の定型集計作業等々)との兼ね合いを含めたスケジュール調整・管理を、支援業務担当もしくは厚生労働省がルールに則って行い、申出者が「自分が順番の何番目で、(最長、最短)何ヶ月後にデータが入手出来るか」おおよその見込みを把握出来るよう希望する。

また申出書を検討している者が、「次の有識者会議で承諾された場合、データ提供までおおよそ(最長)何ヶ月待ちか」が分かるようになっていけば更に実用的だと考える。

③ データ抽出完了から外付けハードディスク到着の間

データ抽出作業が終わってからデータ受取用外付け HDD チェックリストが送付された。少々特殊なスペックを要求されたため、チェックリストに該当する外付け HDD を調達するのに 10 日程を要した。

提供されるデータ量が分かった段階でチェックリストが送付されれば、データ抽出作業完了に間に合うよう器機が調達できるので早めの連絡を希望したい。また、チェックリスト可否の返答が支援業務担当ではなく、データ抽出担当である所為か、可否返答が遅く、待たされた日数だけデータ到着が遅れている。データ抽出担当にも迅速な処理に協力願いたい。

また、3つのボトルネック以外の問題点について以下に挙げる。

- ・申出者提出のマスタと不一致になった保険者や医療機関コードを抽出前に教えていただいたのは、大変助かった。

しかし、保険者マスタや医療機関マスタは殆どの利用者が利用すると思われる。よって、例えば厚生労働省のホームページで公表されている申出依頼テンプレートの注意点(2015年4月1日版)では、目的別DBに医療機関所在地に対応した二次医療圏を振っている。これらの作業に用いられているマスタの閲覧・提供を検討して欲しい。

- ・年度始め、支援業務担当との契約更新時(契約の空き期間)、厚生労働省へ直接

提出した書類が1点のみであるが、支援業務担当へ伝わっていなかった。レセプト情報等申請に必要な書類は一括して支援業務担当が管轄すべきであり、契約の空き期間に関しては、書類提出の猶予を設ける等の措置を講じて欲しい。さもなければ契約の空き期間を作らないで欲しい。

また、もし支援業務担当が別会社が変わるならば、変わった時点で利用者にも報告が欲しい。

## E. 結論

### 1) NDBにおける ID

患者 ID1 と患者 ID2 を用いた名寄せアルゴリズムを開発した。これにより個人単位での集計精度を高めることが可能となる。ただし、同一の氏名・生年月日・性をもつ者が複数いた場合、または、同一の保険者番号・記号番号・生年月日・性である場合(双子など)は、いずれの場合も本アルゴリズムでは同一人物とみなしてしまう点に留意が必要である。この問題に対し、統一個人 ID で同月同一医療機関から複数レセプトが発生していないかどうかなど、より多角的に個人 ID を評価することで、個人集計の精度が一層高まることが期待される。

また、特定のレセプト(例えば入院レセプト)のみを対象とした研究計画にもとづきデータを受領した者が、各自で統合個人 ID を作成する際に、そのほかのレセプト(入院外・調剤レセプトなど)にある患者 ID 情報を加えなかったとしても、同情報を加えた場合と比べて作成された統合個人 ID 数に大きな差異は生じないことが期待される。

## 2) NDB オープンデータの活用

NDB オープンデータとして公表された診療行為件数を、医療需要との関連が強いと考えられる高齢者人口で除した指標を、医療提供量を表す指標として、都道府県で算出し比較した。都道府県内に大きな地域間格差があるであろうことを留意しておく必要がある。

これらの指標により、都道府県別に、必要な医療へのアクセス、新しい技術の普及、施設の整備など医療提供体制の把握の参考となりうる情報を得られることが示唆された。

ただし、NDB オープンデータに全ての診療行為(薬剤含む)が網羅されていないので、医療機能別の利用状況を把握するにはデータベースの改善が必要である。

## 3) NDB データ活用の課題

レセプト情報等提供申し出書類の作成に関しては、ガイドラインに記された通りに進めればさほど困難ではない。

しかし、保険者と二次医療圏、医療機関所在地と二次医療圏、医療機関と病床数のような全ての申請者が利用すると思われるマスタについては、共用出来るマスタが NDB データ提供用限定としてでも提供がされれば、申請者のタイムコスト削減に繋がる。

申出が承諾された後は、迅速な抽出処理が望まれる。

承諾後、データ抽出まで待ち時間を要する場合、おおよその待ち時間が把握出来れば、申出者は状況に沿った研究スケジュールを立てることが出来る。待ちの順番、データ抽出完了までのおおよその

月数、現在の進捗状況等を確認できる仕組みを希望したい。それにより申出者は自身の研究スケジュールに無為な空きを作ることなく、時間配分を最適化することが出来る。

次に、データ抽出フェーズに入った後、抽出データの容量が判別した時点で支援業務担当から申出者へレセプトデータ格納用の HDD チェックリストが送付されれば、データ抽出中に器機の調達が出来、受領までの数日が短縮できる。

これらのデータ抽出作業進捗の見える化、HDD チェックリスト送付は、支援業務担当のみならず、データ抽出作業担当にも協力を願いたい。

その他全般の事務手続きに関しては、年度始めに支援業務担当の空白期間をつくらぬよう、もしくは支援業務担当との契約が切れている間は一切の書類手続きを凍結させる等、一元管理を徹底して欲しい。

## F. 健康危険情報

特になし

## G. 研究発表

特になし

別添 1：(別添 1)指標算出に用いたコード一覧

別添 2：(別添 2)NDB オープンデータグラフ\_都道府県および

別添 3：(別添 3)NDB オープンデータグラフ\_年齢性別

(図 1)

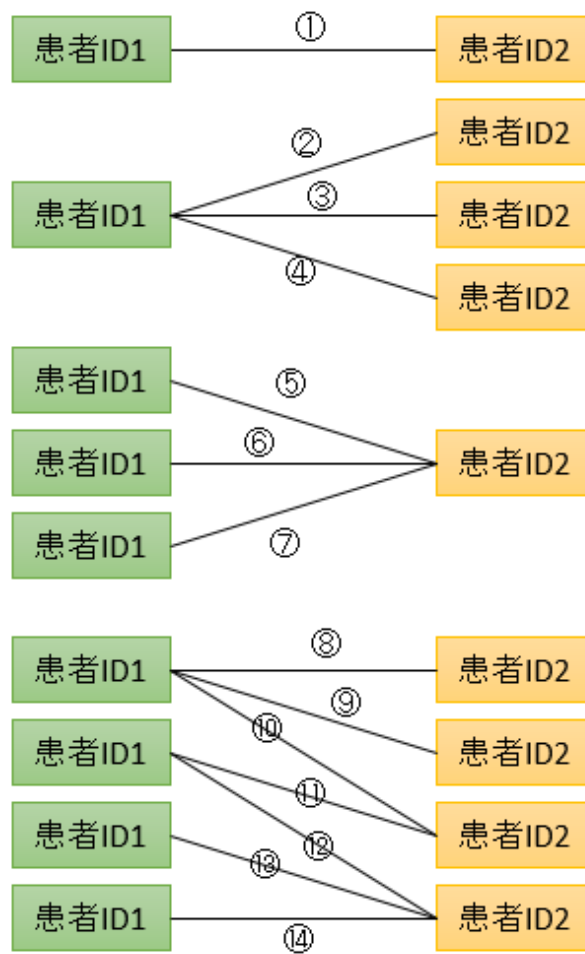


図 1 初期状態イメージ

図 2

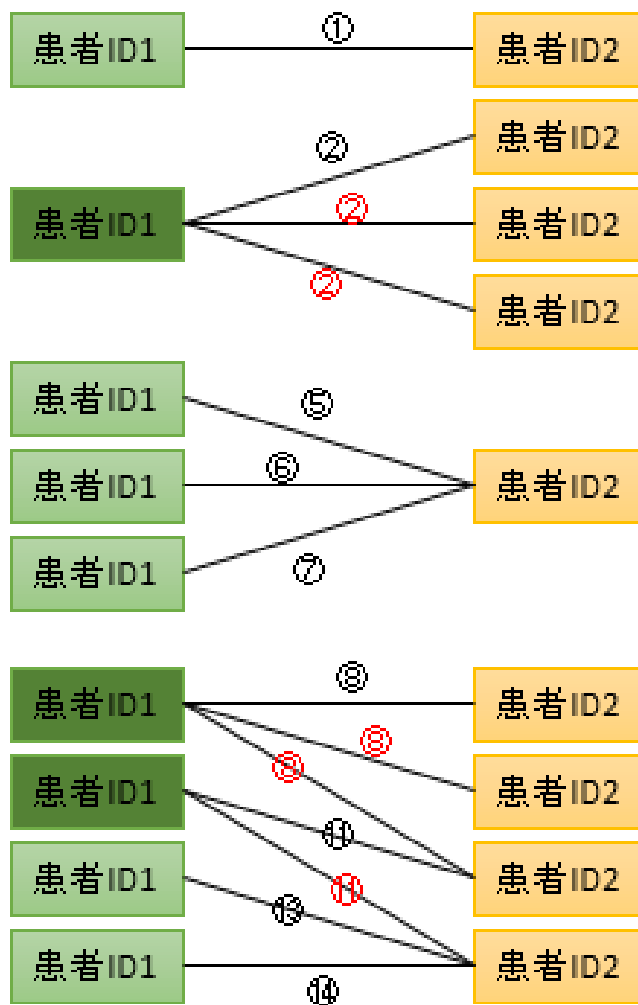


図 2 名寄せ手順1イメージ

図3

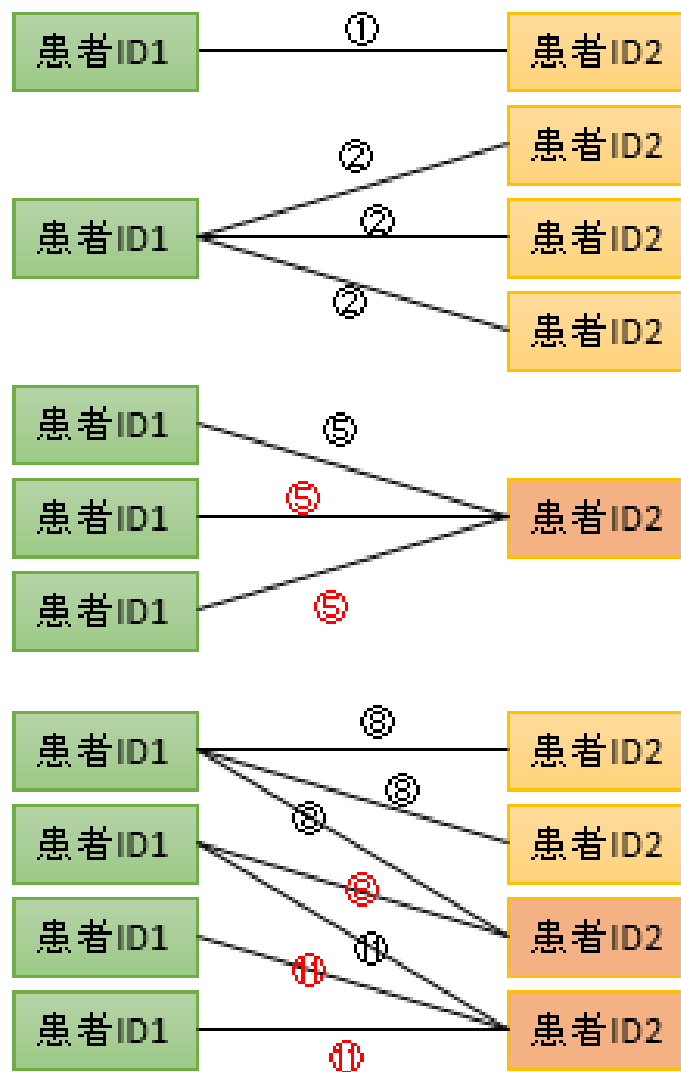


図 3 名寄せ手順 2 イメージ



図 4

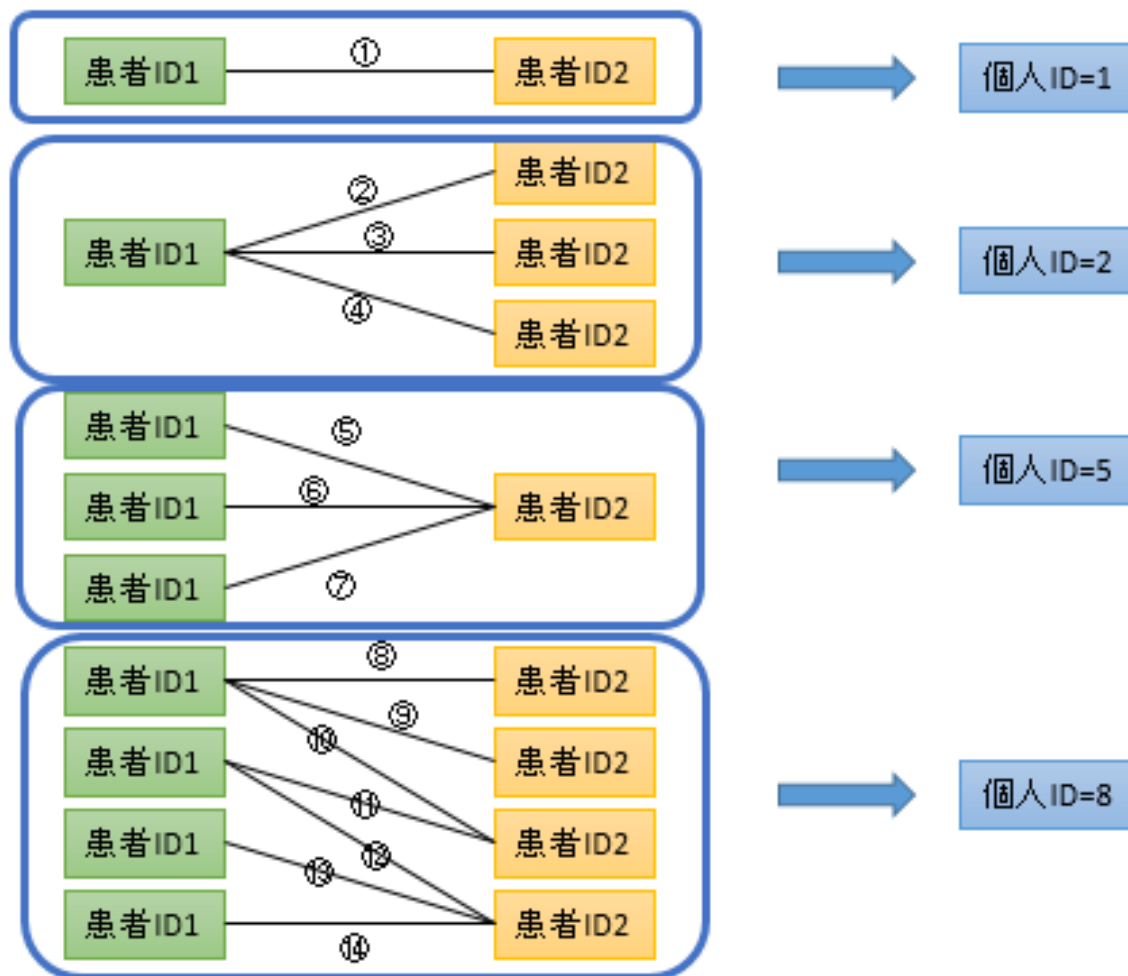


図 4 完成イメージ

表 1 ある統合患者 ID に関する 4 年間の患者 ID1, ID2 の発生月数

医療機関	患者ID1	患者ID2	診療月数
A	X1	Y1	28
A	X2	Y1	20
B	X1	Y2	45
C	X1	Y2	21
D	X1	Y2	44
E	X1	Y2	8
F	X1	Y2	1
G	X1	Y2	1
H	X1	Y2	22
I	X1	Y2	2
J	X1	Y2	1
K	X1	Y2	7
L	X1	Y2	5
M	X1	Y2	5

表2 入院・外来・調剤のレセプトを用いた場合の結果を以下に示す。

ペアID	統合個人ID	患者ID1	患者ID2	患者ID1*	患者ID2
137,903,651	57,382,498	96,819,611	91,746,700	68.73%	59.89%

\*: 複数患者 ID1 発生率=(患者 ID1-個人番号)/個人番号

入院レセプトのみから統合患者 ID を作成した場合の結果を以下に示す。

入院レセプトのみから取得した患者 ID ペア

ペア数	統合個人ID	患者ID1	患者ID2	患者ID1*	患者ID2
19,368,873	16,514,720	18,159,947	17,791,173	9.96%	7.73%

外来・調剤も含めて上記の統合個人 ID に紐づけられた患者 ID1 と ID2

ペア数	統合個人ID	患者ID1	患者ID2	患者ID1*	患者ID2
38,573,851	16,514,720	26,241,390	26,920,821	58.90%	63.01%

入院レセプトのみを用いて行った名寄せの結果

ペアID	統合個人ID	患者ID1	患者ID2	患者ID1*	患者ID2
19,368,873	16,595,369	18,159,947	17,791,173	9.43%	7.21%

表3 来レセプトのみから統合患者 ID を作成した場合の結果を以下に示す。

外来レセプトのみから取得した患者 ID ペア

ペア数	統合個人ID	患者ID1	患者ID2	患者ID1	患者ID2
123,683,799	56,737,219	92,793,837	84,084,530	63.55%	48.20%

入院・調剤も含めて上記の統合個人 ID に紐づけられた患者 ID1 と ID2

ペア数	統合個人ID	患者ID1	患者ID2	患者ID1	患者ID2
137,149,387	56,737,219	96,097,846	91,069,190	69.37%	60.51%

外来レセプトのみを用いて行った名寄せの結果

ペアID	統合個人ID	患者ID1	患者ID2	患者ID1	患者ID2
123,683,799	56,834,026	92,793,837	84,084,530	63.27%	47.95%

表 4 調剤レセプトのみから統合患者 ID を作成した場合の結果を以下に示す。

調剤レセプトのみから取得した患者 ID ペア

ペア数	統合個人ID	患者ID1	患者ID2	患者ID1	患者ID2
95,774,511	53,201,598	82,116,948	65,874,541	54.35%	23.82%

入院・外来も含めて上記の統合個人 ID に紐づけられた患者 ID1 と ID2

ペア数	統合個人ID	患者ID1	患者ID2	患者ID1	患者ID2
131,397,918	53,201,598	91,244,783	86,688,379	71.51%	62.94%

調剤レセプトのみを用いて行った名寄せの結果

ペアID	統合個人ID	患者ID1	患者ID2	患者ID1	患者ID2
95,774,511	53,717,376	82,116,948	65,874,541	52.87%	22.63%

表5 各指標に対するコメント例

[指標 1 グラフ A]

tPA -高齢者人口あたり

指標の概要	年齢(65歳以上)に対するtPA投与について算定された回数 高齢者人口当たりのtPA使用状況について算定された回数
指標をみるこ との意義	医療資源消費の傾向や、心筋梗塞、脳梗塞治療で主に利用される薬剤の使用 状況を確認することができる。
コメント例	65歳から75歳まででは男性の方に多く使われるが、以降は女性に増えている。 年齢人口(寿命)に影響するため、人口当たりなどの別指標も考慮できる。 地域により大きく異なる。ただし、罹患率の違い、罹患率あたりの違い、また治 療法の違い(tPAは血栓溶解療法であるが、手術等による血栓除去など)を併 せて評価する必要がある。

[指標 2 グラフ B]

虚血性心疾患(急性心筋梗塞、狭心症含む)に対する PCI 算定回数 -総人口及び高齢者人  
口あたり

指標の概要	虚血性心疾患に対する治療法の1つである経皮的冠動脈形成術について算定 された回数
指標をみるこ との意義	経皮的冠動脈形成術の適応となる虚血性心疾患患者の分布を間接的に把握 することができる。
コメント例	PCIは男性に多く算定されており、特に65歳以下での性差が大きい。また大規 模PCI実施可能施設が多い東京・大阪・神奈川・愛知のPCI算定回数が少ない 印象がある。 150318310、150260350、150284310はステント治療と併用して行われることに注 意

[指標 3 グラフ C]

安定狭心症に対するステント留置術算定回数 -総人口及び高齢者人口あたり

指標の概要	急性心筋梗塞・不安定狭心症以外の虚血性心疾患に対するステント留置術に ついて算定された回数
指標をみるこ との意義	急性心筋梗塞・不安定狭心症以外の虚血性心疾患の分布を把握することがで きる。
コメント例	ステント治療もPCIと同様に男性に多く算定されており、特に特に65歳以下の差 が大きい。またPCI治療数が多い県では、ステント治療算定数も多い。

[指標 4 グラフ D]

人工股関節置換術 -高齢者人口あたり

指標の概要	人工股関節置換手術(一回目と二回目の計)について算定された回数
指標をみることの意義	人工股関節置換手術の実施の分布を把握することができる。
コメント例	西高東低の傾向が見える。 東北及び岐阜県、静岡県、愛知県、三重県で算定回数が少ない。 全年齢で女性が多い。

[指標 5 グラフ E]

認知症入院 -高齢者人口あたり

指標の概要	精神病棟に入院している認知症患者について算定された回数
指標をみることの意義	認知症の周辺症状(BPSD)など入院が必要となった場合の、認知症患者の受け皿の充実度がわかる。
コメント例	・全国規模で見ると西高東低だが、より多くのニーズが予測される大都市圏・都市部の方が、低い傾向を示している。 ・どの年齢層でも1.5～2倍女性が多い。 政策として、都市部の受け皿の充実を促す際、根拠データとなりうる。 熱心な地域もわかるので、実際の状況を調べて妥当であれば、それらの地域をbest practiceとして呈示できる可能性がある。

[指標 6 グラフ F]

重度認知症患者デイ・ケア料(1日につき) -高齢者人口あたり

指標の概要	重度認知症患者デイ・ケア料(1日につき)について算定された回数
指標をみることの意義	重度認知症患者デイ・ケアを行う施設があることがわかり、認知症患者の受け皿のみならず、介護者のレスパイトケアの充実度がわかる。
コメント例	全国的に少なく、西高東低である。より多くのニーズが予測される大都市圏、都市部の低さが際立つ。特に、東京都で圧倒的に少ないことが一目瞭然でわかる。 政策として、都市部の受け皿の充実を促す際、根拠データとなりうる。 熱心な地域(特に沖縄、山形、大分、広島等)について、実際の状況を調べて妥当であれば、それらの地域をbest practiceとして呈示できる可能性がある。

[指標 7 グラフ G,H,I]

在宅患者訪問診察料 -高齢者人口あたり

指標の概要	定期訪問時について算定された回数 但し、在宅患者訪問診察料は訪問する毎に算定、医総管は施設要件あり、月2回以上訪問した際に月1回だけ算定可能である
指標をみることの意義	医総管(在宅時、特定施設、在宅がん) [医総管なし]在宅患者訪問診察料 においては、医療機関数とあわせて評価することで、在宅医療提供体制のストラクチャー指標としての参考値になると考えられる。
コメント例	大都市圏・一部地域に算定数が多く、医療機関数・規模との関連性が推察される。医総管vs [医総管なし]在宅患者訪問診察料の比率を評価することで、当該地域の提供体制の課題をより詳細に評価できると考えられる。 患者ごとの重複算定が生じる指標であるため、在宅医療のアウトカム指標としては看取り関連加算(看取り加算、死亡診断加算)をあわせて評価する必要がある。 在宅医療の需要は高齢人口との関連が特に深いため、人口動態による医療需要とあわせた評価が重要である。その観点からも、医療圏別・市区町村別の評価が重要であると考えられる。

[指標 8 グラフ J]

訪問看護・指導料 -全年齢及び高齢者人口あたり

指標の概要	医療保険による訪問看護において、訪問毎に算定された回数
指標をみることの意義	訪問看護提供体制のストラクチャー指標の参考値として評価できる。
コメント例	訪問診療、「訪問看護指示料」との関連が少なく、指標の高低には複合的な要因が関与している推察される。 訪問看護においては、介護保険が大部分を占めるため、介護保険とあわせた評価が必須である。

[指標 9 グラフ K]

訪問看護指示料 -全年齢及び高齢者人口あたり

指標の概要	訪問看護指示書(訪問看護内容を医師が規定するために発行される文書)を発行する際に、算定された回数。 但し、医療機関1-6ヶ月毎に更新する必要があるが、更新間隔は不定(医療保険・介護保険双方)である
指標をみることの意義	訪問看護実施のアウトカム指標としての参考値となる。



コメント例	在宅医療指標と相関が疑われ、提供機関数と関連するものと推察される。発行間隔が不定(任意)であることが、指標としての限界である。
-------	---

[指標 10 グラフ L,M,N,O,P]

リハビリテーション -高齢者人口あたり

指標の概要	高齢者人口あたりリハビリテーションについて算定された回数
指標をみることの意義	運動器、心大血管疾患、認知症、脳血管疾患等、がんに関するリハビリテーションの実施状況を見ることができる。
コメント例	認知症患者リハビリテーションを除いて、西高東低、とくに九州地方に多い傾向にある。 九州の中では宮崎県はいずれのリハビリテーション指標も低かった。 脳血管疾患等リハビリテーションにおいて、都道府県で3倍以上の差がみられた。 当施設数や医師数などの資源量との関連が想定される。 該指標群は、プロセス指標に該当する。アウトカム指標との関係を見ることで、地域差解消のあり方について検討することが出来よう。

[指標 11 グラフ Q,R]

IMRTによる体外照射 -高齢者人口あたり

指標の概要	高齢者人口あたり放射線治療管理料(IMRTによる体外照射)について算定された回数 高齢者人口あたり体外照射(IMRT)について算定された回数
指標をみることの意義	先進的ながん治療の実施状況を把握できる。 高度な放射線治療であるIMRTを行える体制の整った医療施設の充実度がわかる。 特に算定には放射線治療の常勤医2名が必要であり、放射線治療専門医などの人的資源を確保できているかの評価に有用と期待される。
コメント例	都道府県格差が著しく、東北、北関東、九州四国に実施が少ない県が見られる一方、石川県、沖縄県が突出して多い。 全国的には西高東低を示すが、中では北陸や沖縄県で高い値を示している。 専門人材と設備整備に地域差があることがうかがえる。 必ずしも都会だけにIMRTの行える体制が集中しているわけではないことが示唆され、さらに現地の状況を調査することで、各地方でのIMRT実施体制の充実を図る上での有用な検討材料となると期待される。 IMRTの対象患者は主に前立腺癌と頭頸部癌であり、これらの腫瘍の有病割合や当該診療科の体制の充実度も、別途数値に影響しうることに留意が必要であろう。

[指標 12 グラフ S]

高額分子標的薬 -高齢者人口あたり

指標の概要	<p>高額分子標的薬であるリツキサソおよびベクティビックスの65歳以上人口千人あたりの算定回数を指標としている</p> <p>年齢(65歳以上)に対する高額な分子標的薬の使用状況について算定された回数</p> <p>高齢者人口当たり的高額分子標的薬使用状況について算定された回数</p>
指標をみることの意義	<p>リツキサソは悪性リンパ腫において、ベクティビックスは大腸がんにおいて用いられている高額分子標的薬である。いずれも発売から一定の期間が経過し、高度専門機関の医師のみならず、化学療法を行う一般病院の医師も広く治療に成熟しており、多くの患者が全国の医療機関で投与を受けている薬剤である。従ってこの指標は、市販後数年が経過し、全国的に広く流通した分子標的薬の使用傾向を反映した指標と考える。</p> <p>医療資源消費の傾向をみることができる。</p>
コメント例	<p>都道府県により算定回数には大きくバラツキがあり、最も多い県と少ない県では約二倍の差がある。しかし、高齢者人口あたり外来抗がん剤算定回数のグラフではそこまでの地域差は見られない。また、高齢者人口あたりの病床数および施設数の地域分布とも一致しない。従って、この地域差を生じる原因は、化学療法を行う施設数の違いのみならず、地域による分子標的薬の選好の違いも大きく影響しているのではないかと推測する。</p> <p>どの年代でも男性が多く消費している。人口当たり、罹患者あたりなど別指標も考慮できるが、この分野において男性による需要の多さが際立つ。</p> <p>地域により大きく異なる。ただし、罹患者率の違い、罹患者あたりの違い、また治療法の違い(レジメンの違い等)を併せて評価する必要がある。</p>

[指標 13 グラフ T]

外来抗がん剤治療 -高齢者人口あたり

指標の概要	<p>外来抗がん剤の高齢者人口千人あたりの算定回数を指標としている</p> <p>高齢者に対する外来抗がん剤治療実施状況について算定された回数</p> <p>高齢者人口あたりの外来抗がん剤治療実施状況について算定された回数</p>
指標をみることの意義	<p>我が国の悪性腫瘍(がん腫特定なし)に対する外来化学療法の地域や年齢による施行状況を見る事ができる。</p> <p>医療資源消費の傾向をみることができる。</p>
コメント例	<p>算定回数はほとんどの都道府県で65歳以上人口千人あたり6~7回となっており、地域差は比較的少なくなっている。男女で算定回数に占める65歳以上の割合に差があるのは、乳がんや婦人科がんなど、罹患する癌腫の性差によるものと考えられる。</p> <p>どの年代でも男性が多く消費している。人口当たり、罹患者あたりなど別指標も考慮できるが、この分野において男性による需要の多さが際立つ。</p> <p>地域により大きく異なる。ただし、罹患者率の違い、罹患者あたりの違い、また治療法の違い(入院治療、外来治療の違い等)を併せて評価する必要がある。</p>

図5 支援業務担当および厚生労働省とのやりとり

年月	From	To	マイルストーン 全申出者共通	マイルストーン 当研究室個別
2015年1月	医療経済学分野	支援業務担当	申出書提出	
2015年2月				
2015年3月	第24回有識者会議 平成26年7月19日～平成27年1月16日受付分の審査			1
2015年4月				
2015年5月	支援業務担当	医療経済学分野	(厚生労働省)承諾通知発出 ※有識者会議から7週間	
2015年6月				マスタ類改訂作業
2015年7月				
2015年8月				
2015年9月				
2015年10月				
2015年11月				
2015年12月	医療経済学分野	支援業務担当		マスタ類再提出
	医療経済学分野	支援業務担当		承諾の付加条件:内部監査(自己点検)規程を改訂
	支援業務担当	医療経済学分野		(倫理審査)倫理委員会承諾書、タイトル不一致のため厚生労働省に受理されず
2016年1月				(倫理審査)新たに倫理審査申請
	医療経済学分野	支援業務担当		(倫理審査)倫理委員会承諾書(表紙写し)を提出
	医療経済学分野	支援業務担当		申出書へ公的研究費を追記
2016年2月	支援業務担当	医療経済学分野		申出書へ公的研究費を追記、厚労省確認
2016年3月				
2016年4月	医療経済学分野	厚生労働省	前回 NDB データ返却	2
	医療経済学分野	厚生労働省		データ早期提供お願い
	厚生労働省	医療経済学分野		データ早期提供お願いへの返答
	支援業務担当	医療経済学分野		保険者マスタ重複コード確認依頼
			※データ返却から2週間	
2016年5月	医療経済学分野	支援業務担当		改訂保険者番号マスタ提出
	医療経済学分野	厚生労働省		データ提供時期問い合わせ
	厚生労働省	医療経済学分野		データ提供時期返答

2016年6月	支援業務担当	医療経済学分野	二次医療圏コードが不明な保険者、医療機関の扱いについて確認依頼
	支援業務担当	医療経済学分野	改訂保険者マスタ、医療機関マスタ提出

2016年7月	支援業務担当	医療経済学分野	データ抽出完了
---------	--------	---------	---------

3

※データ返却から13週間

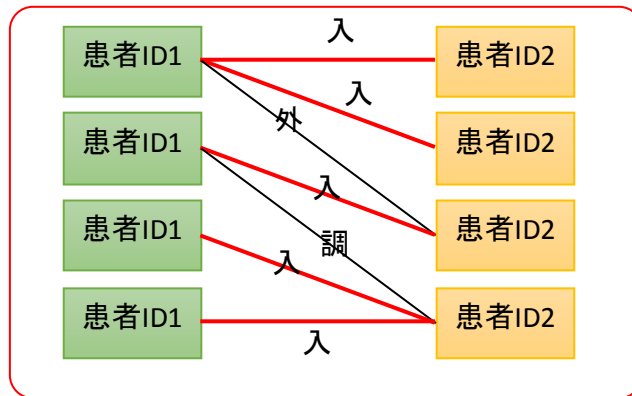
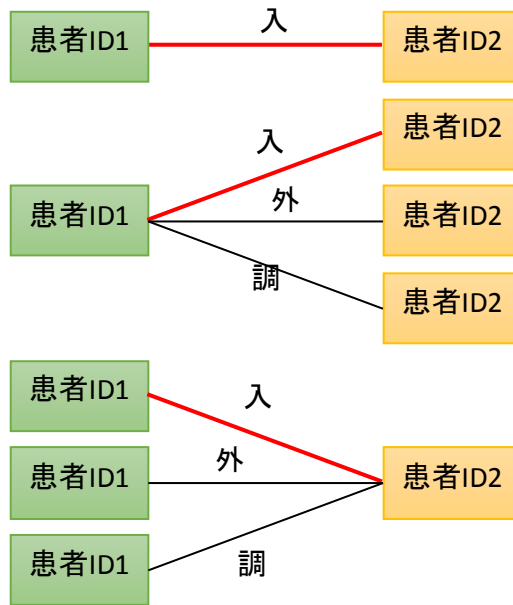
医療経済学分野	支援業務担当		データ受取用外付けHDD準備
医療経済学分野	支援業務担当		データ受取用外付けHDDチェックリスト送付
支援業務担当	医療経済学分野		データ受取用外付けHDDチェックリストに不対応なHDDのため、再検討依頼
医療経済学分野	支援業務担当		データ受取用外付けHDDチェックリスト再送付

2016年8月	医療経済学分野	支援業務担当	データ受取用外付けHDDチェックリスト可否返答について催促メール
---------	---------	--------	----------------------------------

医療経済学分野	支援業務担当	外付けHDD送付
支援業務担当	医療経済学分野	外付けHDD到着

※データ返却から20週間

図 6



厚生労働行政推進調査事業費補助金  
(政策科学総合研究事業(政策科学推進研究事業))

分担研究報告書

レセプト情報・特定健診等情報データベースの情報提供機能の充実と利用

満武 巨裕

一般財団法人 医療経済研究・社会保険福祉協会 医療経済研究機構、副部長

**研究要旨**

本報告書は、今後の日本におけるレセプト情報・特定健診等情報データベース(以下、NDB)の情報提供機能の充実と利用について検討する。日本のNDBは、提供開始(2011年11月)から本年度までの承諾件数が100件を超えたが、日本と同様の診療報酬点数制度を採用している台湾と韓国の年間提供件数が100件を超えている。したがって、研究利用として提供している特別抽出、サンプリングデータ、基本データセット、NDBオープンデータの量と質の増加が必要である。本年度は、第一に基本データセットの作成から見えた課題、第二にNDBオープンデータを利用による国際比較データを作成について検討した。

第一の基本データセットは、一定期間のレセプトを個人毎に統合し、最小限のデータ項目しか含まないデータセットである。ただし、利用者が指定した診療行為コードや医薬品コードの有無のフラグが含まれる。基本データセット作成のためのデータソースは、NDBデータセンターの抽出時間や他の業務への影響等を鑑み、全レセプトデータを提供することはしないとの厚生労働省・保険局の条件のもと、対象年が2009年度から2011年度であり、地域保険をはじめとする都道府県での区分が可能な保険者(協会けんぽ、市町村国保、後期高齢者等)は47都道府県を8ブロックに分け、各ブロックから都道府県をランダムに選定している。職域保険をはじめとする全国にまたがる保険者(組合健保、共済組合等)は、全保険者リストの作成後約30%をランダムに選定している。試行として2利用者にデータを提供した結果、課題となったのが、ユーザーが独自に指定する診療行為コードや医薬品コードの設定である。病名、診療行為、医薬品コードの指定にあたり、診療行為マスタは約6,700種類、医薬品マスタは約20,000種類の中から設定しなければならない。例えば、医薬品の血圧降下剤については100種類以上の医薬品コードが存在する。さらに、設定数については厚生労働省・保険局との協議により256項目としていたことも課題である。結果、医薬品コードと処置コードの指定に関しては、一般名などでグループ化したコードの有無のフラグにより、設定数に対する対策とした。

第二のNDBオープンデータを利用による国際比較データを作成については、OECD(経済協力開発機構)が公表している諸外国の医療の質データを作成できるか検討した。NDBオープンデータは、調査分析ワーキング・グループや厚生労働省・保険局が限られた資

源の中で公開に至った情報であり、その中からどのような新しいエビデンスを作成することができるのか検討した。結果として、CT,MRI,PETの対1000人当たりの試行件数について、OECD加盟国間で比較可能なデータを作成できた。

NDBデータを使いこなすには、基本データセットであってもNDBオープンデータであっても、コードの指定に対して分野横断的な専門性が求められる。今後は、コードの指定などについては、各利用者の経験を公開し共有化することが一つの解決策だと思われる。

厚生労働省が2009年から収集を開始したNDBは、ヘルスケア分野における最大規模のデータベースとなっており、日本の全人口を対象に、保険医療が行われた全記録であるために、そこから作成されたデータは、国内だけではなく諸外国に向けた情報としても価値のあるものと考えられる。

## A. 研究目的

本報告書は、今後の日本におけるレセプト情報・特定健診等情報データベース(以下、NDB)の情報提供機能の充実と利用について検討する。

日本のNDBは、提供開始(2011年11月)から本年度までの承諾件数が100件を超えたが、日本と同様の診療報酬点数制度を採用している台湾と韓国の年間提供件数が100件を超えている状況である。従って、現在NDBの研究利用として提供している特別抽出、サンプリングデータ、基本データセット、NDBオープンデータの量と質の増加が必要である。

本年度は、第一に基本データセットの作成から見えた課題、第二にNDBオープンデータを利用による国際比較データを作成について検討した。

## B. 研究方法

第一の基本データセット作成のデータソースは、厚生労働省・保険局と検討によりNDBデータセンタの抽出時間や他の業務への影響等を鑑み、全レセプトデータを

提供することはしないとの条件のもと、対象年は2009年度から2011年度であり、地域保険をはじめとする都道府県での区分が可能な保険者(協会けんぽ、市町村国保、後期高齢者等)は、47都道府県を8ブロックに分け、各ブロックから都道府県をランダムに選定している。職域保険をはじめとする全国にまたがる保険者(組合健保、共済組合等)は、全保険者リストの作成後約30%をランダムに選定している。ただし、都道府県情報および保険者の法別番号は、元データを保持した。

今回作成した基本データセットは、被保険者番号を基に厚生労働省が暗号化したIDであるID1をベースとして作成した。基本データセットは、一定期間のレセプトを個人ID毎に統合し、最小限のデータ項目(ID、性別、年齢等)で構成されるデータセットをはじめとするが、利用者が設定した診療行為コードや医薬品コードの存在の有無を示すフラグが含まれる。

第二のNDBオープンデータを利用による国際比較データを作成は、OECD、WHO、国連等の国際機関が迅速な提供を各国に

求めている HQ(ヘルスインディケータ:保健医療指標)の中から、新しい医療技術の普及は、医療の質を向上と同時に医療費増加の要因の一つとしても注目されている、CT(Computed Tomography: コンピューター断層撮影)、MRI (Magnetic Resonance Imaging: 磁気共鳴画像)、PET (Positron Emission Tomography: ポジトロン断層撮影)の3つの画像診断技術を対象とした。例えば OECD の OECD.stat (<http://stats.oecd.org/index.aspx>) から一般公開され、国際比較データが作成されている。特に、日本の一人当たり CT と MRI 装置の数は、第一位であるが、日本国民に対して何件の画像検査が行われているかについては、公開されてこなかった(CT に関しては OECD 加盟国の 19 カ国、MRI は 22 カ国がデータを公開している)。

データソースは、NDB オープンデータの E 画像診断の集計表を利用した(第二部:資料編として厚生労働省ホームページからダウンロード可能)。画像診断集計表から、CT(コンピューター断層撮影)、MRI(磁気共鳴コンピューター断層撮影)、PET(ポジトロン断層撮影)の3つを対象とした。

CT は、分類コードが「E200」、診療行為としては「CT撮影(その他)」、「CT撮影(16列以上64列未満マルチスライス型機器)」、「脳槽CT撮影(造影含む)」、「CT撮影(4列以上16列未満マルチスライス型機器)」、「CT撮影(64列以上マルチスライス型機器)」の5つの総計欄を合計した。

MRI は、分類コードが「E202」、診療行為が「MRI撮影(その他)」、「MRI撮影(1.5テスラ以上の機器)」、「MRI撮影(3テスラ以上の機器)」の3つの総計欄を合計した。

PET は、分類コードが「E101-1」「E101-2」「E101-3」、診療行為が「ポジトロン断層撮影(15O標識ガス使用)」、「ポジトロン断層撮影(18FDG使用)」、「ポジトロン断層撮影(13N標識アンモニア剤使用)」、「ポジトロン・コンピューター断層複合撮影(15O標識ガス使用)」、「ポジトロン・コンピューター断層複合撮影(18FDG使用)」、「ポジトロン・磁気共鳴コンピューター断層複合撮影(18FDG使用)」の6つの総計欄を合計した。

## C. 研究結果

第一の基本データセットは、試行として2利用者にデータを提供したが、課題となったのが、ユーザーが独自に指定する診療行為コードや医薬品コードの設定である。病名、診療行為、医薬品コードの指定にあたり、診療行為マスタは約6,700種類、医薬品マスタは約20,000種類の中から設定しなければならない。例えば、医薬品の血圧降下剤については100種類以上の医薬品コードが存在する。さらに、設定数については厚生労働省・保険局との協議により256項目としていたことも課題である。結果、医薬品コードと処置コードの指定に関しては、一般名などでグループ化したコードの有無のフラグにより、設定数に対する対策とした。

第二の NDB オープンデータを利用による国際比較データを作成は、CT の対 1000 人当たりの施行件数を計算したところ、日本は 229 件で 2 位となった(第 1 位が米国の 240 件、3 位がルクセンブルグである)。日本のデータを入院と外来に分けたところ、外来が第 1 位(149 件)となり、入院は第 12 位(80.9 件)



であった。次に、MRIの対1000人当たりの施行件数を計算したところ、日本は111件で第2位であった(1位がトルコの119件、3位が米国の107件である)。日本のデータを入院と外来に分けたところ、外来における日本の順位は1位(91件)であり、入院では15位(20件)となった。また、PETの対1000人当たりの施行件数を計算したところ、日本は3.9件で第6位であった(1位が韓国の7.5件、2位がデンマークの6.5件、3位がイスラエルの5.1件、4位が米国の5.0件、5位がルクセンブルクの4.2件である)。

日本は、諸外国と比べてPET試行件数が中程度といえる。これは、諸外国では悪性新生物が疑われる場合には、CTやMRIよりも優先してPETが行われるというPET Firstという状況がある一方、日本の医療保険では適用されるケースが再発の場合や限定された部位であるという状況を反映していることが考えられる。日本のデータを入院と外来に分けたところ、外来における日本の順位は1位(3.7件)であり、入院では14位(0.26件)であった。

#### D. 考察

第一の基本データセットは、サンプリングデータセットとは異なり、「年度単位で紐付けされたレセプト情報」「入院エピソードごとの情報」といった特徴があり、その他の情報については相当程度削除したデータセットといえる。今後は、元データの抽出率、ユーザーが独自に指定する診療行為コードや医薬品コードの上限数に関して利用者の意見を集めながら、適宜変更を加えていくことが必要である。第二の

NDBオープンデータを利用による国際比較データを作成については、結果として、CT,MRI,PETの対1000人当たりの試行件数について、OECD加盟国間で比較可能なデータを作成できた。しかし、NDBデータを使いこなすには、分野横断的な専門性が求められる。特に、第一の基本データセットと第二のNDBオープンデータの利用においても共通していることが、診療行為コードや医薬品コードの設定である。病名、診療行為、医薬品コードの指定にあたり、傷病名マスタは訳32,000種類、診療行為マスタは約6,700種類、医薬品マスタは約20,000種類の中から設定しなければならない。

現在、先進国は限りある財政状況の中、効率性と公平性を考慮しながら、医療および介護保険システムを運営している。我が国は、これまで、質の高い医療サービスを比較的少ない医療費で提供しているとWHO(世界保健機関)等から評価されてきた。だが、増加傾向が続く医療費・介護費の対GDP比率はOECD(経済協力開発機構)の加盟国中で上位(第3位)となり、今後は一層、現在の医療・介護の質を確保しつつ、費用を適正化しなければならない課題に直面している。こうした状況は多くの先進国およびアジア諸国においても同様であり、このためOECD、WHO、国連等の国際機関は、医療及び介護分野における政策立案に資する国際統計報告として様々なHQ(ヘルスインディケータ:保健医療指標)の迅速な提供を各国に求めているが、日本が国際機関に提出している厚生労働統計分野の項目数は少ない。例えば、2015年にOECDが提出を求めている129

項目の HQ において、日本は 53 項目を提出しているが、加盟 35 ヶ国中 32 位と低い。ちなみに OECD 加盟国の平均提出件数は、89 項目である。そのため、日本は諸外国よりも少ないデータ提出状況を改善することが望ましい。NDB オープンデータからどのような新しいエビデンスを作成することができるのか検討したが、日本の全人口の保険医療記録から作成されたデータは、国内だけではなく諸外国に向けた情報としても価値のあるものと考えられる。特に、国際統計報告は、国連、世界保健機関、OECD(経済協力開発機構)の加盟国統一基準であることから、諸外国の行政関係者や研究者への波及効果も生み出すものであると考える。

## E. 結論

厚生労働省が 2009 年から収集を開始した NDB は、ヘルスケア分野における最大規模のデータベースであり、そこから作成されたデータは、国内だけではなく諸外

国に向けた情報としても価値のあるものと考えられる。今後も、NDB からのデータ提供はとしての特別抽出、サンプリングデータセット、基本データセット、NDB オープンデータについては、量と質の増加が必要である。

## F. 研究発表

1) 満武巨裕、大江和彦、今中雄一：NDB オープンデータを研究利用に活用する：医療技術(CT, MRI, PET)の利用に関する国際比較の試み、社会保険旬報, 第 2661 巻:12-16, 2016 年

2) 満武巨裕(「諸外国の医療ビッグデータ」、第 2 回データヘルス時代の質の高い医療の実現に向けた有識者検討会(平成 28 年 5 月 23 日)、

## G. 知的所有権の取得状況

該当なし

参考資料 1

表 1. 基本データセットのフォーマット(いつ例)

項番	データ項目名(日本語)	項目仕様	条件等記入欄	
1	医療機関都道府県コード		追加	
2	診療年月		追加	
3	ID1	○		
4	ID2	○		
	ID3(研究班作成)	○		
5	男女区分	○		
6	年齢階級コード(H22年次)	○		
7	入院年月日			追加:入院回数を作成する際のKeyとして
	入院回数	○		
8	保険法別番号		職域保険、地域保健、後期高齢者別	
9	保険者都道府県			
10	合計診療実日数	○	今回の入院の合計点数入院日数	
11	合計点数	○	今回の入院の合計点数	
12	医療機関コード:匿名化	○		
	入院回数シーケンス番号	○		
	入院-DPCフラグ	○		
	合計点数	○		
	診療実日数(入院)	○		
13	主病名1	○		b-902で統一
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	主病名(5)		SY主傷病フラグ付の傷病名コードの上位5	
16	診療開始日	○	SY主傷病フラグ付の傷病名コードの上位5の診療開始日コード	
17	転帰区分	○	SY主傷病フラグ付の傷病名コードの上位5の転帰区分	
18	下記のICDに関して、主病名(5)のICD10中分類コード		SY主傷病フラグ付の傷病名コードのICD10中分類	
18項目に まとめ	傷病名:ICD-10 A00-B99	○	感染症および寄生虫症	
	傷病名:ICD-10 C00-D48	○	新生物	
	傷病名:ICD-10 D50-D89	○	血液および造血器の疾患ならびに免疫機構の障害	
	傷病名:ICD-10 E50-E90	○	内分泌、栄養、代謝	
	傷病名:ICD-10 F00-F99	○	精神および行動	
	傷病名:ICD-10 G00-G99	○	神経系	
	傷病名:ICD-10 H00-H59	○	眼及び付属器	
	傷病名:ICD-10 H60-H95	○	耳及び乳様突起	
	傷病名:ICD-10 I00-I99	○	循環器系	
	傷病名:ICD-10 J00-J99	○	呼吸器系	
	傷病名:ICD-10 K00-K93	○	消化器系	
	傷病名:ICD-10 L00-L99	○	皮膚及び皮下組織	
	傷病名:ICD-10 M00-M99	○	筋骨格系、結合組織	
	傷病名:ICD-10 N00-N99	○	泌尿器系	
	傷病名:ICD-10 O00-O99	○	妊娠、分娩および産褥	
	傷病名:ICD-10 P00-P96	○	周産期	
	傷病名:ICD-10 Q00-Q99	○	先天奇形、変形、染色体異常	
	傷病名:ICD-10 R00-R99	○	症状、徴候等、他に分類されないもの	
傷病名:ICD-10 S00-T98	○	損傷、中毒、その他外因		
傷病名:ICD-10 V00-Y98	○	傷病及び死亡の外因		
傷病名:ICD-10 Z00-Z99	○	健康状態に影響を及ぼす要因等		
傷病名:ICD-10 U00-U99	○	特殊目的用コード		

19項目に全傷病名を統合。定義確認	傷病名【高血圧性疾患:非緊急】:ICD-10 I10	○	境界型高血圧症 8832479、高血圧症 8833421、高レープ性高血圧症 8842089、収縮期高血圧症 8842500、若年高血圧症 4010016、若年性矮型型高血圧症 8842500
	傷病名【高血圧性疾患:緊急】:ICD-10 I10	○	悪性高血圧症 8830212、高血圧性緊急症 8842178、高血圧切迫症 8842488
	傷病名【高血圧性心疾患】:ICD-10 I11	○	高血圧性うっ血性心不全 8833422、高血圧性心疾患 4029010、高血圧性心不全8833426
	傷病名【高血圧性腎疾患】:ICD-10 I12	○	高血圧性腎硬化症 4039033、高血圧性腎疾患 4039005、高血圧性腎症 4039006、高血圧性腎不全 8833427
	傷病名【高血圧性心腎疾患】:ICD-10 I13	○	高血圧性心腎症 8833425
	傷病名【腎血管、内分泌性高血圧性】:ICD-10 I15	○	腎血管性高血圧症 8835586、腎実質性高血圧症 8835605、腎性高血圧症 8835614、内分泌性高血圧症 8838336、二次性高血圧症 8838308、副腎性高血圧症 8830699
	傷病名【糖尿病】:ICD-10 E10-E14	○	...
	傷病名【高脂血症】:ICD-10 E78	○	E78に該当する傷病名コードのうち、高HDL血症 2724037、高比重リポ蛋白欠乏症 8833663、低脂血症 8842204、低ベータリポ蛋白血症 8837884、無ベータリポ蛋白血症 8840506、右棘
	傷病名【脳血管障害】:ICD-10 I60-I69	○	...
	傷病名【腎障害:急性腎不全】:ICD-10 N17	○	...
	傷病名【腎障害:慢性腎不全】:ICD-10 N18	○	...
	傷病名【腎障害:腎不全その他】:ICD-10 N19	○	...
	傷病名【カリウム異常】	○	カリウム代謝異常 2768001、高カリウム血症 2767001

1	診療行為【冠動脈造影:右心】	○	160004610 170002310
2	診療行為【冠動脈造影:左心】	○	160004510 170020710
3	診療行為【冠動脈形成】	○	K546 経皮的冠動脈形成術 150153910
4	診療行為【冠動脈形成】	○	K547 経皮的冠動脈粥腫切除術 150260350
5	診療行為【冠動脈形成】	○	K548 経皮的冠動脈形成術(特殊カテーテル:レーザー)15035931
6	診療行為【冠動脈形成】	○	K549 経皮的冠動脈ステント留置術 150263310
7	診療行為【冠動脈形成】	○	K550 冠動脈内血栓溶解療法 160107550
8	診療行為【冠動脈形成】	○	K550 経皮的冠動脈血栓吸引術 150318310
9	診療行為【冠動脈形成】	○	K551 冠動脈形成術(血栓内膜摘除)(1箇所)150145710
10	診療行為【冠動脈形成】	○	K551 冠動脈形成術(血栓内膜摘除)(2箇所以上)150145810
11	診療行為【冠動脈形成】	○	K552 冠動脈、大動脈バイパス移植術(1吻合)150145910
12	診療行為【冠動脈形成】	○	K552 冠動脈、大動脈バイパス移植術(2吻合以上)150146010
13	診療行為【冠動脈形成】	○	K552 冠動脈形成術(血栓内膜摘除)併施加算 150302770
14	診療行為【冠動脈形成】	○	K552 冠動脈、大動脈バイパス移植術(人工心肺不使用)(1吻合)
15	診療行為【冠動脈形成】	○	K553 左室自由壁破裂修復術(冠動脈血行再建術(2吻合以上)を
16	診療行為【在宅人工呼吸治療】	○	C107に対応する診療行為のうち、114005410
17	診療行為【在宅陽圧呼吸治療】	○	C107に対応する診療行為のうち、114009710
18	診療行為【在宅酸素投与】	○	C103 在宅酸素療法指導管理料(チアノーゼ型先天性心疾患)114
19	診療行為【在宅酸素投与】	○	C103 在宅酸素療法指導管理料(その他)114003710
...	...		
	医薬品【糖尿病治療薬】	○	
	医薬品【降圧剤:Caブロッカー】	○	
	医薬品【降圧剤: $\beta$ ブロッカー】	○	
	医薬品【降圧剤:ACE阻害剤】	○	
	医薬品【降圧剤:ARB】	○	
	医薬品【降圧剤:利尿剤】	○	
	医薬品【高脂血症剤】	○	
	医薬品【インスリン】	○	

表 2. 診療コードの例

診療行為コード	診療行為名称
113005910	生活習慣病管理料(処方せんを交付)(糖尿病を主病)
113006110	生活習慣病管理料(処方せんを交付しない)(糖尿病を主病)
113010010	糖尿病合併症管理料
113010370	自己測定血糖値指導加算
113013610	糖尿病透析予防指導管理料
113015610	糖尿病透析予防指導管理料(特定地域)
114005910	血糖自己測定器加算(20回以上)(1型糖尿病の患者を除く)
114006010	血糖自己測定器加算(40回以上)(1型糖尿病の患者を除く)
114007410	血糖自己測定器加算(60回以上)(1型糖尿病の患者を除く)
114009910	血糖自己測定器加算(20回以上)(1型糖尿病・小児低血糖症等)
114010010	血糖自己測定器加算(40回以上)(1型糖尿病・小児低血糖症等)
114010110	血糖自己測定器加算(60回以上)(1型糖尿病・小児低血糖症等)
114010210	血糖自己測定器加算(80回以上)(1型糖尿病・小児低血糖症等)
114010370	経鼻的持続陽圧呼吸療法用治療器加算
114010970	注入器用注射針加算(1型糖尿病、血友病患者又はこれに準ずる患者)
114011070	注入器用注射針加算(その他)
114015510	血糖自己測定器加算(100回以上)(1型糖尿病・小児低血糖症等)
114015610	血糖自己測定器加算(120回以上)(1型糖尿病・小児低血糖症等)

診療行為コード	診療行為名称
113005910	生活習慣病管理料(処方せんを交付)(糖尿病を主病)
113006110	生活習慣病管理料(処方せんを交付しない)(糖尿病を主病)
113010010	糖尿病合併症管理料
113010370	自己測定血糖値指導加算
113013610	糖尿病透析予防指導管理料
113015610	糖尿病透析予防指導管理料(特定地域)
114005910	血糖自己測定器加算(20回以上)(1型糖尿病の患者を除く)
160133330	CPR(赤)
160142350	1.5AG
160151050	グリコアルブミン
160162050	抗GAD抗体
160176950	抗IA-2抗体

表 3. 医薬品コードの例

医薬品コード	医薬品名称	薬価基準コード
610406013	アドメッセン錠1mg	4490004F2010
610406047	ウテロン錠5mg	2590004F1010
610406053	エマベリンLカプセル5mg	2171014N4028
610406055	エマンダキシン錠50mg	1124026F1030
610406073	オタノールカプセル0.25 0.25 $\mu$ g	3112004M1015
610406074	オタノールカプセル0.5 0.5 $\mu$ g	3112004M2011
610406079	ガスター散2%	2325003B2029
610406083	カルシオロールカプセル0.25 0.25 $\mu$ g	3112004M1015
610406084	カルシオロールカプセル0.5 0.5 $\mu$ g	3112004M2011
610406085	カルシタロールカプセル0.25 0.25 $\mu$ g	3112004M1015
610406086	カルシタロールカプセル0.5 0.5 $\mu$ g	3112004M2011
610406089	カルデミン錠0.25 $\mu$ g	3112004F1024
610406090	カルトールカプセル0.25 0.25 $\mu$ g	3112004M1015
610406091	カルトールカプセル0.5 0.5 $\mu$ g	3112004M2011
610406094	カルミサルカプセル0.25 0.25 $\mu$ g	3112004M1015
610406095	カルミサルカプセル0.5 0.5 $\mu$ g	3112004M2011
610406130	サリパラ・コデイン液	2249104S1030
610406157	セオノマル錠5 5mg	2123014F1019
610406158	セオノマル錠10 10mg	2123014F2015
610406161	セクロダン細粒200 200mg	6132005C2025
610406167	セレクナート錠20mg	2149010F1114
610406168	セレクナート錠40mg	2149010F2013
610406170	ソロン錠50 50mg	2329011F1028
610406182	チルミールカプセル50mg	2129003M1013
610406183	チルミールカプセル100mg	2129003M2010
610406193	テンマ末	5100133A1011
610406203	トフィルシン錠50 50mg	1124026F1111
610406207	トルシトリンカプセル0.25 0.25 $\mu$ g	3112004M1015
610406208	トルシトリンカプセル0.5 0.5 $\mu$ g	3112004M2011
610406209	トルバナシン錠50 50mg	1124026F1120
610406227	花扇加エブシ末K	5900001A1025
610406231	バリトゲンHD 98.6%	7212029X1023
610406232	バリトップHD 99%	7212022A1037
610406252	ヒポテリオールカプセル0.25 0.25 $\mu$ g	3112004M1015
610406253	ヒポテリオールカプセル0.5 0.5 $\mu$ g	3112004M2011
610406259	フェニレン錠50mg 鉄50mg	3222013F1050
610406262	フェロチーム錠50 鉄50mg	3222013F1084
610406281	ペルチスタン錠100 100mg	2171010F3010
610406284	ヘルラートL錠10 10mg	2171014G1151
610406290	ポルボノールドライシロップ0.5%	2252006R2022
610406293	マナミンTM散	2339179B1037
610406306	メキシレート錠50 50mg	2129003F1022
610406307	メキシレート錠100 100mg	2129003F2029
610406319	メルDESTカプセル50mg	2129003M1013
610406320	メルDESTカプセル100mg	2129003M2010
610406338	リカマイシンドライシロップ200 200mg	6149001R2027
610406345	リチオマル錠100mg	1179017F1030

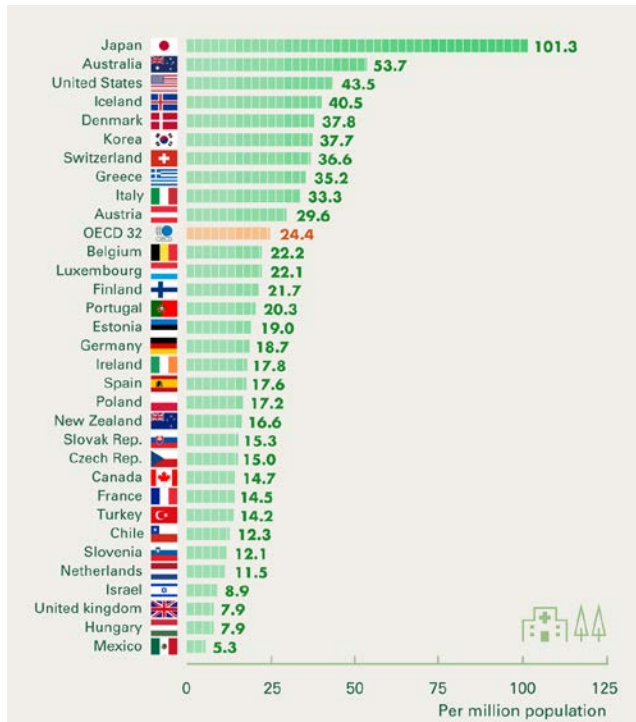


図1. OECD加盟国のCTスキャナー数(左図)および施行件数(右図)。



図2. OECD加盟国のMRIスキャナー数(左図)および施行件数(右図)。

## 研究成果の刊行に関する一覧表

### 書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書 籍 名	出版社名	出版地	出版年	ページ
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
満武巨裕 大江和彦 今中雄一	NDBオープンデータを研究利用に活用する：医療技術(CT, MRI, PET)の利用に関する国際比較の試み	社会保険旬報	第2661巻	12-16	2016
大江和彦	医療情報データベースの基盤整備	情報管理	Vol. 59, No. 5	277-283	2016
松居 宏樹 佐藤 大介 大江 和彦	レセプト情報等オンサイトリサーチセンターにおける、今後の第三者提供の方向性について レセプト情報等オンサイトリサーチセンターにおけるNDBデータの利用 システム環境とNDBの特性に関する報告	医療情報学連合大会論文集	36回1号	138-140	2016
佐藤 大介 大江 和彦	医療データベース利活用の国内基盤の最新状況 レセプト情報等オンサイトリサーチセンターの試行的利用について	日本薬剤疫学会学術総会抄録集	巻22nd	51	2016
大江 和彦	医療ビッグデータとこれからの医療	日本臨床検査自動化学会誌	41巻4号	371	2016