

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

研究代表者 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**

本研究の目的は、超高速レセプトビッグデータ解析基盤を利用することにより、レセプト情報・特定健診等情報データベース（以下、NDB）のデータ（年間370億レコード）及び保険者から収集した75万人規模の医療・介護データを利用して、地域（都道府県・市町村）や保険者ごとの性別・年齢別・疾患別医療費および罹患率を把握し、都道府県・全保険者の医療・介護支出目標を精密に設定する推計方法を開発し、増大する国民医療費の適正化対策としての施策導入を目指す。また、複数年データを活用して時間経過に伴い新たに作成したデータ項目と医療・介護供給体制の情報とリンクさせて、需要および供給体制を把握し、現実のデータの変遷に合わせた予測が可能な先進的なヘルスケアビッグデータサービスを実現する。

具体的には、以下の5つとした。

- 1) 4 医学会と連携した NDB レセプトデータから生活習慣病の発症数・合併症数・重症群の定義を確立を通し、わが国の生活習慣病・癌等の複数疾患の診療の実態を明らかにし、診療エビデンス・パフォーマンスの分析・公開
- 2) 国・都道府県・市町村レベルで、医療費・介護費支出の目標を設定する方法論の確立
- 3) 日本版リスク構造調整方式の確立と財源調整施策の導入の試算
- 4) NDB レセプト情報と特定健診データを活用した新しい1次・2次・3次予防群の確立
- 5) レセプト情報と介護レセプトを活用し訪問診療をはじめとする在宅医療の実態を明らかにするとともに、疾病予防・医療・介護の連携の実態把握

採択額及び研究期間を考慮し、研究内容を重点化した。

ビッグデータ解析基盤の増強については、物理的には、東京大学構内にビッグデータ解析基盤およびソフトウェア開発用のセキュリテイルームを設置した。ハードウェア構成は、サーバ装置群、専用パーティションに設置したクライアント端末構成とした。サーバ装置群は、厚生労働省のガイドラインに基づくレセプト情報等の堅牢な管理と、レセプト情報等に対する解析クエリの実行を担う。クライアント端末は、研究者へのインターフェースであり、レセプト情報等の管理ならびに解析の命令をサーバ装置群に発行し、また、その結果を取得して出力する。

NDBデータの集計・分析としては、疾患毎罹病率、医療費の傾向などの情報を都道府県・二次医療圏別に作成する。次に、患者数が多く大量の医療資源を消費している生活習慣病について、「日本腎臓学会」等と連携し、これらの疾患管理を評価するための指標を同定する。さらに、保険者から収集したデータを利用して、在宅医療（医療保険訪問看護を含む）と介護サービスがどのように併用され、サービスに要する費用等の実態を分析する。特に、在宅医療介護連携の実態を明らかにし、同じ医療や介護が必要とされる高齢者の入院や外来受診、在宅医療の地域比較を行い、現在進行中の医療介護総合改革の適切な推進等に資するエビデンスに基づく政策の推進を図る。

三重県の全ての国民健康保険と後期高齢者のデータ利用し、供給データ（医療機関・介護施設・介護サービス事業所の基本的属性）と組み合わせることで、医療と介護サービスの連携でどのような使われ方をしているか、個人単位で可視化することができた。また、ステークホルダー（主に保険者と地域医師会）用に4種類のデモを作成した。デモは、市町ごとに病名が選択でき、入外来別の患者数と医療費、どこの医療機関を受診したかを可視化した。その結果、地域により疾患別受診状況に差異があることが明らかになった。特に、二次医療圏を超える受診実態が示された。介護サービスの利用についても、地域ごとの特性が示された。本研究により、種々のデータを組み合わせて医療や介護の実態を可視化することで、現在進行中の医療介護総合改革の推進等に資するエビデンスを作成できることを示した。

分担研究者

辻 哲夫・医療経済研究機構・副所長  
喜連川 優・東京大学生産技術研究所・教授  
合田 和生・東京大学生産技術研究所・特任准教授  
伏見 清秀・東京医科歯科大学医療政策学・教授  
辻 一郎・東北大学大学院医学系研究科・教授  
藤森 研司・東北大学大学院・医学系研究科・教授  
興相 貴英・自治医科大学・企画経営部医療情報部・准教授  
中島 直樹・九州大学病院メディカルインフォメーションセンター・教授  
石川ベンジャミン光一・国立がん研究センター

がん対策情報センターがん統計研究部・室長  
飯島 勝矢・東京大学高齢社会総合研究機構・准教授  
吉江 悟・東京大学医学部在宅医療学拠点・特任研究員  
関本 美穂・大阪府済生会吹田病院麻酔科・医長  
印南 一路・慶應義塾大学総合政策学部・教授  
高橋 邦彦・名古屋大学大学院医学系研究科・准教授  
奥村 泰之・医療経済研究機構・主任研究員  
清水沙友里・医療経済研究機構・主任研究員

## A. 研究目的

本研究の目的は、超高速レセプトビッグデータ解析基盤を利用することにより、レセプト情報・特定健診等情報データベース（以下、NDB）のデータ（年間370億レコード）及び保険者から収集した75万人規模の医療・介護データを利用して、地域（都道府県・市町村）や保険者ごとの性別・年齢別・疾患別医療費および罹患率を把握し、都道府県・全保険者の医療・介護支出目標を精密に設定する推計方法を開発し、増大する国民医療費の適正化対策としての施策導入を目指す。また、複数年データを活用して時間経過に伴い新たに作成したデータ項目と医療・介護供給体制の情報とリンクさせて、需要および供給体制を把握し、現実のデータの変遷に合わせた予測が可能な先進的なヘルスケアビッグデータサービスを実現する。

## B. 研究方法

本研究では、1.NDBの全データ、2.三重県下の全国民健康保険および後期高齢者の医療・介護・健康データ（75万人規模）を利用する。

NDBからは2008年度から2014年度迄の全データを取得する。（2010年度の全NDBデータ（電子レセプト：約15億件、特定健診等データ約2000万件）は取得済み）。

三重県の全市町村（国民健康保険）と後期高齢者医療広域連合2008-203年度のデータも収集する。

保険者からは、NDBと同等のデータに加えて、現在のNDBには蓄積されていない、被保険者マスタ、特定健康診査・特定保健指導の対象者ファイル、介護受給者ファイ

ル、介護給付（費）ファイルを収集する。被保険者マスタには、全加入者の基本的属性（性別・年齢）に加え、保険加入日、保険資格喪失日などが含まれ、分析対象群を設定するための必要なファイルである。

特定健康診査・特定保健指導の対象者ファイル（国保システム・特定健診結果等情報作成抽出（受診券情報）ファイル）は、健診の未受診者の基本的属性が含まれる。加えて、特定健診受診者CSVファイル、特定健診結果等情報作成抽出（健診結果情報（横展開））ファイル、特定健診結果等情報作成抽出（保健指導情報）ファイルを手に入る。

介護保険については、介護受給者ファイル（医療費の被保険者マスタに相当）に認定審査結果情報等がふくまれる。介護給付費ファイルは、レセプトに相当する。

使用するデータで示したように、本研究で対象とするデータは膨大であり、従来のデータの扱い方では十分な分析が出来ない。そのためビッグデータ解析基盤を構築する必要がある。既に内閣府Firstおよび厚生労働科学研究費補助金において基礎的な基盤を構築しているが、本研究ではデータがさらに増加するため、ビッグデータ解析基盤の増強を行う。

次に、分析の質を高めるためには原データの質を高める必要がある。一人の患者に関わるレセプトが国内約18万の医療施設（病院約8500、診療所約10万、歯科診療所約7万）および調剤薬局（約5.5万）から発生するため、それらの間の紐づけがうまくいかないことがあることが知られている。本研究においてはNDBと直接保険者から入手したデータを比較すること

等を通じて問題点を洗い出し、解決策を提示する。さらにビッグデータ解析基盤を用いてこれまでより大幅に迅速な基礎統計値を産出する体制を整える。

上記の基盤構築、データ検証を行った後に、レセプトデータを分析することで以下に取り組む。

- ・ビッグデータ解析基盤の増強
- ・全 NDB データの検証と質向上
- ・基礎統計
- ・多次元分析
- ・都道府県・二次医療圏・市町村別分析

はじめに、ビッグデータ解析基盤の増強を行う。物理的には、東京大学構内にビッグデータ解析基盤およびソフトウェア開発用のセキュリティルームを設置する。ハードウェア構成は、サーバ装置群、専用パーティションに設置したクライアント端末構成とする。サーバ装置群は、厚生労働省のガイドラインに基づくレセプト情報等の堅牢な管理と、レセプト情報等に対する解析クエリの実行を担う。クライアント端末は、研究者へのインターフェースであり、レセプト情報等の管理ならびに解析の命令をサーバ装置群に発行し、また、その結果を取得して出力する。

NDBデータの集計・分析としては、はじめに、高速保険レセプト解析基盤の実行性能比較を行った。4種類のクエリ（クエリ1：都道府県ごとの急性心膜炎患者数（対人口比率）、クエリ2：口唇の悪性新生物の都道府県ごとの平均診療費、クエリ3：呼吸器結核に関する年齢ごとのレセプト件数、クエリ4：コレラ患者の都道府県ごとの平均診療費）を実行して、検索時間を比較した。次に、薬剤の処方傾向の分析と

して、処方件数および併用薬や併用禁忌の実態を分析した。

患者数が多く医療資源を消費している生活習慣病について、「日本腎臓学会」等と連携し、患者数、疾病がどのような経過をたどるのか（時系列分析）、疾病に対してどのような診療が行われているか（診療パターン）といった分析を行い、新しい診療ガイドラインを作成するためのエビデンス、疾病を管理するための評価指標の作成につなげる。

保険者から収集したデータを利用して、在宅医療（医療保険訪問看護を含む）と介護サービスがどのように併用され、サービスに要する費用等の実態を分析する。特に、在宅医療介護連携の実態を明らかにし、同じ医療や介護が必要とされる高齢者の入院や外来受診、在宅医療の地域比較を行い、現在進行中の医療介護総合改革の適切な推進等に資するエビデンスに基づく政策の推進を図る。

## C. 研究結果

採択額及び研究期間を考慮し、研究内容を重点化した。

三重県の全国健康保険（29市町）と後期高齢者医療広域連合（75歳以上）の最新の医療データ（電子レセプト）・特定健診データ、介護データはの収集が完了した。ビッグデータ解析基盤への新投資はしないで、セキュリティ環境・ソフトウェア開発、保険者データ（三重）の格納を行った。基礎統計については、可視化ツールを作成し、ビッグデータシステムに実装した。

高速保険レセプト解析基盤の実行性能比較では、従来型と比べて約 100 倍となる高速化を実現できた。

薬剤の処方傾向の分析では、主要抗うつ薬と湿布薬(ケトプロファン、ロキソプロフェンナトリウム水和物処方件数)の年間点数合計と処方日数(365日を超える実態把握)、高血圧治療薬の処方件数と併用薬・併用禁忌の実態を示した。

生活習慣病の分析結果は、例えば慢性腎臓病(CKD)の患者数について約 22.5 万人を抽出した。透析患者については、診療行為名称を特定したところ、NDB から推計した全国の透析患者数は、約 22 万人(生活保護は含まず)であった。また、腎移植後に継続的に透析を行っているの患者が、298 人存在することを確認した。

三重県の全ての国民健康保険と後期高齢者のデータ利用し、供給データ(医療機関・介護施設・介護サービス事業所の基本的属性)と組み合わせることで、医療と介護サービスの連携でどのような使われ方をしているか、個人単位で可視化することができた。具体的には、複数年データを用いて、生活習慣病の高血圧患者の推移を観察し、診断済群(投薬なし、合併症なし)、服薬治療群(投薬あり、合併症なし)、合併症群(投薬なし、合併症あり、透析あり)、透析治療群(投薬あり、合併症あり、透析あり)毎に継時的変化を示した。

保険者から収集したデータを利用したサービスに要する費用等の実態を分析については、ステークスホルダー(主に保険者と地域医師会)用に 4 種類のデモを作成した。デモは、市町ごとに病名が選

択でき、入外来別の患者数と医療費、どこの医療機関を受診したかを可視化した。その結果、地域により疾患別受診状況に差異があることが明らかになった。特に、二次医療圏を超える受診実態が示された。介護サービスの利用についても、地域ごとの特性が示された。

#### D. 考察

我々は、最終的な研究成果の政策への反映方法として以下の 4 つを設定した。

1) 次世代 NDB データ構築と分析による医療費適正化

独自に収集した保険者データと NDB データの突合検証を行うことで、検証、本知見に基づく次世代システム構築への提言を行う。計画通り、検証を実行している。

2) 診療エビデンスの明確化と治療方法の標準化

医学会連携することで、初年度は 2 疾患の診療実態と地域差を明らかにし、患者に対する適切な医療の提供、評価指標の作成を行っている。

3) 在宅医療(訪問看護等を含)と介護サービスの連携強化

在宅の医療・介護の給付実態について、地域ごとの特性分析等を実施。将来、NDB データに介護保険データを含めた次世代システムプロトタイプ構築による、効率的な医療費・介護費の総合的な適正化につなげる。

4) 医療費支出目標の推計、リスク構造調整の検討として、今後の医療保険制度の在り方に関する検討に寄与する。

今回、研究の採択額及び研究期間を考慮

し、研究内容を重点化した。当初の予定とは異なり、データ収集等について想定以上に時間を要した。だが我々は、保険者から独自に収集したデータ（三重県 75 万人規模）を最初に分析、知見を得た上で、次に全日本の NDB データを分析している。これにより抽出時間、処理時間などの見積もりができ、分析の効率化が実現できていると思われる。

高速保険レセプト解析基盤の実行性能比較から、従来型と比べて約 100 倍となる高速化を実現でき、処方傾向の分析等において応用例を示すことができた。レセプト病名については、例えばステージ情報付の傷病名等を有する入力をしている医療機関が少ない、移植のレセプトは紙で請求されることが多いといった理由から、NDB を利用しても全数を把握できない可能性がある。しかし、病名と処置、処方を組み合わせて対象疾患を選定することで治療中の患者の把握率が向上する可能性がある。

今回、ビッグデータ基盤と医療政策・経済等研究を連携させ、研究者、地域医師会、地方自治体、保険者らステークホルダーが実態把握・政策提言を行い社会的にもメリットが得られるようにデータや BI ツールを開発したところ、種々のデータを組み合わせることで、現在進行中の医療介護総合改革の適切な推進等に資するエビデンスを作成できることを示すことができた。今後も、現在の関係者以外にも一定の認識が共有され、諸課題の解決に向けた社会的な合意を形成するための基盤づくりが必要であ

り、現在進行中の医療介護総合改革の推進等に資するエビデンスを作成できることを示した。

## E. 結論

種々のデータを組み合わせて医療や介護の実態を可視化することで、現在進行中の医療介護総合改革等の推進等に資するエビデンスを作成できることを示した。

## F. 健康危険情報

なし

## G. 研究発表

- 1) 「基本データセットの提供について」第 29 回レセプト情報等の提供に関する有識者会議（平成 28 年 3 月 16 日）、<http://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000117367.pdf>
- 2) 満武巨裕：レセプトビッグデータ解析の現状と将来 実験医学 第 34 巻第 5 号：799-804, 2016 年
- 3) 「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）、<http://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000125318.pdf>

## H. 知的財産権の出願・登録状況

該当なし