

201501028A

厚生労働科学研究費補助金  
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クデータ解析基盤の整備  
（課題番号 H27-政策-戦略-014）

平成27年度 総括・分担研究報告書

研究代表者 満武 巨裕

平成28年（2016）年3月

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

厚生労働科学研究費補助金  
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クデータ解析基盤の整備  
（課題番号 H27-政策-戦略-014）

平成27年度 総括・分担研究報告書

研究代表者 満武 巨裕

平成28年（2016）年3月

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

# 目 次

## I. 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クデータ解析基盤の整備

満武巨裕

1

## II. 分担研究報告書

1. 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クデータ解析基盤の整備

満武巨裕、辻 哲夫、喜連川 優、合田 和生、伏見 清秀、辻 一郎、藤森 研司、興梠 貴英、中島 直樹、石川ベンジャミン光一、飯島 勝矢、吉江 悟、関本 美穂、印南 一路、高橋 邦彦、奥村 泰之、清水 沙友里

6

資料 1：研究実施計画書.....13

資料 2：平成 27 年度研究スライド.....39

## III. 研究成果の刊行に関する一覧表

54

## IV. 研究成果の刊行物・別刷

55

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、申請者がこれまで構築してきた高速レセプト・ビッグデータ解析基盤を更に発展させることにより、以下の医療の需要・供給、質、コストが国・地域・医療機関レベルで即座に可視化できるサービスを確立することである。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

結果として、採択額及び研究期間を考慮し、研究内容を重点化した。1)は、4学会から2学会（日本糖尿病学会、日本腎臓学会）に規模を縮小した。2)3)4)は、検討会の開催のみとして概念整理と分析手法の検討にとどめ、データの解析は次年度以降実施することとした。5)は、三重県の国保と後期高齢者をフィールドに医療および介護レセプト情報を連結し、連携の実態把握を行った。

今後も引き続き、最先端のビッグデータ基盤と最先端の医療政策・経済等研究を連携させ、研究者、厚生労働省、地域医師会、地方自治体、保険者らステークスホルダーがデータやBIツールを使い論文化・実態把握・政策提言を行い技術的にも社会的にもメリットが得られるようにする。また、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についても検討を続ける。

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  
興招 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、申請者がこれまで構築してきた超高速レセプトビッグデータ解析基盤を更に発展させることにより、レセプト情報・特定健診等情報データベース（以下、NDB）の全データ（年間370億レコード）及び保険者から申請者らが収集した75万人規模の医療・介護データを利用して、全都道府県・市町村の精緻な医療・

介護費支出の目標設定、財源調整施策としてのリスク構造調整方式の導入、全国民に対する新基準の1次・2次・3次予防群の設定が可能であることを明らかにすることである。こうしたプロセスを経て、現実のデータの変遷に合わせた予測が可能なヘルスケアビッグデータサービスを構築する。

本研究では第一に、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、国民（被保険者）の健康状態・罹患状況に関して、電子レセプトデータの傷病名・診療行為・医薬品情報、介護認定・給付情報、特定健診等情報を利用してエピソード単位で把握する方法論を開発する。

第二に、集計したパラメーターを活用し、保険者の加入者（被保険者）の各種リスク状態を数値化し、予測モデルを構築する。特に、保険者の財政調整に関してエビデンスに基づくリスク構造調整方式が導入可能であることを明らかにする。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のパラメーターを変化させたシミュレーションを行なう。こうしたプロセスを経て、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。

第三に、特定健診データやレセプト情報を活用した後期高齢者の支援金の加算・減

算の評価は、当初の計画を実行できなかった。したがって、特定健診・特定保健指導の評価指標や参酌標準は、保険者の保健活動を評価し、且つ医療費適正化のインセンティブを有するものに再検討する必要がある。そこで、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、保険者の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を行う。

## 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）ファイルを集める。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付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
- ・調整医療費の推計と地域毎医療支出目標額の設定
- ・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
- ・1次・2次・3次予防および予防・医療・介護の連携の実態把握
- ・研究者用データ提供に向けた検討と実装

### 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。保険者データについても作成し、以後更新を重ねる予定である。

今後、多次元分析についても同様に進める予定である。連携学会は、日本腎臓学会、日本糖尿病学会の二つに限定した。すでに定義体の第一バージョンは作成し、保険者のデータを利用して集計を開始した。

都道府県・二次医療圏・市町村別分析、

調整医療費の推計と地域毎医療支出目標額の設定、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については、検討会の開催のみとして概念整理と分析手法の検討にとどめ、データの解析は次年度以降実施することとした。

1次・2次・3次予防および予防・医療・介護の連携の実態把握は、三重県の国保と後期高齢者をフィールドに医療および介護レセプト情報を連結し、連携の実態把握を行った。

研究者用データ提供に向けた検討と実装は、1件の提供を行った。現在、2件目に着手している。

### 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。

大規模データ研究の標準化や教育の充実、人材育成については、医師、医療政策研究者、医療情報研究者達からなる我が班の利用経験に加えて、行政担当者・看護師（保健師）・介護職員等にも資する標準化されたデータの利用に関する教育・人材育成を視野に入れる。H28 年度は、協力保険者の県庁職員、医師会、各市町村国保の保健師に対する教育等を計画している。

今後も引き続き、最先端のビッグデータ基盤と最先端の医療政策・経済等研究を連携させ、研究者、厚生労働省、地域医師会、地方自治体、保険者らステークスホルダーがデータや BI ツールを使い論文化・実態把握・政策提言を行い技術的にも社会的にもメリットが得られるようにする。また、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についても検討を続ける。

## E. 結論

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時間を要したが、今後も引き続き、最先端のビッグデータ基盤と最先端の医療政策・経済等研究を連携させ、研究者、厚生労働省、地域医師会、地方自治体、保険者らステークスホルダーがデータや BI ツールを使い論文化・実態把握・政策提言を行い技術的にも社会的にもメリットが得られるようにする。

## 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



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、申請者がこれまで構築してきた超高速レセプトビッグデータ解析基盤を更に発展させることにより、レセプト情報・特定健診等情報データベース（以下、NDB）の全データ（年間370億レコード）及び保険者から申請者らが収集した75万人規模の医療・介護データを利用して、全都道府県・市町村の精緻な医療・介護費支出の目標設定、財源調整施策としてのリスク構造調整方式の導入、全国民に対する新基準の1次・2次・3次予防群の設定が可能であることを明らかにすることである。こうしたプロセスを経て、現実のデータの変遷に合わせた予測が可能なヘルスケアビッグデータサービスを構築する。

本研究では第一に、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、国民（被保険者）の健康状態・罹患状況に関して、電子レセプトデータの傷病名・診療行為・医薬品情報、介護認定・給付情報、特定健診等情報を利用してエピソード単位で把握する方法論を開発する。

第二に、集計したパラメーターを活用し、保険者の加入者（被保険者）の各種リスク状態を数値化し、予測モデルを構築する。特に、保険者の財政調整に関してエビデンスに基づくリスク構造調整方式が導入可能であることを明らかにする。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のパラメーターを変化させたシミュレーションを行なう。こうしたプロセスを経て、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。

第三に、特定健診・特定保健指導および電子レセプトに基づく、新たな特定健診・特定保健指導の評価指標の作成、特定保健指導の対象群や医療機関との連携を促進する群等の1次・2次・3次予防群の設定が可能であることを明らかにする。

本年度は、採択額及び研究期間を考慮し、研究内容を重点化した。データソースは、NDB データ（2009-2014年度）を入手するため、レセプト情報等の提供に関する有識者会議・審議分科会(7/31)において審査、2/3 に承諾を得た。また、三重県の全国保(29市町)

と後期高齢者から約 75 万人規模の 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）ファイルを収集した。ビッグデータ解析基盤への新投資は不可能であったが、本邦発のテイル集団への詳細分析を実施し、データ解析基盤の整備として BI ツールを作成した。分析対象疾患についても、糖尿病、慢性腎臓病について限定して分析を行った。医療と介護連携については、三重県の全国保と後期高齢者を対象に、医療および介護レセプト情報を連結して分析を行った。医療費支出目標の推計とリスク構造調整については、概念整理と分析手法の検討にとどめて、データの解析は次年度以降とした。

今後も引き続き、最先端のビッグデータ基盤と最先端の医療政策・経済等研究を連携させ、研究者、厚生労働省、地域医師会、地方自治体、保険者らステークスホルダーがデータや BI ツールを使い論文化・実態把握・政策提言を行い技術的にも社会的にもメリットが得られるようにする。また、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についても検討を続ける。

#### 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  
興柁 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

医療分野でのビッグデータの一つに、2009 年度から日本の全患者の保険診療情報を蓄積している国家規模のデータベースがある。これはレセプト情報・特定健診等データベース（以下 NDB）と呼ばれ、このデータを分析し医療政策に利用することが医療の質の改善や医療費適正化の切り札となることが期待されている。

医療費の適正化を目的として、これまで幾多の医療制度改革が行われてきた。例えば2008年度には、特定健康診査・特定保健指導（メタボ健診）や後期高齢者医療制度が導入された。メタボ健診は生活習慣病の一次予防により将来的な医療費を削減する目的で導入されたが、これが本当に生活習慣病を予防し、医療費を削減する効果を持つのかに関するエビデンスは未だ示されていない。

エビデンスを示すことが困難であった原因の一つは、これまで診療の実態や診療行為のアウトカム、医療のコストなどを包括的に分析するための国家規模のデータが存在しなかったことである。レセプト情報は保険者に関わりなく同じフォーマットで作成されるため、この目的にかなうデータとして有望視されてきたが、全国規模のデータを集めるシステムが長らく存在しなかったため、その分析は一部の研究者や研究機関による狭いテーマに限定されてきた。NDBには、全医療機関で提供された医療をカバーするレセプト情報、特定健診情報が収載され、2013年度の社会保障制度改革国民会議の報告書（内閣官房）においても「国が保有するレセプト等データの利活用の促進にも不可欠である」と大きな期待が寄せられている。

また、今後の超高齢社会に向けて目指すべき地域包括ケアシステム構築の中で鍵となる在宅医療についても、現状ではその診療実態はほとんど明らかにされておらず、大きな課題となっている。在宅医療の質を定義する難しさの1つには、その対象となる患者において認知症や終末期など自律的判断を行えない状態にある者が多

く、また治癒を目標に据えられない慢性期医療独特の特徴があることから、診療の評価軸を一律に設定しにくいという点などがあげられるが、現状はそれ以前に、どのような疾患・状態にある患者にどのような在宅医療が提供されているかといった実態さえ明らかにされていない。ビッグデータを活用してまずは実態を確認し、特定の保険者の健診・医療・介護データから在宅医療のエビデンス・パフォーマンスを集積していく端緒を開くことが必要かつ最も効率的である。

本研究では第一に、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、国民（被保険者）の健康状態・罹患状況に関して、電子レセプトデータの傷病名・診療行為・医薬品情報、介護認定・給付情報、特定健診等情報を利用してエピソード単位で把握する方法論を開発する。

第二に、集計したパラメーターを活用し、保険者の加入者（被保険者）の各種リスク状態を数値化し、予測モデルを構築する。特に、保険者の財政調整に関してエビデンスに基づくリスク構造調整方式が導入可能であることを明らかにする。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のパラメーターを変化させたシミュレーションを行なう。こうしたプロセスを経て、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。

第三に、特定健診データやレセプト情報を活用した後期高齢者の支援金の加算・減算の評価は、当初の計画を実行できなかった。したがって、特定健診・特定保健指導の評価指標や参酌標準は、保険者の保健活動を評価し、且つ医療費適正化のインセンティブを有するものに再検討する必要がある。そこで、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、保険者の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を行う。

## 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 2008-2013 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) ファイルを入手する。

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付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
- ・調整医療費の推計と地域毎医療支出目標額の設定
- ・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
- ・1次・2次・3次予防および予防・医療・介護の連携の実態把握
- ・研究者用データ提供に向けた検討と実装

### 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。保険者データについても作成し、以後更新を重ねる予定である。

今後、多次元分析についても同様に進める予定である。連携学会は、日本腎臓学会、

日本糖尿病学会の二つに限定した。すでに定義体の第一バージョンは作成し、保険者のデータを利用して集計を開始した。

都道府県・二次医療圏・市町村別分析、調整医療費の推計と地域毎医療支出目標額の設定、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については、検討会の開催のみとして概念整理と分析手法の検討にとどめ、データの解析は次年度以降実施することとした。

1次・2次・3次予防および予防・医療・介護の連携の実態把握は、三重県の国保と後期高齢者をフィールドに医療および介護レセプト情報を連結し、連携の実態把握を行った。

研究者用データ提供に向けた検討と実装は、1件の提供を行った。現在、2件目に着手している。

### 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。

大規模データ研究の標準化や教育の充実、人材育成については、医師、医療政策研究者、医療情報研究者達からなる我が班の利用経験に加えて、行政担当者・看護師（保健師）・介護職員等にも資する標準化されたデータの利用に関する教育・人材育成を視野に入れる。H28 年度は、協力保険者の県庁職員、医師会、各市町村国保の保健師に対する教育等を計画している。

今後も引き続き、最先端のビッグデータ基盤と最先端の医療政策・経済等研究を連携させ、研究者、厚生労働省、地域医師会、地方自治体、保険者らステークスホルダーがデータや BI ツールを使い論文化・実態把握・政策提言を行い技術的にも社会的にもメリットが得られるようにする。また、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についても検討を続ける。

## E. 結論

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時間を要したが、今後も引き続き、最先端のビッグデータ基盤と最先端の医療政策・経済等研究を連携させ、研究者、厚生労働省、地域医師会、地方自治体、保険者らステークスホルダーがデータや BI ツールを使い論文化・実態把握・政策提言を行い技術的にも社会的にもメリットが得られるようにする。

## 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/000012>

5318. pdf

II. 知的財産権の出願・登録状況  
該当なし

厚生労働科学研究・戦略研究  
健康医療分野における大規模データの分析及び基盤整備に関する研究

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

研究実施計画書



## 内容

0. 研究の名称 .....	2
1. 研究目的 .....	2
2. 研究の背景と意義 .....	2
3. リサーチクエスチョン：クリニカルクエスチョンと医療政策的・経済的クエスチョン.....	5
4. 対象（データ） .....	8
変更点： NDB データは、予定通り申請済み。既に提供を受けている、本研究方法「9) 研究者用データ提供に向けた検討と実装」については、データ延長手続きを行っており 2016 年末まで利用可能。既にデータセットの提供を試行提供（東京大学）しており、今後 2 件目（早稲田大学）に対応する。 .....	9
5. 研究の方法：必要性、目標、手順、人員・組織、スケジュール、成果 .....	10
1) ビッグデータ解析基盤の増強 .....	10
2) 全 NDB データの検証と質向上 .....	12
3) 基礎統計 .....	14
4) 多次元分析(クリニカルクエスチョン) .....	14
5) 都道府県・二次医療圏・市町村別分析(Small Area Analysis: SAA) .....	17
6) 調整医療費の推計と地域毎医療支出目標額の設定.....	18
7) 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 .....	20
8) 1次・2次・3次予防および予防・医療・介護の連携の実態把握.....	22
9) 研究者用データ提供に向けた検討と実装 .....	24
6. スケジュール（研究期間） .....	25

## 0. 研究の名称

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

## 1. 研究目的

本研究の目的は、申請者がこれまで構築してきた高速レセプト・ビッグデータ解析基盤を更に発展させることにより、以下の医療の需要・供給、質、コストが国・地域・医療機関レベルで即座に可視化できるサービスを確立する。

- 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

### 変更点

- 1)は、4 学会から 2 学会（日本糖尿病学会、日本腎臓学会）に規模を縮小した。
- 2)は、検討会の開催のみとして、今年度中に 1 回の開催予定である。集計結果を基に検討する必要があるため、2 月初旬を予定している。
- 3)は、2)と同様である。

## 2. 研究の背景と意義

### (背景)

医療分野でのビッグデータの一つに、2009 年度から日本の全患者の保険診療情報を蓄積している国家規模のデータベースがある。これはレセプト情報・特定健診等データベース（以下 NDB）と呼ばれ、このデータを分析し医療政策に利用することが医療の質の改善や医療費適正化の切り札となることが期待されている。

医療費の適正化を目的として、これまで幾多の医療制度改革が行われてきた。例えば 2008 年度には、特定健康診査・特定保健指導（メタボ健診）や後期高齢者医療制度が導入された。メタボ健診は生活習慣病の一次予防により将来的な医療費を削減する目的で導入されたが、これが本当に生活習慣病を予防し、医療費を削減する効果を持つのかに関するエビデンスは未だ示されていない。

エビデンスを示すことが困難であった原因の一つは、これまで診療の実態や診療行為のアウトカム、医療のコストなどを包括的に分析するための国家規模のデータが存在しなかったことである。レセプト情報は保険者に関わりなく同じフォーマットで作成されるため、この目的にかなうデータとして有望視されてきたが、全国規模のデータを集めるシステムが長らく存在しなかったため、その分析は一部の研究者や研究機関による狭いテーマに限

定されてきた。NDB には、全医療機関で提供された医療をカバーするレセプト情報、特定健診情報が収録され、2013 年度の社会保障制度改革国民会議の報告書（内閣官房）においても「国が保有するレセプト等データの利活用の促進にも不可欠である」と大きな期待が寄せられている。

また、今後の超高齢社会に向けて目指すべき地域包括ケアシステム構築の中で鍵となる在宅医療についても、現状ではその診療実態はほとんど明らかにされておらず、大きな課題となっている。在宅医療の質を定義する難しさの 1 つには、その対象となる患者において認知症や終末期など自律的判断を行えない状態にある者が多く、また治癒を目標に据えられない慢性期医療独特の特徴があることから、診療の評価軸を一律に設定しにくいという点などがあげられるが、現状はそれ以前に、どのような疾患・状態にある患者にどのような在宅医療が提供されているかといった実態さえ明らかにされていない。ビッグデータを活用してまずは実態を確認し、特定の保険者の健診・医療・介護データから在宅医療のエビデンス・パフォーマンスを集積していく端緒を開くことが必要かつ最も効率的である。

#### （意義）

ビッグデータの利用環境が整備される中、重要課題はビッグデータの具体的な活用事例を提示することである。ビッグデータがその価値を発揮するには明確に定義された「検証すべき仮説」が必要であり、仮説思考に基づかないビッグデータは単なる大量のデータにすぎず、戦略なきデータ解析から得られる情報は往々にして利用価値の低いものになりがちである。そこで本研究ではまず、NDB やそれに準ずるビッグデータ（介護給付情報、被保険者マスタ情報等）の分析が、医療の需要と供給、質、コストに関してどのような情報を提供できるのかを検討する。ビッグデータの解析から得られた情報を社会に還元する方策について具体的事例を示し、そのリスクと便益を検証することは、NDB あるいはヘルスケア・ビッグデータの管理・運営ならびに利活用のあり方に資する有用な知見をもたらす。

申請者はこれまでに我が国国民全体を対象として年間約 400 億件のレコードをリアルタイムに解析し可視化し得る高速レセプト・ビッグデータ解析基盤を構築してきた。国全体を俯瞰可能なレセプト解析システムは、世界的に見ても他に類がなく、本研究においては、当該基盤の更なる発展と広汎な利活用を通じて、NDB 全データの精度検証情報の公開、ダッシュボード機能等を利用した診療情報の集計結果の公表、研究者用データセットの整備、更なる研究利用の振興と裾野を広げることにより、我が国のヘルスケア・ビッグデータの利活用を世界トップレベルに引き上げることができると確信している。

本研究の政策的な意義について、以下に述べる。内閣官房（社会保障制度改革推進本部）は「医療・介護情報の活用による改革の推進に関する専門調査会」において、医療費を適正化するための施策として都道府県ごとに医療支出目標を設定することを検討している。ヘルスケア・ビッグデータの解析により地域の医療の需給バランスが明らかになれば、医療支出目標の設定並びに適切な医療資源配置について具体的な方法論を提示・確立できる。

過去に医療制度改革推進本部（厚生労働省）において性別と年齢によるリスク構造調整手法が保険者の財源調整方法が検討されたが、実際の推計は行われなかった。2015 年の通常国会において国民健康保険の運営を市町村から都道府県へ移管する法案が提出される予定だが、移管に当たり全国知事会等は財政基盤強化策・政府部門の国費投入の規模の

提示（エビデンス）を求めている。また、都道府県の運営にあたっては、医療費適正化の実施が各都道府県に課せられること、そのためにリスクを加味した適正な医療費の目安や予測の方法が求められる。すなわち現在の保険者間の財源調整は、前期高齢者では「前期高齢者」と「それ以外」の区分の加入率のみに基づく調整であり、後期高齢者では医療費が発生した後の事後的な調整に過ぎない。各都道府県が保険者として医療費の適正化に取り組むためには、保険者毎のリスクプロファイルとそれに基づく適正医療費の目安、さらには公平性の高い財政調整措置の方法論が求められる。そこで、本研究より提示する諸外国のモデルに基づいた地域カテゴリ別のリスク構造の推計結果を、全国町村会、全国市町会、全国知事会、地域保険者、都道府県国保連合会、各自治体の首長へ情報提供することで、国民健康保険広域化の検討材料とする。

理論的には医療費は患者の属性と医学的な要因だけで決定されるべきだが、実際には医療費は医師による診療パターンの違いや医師密度、医療機関の利用のしやすさ、医療機器や設備の配置などさまざまな要因の影響を受ける。また既存の研究により、医療費のバラツキには、医学的な要因よりもその他の医療提供側の要因や社会的要因の寄与がはるかに大きいことが示されている。このような場合、医療費を適正化するには医療費が地域により異なることを示すだけでは不十分で、そのバラツキが何から生じるのかが明らかにされない限り適正化の対策はあり得ない。また、在宅医療との連携という今後の政策転換においても、在宅医療の実態は明らかになっていない。即ち、ビッグデータの分析の範囲は医療費に限らず、個々の疾患、診療行為、あるいは診療パターンと地域の医師密度や医療資源との関係など多角的な分析が必要となる。ヘルスケア・ビッグデータとさまざまな政府統計や医療機関の持つデータなどと組み合わせることにより、より具体的な医療費・介護費適正化の対策が立てられ、特にわが国における診療の標準化の取り組みは緒に就いたばかりであり、今後診療パターンや治療の標準化の実態を地域レベル、医療機関レベルで評価することは、医療の質を改善し、在宅医療をはじめとする医療資源を将来に効果的に配分するための欠かせない取り組みと言える。

診療の質の評価を改善に結びつける方策は大きく2つある。ひとつは選択による改善であり、診療情報の分析結果を公表し、患者が医療機関を選択する際の参考としてもらうやり方で、質の低い医療機関が淘汰されることによる質の改善である。もうひとつは医療者による自主的な改善であり、これは全ての医療者の自主的な努力により全体的に質が改善されることを目的とする。医療の質の評価と結果の公表が古くから行われてきた米国をはじめとする欧米諸国の研究によると、診療パフォーマンスの公表が患者が医療機関を選択する際に利用されることはほとんどないが、医療者は公表結果を非常に気にするため、結果的に診療の標準化や質の改善等がみられるという。わが国では診療パフォーマンスの評価は始まったばかりであり、データの利用可能性の点からこれまで専らDPCデータによる入院診療を対象とした評価が行われてきた。これに対して医療費の約半分を費やす外来診療の分析は、これまでほとんど行われてこなかったが、生活習慣病の管理を中心とするプライマリ・ケアおよび在宅医療における診療の標準化は喫緊の課題であり、これらの診療を評価する方法論の開発や、結果のフィードバックを通じた診療の質の改善などは、わが国の診療の質の改善と政策転換に大きな威力を発揮するはずである。

ただし、NDB データを使って医療機関を特定することは禁止されているために、提供単位としては県・二次医療圏レベルの情報提供（フィードバック）となる。ただし、本研究では500万人規模の保険者をフィールドとして、医療・介護・健康データの取得を済んでいる。保険者には、市町村および医療機関別の情報提供（フィードバック）を行うもので