



ESRI Research Note No.65

医療の質の変化を反映した価格の把握手法に関する研究
—レセプトデータ(悉皆)による試算—

西崎寿美、近藤雄介、大里隆也、菊川康彬

May 2022



内閣府経済社会総合研究所
Economic and Social Research Institute
Cabinet Office
Tokyo, Japan

ESRI Research Note は、すべて研究者個人の責任で執筆されており、内閣府経済社会総合研究所の見解を示すものではありません (問い合わせ先：<https://form.cao.go.jp/esri/opinion-0002.html>)。

ESRI リサーチ・ノート・シリーズは、内閣府経済社会総合研究所内の議論の一端を公開するために取りまとめられた資料であり、学界、研究機関等の関係する方々から幅広くコメントを頂き、今後の研究に役立てることを意図して発表しております。

資料は、すべて研究者個人の責任で執筆されており、内閣府経済社会総合研究所の見解を示すものではありません。

The views expressed in “ESRI Research Note” are those of the authors and not those of the Economic and Social Research Institute, the Cabinet Office, or the Government of Japan.

医療の質の変化を反映した価格の把握手法¹に関する研究
—レセプトデータ（悉皆）による試算—

西崎寿美、近藤雄介、大里隆也、菊川康彬*

要旨

内閣府経済社会総合研究所では、GDP 統計の推計の精度向上を図ることを目的に、医療の質の変化を反映した価格の把握手法について研究している。本稿は、「レセプト情報・特定健診等情報データベース」（以下、NDB という。）のレセプトデータ（悉皆）を使用し、傷病ベースでの医療デフレーターを推計した初めての試みである。推計結果について考察するとともに、推計の過程で明らかとなったレセプトデータを分析データとして使用する上での留意点やその課題等について整理した。

目次

1. 概要.....	2
2. NDB の悉皆データを使用した傷病ベースの医療デフレーターの推計方法.....	4
2. 1. NDB の悉皆データとサンプリング・データの相違.....	4
2. 2. 患者の特定（名寄せ）について.....	5
2. 3. 傷病の特定について.....	8
3. 傷病ベースの医療デフレーターの算出結果.....	11
4. 考察と課題.....	15
参考文献.....	17
付図.....	18
補論A. 医療給付実態調査との比較.....	19
補論B. 電子レセプトの課題.....	20

- * 西崎寿美 : 内閣府経済社会総合研究所上席主任研究官
 近藤雄介 : 内閣府経済社会総合研究所政策調査員
 (三井住友海上火災保険株式会社より内閣府に派遣)
 大里隆也 : 内閣府経済社会総合研究所研究協力者、株式会社帝国データバンク
 菊川康彬 : 内閣府経済社会総合研究所研究協力者、株式会社帝国データバンク

¹ 本稿の作成にあたっては、原湖楠研究員（アリゾナ大学）より有益なアドバイスを頂くとともに、桑原進総括政策研究官をはじめとした内閣府経済社会総合研究所の職員からも有益なコメントを頂いた。記して感謝したい。なお、本稿の内容は、筆者が属する組織の公式の見解を示すものではなく、あり得べき誤り等内容に関してのすべての責任は筆者にある。

1. 概要

経済社会総合研究所では、「統計改革の基本方針」（平成 28 年 12 月 21 日経済財政諮問会議決定）、及び「公的統計の整備に関する基本的な計画」（平成 30 年 3 月 6 日閣議決定）などに基づき、国民経済計算（System of national accounts、以下「SNA」という。）の推計の精度向上を目的とした取組みの一つとして、関係府省と連携し、医療の質の変化を反映した価格の把握手法について研究を行っている²。

2017 年度より、まずは、SNA における医療の質の調整に関する欧米諸国での研究の蓄積および具体的な採用状況を整理することから始めた³。これによると、国際機関では、「傷病の治療」を医療サービスのアウトプットと定義していることから、日本においては、傷病分類ごとの治療コストを計測し、傷病ベースでの医療デフレーターを推計することが、日本の国民経済計算（JSNA）との親和性が高く望ましい。このためには、患者ごとの傷病の特定、診療行為の記録が必要となるが、近年、保険請求のためのレセプトの電子化、第三者提供が進められつつあることから、初めての試みとして、レセプトデータを利用した傷病ベースでの医療デフレーターを推計作業を開始した。まずは NDB のサンプリングデータを利用し、傷病の分類は主に ICD-10（2013 年版）に準拠した「疾病、傷害及び死因の統計分類」を使い⁴、傷病分類ごとの治療コストを計測し、傷病ベースでの医療デフレーターを試算した。

しかしながら、NDB のサンプリングデータでは、傷病分類ごとの治療コストは、レセプト 1 件当たりの平均単価でしか計測することができなかった⁵。これを、国際機関が推奨する「傷病の治療」の概念に近づけるために、本稿では、新たに厚生労働省より提供された NDB の悉皆データを用いて患者一人につき一つの傷病を特定し、その 1 年間の平均診療報酬点数を傷病ごとの価格と定義した傷病ベースでの医療デフレーターを推計することとした。また、この推計を行う上で必要な、患者を特定するための名寄せの手法と患者の傷病を特定するための手法について検討を行った。

推計結果をみると、傷病ベースでの医療デフレーターは緩やかに上昇していることがわかった。傷病別にみると、「循環器系の疾患」の貢献が最も大きく、これは、「循環器系の疾患」が、高齢者ほど併存疾患の増加などにより治療費が上昇することによるものと思われる。

しかしながら、この推計結果が質の変化を十分に反映したものであるかについては、まだ多くの課題が残されていることから評価が難しい。第一に、レセプトは保険診療の請求のためのものであり、分析のために設計されたものではないことから、患者や傷病を特定することが困難であるなど、分析データとして多くの問題が存在することがあげられる。具体的には、今回の推計は、最も診療報酬点数が高い月のレセプトで定義されている主傷病に患者の 1 年間の治療費を全て割り振るなど、非常に大胆な仮定を置いて行ったものであり、こうし

² 問題意識の詳細については、石橋等(2020)を参照。

³ 野口等(2019)

⁴ これによって均質かつ相互に排他的であり、網羅的に分類を行うことができると仮定。

⁵ レセプトに含まれるデータとその課題については、石橋等(2020)を参照。

た点を十分留意しなければならない。したがって、傷病ごとの治療コストを計測するためのデータがレセプトデータ以外にない現状において、「傷病の治療」の価格を精緻化することは難しいといわざるをえない。第二に、傷病ベースでの医療デフレーターが上昇していることの解釈である。「傷病の治療」の価格が上昇していたとしても、治療の成果（アウトカム）が向上しているのであれば、価格の上昇を過剰に推計しないよう、品質調整を行うことが必要となる。治療の成果（生存率、平均在院日数、QOLなど）の計測については、現在、日本も含め各国において指標の選定、傷病別での計測手法等、検討が進められているところである。

このように今回の推計は、国際機関で推奨されている「傷病の治療」を医療サービスのアウトプットととらえ、レセプトデータ（悉皆）を使用した「手探り」の手法で、傷病ベースの医療デフレーターを計測した初めての試みである。今後、この計測方法を発展させ、SNAに実装するためには、より頑健な方法で患者を特定し、合併症も含め、きめ細かい傷病に紐づいた患者の診療行為が記録されたデータセットが必要となる。未だ遅れているカルテの電子化を進めることが理想であるが、今回の推計同様、保険請求のためのレセプトデータを利用する場合、患者や傷病を特定できるよう、今後、レセプトデータを整備することが最低限必要であろう。また、レセプトデータの第三者提供を進めることで、レセプトデータを分析データとして利用する上でのノウハウの蓄積が期待できる。

もちろん、「傷病の治療」の価格についての研究は、ICD-10コードなど国際的に同一の疾病分類が普及し、疾病ごとの診療行為、薬剤などの標準治療の研究、国際比較が進むなかで、重要な研究テーマの一つである。治療の成果（アウトカム）の研究とも並行して、費用便益分析も盛んに行われており、今後とも医療の質の向上、効率化に向け、進めていくべき研究であることに変わりはない。

2. NDB の悉皆データを使用した傷病ベースの医療デフレーター の推計

内閣府では、これまで「医療の質の向上を反映した価格」を把握するため、医療のアウトプットを「傷病の治療」と定義した上で、患者一人の特定の傷病を治療するために要した初診から治癒までの診療行為全体の合計を価格として計測することを追求してきた。

こうした傷病ベースでの推計を行うためには、患者一人ひとりの医療機関にかかった全ての経歴、すなわち、いつ、どのような傷病で、どの医療機関で、どのような診療行為、医薬品を受けとったか、を記録したデータが必要となる。電子カルテが普及していない日本においてこれに相当するデータとしては、NDBに含まれる、ほぼすべての被保険者をカバーした保険請求書である電子レセプトのみが利用可能である。このため、内閣府では、2018年度よりNDB⁶を使用した傷病ベースでの医療デフレーター の推計を開始した。

まずは、NDBの特別抽出によるサンプリングデータによる推計を行ったが、サンプリングデータでは、患者を追跡することができないことから、結局はレセプト1枚ごとの医療費を価格と定義し、推計せざるを得なかった。しかし、今回は、NDBの悉皆データを使用することにより、患者を特定し、その患者の1年間のレセプトを集計することができるようになったことから、より「傷病の治療」の価格に近似した推計が可能となった。

一方、レセプトに含まれるデータを利用する過程において、保険請求を目的とするレセプトデータでは、患者を特定し、治療した傷病を特定することには限界があることも判明した。本章では、今回の推計において、レセプトデータを利用し、どのように患者を特定し、どのように患者の傷病を判断し、傷病ごとの価格をいかに合算してデフレーターを推計したのか、その手法について述べる。1節でNDBのサンプリングデータと悉皆データの推計の相違点をまとめ、2節で患者の特定コードに関する言及と名寄せ手法について、3節で傷病の特定方法について説明する。

2.1. NDB の悉皆データを用いた推計とサンプリングデータを用いた推計の相違

これまでに研究を進めてきた傷病ベースでの医療デフレーター の推計を、新たに入手したNDBの悉皆データを使用して行った。傷病ごとの価格を算出する上で、前回のサンプリングデータと今回の悉皆データとの重要な違いは次の2点である。

一点目は、患者の追跡が可能かどうかである。サンプリングデータには患者を特定するIDが含まれず、複数月に渡り同一患者を追跡することができない。このため、「傷病の治療」、すなわち、理想的には、患者一人の特定の傷病を治療するために要した初診から治癒までの診療行為全体の合計を価格と認識する、という定義とは大きく異なり、レセプト1件あたりの平均診療報酬点数を価格として計測せざるを得なかった。しかし、悉皆データでは、異な

⁶ NDBのレコード等解説については、石橋等(2020)を参照。

る月でも患者を特定する ID が存在することから、同一患者を追跡することができるため、患者一人の一定期間における治療コストを計測することが可能となる⁷。本稿では米国 BEA を参考に、患者一人の1年間の治療コストを「傷病の治療」として計測することとした。

二点目は、患者の傷病の特定方法である。サンプリングデータではレセプト1件あたりを価格としたため、そのレセプトに記載された主傷病を採用して ICD-10 コードの傷病分類に割り振った。しかし、悉皆データにおいては、ある年のデフレーターを算出するために、12 か月分存在する月次のレセプトデータをすべて統合して活用する。その際に問題となるのが、12 か月の間に同一の患者が複数回受診した場合、受診時点によって主傷病が異なるという状況が生まれることである。そのため、患者ごとにある年の主傷病を定義する際に、どの時点のレセプト情報による主傷病を採用するのかを検討する必要がある。本研究では、第一案として、ある年における初回の受診時の主傷病を年次の主傷病として定義する方法をとった。しかし、この方法では季節性が存在するような傷病において、実態との乖離が大きくなりやすいことが確認されたため、改善案として診療報酬点数の合計が最も高い月における主傷病を年次の主傷病として定義することとした。

2.2. 患者の特定（名寄せ）手法について

NDB の悉皆データを活用し、傷病ベースの医療デフレーターを正確に算出するためには、母数となる患者数の精緻化は必須である。現行ではマイナンバー等の生涯不変の個人に固有かつ不変な情報は存在しないことから、レセプト同士を紐づけるために用意された2種類の個人 ID である「ID1」ならびに「ID2」を用いて名寄せを行う必要がある。ID1 は「保険番号+被保険者の記号・番号+生年月日+性別」から、ID2 は「氏名+生年月日+性別」から生成されており、いずれかの ID のみでは、保険離脱（失職、新卒、定年退職、後期高齢者医療制度への移行）や、結婚などの氏名が変わることにより、追跡が不可能となる。このため、ID1 と ID2 のいずれも活用した名寄せのロジックの開発が進められている。

奈良県立医科大学の研究グループでは、ID1/ID2・転帰区分を考慮した名寄せロジックの開発と検証が精力的に行われている。その手法を紹介するとともに、それを参考に医療デフレーターを推計する上で適当な名寄せロジックについて検討する。

まず、奈良県立医科大学の研究グループが開発した名寄せロジック（以下で、名寄せロジック A と呼ぶ）の概要は次の通りである。

⁷ ただし、保険離脱（失職、新卒、定年退職、後期高齢者医療制度への移行）により、単独の ID では患者の追跡は不可能であり、複数の ID を利用しての患者の名寄せが検討されている。

[名寄せロジック A]

- ① ID1 と ID2 が同一である場合は、同一の患者とする。
- ② ID1 で名寄せを行い、ID1 で紐づかない場合は ID2 での紐づけを行う。
 - ・ ID1 と ID2 のそれぞれで紐づけが可能な場合は、ID1 を優先して紐づけを行う
 - ・ ID1 が紐づけ可能な場合でも、同月内に異なる ID2 が複数存在していた場合、紐づけを実施しない
 - ・ 紐づけ可能な ID2 が紐づけ可能な場合でも、同月内に異なる ID1 が複数存在していた場合、紐づけを実施しない
- ③ 転帰区分：死亡となった月以降の名寄せは実施しない

図表 1 名寄せが困難なケース

① ID2 でのみ紐づくが、ID1 が後月に2種類以上存在	② ID2 で紐づくが、ID1 が前月に2種類以上存在	③ ID1 が一意であるが、ID2 が3種類以上存在																																								
<p>2018年1月</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>xxxxx1</td><td>xxxxx2</td></tr> </table> <p>2018年2月</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>yyyyyy</td><td>xxxxx2</td></tr> </table> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>zzzzzz</td><td>xxxxx2</td></tr> </table>	ID1	ID2	xxxxx1	xxxxx2	ID1	ID2	yyyyyy	xxxxx2	ID1	ID2	zzzzzz	xxxxx2	<p>2018年1月</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>yyyyyy</td><td>xxxxx2</td></tr> </table> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>zzzzzz</td><td>xxxxx2</td></tr> </table> <p>2018年2月</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>xxxxxx1</td><td>xxxxx2</td></tr> </table>	ID1	ID2	yyyyyy	xxxxx2	ID1	ID2	zzzzzz	xxxxx2	ID1	ID2	xxxxxx1	xxxxx2	<p>2018年1月</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>xxxxx1</td><td>xxxxx2</td></tr> <tr><td>xxxxx1</td><td>xxxxx3</td></tr> <tr><td>xxxxx1</td><td>xxxxx4</td></tr> </table> <p>2018年2月</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>ID1</td><td>ID2</td></tr> <tr><td>xxxxx1</td><td>xxxxx2</td></tr> <tr><td>xxxxx1</td><td>xxxxx3</td></tr> <tr><td>xxxxx1</td><td>xxxxx4</td></tr> </table>	ID1	ID2	xxxxx1	xxxxx2	xxxxx1	xxxxx3	xxxxx1	xxxxx4	ID1	ID2	xxxxx1	xxxxx2	xxxxx1	xxxxx3	xxxxx1	xxxxx4
ID1	ID2																																									
xxxxx1	xxxxx2																																									
ID1	ID2																																									
yyyyyy	xxxxx2																																									
ID1	ID2																																									
zzzzzz	xxxxx2																																									
ID1	ID2																																									
yyyyyy	xxxxx2																																									
ID1	ID2																																									
zzzzzz	xxxxx2																																									
ID1	ID2																																									
xxxxxx1	xxxxx2																																									
ID1	ID2																																									
xxxxx1	xxxxx2																																									
xxxxx1	xxxxx3																																									
xxxxx1	xxxxx4																																									
ID1	ID2																																									
xxxxx1	xxxxx2																																									
xxxxx1	xxxxx3																																									
xxxxx1	xxxxx4																																									

名寄せのロジック (A) では、図表 1 のような「不明確な名寄せのレアケース」では、同一患者とならない。このことから、名寄せのロジック (A) は ID1・ID2 において複数の名寄せの可能性が存在する場合には名寄せをしない、「疑わしきは名寄せせず」のロジックといえる。ただし、久保 (2017) によると、ID1 のみでの患者数から名寄せロジックを利用することで約 5.2%、約 600 万人の統合が行われており、名寄せの効果は絶大であるといえる。

しかしながら、本研究はデフレーターを傷病ごとに算出することを目的としており、患者が分散することで傷病を過小評価してしまう可能性は排除したい。また、デフレーターの算出を目的とし、暦年単位での統合となることから、先行研究のように存在するすべてのレセプトの統合は実施しない。そのため、本研究においては、いずれかの ID が一致する場合には患者数を統合し、患者数を減少させる手法を実施することとする。また、暦年単位での算出のため、各年の 1 月時点のレセプトでの ID を基盤とした名寄せ（以下、名寄せロジック B と呼ぶ）を行っていく。

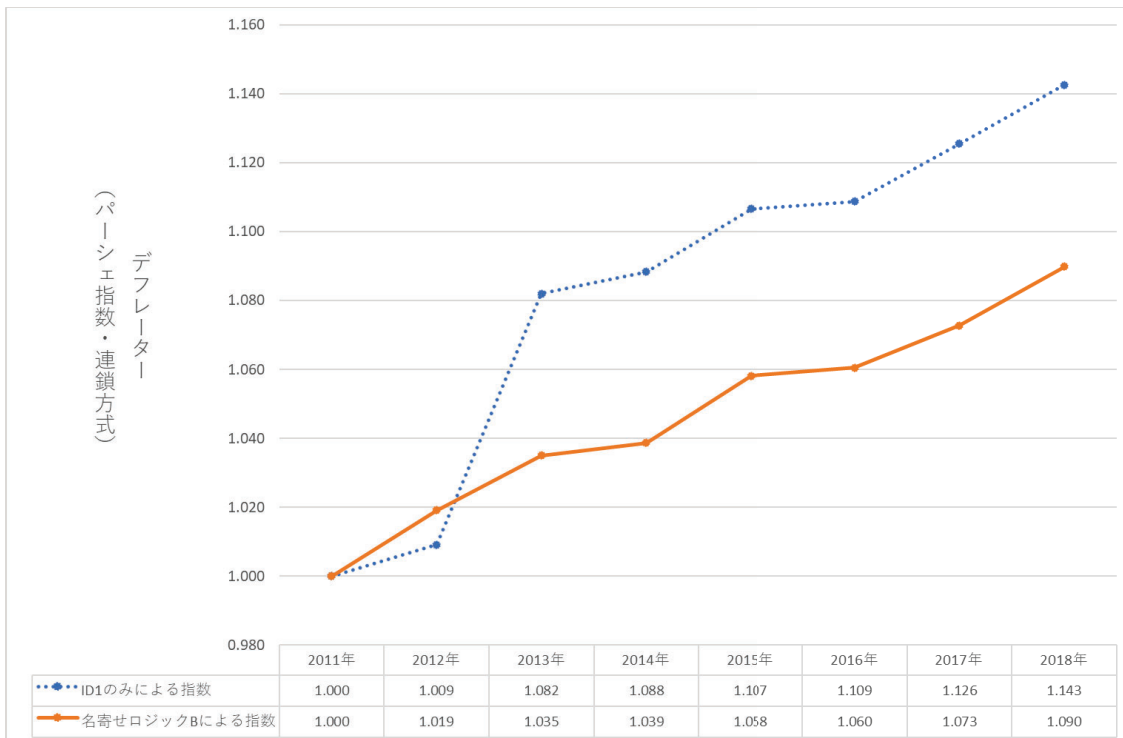
[名寄せロジック B]

- ① 1月時点に存在している ID1/ID2 のペアを軸とする。
- ② 翌月データと ID1 で名寄せを行い、紐づく場合は同一患者とする。複数 ID2 が存在していたとしても、ID1 が同一の場合は1患者扱いとする。
- ③ ID1 が紐づかない場合、ID2 での名寄せを行う。同一 ID2 で ID1 が複数存在している場合も同一患者とみなす。
- ④ 転帰区分が死亡となった月から名寄せは実施しない。

図表2 「ID1」での件数表と「ID1+ID2+転帰区分」名寄せでの件数表の比較

「ID1」での件数表				名寄せロジック(B)での件数表			
診療年	患者数 (千人)	合計点数 (百万点)	平均点数	診療年	患者数 (千人)	合計点数 (百万点)	平均点数
2011	125,035	2,419,276	19,349	2011	109,987	2,419,276	21,996
2012	125,956	2,499,901	19,847	2012	110,153	2,499,901	22,695
2013	119,707	2,553,281	21,329	2013	109,742	2,553,281	23,266
2014	119,791	2,598,456	21,692	2014	110,017	2,598,456	23,619
2015	120,496	2,676,186	22,210	2015	110,523	2,676,186	24,214
2016	121,532	2,731,568	22,476	2016	111,335	2,731,568	24,535
2017	120,469	2,768,791	22,983	2017	110,789	2,768,791	24,991
2018	121,068	2,828,998	23,367	2018	111,328	2,828,998	25,411

図表3 「ID1」のみと「ID1/ID2+転帰区分」名寄せでのデフレーター



このロジックでは、図表1の①～③のケースのすべてのケースで名寄せが行われる。そのため、名寄せロジック(B)は「疑わしきは同一とする」のロジックといえる。図表2は名寄せロジックBによる名寄せを行った結果である。患者数をID1と比較した場合、2015年においてはID1のみで1億2050万、名寄せでは1億1052万と、患者数が1割程度減少することとなった。減少率からも、名寄せロジックAよりも統合される患者数は名寄せロジックBの方が多くなることがわかる。

図表3では、ID1のみでの患者特定によるデフレーターと名寄せロジック(B)を用いた患者特定によるデフレーターを算出している。比較すると、ID1における2012年から2013年での急拡大が、名寄せ実施後には解消されていることが分かる。ID1の急上昇している要因が、図表2での患者数が大幅に減少し、平均点数を押し上げていることと考えられ、名寄せが機能していることを示している。また、2012年から2013年にかけて平均点数が大きく増加している傷病についても変化幅が縮小していることがみてとれた⁸。

2.3. 傷病の特定方法について

次に、名寄せしたレセプトからそれぞれの患者における年間での主傷病を決定する。レセプトは保険診療の請求書であり、審査では個々の医療行為や薬剤に対する適応症が存在するか否かが判断のポイントになるので、全体としての傷病名の重みづけは重要ではない。また、月次のレセプトには、複数の傷病を患っていてもいずれが主な傷病かを特定するための主傷病フラグが存在しているが、年間でレセプトを名寄せした場合に、月で異なる傷病が主傷病とされる場合がある。

以上のことから、レセプトから傷病名を使用した分析を行うには、その分析の目的に応じた傷病名の特定作業が必要である。デフレーター算出にレセプトを用いる場合は、病名を除くすべての傷病名を均等の重みで扱わなければならない。特に傷病別の医療費分析では、傷病名に重みづけがない以上、疑い病名を除くすべての傷病名に均等に医療費を配分せざるを得ないことになる。

2.3.1. 発生時期優先から点数優先への変更

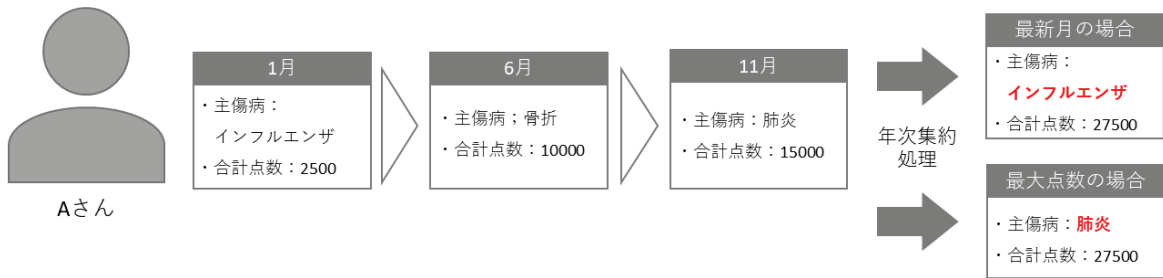
我々は、年間の主傷病の特定に1～12月にかけての最初の月のレセプトの主傷病を用いた。図表4ではそのイメージであり、インフルエンザが1月に罹患していることから、この患者の年間の主傷病と特定している。

しかしながら、この方法では季節性のある傷病にバイアスが大きく存在していることが判明した。例えば、インフルエンザは1～3月に最も流行するため、最初の月の傷病を割り当ててしまうと、インフルエンザを疾患した患者のほとんどをインフルエンザとして年間

⁸ なお、循環器系の平均点数で確認したところ、ID1では2012～2013年8.4%増加しているのに対し、名寄せロジック(B)の場合、2012～2013年1.7%の増加と大きく縮小した。

の主傷病を特定してしまい、インフルエンザによる点数や患者数を過大に評価することとなる。「インフルエンザ及び肺炎」の変動が大きいことを検証するために、インフルエンザに罹患したと判別された患者について、年間のインフルエンザ以外での診療報酬点数を集計し、結果を図表5に記載した。

図表4 主傷病の特定手法



図表5 インフルエンザ判別患者のインフルエンザ以外の点数の集計

合計点数 (単位：百万点)							
審査年	1_インフルエンザ			2_非インフルエンザ			総計
	DPC	MED		DPC	MED		
2011	4,801	180	4,622	17,928	3,804	14,124	22,730
2012	5,740	210	5,531	21,985	4,843	17,142	27,725
2013	5,543	221	5,322	24,838	5,809	19,028	30,381
2014	5,911	207	5,703	21,746	5,043	16,703	27,657
2015	6,650	335	6,315	34,078	7,931	26,147	40,728
2016	6,002	228	5,774	21,508	5,213	16,295	27,509
2017	7,133	340	6,793	32,828	8,135	24,693	39,961
2018	9,500	445	9,055	46,652	12,072	34,580	56,152

点数比率							
審査年	1_インフルエンザ			2_非インフルエンザ			総計
	DPC	MED		DPC	MED		
2011	21%	1%	20%	79%	17%	62%	100%
2012	21%	1%	20%	79%	17%	62%	100%
2013	18%	1%	18%	82%	19%	63%	100%
2014	21%	1%	21%	79%	18%	60%	100%
2015	16%	1%	16%	84%	19%	64%	100%
2016	22%	1%	21%	78%	19%	59%	100%
2017	18%	1%	17%	82%	20%	62%	100%
2018	17%	1%	16%	83%	21%	62%	100%

なお、インフルエンザとしての対象主傷病は、「J09_特定のインフルエンザウイルスが分離されたインフルエンザ」、「J10_その他のインフルエンザウイルスが分離されたインフルエンザ」、「J11_インフルエンザ、インフルエンザウイルスが分離されないもの」、「J14_インフルエンザ菌による肺炎」とした。

図表 5 から、年間での合計点数の内訳をみると、2011～2018 年においてインフルエンザ関連で 16～22%、非インフルエンザ関連が 78～84%となっており、インフルエンザ以外の傷病が点数のほとんどを占めていることがわかる。このことから、名寄せした患者の最初の月の主傷病を年間を通じた特定の傷病と定義することは不適當であることがわかった。

我々は、この季節性を排除するために、点数が最も高い月の主傷病を、当該患者の主傷病として定義することとした⁹。図表 4 でのケースでは、最大点数の月の主傷病を採択し、肺炎が主傷病となる。

2.3.2. 細分類単位「部位不明」の減少（ICD-10 の 5 桁分類から 3 桁分類への変更）

次に、傷病分類の粒度を決定する。本研究では、世界保健機関(WHO)により公表されている「疾病及び関連保険問題の国際統計分類」に準じて定められた ICD-10 分類¹⁰を利用した傷病分類を採用している。

傷病分類の粒度を決めるために注目しなくてはならないのが、ICD-10 における細分類単位での「部位不明」件数が年々減少していることである。例えば、ICD-10 の 3 桁分類「C34: 気管支及び肺の悪性新生物」をみると、全体の患者数が 2011 年の 44 万人から 2018 年には 49 万人に増加しているにも関わらず、その中の 5 桁分類「C34.9: 部位不明」は 2011 年の 33 万人から 2018 年の 30 万人に減少しており、部位の特定が進んでいることがうかがえる。詳細な部位の特定が増加することは歓迎すべきことではあるが、収録率が変動してしまう場合、当初からの収録率が異なることで算出されるデフレーターは大きく異なってしまう。

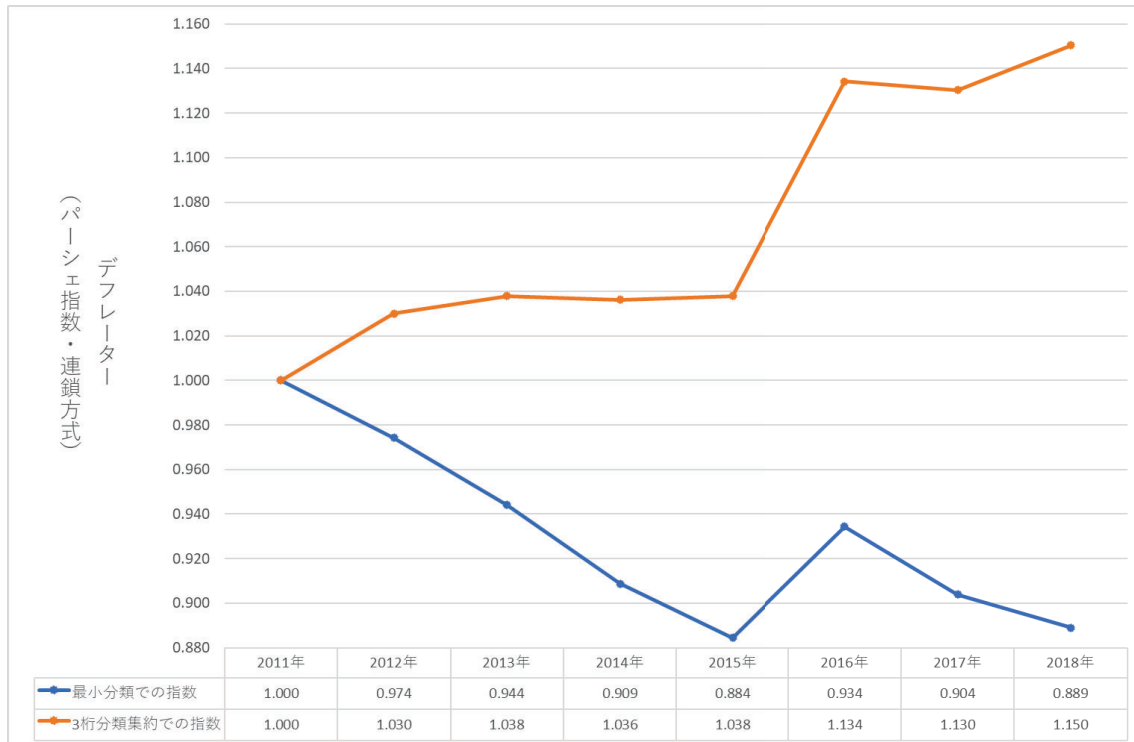
図表 6 では、「C34: 気管支及び肺の悪性新生物」における 5 桁分類でのデフレーターと 3 桁集約したデフレーターを算出した結果である。5 桁分類では 1 を下回るのに対し、3 桁分類による「気管支及び肺の悪性新生物」のデフレーターは上昇しており、分類粒度の違いでデフレーターに乖離があることがわかる。したがって、我々は収録率によるバイアスを除くために、3 桁分類に集約してデフレーターを算出することとした¹¹。

⁹ 発生時期と点数による傷病の特定の違いによるデフレターの比較は付図 1 に記載。

¹⁰ 石橋(2020)参照。

¹¹ 発生時期と点数による傷病の特定の違いによるデフレターの比較は付図 1 に記載。

図表6 「気管支及び肺の悪性新生物」の5桁分類と3桁分類でのデフレーター比較



3. 傷病ベースの医療デフレターの算出結果

第2章で検討したNDBの悉皆データによるデフレーター算出方法を用いて、本章では、1患者1傷病を特定し、1年間の平均診療報酬点数を傷病ごとの価格と定義し、傷病ベースでの医療デフレターを試算する。

まずは、NDB全体の患者数と一人当たり平均点数の推移を図表7に示した。図表7によると、患者数はほぼ横ばいであるが、一人当たり平均点数は上昇傾向にあることがわかる。これを傷病別に求めたのが図表8である。図表8によると、合計点数が高い傷病において患者数が増えているのは「新生物<腫瘍>」、「神経系の疾患」、「内分泌、栄養及び代謝疾患」、平均点数が上昇しているのは「循環器系の疾患」、「損傷、中毒及びその他の外因の影響」、「筋骨格系及び結合組織の疾患」、平均点数が高いのは「新生物<腫瘍>」、「循環器系の疾患」、「神経系の疾患」となっている。

次に、NDBのデータ収集率が高い2011年を基準として、パーシェ（連鎖基準）価格指数による、ICD-10の3桁分類に集約した傷病ベースの医療デフレーターを算出した。図表9をみると、パーシェ（連鎖基準）価格指数は緩やかに上昇していることがわかる。

図表7 患者数及び患者一人当たり診療報酬平均点数の推移

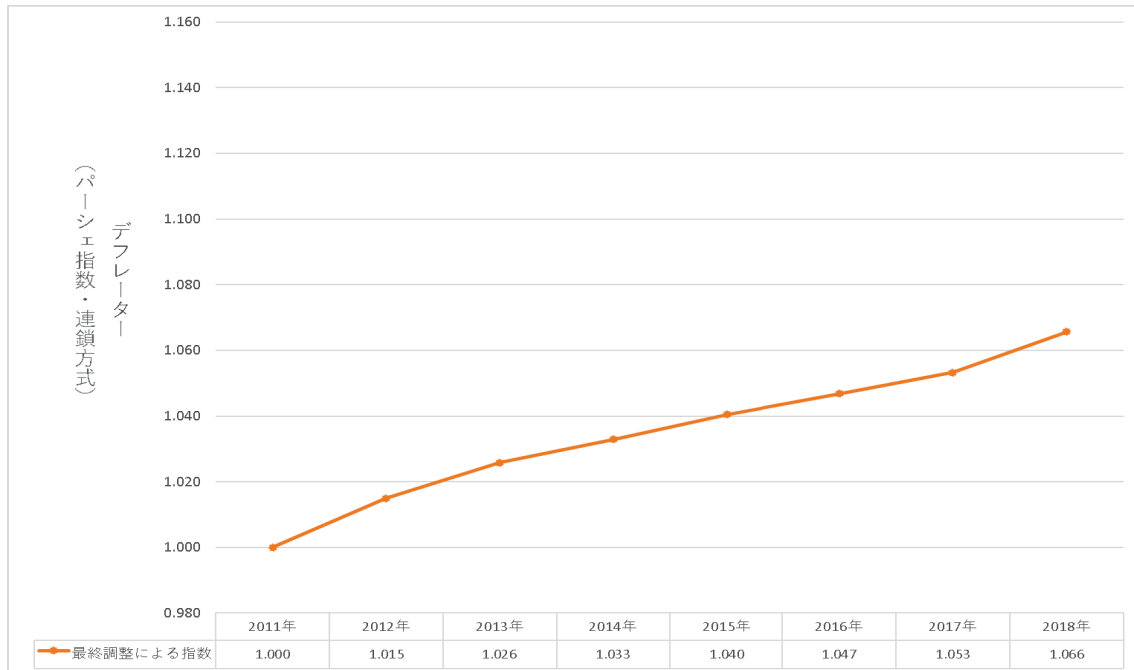
NDB全件			
診療年	患者数 (千人)	合計点数 (百万点)	1人当たり 平均点数
2011	109,987	2,419,276	21,996
2012	110,153	2,499,901	22,695
2013	109,742	2,553,281	23,266
2014	110,017	2,598,456	23,619
2015	110,523	2,676,186	24,214
2016	111,335	2,731,568	24,535
2017	110,789	2,768,791	24,991
2018	111,328	2,828,998	25,411

図表8 傷病別患者数および患者一人当たり診療報酬平均点数の推移¹²

コード名	患者数 (千人)			平均点数		
	2011年	2015年	2018年	2011年	2015年	2018年
感染症及び寄生虫症 (A00 - B99)	5,654	5,704	5,199	12,979	13,991	14,342
新生物<腫瘍> (C00 - D48)	6,405	7,002	7,264	55,771	57,821	60,855
血液及び造血器の疾患並びに免疫機構の障害 (D50 - D89)	736	757	761	40,398	39,852	41,307
内分泌、栄養及び代謝疾患 (E00 - E90)	6,667	7,008	7,252	18,917	19,479	19,353
精神及び行動の障害 (F00 - F99)	2,494	2,570	2,643	49,160	49,239	48,302
神経系の疾患 (G00 - G99)	2,121	2,327	2,391	43,111	46,073	49,350
眼及び付属器の疾患 (H00 - H59)	9,522	9,555	9,469	9,319	10,677	11,404
耳及び乳様突起の疾患 (H60 - H95)	2,065	2,138	2,050	9,342	9,666	10,150
循環器系の疾患 (I00 - I99)	10,638	10,662	10,780	41,548	45,148	47,173
呼吸器系の疾患 (J00 - J99)	23,745	22,524	23,861	9,516	10,596	10,492
消化器系の疾患 (K00 - K93)	8,496	8,779	8,305	22,736	23,892	25,656
皮膚及び皮下組織の疾患 (L00 - L99)	6,009	6,256	6,043	8,150	8,744	9,466
筋骨格系及び結合組織の疾患 (M00 - M99)	8,467	8,757	8,888	20,352	22,554	24,636
腎尿路生殖器系の疾患 (N00 - N99)	4,147	4,333	4,481	33,391	34,477	35,370
妊娠、分娩及び産後<褥> (O00 - O99)	779	846	792	25,520	27,236	29,025
周産期に発生した病態 (P00 - P96)	185	214	220	72,568	77,644	78,401
先天奇形、変形及び染色体異常 (Q00 - Q99)	270	313	343	47,580	49,758	49,879
症状、徴候及び異常臨床所見 異常検査所見で 他に分類されないもの (R00 - R99)	2,954	3,265	3,346	15,235	15,375	15,088
損傷、中毒及びその他の外因の影響 (S00 - T98)	5,944	6,381	6,522	27,870	33,023	36,133
傷病及び死亡の外因 (V01 - Y98)	0	0	0	17,791	17,233	24,302
健康状態に影響を及ぼす要因及び 保健サービスの利用 (Z00 - Z99)	100	140	154	37,507	41,562	40,603
特殊目的用コード (U00 - U99)	0	0	0	130,487	26,930	84,051
分類不明	2,588	991	566	11,309	15,188	17,501

¹² 最もレセプトの点数が高い月の主傷病に患者の1年間の治療費を全て割り振って計測した結果であり、十分留意が必要。

図表9 傷病ベースでの医療デフレーターへの推計結果

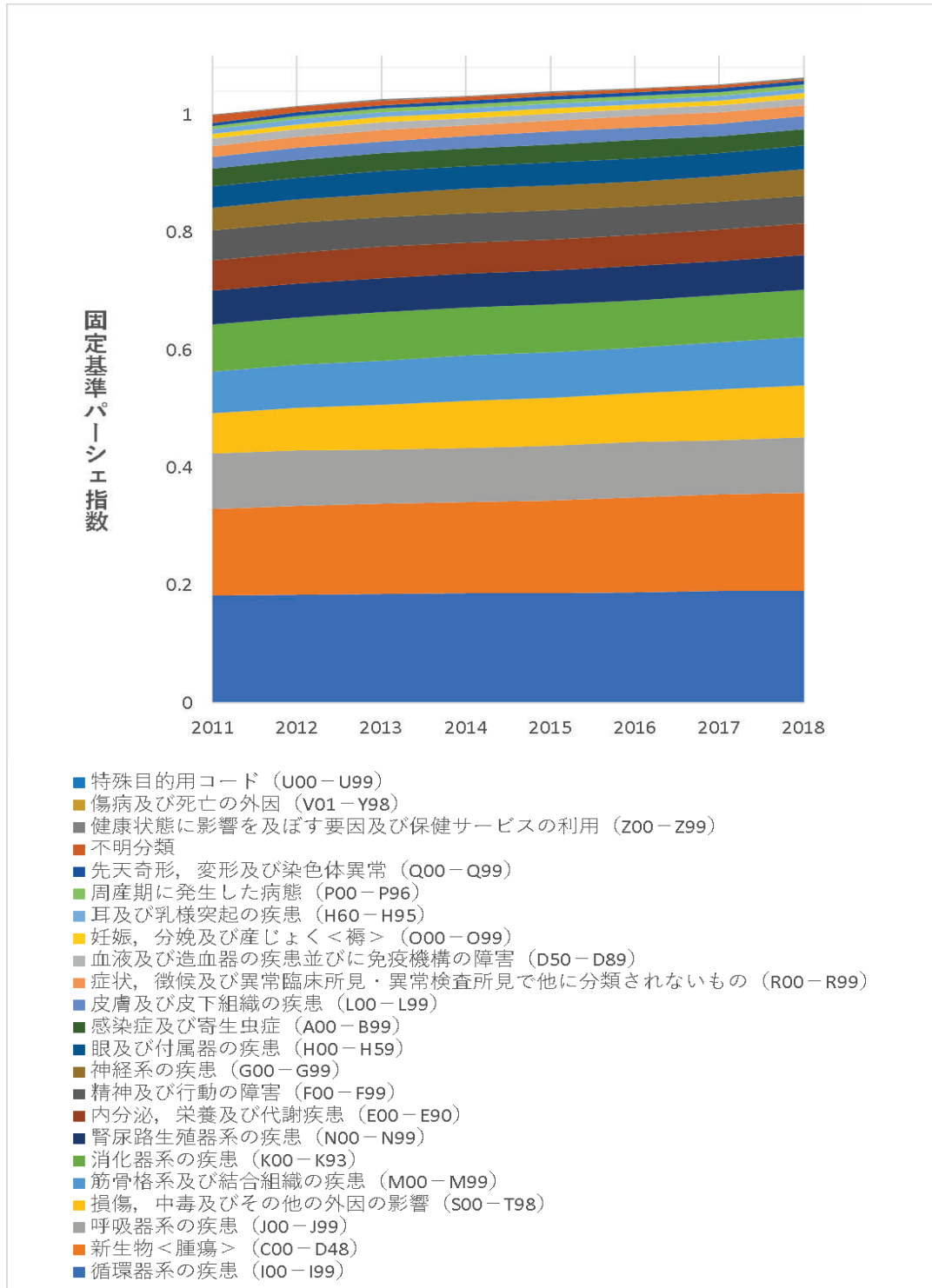


デフレーターへの傷病ごとの寄与度を求めたのが図表10である。なお、傷病ごとの割合を示すために、パーシェ指数の連鎖方式ではなく固定方式での算出となっていることを了承されたい。

上昇要因について、傷病別にみると貢献が大きいのは、「循環器系の疾患」、「新生物〈腫瘍〉」、「呼吸器系の疾患」である。いずれも合計点数（治療費総額）が大きい疾病であり、特に「循環器系の疾患」は、治療費総額が最も高い上に、患者一人当たりの平均単価の上昇が大きいことから、貢献がもっと大きくなっている。その要因としては、循環器疾患（特に、高血圧や心不全などの不整脈、大動脈弁狭窄症などの弁膜症）は、高齢になるにつれて併存疾患が多くなりやすく、治療費が高くなるためではないか¹³と想定することができる。ただし、全体としてみると、特定傷病の拡大ではなく、多くの傷病で拡大傾向にあるといえる。

13 補論A参照

図表 10 ICD-10 章分類における傷病別デフレーターの内訳



4. 考察と課題

今回の推計は、国際機関で推奨されている「傷病の治療」を医療サービスのアウトプットととらえ、NDBのレセプトデータ（悉皆）を使用し、傷病ベースの医療デフレーターを計測したはじめての試みである。これまでに経験したことのない膨大なデータ容量¹⁴、ファイル数¹⁵からなるレセプトデータ（悉皆）を取り扱う過程は、想定外の困難に遭遇しながらの、まさに手探りの状況であったが、はじめてレセプトデータ（悉皆）を扱った意義は大きいものであった。それとともに、保険診療の請求を目的とするレセプトデータを分析に使用することの限界と多くの課題が浮き彫りとなった。

特に困難な点は、患者を特定するために多大な処理が必要なことである。現行のレセプトではマイナンバー等の個人に固有かつ不変な情報はなく、突合は主としてID1と呼ばれる保険者番号＋被保険者番号＋生年月日＋性別をハッシュ値とするものであったが、保険離脱（失職、新卒、定年退職、75歳の後期高齢者医療制度への移行）により追跡することができない。このため、今回は、先行研究を参考に、患者氏名＋生年月日＋性別のハッシュ値であるID2とID1を接合させることで名寄せを行った。しかし、確定された名寄せの手法があるわけではなく、各研究においても個々の目的に応じた手法が採用されているのが現状である。このように、デフレーターの算出において分母となる患者数が正確に把握できないということは、正確なデフレーターを算出する上で大きな問題であるといえよう。

レセプトデータによる患者の主傷病の特定については、さらに深刻な課題がある。まず、レセプトでは傷病名の間で重みづけが不足している。これは、出来高払いである保険請求（レセプト）では、個々の医療行為や薬剤に対する適応性が存在するか否かが判断のポイントとなるので、全体としての傷病名の重みづけが重要視されてこなかったためと思料できる。また、ICD-10コードにより傷病を分類しているが、ICD-10の5桁分類まで細分化すると、患者数の変動が大きいことがわかった。これは、中分類の傷病（例えば呼吸器系の疾患）における「部位不明」が、部位の特定が進むことでその割合を低下させていることが大きな要因となっている。こうした動きは、レセプトの傷病名の明確化が進めば安定するものと考えられるが、いまだその途上といえる。さらに、医科入院外レセプトではすでに治癒している傷病名の記載が残っているケースが多いことも指摘されている。今回の計測では、最も診療報酬点数が高い月のレセプトで定義されている主傷病に患者の1年間の治療費を全て振り分けるといふ、非常に大胆な仮定を置いて実施した。このため、「傷病の治療」にかかった1年間の平均単価という概念からは、大きく逸脱している可能性がある。

¹⁴ 2009年4月～2019年3月で約15TB、バックアップ用ファイル、データの分析処理用の容量も確保すると、最低限でも60TBは必要。

¹⁵ 約1万ファイル

また、そもそも主傷病のみで分類してよいのかという点についても議論がある。すなわち、高齢者の増加により、生活習慣病など多くの併存疾患をもつ患者が急増しており、そのことが傷病ごとの単価を引き上げているものと想定される。このため、併存疾患も含めた傷病の細分化による均質化が必要であるが、併存疾患をレセプトから読み込み、場合分けすることは現状では困難である。

さらにレセプトデータの課題としては、NDBは「高齢者の医療の確保に関する法律（いわゆる高確法）」により、2009年度からデータの収集を開始したため、2009年度以前に遡及することができないほか、実務上の課題としては、容量が大きい、ファイル数が多い、などがあげられる。

こうした推計に使用したレセプトデータの課題とともに、傷病ベースでの医療デフレーターが上昇していることの解釈については、もう一つ大きな留意点がある。それは、「傷病の治療」の価格が上昇していたとしても、治療の成果（アウトカム）が向上しているのであれば、価格の上昇を過剰に推計しないよう、品質調整を行うことが必要となることである。例えば、「新生物<腫瘍>」の患者一人当たりの平均単価は上昇しているが、「新生物<腫瘍>」患者の5年生存率は年々上昇しており、治療の成果は向上しているといえる。こうした、治療の成果（生存率、平均在院日数、QOLなど）の計測については、現在、日本も含め各国において指標の選定、傷病別での計測手法等、検討が進められているところである。

このように今回の推計は、国際機関で推奨されている「傷病の治療」を医療サービスのアウトプットととらえ、レセプトデータ（悉皆）を使用した「手探り」の手法で、傷病ベースの医療デフレーターを計測したはじめての試みである。今後、この計測方法を発展させ、SNAに実装するためには、より頑健な方法で患者を特定し、合併症も含め、きめ細かい傷病に紐づいた患者の診療行為が記録されたデータセットが必要となる。未だ遅れているカルテの電子化を進めることが理想であるが、今回の推計同様、保険請求のためのレセプトデータを利用する場合、患者や傷病を特定できるよう、今後、レセプトデータを整備することが最低限必要であろう。また、レセプトデータの第三者提供を進めることで、レセプトデータを分析データとして利用する上でのノウハウの蓄積が期待できる。

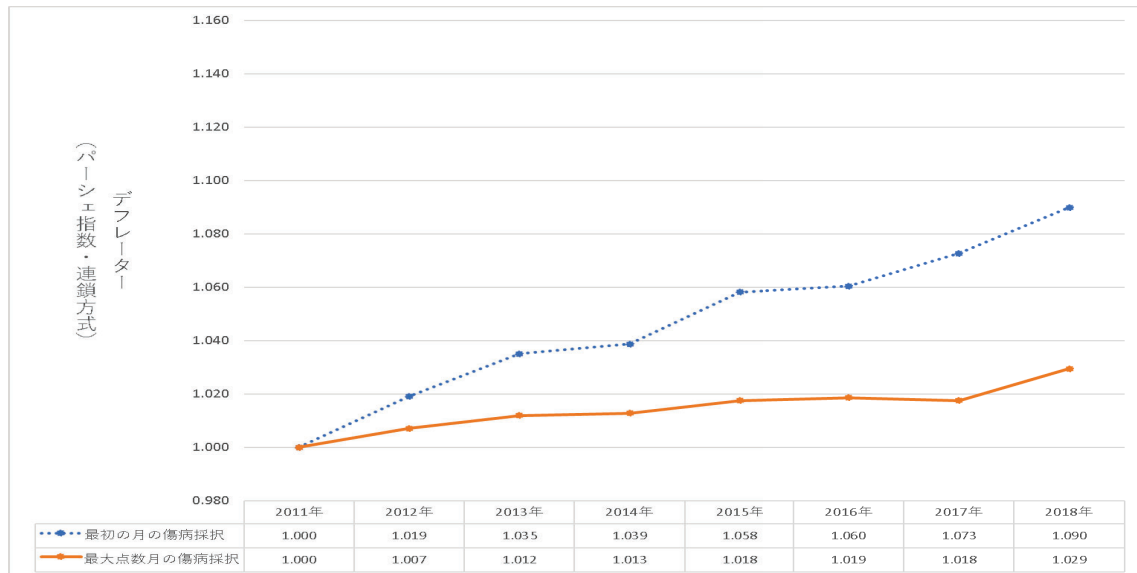
もちろん、「傷病の治療」の価格についての研究は、ICD-10コードなど国際的に同一の疾病分類が普及し、疾病ごとの価格や診療行為、薬剤などの国際比較が進み、標準治療の研究が進むなかで、重要な研究テーマである。治療の成果（アウトカム）の研究とも並行して、費用便益分析も盛んに行われている。今後とも医療の質の向上、効率化に向け、研究を進めていくべきであろう。

参考文献

- Aizcorbe, A. M. (2013). Recent Research on Disease—Based Price Indexes: Where Do We Stand?" *Survey of Current Business*, 93: 9–13
- Aizcorbe, A. M., & Highfill, T. (2015). Medical Care Expenditure Indexes for the US, 1980 - 2006. *BEA Working Papers*, 0121, Bureau of Economic Analysis.
- Eurostat. (2016). *Handbook on prices and volume measures in national accounts*. Luxembourg: Publications Office of the European Union,
- Eurostat. (2018). *Final Report of the Task Force “Price and Volume Measures for Service Activities”*.
- Schreyer, P. (2010). Towards Measuring the Volume Output of Education and Health Services: A Handbook. OECD Statistics Working Papers, vol. 2. Paris, OECD.
- 石橋尚人, 丸山雅章, 桑原進, 石井達也, 川崎暁, 西崎寿美, 村舘靖之, 大里隆也, 菊川康彬 (2020) 「医療の質の変化を反映した価格の把握手法に関する研究—推計法の検討とレセプトデータによる試算—」 ESRI Research Note No. 57
- 奥村泰之 (2017) 「JMDC Claims データベースとレセプト情報・特定健診等情報データベースを活用した臨床疫学研究の事例と留意点」
- 奥村泰之, 佐方信夫, 清水沙友里, 松居宏樹 (2017) 「ナショナルデータベースの学術利用促進に向けて：レセプトの落とし穴」 *Monthly IHEP* No. 268, 16-25.
- 久保慎一郎, 野田龍也, 西岡祐一, 明神大也, 東野恒之, 松居宏樹, 加藤源太, 今村知明 (2017) 「レセプト情報・特定健診等情報データベース (NDB) 利用促進に向けた取り組み - 患者突合 (名寄せ) の手法開発と検証 - 」 奈良県立医科大学
- 厚生労働省 (2013) 「レセプト情報等のデータ構造について」
- 国立保健医療科学院 (2013) 「レセプト分析法マニュアル」
- 杉原茂, 川渕孝一, 池本靖子, 今村育未 (2017). 「Measuring Health Care Output」 ESRI Research Note No.33
- 杉原茂, 市川恭子, 今井健太郎, 野口良平, 岡崎康平, 小池健太 (2018). 「医療の質の変化を反映した実質アウトプット・価格の把握～方法論の整理～」 ESRI Research Note No.36
- 野口良平, 市川恭子, 藤森裕美, 岡崎康平, 小池健太, 石橋尚人 (2019) 「諸外国における医療・介護の質の変化を反映した価格・実質アウトプットの把握手法～各国ヒアリングの結果～」 ESRI Research Note No. 48
- 藤森研司 (2016). 「レセプトデータベース (NDB) の現状とその活用に対する課題」 『医療と社会』, 26(1), 15-24.

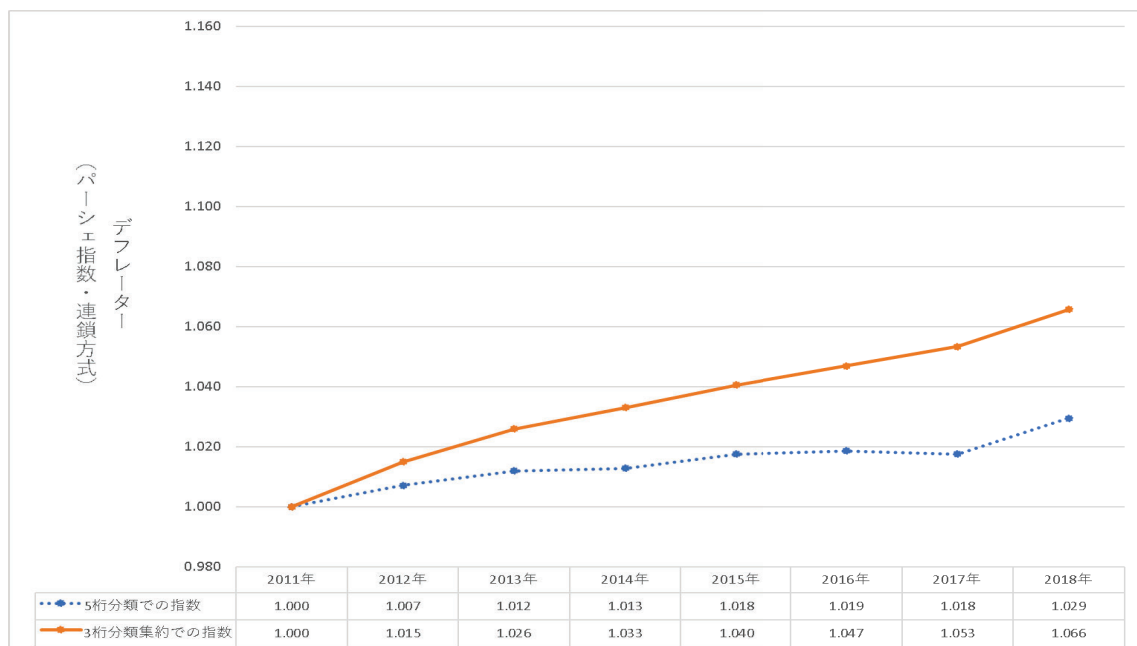
付図1 発生時期と点数による傷病の特定の違いによるデフレーター比較

傷病の特定を最初の月の主傷病としたデフレーターと最大点数の月の主傷病としたデフレーターとでは乖離が生じている。この要因の一つとして、季節性傷病によるバイアスが考えられる。



付図2 傷病分類の粒度の違い（ICD-10の5桁分類と3桁分類）によるデフレーター比較

「C34:気管支及び肺の悪性新生物」で確認したように、粒度によるデータバイアスを除いたことで、全体的な水準として拡大している。



補論A 医療給付実態調査との比較

厚生労働省においても医療保険制度の健全な運営をはかるための基礎資料として、年に1回、レセプトデータを基に、医療保険制度加入者の受診状況（診療種別、受診月、医療費、主疾病等）のデータを、制度別、年齢別、疾病分類別等様々な切り口から観察している。

図表13では、2019年度医療給付実態調査のうち、協会（一般）の年齢階級別、疾病別、入院+食事・生活療養、1人当たり医療費の集計値と今回の推計結果を比較してみた。今回の推計では、最もレセプトの点数が高い月の主傷病に患者の1年間の治療費を全て振り分けるという、非常に大胆な仮定を置いて実施したため、疾病別の医療費水準には幅をもって見る必要がある。しかし、大きな傾向としては、①新生物（腫瘍）、循環器系の疾患の1人当たり医療費が高いこと、②特に循環器系の疾患は高齢になるほど大幅に医療費が高くなること、が見てとれることから、今回の推計結果と概ね同じ方向であるといえる。

図表13 年齢階級別、疾病別、入院+食事・生活療養、1人当たり医療費（協会（一般））

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	(単位 円)
	総数	感染症及び寄生虫症	新生物<腫瘍>	血液及び造血器の疾患並びに免疫機構の障害	内分泌、栄養及び代謝疾患	精神及び行動の障害	神経系の疾患	眼及び付属器の疾患	耳及び乳突突起の疾患	循環器系の疾患	呼吸器系の疾患	消化器系の疾患	皮膚及び皮下組織の疾患	筋骨格系及び結合組織の疾患	腎泌尿生殖器系の疾患	妊娠、分娩及び産じょく	周産期に発生した病態	先天奇形、変形及び染色体異常	症状、徴候及び異常所見・異常検査所見で他に分類されない	損傷、中毒及びその他の他の影響	特殊目的用コード	
総数	53,363	830	12,525	437	1,123	2,163	2,491	951	316	9,776	2,602	3,803	398	4,325	2,011	2,186	1,794	1,377	331	3,923	0	
0-4	97,583	2,447	3,950	905	1,162	186	3,926	304	933	1,401	15,163	2,658	760	3,030	1,801	33	37,096	18,192	1,569	2,067	-	
5-9	20,258	771	2,629	580	668	320	2,628	368	498	565	3,129	1,274	337	1,024	794	0	204	2,397	219	1,853	-	
10-14	19,808	435	2,458	367	803	1,805	2,445	229	305	825	1,354	1,520	220	1,711	636	11	102	1,774	173	2,637	-	
15-19	22,988	451	2,307	425	442	2,284	2,693	209	173	926	1,755	1,430	222	1,985	536	322	50	1,052	160	5,567	-	
20-24	19,772	600	1,844	294	411	2,314	2,121	186	148	858	1,652	1,581	211	897	640	2,441	55	556	171	2,793	0	
25-29	25,128	544	2,597	357	444	2,489	1,970	183	151	1,156	1,316	1,915	220	930	1,050	7,056	99	433	183	2,034	-	
30-34	30,237	458	3,699	385	552	2,531	1,811	224	168	1,768	1,207	2,194	221	1,146	1,404	9,925	55	340	182	1,926	-	
35-39	30,656	421	5,483	302	635	2,176	1,434	272	210	2,848	1,135	2,511	229	1,474	1,578	7,463	13	275	166	1,989	-	
40-44	31,760	455	8,225	231	752	1,784	1,413	394	219	5,068	1,130	2,857	296	2,051	1,818	2,255	4	267	158	2,371	-	
45-49	40,181	542	11,805	265	961	1,789	1,574	604	232	8,739	1,199	3,611	371	2,967	2,167	100	-	230	178	2,811	-	
50-54	54,957	695	15,436	363	1,330	2,161	2,073	1,061	284	13,330	1,481	4,683	467	4,959	2,291	8	3	234	258	3,844	-	
55-59	76,445	973	21,927	414	1,705	2,642	2,744	1,718	362	18,810	2,047	5,995	531	7,911	2,676	0	-	240	338	5,403	-	
60-64	104,859	1,301	31,383	601	2,112	2,917	35,883	2,499	441	26,182	3,018	7,613	643	11,213	3,565	0	-	234	477	7,075	-	
65-69	143,901	1,703	44,063	790	2,836	3,531	4,623	3,656	527	36,947	4,611	9,711	734	14,802	4,808	-	-	218	667	9,675	-	
70-75	204,836	2,731	58,403	1,136	4,244	4,359	7,254	5,571	648	54,633	7,869	12,914	1,084	20,328	7,416	-	-	191	1,053	15,003	-	

注) 1. 「特殊目的用コード」は重症急性呼吸器症候群（SARS）及びその他の疾病を表す。
 2. 疾病コードに関する記載が一部なかったため、総数の数値が表2-6-2の「入院+食事・生活療養」の数値と一致するように補正している。

出典) 医療給付実態調査（2019年度）表2-6-3

補論B 電子レセプトの課題

電子レセプトは保険診療の請求のためのものであり、分析のために設計されたものではない。一方、これほどに電子レセプトが普及し、一定程度の分析が可能となった現在では、単に請求用のデータに終わるのではなく、多面的な分析により我が国の医療提供体制の検討や臨床疫学的な研究に用いられるべきであろう。以下では、奥村 (2017) や奥村ら (2017) などを参考に、電子レセプトを様々な研究に活用する上で検討すべき 9 つの課題について整理した。

課題1. 入退院日の不確実性

レセプトには入院日と退院日が記録されているが、正確な在院日数や再入院率を算出することは容易ではない。

DPC レセプト	医科入院レセプト
<ul style="list-style-type: none"> ・DPC 算出対象外の病棟から DPC 算出対象の病棟に転棟した場合、「転棟日」がレセプトの「入院日」に記録される。 ・DPC 算出対象の病棟に入院していても特定入院期間を超えた場合は「特定入院期間の最終日」が「退院日」に記録される。 ・同一傷病で 7 日以内に再入院した場合は一連の入院とみなした「入院日」「退院日」が記録される。 	<ul style="list-style-type: none"> ・「入金基本料の起算日」が医科入院レセプトの「入院日」に記録される。 ・退院日から原則 3 か月未満に同一傷病で同一医療機関に再入院した場合、「前回の入院日」が記録される。 ・「退院日」は文字データとして記録されているため NDB では使用できない。

課題 2. 傷病情報が DPC と医科入院で不一致

傷病情報が DPC と医科入院では一致しないため、傷病別の入院患者数を算出する際は注意が必要である。※今回の分析では、医科入院、医科入院外レセプトの傷病名コードに対応する ICD-10 コードを付与するための作業を行った。¹⁶

DPC レセプト	医科入院レセプト
<ul style="list-style-type: none"> ・傷病名区分により「医療資源を最も投入した傷病名」「副傷病名」「主傷病名」「入院の契機となった傷病名」「医療資源を2番目に投入した傷病名」「入院時併存傷病名」「入院後発症傷病名」の7つに分類される。 ・傷病名コードと ICD-10 コードが両方存在しどちらを使用するかという意思決定が必要である。 	<ul style="list-style-type: none"> ・傷病名区分コードがない。 ・「医療資源を最も投入した傷病名」の医科レセプトからの類推は難しい。

参考) 傷病情報に関する DPC と医科入院の相違

項目	DPC レセプト	医科入院レセプト
傷病名コード	○	○
ICD-10 コード	○	×
傷病名区分コード	○	×
診療開始日	×	○
主傷病コード	×	○

¹⁶ 厚生労働省が運営する診療情報提供サービスの HP、または社会保険診療報酬支払基金の HP から「傷病名マスター」と「旧傷病名管理ファイル」をダウンロードして行う。

課題 3. 医科入院外の治療継続に関する不確実性

医科入院外レセプトではすでに治癒している傷病名の記載が残っているケースが多く、真に治療が継続しているか不明である。

医科入院外レセプトの問題点	事例
<ul style="list-style-type: none"> ・すでに治癒している傷病名の記載が残っているケースも多い。 ・レセプトから傷病別の外来患者数の正確な算出は難しい。 	<ul style="list-style-type: none"> ・医療経済研究機構の研究グループの検証によると、インフルエンザ（JCD-10 コード：J10）の患者 31 万人を 1 年間追跡した結果 <ul style="list-style-type: none"> 1 月：診療開始日が 1 か月を超える人の月別割合 … 0.4% (MIN) 8 月：診療開始日が 1 か月を超える人の月別割合 … 93.8% (MAX)

課題 4. 診療科情報の欠測

レセプトには診療科情報が記録されているものの、記入率が低いなど、診療科別の分析は行える状況にはない。

- ・レセプトにおいて診療科名の記録は必須ではない
- ・診療科名の記入率は 50%程度
- ・各地方厚生局が保有している情報から医療機関マスターを作成して、NDB に診療科情報を付与するよう利用申出することも可能ではあるが非現実的
- ・医療機関が地方厚生局に届け出ている診療科名が統一されていないうえに実際の診療科名と一致する保証がない

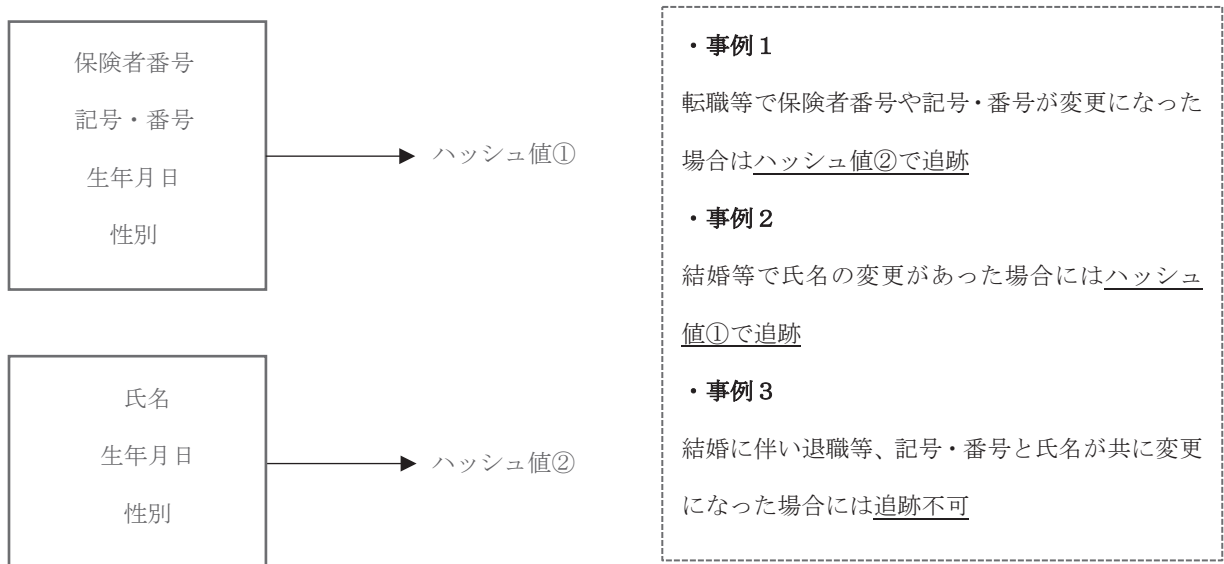
課題 5. 保険離脱による患者の追跡は不可能

NDB には患者の追跡可能性を確保するために 2 種類の ID が格納されているが、保険離脱により追跡が不可能である。※保険離脱については、失職、新卒、定年退職、75 歳の誕生日における後期高齢者医療制度への移行の際に生じうる。年齢層や追跡期間を制限するなど注意が必要である。

ID1	ID2
<ul style="list-style-type: none"> ・「保険者番号+被保険者の記号・番号+生年月日+性別」から生成 ・保険離脱の影響を受ける ・同性の双子を区別できない ・現状 ID1 を使用する事例が多い 	<ul style="list-style-type: none"> ・「氏名+生年月日+性別」から生成 ・医療機関間で生じる氏名の表記揺れに影響を受ける ・生年月日が同一で同性の同姓同名を区別できない

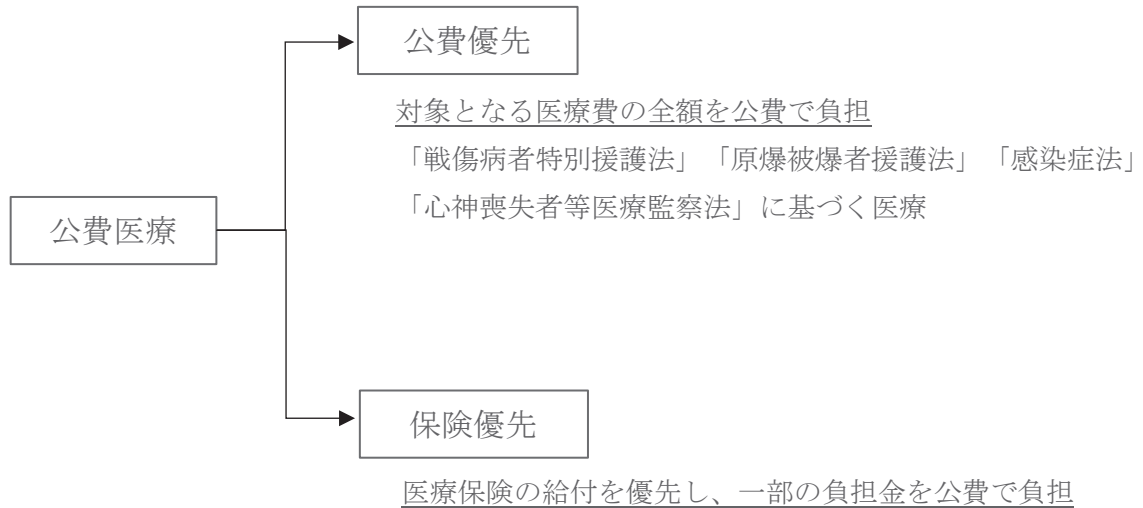
○分析目的に応じてハッシュ値の使い分けが必要

ハッシュ値とは固定長の疑似乱数。同一の入力に対しては同じ値が返ってくる



課題6. レセプトの欠測問題

NDB において「公費優先」「生活保護受給者」に関するレセプト情報の取得は困難



- ・ NDB には公費医療のレセプトは格納されていない
- ・ 生活保護受給者の約 98%は被用者保険に加入していない

課題7. 自由診療のレセプトの欠測

NDB には自由診療のレセプトは格納されていない

- ・ 産婦人科や先進医療など自由診療と保険診療が混在する症例には注意が必要

[事例]

自然分娩 … 自由診療 → 入院料は NDB に含まれない

帝王切開 … 保険診療 → 入院料は NDB に含まれる

課題 8. 医療機関情報の変更の不確実性

同一の診療所に継続的に受診していることをレセプトから同定することは難しい

- ・レセプトには医療機関情報が記録されており、都道府県コード（2桁）＋医療機関コード（7桁）でユニークなコードとなっている
- ・医療機関の移転や統合、運営母体の変更がある場合は医療機関のコードも変更になるため、同じ医療機関に継続的に受診していても同一の医療機関コードが記録されているとは限らない
- ・一定以上の規模の病院は絶対数が少ないため変更を考慮したマスターを作成すれば対処が可能であるが、診療所の場合は変更の絶対数が多いため変更を考慮したマスターを作成することそのものが事実上不可能である

課題 9. 診療行為の点数および回数情報の復元

レセプトの記載だけでは診療行為の実際の点数と回数がわからないため、医療診療行為マスターから点数を取得する必要がある

レセプト上の記載			電子レセプト上の記録		
再診料			...	112007410,	
地域医療貢献加算			...	112015670,	
明細書発行体制加算 73点 × 2			...	112015770,	73, 2
Excel 上でファイルを展開した場合			実際の点数と回数（分析時に必要な形）		
コード（診療行為）	点数	回数	コード（診療行為）	点数	回数
112007410			112007410	69	2
112015670			112015670	3	2
112015770	73	2	112015770	1	2
			※医科診療行為マスターから点数を取得する必要あり		

医科診療行為マスターの中身

s_ALL20190408.csv - Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	0	S	111000110	3	初診料	7	シヨシリョウ	0	0		3	282	0	0	110
2	0	S	111000370	9	乳幼児加算	10	ニョコガジ カ	0	0		3	75	0	1	110
3	0	S	111000470	10	乳幼児育児	20	ニョコガジ イ	0	0		3	130	2	1	130
4	0	S	111000570	9	時間外加算	10	ジカガ イカ	0	0		3	85	0	0	110
5	0	S	111000670	8	休日加算 (9	キョジ カカ	0	0		3	250	0	0	110
6	0	S	111000770	8	深夜加算 (6	シヤカ	0	0		3	480	0	0	110
7	0	S	111000870	15	時間外特例	20	ジカガ イカ	0	0		3	230	0	0	110
8	0	S	111010770	28	乳幼児時間	20	ニョコガジ ジ	0	0		3	230	2	1	130
9	0	S	111011570	14	乳幼児夜間	13	ニョコガジ ヤ	0	0		3	200	0	1	110
10	0	S	111011670	14	乳幼児休日	16	ニョコガジ キ	0	0		3	365	0	1	110
11	0	S	111011770	14	乳幼児深夜	13	ニョコガジ シ	0	0		3	695	0	1	110
12	0	S	111011810	18	初診料 (同	7	シヨシリョウ	0	0		3	141	0	0	110
13	0	S	111011970	12	乳幼児時間	17	ニョコガジ ジ	0	0		3	200	0	1	110
14	0	S	111012070	11	乳幼児休日	16	ニョコガジ キ	0	0		3	365	0	1	110
15	0	S	111012170	11	乳幼児深夜	13	ニョコガジ シ	0	0		3	695	0	1	110
16	0	S	111012270	18	乳幼児時間	20	ニョコガジ ジ	0	0		3	345	0	1	110
17	0	S	111012470	12	夜間・早朝	13	ヤンツョウカ	0	0		3	50	0	0	110
18	0	S	111012510	17	初診料 (文	7	シヨシリョウ	0	0		3	209	0	0	110
19	0	S	111012610	31	初診料 (同	7	シヨシリョウ	0	0		3	104	0	0	110
20	0	S	111012710	8	特定要結率	17	トケイケツ	0	0		3	209	0	0	110

※ヘッダー情報は含まれていない。診療報酬情報提供サービスのファイルダウンロードページの「データ内容に関する説明はこちら」から仕様を確認することができる。

医科診療行為マスターのヘッダー情報 (PDF ファイルでの公開のみ)

ファイルレイアウト(医科診療行為マスター)1/3

No.	データ項目名	属性	最長桁数	最長バイト数	備考
1	変更区分	数字型	固定長	1	1
2	マスター種別	文字型	固定長	1	1 S:固定
3	診療行為コード	数字型	固定長	9	9
	診療行為省略名称				
4	省略漢字有効桁数	数字型	可変長	2	2
5	省略漢字名称	文字型	可変長	32	64
6	省略カナ有効桁数	数字型	可変長	2	2
7	省略カナ名称	文字型	可変長	20	20
8	データ規格コード	数字型	可変長	3	3
	データ規格名				
9	漢字有効桁数	数字型	固定長	1	1
10	漢字名称	文字型	可変長	6	12
	新又は現点数				
11	点数識別	数字型	固定長	1	1
12	新又は現点数	数字型	可変長	10	10 整数部「7桁」、小数点「1桁」及び小数部「2桁」の組合せで設定
13	入外適用区分	数字型	固定長	1	1
14	後期高齢者医療適用区分	数字型	固定長	1	1
15	点数欄集計先識別(入院外)	数字型	可変長	3	3
16	包括対象検査	数字型	可変長	2	2
17	予備	数字型	固定長	1	1 未使用(全て0)
18	DPC適用区分	数字型	固定長	1	1
19	病院・診療所区分	数字型	固定長	1	1
20	画像等手術支援加算	数字型	固定長	1	1

分析用データセット作成に必要な作業

- ・各時点（平成 24, 26, 28, 30 年改訂）の医科診療行為マスターを取得し、整備する
- ・「診療行為コード」「（診療行為省略名称）漢字名称」「新又は現点数」のカラムを抽出し、「改定年」のカラムを追加する

改定年	診療行為コード	診療行為名	点数
2012	111000110	初診	270
2014	111000110	初診	282
2016	111000110	初診	282
2018	111000110	初診料	282
2012	111000370	初診（乳幼児）加算	75
2014	111000370	初診（乳幼児）加算	75
2016	111000370	初診（乳幼児）加算	75
2018	111000370	乳幼児加算（初診）	75
2012	111000470	乳幼児育児栄養指導料	130
2014	111000470	乳幼児育児栄養指導料	130
2016	111000470	乳幼児育児栄養指導料	130
2018	111000470	乳幼児育児栄養指導料	130
2012	111000570	初診（時間外）加算	85
2014	111000570	初診（時間外）加算	85
2016	111000570	初診（時間外）加算	85
2018	111000570	時間外加算（初診）	85