

より良い保健・医療・介護制度等に貢献する 質の高い医療経済学研究のための データベース構築に向けた要望書

医療経済学会 学術推進委員会
質の高い医療経済学的研究のためのデータベース検討事業
ワーキンググループ (WG)

はじめに

我が国の医療・介護レセプトは、詳細な医療・介護サービスの内容がわかる世界にも稀なデータベースとしてレセプト情報・特定健診等情報データベース (NDB)、診療群分類包括評価 (DPC)、国保データベース (KDB)、介護 DB など、研究に用いられる環境が整備されてきた。現在利用しうるデータを駆使して、医療経済学的な研究成果が生まれ、政策立案者など社会への還元がなされるようになってきている。

一方、本来、報酬請求の目的のレセプトデータであるため多くの限界があり、諸外国と比した時、我が国の現状は、政策立案や評価に役立つ、質の高い医療経済学研究が出来るデータ環境とは言いがたい。今後、我が国の保健・医療・介護等の社会保障政策の立案や遂行状況のモニタリング、見直しに有用なニーズの分布やメカニズムを明らかにし、政策・プログラムのプロセス・アウトカム評価などができる総合的なデータ環境の整備が必要である。多くの制約がある現状を改善しようと、デジタル庁も設置され、個人情報保護に配慮しつつ、社会にとって有用なデータベースの利活用環境づくりに向け、関係省庁や機関における論議がなされている。

そこで、科学的エビデンスの社会還元を通じて健康社会を実現することを目的に、保健・医療・介護等の社会保障制度・業務の継続維持・発展に寄与する医療経済学研究を可能とするとともに、国・自治体・医療介護等の現場・各種事業者・一般市民・患者等の全てのステークホルダーに歓迎されるデータベースづくりに向けて、医療経済学会学術推進委員会として、現在のデータベースやデータ環境の課題と、あるべき姿について検討した。

本要望書は、論議の中で得られた関係者のコンセンサス (合意) をまとめたものである。

現状のデータベースの課題

1. 背景要因やアウトカム変数が乏しい

心身機能や疾病・死亡などの健康アウトカムには、予防・診断・治療・リハビリテーション・介護のサービスだけでなく、患者や集団・地域が持つ背景・属性要因が影響している。例えば、低所得者集団や一部の地域に不健康が多い健康格差が日本にもあることから、「健康格差の縮小」は、「健康日本21」の基本的方向として示されている。となると、社会経済的要因や世帯状況などの背景・属性要因がわかるデータが結合され、それらを考慮した分析が必要となる。また、それらの代理変数としては、加入している健康保険の種類（国民健康保険、健康保険組合、協会けんぽなど）がわかる被保険者台帳情報も利用しうる。しかし、現状では、これらのデータの結合は容易でない。

現状では、健康アウトカム変数になるメンタルヘルス等の健康状態・機能状態や、外泊・退院等に関するデータが乏しい。死亡がわかるデータはレセプトデータには含まれておらず、死亡小票データは結合可能になったが、実際に結合するのは相当の経験者でなければ容易でないのが実情である。

一方、海外では、これらのデータも結合されたデータベースが整備されている国が多数に上る。ただし日本にも、世界に誇る差別化要素が存在する。介護保険に係る電子データ（要介護認定に係る心身情報データ、多種多様な介護サービスに係る介護レセプトデータ）が長期間蓄積されており、これは他国にはない。上述の高齢者状態を把握する上での医療と介護の両データを統合しての様々な研究が日本の独壇場として展開できる可能性がある。

2. サービス供給者データが乏しい

サービスの内容決定とその質に対して、供給者の与える影響は非常に大きい。しかしながら、現状では、サービス供給者のデータは、医療機関レベル、医師レベル共に非常に乏しい。

各地方厚生局では医療機関コードを管理しており、医療機関ごとの情報が公開されている。ところが、NDBにおいては医療機関コードが公開されていない。一方、医療機関ベースで収集されているレセプトデータベースについては、医療機関が識別出来ることもあるが、複数の医療機関を受診する場合の情報は得られない。したがって、現状では複数の医療機関を受診を医療機関の情報と共に分析する事は非常に難しい。

科学的な根拠に基づいて診療ガイドラインが整備されている現在であって

も、実際の治療選択は個々の患者の状況に基づいて、医師の選択の自主性に基づいて決定されている。このような個々の医師の治療内容決定が、医療の質に影響を与えうることは、過去の研究でも数多く指摘されている。

医師ごとの分析を行うためには、医療機関の識別だけでなく、医師の識別が可能な情報が必要となる。たとえば、アメリカの公的医療であるメディケアとメディケイドを管理するCMS(Centers for Medicare & Medicaid Services)では、ダウンロード可能なデータベースとして全国の医師の情報を公開している。

生産年齢人口が激減する一方で需要が増加する医療介護分野において各職種のマンパワーを効率的に配分することが不可欠であるが、医師のみならず様々な専門職についても就労状況や賃金といった情報がデータベース化されたものは存在しない。

3. 縦断データが入手困難

一時点の横断データでは、因果の向きを特定できない。因果の向きを特定するには、時間的前後関係を確認する必要があるため、同一の個人を追跡して健康状態の変化を捉えた縦断データが必要である。観察開始の一時点のベースラインデータと、数年後あるいは数年間の健康アウトカム変数を含むデータとは、別ファイルになっている。そのため、これらのデータを結合しなければ縦断データは得られない。しかし、個人情報保護のため、あるいは異なる社会保険制度では異なる被保険者番号が用いられているため、同一個人の別ファイルを結合することは、理論上は可能でも、実際に結合するのは容易でない。

例えば、同じ「被保険者番号」という変数名であっても、医療保険(協会けんぽ・健保組合・国民健康保険・後期高齢者医療保険制度)と介護保険の各保険制度で、それぞれで異なる「被保険者番号」が割り当てられており、「被保険者番号」だけでは、これらを結合することはできない現状にある。

4. 他データとの連結・リンケージに制約

現状の医療・介護系データの問題点として、他のデータとの連結・リンケージに関する様々な制約がかかる点があげられる。NDB・KDBといったレセプトデータについては、その性質上個人に関する社会・経済的な情報がほとんど含まれていないが、これらを補完するために他の情報源と連結することは禁じられており、そのことは利用規約に明記されている。NDB・KDB以外の、統計法の目

的外申請によって提供されるデータに関しては、連結が明示的に禁じられているわけではないが、連結のための情報が存在しないなどの理由で事実上制約がかかる場合も多い。

近年では、NDB・KDBが他のデータと連結済みの状態で提供されるようになるなど、状況は一部で改善しつつあるが、諸外国と比較すると見劣りする部分がある。例えばデンマークなどでは、公的業務情報のほとんどが連結された状態で研究者に提供されるため、個人に関する情報が教育・就業・医療など多様な社会活動をまたぐ形で分析可能になっている。よりデータに即した政策提言を行うためには、多様なデータのリンケージは、将来的に不可欠の課題と言えるだろう。

5. データの利用環境に関する制約

近年、学部・大学院レベルでのデータサイエンス教育が活発に行われ、データ分析についての能力を持つものは増えている。社会的に重要なリサーチクエストとデータ分析を行う能力があったとしても、データ分析についての準備に労力を割かざるを得ないことが少なくない。

まず、それぞれのデータベースに何がどのような形で格納されているかの情報の整理について、標準化された内容でまとまった形で公開されることは少ない。また、地方厚生局の医療機関の情報など、データが公開されていても地域によってフォーマットが異なり収集に手間取る場合がある。

徐々に利用可能なデータベースが増えてきた現状でも、クラウド上でのデータベース分析の構築は研究者に任されていることが多く、そもそもクラウド上での分析が許されない場合もある。そのため、分析環境の構築自体がデータ分析のボトルネックとなっている。

あるべきデータベース

近年では、医療・介護に関する大規模行政管理情報についてその突合が進んでいる。NDB・介護DB・DPCといったレセプトデータベースに加え、健診・介護情報（LIFEデータ）、難病DB、小児慢性特定疾病DB、がん登録DBなどの突合はほぼ完了している。これは、保健・医療・介護の状況の全容を生涯に通じて観察可能なデータベースが整備されつつあることを示している。こうしたデータベース整備に対する行政関係者の不断の努力には敬意を表したい。

上で述べた現状での課題を基に、今後の医療・介護データに関するあるべき

姿を列挙する。

1. 背景要因やアウトカム変数の充実
サービス需要者のデータとして充実を目指すべきもの
 - 社会経済要因（所得・教育・就業状況等）
 - メンタルヘルス関連指標・機能状態
 - 外泊及び退院等の詳しい転帰
 - 疾患を持つものや受診者以外の人々を分析するための被保険者の情報

2. サービス供給者データの充実
 - 医療・介護施設の所在地や機能
 - 診療に関わった医師に関する情報
 - 医師だけでなく、医療・介護に関わる専門職の人材資源・働き方（就労時間、報酬等）を含めた情報
 - 高齢者の保健事業と介護予防の一体的実施事業及び介護予防サービスに関する実施情報

3. 縦断データが容易に入手可能となること
 - 同一意思決定主体の追跡可能なデータの整備
 - 同一主体の異時点でのデータを容易に結合可能にする ID 等の整備

4. 他データとの連結に関する制約の軽減
 - 個々の意思決定主体（個人・専門職個人・医療介護施設）、地域（都道府県・二次医療圏・市町村）、保険者の情報をそれぞれ結合可能にする ID 等の整備
 - 行政管理情報と官庁統計調査との連結

5. データの利用環境に関する制約の軽減
 - 同一変数のフォーマットの統一
 - 個々のデータベースの内容と変数に関する情報の整理と公開
 - クラウド上で分析可能な環境の整備

当然のことながら、短時間で上記のあるべき姿が実現されることはない。特にデータ相互の連結に関しては、個別データの収集目的や管理者など法律レベルでの制約がある場合も多い。個々の意思決定主体の詳細な情報を得ることは、個別具体性への興味からでは決してない。その目的は、個々の主体を識別

した分析によって、各主体それぞれの特性を調整し厳密な分析をすること、及び個別主体の集団的特性の影響を分析することである。

こうした分析を蓄積することで、社会保険料及び税、医療・介護従事者の労働力といった貴重な資源を有効に活用し、国民全体の健康やウェルビーイングの改善することが医療経済研究の目的であり、そのための関係者との協働や国民理解の醸成に努力したいと考える。

補足：データ連結へのそこから生じる問題点

以下ではデータ連結への制約とそこから生じる問題点を具体的に提示する。

まず一つ目の問題として、一般に開示されているような情報とであっても、レセプトデータが連結することができないことがあげられる。例えば、県や市町村レベルの人口といった基礎統計との連結も禁じられている。このことが原因として起こった問題として、市町村における COVID-19 の検査陽性者数は日々報道されていた公開情報であるが、これらがレセプトデータと結合できず、感染拡大と医療・介護費の関係を分析することができなかったという事例がある。

第二の問題としては、本来介護・医療サービスの提供時には明らかになっているにもかかわらず、記録上の問題でそもそもデータとして整備されていない情報があることがあげられる、こうした情報の例としては、サービス提供機関・提供者の ID があげられる。医療施設や医師の ID については NDB に含まれるべき情報だろう。さらに、看護師、保健師、リハビリテーション専門職や介護従事者などについては、機関ごとの従事者数などは把握できていても、個人レベルの ID については、そもそも国が保有するデータベースが存在していない。こうした提供サイドについての情報整備が進み、さらに個別従事者の賃金や労働時間、教育水準や就労年数などの情報が蓄積されて他のデータとリンクが可能になれば、医療・介護分野における労働市場の分析が可能になるだろう。

第三の問題として、レセプトデータと官庁統計個票とのリンクができないうことがあげられる。例えば厚生労働省による中高年縦断調査、国民生活基礎調査などの官庁統計調査は、家族構成、経済状況といった情報を広範に含んだデータとして様々な研究の情報源となってきた。これらと NDB・KDB が連結できれば、社会・経済状況によってどのような医療・介護支出の違いがあらわれるのか、といった要因分析が可能となり、より効果的な医療・介護提供体制のために何が必要かを考えるヒントが得られるだろう。また、文部科学省による「全国学力・学習状況調査」がレセプトデータと連結できれば、子供の健康が認知・非認知能力・学力などの人的資本にどのように影響するかが分析可能になる。

こうしたリンクに関しては、官庁統計のようなサーベイデータとレセプトのような行政管理情報の間には、法律上の制約のため、本人同意がないと結合できないという点が困難をもたらしている。個人情報保護と、公益性が高い研究に対してデータを提供していく方策について、今後議論が進んでいくことが望ましい。

四つ目の問題として、他の行政情報との連結ができないということがあげられる。他の行政情報としてはいくつかの候補があるが、まずは介護保険・医療保険の保険者台帳との連結ができないという問題がある。この結果、保険対象サービスの利用がゼロであるケースについて、保険給付がないのか、それとも調査対象者が被保険者でないのかが識別できていない。このことは、一人当たりの医療・介護費について考える際に、分母が定まらず精度の高い推定ができないといった困難をもたらしている。

レセプトデータとリンケージされれば学術研究の可能性を拡大させうる行政情報は、他にも数多い。たとえば住民基本台帳との連結ができれば、家族構成と医療・介護の関係について様々な分析が可能になる。税務関連情報との連結ができれば、所得・資産との関係も分析可能になり、医療・介護費に関する制度設計に関しても様々な研究が進むものと期待できる。こうした連結に際しては、社会保障・税番号（マイナンバー）を識別情報として用いるのが効率的であろう。ただし、個人情報保護に重点を置く「行政機関個人情報保護法」のもとで集められている行政情報を結合させることは現状では難しく、官庁統計個票とのリンケージと同様に、個人情報保護と学術研究との関連について活発な議論が待たれる。

他にも、特定健診情報はNDBに格納されているが、学校健診情報のように現在では提供されていない情報もあり、こうした情報が利用可能な価値になり、より充実したデータが提供されれば、学術研究にとって大きな刺激となるだろう。

質の高い医療経済学的研究のためのデータベース検討事業 ワーキンググループ 構成員名簿

ワーキンググループ メンバー

飯塚 敏晃	東京大学大学院経済学研究科 教授
泉田 信行	国立社会保障・人口問題研究所 部長
岸田 研作	岡山大学大学院 社会文化科学学域 教授
後藤 励	慶應義塾大学大学院 経営管理研究科 教授
近藤 克則	千葉大学予防医学センター 特任教授
菅原 慎矢	東京理科大学 経営学部 准教授
高久 玲音	一橋大学大学院経済学研究科 教授
野口 晴子	早稲田大学 政治経済学術院 教授
野田 龍也	奈良県立医科大学 医学部医学科公衆衛生学 准教授
富 蓉	早稲田大学 商学部 准教授

オブザーバー

橋本 英樹	東京大学大学院医学系研究科 教授
伏見 清秀	東京医科歯科大学大学院医療政策情報学分野 教授