
公募シンポジウム

公募シンポジウム7

レセプト情報等データベース（NDB）の利活用について：これまでの動向の整理

2018年11月25日(日) 09:00 ～ 11:00 A会場 (3F メインホール)

[4-A-1-1] レセプト情報等データベース（NDB）の利活用について：これまでの動向の整理

○加藤 源太（京都大学医学部附属病院 診療報酬センター）

レセプト情報等データベース（NDB）は、2011年の第三者提供開始以降、データ提供承諾件数は150件を超えており、徐々に提供実績が拡大しているところである。また、レセプト情報等オンサイトリサーチセンター（東京）および（京都）（以後「オンサイトセンター」とする）では2015年度より試行的利用が開始されており、2018年度には第三者提供に向けた実証実験が予定されている。一方で、次世代医療基盤法の施行にみられるように、保健医療データの有効な利活用は引き続き重要な社会的課題として位置づけられている。本シンポジウムは、レセプトデータの利用のみならず運用にも関与している日本および海外の関係者から最新の動向について講演いただくとともに、パネルディスカッションを通じて、現存する課題とその克服について議論を深め、レセプトデータ活用環境の充実化の一助とすることを目的として企画したものである。当演題では、日本におけるこれまでのNDBの利活用について紹介するとともに、オンサイトセンターでの取り組みや海外の動向等についての概要を整理し報告する予定である。

レセプト情報等データベース(NDB)の利活用について

- これまでの動向の整理 -

加藤源太^{*1}^{*1} 京都大学医学部附属病院 診療報酬センター

About the effective use of the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) - Summary of the trends until now -

Genta KATO ^{*1}^{*1} Solutions Center for Health Insurance Claims, Kyoto University Hospital

The National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) has been open to researchers since 2011, used for more than 150 research cases and becoming popular year by year. In a similar way, NDB On-Site Research Center (Tokyo/Kyoto) was tentatively open since 2015, and would experimentally be open to common users in 2018. At the same time, the effective use of healthcare data is still considered as a vitally important social challenges. In this paper, we discuss the trends concerning not only the effective use of NDB but also other countries' healthcare big data by inviting front persons who are engaged in data managing works, data providing works etc. As an introduction, I introduce the history of the use of NDB including the improvements concerning usability, and summarize the various effort at NDB On-Site Research Center by citing the documents of the board council meeting of experts. After that, the trends of the healthcare data provision for researchers or policy makers at other countries would be discussed.

Keywords: National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB), NDB On-Site Research Center, Secondary Data Use

1. 目的

レセプト情報・特定健診等情報データベース(NDB)は、悉皆性の高いヘルスケアデータとして注目を集めているデータであり、平成 23 年からは研究目的での第三者提供が行われているが、データ分析にあたってはレセプト情報及びその分析に対する様々な知識が前提として求められ、これがデータ分析の推進・活性化を阻む要因の一つとされている。一方で、第三者提供開始より現在に至るまで厚生労働省では NDB の利用者向けに様々な利便性向上の施策を打ち出しているが、それらは利用者において必ずしも十分には情報が行き渡っておらず、これらの情報を把握している利用者としてでない利用者として、NDB の利用に関するリテラシーに差が生じている恐れがある。また、こうした利用者支援の動向が海外におけるヘルスケアデータの利用動向と比較して、どの程度充実しているのか否かが十分に議論されているわけではない。本稿はこうした現況を踏まえて、NDB の第三者利用に関して利用者向けに打ち出された昨今の利便性向上策の動向について整理を行いつつ、海外の動向を踏まえながら NDB の利用動向を相対化することを目的とするものである。

2. 方法

NDB を利用する際に最も参考になる情報は、「レセプト情報等の提供に関する有識者会議」(以下「有識者会議」)における資料ならびに議論、および関連ガイドラインであり、これらは厚生労働省ホームページにおいて一か所のサイトにまとめられている¹⁾。本稿ではここ数年の有識者会議の議論を中心に、NDB の第三者利用の変遷や、これまでに改善されてきた事項を再確認する。あわせて、資料を引きながら、海外のヘルスケアデータの二次利用の動向を整理して、日本における動向を相対的に評価する。

3. NDB における利便性改善の変遷

平成 23 年より、研究者等の利用に対して NDB の第三者提

供が開始されている。その後 7 年の経過において、NDB の利用を取り巻く環境の変化がどう生じてきたかについて、以下に個々の事例に基づいて紹介を行っていく。

3.1 NDB 第三者提供の件数の推移

NDB データの第三者提供が開始された平成 23 年度は、総申出件数 43 件に対して承諾件数が 6 件と低調であった。平成 24 年度には状況はやや改善したもの、総申出件数 23 件に対して承諾件数は 13 件という状況であった。しかし、平成 25 年度には承諾件数は 22 件と、20 件を超えるようになり、その後平成 26 年度の 20 件を経て、平成 27 年度には 34 件、平成 28 年度には 42 件、平成 29 年度にも 41 件と、承諾件数が徐々に増加してきている。また、有識者会議での審査においてデータ提供が承諾される割合も、平成 25 年度以降は 90%を超えており、申出に際しての事前審査の手続き等が徐々に定式化・円滑化してきていることが示唆される²⁾。この間の経緯として、NDB データのサーバ増強などといったハード面での充実化、有識者会議における審査に特化した分科会の設置(平成 25 年以降)や事務手続きの一部外部委託化といったソフト面での改善が、承諾件数の増加に貢献していると考えられる。

3.2 NDB オープンデータ

NDB オープンデータは、NDB データの実操作が困難な者にあっても気軽にデータを有効に活用できるようにすることを目的に典型的かつ一般的な観点から NDB データを集計し公表してはどうか、という議論を踏まえて作成されたデータである³⁾。具体的には医科診療報酬点数表項目、歯科傷病、特定健診集計結果、薬剤データについて、主に単純集計に結果が得られる項目について、性・年齢階級別、及び都道府県別の集計結果が公表されている。これらオープンデータを適切に活用することが出来れば、例えば非常に基本的な事項に関する集計結果などは個別に NDB データの提供依頼申

出を行わずとも知見を得ることが可能となった。また、民間を含む利用者において集計項目等に関する要望を自由に提言することができ、これら要望に対しては対応可能なものから次回以降のオープンデータにて対応する体制が整えられている。たとえば、第 1 回 NDB オープンデータが公表された平成 28 年 10 月以降、利用者からのフィードバックに基づき、平成 29 年に公表された第 2 回 NDB オープンデータでは、医科診療行為における特定の加算項目の集計、薬剤情報における薬効分類別の上位 100 位まで(第 1 回では 30 位まで)の公表範囲の拡大、等の反映がなされた。

3.3 申出範囲の追加(市区町村)、および市区町村単位で結果を公表する際の公表基準の緩和

NDB データの利用成果を公表に際しては、個人が特定されるリスクが高くないよう、米国 CMS における cell size suppression policy を援用して患者等の集計値が 10 未満の場合は公表してはならないようルールが定められたが、市区町村を集計単位とする場合は、集計単位が小規模となるため個人特定のリスクがさらに高まることを懸念し、集計値が 100 未満の場合は公表してはならない、というより厳しいルールが設定されていた⁴⁾。しかし、一般的な行政単位である市区町村を単位とする際に、100 未満の集計値をすべてマスクするとすると、多くの重要な集計値が表現できなくなり、集計結果を十分に活用できないことが懸念される。そこで、市区町村を集計単位とする際の「100 未満の集計値は公表しない」というルールを、現在の市区町村の人口分布等を踏まえて再検討し、

1. 人口 2,000 人未満の市区町村では、患者等の数を公表しない
2. 人口 2,000 人以上 25,000 人未満の市区町村では、患者等の数が 20 未満になる集計単位が含まれない
3. 人口 25,000 人以上の市区町村では、患者等の数が 10 未満になる集計単位が含まれない

というルールへと改められた。また、こうした基準緩和と同時に、提供依頼申出者の範囲に市区町村を加えることも併せて決定された⁵⁾。平成 29 年 5 月時点で市区町村からの申出も 3 件みられ、これら申出者による NDB データの分析結果が、地域医療の充実化に向けた施策へと展開されることが期待される。

3.4 サンプルングデータセットの充実化

サンプルングデータセットは、医科入院、医科入院外、調剤、DPC の 4 種類のレセプトについて、10 月診療分のデータに対してランダムサンプルングを施すとともに、出現頻度の少ない傷病名や診療行為等の情報、および都道府県など地域情報をマスクするなどして匿名性を高めたデータセットであり、第三者提供が開始された翌年、平成 24 年 6 月より提供が開始された。特別抽出データの場合は、最小限のデータ提供しか認められておらず、探索的な利用も禁じているのに対し、サンプルングデータセットでは匿名性を高めることで安全性を確保していることから、ある程度の探索的研究を利用者に認めている。このサンプルングデータセットは、単月データという性質上、特定の事例を時系列で追いかけて評価する研究には使用できないが、あらかじめパターン化されたデータセットゆえにデータ抽出作業等に要する時間が短く、迅速にデータを提供できるという利点や、利用者の側も入手した NDB データの紐付けや成型等の作業が比較的容易である、という利点が

あり、患者数の評価など、大まかな動向を探ることには適したデータである。こうした背景を鑑み、研究者が更に利用しやすいよう、

1. 小児および高齢者の年齢階級区分について、従来よりも精緻化させた区分を別に設ける
2. 「疑い病名」及び「主傷病」を簡単に判断できるフラグをあらかじめ導入する
3. 従来は 10 月診療分のみだったところを、受療行動や罹患率に季節変動が生じることを鑑み、「1 月診療分」「4 月診療分」「7 月診療分」のサンプルングデータセットを新たに導入する

などといった改善が図られた⁶⁾。平成 30 年 2 月時点では、研究目的でデータ提供の承諾を受けている 108 件のうちの 22 件、全体の 20%弱がサンプルングデータセットの利用を希望しており、サンプルングデータセットへの一定程度のニーズが利用者のあいだで確立しているものと思われる。

3.5 レセプトデータ・特定健診等データ突合率の改善

NDB は、「レセプトデータ」と「特定健診および特定保健指導データ」を格納したデータベースであるが、両者は別々の経路で国管理のサーバに格納されている。格納に際しては受診者の生年月日、性および被保険者証等記号・番号を用いたハッシュ ID を作成することで、匿名性を維持しつつ受診者の同一性を確保する、というデータ格納フローとなっている。ところが、一部の保険者においてレセプトデータは「全角」、特定健診等データでは「半角」で出力される場合があり、この場合は同一番号であっても異なるハッシュ ID が発出されてしまうため、突合率が保険者によって大きく異なる、という事象がみられた。これに対して、「半角/全角」「ゼロ埋めあり/ゼロ埋めなし」といった表記の揺れに対応できるよう、システム改修が行われた。具体的には、表記揺れを変換したうえで新しいハッシュ ID (ID3) を発出し、これを過去の NDB データにも遡って付与することで、レセプトデータと特定健診等データ間の同一人認識が可能になるよう対応した。これによって、平成 27 年度の特定健診等データと ID3 をさかのぼって設定した同年度のレセプトデータを突合させたところ、従来の ID では全体で 25.7%に過ぎなかった突合率が、87.6%にまで改善するとともに、「国保」「協会けんぽ」「健保組合」「共済組合」といった保険者種別ごとの突合率のばらつきも、ID3 を使用することで全ての種別で 80%以上の突合率を達成するなど、大幅な改善が認められた⁷⁾。これにより、メタボリックシンドロームに着目した個々の被保険者における医療費・医療受療動向の推移の評価を、より精緻に行えることが期待される。

3.6 オンサイトリサーチセンター

東京大学ならびに京都大学において 2015 年度より試行利用が開始されたオンサイトリサーチセンターは、その後パフォーマンス検証等を経て、現在、両施設とも学内の研究シードに対し限定的かつ試行的に門戸を開き、その運用について検証を行っている。また、厚生労働省の所管するオンサイトリサーチセンターにおいては、現在、完全な第三者向けの開放を視野に入れた試行運用が開始されているところである⁸⁾。

いずれにおいても、一定程度の成果の導出は実現しているものの、NDB についてのリテラシーを備えた施設スタッフや

専門の研究者がかなりの負担を負ったうえで初めて研究が成立しているような状況であり、NDB に対するリテラシーを十分に備えていない利用者が直ぐにオンサイトにアウトして知見を得られる状況にないことが指摘されている。サーバのスペック等、オンサイトリサーチセンターのハード面での機能充実にも増して、NDB 利用を志す者が相応の負担程度で知見が得られるような、研究支援体制の充実化が今後の大きな課題となっていくであろう。

4. 海外の動向について

NDB は大きく分けて「レセプト情報」および「特定健診情報、特定保健指導情報」より成るデータベースである。各国間での医療制度の違いもあり、海外において、日本の NDB と同一のデータ項目が採用されているデータベースは存在しない。一方で、NDB のように相対的に公的な機関により管理され、国レベルでデータ収集が行われており、かつ日本のレセプトに近い項目のデータがそろっている海外の大規模ヘルスケアデータベースはいくつか存在する。本稿では、主に「第2回医療・介護データ等の解析基盤に関する有識者会議」における公開資料等^{10) 11)}を引きながら、海外の動向について紹介する。

4.1 今回紹介の対象とする事例

本稿では、「第2回医療・介護データ等の解析基盤に関する有識者会議」における公開資料に則り、以下各国・組織およびデータベースを紹介する。

表1 今回紹介対象とする事例

	データ提供機関		主たる対象データ
アメリカ	省庁傘下組織および研究支援グループ	CMS ResDAC	レセプトデータ
イギリス	省庁傘下組織	CPRD	臨床データ
フランス	保険者	CNAMTS	レセプトデータ
韓国	保険者	NHIS	レセプトデータ
台湾	省庁	衛生福利部	レセプトデータ

4.2 データ提供件数、利用料の徴収および民間提供の可否

各事例における、データ提供件数、利用料の徴収および民間利用の可否については、以下の通りであった。

表2 今回紹介対象とする事例

	年間データ提供件数	利用料の徴収	民間提供の可否
アメリカ	300～400件	あり	可
イギリス	不明、但し論文刊行数は累計2000本以上	あり	可
フランス	不明、利用者は年50名以上	なし	可
韓国	700件以上	あり	不可
台湾	2011年以降1000件以上	あり	可

4.3 海外動向の小括

各国とも歴史的、地理的、人口動態的要因等、さまざまな背景の下で医療制度は様々であり、大規模ヘルスケアデータの質や収集経路、国民に対するカバー率等も、それぞれに大きく異なっている。とはいえ、NDB データの提供件数が年々増加していることはこれまでに述べたとおりであるが、それでも直近で年間 50 件を超えることはなく、今回紹介したいずれの事例と比較しても、件数は下回っていることが確認できた。一方で、多くの海外事例においては利用料を徴収する仕組みすでに成立しており、アメリカ CMS やイギリス CPRD などでは、これを原資としてデータ提供環境の充実化を図られているとのことである¹⁰⁾。各国において、データ利活用の充実化においては相応の資源が投入されていることが伺われる。

5. 結論

ここまで、主に「レセプト情報等の提供に関する有識者会議」資料を引きながら、レセプト情報等データベース(NDB)の利活用について、これまでの動向の整理を行ってきた。平成23年以降、すでに7年が経過しており、提供件数の増加とともに、データ利用に資する様々な改善が認められた。一方で、海外の事例と比較すると、必ずしもデータ利用件数自体は多いとはいえなかった。今後は、海外の先進事例なども参考にしながら、改善を継続していくことが重要であろう^{11) 12)}。

謝辞

本稿で述べた各種事象の検証等は、日本医療研究開発機構平成30年度臨床研究等ICT基盤構築・人工知能実装研究事業「NDB データによる理想的な健康医療ビッグデータ活用モデルの確立に関する研究」(2018年度)、厚生労働行政推進調査事業費補助金「医療費適正化に向けた生活保護受給者の医薬品処方および生活習慣病の実態調査:大規模レセプト分析」(2017年度～)、厚生労働行政推進調査事業費補助金「患者調査等、各種基幹統計調査におけるNDBデータの利用可能性に関する評価」(2017～18年度)ならびにJPSS 科研費18H04076「レセプトデータベース(NDB)の利用を容易にするための包括的支援システムの開発」(2018年度～)で得た資金を活用して行うことができた。ここに御礼申し上げる。

参考文献

- 1) レセプト情報・特定健診等情報の提供に関するホームページ. 厚生労働省保険局, 2018.
[http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryoku/iryokuhoken/reseputo/index.html (cited 2018-Sep-5)].
- 2) 厚生労働省保険局. 第41回レセプト情報等の提供に関する有識者会議・第三者提供の現状について(報告). 厚生労働省保険局, 2018.
[<https://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000211810.pdf> (cited 2018-Sep-5)].
- 3) 厚生労働省保険局. 第1回NDBオープンデータ【解説編】. 厚生労働省保険局, 2016.
[<http://www.mhlw.go.jp/file/06-Seisakujouhou-12400000-Hokenkyoku/0000141549.pdf> (cited 2018-Sep-5)].
- 4) 厚生労働省保険局. 第6回レセプト情報等の提供に関する有識者会議・公表形式についての一応の基準(案). 厚生労働省保険局, 2011.
[<https://www.mhlw.go.jp/stf/shingi/2r9852000001g288-att/2r9852000001g2dq.pdf> (cited 2018-Sep-5)].
- 5) 厚生労働省保険局. 第26回レセプト情報等の提供に関する有識者会議・成果物の公表基準に関する検討. 厚生労働省保険局,

2015.
[https://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000108547.pdf (cited 2018-Sep-5)].
- 6) 厚生労働省保険局. 第 26 回レセプト情報等の提供に関する有識者会議・サンプリングデータセットの内容充実について. 厚生労働省保険局, 2015.
[https://www.mhlw.go.jp/content/12401000/000332313.pdf (cited 2018-Sep-5)].
 - 7) 厚生労働省保険局. 第 38 回レセプト情報等の提供に関する有識者会議・レセプトデータと特定健診等データの突合について(報告). 厚生労働省保険局, 2017.
[https://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000174510.pdf (cited 2018-Sep-5)].
 - 8) 松居宏樹, 小林康毅, 大江和彦, 橋本英樹, 康永秀生. 第 41 回レセプト情報等の提供に関する有識者会議・レセプト情報等オンラインリサーチセンターの現況について. 厚生労働省保険局, 2018.
[https://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000211811.pdf (cited 2018-Sep-5)].
 - 9) 東尚弘. 第 42 回レセプト情報等の提供に関する有識者会議・オンラインセンターを利用して. 厚生労働省保険局, 2018.
[https://www.mhlw.go.jp/content/12401000/000347518.pdf (cited 2018-Sep-5)].
 - 10) 加藤源太. 第 2 回医療・介護データ等の解析基盤に関する有識者会議・海外におけるレセプトデータ等の利活用の動向について. 厚生労働省保険局, 2018.
[https://www.mhlw.go.jp/file/05-Shingikai-12401000-Hokenkyoku-Soumuka/0000209440.pdf (cited 2018-Sep-5)].
 - 11) 加藤源太・大寺祥佑. レセプト情報をはじめとする保健医療データの二次利用を可能にする体制についての研究: 海外事例を手掛かりに. 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))「患者調査等、各種基幹統計調査における NDB データの利用可能性に関する評価」(代表: 加藤源太, 課題番号 H29-政策-指定-005), 2018. 72-80.
 - 12) 加藤源太. レセプト情報・特定健診等情報データベース(NDB)の第三者利用に関する利用者に向けた利便性向上策の動向について. 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))「医療費適正化に向けた生活保護受給者の医薬品処方および生活習慣病の実態調査: 大規模レセプト分析」(代表: 高橋由光, 課題番号 H29-政策-指定-007), 2018. 59-66.