

平成 29 年度 厚生労働行政推進調査事業費補助金
(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 (H29-政策-指定-007)

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

研究分担者 加藤源太 京都大学医学部附属病院 診療報酬センター准教授

研究要旨

レセプト情報・特定健診等情報データベース (NDB) の第三者利用においては、2011 年の開始以降現在に至るまで、NDB データの利用者向けに様々な利便性向上の施策が打ち出されてきたところである。しかし、それらは「レセプト情報等の提供に関する有識者会議」での会議資料ならびにウェブサイトの議事録情報などにおいて発表されるのみであり、利用者間での NDB に関する知識に格差が生じている恐れがある。こうした背景のもと、本研究では 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った。結果として、広く一般に開放されている NDB オープンデータの公表ならびに利用者のフィードバックに基づいたデータの質の改善、申出範囲の拡大や実用性を踏まえた公表基準の現実化、サンプリングデータセットの質の向上やデータの紐付け率の改善、サーバ増強によるデータ提供運用体制の強化などといった NDB データ提供体制の拡充が図られており、NDB データの利用者も年々増加してきていることが確認された。一方で、NDB を利用した研究成果はまだそれほど充実した質量を確保できてはならず、NDB の第三者利用に関する利便性向上の施策を今後も積極的に推進させていく必要があるとともに、そうした施策を積極的に開示して利用者の裾野を広げていく努力も、行政と研究者とで力を合わせて推進していく必要がある。

A. 研究目的

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 データの

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「レセプト情報等の提供に関する有識者会議」での会議資料ならびにウェブサイトの議事録情報などにおいて発表されるのみとなっている。このように、それらの有益な情報が利用者の目線に立って再構成されているわけではないため、これらの情報を把握している利用者とはそうでない利用者とは、NDB 利用に関するリテラシーに差が生じている恐れがあると同時に、期待される研究成果の質量に影響を及ぼしている可能性が無いとは言えない状況にある。本研究で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、NDB の利用を検討する者がその概要を容易に把握できるようにすることを目的とするものである。

B. 研究方法

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。ただし、公開資料は「議事進行のためのたたき台」として使用されている場合があり、実際の議論では公開資料に記されていることが否定されていたり、あるいは修正が加えられていたりすることがある。また、有識者会議の資料はあくまでも時系列に沿ってリストアップされているため、NDB に関する各々のトピックが数回の有識者会議を隔てて出現することも少なくなく、かなり丹念に資

料を読み込まないと、各トピックの連続性を追うことができない場合がある。こうした状況を踏まえ、当研究ではここ 2~3 年の有識者会議の議論を中心にレビューしながら、NDB の利用を検討する者にとって関係が深いトピックごとに、これまで提示されてきた昨今の利便性向上策の動向を再確認し、整理することとする。

C. 研究結果

ここ最近の有識者会議での議論から、複数回にまたがって議論されているトピックが幾つか見いだされた。本章において、それらトピックを列記しつつ、利便性向上に関する議論の動向を以下に確認する。なお、オンラインリサーチセンターに関する議論は、まだ試行利用にとどまっており多くの利用者が利用できる環境が十分に整っていないことを考慮し、本稿では言及しない。カッコ内の数値は、その議論がなされた有識者会議の開催回の情報を示すものである。

1) NDB オープンデータ

NDB オープンデータは、NDB データの有用性を活かすため、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り²、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（第 33 回）。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、ここから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。第 1 回 NDB オープンデータが公表された 2016 年 10 月以降、利用者からのフィードバックに基づき、2017 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。

2) 市区町村の申出範囲への追加、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するにあたっては、その公表結果によって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（第 6 回）が、同時に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（第 6 回）。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

るとなると、多くの重要な集計値が表現できなくなり、集計結果を十全に活用できないことが懸念される。第 26 回の有識者会議においては、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、

- 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
- 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
- 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（第 26 回）。2017 年 5 月時点で市区町村からの申出も 3 件みられ（第 37 回）、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) 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、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、2012 年 6 月より提供が開始されたデータセットである

(第9回)。特別抽出データの場合は、利用に際しては最小限のデータ提供しか認めておらず、探索的な利用も禁じているのに対し、サンプリ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、ある程度の探索的研究を利用者に認めている点が異なっている。このサンプリ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使用することは不可能である。一方で、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それほど大仰なものとならない、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、

- 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。
- 「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。
- 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」のサンプリ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(第26回)。2018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%弱がサンプリングデータセットの利用を希望しており(第40回)、サンプ

リ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。

4) 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(第1回)。また、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はレセプトデータは「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。これに対しては、「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、2015年度の特定健診等データとID3をさかのぼって設定した2015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。これにより、NDBの本来目的である、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。

5) データ提供承諾件数の増加

NDBデータの第三者提供が開始された2011年度は、総申出件数43件に対して承諾件数が6件と非常に低調であった。2012年度には状況はやや改善したものの、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、2013年度は承諾件数は22件と、20件を超えるようになり、2014年度の20件を経て、2015年度には34件、2016年度には42件と、承諾件数が増加の傾向を示すようになってきた。また、有識者会議での審査において、全ての提供依頼申出のうちデータ提供が承諾される割合も、2013年度以降は90%を超えており（第37回）、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。

D. 考察

ここ2年余りの有識者会議の議論を中心に、NDBの利用を検討する者にとって関係が深いトピックについて、レビューを行ってきた。広く一般に開放されているNDBオープンデータの公表ならびに利用者のフィードバックに基づいたデータの質の改善、サン

プリングデータセットの質の向上やサーバ増強によるデータ提供運用体制の強化などといったNDBデータ提供体制の拡充が図られており、NDBデータの利用者も年々増加してきている。一方で、NDBのデータ構築が開始されてから間もなく10年、第三者提供が開始されてから7年が経過しており、保健医療データの利活用に対する社会からの要請も、その度合いを増してきている。「経済財政運営と改革の基本方針」（「骨太の方針」）では、年余にわたりNDBを含むレセプトデータを活用した、医療費の適正化についての提言が繰り返されている^{3,4}。NDBの第三者提供にあたり当初から掲げられてきた「医療サービスの質の向上等を目指して正確なエビデンスに基づく施策を推進するにあたって有益となる分析・研究、学術研究の発展に資するような研究」⁵への貢献の程度で言えば、2017年1月時点でデータ利用申出者が厚生労働省に開示した研究成果において、査読付きの学術論文は19編（うち英文9編）とのことであった⁶。単純に比較できるものではないが、イギリスにおいて特定の医師からの報告に基づく患者診療データを核として構築されたデータベースClinical Practice Research Datalink (CPRD)からは1,800本以上の研究成果が刊行されており⁷、悉皆性において勝る日本のNDBにはデータの利活用を推進する施策を今後も充実化させていくことで、保健医療領域において更なる成果を導出できる余地があるかと思われる。そのためにも、NDBの第三者利用に関する利便性向上の施策を今後も積極的に推進させていく必要があるとともに、そうした施策を積極的に開示して利用者の裾野を広げていく努力も、行政と

研究者とで力を合わせて推進していく必要がある。

E. 結論

NDB の第三者利用に関する利便性向上策については、利用者や利用を検討する者の声も踏まえながら、各トピックについて漸次推進・改善されてきていることが確認された。

G. 研究発表

1. 論文、書籍発表

- 1) Kubo Shinichiro、Noda Tatsuya、Myojin Tomoya、Nishioka Yuichi、Higashino Tsuneyuki、Matsui Hiroki、Kato Genta、Imamura Tomoaki. National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB): Outline and Patient-Matching Technique. *bioRxiv*, 4, 2018, DOI: <https://doi.org/10.1101/280008>.
- 2) 加藤源太、中山健夫. 1 レセプト情報・特定健診等情報データベース (NDB) を用いた研究の動向 2) NDB データを用いた臨床研究. *Progress in Medicine*, 38, 123-126, 2018.
- 3) 加藤源太. レセプト情報・特定健診等情報データベース (NDB) 利活用の歩み. *生体医工学*, 55, 143-150, 2017.
- 4) 野田龍也, 久保慎一郎, 明神大也, 西岡祐一, 東野恒之, 松居宏樹, 加藤源太, 今村知明. レセプト情報・特定健

診等情報データベース (NDB) における患者突合 (名寄せ) 手法の改良と検証. *厚生の指標*, 64, 10-15, 2017.

- 5) Yamasaki Daisuke、Tanabe Masaki、Muraki Yuichi、Kato Genta、Ohmagari Norio、Yagi Tetsuya. The first report of Japanese antimicrobial use measured by national database based on health insurance claims data (2011-2013): comparison with sales data, and trend analysis stratified by antimicrobial category and age group. *Infection*, 46, 207-214, 2017.
- 6) Nakayama Takeo, on behalf of BiDAME: Big Data Analysis of Medical Care for the Elderly in Kyoto, Imanaka Yuichi, Okuno Yasushi, Kato Genta, Kuroda Tomohiro, Goto Rei, Tanaka Shiro, Tamura Hiroshi, Fukuhara Shunichi, Fukuma Shingo, Muto Manabu, Yanagita Motoko, Yamamoto Yosuke. Analysis of the evidence-practice gap to facilitate proper medical care for the elderly: investigation, using databases, of utilization measures for National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB). *Environmental Health and Preventive Medicine*, 22, doi:10.1186/s12199-017-0644-5, 2017.
- 7) 久保慎一郎. 野田龍也. 明神大也. 東野恒之. 松居宏樹. 加藤源太. 今村知明. レセプト情報・特定健診等情報デー

データベース (NDB) の臨床研究における名寄せの必要性和留意点. 日本健康開発雑誌, 38, 11-19, 2017.

- 8) 浦西友樹, 丸山裕, 内藤知佐子, 岡本和也, 田村寛, 加藤源太, 黒田知宏. 失敗を可視化する採血トレーナ. 日本バーチャルリアリティ学会論文誌, 22, 217-227, 2017.
2. 学会発表等
- 1) 岩尾友秀, 平木秀輔, 大寺祥佑, 酒井未知, 田村寛, 加藤源太, 黒田知宏. レセプト情報・特定健診等情報データベース (NDB) を対象とした疫学研究に適した分析用データベースの構築、第 11 回 IT ヘルスケア学術大会: 名古屋、2017 年 5 月 27 日.
 - 2) 岩尾友秀, 大寺祥佑, 酒井未知, 平木秀輔, 大鶴繁, 近藤英治, 加藤源太, 田村寛, 黒田知宏. A reconstruction method of health insurance claims database for epidemiological research、生体医工学シンポジウム 2017: 上田、2017 年 9 月 15 日.
 - 3) 加藤源太, 酒井未知, 大寺祥佑, 下垣徹, 松居宏樹, 野田龍也, 康永秀生, 今村知明, 黒田知宏. 新たなエビデンス創出のための次世代 NDB データ研究 基盤構築に関する研究: 疫学研究への活用可能性について、第 1 回日本臨床疫学会年次学術大会: 東京、2017 年 9 月 29 日.
 - 4) 加藤源太, 酒井未知, 大寺祥佑, 下垣徹, 松居宏樹, 野田龍也, 康永秀生, 今村知明, 黒田知宏. 新たなエビデンス創出のための次世代 NDB データ研究 基盤構築に関する研究: 概要報告、第 76 回日本公衆衛生学会総会: 鹿児島、2017 年 11 月 1 日.
 - 5) 酒井未知, 大寺祥佑, 岩尾友秀, ネフ由紀子, 加藤源太, 黒田知宏, 高橋由光, 中山健夫. 大規模レセプトデータベースを用いた高齢者終末期医療の実態解明、第 12 回医療経済学会学術大会: 東京、2017 年 9 月 1 日.
 - 6) 加藤源太, 趙晃済, 中谷友香, 大鶴繁, 吉村健佑. NDB オープンデータを用いた侵襲度の高い救命治療の実態把握ー既存データを基準とした妥当性の検証ー、第 45 回日本救急医学会年次学術大会: 大阪、2017 年 10 月 24 日.
 - 7) 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. 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ー患者突合(名寄せ)の手法開発と検証ー、第 37 回医療情報学連合大会: 大阪、2017 年 11 月 20 日.
 - 8) 加藤源太, 酒井未知, 大寺祥佑, 下垣徹, 松居宏樹, 野田龍也, 康永秀生, 今村知明, 黒田知宏. 新たなエビデンス創出のための次世代 NDB データ研究 基盤構築に関する研究: 新たなシステム下での検索速度等に関する評価、第 37 回医療情報学連合大会: 大阪、2017 年 11 月 20 日.
 - 9) 大寺祥佑, 酒井未知, 加藤源太, 黒田知宏. NDB オンサイトリサーチセンター (京都) における運用の報告、第 37 回医療情報学連合大会: 大阪、2017 年 11 月 20 日.
 - 10) Masaki Tanabe, Yuichi Muraki,

Daisuke Yamasaki, Genta Kato and Tetsuya Yagi. Geographical analysis of Antimicrobial Consumption Surveillance using the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB JAPAN) 2011-2013, ID WEEK 2017: San Diego, 2017年10月4日.

- 11) Shosuke Ohtera, Michi Sakai, Tomohide Iwao, Yukiko Neff, Yoshimitsu Takahashi, Genta Kato, Takeo Nakayama. ANALYSIS OF STATIN PRESCRIPTION FOR DYSLIPIDEMIA WITH THE NATIONWIDE HEALTH INSURANCE CLAIMS DATA IN JAPAN: A REPEATED CROSS-SECTIONAL STUDY, ISPOR 22nd Annual International Meeting: Boston, 2017年5月23日.

H. 知的財産権の出願・登録状況（予定を含む）

- | | |
|-----------|----|
| 1. 特許取得 | なし |
| 2. 実用新案登録 | なし |
| 3. その他 | なし |

参考文献

1. 厚生労働省保険局、「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」
 (http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/iryuu_hoken/reseputo/index.html, 2018年5月21日確認).

2. 厚生労働省保険局、「第1回NDBオープンデータ【解説編】」
 (<http://www.mhlw.go.jp/file/06-Seisakujouhou-12400000-Hokenkyoku/0000141549.pdf>, 2018年5月21日確認).
3. 内閣府、「経済財政運営と改革の基本方針2014 ～デフレから好循環拡大へ～」
 (http://www5.cao.go.jp/keizai-shimon/kaigi/cabinet/2014/2014_basicpolicies_01.pdf, 2018年5月21日確認).
4. 内閣府、「経済財政運営と改革の基本方針2016 ～600兆円経済への道筋～」
 (http://www5.cao.go.jp/keizai-shimon/kaigi/cabinet/2016/2016_basicpolicies_ja.pdf, 2018年5月21日確認).
5. 厚生労働省保険局、「『医療サービスの質の向上等のためのレセプト情報等の活用に関する検討会』報告書」
 (<http://www.mhlw.go.jp/shingi/2008/01/dl/s0130-16a.pdf>, 2018年5月21日確認).
6. 奥村泰之、佐方信夫、清水沙友里、松居宏樹、「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 2017年10月号、16-25.
7. CPRD, CPRD: Clinical Practice Research Datalink
 (<https://www.cprd.com/intro.asp> , 2018年5月21日確認).