

公募企画

公募企画シンポジウム1

レセプト情報等オンサイトリサーチセンターの現況および今後について – これまでの進捗の報告、および今後の第三者利用について–

2017年11月21日(火) 08:30 ~ 10:00 E会場 (10F 会議室1003)

[2-E-1-PS1-1] 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

大寺 祥佑¹, 酒井 未知¹, 加藤 源太², 黒田 知宏¹ (1.京都大学医学部附属病院医療情報企画部, 2.京都大学医学部附属病院診療報酬センター)

レセプト情報・特定健診等データベース（以下、NDB）は、医療保険請求に関する国内最大のデータソースである。医療ビッグデータの活用に対する社会的関心が高まる中、NDBを用いた医療の実態把握や新たなエビデンスの創出に期待が寄せられている。しかしNDBを利用するには適切なセキュリティ環境の確保が求められ、研究者等にとっては大きな参入障壁となってきた。そこで、あらかじめ万全なセキュリティ環境を整えてNDBのデータ利用前に生じる負担を軽減することで、より多くの研究者等が利用できる仕組みを目指して、レセプト情報等オンサイトリサーチセンターが設置された。

2016年2月、京都大学でレセプト情報等オンサイトリサーチセンター（京都）（以下、オンサイト京都）の試行的利用が開始された。オンサイト京都では、NDB分析基盤のパフォーマンス検証や運用体制の構築に向けた検討を行ってきた。分析基盤のパフォーマンス検証では、操作環境の各種アプリケーションによるデータ抽出や統計解析における特徴を整理した上で、過去にNDB特別抽出によるデータ提供を受けた疫学研究と同様の分析をオンサイト京都でも実施し、分析所要時間や操作感覚等について特別抽出を利用した場合とオンサイト京都を利用した場合の違いを調べた。運用体制の構築に向けては、今後オンサイト京都を学外の研究者等が利用することを想定して、実際に発生しうる設備破損等の施設利用に絡む問題について学内の関連部署および厚生労働省との調整を行い、規程類を整備した。加えて、レセプト情報等オンサイトリサーチセンターの持続的かつ円滑な運用を可能とするために必要な資源の試算も行った。

当セッションでは、これまでのオンサイト京都における試行利用の経過を報告するとともに、レセプト情報等オンサイトリサーチセンターの特性を活かしたNDBの利用促進に関する今後の方策について議論したい。

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

大寺祥佑^{*1}、酒井未知^{*1}、
加藤源太^{*2}、黒田 知宏^{*1}

*1 京都大学医学部附属病院医療情報企画部、

*2 京都大学医学部附属病院診療報酬センター

Progress in Usage of National Insurance Claims Database (NDB)

On-site Research Center in Kyoto

Shosuke Ohtera ^{*1}, Michi Sakai ^{*1}, Genta Kato^{*2}, Tomohiro Kuroda ^{*1}

*1 Division of Medical Information Technology and Administration Planning, Kyoto University Hospital

, *2 Solutions Center for Health Insurance Claims, Kyoto University Hospital

The National Insurance Claims Database (NDB) is the largest data source on medical insurance claims in Japan. As social interest in utilization of medical big data has increased, expectations are being made for grasping the actual condition of medical treatment using the NDB and creating new evidence. However, rigorous security environment is required to use the NDB and it has become a big entry barrier for researchers. The on-site research centers were set up aiming at a system that can be used by more researchers by preparing strict security environment beforehand and reducing the burden caused before using the NDB data. In 2016, the trial use of the NDB On-site Research Center in Kyoto was started at Kyoto University and conducted performance test for analytical platform of the NDB data and establishment of the operational framework. In the performance test, we analyzed the characteristics of data extraction and statistical analysis by applications in the system and analyzed the data according to the plan of epidemiological research acknowledged as a simulation task in the trial use of the On-site Research Center. For building the operation framework when researchers out of the Kyoto University use the On-site Research Center in the future, the related departments in the university and the Ministry of Health, Labor and Welfare should adjust the regulations. In addition, we also tried to estimate the resources to enable sustainable and smooth operation of the On-site Research Center. In this session, we will report on the progress of trial use and discuss future strategies for promoting the use of NDB in the On-site Research Center.

Keywords: Insurance, Administrative Claims, Database

1. はじめに

レセプト情報・特定健診等情報データベース (national insurance claims database: NDB) は、高齢者の医療の確保に関する法律に基づき、2009 年 4 月診療分以降の電子化されたレセプト情報、ならびに 2008 年度分以降の特定健康診査、特定保健指導の情報が蓄積されたデータベースである。¹⁾ NDB が構築された本来の目的は医療費適正化計画のための調査等への活用である一方で、医療サービスの質向上のためのエビデンスに基づいた施策の推進に向けた調査や学術研究への活用も認められており、研究者等の第三者に対するデータの提供が行われてきた。

NDB に含まれる情報は、個人情報に準ずる情報として取り扱われなければならない。そのためデータ提供を受ける場合は、データ漏洩等に十分配慮した高いセキュリティ要件を満たすことが求められる。一方、安全なデータ利用環境の構築は研究者等の負担となるため、NDB のデータ利用を促進する上での障壁となることが懸念された。そこで NDB のデータ利用を希望する者がデータを受け取って自ら分析環境を整える負担を減らすため、セキュリティ環境があらかじめ整備された場所にデータ利用者が出向いて、そこで分析を行えるような体制の実現をめざしてレセプト情報等オンサイトリサーチセンター(以下、NDB オンサイト)の整備が進められた。この NDB オンサイトは、2014 年度の公募で採択を受けた東京大学と京都大学に設置された²⁾。

東京大学に設置されたレセプト情報等オンサイトリサーチセンター(東京)、ならびに京都大学に設置されたレセプト情報等オンサイトリサーチセンター(京都)(以下、オンサイト京

都)は 2015 年度に試行利用を開始した。それ以降、オンサイト京都では、京都大学医学部附属病院および同大学大学院医学研究科の教職員らが運用体制の整備を行うとともに、NDB のオンサイト利用におけるパフォーマンス検証等を行ってきた。

2. レセプト情報等オンサイトリサーチセンター(京都)の運用体制およびデータ利用環境

オンサイト京都では京都大学医学部附属病院にて定めた内規に基づき、レセプト情報等オンサイトリサーチセンター(京都)運用部を設置し、当該施設の運営、レセプト情報等の第三者へのデータ提供制度に関する普及・啓発、レセプト情報等に関する研究等を行ってきた。運用部ならびに運用部の円滑な運営に関する審議を行う運営委員会は、医学部附属病院ならびに大学院医学研究科社会健康医学系専攻(公衆衛生大学院)の教職員によって構成されている。現在は厚生労働省が定めたオンサイトリサーチセンターにおけるレセプト情報等の試行的利用規約、レセプト情報・特定健診等情報の提供に関するガイドライン、および同省と京都大学が定めた運用管理規定に則って運用を行っている。施設責任者および運用責任者は、運用管理規定に添付された自己点検手順書に従って運用の適切性に関する点検を定期的に行い、その結果を運営委員会に報告している。

オンサイト京都には、NDB 分析用の端末が 2 台設置されている。電子錠によって入退室可能な部屋の中に、ロッカーなどが置かれた前室と、鍵付きのドアにより入退室可能な 2 つの領域(各領域は壁で区切られている)を用意し、それぞれの領域に端末を 1 台ずつ配置している。前室と端末が置かれた

各領域には防犯カメラが設置され、部屋の内部の様子が常時記録されている。端末は回線によって厚生労働省のデータセンターと接続され、NDBの中で第三者利用が認められているデータを操作することができる。サーバーの表領域は150GB、ネットワークの保存領域は1TB、ローカル領域のハードディスクは2TB、メモリは16GBとなっており、端末内には分析用のアプリケーションとして、NDBのデータをSQLで操作するためのインターフェースであるSQL*Plus、GUIでマウス操作が可能なOracle Business Intelligence (BI)、統計解析ソフトのSASとRが用意されている。またRにはNDBのデータ操作を可能にするOracle R Enterprise (ORE)が導入されているほか、Rの統合開発環境であるRStudioやテキストエディタのサクラエディタも使用することができる。

3. NDBのオンサイト利用におけるパフォーマンスの検証

昨年度、オンサイト京都ではNDBのオンサイト利用における分析環境の操作性や処理能力に関するパフォーマンスの検証を実施した。操作性についてはOracle BIがマウス操作可能であるのに対して、SQL*Plus、R、SASはコマンドラインによる操作となるため不慣れた利用者にとってはハードルが高いと考えられた。

Oracle BIの機能には、条件指定によりレセプト帳票の抽出や行数の集計を行うことができる「定形帳票」に加えて、テーブルの結合や計算式の設定が可能な「自由分析」がある。定形帳票は操作が簡便である反面、複雑な分析を行うことはできない。一方で自由分析は複雑な分析も行うことができるが、操作手順を覚えるには一定以上のリテラシーが必要であると考えられた。SQL*PlusはSQLを実行してNDBのデータを抽出、集計することができる。テキストエディタとの連携やファイル実行も可能である。RはOREの利用によってNDBのデータを抽出、分析することができる。OREには統計解析用の関数も用意されている。RではCRANのパッケージを使用することも可能で、ローカル領域に保存したNDBのデータを用いて統計解析を実行できる。SASはNDBへの接続を可能とするインターフェースが用意されていないため、ローカル領域のデータを用いた統計解析のみが可能である。

オンサイト京都では端末の操作性に加えて、データの集計や抽出に関する処理能力の検証も行った。まず、前立腺がんまたは前立腺肥大の患者の1年分の医科レセプトの件数を集計した結果、Oracle BIの定形帳票、自由分析、SQL Plus、OREのそれぞれで一定時間内に処理が完了した(各6分、1分、2分、12分)。傷病名に「高血圧」を含む医科レセプトが発行された患者のID1の数の集計をSQL*Plusで行った結果、ID1の数の増加にともなって処理時間が概ね線形に増加することがわかった(1年分の約2700万人で約2分、3年分の4200万人で約8分、6年分の5700万人で約17分)。また、これと同じ処理をオンサイト京都の端末2台同時に行うと処理時間が倍増することがわかった。次に、NDBのデータをローカル領域にダウンロードしたところ、傷病名に「高血圧」を含む医科レセプトの患者ID1、レセプトID、年齢、性別、診療年月、医療機関所在地(都道府県)、点数の7列では、DPCレセプトの4年分(約2700万行、1.4GB)の処理時間がSQL*Plusで8時間、OREで25分であった。また条件で医科レセプトの6年分(約15億行、25.4GB)であった場合、SQL*Plusの処理時間が100時間で、OREはメモリの上限(16GB)に達したために処理が中断された。最後に、傷病名

に「高血圧」を含むDPCレセプトの2年分(1200万行、11列、1.6GB)のデータを用いて重回帰分析を行った。ローカル領域におけるSASによる処理は1分以内に完了したのに対し、Rでは16GBのメモリ上限に達してしまい処理が中断した。一方で同様の処理をOREで実行すると約5分で処理を完了することができた。

以上より、NDBオンサイトの操作性や処理能力に関する一定の知見を得たものと考えている。なお、この検証結果は2016年7月に開催された第32回レセプト情報等の提供に関する有識者会議にて報告した³⁾。

4. オンサイトにおける臨床研究実施の検証

現在、オンサイト京都では試行利用の一貫として有識者会議の承諾を得た模擬申出課題の臨床研究に取り組んでいる。具体的には、高齢者に対して不適切な処方と疑われる薬剤の使用実態を明らかにするため、高齢者の安全な薬物療法ガイドライン(STOPPリスト)およびピアーズリストに挙げられる薬剤の処方割合等を調べる研究である。本研究はオンサイト京都の利用者を含む研究グループが、NDBの特別抽出データの提供を受けて既に実施しているものである。そのため、この検証では、オンサイトと特別抽出で同じ研究方法を用いた場合の各々の利点欠点について明らかにすることを主な目的としている。

同様の目的で、オンサイト京都の利用者を含む研究グループが既にNDBの特別抽出やサンプリングデータで実施した他の研究テーマについても検証を予定している。このようにして試行利用期間中に得られた知見を蓄積し、広く社会に普及することで、今後NDBオンサイトを利用する研究者にとって有用な情報資源となることを期待している。

5. NDBオンサイトの今後の展望

これまでの試行利用を通して得た知見をもとに、NDBオンサイトの将来像を考えたい。NDBのデータ利用者となりうるのは、現状、政策立案者や研究者に限られる。よって今後NDBをオンサイトで利用することが想定されるのも、同様の部類に属する人たちであろうと思われる。NDBオンサイトのパフォーマンス検証や現在実施中の臨床研究を通じて、オンサイトを利用する際には一般的な知識や技術に加えて、オンサイトに特有の知識や技術も必要であることがわかってきた。一般的な知識や技術とは、NDBに含まれるレセプト情報や特定健診情報、それらを用いた研究手法に関する知識、データベースや統計解析に関する知識や技術である。一方、オンサイトに特有の知識や技術には、オンサイトの端末スペックや処理能力に関する知識、オンサイトで利用可能なアプリケーション(Oracle Database、SQL*Plus、Oracle BI、R、ORE、SAS)に関する知識や技術などが含まれる。これらの知識や技術を十分に持つ利用者であれば、短期間に成果を得ることができる。運用面とあわせて考えると、そのような利用者が増えれば増えるほど、オンサイト利用の効率化が期待できる。一方で、利用者の知識や技術が不十分である場合には、成果を得るまでに時間がかかるか、あるいは何らかの支援が必要となる可能性がある。

オンサイト京都での経験を踏まえて考えると、すべての利用者に上述の知識や技術を完全に習得させることは非現実的である。将来的に実現可能なオンサイトの利用方法として、複数名からなるチームを組める研究者であれば、自分の専門分野以外の知識や技術を補えるメンバーとともに研究を進められるのが理想的であろうと考える。また研究チームの構築が

難しく、単独あるいは少数で行わなければならない研究の場合には、適宜支援を受けられるリソースを確保しておく必要があるだろう。そのようなリソースの準備を今後どのように行うかは重要な課題であり、立場によって様々な考え方ができるように思うが、少なくとも試行利用を通じて蓄積した各オンサイトからの知見をうまく活用する方策を検討する価値はあるだろう。

今後、試行利用に続いて、NDB オンサイトの第三者利用の開始が見込まれている。オンサイト京都の関係者が把握しているだけでも、NDB オンサイトの第三者利用に関する問い合わせが2016年度京都大学内外から10件以上あった。問い合わせの内容や研究計画の質は様々であるが、中にはリサーチセッションが明確で実現可能性が高いと思われる研究も含まれていた。第三者提供が開始された場合に発生すると考えられる業務として、研究の実施可能性に関する検証やFAQ対応、文書整備や契約手続、オンサイトの利用に関する管理(訪問日調整、生体認証登録、鍵貸出、手順書整備、自己点検などのリスク管理)、利用契約終了手続きが考えられる。これらの業務に対応できる人的資源の確保が必要であることは自明である。また第三者利用の開始にともなって、各オンサイトを設置する法人(大学)と厚生労働省の間での取り決めを見直さなければならない場合もある。例えば、第三者がオンサイトを利用しているときに発生した設備の破損等の損害に対する復旧にかかる補償をどのように行うのかについては、関連文書の改訂が検討されている。

6. まとめ

本稿ではオンサイト京都の設置から試行利用の開始、運用体制の整備、NDBのオンサイト利用に関する検証に関するこれまでの経過をまとめ、第三者利用の開始を含めたNDBオンサイトの今後の展望について記述した。当日のセッションではNDBオンサイトの関係者らとともに、本稿で示した課題や展望について、さらに発展した議論となることを期待したい。

7. 研究資金

本稿で述べたオンサイトの運用および各検証を行うにあたり、厚生労働科学研究費補助金戦略型研究「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」(2014年度-2016年度)および日本医療研究開発機構臨床研究等ICT基盤構築研究事業「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」(2016年度-2017年度)より資金を得た。

参考文献

- 1) レセプト情報等の提供に関する有識者会議(第1回)資料. 厚生労働省, 2010.
[<http://www.mhlw.go.jp/stf/shingi/2r9852000000thao.html> (cited 2017-Sep-14)]
- 2) 第23回レセプト情報等の提供に関する有識者会議資料. 厚生労働省, 2014.
[<http://www.mhlw.go.jp/stf/shingi2/0000072558.html> (cited 2017-Sep-14)]
- 3) 第32回レセプト情報等の提供に関する有識者会議資料. 厚生労働省, 2016.
[<http://www.mhlw.go.jp/stf/shingi2/0000131575.html> (cited 2017-Sep-14)]