

大規模レセプト情報の二次利用を取り巻く環境

鈴木 里彩*, 小西 孝明*

抄 録

匿名医療保険等関連情報データベース（NDB, National Database of Health Insurance Claims）は、医療レセプト情報ならびに特定健診・特定保健指導情報等によって構成された医療ビッグデータである。2008年に医療費適正化計画の作成、実施及び評価のために構築され、2011年から二次利用のための第三者提供が開始された。2023年12月現在、約260億件の医療レセプト情報と約3.8億件の特定健診・特定保健指導情報が格納されている。また、2016年からNDBオープンデータが公開されており、基礎的な集計表が無料で即時ダウンロード可能である。厚生労働省では、第三者提供の迅速化、クラウド上での解析環境の開発、連結解析対象となるデータベースの拡大、NDB収載情報の拡大、申請・審査プロセスの合理化、研究者支援体制の構築等、二次利用の利便性の向上に取り組んでいる。

キーワード：データベース、NDB、クラウド、レセプト、二次利用

社会保障研究 2024, vol.9, no.1, pp.54-65.

I NDBの概要

1 データベースの概要

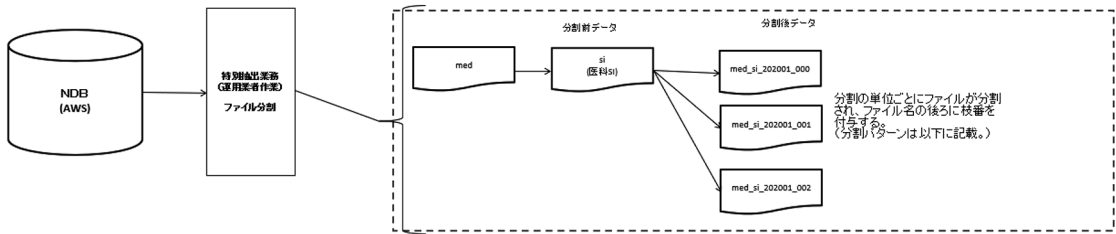
本邦において国が保有する大規模レセプト情報の一つに、匿名医療保険等関連情報データベース（NDB, National Database of Health Insurance Claims）がある。高齢者の医療の確保に関する法律（昭和57年法律第80号。以下、「高確法」という。）第16条に基づいて、医療費適正化計画の作成、実施及び評価のために厚生労働省が2008年に構築した。同第16条の2に基づいて、厳格な運用の下にデータの第三者提供が行われており、研究目的による二次利用が可能である。

NDBには保険診療にまつわるさまざまなデータが含まれており、国民皆保険下にある本邦では、国民の全数に近い医療動向を把握することができる。保健・医療を対象とした政策立案や研究の遂行にあたり、貴重な医療ビッグデータであると位置づけられる一方で、機微性の高い情報であることから、利用者には高いセキュリティ環境の担保が求められることや、その膨大なデータ量からハンドリングが難しいことなど、利活用の推進に向けた課題も残っている。

2 データベースの内容と構成

NDBは保険診療における電子化された医療レセプト情報ならびに特定健診・特定保健指導情報

* 厚生労働省保険局医療介護連携政策課保険データ企画室 室長補佐



【提供データ サンプル】

```

ir
00001,Rzqiwk0gc1,,,IR,1,01,1,01,23456,abcdfg,202002,00,01,01,202001,202004
00020,Rzqiwk0gc2,,,IR,1,01,1,01,23457,abcdfh,202002,00,01,01,202001,202004
00039,Rzqiwk0gc3,,,IR,1,01,1,01,23458,abcdfi,202002,00,01,01,202001,202004
00056,Rzqiwk0gc4,,,IR,1,01,1,01,23459,abcdfj,202002,00,01,01,202001,202004

iy
00011,Rzqiwk0gc1,,,IY,,,999999999,1,00000,,,,,,,,,,,,,,,,,,,,,202001,202004
00029,Rzqiwk0gc2,,,IY,,,999999999,1,00000,,,,,,,,,,,,,,,,,,,,,202001,202004
00047,Rzqiwk0gc3,,,IY,,,999999999,1,00000,,,,,,,,,,,,,,,,,,,,,202001,202004
00056,Rzqiwk0gc4,,,IY,,,999999999,1,00000,,,,,,,,,,,,,,,,,,,,,202001,202004

re
00003,Rzqiwk0gc1,,,RE,123(略)abc,123(略)abczzz(略)111,,,202001,,,,,,,,,,,,,202004
00021,Rzqiwk0gc2,,,RE,456(略)def,456(略)defxxx(略)222,,,202001,,,,,,,,,,,,,202004
00039,Rzqiwk0gc3,,,RE,789(略)ghi,789(略)ghiyzz(略)333,,,202001,,,,,,,,,,,,,202004
00057,Rzqiwk0gc4,,,RE,123(略)kl,123(略)klwww(略)444,,,202001,,,,,,,,,,,,,202004
    
```

出所：申出テンプレートの注意点。

図1 NDBデータの提供形式

等によって構成されている。NDB格納時に後述する個人特定リスクを減ずる処理をされた上で、レコード毎のcsvファイル形式で提供される。

医療レセプト情報は以下のような流れでNDBに収載される。保険診療を行った医療機関は、診療報酬点数表に基づいて計算した診療報酬（医療費）を毎月の月末に患者一人ずつ集計して明細書を作成し、審査支払機関を経由して保険者へ診療報酬を請求する。この明細書をレセプト（診療報酬請求明細書）という。傷病名、診療開始日、投薬、処置、手術、検査等の情報を含んだレセプトは、患者別、月別、受診した診療所及び病院別、入院外来別にそれぞれ1部ずつ発行される。診療報酬請求のために提出された医療レセプトは、氏名等の個人を識別する情報が削除された後、医科、DPC、調剤、歯科の4種類のレセプトのそれぞれについて、情報の種別毎のレコードに区分して格納される。診療月から約3ヶ月後に格納されており、その件数は年間約20億件にのぼる。2023年12月時点で、2009年4月診療分からの約260億件のレセプトが格納されている。なお、レセプトは保

険診療に関する診療報酬明細書であるため、現状保険収載されていない妊娠/分娩等に関する医療行為や、業務上の疾病、予防医療等の保険外の診療はNDBには収載されない。また、検査値やバイタルサイン等の臨床指標も格納されていない。

特定健康診査（以下「健診」という。）および特定保健指導については、2008年度から40歳以上75歳未満の被保険者・被扶養者を対象として、内臓脂肪型肥満に着目した事業の実施が医療保険者に義務付けられているものである。特定健診の結果、生活習慣病の予防効果が多く期待できる者に対して特定保健指導の実施が義務付けられている。それらの特定健診・特定保健指導について、実施日、健診結果、問診結果等の情報が支払基金に法定報告され、個人情報等を削除したうえでNDBに格納される。2023年12月時点で、2008年度実施分からの約3.8億件が格納されている。

情報の安全性を高めるため、NDB格納時に、一部の項目は通番化、階級化等の匿名化処理が行われている。氏名や医療機関名などの固有名詞は削除されて一定のルールに基づくIDに置き換えら

表1 NDBに格納されているID

ID	元となる情報	特徴	格納期間	主な利用方法
ID1	保険者番号, 被保険者証の記号・番号, 生年月日, 性別	保険者の変更や誤記により 紐付けができなくなる 可能性がある	レセプト: 2009年4月診療分～ 特定健診: 2008年度実施分～	レセプト情報間の紐付け レセプト情報と特定健診・保健指 導情報との紐付け
ID2	氏名, 生年月日, 性別	氏名の変更や誤記により 紐付けができなくなる 可能性がある	レセプト: 2009年4月診療分～ 特定健診: 2008年度実施分～	レセプト情報間の紐付け
ID4	カナ氏名, 生年月日, 性別	氏名の変更や誤記により 紐付けができなくなる 可能性がある	レセプト: 2018年4月診療分～ (介護: 2020年3月給付分～) (DPC: 2020年4月診療分～)	NDBデータとほかの医療・介護等 データとの紐付け
ID5	被保険者番号の履歴	保険者の変更により紐付け できなくなるID1, ID1Nの 課題に対応	レセプト: 2022年2月診療分～ 特定健診: 2021年度実施分～	レセプト情報間の紐付け レセプト情報と特定健診・保健指 導情報との紐付け NDBデータとほかの医療・介護等 データとの紐付け

出所:「NDBの利用を検討している方へのマニュアル2023年10月版」より一部改変。

れる。医療機関コード、健診実施機関コード、保険者番号等の機関を一意に特定できるコードは、機関毎に一意となるような通番化を施されて格納される。年齢は、提供時に原則として階層化される。そのほか、生年月日については日の情報が削除されている。一方、オンライン資格確認システムを活用して、2022年より患者居住地の郵便番号や限度額適用認定証区分がレセプト情報に付加されるようになった。上記にあげた項目については、研究計画において活用する必要性があると認められた場合に二次利用目的に第三者提供されており、研究目的に応じて適切に粒度を下げた状態で提供することが可能である。

NDBは個人を特定可能な情報を削除しているため、異なる医療レセプトや特定健診・特定保健指導情報を連結するには、ID等による「名寄せ」が必要となる。名寄せを可能とするため、運用開始当初から、氏名・性別・生年月日に基づくIDと、保険者番号・被保険者番号・性別・生年月日に基づくIDがNDBには格納されていた。しかし、結婚等による改姓や転職による保険者変更により、同一個人であっても同一IDが付与されず名寄せができないことが問題視されていた。そのため、2022年から、医療保険の資格を確認するオンライン資格確認システム基盤を活用して、個人の持つ最も古い被保険者番号に基づくID (ID5) の格納を開始した。これによって、個人追跡の精度が向上することが期待される。

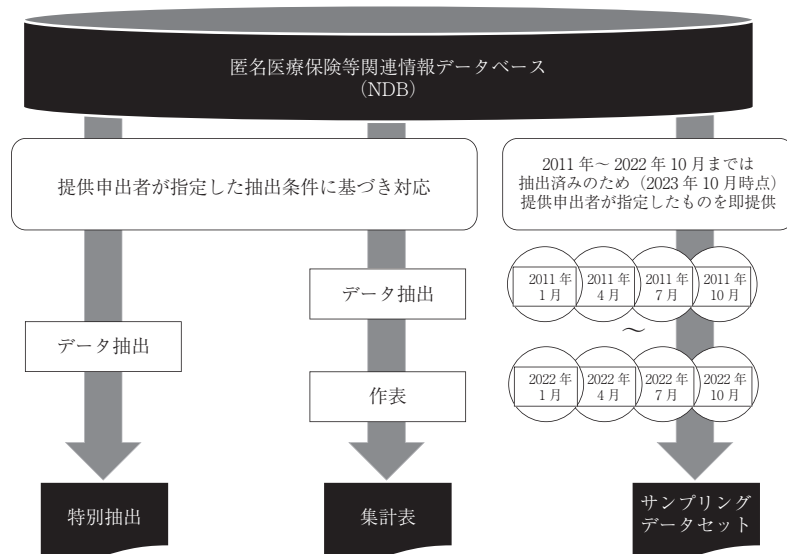
このほか、NDBに格納・提供される際の加工については、厚生労働省の「**【NDB】匿名医療保険等関連情報データベースの利用に関するホームページ**」内、「申請に必要な書類」の「申出依頼テンプレートの注意点」に詳述されている。

3 NDBの二次利用

2008年のNDB構築時、NDBデータの利用は厚生労働省内の医療費適正化を目的とするものに限られていた。その後、データ利活用促進の一環として、2011年には提供に関するガイドラインの初版が作成され、大学の研究者等による二次利用のためのNDBデータの第三者提供が開始された。

2019年の健康保険法（大正11年法律第70号）等改正により、第三者提供は法定化され、研究計画に公益性があること、商用利用でないこと等を条件として民間事業者へのデータ提供も可能となった。また、同改正で介護保険総合データベース（介護レセプトを格納）や匿名診療等関連情報（DPCデータを格納）との連結解析も可能となった。

データ提供の可否については、高確法第16条の2の第3項により、「厚生労働大臣は予め社会保障審議会の意見を聴かなければならない」とされており、社会保障審議会医療保険部会の下に設置される保険者や研究者などの有識者から構成される専門委員会が合議によって意見を述べ、最終的に厚生労働大臣が判断している。審査事項及び審査



出所：「NDBの利用を検討している方へのマニュアル2023年10月版」より一部改変。

図2 2023年10月時点のNDB提供データの種類

基準は「NDBの利用に関するガイドライン」に示されており、提供申出者等の氏名等、利用目的、希望するデータの概要と必要性、研究体制等、安全管理対策、結果の公表予定、その他必要な事項に区分されている。

この制度のもと、利用目的に応じた複数の形式(集計表、サンプリングデータセット、特別抽出)により、年間70件程度の第三者提供が行われている。

先述したとおり、NDBは保険医療を受けるすべての国民の医療レセプトが集積された機微性の高い情報である。提供したデータが不正に二次利用されることのないよう、法令、規約、ガイドライン等にさまざまな安全管理措置や規定が定められている。法令上は、高確法において、NDBデータを利用するものは他の情報と照合してはならないこと(第十六条の三 照合等の禁止)、利用する必要がなくなった時はデータを遅滞なく消去しなければならないこと(第十六条の四 消去)、漏洩等の防止その他の安全管理を講じなければならないこと(第十六条の五 安全管理措置)、知りえた内容を不正な目的に使用してはならないこと(第十六条の六 利用者の義務)、厚生労働大臣は利用

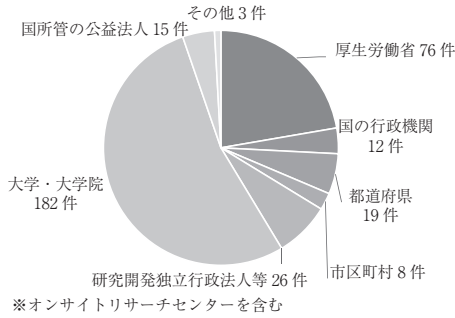
場所等に立ち入って検査できること(第十六条の七 立入検査等)、以上の項目に違反している場合は厚生労働大臣が是正を命じられること(第十六条の八 是正命令)が定められている。また、規約やガイドラインにおいて、違反に対する措置内容として一定期間のデータ利用停止や新規提供禁止等が定められ、厳格な運用が行われている。

4 NDBオープンデータの公開

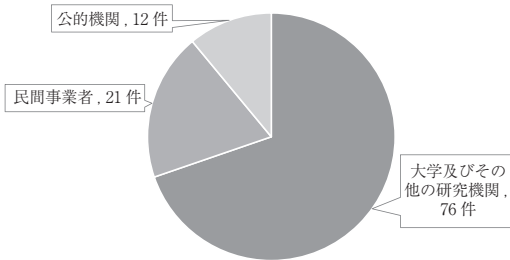
上記のようなNDBデータ提供が特定の施策や研究等の利用を目的として個別に行われている一方で、厚生労働省はNDBデータから汎用性の高い基礎的な集計表を作成して、NDBオープンデータとして厚生労働省のホームページ上で公開している。本集計表はExcelファイルとして即時ダウンロード可能で、同ホームページには年間約44万回のアクセスがある。

まず2016年に、医科(入院、入院外)、歯科、調剤、DPCの各レセプトおよび特定健診データから都道府県別及び性年齢階級別に集計したデータの公開を開始した。2019年からは二次医療圏別、診療月別の集計が追加された。オープンデータの詳細な集計内容については、NDBオープンデータの

【提供依頼申出者の区分（件数）（2020年9月審査分まで）】



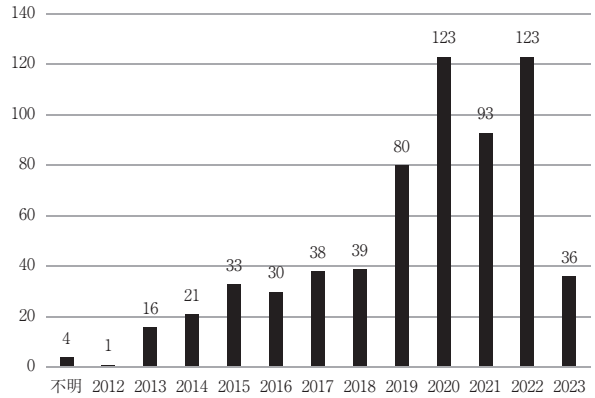
【提供申出者の区分（件数）（2020年12月審査分から）】



※担当者の所属施設を集計（1申出につき主たる研究機関1つをカウント）
※オンラインリサーチセンターを含む

出所：第16回匿名医療情報等の提供に関する専門委員会資料3より一部改変。

発表年ごとの成果物の集計



※申出者より成果物として提出された公表物について発表年で集計
※2023年は5月時点で確認

図3 第三者提供の提供申出者の区分と成果物の集計

ホームページからリンクされている各回のオープンデータの公表ページに「解説書」として提示している。

2021年には、excel形式のデータ掲載だけでなく、BIツールを用いてデータをビジュアライズできるようにしたNDBオープンデータ分析サイトを公開した。過去分のデータについても遡ってBIツールで操作できるようになっている。

2024年以降、これまで算定回数の集計を基本としていたところ、ID5を用いた算定人数の集計も開始する予定である。今後も、ホームページ上で収集している意見や有識者からの指摘を踏まえ、必要性を吟味しながら随時集計内容を拡充していく予定である。

II NDB利活用の現状と課題

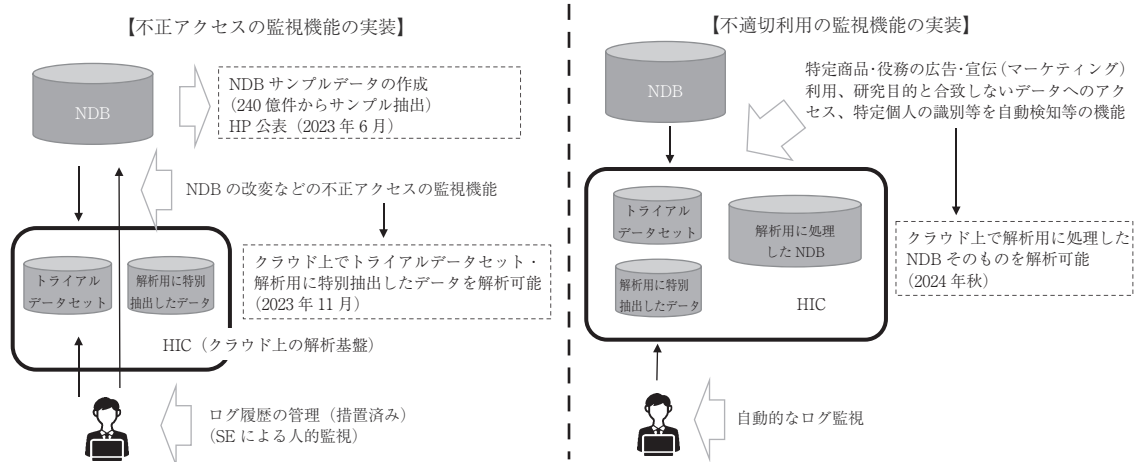
1 二次利用の現状

NDBはこれまで約500件の第三者提供を行っており、その多くは大学等の研究機関に対する提供である。主たる利用方法はその悉皆性を活かした記述疫学研究であるが、治療や背景因子等の比較研究も実施されている。NDBデータを利用した研究成果の公表も徐々に増えており、年ごとの調査で、2022年度には約150件の成果物の報告があった。

2 NDB二次利用の抱える課題

利活用が進む一方で、データ提供に要する時間の長さが課題となっている。承諾からデータ提供まで平均330日を要し、その大部分が研究計画に対し必要最小限のデータを提供するための調整/

- 直ちに、サンプルデータ、トライアルデータセットの作成、セキュリティの担保されたクラウドの実装に取り組み、
 - ・2023年6月、NDBサンプルデータを厚生省HPに公表
 - ・2023年11月、リモートアクセスでトライアルデータセット・解析用に特別抽出したデータを解析可能
 - さらに、不適切利用等の監視機能やポータルサイトの機能拡充を開発・実装の上、
 - ・2024年秋、リモートアクセスの解析データを拡大
 - ・申請からデータ提供まで平均390日の現状に対し、申請メ切を毎月設定し、申請から原則7日で処理
- ※申請が月5件程度であることを踏まえ、当月月1回を設定するが、今後申請件数が増えれば複数回設定する



出所：第165回社会保障審議会医療保険部会資料2より一部改変。

図4 NDBデータ提供の抜本的見直し（概要）

審議と、研究計画をデータの抽出条件に落とし込むためのコーディネートと、抽出作業に費やされている。

NDBは診療報酬請求のため集められた各月のレセプトを格納した二次利用目的のデータベースであり、研究への一次利用目的に構築されたレジストリーと異なり複雑な構造を有している。このため、NDBを用いた分析を行うには、一般的なデータベースハンドリングや統計解析、公衆衛生・疫学等の知識に加えて、レセプト構造の理解、コードの解釈のための診療報酬の知識が必要である。さらに、NDBを研究活用のためのパネルデータとして活用するためには、レセプト単位で作成されるデータを個人単位化する等の前処理が必要となる。

また、ハード面では、研究に携わる者のみが入室できる区画を用意し入室管理を行う等の安全管理対策が求められている上、受領するデータ量によっては高額のサーバーが必要になる等、環境

整備のハードルも高い。

3 「NDBの迅速提供に向けたスキーム再構築に資する研究」

こうした現状を受けて、2023年6月の規制改革実施計画において、「厚生労働省はNDBについて、利用申請から申請者が実際にデータの利用を開始し得るまでに要する期間について、平均で390日を要する現状から原則7日（研究者等側の都合に要した期間は除く。）とすること」とされ、NDB提供体制は抜本的な見直しを求められている。

この見直しのため、厚生労働省ではNDBを利用する研究者等による特別研究「NDBの迅速提供に向けたスキーム再構築に資する研究」を実施し、NDBデータの利活用の推進のために、対応が必要な政策・技術的課題の検討と、将来的なNDBの在り方も含めた課題解決に向けた具体的提言を仰いでいる。2023年12月に行われた匿名医療等関連情報提供に関する専門委員会では、本研究の中間

報告として課題解決のための方向性が示された。以下にその一部を抜粋する。厚生労働省では、研究班からの提言を踏まえて、引き続き具体的な運用改善に向けた検討を行う。

(1) 分析に適したデータセットと提供方法に関する指摘

NDBを研究活用のためのパネルデータとして活用するためには、レセプト単位で作成されるデータを個人単位化する等の前処理が必要となる。このようなNDBの利便性を向上させるための前処理の一部は、これまでも厚生労働省において行われてきているが、NDB分析に関する知見は今も発展段階であり、今後広く活用するためには、こうした個人単位化するために必要な情報の付与やデータベース構造の見直し等のデータベースを扱いやすくする不断の見直しが必要である。そのためには、研究者は利活用を通じて得られたNDBの分析に関する知見を厚生労働省に還元し、厚生労働省はその知見に基づき、次項で提案するような迅速提供のためのデータの整備を優先的に行いつつ、今後も新たに得られた知見を優先度に応じて発展させるようなアジャイル型のデータベースの発展させていくことが望まれる。

(中略)

原則7日での提供にむけては、これまで、研究に必要なNDBデータのみを抽出したデータセットが研究者に提供されていたため、抽出から提供までの時間がかかる要因となっていた。他方、提供期間短縮のために、全ての研究者が全てのデータに実質的な制限なくアクセス可能とすることは、個人特定性のリスクやインフラ維持のコストを考えたときに現実的なものとは言えない。研究利用の観点においては、全ての研究において必ずしもNDBに格納された十数年分の全てレセプトデータが必要ではなく、一定期間のデータについて、1/1000を抽出した場合でも外来入院合わせ

て年間で延べ200万レセプト程度のデータとなるため、一定の探索的な解析を含め、多くの研究には十分な数となる。この十分量のデータについて、事前のデータ処理やHIC上での提供を合わせて行うことで、個人特定性のリスクを減じることができ、コストも低減できると考えられる。従って、症例の縦断的追跡を可能となるような情報を付与・加工し、1~数年分の十分なデータ件数を確保したデータの提供を広く行い簡易な審査によって提供することで、短期間で安全性の高いNDBデータ提供が実現するものと考えられる。前述の加工済みデータよりも多くのデータが必要な研究については、追加的な審査の上で、提供された環境に追加的なデータ利用を可能とすることにより、初回の利用から切れ目無い研究環境を構築できると考えられる。

NDBの全データが研究実施に必ずしも必要ではない点に触れたうえで、一定の探索的解析を含む多くの研究に十分なデータ件数を確保した、1~数年分の縦断的追跡研究を可能とするデータセットの開発が提案されている。事前のデータ処理等により個人特定性のリスクを下げた上で、簡易な審査でこのデータセットを提供することにより、短期間で安全性の高いNDBデータ提供が可能となると考えられる。また、追加的な審査を経てより広範なデータの利用を認めることで、切れ目のない研究が可能であるとしている。

(2) 研究者支援に関する指摘

このほか、一部の研究者は頻回にNDBを利用しているが、研究を始める際に様々な課題があり、新規でNDB研究を行おうとする者は多くはない現状がある。新規参入者を増加させるためにも、利便性の高い分析環境の構築とともに、研究者への支援体制も必要である。研究を始める上で特に困難となるのは、アクセスするNDBのデータ構造に関する理解や、

それに基づく前処理手法、マスタ整備等がある。こうした課題に対しては、これまで複数の分担研究者がNDBに関するe-learning等を開発し、データベースを再構築するための提案等が行われてきている。こうしたNDBに関する分析等の知見を一元的に集約したプラットフォーム環境の構築が期待される。このほか、上記のような仕組みを前提とした申請・審査体制の再構築も必要である。複数のデータベースの連結解析に係る申請の際に同一の書類を複数の事務局に提出しなければならない、必ずしも個人特定性のリスクが低いデータセットの利用にも厳格に審査している、等の課題があり、データ利用における安全性確保、特にHIC特有の解析環境の課題やアクセス向上による新たなリスクへの対応等も前提としつつ、合理的でない申請業務は書類審査の実現等、研究者の負担軽減の観点に立った運用改善が必要である。

研究者支援に関しては、初めてNDBデータを用いた研究を行うとする者にとって、データ構造に関する理解や、データ前処理手法、マスタ整備が障壁となることが述べられている。新規参入者を増やすために、これらに関する知見をまとめたプ

ラットフォームの構築が提案されている。また、複数データベースを利用する際の複雑な申請手続や、すべてのデータセットについて行われている会議審査等について、研究者の負担軽減の観点に立った運用改善が必要であるとしている。

III 医療・介護等のデータベースの二次利用を取り巻く今後の展開

1 クラウド基盤の整備

厚生労働省では、媒体による提供を基本としているNDB等のデータベースの二次利用をクラウド環境で利用可能とする、医療・介護データ等解析基盤（HIC, Healthcare Intelligence Cloud）を段階的に整備している。このようなリモートアクセス環境上でのデータ提供は、海外においても米国CMS（Center for Medicare and Medicaid Services Data）のVRDC（Virtual Research Data Center）、英国DARS（Data Access Request Service）のSDE（Secure Data Environment）、韓国HIRA（Health Insurance Review and Assessment）、デンマークForskertjeneste等で開始されている。

現在、HIC上で提供可能なのは、サンプリングデータセットからさらに希少レセプトを除外したトライアルデータセットと特別抽出の一部（NDB

表2 HICとNDBの主な安全管理措置の比較

	HICガイドライン	NDBガイドライン（特別抽出の場合）
組織的措置	運営管理規定を整備	運営管理規定を整備
人的措置	法令・契約違反者等は利用不可	法令・契約違反者等は利用不可
入退室管理	利用場所の施錠 職員等以外の入室禁止	利用場所の施錠 入退室のチェック、取扱者以外の入室禁止
盗難・覗き見防止	クリアスクリーン等による覗き見防止 スクリーンショットやスマートフォン等での撮影禁止 利用端末を追跡・遠隔からの命令等により消去する機能	クリアスクリーン等による覗き見防止 スクリーンショットやスマートフォン等での撮影禁止 窃盗防止用チェーン等設置による盗難防止
認証・識別	パスワードの規則遵守 二重認証	パスワードの規則遵守 二重認証
不正ソフトウェア対策	セキュリティ対策のアップグレード 不正なログオン等が認められれば、サービスの利用停止	セキュリティ対策のアップグレード
ネットワーク対策	公衆無線LANに接続しないこと 無線LANの不正アクセス対策 その他、ネットワークの規定	外部ネットワークに接続しないこと
消去	(運用保守業者が削除)	復元不可能な手段で廃棄すること
ログ管理	(運用保守業者が管理)	アクセスログの確認・管理

出所：第15回匿名医療・介護情報等の提供に関する委員会資料1。

と介護DBの連結案件を含む)のみであるが、将来的にはさまざまな医療・介護データ等の解析基盤として活用される構想があり、検討が続けられている。

HICを利用する場合、研究者は厚生労働省管理下のシンクライアント環境でデータを取り扱うことになるため、取扱者のみに入室が限定された区画の用意や、厳格な入退室管理といった、一部の安全管理措置が緩和され、研究環境を用意する負担が軽減される。HICは本格運用が始まったところであり、継続的に利用環境の改善が図られる見込みである。

2 連結解析のさらなる拡大

NDBとNDB以外のデータベースとの連結解析も拡大されつつある。これまで、2020年に介護保険総合データベース、2022年に匿名診療等関連情報がNDBと連結解析可能になっている。2024年4月には匿名感染症等関連情報、5月までに次世代医療基盤法に基づく匿名加工医療情報も連結解析が可能となる予定である。このほか、障害福祉、予防接種、指定難病、小児慢性特定疾病のデータベースについて、連結解析を可能とする法案が成立し、施行に向けて検討中である。また、全国がん登録に係るデータベースについても、NDBとの

連結解析の検討がなされている。抽出したデータの連結解析が可能となるのと同時に、HICのような共通のクラウド基盤で複数のデータベースのデータを用いた解析が可能となる構想についても検討が進められており、今後、厚生労働省の有するデータベースを用いた二次利用の幅が大きく広がると考えられる。

3 NDB記載情報の拡大

2024年から、死亡の事実や死因等の死亡情報のNDB記載が開始される。死亡というハードアウトカムが把握できるようになることから、効果比較研究の幅が広がることが期待される。また、重篤な併存疾患を持つ集団において、正しく死因を把握することで、より正確な治療介入効果の検証に資すると考えられる。

その他、2024年以降、生活保護の医療扶助についてオンライン資格確認が開始されるのと同時に、健診情報がNDBの記載対象となるほか、訪問看護レセプトのオンライン請求開始に伴うNDB記載等、NDB本体データの充実をめざして記載情報が随時拡大されていく予定である。正式な記載・提供の時期については、社会保障審議会医療保険部会や匿名医療情報等の提供に関する専門委員会において、都度公表予定である。

区分	DB名	主なデータ	NDBとの連結の意義・必要性	連結の検討状況等
公的	介護DB	・介護レセプト ・要介護認定情報	要介護者の治療前後の医療・介護サービスの利用状況の把握・分析に資する。	2020年10月開始
	DPCDB	・DPCデータ (診療情報、請求情報)	急性期病院の入院患者の状態の把握が可能となり、急性期医療の治療実態の分析に資する。	2022年4月開始
	障害福祉DB	・給付費等明細書情報 ・障害支援区分認定情報	障害者の治療前後の医療・障害福祉サービスの利用状況の把握・分析に資する。	連結解析を可能とする法案が成立。施行に向けて検討中。
	予防接種DB	・予防接種記録 ・副反応疑い報告	予防接種の有無を比較した、ワクチンの有効性・安全性に関する調査・分析に資する。	連結解析を可能とする法案が成立。施行に向けて検討中。
	感染症DB	・発生届情報	感染症の治療実態と予後の把握・分析に資する。	連結解析を可能とする法案が成立。2024年4月施行予定。
	難病DB	・臨床調査個人票	網羅的・経時的な治療情報を得ることが可能となり、より詳細な治療実態の把握・分析に資する。	連結解析を可能とする法案が成立。連結に向けて検討中。
	小慢DB	・医療意見書		
民間	全国がん登録DB	・届出対象情報 ・死亡情報	各種がんの各ステージ分類毎による治療実態と予後の把握・分析に資する。	引き続き検討中
	次世代DB	・医療機関の診療情報	医療情報と連結・分析を可能にすることにより医療分野の研究開発を促進する。	連結解析を可能とする法案が成立。2024年4月施行予定。

出所：第15回匿名医療・介護情報等の提供に関する委員会資料1より一部改変。

図5 医療・介護データ等の連結解析に向けて

4 申請・審査プロセスの合理化等

現在、NDBの申請プロセスは、メールにおける多数の文書のやりとりを基本としている。「NDBの迅速提供に向けたスキーム再構築に資する研究」中間報告においても利用者の負担軽減の観点に立った運用改善が求められており、規制改革実施計画でも触れられているような手続のオンライン化を活用した、手続の負荷軽減と迅速化が検討されている。

また、現在、厚生労働省の医療・介護データ等を複数連結して二次利用しようとする場合、利用しようとするすべてのデータベースの窓口に対して申出手続を行う必要がある。連結解析を行おうとする申出者にとっては、類似の書類の提出を複数の窓口に対して行わなければいけないため、負担感が大きい。

こうした研究者の利便性を向上するための一元的な利用申請の受付・審査体制のあり方等については、医療等情報の二次利用に関するワーキンググループにおいて今後検討が行われる予定である。

5 研究者支援体制の構築

課題の項でも述べた通り、NDBは研究用の一次利用データベースと異なり構造に特徴があるほか、データを解釈するためには診療報酬体系や記録条件仕様の知識が必要となる。このほか、データ構造に適した前処理を行うことや、マスタの整備が初学者にとってハードルとなることが研究班の中間報告書でも触れられている。

データの二次利用を進めるためには、こうした多くの研究者に共通する課題について厚生労働省として一定の支援体制を構築することが必要である。中間報告で提案されたとおり、これまでに公表されている教育資材や研究成果、共有可能な知財を集約するほか、規制改革実施計画で触れられている専門家集団による研究者支援についても体制を整え、申請時に利用できるポータルサイト（開発中）上にプラットフォームを構築し、一元的に提供できるような環境の検討が進められている。

IV まとめ

NDBは世界でも有数の規模と悉皆性を誇るデータベースであり、今後、死亡情報等の取載拡大や、ほかの医療・介護等データと連結が可能になる等、これまで得にくかった背景情報やアウトカム情報についても充実が図られる見込みである。使いやすいクラウド連携基盤の開発により、厚生労働省の所管する公的データベース同士の連携は益々進むことが期待される。公的データベース以外との連結については、次世代医療基盤法に基づく認定作成事業者の提供する匿名加工情報との連結が2024年春から始まるところであり、議論は緒についたところといえる。引き続き、医療情報の安全な運用を語りながら今後の政策研究・疫学研究に資するデータ利活用の在り方について検討を重ねる。

参考文献

- 【NDB】匿名医療保険等関連情報データベースの利用に関するホームページ https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryuu/iryuhoken/reseputo/index.html（2024年1月19日最終確認）。
- 【NDB】NDBオープンデータ：<https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000177182.html>（2024年1月17日最終確認）。
- NDBの利用を検討している方へのマニュアル 2023年10月版：<https://www.mhlw.go.jp/content/12400000/001162218.pdf>（2024年1月17日最終確認）。
- 規制改革実施計画（令和5年6月16日）pp.66-69：https://www8.cao.go.jp/kisei-kaikaku/kisei/publication/program/230616/01_program.pdf（2024年1月19日最終確認）。
- 申出依頼テンプレートの注意点：<https://www.mhlw.go.jp/content/12400000/001168581.xlsx>（2024年1月18日最終確認）。
- 第2回健康・医療・介護情報利活用検討会 医療等情報の二次利用に関するワーキンググループ資料について 資料4 医療等情報の二次利用に係る基本的な考え方・論点（案）：<https://www.mhlw.go.jp/content/10808000/001188761.pdf>（2024年1月19日最終確認）。
- 第15回匿名医療・介護情報等の提供に関する委員会資料1 新たに連結可能となる医療介護データ等との連結に係る審査方針について：<https://www.mhlw.go.jp/content/12301000/001206825.pdf>（2024年2月13日最終確認）。

第16回匿名医療情報等の提供に関する専門委員会 資料3 匿名レセプト情報等の第三者提供の現状について（報告）：[https:// www. mhlw. go. jp/ content/ 12400000/001101485.pdf](https://www.mhlw.go.jp/content/12400000/001101485.pdf)（2024年1月17日最終確認）。

第18回匿名医療情報等の提供に関する専門委員会 資料1 NDBの迅速提供に向けたスキーム再構築に資する研究中間報告書：[https:// www. mhlw. go. jp/ content/12400000/001176009.pdf](https://www.mhlw.go.jp/content/12400000/001176009.pdf)（2024年1月17日最終確認）。

第20回匿名医療情報等の提供に関する専門委員会 資料1 NDB提供体制の見直しについて（案）：<https://>

[www. mhlw. go. jp/ content/ 12400000/ 001214757. pdf](https://www.mhlw.go.jp/content/12400000/001214757.pdf)（2024年3月7日最終確認）。

第165回社会保障審議会医療保険部会 資料2 NDBデータの利活用の更なる促進について：[https:// www. mhlw. go. jp/ content/ 12401000/ 001114695. pdf](https://www.mhlw.go.jp/content/12401000/001114695.pdf)（2024年1月17日最終確認）。

（すずき・りさ）
（こにし・たかあき）

Secondary use of large-scale claim databases in Japan

SUZUKI Risa*, KONISHI Takaaki*

Abstract

National Database of Health Insurance Claims (NDB) is a large-scale healthcare database that contains anonymized information of almost all population in Japan. The Ministry of Health, Labour and Welfare (MHLW) established NDB for medical expenditure optimization in 2008 and started providing NDB data to third parties for secondary use in 2011. In 2019, the MHLW also began providing NDB data to private businesses, given the public benefit of the research and non-commercial use of the data. As of December 2023, NDB stored approximately 26 billion claims data (such as diagnosis, drugs use, and interventions) and 380 million checkups data (such as results of laboratory tests and medical interviews). More than 500 NDB datasets have been provided for secondary use and approximately 150 publications issued in 2022. NDB Open Data also became available since 2016 for free and immediate download of summary tables. The MHLW is conducting an official research project on prompt provision of NDB data. Furthermore, the MHLW is developing a cloud-based analysis environment, extending interlinked analysis, expanding NDB-stored data, streamlining the application, and establishing a support system for researchers.

Keywords : Administrative database, Checkups, Claims, NDB, Secondary use

* M.D., Ph.D., Insurance Data Planning Office, Division for Health Care and Long-term Care Integration, Health Insurance Bureau, Ministry of Health, Labour and Welfare