

令和6年3月6日

資料1

第20回 匿名医療情報等の提供に関する専門委員会

NDB提供体制の見直しについて（案）

厚生労働省

保険局医療介護連携政策課

Ministry of Health, Labour and Welfare of Japan

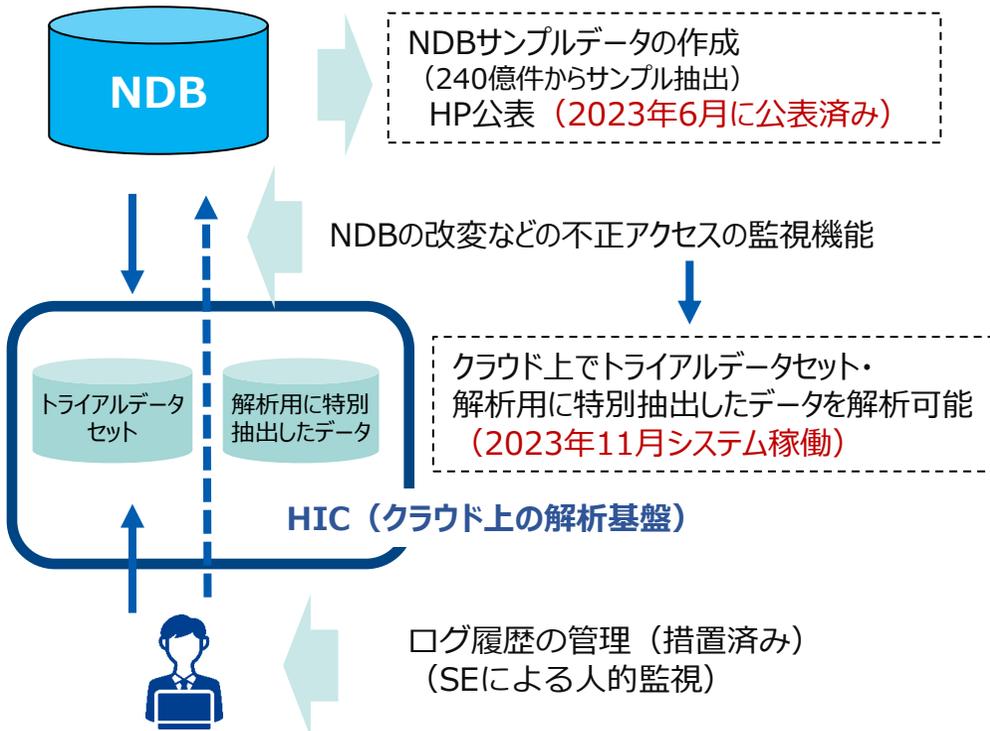
NDBデータ提供の抜本的見直し（概要）

令和5年6月29日
第165回社会保障審議会医療保険部会

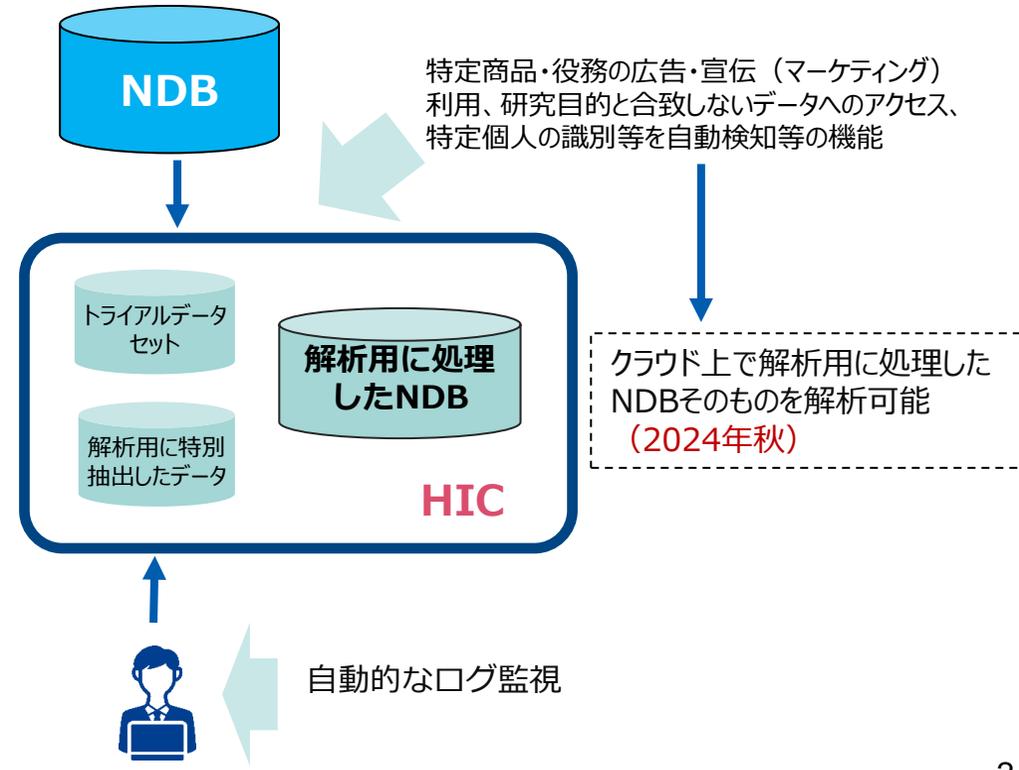
資料2
更新

- **直ちに**、サンプルデータの作成、トライアルデータセットの作成、不正アクセス監視機能の実装に取り組み、
 - ・ **2023年6月**、NDBサンプルデータを厚労省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×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、当面月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

対応済 【不正アクセスの監視機能の実装】



【不適切利用の監視機能の実装】



「NDBの迅速提供に向けたスキーム再構築に資する研究」中間報告の要旨

- NDBの利活用促進のために必要な機能、具備すべき安全管理措置、研究者への支援のあり方等を検討するため、NDBや審査実務に精通した研究者により、令和5年度厚生労働科学研究事業「NDBの迅速提供に向けたスキーム再構築に資する研究」を実施している。
- 第18回匿名医療情報等の提供に関する専門委員会において、当該研究事業の中間報告が行われた。

縦断的追跡を可能とするデータセットについて

- 現状では、研究に必要なデータのみを調整・抽出することが、提供までに時間がかかる要因となっている。しかし、NDB全ての膨大なレセプトデータからのデータ抽出は、必ずしも全ての研究において必要ではない。探索的解析に十分なデータ件数を含み、個人特定性リスクを減じる加工をした、1～数年分の縦断的追跡を可能とするデータセットを作成し、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。

審査の効率化について

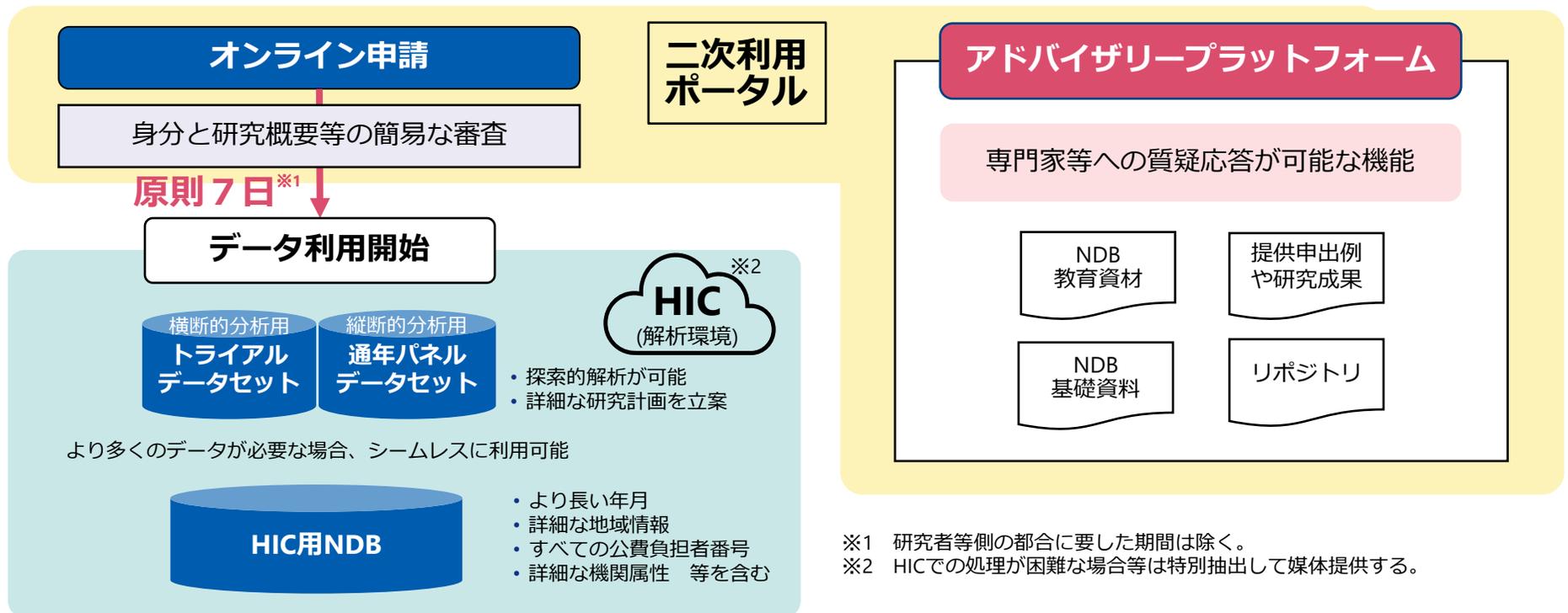
- 個人特定性のリスクの低いデータセットの利用に係る審査については、データ利用の安全性確保を前提としつつ、書類審査の実現等、研究者の負担軽減の観点に立った運用の工夫が必要である。
- 上述のデータセットよりも多くのデータが必要な研究については、追加的な審査の上で、研究者に提供されたクラウド環境において追加的なデータ利用を可能とする。これにより探索的解析から切れ目ない研究環境を構築できると考えられる。

研究者支援について

- 新規参加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。
- 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。

NDBデータの第三者提供フロー見直しの全体像（案）

中間報告で指摘された課題	対応方針
研究ごとに必要なデータを調整・抽出することに時間を要している。	縦断的追跡研究に十分なデータ件数を確保した 通年パネルデータセット をあらかじめHIC環境に配置して迅速に提供する。
安全性確保を前提としつつ、書類審査の導入等、研究者の負担軽減の観点に立った運用の工夫が必要である。	トライアルデータセット、通年パネルデータセット等については、原則として オンラインでの簡易な書類審査 により利用可能とする。
探索的解析に基づいて研究を実施することが困難である。	迅速提供可能なデータを用いた探索的解析に基づく研究の詳細に係る 追加的な審査 を経て、より 広範なNDBデータの利用 を可能とする。
研究者への専門家等による支援体制が不十分である。	NDBに関する知見を集約し、研究者を専門家等が支援する機能等を備えた アドバイザリープラットフォーム を構築する。
申請手続きや研究に必要な情報が集約されていない。	オンライン申請手続きやアドバイザリープラットフォームについて、開発を進めている 二次利用ポータル に機能を 集約 する。



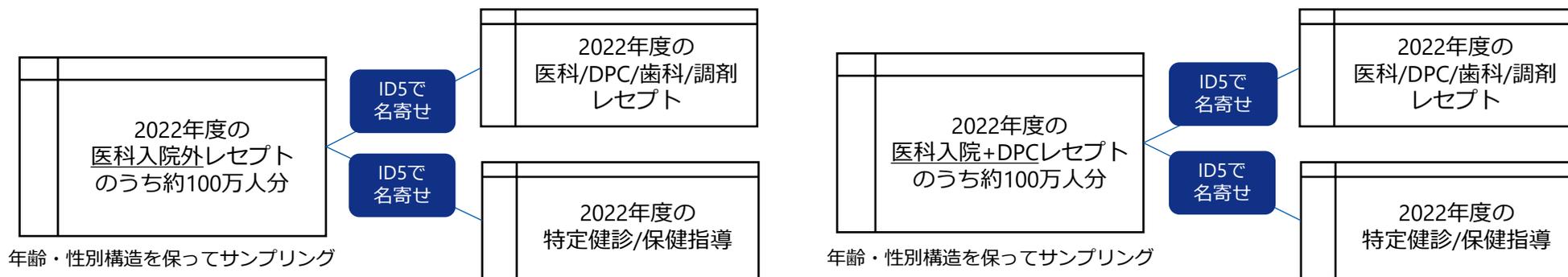
通年パネルデータセットの仕様（案）

作成コンセプト

- 正確性の高いIDによって名寄せを行った、年単位の縦断的追跡を可能とするデータセットを作成する。外来診療と入院診療のいずれも研究対象にできるようにし、HIC利用環境で解析可能なデータサイズとなるよう調整する。
- 簡易な審査でアクセス可能とするため、情報の安全性を高める。
- HIC用NDBの利用に向けてレセプト構造の理解の助けとなるようにするため、NDB本体の構造を保ちつつ、探索的解析に必要な情報を付加する。

仕様の案

- 以下のようなイメージの、2種類の通年パネルデータセットを作成する。



- 機関を表すコード（※1）については、匿名化コードを提供する。
- 医療機関の属性（病院/診療所）と二次医療圏に係るフラグを付与する。郵便番号上二桁（※2）と限度額適用認定証区分を提供し、公費負担者番号上二桁（法別番号）のうち該当者が多い番号（※3）を表示する。

※1 保険者番号、医療機関コード、調剤薬局コード、健診実施機関番号、保健指導実施機関番号。

※2 概ね都道府県に相当する地域番号。

※3 28（感染症法・一類感染症等の入院）、54（難病法・特定医療）、21（障害者総合支援法・精神通院）等。

審査の簡素化に向けた書類審査の導入（案）

- ログ監視機能が実装されたHIC上における個人特定性リスクの低いデータ（※1）の利用に係る申請においては、省令で定められている必要最小限の項目について、ポータルを活用したオンラインでの審査を導入し、簡略化してはどうか。
- HIC用NDBの利用や特別抽出に係る追加的審査においては、引き続き個人特定性等について審査することとし、ポータルを活用したオンラインでの審査を基本としてはどうか。

審査基準	観点	個人特定性リスクの低いデータ利用に係る審査	広範なデータ利用に係る追加的審査
(1)	取扱者の氏名・所属等	○	
(2)	利用目的	○	
(3)	研究の概要	○	
	研究の詳細（※2）		○
(4)	研究体制等	○	
(5)	安全管理対策	○	
(6)	公表予定	○	
(7)	倫理審査通過の有無		○

※1 ランダムサンプリングや個人特定の可能性がある項目のマスク等の処理を行ったトライアルデータセットや通年パネルデータセット等。

※2 研究計画の具体性、抽出内容との整合性、個人特定の可能性がある項目や幅広いデータの利用希望の妥当性等を審査する。過去の審議で論点となった事項がある場合等は審査会による審査の対象とする。

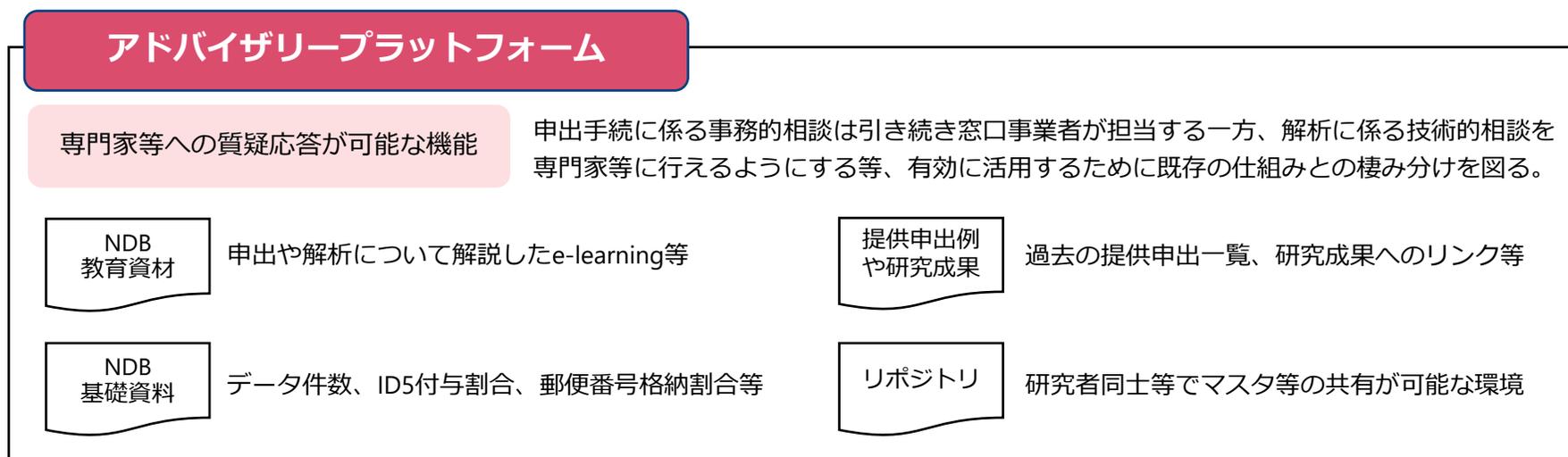
アドバイザープラットフォームの構成（案）

これまでの議論で期待される支援体制

- 規制改革実施計画（令和5年6月16日閣議決定）において「研究者等が希望する場合に、NDBデータの専門家等が抽出条件のアドバイスを行う等の支援体制を構築すること」とされている。
- 特別研究班中間報告において、NDBデータを用いた研究への新規参入者増加のために、NDBのデータ構造、前処理手法、マスタ等に関する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ている。

アドバイザープラットフォームの概要

- NDBデータの提供申請前に閲覧可能なポータルサイト（二次利用ポータル）上に、アドバイザープラットフォームを配置することで、NDBデータを利用した研究を行おうとする研究者の支援を図ってはどうか。
- 以下のようなコンテンツを中心に、具体的内容について引き続き検討する。



参考資料



(参考) 規制改革実施計画 (令和5年6月16日閣議決定)

II 実施事項 3. 個別分野の取組 <医療・介護・感染症対策分野>

(1) デジタルヘルスの推進① – データの利活用基盤の整備 – 2 NDBの利活用の容易化等

- 厚生労働省は、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「高齢者医療確保法」という。）に基づくNDBに収載されたデータ（以下「NDBデータ」という。）の大学、民間事業者等の研究者その他の利用者（以下本項において「研究者等」という。）への提供（高齢者医療確保法第16条の2）等の迅速化及び円滑化を図り、医療サービスの質の向上につなげていくため、以下の措置を講ずる。
 - a. (サンプルデータの公開：対応済み)
 - b. (トライアルデータセットをHIC上で解析可能とし、当該データの審査を簡略化する：対応済み)
 - c. 厚生労働省は、**解析用に処理したNDBデータ（ブラックリスト方式で個人特定の可能性のある項目を匿名化する等の処理をしたもの）に対するリモートアクセスを、以下の点に留意しつつ可能とする。**あわせて、**専門委員会による審査の効率化等を行い、**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。また、現状の申請件数を踏まえ、当面月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。
 - 特定の商品又は役務の広告又は宣伝を目的とする利用、承諾された利用目的以外の利用、特定の個人を識別する目的での利用その他の不適切利用をオンラインで監視可能な解析環境を構築すること。なお、研究者等がNDBデータを利用する場合を含め研究を行うに当たっては、探索・試行的なデータ解析を行うことが通常であることに留意する。
 - 研究者等による、厚生労働省等に対する**リモートアクセスの申請手続等をオンラインで行うことを可能とすること。**
 - 研究者等が希望する場合に、**NDBデータの専門家等が抽出条件のアドバイスを行う等の支援体制を構築するとともに構造化されたデータを整備するなど解析環境を整備すること。**

「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①

はじめに

令和5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提供体制の抜本的な見直しが進められる。本研究は、NDBデータの利活用の推進のために、令和5年度において対応が必要な政策・技術的課題の検討と、将来的なNDBのあり方も含めた課題解決に向けた具体的提言を行い、厚生労働行政に直接的に寄与することを目的としている。

NDBデータの提供については、現在運用が開始されつつあるHICでの提供の実現によって、新たにNDBを用いた研究を始めようとする者の初期投資を大幅に削減することが期待され、今後規制改革実施計画においてとりまとめられたような迅速なデータ提供が実現すれば、劇的に研究促進につながることを期待される。他方、単なるアクセス向上のみを目的とするのではなく、NDBが有する個人特定性等の課題を踏まえつつ、研究に必要な情報を迅速に提供し、結果として、公衆衛生の向上や厚生労働行政に資するNDB利活用を実現させることが重要である。そこで、本研究班では、短期的なアクセス向上への対応にとどまらず、NDBが政策や研究の発展に資するために長期的に目指すべき方向性となることを目的として、まず、NDBの特徴を踏まえた利活用のあり方を検討し、効率的な提供に資するデータの仕様及びそのための提供体制を議論し、更に、現在の審査体制の課題等について、厚生労働省において提供業務に携わったことのある複数の研究協力者に参画頂き、検討を行っている。本中間報告においては、NDBの特徴を踏まえた利活用のあり方や令和6年度に開始するデータ提供内容についての具体案について、専門委員会での議論に資するよう報告する。

NDBの特徴を踏まえた利活用のあり方

NDBは膨大で悉皆的なデータを保有しているという利点が最大の特徴であり、患者重症度等の調整が困難であるという弱点はあるものの、この利点を活かした厚生労働行政の評価や政策利用の手法開発等を推進することが必要である。また、NDBは診療報酬請求のため集められた各月のレセプト毎のデータベースであり、必ずしも研究のための利便性を考慮して構築されたものではない。このため、一般的なデータベースハンドリングや統計解析、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。次項で詳述するが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、今後広く利活用するためには、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。

「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②

規制改革実施計画の対応に向けたデータの仕様と提供体制

NDBを用いた多くの研究は、レセプト単位ではなく個人単位で行われるため、NDBデータが提供された初期段階において、個人単位のパネルデータにするために必要な情報の付与やその他研究に必要なデータを作成するために、データベース構造を再成形する処理等の前処理が実施されることが一般的である。迅速かつ効果的なNDBの政策活用や研究利用を発展させるためには、現時点で妥当と考えられる情報の事前の付与を行いつつ、将来的には、利用しやすい提供用のデータベース構造を設計可能か等も検討を進めることが望まれる。

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。

その他

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。