



NDBデータの利活用の更なる促進について

【NDBの提供形式】

	特別抽出（本省利用）	特別抽出（第三者提供）	オープンデータ	サンプリングデータセット
内容	政策の企画立案のため、厚労省の指定に沿ってデータ抽出を行い、個人単位のデータを利用。	審議会による提供承諾を得た研究において、研究者の依頼に応じたデータ抽出を行い、個人単位のデータを提供	診療行為、処方薬、健診項目等について、全国レベルで集計を行った集計表（csvファイル形式）	1ヶ月分（1月/4月/7月/10月診療分）の匿名レセプトデータから、入院診療10%、外来診療1%を抽出し、高額レセプトの削除等の匿名化処理を行ったプリセットデータ
利用方法等	<ul style="list-style-type: none"> 患者個人単位の経時的な追跡など、<u>複雑な解析が可能</u> 厚労省の指定に沿って、事務局で集計表までを作成。 審議会の審査は不要。 	<ul style="list-style-type: none"> 患者個人単位の経時的な追跡など、<u>複雑な解析が可能</u>（NDBデータに一定の見識を有した研究者が利用） 利用者希望に応じ、事務局で集計表までを作成し提供（NDBのデータ分析ノウハウのない研究者に有用） 提供日数は平均330日（2021年度）。提供件数は年間約30~50件。 	<ul style="list-style-type: none"> 厚労省HP上に公開。利用者は審議会による審査を経ずに自由に即時ダウンロード・利用可能 初診料・再診料、特定の処置、手術等の診療行為、利用頻度の高い医薬品等の利用傾向把握に有用。 全国レベルの集計に加え、診療月・性年齢・都道府県・二次医療圏別の層別集計結果も公開。より細かい傾向の把握も可能。 オープンデータへのアクセス件数は、年間約44万回（2021年度） 	<ul style="list-style-type: none"> 審議会の承諾が必要。<u>プリセットデータをそのまま提供するため、提供日数が短縮。</u> 提供件数は、年間約10件。承諾された研究目的の範囲内で、<u>探索的研究を行うことが可能であり、自由度は高い</u> 提供日数は、データセット再作成やNDBクラウド化による抽出停止期間の影響を受けた期間は、平均109日（2020年度）と平均182日（2021年度）であったが、これらの影響のない期間は平均52日（2019年度）。

【NDB抽出件数等】

	平成31年度	令和2年度	令和3年度	令和4年度	令和5年度予算
抽出件数	212件 本省利用 140件 第三者提供 72件	188件 本省利用 144件 第三者提供 44件	292件 本省利用 218件 第三者提供 74件	327件 本省利用 255件 第三者提供 72件	—
抽出量	計149TB 本省利用 77TB 第三者提供 72TB	計151TB 本省利用 92TB 第三者提供 59TB	計271TB 本省利用 136TB 第三者提供135TB	計188TB 本省利用 69TB 第三者提供 119TB	—
事業費 ※R4年度～予算ベース	2.9億円	3.2億円	4.6億円	9.2億円 ※令和4年4月からクラウド化に伴う利用料を含む	9.6億円

(参考) NDB利活用の促進に係る指摘等について その1

○デジタル田園都市国家構想総合戦略(令和4年12月23日閣議決定)

- レセプト情報等を活用した分析の体制整備の推進・保健医療に関するビッグデータの効果的な利活用を推進するため、NDB(レセプト情報・特定健診等情報データベース)や、NDBと介護DB(介護保険総合データベース)の情報の連結解析を可能とするために構築した医療介護連結解析基盤(HIC)について、ポータル機能や探索的利用環境の運用開始といった機能の充実や利便性の向上等を図る。

○規制改革推進に関する中間答申(令和4年12月22日規制改革推進会議)

II 重点分野

- 「社会課題の解決」をフロンティアとして経済成長を実現すべく、会議として来夏の答申の取りまとめに向けては、(1)デジタル時代の規制改革、(2)GX推進のための規制改革、(3)人口減少等に対応した規制改革、(4)地方の活性化を図る規制改革の4つの分野において重点的に取り組み、これを突破口として規制改革の展開を図る。
- (略) こうした考えの下、以下の事項を突破口にデジタル時代の規制改革を進める。
 - (略)
 - 医療データの利活用促進

○規制改革推進会議 医療・介護・感染症対策WG(令和4年11月7日)

※出席者：佐藤座長(一橋大学教授)、武井座長代理(西村あさひ法律事務所)、杉本委員(日本大学教授)、中室委員(慶應義塾大学教授)、菅原委員(経済同友会常務理事)、印南委員(慶應義塾大学教授)、大石委員(メディヴァ社長)、大浦委員(おとなの学校取締役)、大橋委員(東京大学副学長)、佐々木委員(悠翔会理事長)、落合委員(渥美坂井法律事務所)、戸田委員(日本電気常務理事)、林委員(桜坂法律事務所)、説明者：宮田俊男氏(早稲田大学理工学術院教授)、他

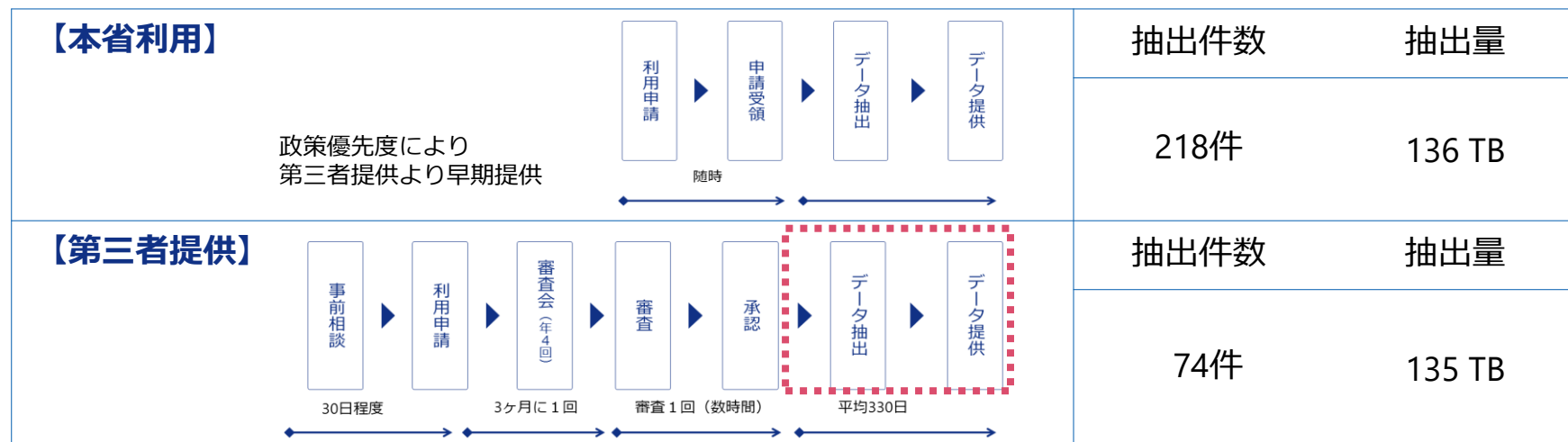
議題：医療データ等の利活用の促進について

(1) NDBの利活用容易化について

厚生労働省、日本製薬工業協会、健康保険組合連合会からヒアリング

NDBデータの提供について

【利用申請からデータ提供まで】



※令和3年度実績値

【第三者提供における承諾からデータ提供の内訳】

※令和3年度に承諾を受けて翌年度（12月時点）に提供した4件の実績

	研究者の書類提出 ※1	①抽出条件の確定		②抽出待ち ※2	抽出作業 ※3	手数料納付等 ※4	媒体発送 ※5	合計
		コーディネータ	DB技術者					
平均日数	32	151	9	55	44	37	3	331

課題：

- ・研究者とデータベース技術者（SE）をつなぐコーディネータの不足
- ・研究者のNDBデータリテラシーの不足

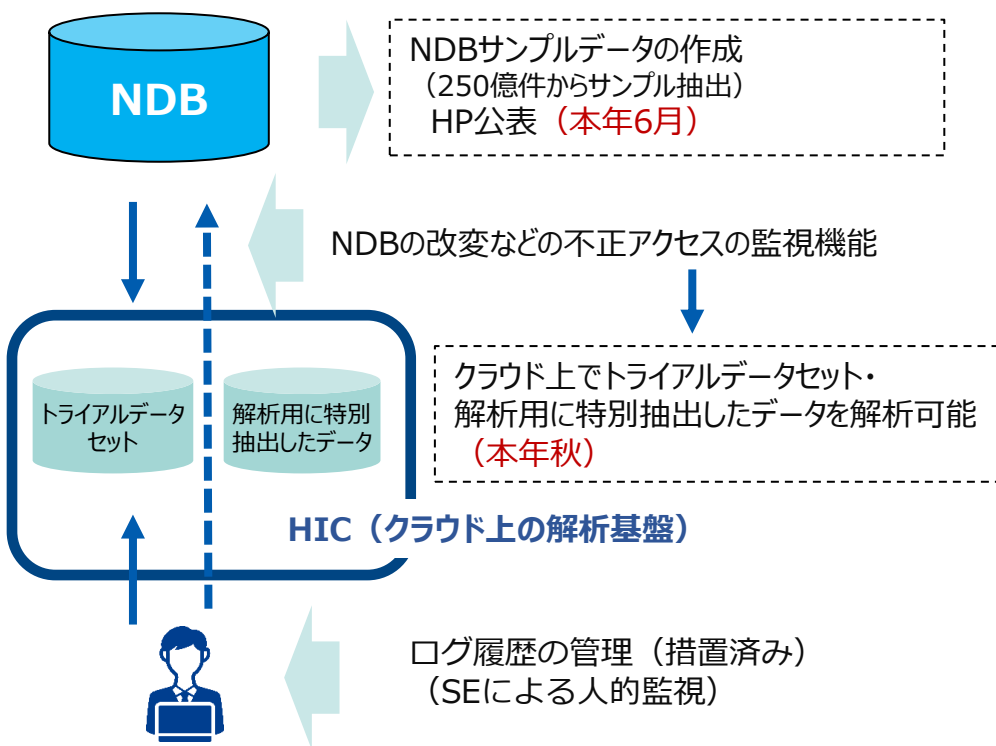
- ・物理的制限（サーバー数）
- ・人的制限（技術者（SE）の数）

- ※1 研究者の書類提出（データ抽出依頼書・誓約書）
- ※2 抽出作業まわりの滞留期間
- ※3 プログラムの動作点検・修正対応を含む
- ※4 手数料通知の発送 + 研究者からの手数料納付
- ※5 R5年度中のHICの運用開始により、クラウド上での利用が可能に

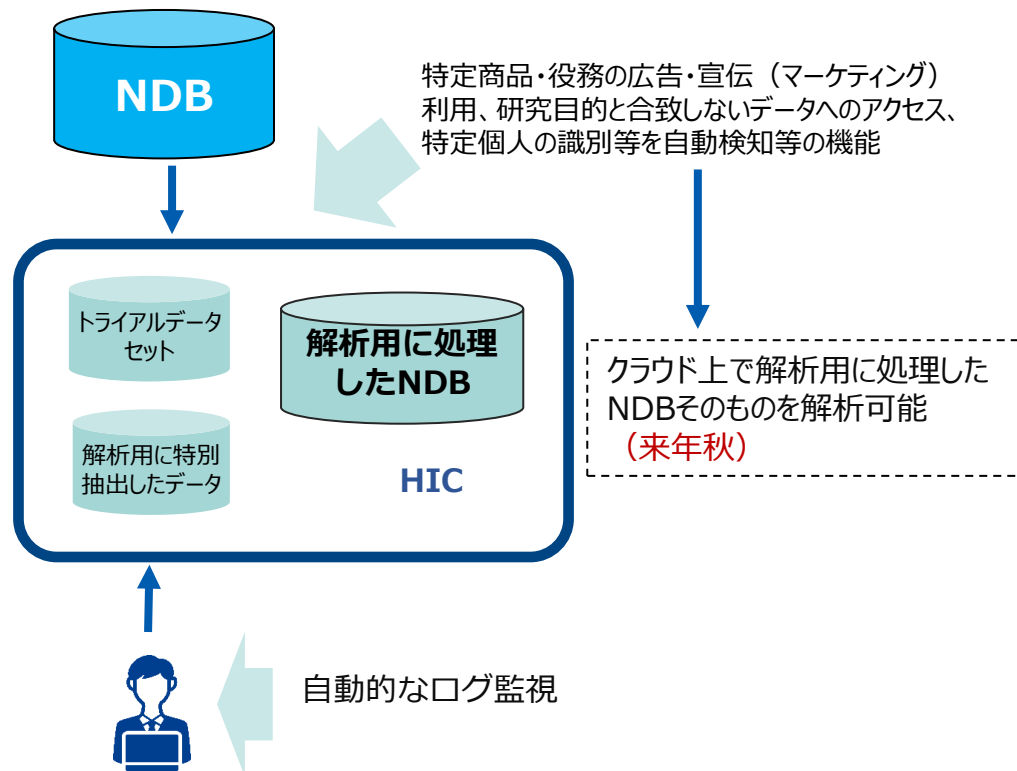
NDBデータ提供の抜本的見直し（概要）

- **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、
 - ・ **本年6月**、NDBサンプルデータを厚労省HPに公表
 - ・ **本年秋**、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、
 - ・ **来年秋**、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

【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サンプルデータ（データ見本）の公開

NDBサンプルデータ（データ見本）

サンプルデータ（データ見本）は、サンプリングデータセットから100～200レセプト程度（数MB）を抽出し、個人特定につながるようなコードを加工したうえで、厚労省のHP上に公開する

	サンプリングデータセット	トライアルデータセット (検討中)	サンプルデータ（データ見本）
内容	1ヶ月分（1月/4月/7月/10月診療分）の匿名レセプトデータから、入院診療10%、外来診療1%を抽出し、高額レセプトの削除等の匿名化処理を行ったプリセットデータ	サンプリングデータセットから、さらに低頻度のコードが含まれるレセプト自体を削除したプリセットデータ	サンプリングデータセットから100-200レセプトを抽出したもの
目的	<ul style="list-style-type: none">承諾された研究目的の範囲内で探索的研究を行うことが可能	<ul style="list-style-type: none">承諾された研究目的の範囲内で探索的な研究を行うことが可能。本格利用前のデータ構造の理解	<ul style="list-style-type: none">本格利用前のデータ構造の理解
利用方法等	<ul style="list-style-type: none">審議会の承諾が必要承諾後、50日程度で提供数GB/月	<ul style="list-style-type: none">審議会の承諾が必要数GB/月	<ul style="list-style-type: none">審議会の承諾は不要自由にダウンロード可厚生労働省のHP上に公開、数MB

NDBサンプルデータ（データ見本）のイメージ

医科入院外レセプトの例

NDBの格納項目



高確法における匿名加工

- ① 提供の都度、ハッシュ化のアルゴリズムを変える
- ② 特異な情報を削除等
→100歳以上一括りなど
- ③ 機微な情報の加工
→医療機関コード等は匿名で提供等



レセプトデータ（イメージ）

1,2,0,IR,1,13,1,9999913,,**サンプル病院**,202005,00,03-9999-9999
 1,3,0,RE,9,1117,202004,**サンプル**
 丸,1,19490208,,20200304,,,29,,0009,,,,,
 10,,,,,
 1,4,0,HO,06132013,1 2 3 4 5 6 7,
 1,3,82255,,6,3990,,,57600,,,
 1,6,0,SY,1509003,20151102,1,,,01,
 1,7,0,SY,8841698,20151110,1,,,
 1,8,0,SY,2500015,20151110,1,,,
 1,9,0,SY,1509003,20151217,1,8048,,01,
 1,10,0,SY,8842624,20151222,1,,,01,
 1,11,0,SY,1101003,20151222,1,,,
 1,12,0,SY,6869044,20160530,1,,,

カバ番号等: sample-ika-079	受付番号: 2205-00-000-000000-00-0000	ページ番号: 000,001-000
診療報酬明細書 (医科入院外)	平成 22年 4月分	院号: 9999913
公費①	公費②	公費③
サンプル 79	保険者番号 06132013	記号・番号 123456779
保険医療機関名 東京都港区新橋	被保険者証記号番号	
氏名	生年月日	性別
傷病名 (主) 糖尿病	傷病行為コード	転帰
請求点数	診療行為	診療行為コード

- 診療年月
- 医療機関コード
- 保険者番号
- 保険医療機関名
- 被保険者証記号番号
- 氏名
- 生年月日
- 性別
- 保険者番号、被保険者番号記号番号、生年月日、性別をハッシュ化
- 氏名、生年月日、性別をハッシュ化
- 傷病名 → 傷病行為コード、転帰
- 診療行為 → 診療行為コード
12再診
13医学管理
14在宅
20投薬等
- 請求点数

郵便番号
市町村コード

高額療養費
自己負担限度額
(適用区分)

公費負担医療
レセプト
公費負担者番号

レセプト情報がどのようにNDBに格納されているのかがイメージしにくく、利活用の障壁となっている

(参考) NDBサンプリングデータセットの仕様

1ヶ月分で5-6GB

各年度の1月、4月、7月、10月について予め作成されており、専門委員会の審査後50日程度で提供可能

1. 高額レセプトの除去

「医科入院」では700,000点以上、「医科入院外」「調剤」では50,000点以上の高額レセプトを除外



2. 傷病名コード、医薬品コード、診療行為コード、診断群分類について、出現回数の少ないコードを匿名化

「DPC」「調剤」では出現回数の少ないものから総出現回数の0.1%

「医科入院」「医科入院外」では0.01%に達するまで匿名化（コードを9999999等に変更）し
点数を空欄とする



3. 母集団から性年齢別比を変えずに一部を抽出

「医科入院」「DPC」の10%、「医科入院外」「調剤」の1%

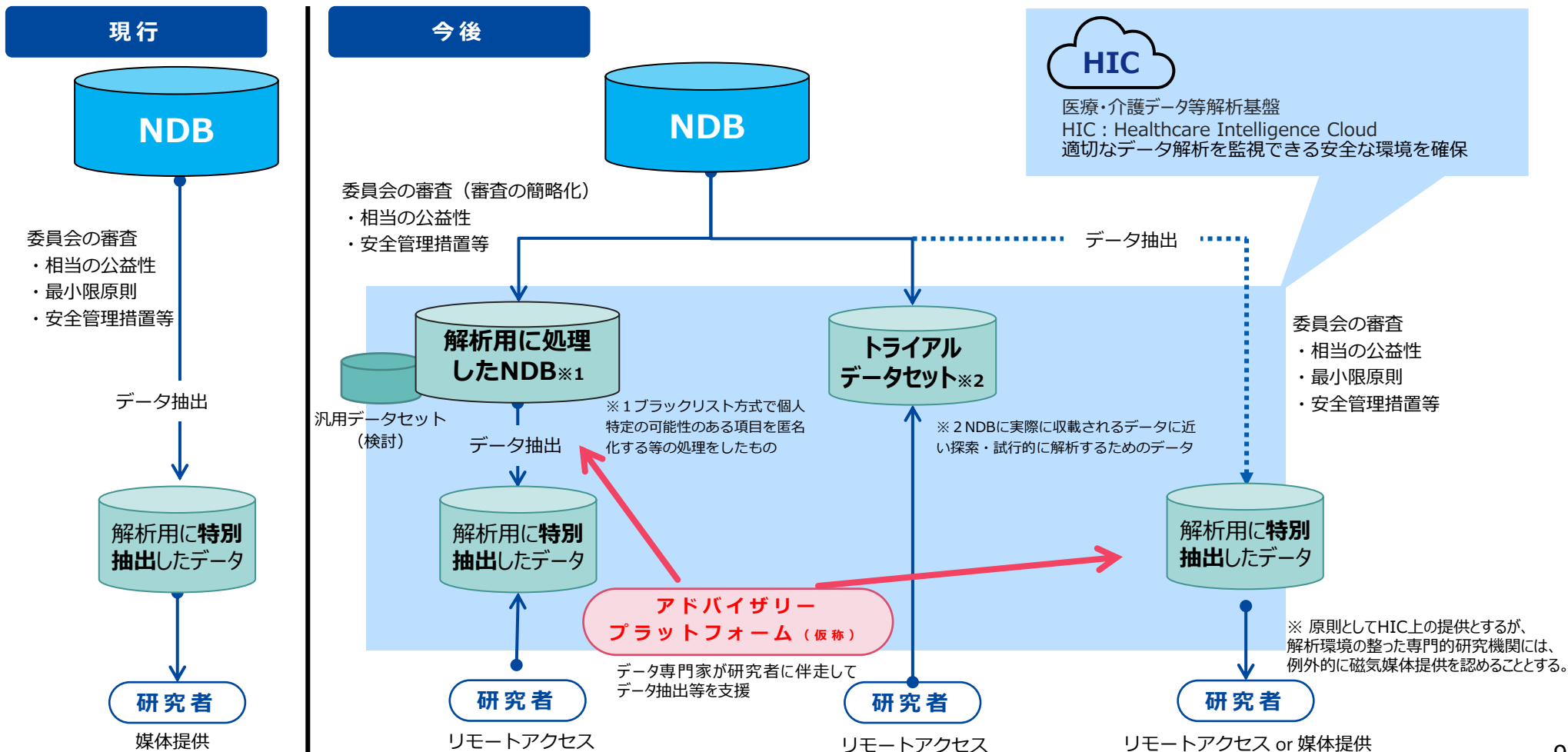


4. 機微なコードについて空欄化

患者の住所地、都道府県、保険者、「DPC」の医療機関別係数とその係数の影響を受ける情報
公費（KO）レコードの負担者番号等

医療・介護データ等解析基盤（HIC）の機能拡充（案）

- ▶ クラウド環境で整備中の医療・介護データ等解析基盤（HIC）の機能を拡充し、解析用に処理したNDB（ブラックリスト方式で個人特定の可能性がある項目を匿名化する等の処理をしたもの）にリモートアクセスする仕組みを構築。
- ▶ NDBの解析用の処理内容、HIC上での適切なデータ解析を監視できる安全な環境の確保、研究者が負担するHIC利用料等について検討のうえ、システム改修を行い、令和6年秋に試行運用を開始



令和5年度厚生労働科学研究について

- 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など技術的課題を検討するため、令和5年度中に、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。
- NDBオンサイトセンターの利用者、NDBで研究実績がある研究者、審査実務に精通した研究者で体制を構築し、レセプトデータの取扱経験のある者へのヒアリング等を通じて、データユーザーの観点、審査実務の観点からNDB利活用促進に向けた検討を進める。

<主な検討事項>

以下について、ヒアリングやこれまで集積された知見等を踏まえ、具体的な提言を行う。

- HICにおける効率的なデータ提供のあり方
- 新たな提供体系において必要な安全管理措置
- データ活用を促進するための必要な支援のあり方

<研究体制>

◎山本 隆一	医療情報システム開発センター
今中 雄一	京都大学
今村 知明	公立大学法人 奈良県立医科大学
松田 晋哉	産業医科大学
康永 秀生	東京大学
満武 巨裕	医療経済研究機構

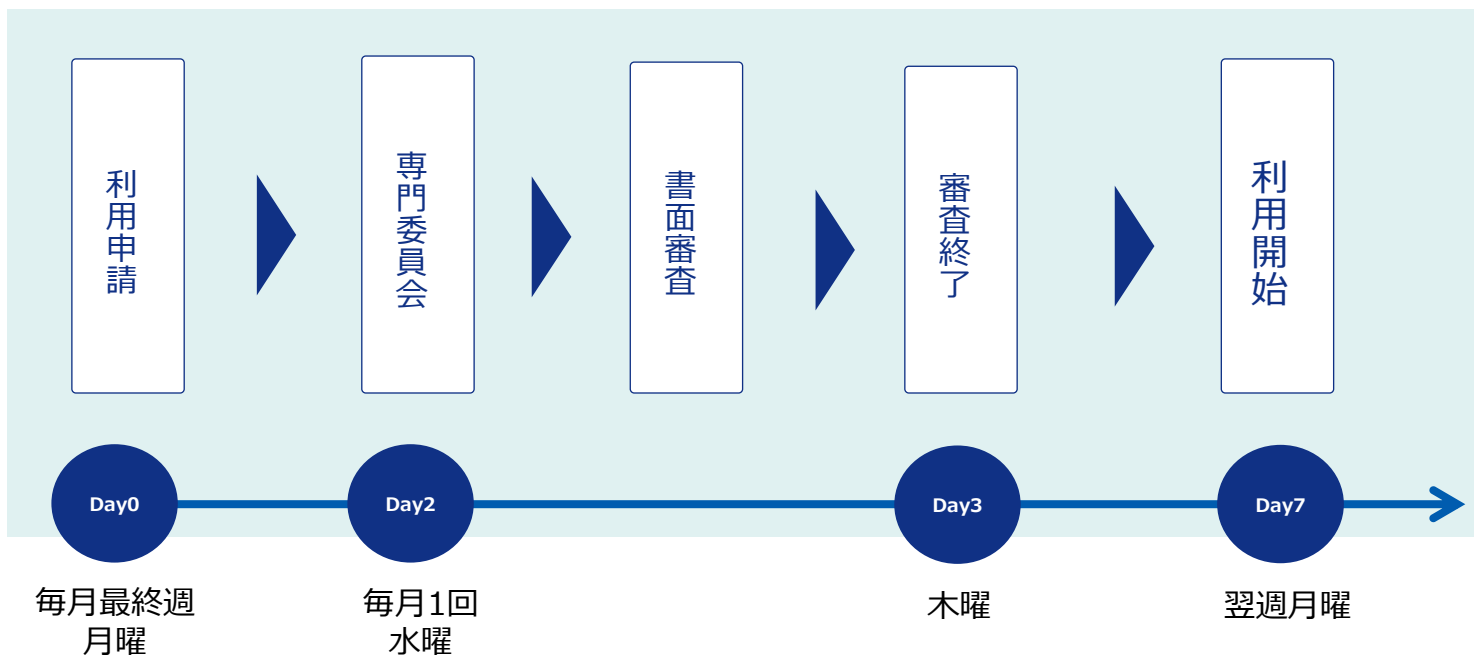
◎：研究代表

HIC機能拡充に伴う審査方法の見直し（案）

- 申請からデータ提供まで平均390日の現状に対し、申請〆切を毎月設定し、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

最短7日のイメージ※1

※1 研究者側の都合に要した期間は除く



- ポータルサイト上で申請

【承諾の場合】

- ポータルサイト上でID/PW提供
- 手数料の納付が必要

【不承諾・継続審査】

- ポータルサイト上で理由を通知
- 研究内容が不明確で判断できない場合等は対面審査

※ 例外的に磁気媒体の利用審査を行う場合は、対面審査

(参考) NDB利用の安全管理措置

1 NDBの収載

- 厚生労働省は、高齢者の医療の確保に関する法律（以下「高確法」）に基づき、医療費適正化計画の作成・実施・評価に資するため、保険者等からデータの提供を受け、NDBに収載している。
- 厚生労働省は、医療費適正化計画の作成・実施・評価に資する目的で、自ら調査・分析を行うとともに、国民保健の向上に資する目的で相当の公益性を有する分析等を行う者に対して提供することができる。

2 NDBの安全管理措置

- NDBは、高確法に基づき、個人特定できないよう氏名等を削除し、匿名化した形でデータを収集している。
- NDB利用者は、高確法に基づき、他の情報と照合等の禁止義務、利用後のデータ消去、安全管理措置、不当な目的利用等の禁止などの義務が課されている。
 - ※ NDB利用の成果物は、公表前に、個人情報保護の観点から、希少疾患など個人が特定されないか（最小集計単位が10未満となっているか）を確認などを講じている
- 厚生労働省は、法令違反などの疑いがある場合には、高確法に基づく立入検査・是正命令を行うことが可能。
 - ※ 不当な目的利用等の禁止義務や是正命令に違反した者には、罰則あり（1年以下の懲役・50万円以下の罰金）
 - ※ このほか、NDBの不適切利用が発生した場合、厚生労働省は、ガイドライン・利用規約に基づく利用停止等の措置を行うことが可能

(参考) N D B 利活用に関する閣議決定

○経済財政運営と改革の基本方針2023（令和5年6月16日閣議決定）

（社会保障分野における経済・財政一体改革の強化・推進）

- 医療DX推進本部において策定した工程表※1に基づき、医療DXの推進に向けた取組について必要な支援を行いつつ政府を挙げて確実に実現する。マイナンバーカードによるオンライン資格確認の用途拡大や正確なデータ登録の取組を進め、2024年秋に健康保険証を廃止する。レセプト・特定健診情報等に加え、介護保険、母子保健、予防接種、電子処方箋、電子カルテ等の医療介護全般にわたる情報を共有・交換できる「全国医療情報プラットフォーム」の創設及び電子カルテ情報の標準化等を進めるとともに、PHRとして本人が検査結果等を確認し、自らの健康づくりに活用できる仕組みを整備する。その他、新しい医療技術の開発や創薬のための医療情報の二次利活用、「診療報酬改定DX」による医療機関等の間接コスト等の軽減を進める。その際、医療DXに関連するシステム開発・運用主体の体制整備、電子処方箋の全国的な普及拡大に向けた環境整備、標準型電子カルテの整備、医療機関等におけるサイバーセキュリティ対策等を着実に実施する。

※1 「医療DXの推進に関する工程表」（令和5年6月2日医療DX推進本部決定）。

(参考) NDB利活用に関する閣議決定

○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に収録されたデータ(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. 厚生労働省は、NDBデータの利用を行おうとする者に対して、NDBデータの項目及びその構造等の理解を助け、NDBデータを効率的に解析し得るよう、そのサンプルデータを公開する。
 - b. 厚生労働省は、NDBデータの利用を行おうとする者が探索・試行的にデータ解析することを可能とするため、トライアルデータセット(NDBの各年1月、4月、7月及び10月分から無作為に数%程度抽出する等の処理をしたものをいう。以下同じ。)又は特別抽出(研究者等の指定した抽出条件に従ってNDBデータをNDBから抽出することをいう。)の承認を受け当該研究者等に提供されたデータに対する医療・介護データ等解析基盤(HIC: Healthcare Intelligence Cloud)を通じたリモートアクセス(国が指定する特定の施設に限定せず、研究者等の自宅や研究室等からセキュリティレベルを保ったまま調査票情報等を格納するシステムにアクセスし、分析・集計を行うことができるアクセス方式をいう。以下同じ。)による解析を可能とする。なお、トライアルデータセットの利用申請に関する審査については、匿名医療情報等の提供に関する専門委員会(以下「専門委員会」という。)における審査項目を減らすなど、審査を簡略化するものとする。
 - 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。

(参考) NDB利活用に関する閣議決定

○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

- (略)
- d. 厚生労働省は、NDBデータの利用の要件として高齢者医療確保法第16条の2に定める「相当の公益性を有すると認められる業務」について、特定の商品等の広告・宣伝を除く、医薬品や医療機器の創出又は改善に資する調査、研究又は開発（製薬企業を含む民間事業者等による医薬品安全性調査、市販後の有害事象のエビデンス収集等の研究を含む。）に利用可能であることを明確化する。
- e. 厚生労働省は、NDBデータの利用に関して、「匿名レセプト情報・匿名特定健診等情報の提供に関するガイドライン」（令和2年10月厚生労働省）において利用を行った研究者等に対して「他の研究や政策利用等を阻害するような特許の取得を禁止する」とされていることについて、当該記載は特許法（昭和34年法律第121号）第32条の不特許事由と同様の趣旨であり、NDBデータの利用による研究を基礎とする場合であっても、公の秩序、善良の風俗又は公衆の衛生を害するおそれがない限り、特許を受けることが可能であることを明確化する。
- f. 厚生労働省は、学会誌への投稿手続が進行している場合など一定の場合は利用期間の延長が可能であることを明確化する。あわせて、利用期間の延長手続によって延長可能な期間が運用上1年以下となっている現状に対し、必要に応じて2年以上の延長が認められることを明確化する。
- g. 厚生労働省は、NDBオープンデータ（診療行為、処方薬、健診項目等について、全国レベルで集計を行った集計表をいう。）について、個人情報の保護等を引き続き図るとともに、利用による研究等を精密化する観点から、①各セルにおける患者数の該当数値が0である場合にも公表データにマスキングを行っている現状について、当該セルの該当数値が0であることを明らかにすることによって、他のセルにおいて10未満の患者数であることが必然的に明らかとなるケースを自動的に判別するプログラム開発を行うことによって、マスキングの範囲を減らす運用を実施し、あわせて、②各セルにおける医療機関等の施設数が3未満となる場合についてマスキングを行っている現状について、マスキングを行わないことを基本的に検討し、結論を得る。
- h. (略)

d～f：令和5年秋措置 g：①令和5年度措置、②令和5年度検討・結論