

# NDBサンプルデータ（データ見本）について

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

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4月19日

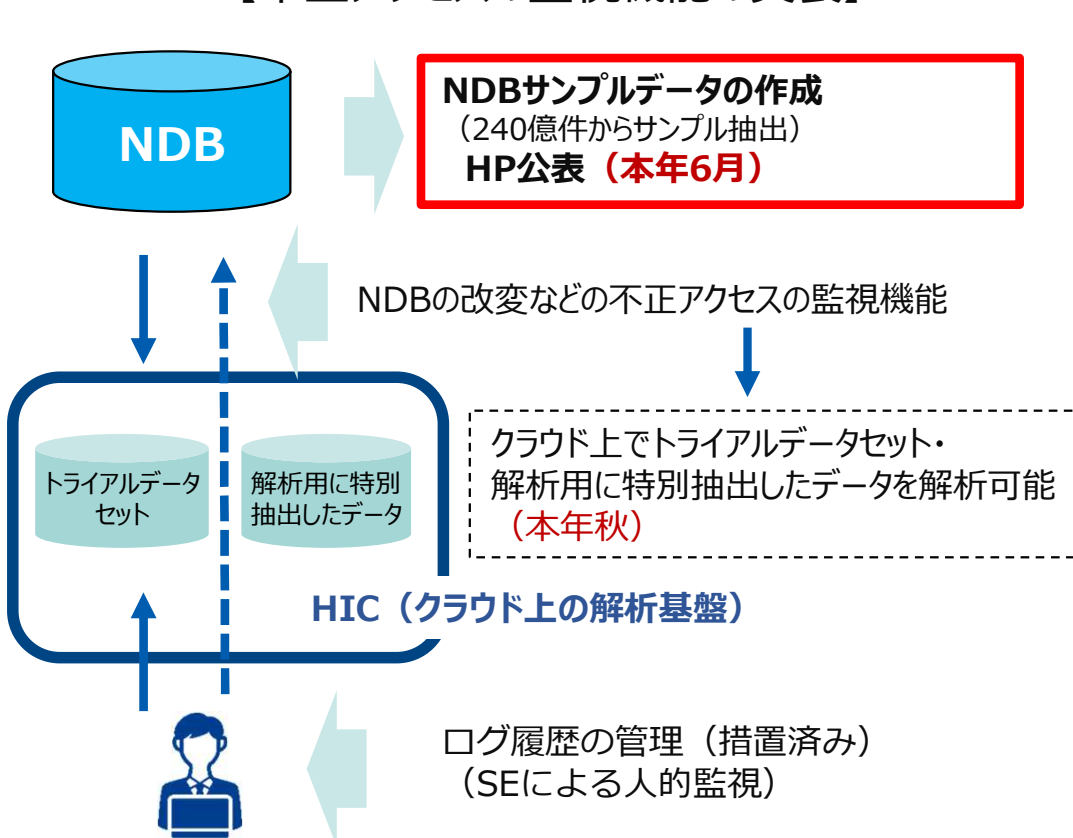
資料1

第15回匿名医療情報等の提供に関する専門委員会

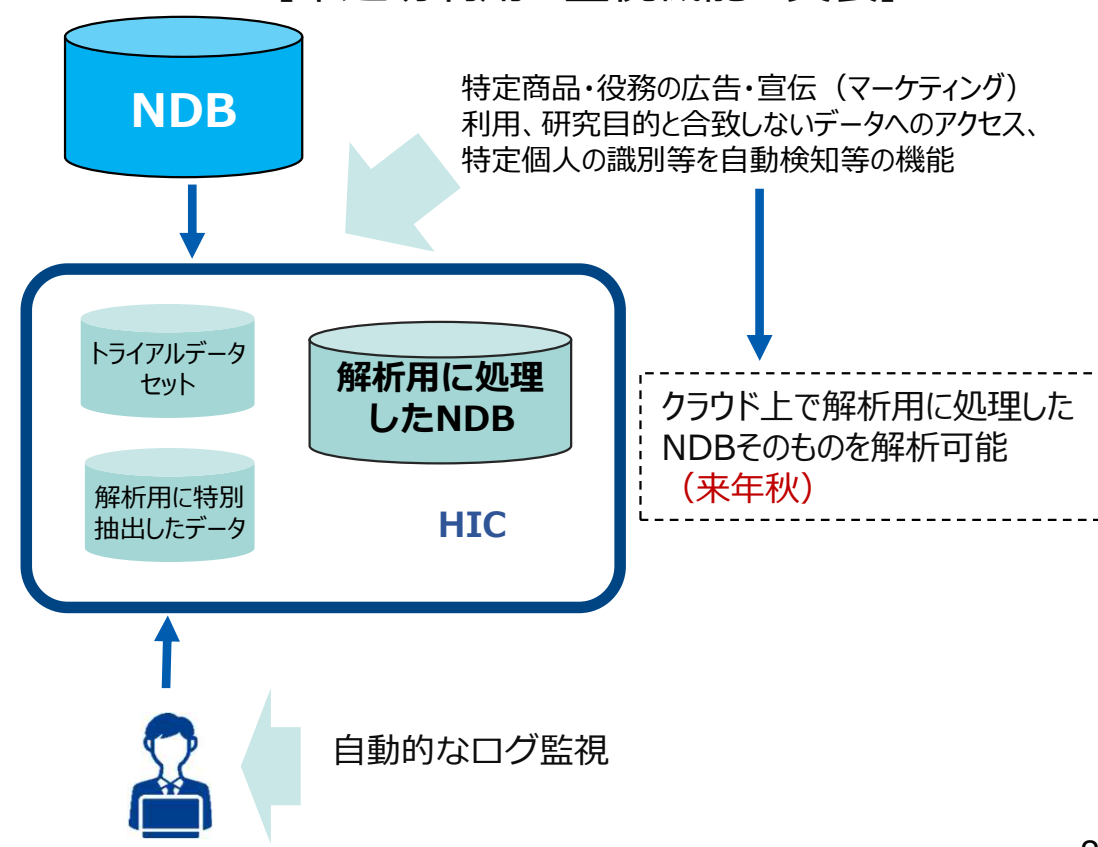
- **直ちに**、サンプルデータの作成、トライアルデータセットの作成、不正アクセス監視機能の実装に取り組み、
  - ・ **本年6月**、NDBサンプルデータを厚労省HPに公表
  - ・ **本年秋**、リモートアクセスでトライアルデータセット・解析用に特別抽出したデータを解析可能
- さらに、不適切利用等の監視機能やポータルサイトの機能拡充を開発・実装の上、
  - ・ **来年秋**、リモートアクセスの解析データを拡大
  - ・ **申請からデータ提供まで平均390日の現状に対し、申請〆切を毎月設定し、申請から最短7日で処理**

※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"><li>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</li></ul>	<ul style="list-style-type: none"><li>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。</li><li>本格利用前のデータ構造の理解</li></ul>	<ul style="list-style-type: none"><li>本格利用前のデータ構造の理解</li></ul>
利用方法等	<ul style="list-style-type: none"><li>審議会の承諾が必要</li><li>承諾後、50日程度で提供</li><li>数GB/月</li></ul>	<ul style="list-style-type: none"><li>審査会の承諾が必要</li><li>数GB/月</li></ul>	<ul style="list-style-type: none"><li>審査会の承諾は不要</li><li>自由にダウンロード可</li><li>厚生労働省のHP上に公開、数MB</li></ul>

## (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