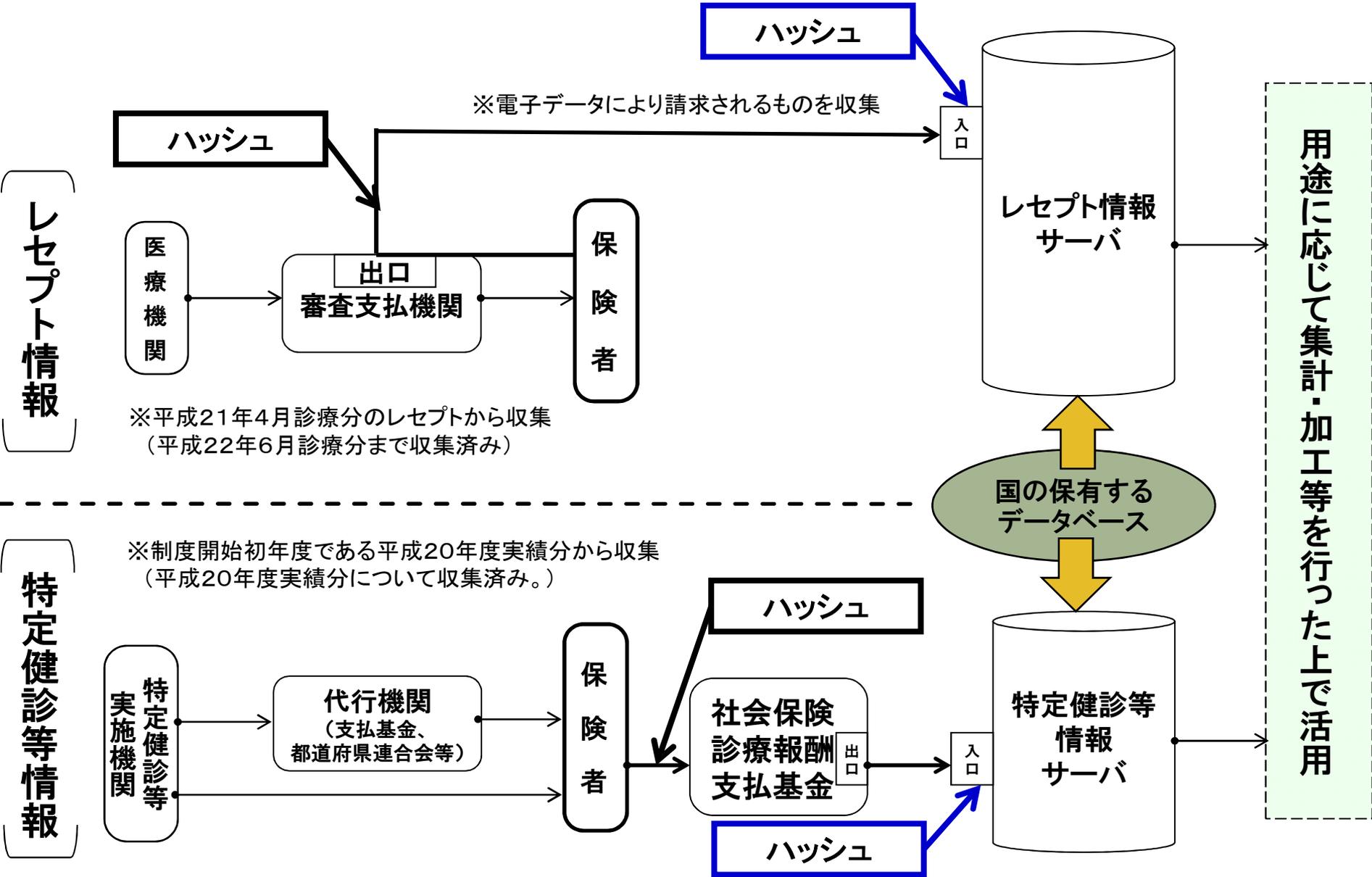


指定難病患者データベースと小児慢性特定疾病 児童等データベースの利活用の在り方について — その他のHEALTH データベースの利活用と課題から —

レセプト情報・特定健診等情報の収集経路



NDBの利用利活用

(平成20年度検討会報告を踏まえた仕組み)

高齢者医療確保法に基づく利用

厚生労働省保険局総務課
医療費適正化対策推進室

都道府県

医療費適正化計画の作成等
のための調査及び分析等

国による分析等



結果の公表

国が公表する結果のほか、都道府県が、国に対し、医療費適正化計画の評価等に必要な情報の提供を要請し、入手



都道府県による分析等

左記の本来目的以外の利用

厚生労働省内の他部局・他課室
関係省庁・自治体

左記以外の主体
(研究機関等)

医療サービスの質の向上等
を目指した正確なエビデンス
に基づく施策の推進

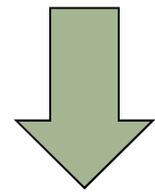
- 感染症などの疾患の実態把握に基づく施策
- 介護給付費と医療費の実態把握に基づく施策等

※所掌事務の遂行に必要な範囲内
であることが前提

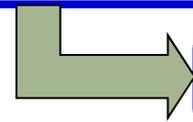


有識者会議における審査
※データ利用の目的や必要性について審査
※データ利用の目的としての「公益性の確保」が必要

- 左記のような施策に有益な分析・研究
- 学術研究の発展に資する目的で行う分析・研究



データ提供の可否について大臣に助言



大臣決定

レセプト情報・特定健診等データ提供に関するガイドライン

- **利用者の限定： 国、都道府県、研究開発独立行政法人（PMDAを含む）、医師会等、国の研究助成金を受けているもの。**
- **公益性の確認**
- **必要最小限のデータ要求であること。**
- **研究結果の公表ルール：公表形態はレセプト件数10以下、医療機関数で3以下に限定されてはいけない。**
- **安全管理基準：医療情報システムの安全管理に準ずる。**

研究者にとって、電子化情報の安全管理は容易ではない

NDBの現状

- **120億件以上のレセプトデータと2億件の特定健診特定保健指導データ**
- **特別抽出データの提供**
- **サンプリングデータセット:**
 - **外来の1%および入院の10%のレセプトベースのサンプリング**
 - **一ヶ月分のデータ(1月、4月、8月、10月)**
ただし医科と薬科の連結データは薬科の翌月分も含む
 - **出現頻度0.1%医科の病名、医療行為はダミーに置き換え**
- **ベーシックデータセット:**
 - **患者ベースで5%にサンプリング、同一患者のレセプトは連結している。**
- **特別抽出、サンプリングデータセットを中心に約150の研究プロジェクトに提供。**
- **100以上の査読付き学術論文がすでに発表されている。**
- **NDBオープンデータの公開(2016~)**



Mid-net Project

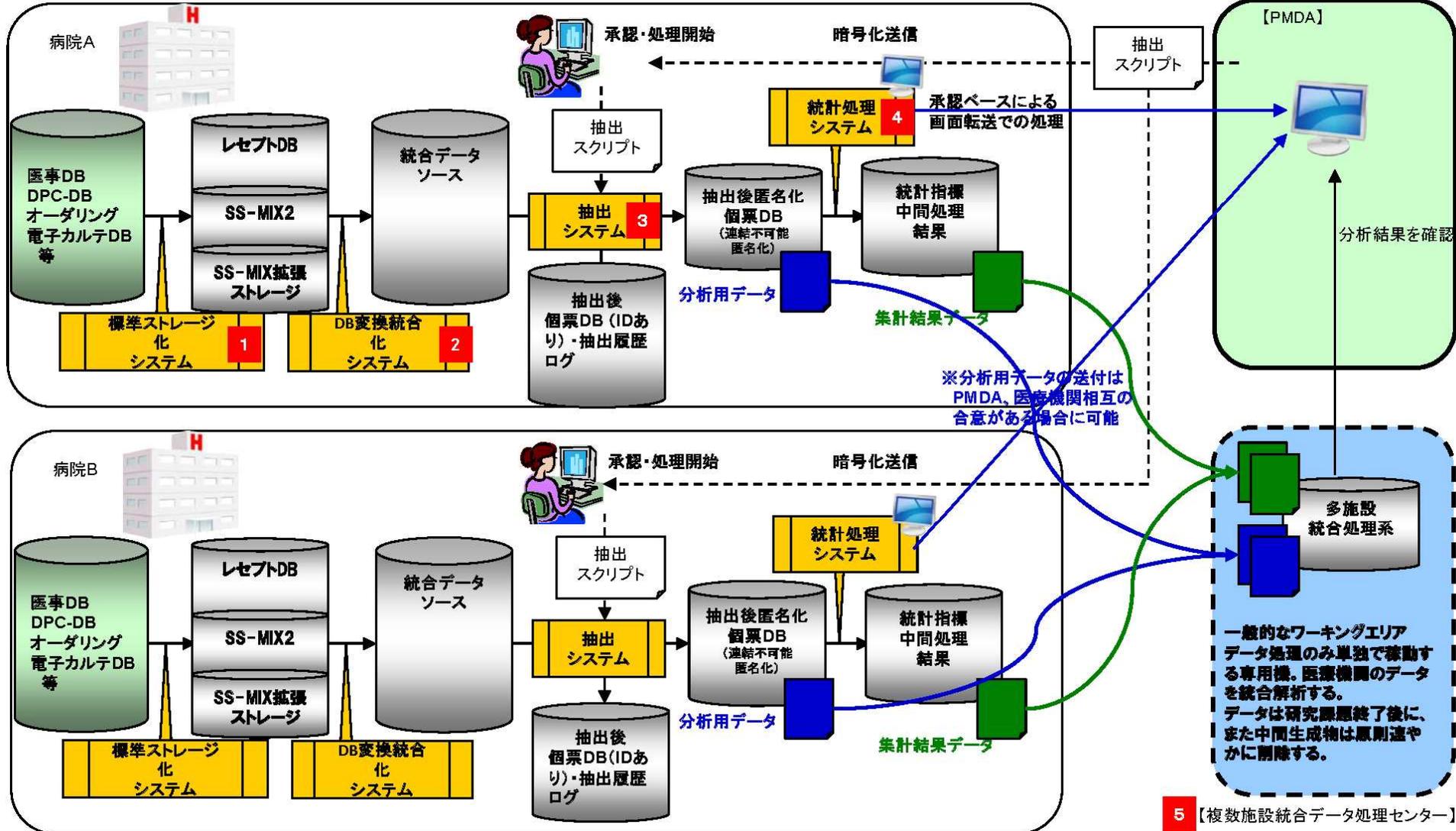
【医療機関】

集計結果データ

■ 統合データソースから抽出した集計結果情報(主に、発生頻度、件数、記述統計、クロス集計結果等)

分析用データ

■ 統合データソースから抽出した分析用データ(主に、データをマージ・ソート処理が必要な統計処理等)



5 【複数施設統合データ処理センター】

利活用の流れと利活用ルール of 主な検討事項

利活用の流れ

利活用ルール of 主な検討事項

利活用の申出

有識者会議による事前審査

オンサイトセンターの利用

データの移動

解析結果の公表審査

利活用の終了

- 利活用が認められる目的の範囲について
- 有識者会議で事前に審査を行う事項について
- 承認された利活用の公表について
- 利活用の形態について(利活用者が自ら解析する形態か、PMDAに解析を委託する形態か)
- オンサイトセンターの設置について
- 外部へのデータの移動の可否について
- データ移動の手続きについて
- 公表審査の手続き・公表基準について
- 公衆衛生等の観点から公表が必要な場合の手続きについて
- 利活用終了後のデータの削除・保管について
- 利活用者がデータ保管を希望する場合の取扱いについて
- データの漏洩等の不適切な利活用への対応について

PMDAの有識者会議における事前審査

○MID-NETが、機微性の高い電子診療情報を扱うことや、公費を投じて構築されたものであることを踏まえ、利活用が適切なものであることを事前に有識者会議で審査を行う。

(事前審査の主な項目)

- **目的**:利活用の目的に合致しているか。
- **内容**:利活用の内容が目的を逸脱せず、かつ実施可能であるか。
- **期間**:利活用の期間が内容に照らして適切か。
- **利活用する情報の範囲**:目的に照らして必要最小限であるか。
- **情報の管理方法**:利活用者の所属する組織、データを取り扱う環境等において「医療情報システムの安全管理に関するガイドライン」に準じた措置が講じられているか。
- **結果の公表内容・公表方法**:目的に合致しているか、第三者によって個人や医療機関等が特定されないように措置が講じられているか。
- **利益相反の管理**:利活用者の所属する組織において適切に利益相反が管理されているか。

○透明性を確保する観点から、利活用が承認されれば、利活用の情報(利活用者、利活用名、調査・研究の概要等)をPMDAのホームページに公表する。

次世代医療基盤法



丁寧なオフアウト



PHR事業者

医療情報取扱事業者(医療機関等)

医療情報は独自の定義
死者の情報・遺伝情報への対応？

医療情報

認定匿名加工医療情報作成事業者

認定医療情報等取扱受託事業者

匿名加工もある程度は独自の定義
有用性の確保を配慮？

匿名加工医療情報

匿名加工医療情報取扱事業者



次世代医療基盤法における匿名加工医療情報の検討

匿名加工医療情報基準

匿名加工情報基準

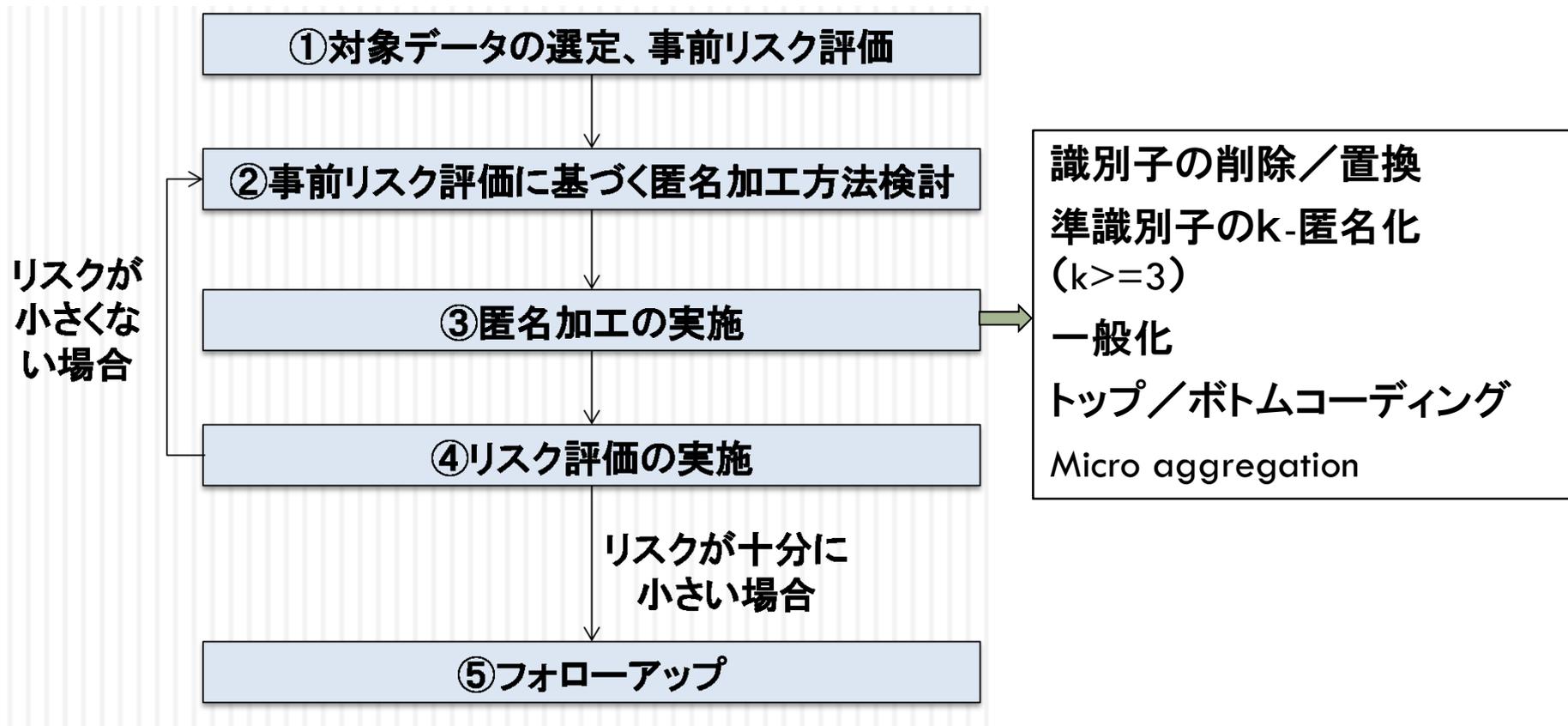
**個人情報委員会の匿名加工基準は
一般人基準**

**次世代医療基盤法の匿名
加工医療情報は一般の医療
従事者・医学研究者基準**

**個人情報委員会の基準に加えて
リスクベースの対策強化を求
めている。**

次世代医療基盤法における匿名加工医療情報の検討

データベースの特性評価



次世代医療基盤法における匿名加工医療情報の検討

分類	概要
識別子	個人に直接紐づく情報(氏名、被保険者番号等)
準識別子	複数を組み合わせることで個人の特定が可能な情報(生年月日、住所、所属組織等) ※医療機関コードは準識別子に該当すると考えられる。
静的属性	不変性が高い情報(身長、血液型、アレルギー、受診日等の日付等) 障害等の外見的に特徴に関する情報 ※不変性の高い慢性疾病情報の取扱いは要検討
半静的属性	一定期間、普遍性がある情報(体重等) 疾病、処置、投薬等の情報はここに該当する想定
動的属性	常に変化する情報(検査値、食事、その他診療に関する情報等)

分類	匿名加工方法の例
識別子	削除、もしくは非可逆な仮名化
準識別子	k-匿名性を満たすように一般化(生年月日→生年、住所→都道府県等)あるいはマイクロアグリゲーション データ項目削除を実施 医療機関コード等は属性(地理、規模等)を付加して特定できない形にコード変換
静的属性	数値はトップ・ボトムコーディング 一般化あるいはマイクロアグリゲーション 受診日等については、一般化やオフセット
半静的属性	数値はトップ・ボトムコーディング 機微な疾病等については必要ない場合は削除
動的属性	匿名加工不要であるが、必要に応じて数値はトップ・ボトムコーディング 異常値の重要性を考慮し、値の分布を見て上下数%を丸める等の加工を想定

ゲノム情報のリスクの低減

ゲノムデータ(細胞から採取されたデオキシリボ核酸(別名DNA)を構成する塩基の配列を文字列で表記したもの)のうち、全核ゲノムシーケンスデータ、全エクソームシーケンスデータ、全ゲノム塩基多型(single nucleotide polymorphism: SNP)データ、互いに独立な40箇所以上のSNPから構成されるシーケンスデータ、9座位以上の4塩基単位の繰り返し配列(short tandem repeat: STR)等の遺伝型情報により本人を認証することができるようにしたもの(個人情報の保護に関する法律についてのガイドライン(通則編)9頁)

- 「個人情報の保護に関する法律についてのガイドライン(通則編)」において上記が個人識別符号として規定されている。一連のシーケンスデータの中にこれらの要素が含まれている場合は個人識別情報として削除する必要がある。
- 上記以外のゲノムデータは個人識別符号には該当しないが、静的属性として、そのリスクを踏まえて匿名加工を行う必要がある。外見上の特徴を持つ疾患の発現可能性があるSNPであれば、そうでないSNPよりもリスクが高くなると考えられる。「互いに独立な30未満のSNPから構成されるシーケンスデータ、がん細胞等の体細胞変異、単一遺伝子疾患の原因遺伝子の(生殖細胞系列の)ホットスポット変異」については「個人識別性がほぼ無いと判断できる」レベルとされており、これらを考慮して匿名加工の方法を検討する必要がある。
- レアバリエントの中で、臨床的意義が明らかな希少性の高い難病等の原因変異については、他の情報との突合により容易に個人識別が可能なものとして、データの取扱には十分注意する必要がある。
- ゲノムデータに所見等を加えたゲノム情報については、発現率等の確率はあるものの静的属性として匿名加工を行うことが必要と考えられる。
- ALLELEの内、特徴的な家系図を示す場合においても、静的属性としてリスクに基づく匿名加工について検討する。

指定難病患者データベースと小児慢性特定疾病児童等データベースの 当面の利活用の在り方に係る論点(案)

基本方針において、指定難病患者データベースと小児慢性特定疾病児童等データベースの構築の目的が、個人情報の保護に万全を期し、難病、小児慢性特定疾病の調査及び研究の推進に資することとされていることを踏まえ、以下の点を整理してはどうか。

1. 情報の提供先についてどう考えるか？

- ① 厚生労働省または厚生労働省が補助を行う研究班に限定する。
- ② ①に加えて、都道府県、大学、国立研究開発法人、国立研究所、研究の実施に要する費用を国の行政機関から補助されているものも対象とする。
- ③ ②に加えて、企業等も研究の内容によっては認める。

2. 提供する情報の範囲についてどう考えるか？

- ① 個人が特定されない情報のみを提供する。
- ② 希少性から個人が特定されやすいという難病、小児慢性特定疾病特有の性質に配慮しつつ、原則として、匿名化した上で、研究内容等から判断して必要最小限の情報を提供する。その際、研究機関における情報の管理方法・体制や研究成果の公表方法等が個人情報保護の観点から適正かどうか審査をする。

3. 利活用の目的

- ① 難病、小児慢性特定疾病の研究の基礎資料としての利用に限定する。
- ② ①に加えて、臨床研究等への利活用も認める。

4. 審査会の設置

- ① 全件について審査会で審査する。
- ② 厚生労働省が補助を行う研究班に個人が特定されない情報のみを提供する場合など、審査会を省略できる場合を設ける。
- ③ 運用基準を明確化して個別の審査は行わない(審査会を設置しない)。

難病DB、小慢DBの提供を行うために(ご提案)

- **1. 情報の提供先についてどう考えるか？**
 - 行個法の制限*の中で提供可能な範囲を含め提供に係わる運用規則を制定。
 - 厚労省の補助に限ることについては慎重な議論が必要
- **2. 提供する情報の範囲についてどう考えるか？**
 - これまでに提供している厚労省の所管しているDBとの違いを明確にする。次スライド
- **3. 利活用の目的**
 - 次スライド
- **4. 審査会の設置**
 - ルール制定: 個人情報法、DBの利活用の専門家を含めた検討会で行うこととしてはどうか。
 - 提供審査: 提供主体は厚生労働省なので、外部審査委員会としてアドバイザリーボードの位置付けがよいか。

* 行個法 第四十四条の七

2 第四十四条の五第二項第三号の提案に係る行政機関非識別加工情報の本人の数が、行政機関非識別加工情報の効果的な活用の観点からみて個人情報保護委員会規則で定める数以上であり、かつ、提案に係る個人情報ファイルを構成する保有個人情報の本人の数以下であること。

行政機関の保有する個人情報の保護に関する法律第四章の二の規定による行政機関非識別加工情報の提供に関する規則
第5条 法第四十四条の七第一項第二号の個人情報保護委員会規則で定める数は、千人とする。

参考：(私見)難病DB、小慢DBの特殊性？

- 症例数の少ない疾患が含まれる。
例えばNDBでは人口2000以下の市町村別データは提供されない。
行個法・個人情報規則で、1000人以下の情報は非識別加工できない。
- 収容している疾患に関しては網羅性が高い。
- 遺伝性疾患が含まれる。
個人情報保護法制だけで十分か？
- 小慢DBから難病DBに同一患者が移行する場合がある。小慢DBと難病DBの結合が必要か？
目的・結合範囲の合理性、情報管理の安全性の確認を前提に容認すべき
- 網羅性は低いが、詳細な情報を格納しているコホート研究が別に存在する
場合があり、他のデータベースとの結合が必要な場合があると思われる。
結合をあらかじめ認めるか？（NDB、がん登録では禁止）
目的・結合範囲の合理性、情報管理の安全性の確認を前提に容認すべき