

次世代医療基盤法検討WG 提出資料

医療データに関する海外事例調査

株式会社野村総合研究所
コンサルティング事業本部
ヘルスケア・サービスコンサルティング部

令和4年2月15日

NRI

Share the Next Values!






はじめに




- 次世代医療基盤法の今後を検討するにあたって、先進的な事例、もしくは今後の論点において参考となる事例を、海外調査からご紹介させて頂く。医療データの二次利用が進んでいると言われているイギリス、アメリカ、フィンランドを中心に調査した。
- 調査項目は下記の通り。

項目	論点
匿名加工医療情報の主要DB	<ul style="list-style-type: none">• 主要なDBの規模、収集項目、収集方法、運営主体、用途等で国内外の違い• 諸外国における民間DBのビジネスモデル、実績
公的DBにおける死亡情報の取り扱い	<ul style="list-style-type: none">• 諸外国において、死亡情報（死亡日・死因等）はどの程度活用されているのか• されているのであれば、その実例と収集・結合方法
匿名加工医療情報のユースケース	<ul style="list-style-type: none">• 諸外国での代表的な用途や件数
RWDの薬事目的での利用に関する状況	<ul style="list-style-type: none">• 諸外国において、薬事承認において、使用されているRWDは匿名情報か否か、及びその課題
匿名加工医療情報の扱い・定義	<ul style="list-style-type: none">• 匿名化の基準の国内外での違い
ゲノム情報の取り扱い	<ul style="list-style-type: none">• 二次利用において、ゲノム情報と他の医療情報（レセプト、電カル等）との取扱いの違い• ゲノム情報の利活用に関する同意の必要性に関する国内外の違い• 用途等の制限に関する国内外の違い

各国とも国主導で大規模なDBを保有している。イギリス、フィンランドのDBは商用利用も可能であり、匿名加工医療情報の利活用も盛んに行われている。

	 アメリカ	 イギリス	 フィンランド	
DB名	CMS data	NHS Digital	CPRD	KanTa
運営主体	CMS(Centers for Medicare & Medicaid Services)：保健福祉省傘下の公的機関。二次利用は、ResDAC (ミネソタ大学、ボストン大学、ダートマス医科大学、及びモアハウス医科大学のコンソーシアム)が運営	NHS Digital: NHS傘下の、医療と社会福祉に関連するデータ集積、加工、配布、分析を行うことを目的として公的組織で、2013年に設立。	CPRD：保健省の執行機関であるMedicines & Healthcare products Regulatory Agency (MHRA)と国の臨床研究補助金の助成機関であるNational Institute for Health Research (NIHR)との共同出資で設立	社会保険庁(Kela)
概要	メディケア受給者の主にはレセプト情報の二次利用を目的としたDB	二次医療圏以上のHPの診療システムを介したDB。 <div style="border: 1px solid black; padding: 2px; text-align: center;">両者のデータは要請に応じて、結合することも可能</div>	一次医療圏のGPの診療システムを介したDB。	地域を横断して、中央的にデータを収集・公開する全国規模のEHR。二次利用もされている。
顕名/匿名	顕名/匿名	顕名/匿名（利用例の79%が匿名データ）	匿名	顕名/匿名
規模	5,300万人分 (メディケア受給者)	-	6,000万人分	570万人分 (全国民以上)
データ提出義務	自動的	義務（法律上は提出を強制する権利を有する）	義務ではない（同意したGPからデータ収集。以下便益を訴求） ・治験参加による収入増加 ・医療品質向上のためのレポート受領（ベンチマーク等） ・治療ガイドラインへの反映	義務
主な項目	レセプト	電子カルテ レセプト	電子カルテ レセプト	電子カルテ レセプト
用途制限	個人識別情報、識別不能個人レベルデータは研究目的限定	商用利用も可	商用利用も可	商用利用も可

英国、米国のEHRデータは死亡情報を含んだ匿名加工医療情報を提供している。
 フィンランドは要請に応じて、死亡情報と医療情報との結合が可能。

	 アメリカ	 イギリス	 フィンランド
DB名	CMS data	・NHS Digital ・CPRD	KanTa
死亡情報の有無			
死亡日	○	○	×（但し、要請があれば結合可能）
死因	○	○	×（但し、要請があれば結合可能）
死亡情報の連結方法	複数のソースから死亡情報は取得すると明記。主要なものとして以下が例示。 ・メディケアの保険適格性の管理をする Medicare Common Working Fileからの請求データ ・故人の家族によって提出された死亡に関するオンラインデータ ・社会保障局（SSA）、Railroad Retirement Board (RRB) から集められた受給者に関する情報	Local Registration Serviceが Registration Online Systemに死亡情報を登録。その情報が、検証された後、ONS（Office for National Statistics：国家統計局）に蓄積され、ONSからCPRDのデータに連結される。 ※死因のみは、医師のMedical Certificate of Cause of Deathを基にしている	フィンランドにおいては、複数のレジストリからデータを抽出して研究に活用する場合には、Findata※に依頼。死亡日データはDigital and Population Information Agency (DVV)、死因についてはStatistics Finlandがそれぞれ保有しており、Findataが窓口となって、KanTaと死亡情報データを結合し、提供する
注意	死亡日の正確性が検証されているのは99%。 死亡情報個人情報を含むRIF、識別不能個人レベルデータのLDSに死亡情報は含まれる。	新型コロナウイルスの影響で、死因のデータの結合に大幅な遅れが生じている	※FinDataは2019年に施行されたAct on the Secondary Use of Health and Social Dataに基づいて、国立健康福祉研究所(THL)の下に設置されたデータ利用許認可機関。

米国では民間のデータベースも広く利用されており、例えば、がん研究に特化した匿名化データベースを構築しているFlatiron社が有名で、約300万人分の匿名化されたDBを有している。

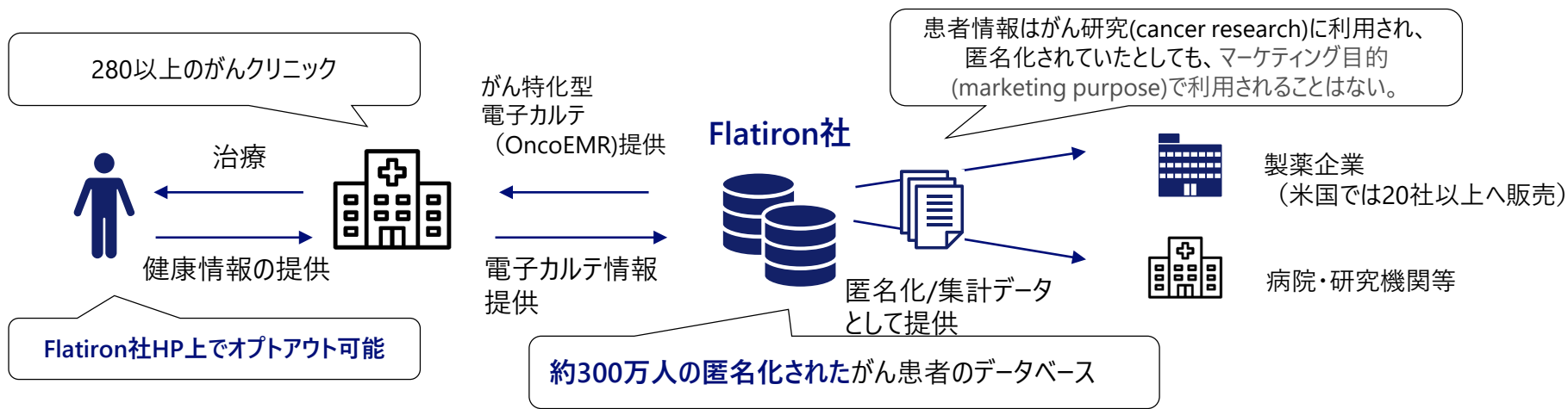
Flatiron社概要

創立年度	2012年
調達資金	\$313 million
株主	Roche (100%) 2018年2月より 2018年2月に2,000億円(19億ドル) で買収

Flatiron社データベース概要

データ利用者	病院・研究機関・製薬会社等
データ提供者	280以上のがん治療施設
保有データベース	約300万人のがん患者の匿名化されたデータベース
薬事承認事例	医薬品14品目の承認申請データに使用実績がある

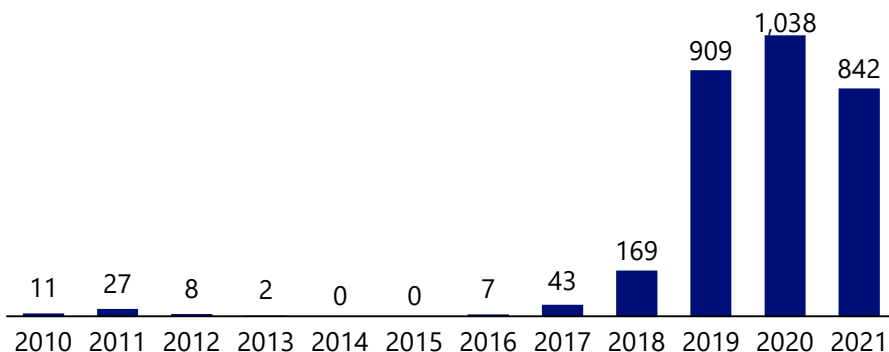
Flatiron社によるデータ収集～提供の流れ



匿名加工医療情報のユースケース 英国（NHS Digital）の例

英国NHS Digitalのデータは、公表されているだけで2010年から3,000件以上の利用実績があり、近年著しい増加が見られる。幅広いタイプの利用者が存在するが、NHS組織、アカデミア、地方政府が多い。

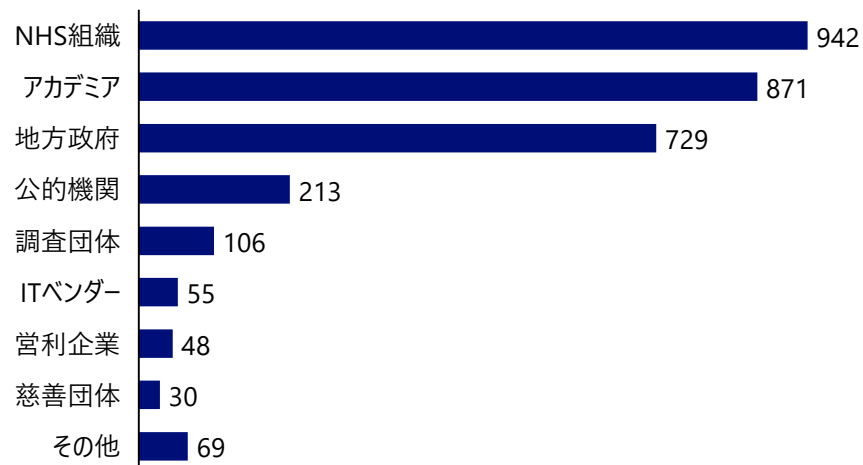
年別利用契約数（N=3,063）



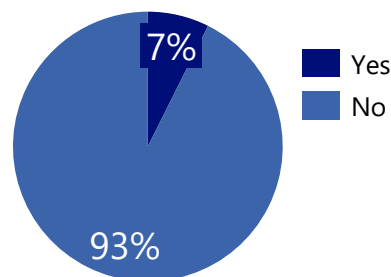
2019年から利用契約数が急増している理由として、以下が挙げられる

- 2018年5月に施行されたData Protection Actにより、データの取り扱いに関する法的な解釈が明確になったことで、NHS Digitalのデータ利用が加速した
- 2019年にシステム（データ・プロセッシング・サービス）が刷新され、医療機関からNHS Digitalへのデータ送信が簡易化された。また、NHS Digitalから利活用者側に対してデータをより早く、豊富なデータ項目を柔軟に提供できるようになった。

申請主体別利用契約数（N=3,063） 2010年～2022年



契約数別商用利用(Commercial Use)割合（N=3,063）



商用利用といっても、保険、マーケティング、製品・サービスのプロモーション・販売、マーケティングリサーチ、広告・宣伝目的で患者データが販売されることはない

匿名加工医療情報のユースケース 英国（NHS Digital）の例

地方政府は地域の公衆衛生施策のための分析のために利用しているケースが多い。
製薬企業は足元ではコロナワクチンの治験の計画策定のために利用していることが多い。

NHS Digital匿名加工医療情報の利活用例

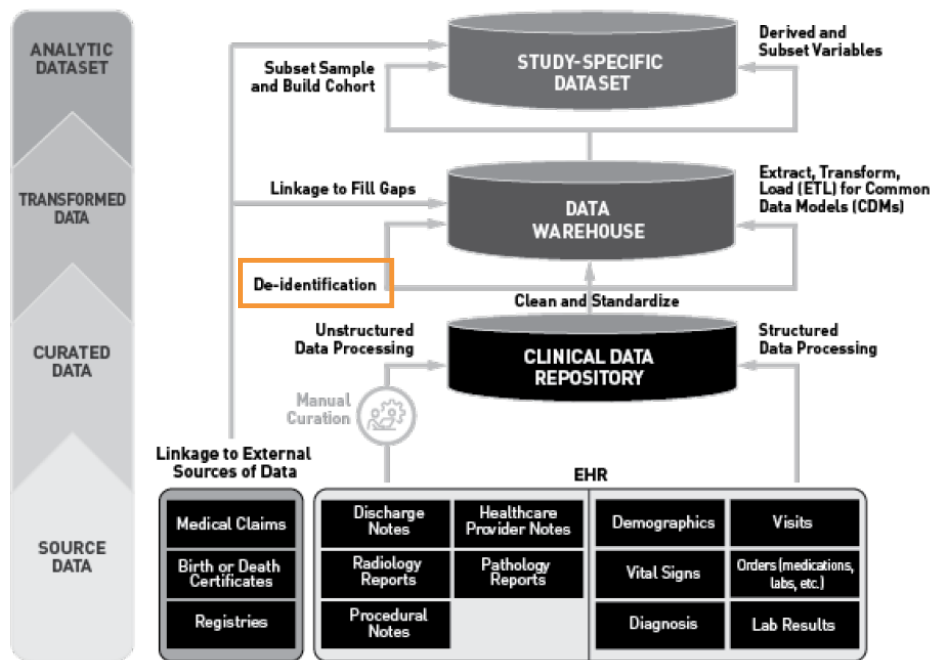
利用者種類	利用者	商用利用	匿名/顕名	利用目的
地方政府	BRIGHTON AND HOVE CITY COUNCIL	－	匿名	疾病の発症率・有病率・治療パターン・再入院率・地域別・人口別の治療へのアクセス等に関する傾向に関して分析し、公衆衛生の向上・達成のため利用
	LONDON BOROUGH OF ENFIELD	－	匿名	出生・死亡データ（特定可能情報）を、地域における死亡率の分析や施策の評価のため利用
	SALFORD CITY COUNCIL	－	匿名	仮名化されたレセプトデータを、疾患の発症率、公衆衛生に対するリスク・予防ケアサービスの需要や利用・人口における健康アウトカムの違い・地域間の疾患のリスク要因の違い等について分析、公衆衛生の向上のため利用
	DUDLEY METROPOLITAN BOROUGH COUNCIL	－	匿名	NHS Digitalが収集しているCOVID-19に関する患者情報（診断・治療・ワクチン接種状況等）を、COVID-19の感染拡大の防止に資する施策検討のため利用
ITベンダー	IQVIA	○	匿名	仮名化された医療データを統計的に分析・視覚化等の加工
	Methods Analytics	○	匿名	データを活用した統計情報の提供
	Health IQ	○	匿名	疾病の負担や治療の効率化改善等を目的とした研究の実施
	Harvey Walsh	○	匿名	データを分析した各種レポート等の作成
非営利団体	RCOG	－	匿名・顕名	英国の出産に関するコホート研究の実施
	Health Data Research UK	－	匿名	COVID-19の心血管疾患への影響を明らかにする
	UK Biobank	－	顕名	ガン、心臓病、脳卒中、糖尿病等の予防・診断の向上
製薬企業	Novavax	○	匿名	COVID-19のワクチン治験参加候補者数の事前調査
	AstraZeneca	○	匿名	COVID-19のワクチン治験参加候補者数の事前調査

RWDの薬事目的での利用に関する状況

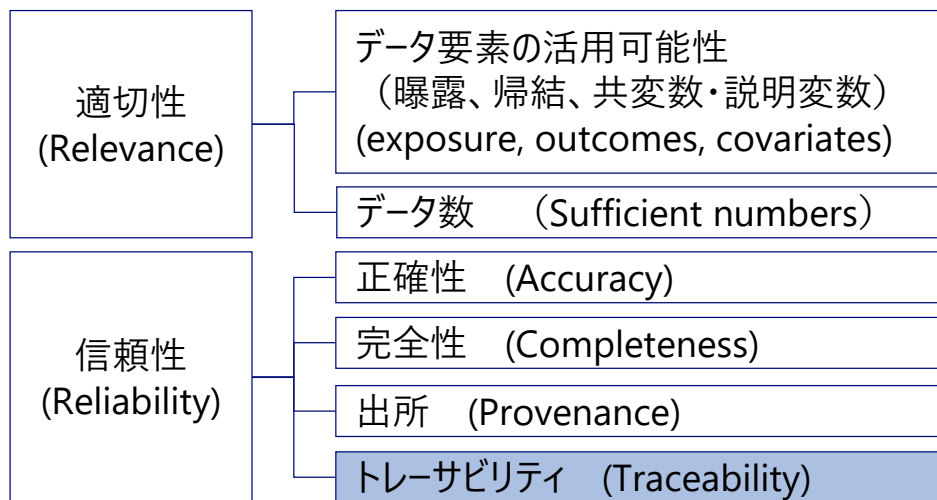
米国FDAはドラフトガイドラインにて、薬事目的においてDe-identifiedされたRWDデータの活用も想定している。一方で、薬事利用時には患者が再識別できることが望ましいとの記載もあり、不可逆的な匿名加工データの扱いについては今後の検討を見守る必要がある。

- 2016年12月に署名された 21st Century Cure Actを受け、米国FDAはRWDの活用可能性の評価プログラムのフレームワークを作り、その一環でドラフト版ではあるが、RWDの利用に関する（拘束力のない）ガイドラインを公表している。

FDAが想定するEHRデータの成形フロー例



FDAが考える薬事での意思決定のためのRWDの重要要素



FDAは、患者データは、非識別化（De-identification）しつつも、オリジナルデータの中でユニークな患者を再識別できる特徴を有することを推奨するとの記載がある。

FDAのガイドラインでは非識別化（De-Identification）は必要に応じて実施することとの記載がある

匿名加工、仮名加工の定義は各国で大幅に異なり、明確には定義されていないことも多い。

匿名加工基準、仮名加工基準の考え方

	 アメリカ	 イギリス	 フィンランド
匿名加工基準	HIPAAでは、匿名化の要件は、統計学や科学的知識を持つ専門家が匿名化処理し、処理後のデータが個人を再識別できないことを書面にて入手すること、又は、個人又は個人の親族、雇用主、家族の識別子(18項目)の情報を取り除くこと(かつ、Covered Entityが残りの情報を使用して個人を特定できないこと)とされる。また、制限されたデータセット(Limited Dataset)を、研究、公衆衛生、ヘルスケア業務向けにのみ利用及び開示できると規定されている。	英国の個人情報保護法Data Protection Act2018では、匿名情報や匿名化を明確に定義してはいない。2012年11月に英国の個人情報保護法遵守の監督機関ICOが匿名化に関する行動規範 (code of practice)を公表しており、現在その更新版を準備中。そのなかで、匿名化の概念や考え方、匿名化が十分にされたか(識別可能性)のチェックフローなどを示している。	法的な定義は明確でないが、FINNISH SOCIAL SCIENCE DATA ARCHIVEの説明に準ずるとのこと。以下概要。 匿名データ:提供されたデータ、あるいは追加のデータと組み合わせたものから、相当程度の努力(reasonable effort)をもってしても個々のデータ単位(個人)を再識別することはできないもの。完全に匿名のデータは存在しないが、適切に実行された手順では相当な努力をもってしても個人を特定できない結果を得ることができる。匿名化とは、匿名性を達成するために使用される様々な技術やルールのことをさす。
仮名加工基準	米国の法令には仮名加工の概念は登場しない。	匿名加工と同様に、Data Protection Act2018では、仮名情報や仮名化を明確に定義してはいない。前述の行動規範 (code of practice)では仮名化の概念やケースを紹介しており、現在その更新版を準備中。	上記FINNISH SOCIAL SCIENCE DATA ARCHIVEの説明によれば、別のデータ単位は、追加の別個の情報なしに仮名データにもとづいて再識別することはできない、仮名化は、識別子を削除あるいは、識別子を仮名あるいはコードに置き換えた上で、(元の識別子と置き換えられたもの)それぞれは個別に保管され、技術的および組織的な手段によって保護される。追加の識別情報が存在する限り、そのデータは仮名状態にとどまる。

米国では、患者の同意無しに個人情報を提供するには匿名化が必要で、除去すべき識別子は下記の通り。また、匿名化とは別に、個人の許諾を得ずに保護対象の保健情報を使用又は開示することが許可される「Limited Dataset(限定されたデータセット)」という概念が存在。

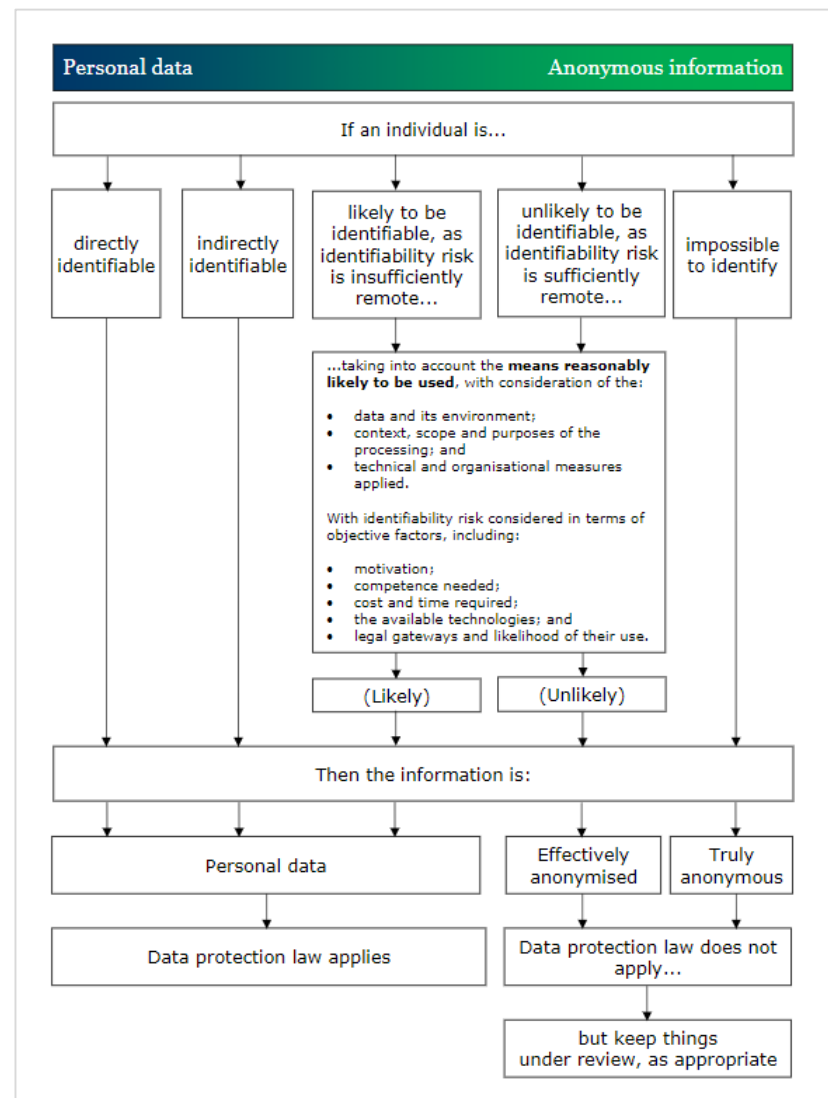
- 研究や公衆衛生等の目的の場合は、匿名化よりも除去すべき情報が少ない「Limited Dataset(限定されたデータセット)」を本人同意なしで使用又は開示が可能。開示する場合は、開示先とDate Use Agreementを締結し、開示先から上記で記載している目的以外にはデータを使用しないという確約を得る必要がある。

	De-identification	Limited Dataset
名前	✓	✓
住所・郵便番号	✓	✓
生年月日や入院等全ての日付データ	✓	
電話番号	✓	✓
ファックス番号	✓	✓
電子メールアドレス	✓	✓
社会保障番号	✓	✓
医療記録番号	✓	✓
ヘルスプラン受診番号	✓	✓
アカウント番号	✓	✓
ライセンス番号	✓	✓
ナンバープレートを含む車両識別子とシリアル番号	✓	✓
デバイス識別子とシリアル番号	✓	✓
URL	✓	✓
IP番号	✓	✓
指紋及び声紋を含めた生体認証識別子	✓	✓
フルフェイスの写真画像及び同等の画像	✓	✓
その他のユニークなID番号、文字、コードなど	✓	

英国の個人情報保護法の遵守監督機関ICOは匿名化・仮名化に関する行動規範を公表している。

- 英国の個人情報保護法Data Protection Act2018のSection3.では、「個人情報」とは、識別された、または識別可能な「生きている個人」に関連する情報としている。
- 「識別可能な生きている個人」とは、次のような項目を参照することで、直接または間接的に識別可能な「生きている個人」を指している。
 - (a) an identifier such as a name, an identification number, location data or an online identifier, or
 - (b) one or more factors specific to the physical, physiological, genetic, mental,
- 上記のDPA2018では、匿名情報や匿名化を明確に定義してはいないが、英国の個人情報保護法の遵守監督機関であるICO (Information Commissioner’s Office) は2012年に匿名化に関する行動規範 (code of practice) を発表しており、現在その更新版を準備中。そのなかで、データセットから直接識別子を削除するだけでは、効果的な匿名化を保証するには不十分であるなど、匿名化の概念や考え方、また匿名化が十分になされたか (識別可能性) のチェックフローなどを示している。




識別可能性評価のチェックフロー



ゲノム情報の取り扱い

諸外国においても、国が運営するゲノムデータは独立したバイオバンクに格納されており、国家レベルのEHRデータベースとは分離されているものの、バイオバンク側でEHRデータを有している。

- 逆に、国主導のEHRデータベース（英: NHS Digital、CRPD、米: CMS medical data、フィンランド: KanTa）にはゲノムデータは含まれない。




	 アメリカ	 イギリス	 フィンランド
DB名	All of Us	Genomics England	THL Biobank
DB概要	オバマ大統領の指揮の下、2015年にスタートしたPrecision Medicine Initiativeを引き継ぐもの。NIHにより、運営されている。	保健省が主体となって設立された国営企業であるGenomics Englandが運営。2012年、当時のキャメロン首相が、10万ゲノムプロジェクトの実施を公表し、それに基づいて設立された。	国立健康福祉研究所(THL)の一環として2014年に創設されたバイオバンク。
規模	First step（同意やサンプルの提供等）を完了した参加者：31万6千人分以上（2022年1月26日時点） 100万人規模を目指している。	2018年には10万ゲノムの解読が完了し、500万ゲノムプロジェクトに拡大中	22万人 （2022年1月10日時点）
対象者	人口構成を反映した一般市民 （特別な割付はない）	希少難病とがんの患者 ※UK BioBankは一般市民	人口構成を反映した一般市民 （特別な割付はない）
EHRのDBとの関連性	なし （All of Us側で、EHRデータを保有し、ゲノムデータとセットで提供している。但し、All of Us全参加者のEHRデータが存在するわけではなく、27万3千人（86%）分に留まる）	あり （Genomics England側で、EHRデータを保有し、ゲノムデータとセットで提供している。EHRデータはNHS Digitalからリンクされている）	なし （THL Biobank側で、登録者の病歴等のデータは保有しているとの記載あり）

※なお、上記のバイオバンクは、各国の代表的、且つ国主導のバイオバンクであり、他にもバイオバンクは複数存在する

出所） Genomics England, All of us, THL Biobank公開情報、




「諸外国におけるゲノム医療の制度・体制・運用等に関する調査（概要版）」2019年12月内閣官房健康・医療戦略室委託事業デロイト・トーマツコンサルティング合同会社

諸外国において、ゲノムデータの利用については個人同意を前提としている。

	 アメリカ	 イギリス	 フィンランド
DB名	All of Us	Genomics England	THL Biobank
ICのカバー対象	<p>包括的なICを取得</p> <ul style="list-style-type: none"> 取得したサンプルに加え、EHR等診療情報、外部から取得可能なデータ、住んでいる地域等のデータを提供することの同意を取得 All of Usがデータを活用することはもちろん、外部機関がデータを活用することについての同意を取得 当人のサンプルが枯渇しそうな場合は、サンプルの提供を再度依頼する可能性についての同意を取得 幅広い研究トピックに使える旨の同意を取得 	<p>包括的なICを取得</p> <ul style="list-style-type: none"> 取得したサンプルに加え、EHR等診療情報、サンプル取得以前に取得した腫瘍細胞、ならびに死後を含む将来のデータへのアクセスについての同意を取得 Genomics Englandがデータを活用することはもちろん、外部機関がデータを活用することについての同意を取得 サンプルから研究者が収入を得る可能性があること、サンプル提供者が新しい治療の開発から財務的メリットを享受できないことへの同意を取得 	<p>包括的なICを取得</p> <ul style="list-style-type: none"> サンプル、及びサンプルやドナーに関する情報のバイオバンク内への格納、及び研究への利活用の同意を取得 追加情報のリクエストがあった場合、またサンプルの分析によって、ドナーの将来の検査、診断の必要性に関するいかなる情報が明らかになった場合の、コンタクトに関する同意 国のレジスタデータ等とリンケージされる可能性があることの同意を取得 個人識別番号がない状態で研究者に引き渡されることの同意を取得
ICの取得プロセス	<ul style="list-style-type: none"> ICは、以下いずれかの方法で取得 <ul style="list-style-type: none"> 対象者がプログラムWeb サイトやモバイルアプリを介したeConsentプラットフォーム上で承諾 医療機関等で、直接トレーニングを受けたボランティアが対象者に説明をした上で、対象者が承諾 	<p>対象者が施設を訪問した際に、必ず質問に答えられるスタッフがいる状態で参加者にプロジェクトの説明を行い、同意書にサインをもらう。</p>	詳細不明
オプトアウト	参加する際にはICに同意したとしても、いつでもオプトアウトすることが可能	参加する際にはICに同意したとしても、いつでもオプトアウトすることが可能	参加する際にはICに同意したとしても、いつでもオプトアウトすることが可能

出所) Genomics England, All of us, THL Biobank公開情報、
「諸外国におけるゲノム医療の制度・体制・運用等に関する調査（概要版）」2019年12月内閣官房健康・医療戦略室委託事業デロイト・マツコンサルティング合同会社

調査対象各国のバイオバンクは民間企業の利用も可能である。

		 アメリカ	 イギリス	 フィンランド
DB名		All of Us	Genomics England	THL Biobank
収集データ	サンプル	・血液、唾液、尿等の生体サンプル	・血液、唾液、尿、DNAサンプル ・がん患者の組織サンプル	・血液、唾液、細胞、DNAサンプル
	問診情報 調査票	・医療履歴と生活習慣に関する調査票	・生活習慣に関する調査票	・生活習慣に関する調査票
	電子カルテ、 診療録データ	・病院等の医療機関の電子カルテ情報	・主治医が登録する診療データとEHRの 時系列データ	・検査結果（コレステロール値等）、及 びレジスタデータ（医療機関への入院や 処方等）
	その他	・Fitbit等のウェアラブルデバイスで収集さ れた情報	・視覚・脳画像データなどの追加データ	
データの保管方法		ゲノムに紐づくメタデータは匿名化された 上で、個人情報保護規格に従った安全 なクラウド環境（Data Research Centerのデータベース）に保存されている	ゲノムに紐づくメタデータは仮名化され、 IDと参加者のマッチング情報はGeCIP （Genomics England Clinical Interpretation Partnership）のリサー チドメインに保存されている	FIMEA（Finnish Medicines Agency） の監督下におかれ、ゲノムデータはTHLに よって管理されるゲノムセンターに保管す ることが求められる
データの提供範囲		All of Usリサーチプログラムに許可された 研究者のみが利用可能であるが、製薬 企業等の民間企業の研究者も許可対 象となりえる。	研究機関も民間企業もデータ利用が可 能である。患者のサンプルは外部に提供 していない。制限ありのデータへのアクセ スには会費が必要。	研究機関も民間企業もデータ利用が可 能である。
予算規模		2016年から10年で15億ドル	中央政府は2015年から5年で2.5億ポ ンド。その他にもNHS、財団等からの数 千万ポンドが拠出	7,190万ユーロ（2021年）

出所) Genomics England, All of us, THL Biobank公開情報、

「諸外国におけるゲノム医療の制度・体制・運用等に関する調査（概要版）」2019年12月内閣官房健康・医療戦略室委託事業デロイト・トーマツコンサルティング合同会社

The text is framed by two decorative swooshes. The top swoosh is a gradient bar transitioning from blue on the left to red on the right. The bottom swoosh is a solid blue bar.

Share the Next Values!