

令和4年厚生労働科学研究費 補助金  
政策科学総合研究事業（臨床研究等ICT基盤構築・人工知能実装研究事業）  
総括研究報告書  
大規模データの利活用研究の加速のための研究

研究分担者 佐々木香織（北海道公立大学法人札幌医科大学 医療人育成センター 教授）

#### 研究要旨

2018年より次世代医療基盤法が施行され、新たに設立された認定事業者を通じた医療情報の利活用の促進が期待されている。しかし機械学習の利用や行政情報との連携は具体的には未整備なため、これらを視野に入れた認定事業者のあり方について考慮する必要があるだろう。そこで本研究は個人情報保護と市民社会の関係性において類似点が日本と多く、且つ機械学習と行政情報の利活用を含めた医療情報利活用の制度整備が進んでいる英国の事例から学ぶ。そこから認定事業者にかかる施策提言に発展させる予定である。

#### A. 研究目的

2018年より次世代医療基盤法が施行され、医療情報の利活用の促進が期待されている。同法で想定している利活用とは、認定事業者が提供する匿名加工医療情報を用いた研究や開発を進展することである。しかし匿名加工情報を用いた機械学習においては、懐疑的な見方もでている。機械学習手法は多数の変数を要求する傾向があるが、匿名加工による変数削減とリスクのトレードオフの関係下で変数の数に制約を課せられる可能性があり、その様な要求に応えることが厳しいと推測されるからである。更に認定事業者は行政情報も扱える設計だが、医療情報との連携は未整備である。したがって機械学習と行政情報の利活用を視野に入れた発展が必要と考えられる。

海外でも医療情報の機械学習や行政情報と連携した利活用への期待は高いが、各国で市民社会における個人情報への態度が異なるため、様々な体制が整備されている。英国は個人情報保護と市民社会の関係性が日本と類似する点が多い一方、機械学習を含めた医療情報利活用の体制整備が進んでいる。そこで本分担研究は、英国の事例から本邦が学ぶべき点を探究することを目的とする。

#### B. 研究方法

本研究は文献調査と令和4年9月に英国の関連団体——NHS Digital (England); Scottish Safe Havens; Public Health Scotland; University of the West England; Imperial College London; Small Area Health Statistics Unit; Clinical Practice Research Da

talink——に実施した聞き取り調査からなる。

#### C. 研究結果

英国はEngland, Wales, Scotland, Northern Irelandの連合王国であるため、公的医療サービスは四王国に分かれて運営される。そのうち本研究は紙幅の関係上Scotlandの医療情報利活用を支える公的機関(i.e., Scottish Safe Haven)に焦点をあてる。なぜならScotlandが全英で利活用の体制整備を最も発展させたからである。また一次医療圏のデータ利活用を扱うCPRDにも触れる。全英レベルの組織で自主独立運営という日本の認定事業者との類似性と、近年設立の他組織と異なり30年の歴史という経験知から、参考にすべき点が多いからである。

##### (1) Scottish Safe Haven\*

英国では主に各王国のNational Health Service (NHS)という公的機関を通じて、医療サービスが提供される。Scotland王国のNHS Scotlandでは、医療情報の利活用の体制を整備するにあたり、2008年に組織された官民連携の研究グループScottish Informatics Programme(SHIP)が制作したBlueprintという報告書\*\*にもとづいて、新たにSafe Havenという組織を設立させた。それはNHSの編成に倣い国と地方に分かれている。具体的には、国家レベル組織のNHS Scotland内にあるInformatic Service Division下にNational Safe Havenを配置し、Aberdeen, Dundee, Edinburgh, Glasgowの四地域にある大学附属病院を擁する地方単位のNHSには、四つの地域Safe Havenを配置している。

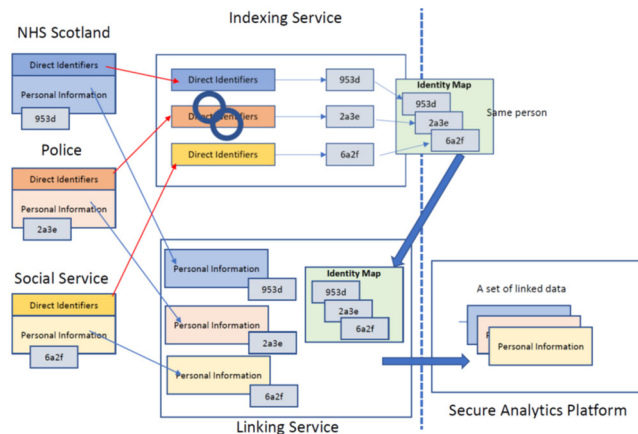
ここで注目すべき点は二つある。一つがTrusted Research Environment (TRE)という環境の整備だ。それはScottish Safe Havenが匿名加工情報を研究者に提供するのではなく、TREという研究目的の安全なIT環境と仮名化した医療情報の二つを提供し、TREという環境下で提供したデータを用いた研究を行ってもらうサービス体制である。それにより機械学習を含めて匿名加工情報では研究の深化や発展が難しい案件を扱うことが可能となる。

しかし仮名情報の使用を鑑み、TREの安全性を盤石にする必要がある。そこでScottish Safe Havenでは英国のOffice for National Statisticsが考案したFive Safe Modelsを発展的に採用し、安全性を担保している。その概要\*は以下の通りだ。

1. **安全な利用者(Safe People):** データ利用は定められたガイドラインと規則に従うことができる研究者に限る。
2. **安全なプロジェクト(Safe projects):** プロジェクトは公共の利益に資するもので、データを適切に扱うことを確約するものに限る。また患者と市民参画、意思決定過程とデータの利用に関する透明性を担保できるものに限る。
3. **安全な施設/環境(Safe settings):** データが安全に保管され、利用者が安全にデータ分析・解析できるTREを提供。且つ利用者の作業を追跡して点検ができるように設定。
4. **安全なデータ(Safe data):** 研究者が分析する過程でデータ対象者を直接特定できるような情報を削除するなどの保護。それによりEU離脱前に批准した通称GDPR (General Data Protection Regulation=一般データ保護規則)に則った形でデータ秘匿性を担保。
5. **安全な分析結果(Safe outputs):** データや成果物の持出しは、個人情報漏洩や特定するリスク点検後に許可されたものだけに限定。

第二は、上述のBlueprintに基づいたデータ突合にかかる制度設計である。各Safe Havenでは突合にあたりIndexerとLinkerの組織を分けることで、個人情報保護を担保している。しかし行政機関など所管が異なるデータの突合の際にはSafe Haven外の第三者機関とも連携 (e.g., National Record of Scotland でのIndexing; Public Health England 内のeDRISでのlinking) している。下図\*\*\*のよう

況に関する情報を扱う機関(Linker)が別個なため、秘匿性を極めて高くした設計となっている。



\* 本セクションはScottish Safe HavenのコアメンバーであるEmily Jeffers on 教授への聞き取り(2022年9月15日) と、その際に提示された彼女の2021年度University of Dundee 大学院講義のスライドをもとにしている

\*\* SHIP (2012) Blueprint for healthcare research in Scotland. [Online] [http://www.scot-ship.ac.uk/sites/default/files/Reports/SHIP\\_BLUEPRINT\\_DOCUMENT\\_final\\_100712.pdf](http://www.scot-ship.ac.uk/sites/default/files/Reports/SHIP_BLUEPRINT_DOCUMENT_final_100712.pdf)

\*\*\* 上註\*の講義スライドを、研究代表者の木村が構成し筆者が一部改変

## (2) Clinical Practice Research Datalink (CPRD)\*\*\*\*

GPRDもScottish Safe Havenと同様の背景でTREへの移行が進み、ほぼ同様のFive Safe Modelを採用している。それに伴い適切なTREの構築と利用者への教育を検討課題として挙げている。また検討の際に特に問題視されているのが、NPOとして受益者負担で運営されている一方で、データの収集・保守・管理や匿名加工といった作業には無料か少額の費用で済むと考えている人が一般市民のみならずCPRDの利用者の中にも一定数いるため、費用への不満がCPRDの設立から三十数年を経ても散見される点である。そこで以下のバランスが今後の発展に重要だとされている――1. 公益に資する研究を行うニーズに、どこまで応えて研究環境を提供するか、2. データ提供の対価として、どの様な価格設定が適切で研究者が負担可能な価格ほどの位か。

\*\*\*\*本セクションは、CPRD Deputy Director Dr Puja Mylesへの聞き取り(2022年9月23日) をもとにまとめている。

## E. 考察

本邦が学ぶべき点は以下の四点に収斂するだろう。

- TRE整備のあり方(ニーズと価格の考慮を含む)
- Five Safe Modelsのあり方
- IndexerとLinkerの分離の可否とそのあり方
- 利用者への教育体制の検討

## G. 研究発表(学会発表)

木村映善, 佐々木香織, 伊藤伸介「英国における行政記録情報と医療情報をリンケージした利活用状況について～英国における機械学習での利用を前提

とした医療情報の利活用状況の調査から～」『大規模データの公開におけるプライバシー保護の理論と応用』公的統計マイクロデータ研究コンソーシアム/統計数理研究所 共済 2022年12月9日