

追跡終了後コホート研究を用いた共通化データベース基盤整備と その活用に関する研究：進捗報告

研究代表者 玉腰暁子（北海道大学大学院医学研究科・教授）

研究要旨

国内で実施され追跡を終了した複数のコホート研究情報を共通化し、その利用環境を整え、将来にわたって向後終了するコホート研究も組み入れ可能な体制を構築するために必要な事項を検討する。そのため、初年度は、米国のデータ利用環境、統計法に基づき実施された国内統計資料の利用のあり方、データ二次利用にまつわる倫理的課題と知財、研究倫理教育のあり方について、現状を確認し、課題を整理するとともに、各コホートが持つデータを公開した場合の一意性の問題について、実データを基に検討した。その結果、統計法の規定から、人口動態統計資料から得られた死因情報をコホート研究のアウトカムとして公開することはできないことが明らかとなった。そのため、コホート研究データをアーカイブ化し二次利用を進めるためには、死因情報のソースとなる人口動態統計調査の統計法内における位置づけを変更することが必要と考えられた。また、二次利用に際しては、対象者の同意の範囲への考慮とそれに伴う匿名化情報としての取り扱いはもちろん、二次利用者の責務を意識することが重要である。さらに、特に試料の利用にあたっては、知財に関し、関係者で事前に十分に検討し、ルールを明確にしておくことが必要と考えられた。なお、多くの項目を組み合わせれば一意性があるものとの前提で、二次利用対応を考えていくことが必要である。

分担研究者

磯博康（大阪大学大学院医学系研究科・教授）

大橋靖雄（東京大学大学院医学系研究科・教授）

祖父江友孝（大阪大学大学院医学系研究科・教授）

辻一郎（東北大学大学院医学系研究科・教授）

A. 研究目的

本研究では、国内で実施され追跡を終了した複数のコホート研究情報を共通化し、その利用環境を整え、将来にわたって向後終了するコホート研究も組み入れ可能な体制を構築するために必要な事項を検討する。初年度は、米国のデータ利用環境、統計法に基づき実施された国内統計資料の利用のあり方、データ二次利用にまつわる倫理的課題と知財、研究倫理教育のあり方について、現状を確認し、課題を整理した。また、各コホートが持つデータを公

開した場合の一意性の問題について、実データを基に検討した。

B. 研究方法

1. 各分野の専門家から現状を伺い、追跡を終了したコホート研究データを二次利用する際の課題を整理した。
2. 三府県コホートデータと JALS 死亡者データを用い、それぞれどのような頻度で一意性が見られるかを検討した。

C. 研究結果

1. 資料・データベースの二次利用の現状と課題
 - 1-1. 米国における二次データ利用の現状
米国では、電子カルテデータ、レセプトデータ、患者レジストリデータのみならず、コホート研究デ

ータについても公開が始まっている。これらのデータを用いることで、コホート研究を1から始めるのに比べ少ない労力でデータセットを作ることが可能である。また、NHLBI から研究費を受けた行われたRCTとコホート研究を含めた全てのデータを最大利用することを意図して、Biologic Specimen and Data Repository Information Coordinating Centerを事務局としたデータデポジトリが行われている。データ提供にあたっては、インフォームド・コンセントに沿った形にすること、個人同定情報は削除すること、地域情報は外すことなどがガイドラインで規定されている。このデポジトリでは、データあるいは生体試料二次利用希望者はHP上で、各研究の詳細を確認し、自身の用いたい研究に対しリクエストをすることができるが、その際、倫理審査を受けておくなどの規則も定められている。このシステムにより、すでに他国からの申請も含め多くの研究が実施されている。

1-2. 日本における公的統計の利用

公的統計は、その体系的かつ効率的な整備及びその有用性の確保を図ることを目的とする統計法で規定されている。2007年の全部改正にあたっては、社会の情報基盤としての統計というコンセプト、そしてそれを実践するための統計データの利用促進、特に二次利用に積極的に取り組むことが盛り込まれた。そのため現在では、オーダーメイド集計や匿名データの利用が推奨されている。なお、ほぼ全てのコホート研究で死因情報は、人口動態統計調査の利用手続きを経て入手されていることから、この情報を更に別の二次利用(提供)に供することは現行の統計法の下では不可能である。

1-3. 試料・情報の二次利用に関する説明と同意

新しく開始される研究では、二次利用まで見越した同意を取得することが望ましい。当初の研究会市時点では具体的な二次利用の詳細は決まっていないうことも多いが、広範同意で二次利用に関する同意も得、実際に二次利用を進める際には倫理審査委員会など第三者の意見も聞いたうえで実施にあたってはオプト・アウト方式(研究に関する情報を公開し、拒否権を保障する)を採用することが推奨され

る。それに対し、既に進んでいるコホート研究では、そのような二次利用を想定していないことが多く、一次同意の範囲では二次利用をカバーできていない。そのような試料・情報であっても有効に研究に活用するためには、対象地域や対象者に対する情報公開を多面的かつ十分に行ったうえで拒否権を保障することが重要であると考えられる。また、二次利用者も研究にあたっては情報公開をしっかりと行うとともに、試料・情報提供者に対し、研究成果に関する報告や有用な個別結果の返却など、できるだけ利益を還元するよう務めることも必要であろう。

1-4. 試料・情報二次利用に必要な研究倫理教育

二次利用に供される試料・情報の特徴を考えると、データアーカイブを用いて研究を行う二次利用者に対しては、以下の点を目指した研究倫理教育が必要と考えられた。

- 一次収集者の苦勞を知る
- 試料・情報の背後には一人ひとりの生身の人間がいることを知る
- ヒト試料研究一般に求められる倫理性を理解する
 - ヒト試料の地位
 - 同意取得の原則
 - 同意原則の緩和が許容されるための条件と発生責務
 - 個人情報の安全管理措置
- 各ステークホルダーに課せられる倫理的責務を知る
- 知財・オーサーシップ等についての適切な知識を持つ
- データ不正等の罪深さを知る
- 情報倫理について理解する

1-5. 試料・情報二次利用に関連する知財

ヒトに由来する試料は、細胞や血液といった形でそのヒトの中にあるときは、当然そのヒトに帰属する。しかし、研究利用に関する説明が行われ、同意を得た上で採取された細胞や血液は、インフォームド・コンセントで示された条件で使用するという約束の下でドナーが譲渡したものと考えることができる。また、その試料・情報を用いて研究が行われ

た場合、研究によって試料・情報に付加価値がつく。したがって、いったん研究に提供された試料・情報を用いて行われた研究の成果として知的財産権が発生した場合、その権利は試料・情報提供者ではなく、研究者(研究機関)に帰属すると考えるのが適切である(対象者への説明時点でももちろん述べておくべき内容である)。バンクや共同研究の形で提供される試料・情報を用いた二次利用で開発された研究成果に関する特許権の分配方法については、事前に取り決めを交わしておくのが望ましい。一次情報収集者の努力、バンク運用者の作業を考えると、そこにどう報いるかを二次利用者は常に意識することも重要であろう。このことは特許権の配分のみならず、論文の共著者に加えるかどうか、謝辞に述べるかどうかも含めて求められる。

2. 一意性の検討

2-1. 三府県コホート対象者に関する検討

三府県コホート対象者10万人強について204変数を分析対象として検討したところ、変数を1つずつ個別に見た場合の一意性は小さかったが、全変数を組み合わせた場合、一意であるレコード数は対象者の約99.98%であった。一変数で見た場合、一意性には、10年観察終了日のように、分類数が大きいことでそれぞれに振り分けられる個体数が少なくなるため生じる一意と、出産人数において出産人数が20人というように、疫学的にまれな属性の個体が存在したために生じる一意の大きく2パターンが存在した。

2-2. JALS 対象者中の死亡者に関する検討

複数のコホートが共同で実施しているJALSのデータを用いて、死亡者約5000人の性、ベースライン時年齢、死亡時年齢、地域、死亡年データによる個人特定可能性につき検討した。性、ベースライン時年齢単独での特定性は低かったが、死亡時年齢が40歳代もしくは90歳代後半である場合や小規模コホートの対象者で特定可能性が高くなった。さらに死亡年月日を月までで丸めた場合でも個人が一意に特定される場合が観察された。また、これら変数を2つ以上組み合わせた場合には、一意に特定されるものの数が上昇した。

D. 考察

1. 試料・情報二次利用の現状と課題

日本よりデータベースの二次利用が進んでいる米国において、研究に用いられているデータベースの種類、活用事例とその成果、二次利用を進めるにあたっての留意点等について確認した。日本では、二次データの公開システムとしてバイオサイエンスデータベースセンター(NBDC)が立ち上がったばかりであり、仕組みそのものに関する議論が十分に行われてはいない。今後、日本においても公的研究費を受けた研究を適切に二次利用することが求められるが、そのためには、個人情報保護と研究活用とのバランス、事務手続きの標準化・単純化と必要経費、データの適正使用と質保障のためのサポート、共通化によるデータマニピュレーション、情報のロスとデータ容易使用のバランス、共通化プロセスの透明化と公正なシェアの仕組み等につき、議論を重ねていくことが必要と考えられた。

国内の統計行政に関しては、統計法で規定されている。改正統計法は、社会基盤としての統計をコンセプトに、公的統計の体系的かつ効率的な整備及びその有用性の確保を図ることが目的とされている。公的統計の利用に関しても範囲が拡大され、秘匿しない調査票情報の提供、秘匿処理した調査票情報の利用(匿名データの利用)、行政機関に特別集計を依頼(オーダーメイド集計)が可能である。なお、提供された調査票情報や匿名データなどを統計の作成に関連する目的以外に利用・提供した者や、守秘義務規定に違反した者に対して、罰則が定められている。したがって、コホート研究データをアーカイブ化し二次利用を進めるためには、アウトカムに用いている死因情報のソースとなる人口動態統計調査の統計法内における位置づけを変更することが必要と考えられた。

研究倫理の観点からは、試料・情報の二次利用に際しては、対象者の同意の範囲への考慮とそれに伴う匿名化情報としての取り扱いはもちろん、二次利用者の責務を意識することが重要である。したがって、アーカイブやバンクによるデータ二次利用を行う研究者に対する研究倫理教育では、通常教育研

修内容に加え、二次利用に関する同意のあり方、オープンシップや知財に関する内容を含めることが必要と考えられた。また、アーカイブやバンクによるデータ二次利用を進めるためには、その後が生じ得る知的財産検討に関し、事前に関係者間で十分に検討することが必要と考えられた。

2. 一意性の検討

一変数で見た場合、一意性には、分類数が大きいことでそれぞれに振り分けられる個体数が少なくなるため生じる一意と、疫学的にまれな属性の個体が存在したために生じる一意の大きく2パターンが考えられた。前者に対しては例えば日付データを月までにするなどにより分類数を減らすことで一意性を減少させることが可能であり、後者に対しては一定値以上（以下）については直接表示せず、無限までの片側区間で表示するといった方法により一意性の減少が図られる。しかし、すべての変数を組み合わせた場合には全レコード数の約99.98%が一意であったが、変数が多くなる（質問項目が多い）ことによる一意性は容易に避けられるものではない。また、10万人規模のデータではまるめ処理などによりある程度の一意性の減少がみられるが、規模が小さくなると一意性が上がる可能性も高い。したがって、多くの項目の組み合わせを行う場合には、一意性があるものとの前提で対応を考える必要がある。

また、複数の比較的小規模なコホートをあわせて実施されているコホート研究における死亡者データで検討したところ、地域の情報が特に個人の特定に繋がりがやすいこと、死亡時年齢が若年もしくは高齢である場合に一意性が高じることが明らかとなった。また、死亡年月日については、月までに丸めた場合でも個人が一意に特定される場合があることから、可能な限り死亡年までの情報とすることが求められる。さらに死亡者につき用いる情報（性、ベースライン時年齢、死亡時年齢、地域、死亡年）を一変数ではなく二変数以上とした場合には、組み合わせによって一意性が上昇した。したがって、分類を粗くするなどの対応が必要と考えられた。

E. 結論

米国のデータ利用環境、統計法に基づき実施された国内統計資料の利用のあり方、データ二次利用にまつわる倫理的課題と知財、研究倫理教育のあり方について、現状を確認し、課題を整理するとともに、各コホートが持つデータを公開した場合の一意性の問題について、実データを基に検討した。その結果、統計法の規定から、人口動態統計資料から得られた死因情報をコホート研究のアウトカムとして公開することはできないことが明らかとなった。そのため、コホート研究データをアーカイブ化し二次利用を進めるためには、死因情報のソースとなる人口動態統計調査の統計法内における位置づけを変更することが必要と考えられた。また、二次利用に際しては、対象者の同意の範囲への考慮とそれに伴う匿名化情報としての取り扱いはもちろん、二次利用者の責務を意識することが重要である。さらに、特に試料の利用にあたっては、知財に関し、関係者で事前に十分に検討し、ルールを明確にしておくことが必要と考えられた。なお、多くの項目を組み合わせれば一意性があるものとの前提で、二次利用対応を考えていくことが必要である。

F. 健康危機情報

なし

G. 研究発表

1. 論文発表
なし
2. 学会発表
なし

H. 知的財産権の出願・登録状況（予定を含む）

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし