

201508003A

厚生労働科学研究費補助金

循環器疾患・糖尿病等生活習慣病対策総合研究事業

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

平成 27 年度 総括・分担研究報告書

平成 28 (2016) 年3月

研究代表者 玉腰 暁子

目 次

I. 総括研究報告

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…………… 1

玉腰暁子

II. 分担研究報告

疫学追跡終了後コホートデータの共通利用（アーカイブ化）の際の死因データ利用に関する検討…………… 7

大橋靖雄、原田亜紀子

アーカイブデータ提供のためのガイダンスに含める必要がある項目の整理…………… 12

玉腰暁子、祖父江友孝、他

コホート研究で人口動態統計資料から得た死因情報をデータアーカイブに付加して提供するための統計法解釈に関する検討…………… 14

玉腰暁子、大橋靖雄、他

個人情報保護法改正に伴う疫学研究への影響に関する検討…………… 17

磯博康、辻一郎、他

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

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

疫学研究により得られたデータを広く共有化するためのシステムであるデータアーカイブ化に向け検討を行った。統計法の規定上、死因情報を付加してのデータ公開・二次利用は認められていないことなどを踏まえて、現制度下での運用方法につき、分散型ネットワークについて国内外の実例を収集し、疫学研究追跡終了後のコホートデータの共通利用を行う上での課題を検討した。また、データアーカイブの利活用を進めることを念頭に、データ提供の際に従うべきガイダンスに含める必要がある項目を A 対象者の個人情報保護、B インフォームド・コンセントと倫理審査、C 知的財産権の帰属、D 寄託する項目/しない項目の判断、E データ提供先の制限、の 5 点に整理した。

一方で、エンドポイントとしての死因情報は非常に価値が高いものであることから、追跡が終了したコホート研究の二次利用を進めるために、死因情報のソースとなる人口動態統計調査の有効活用の方策を模索することが望まれ、その際の論点を 2 つ提示した。最後に、改正された個人情報保護法において、病歴が要配慮個人情報に位置づけられたことから、以前から行われているコホート研究において生じうる影響を検討した。その結果、追跡情報の入手を医療機関等から受けている場合に、過剰反応により情報が提供されないことが懸念された。

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(中央大学理工学部・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こと

を目的としている。統計法の規定により、人口動態統計情報を利用して把握している死因情報は、データアーカイブに付加して二次利用できないことから、昨年度に引き続き①現制度化での利用方法の提案を行った。また、死因情報の付加如何にかかわらず、人々の生活と健康に関する情報を取り扱うコホート研究では対象者の個人情報保護等に関する配慮が求められることから、②アーカイブ化を進めるに際して必要なガイダンス項目を整理した。一方で、現状では死因情報の二次利用は認められていないものの、追跡が終了したコホート研究を二次利用するにはその価値は極めて高いこと

から、③付加して提供することを可能にするための統計法解釈の論点を提示した。最後に、本研究で取り扱っているコホート研究は指針や個人情報保護法施行以前から開始されていることから、2015年に行われた個人情報保護法改正により病歴が要配慮個人情報となったことによる研究への影響を検討した。

B. 研究方法

①現制度下で二次的に死因情報を利用するため、国内外の医療（臨床）データ連携、疫学共同研究などにおける分散型ネットワークの先行事例を収集した。

②昨年度までの研究成果や先行する分野のデータアーカイブセンターのガイドライン等を参考に、特に研究開始時期のデータ取り扱いが現在のものと異なり、かつ多数を取り扱う疫学研究データのアーカイブ化の場合にガイダンスに含めておくべき項目を検討した。

③統計法の位置づけと人口動態統計資料の利用価値の観点から、情報関係法制に詳しい法律学の専門家である友岡史仁日本大学法学部教授を交え、検討を行った。

④個人情報保護法の条文、特に要配慮個人情報に関する規定を確認するとともに、疫学研究者、特にコホート研究の代表者から現在の運用等に関する情報を得るとともに議論を行った。

C. 研究結果

①現制度化での二次的な死因情報の利用

分散型ネットワークとは、データを保持者のもとに置いたまま、必要と判断した情報だけを選択的に共有させる方法である。近年整備されつつある分散型ネットワークについて、医療（臨床）データ連携、国際疫学共同研究での5つの先行事例を収集し検討を行ったところ、医療（臨床）と疫学研究の事例では、データ連携（交換）の目的（必要とする背景）、運用方法などが異なっていることが明らかとなった。

②アーカイブデータ提供のためのガイダンス項目

含める必要があると考えられたのは、以下の項目であった。

A 対象者の個人情報保護：実データを用いた我々の過去2年間の検討では調査項目が多岐に渡り、かつ対象者数が多いコホート研究では、データの一意性はあるものと考えた方がよいと結論した。回答者のプライバシー保護の観点から、調査地域を粗く束ねるなどの手段を用いて、個々の回答者の識別を不可能にすることが望まれる。

B インフォームド・コンセントと倫理審査：新たに開始される研究では、対象者からデータ寄託に関する同意が得られていることを原則とした上で、寄託者の所属機関で倫理審査を受けることが可能である。しかし、既に追跡が終わったコホート研究の場合、その開始は今から20年以上前であると考えられ、研究の説明同意プロセスは現在と異なる。そのため、対象者の同意あるいは再同意の取得を前提とすることは困難であり、所属機関での倫理審査に加え何らかの形で二次利用に関する広報・情報公開を行うことが適切と考えられた。

C 知的財産権の帰属：生体試料を伴わない情報のみのアーカイブでは、知的財産権の発生は稀と思われるが、生体試料を収集・保管するコホート研究では、研究終了時にも試料が残ることが想定される。二次利用者とデータ寄託者の間の分配方法については、一定のルールが定まっていないことから、利用開始時に決めを交わしておくことが望ましい。

D 寄託する項目/しない項目の判断：個人同定の可能性が特に高い項目は、丸め処理をするか、提供しないことが望まれる。また、調査方法や設定が特殊である等何らかの理由により、研究内容を十分に理解しない二次利用者の解析・解釈によって誤った結論を引き出すおそれがある項目についても、提供しないことが推奨される。

E データ提供先の制限：二次利用者の資格、利用上の制約等については、各研究関係者の意向を反映するような区分けの周知とそれに沿った対応が求められると考えられた。

③統計法解釈に関する検討

コホート研究に不可欠とされる死因情報を人口動態統計資料から得ることは統計法上の目的外利用を意味

するため、その活用に当たり、統計法上設けられた諸
手続といった一定の制約が生じている。しかし、データ
の二次利用により、公費を投入し多くの人々の協力を
得て収集された試料・情報の有効利用、若手研究者の
育成、データの検証等幅広い研究利用とその成果に
直結し、ひいてはより正確な健康関連要因の解明に具
体的に貢献しうることを考え合わせれば、統計法の規
定による制約を受けない形で、死因も含めたより広い
形による該当データの利用可能性が切望され、当該情
報を長期的かつ利用者の範囲を広げてコホート研究に
有効活用する道を何らかの形で法制度上担保する必
要性を論じておく意義は、十分にあり得ると思われる。

この場合の論点として、A 統計法の解釈運用を変更
することで足りるのか、B 統計法とは異なる別立法を行
う必要があるのか、という二点が考えられる。A の場
合は、上記の制度趣旨と過去の運用の経緯に照らして
乗り越えるべき解釈論を構築する必要があると考えら
れ、その意味で、死因情報以外の件についても考慮に
入れた制度全体の観点から議論を要するものと思われ
る。対して B の場合は、A とは異なり立法的措置により
解決するため、特別法として位置付けられ、法解釈上
の齟齬はきたさないとされる一方、既存の仕組みと
は異なる新たな制度設計を求められる分、個人情報保
護に照らした保護に係る必要なスタンダードを充足す
る詳細な議論を要する。A または B のいずれを選択す
るにせよ、新たな立法の煩雑さや既存の仕組みの有効
活用という効率性の視点からすれば、現時点では A の
選択肢が方向性として現実的と考えられた。

④病歴が要配慮個人情報に位置づけられたことによる 研究への影響

今回の個人情報保護法改正で定められた「要配慮
個人情報」は、その取り扱いによって差別や偏見、その
他の不利益が生じるおそれがあるため、特に慎重な取
扱いが求められる個人情報とされており、この中に本
人の人種、信条、社会的身分、犯罪の経歴、犯罪によ
り害を被った事実とともに病歴が加えられた。要配慮個
人情報では、取得にあたっては原則として本人の同意
を得ること、オプトアウト手続きによる第三者提供が認

められないこととされている。ただし、関連性を有する範
囲内で利用目的を変更したり、匿名加工情報に加工し
て第三者に提供したりすることは可能である。

指針等の整備前から開始されたコホート研究であつ
ても、開始以降に倫理審査を受けるなど、研究者は可
能な限り対象者の個人情報保護に配慮して研究が行
われていた。また、研究開始時に得ている病歴に関し
ては、本人の自己申告によるものが多く、また医療機関
や健診機関から得ている場合でも、その利用に関して
は何らかの説明が行われていた。しかし、対象者が多
いこと、開始時期が古く対象者が研究に参加している
ことを記憶していない場合もあること、遡っての同意手
続きによって対象者の偏りが懸念されることなどにより、
開始時点の説明が現在の基準に照らした場合に不十分
であっても改めての同意手続き等を経ることなく、病
院から追跡情報として疾病情報を受けている研究も存
在した。したがって、今回、病歴が第三者提供に際して
対象者からの同意が必須である要配慮個人情報と位
置づけられたことから、医療機関等の過剰反応によりコ
ホート研究に重要な追跡情報を得る際に支障を来すこ
とが懸念された。また、仮に、改めて再同意手続きをと
り同意が得られた者のみを観察対象とした場合には、
対象者の偏りを生じることから研究の価値を著しく損
ね、ひいては公衆衛生上の不利益につながるものが危
惧された。さらに、疾病登録による罹患率等の把握の
ための疫学研究、特定の疾患の患者の予後及び予後
要因を明らかにする疫学研究（臨床研究）、複数の既
存資料を個人情報とキーとして突合する疫学研究、構
成員全員の健康増進対策を検証・評価する研究、定
期健康診断のデータを用いた疫学研究などで、倫理
指針の例外規定が正しく認知されないと影響が出るこ
とが懸念された。

一方、本研究班で検討している追跡終了後のコホー
ト研究をアーカイブ化し、二次利用に用いる際には、②
に示したように対象者を遡れないような匿名化措置が
必要と考えられることから、病歴を付加して提供す
ことは差し支えないと考えられた。ただし、非常に稀な病
気の場合等は、詳細な病名ではなく、ある程度まとめた
疾患概念として示すなどの配慮を行うことが求められ

る。また、一口に病歴といっても、医師が記載した診断書病名のみならず、本人申告や検査値やその組み合わせ等のバリエーションが考えられることから、どこまでを要配慮個人情報としての病歴と定めるかの合意が必要と考えられた。

D. 考察

国民の税金を投入し、多くの人手と長期の追跡を経て構築されたコホート研究データをアーカイブ化し、広く二次利用を可能とすることは、研究の透明性確保、第三者による研究結果の検証、若手の育成に寄与するのみならず、研究の無駄・重複を減らし、必要な公費・労力を新しい有意義な研究に向けるという意義もある。多くのコホート研究では、死因情報を人口動態統計資料から得ており、統計法の規定上、現状では非常に重要な情報であるにもかかわらずアーカイブデータに死因情報を付加して二次利用に供することはできない。そこで、現制度下での運用方法につき、昨年度に引き続き今年度は、後者の分散型ネットワークについて国内外の実例を収集し、疫学研究追跡終了後のコホートデータの共通利用を行う上での課題を検討した。その結果、医療（臨床）データ連携と疫学研究の事例では、それぞれ、データ連携（交換）の目的（必要とする背景）、運用方法などが異なっており、疫学研究のデータ連携では、様々な統計解析を柔軟に行えること、ローカルサイト（大学、研究所、医療機関を想定）へアクセスを行う上でのセキュリティ、さらには、人口動態統計の二次利用をローカルサイト単位で各々申請し、死因を付与したデータを統合して用いることや、施設外部の PC 上の RAM への一時的な書き出し（保存はされない、電子データの「一時的蓄積」）の可否など、法令解釈の問題にも留意する必要があると思われた。

また、前述したようにアーカイブ化により二次利用環境を整える意義は大きいですが、疫学研究では生活習慣や生活環境のみならず、個人の心身面の健康状態情報を収集することから、その取り扱いには慎重であることが望まれる。今回、特に追跡の終わったコホート研究の利活用を進めることを念頭に、データ提供の際に従うべきガイダンスに含める必要がある項目を5点に整

理した。今後、追跡が終了したコホート研究データの寄託を進める際の指標になるものと考えられる。

一方でコホート研究にとって死因情報は非常に重要なものであり、利用できないことによりアーカイブデータの二次利用の価値は低下すると考えられる。また、米国では National Death Index (NDI) という、厚生省 (U.S. Department of Health and Human Services) の下部機関が、研究目的での生存・死亡確認情報（死亡時には、死亡年月日や死因などを含む）の提供を行っている。研究者は、調査対象者リスト（氏名、性、生年月日、住所、社会保障番号など）を提出し、審査にパスすると、有料（基本料 350 ドル＋対象者1人1年あたり 15 セント）で、上記情報が提供される。これにより、米国の疫学研究・臨床研究のレベルと即時性は飛躍的に向上し、医学研究や医薬品開発において国際的に有利な地位を確保することができたといえ、今のままでは日本の疫学研究は後塵を拝する。そこで、法の解釈、または立法により、コホート研究において重要な追跡情報となる死因を公的調査情報（人口動態統計資料）から得た上で二次的利用を可能にできないかを検討した。しかし、これには統計法の従来の解釈経緯などといった従前のコホート研究とは異なる方向性が必要とされており、今後も継続して、法学分野からの知見に照らし、より視野を広げた説得力ある議論を展開する必要があると考えられた。

最後に、改正された個人情報保護法において、病歴が要配慮個人情報に位置づけられたことから、以前から行われているコホート研究において生じうる影響を検討した。その結果、追跡情報の入手を医療機関等から受けている場合に、過剰反応により情報が提供されないことが懸念された。またその他にもいくつかの研究事例で、病歴が要配慮個人情報となったことにより影響を受けることが懸念された。したがって、研究者が適切な個人情報保護体制を整えることはもちろんであるが、その下で、これまでと同様に利用が可能になるようガイダンスの整備等を進めることが重要と考えられる。今までにも追跡のために対象者の異動情報を自治体から得る際、ごく一部の自治体からではあるが個人情報に対する過度な対応から情報提供がなされなかったという

事例が生じていることから、個人情報保護法改正を受けて行われる指針やガイダンス改正作業ではこの点に十分配慮したものとなるよう働きかけることが重要である。

E. 結論

疫学研究により得られたデータを広く共有化するためのシステムであるデータアーカイブ化に向け検討を行った。統計法の規定上、死因情報を付加してのデータ公開・二次利用は認められていないことなどを踏まえ、現制度下での運用方法につき、分散型ネットワークについて国内外の実例を収集し、疫学研究追跡終了後のコホートデータの共通利用を行う上での課題を検討した。また、データアーカイブの利活用を進めることを念頭に、データ提供の際に従うべきガイダンスに含める必要がある項目を A 対象者の個人情報保護、B インフォームド・コンセントと倫理審査、C 知的財産権の帰属、D 寄託する項目/しない項目の判断、E データ提供先の制限、の 5 点に整理した。

一方で、エンドポイントとしての死因情報は非常に価値が高いものであることから、追跡が終了したコホート研究の二次利用を進めるために、死因情報のソースとなる人口動態統計調査の有効活用の方策を模索することが望まれ、その際の論点を 2 つ提示した。最後に、

改正された個人情報保護法において、病歴が要配慮個人情報に位置づけられたことから、以前から行われているコホート研究において生じうる影響を検討した。その結果、追跡情報の入手を医療機関等から受けている場合に、過剰反応により情報が提供されないことが懸念された。

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

分担報告書

疫学追跡終了後コホートデータの共通利用(アーカイブ化)の際の死因データ利用に関する検討

研究分担者 大橋 靖雄 中央大学理工学部人間総合理工学科生物統計学
研究協力者 原田亜紀子 東京大学大学院医学系研究科生物統計学

研究要旨

人口動態統計二次利用申請により照合した死因情報を付与した形でのデータアーカイブ化が難しい現状を鑑みて、現制度下での運用案を検討した。一つ目はアーカイブセンターなどで死因情報以外のデータを集約管理し、必要に応じて従来通りの死因照合作業を行い、死因を付加したデータセットを作成し解析を行う方法である。二つ目は、データを保持者のもとに置いたまま、必要と判断した情報だけを選択的に共有させる分散型ネットワークによる方法である。今年度は、後者の分散型ネットワークについて国内外の実例を収集し、疫学研究追跡終了後のコホートデータの共通利用を行う上での課題を検討した。

A. 目的

医療(臨床)データ連携において環境が整備されつつある分散型ネットワークの事例を収集し、疫学研究追跡終了後のコホートデータの共通利用を行う上での課題を検討する。

B. 方法

国内外の医療(臨床)データ連携、疫学共同研究などにおける分散型ネットワークの先行事例を収集した。

C. 結果

1. 国内:医療(臨床)データ連携

1) Standardized Structured Medical record Information eXchange (SS-MIX)

医療機関では電子カルテ・オーダエントリーを中心として、調剤システム、臨床検査システム、画像診断情報システム等、様々な部門システムが稼働し、各々のシステム間で情報がやり取りされている。SS-MIX は、厚生労働省電子的診療情報交換推進事業(Standardized Structured Medical record Information eXchange)で策

定された『電子的診療情報を他システムとの交換や地域医療連携で利用するために、診療情報を標準的な形式で蓄積・管理するデータとして保存できる領域』の仕様のことである。SS-MIX では、これらの医療情報を「標準化ストレージ」というツールに各情報を標準化した形式で格納・蓄積することにより、複数ベンダー間・複数システム間の相互運用性を高めることを目指している。標準化ストレージの構造は、図 1 に示す通り、コンピュータの一般的なファイル格納形式と同様に階層化されたフォルダのディレクトリー構造を用いている。電子カルテシステムなどではほとんどの業務で個々の患者を軸として診療情報が格納されているため、標準化ストレージにおいても患者にひも付く各種の情報をフォルダの階層構造にルールを決めて格納している。フォルダの階層構造は、まず各医療施設用のルートフォルダ(医療施設 ID)を置き、配下に患者 ID 先頭 3 文字、患者 ID 4~6 文字、患者 ID、診療日、データ種別フォルダの順に設け、このフォルダの中に HL7 ベースの標準化された各種データファイル群が格納される。ネットワーク内で「院内リポジトリ」及び「公開リポジトリ」に標準

化ストレージを適用し、それ以外の項目について、拡張ストレージを設け統合することで、標準化ストレージを核にし、施設間連携を構築し、診療情報の研究利用、地域医療連携、災害時連携などを目指している(図 2)。

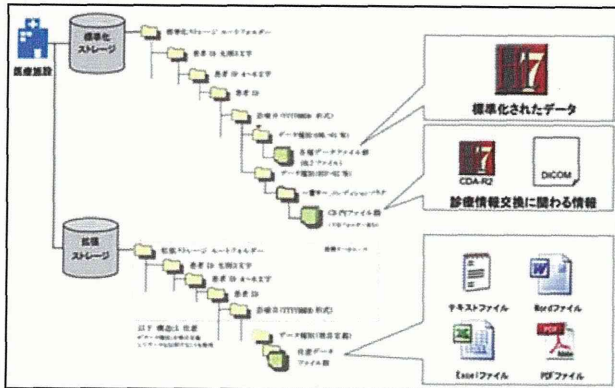


図1 SS-MIX 情報格納ルール

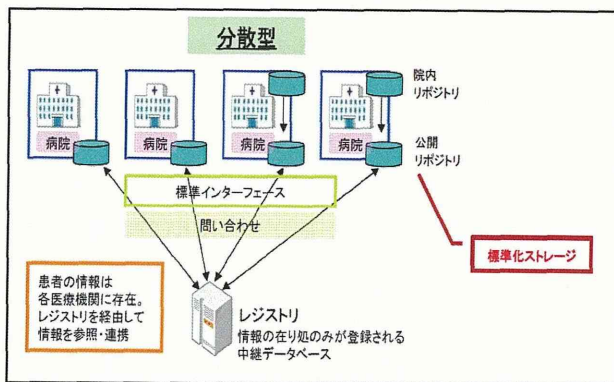
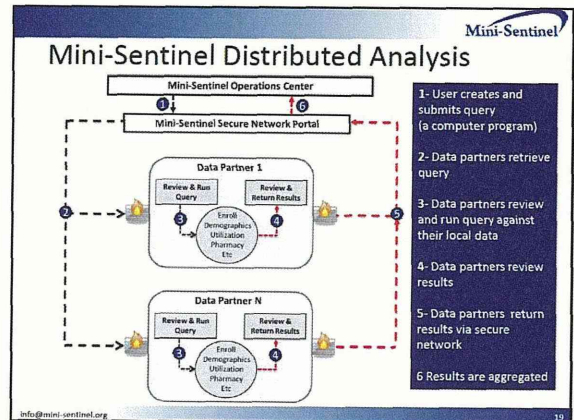


図2 分散型ネットワーク

2. 米国:医療(臨床)データ連携

1) Sentinel Initiative データ交換モデル

米国では 2008 年より、Food and Drug Administration Amendments Act (FDA 改革法)によって FDA 承認医薬品および医療製品の安全性モニタリングのため、既存のデータベースを用いたアクティブ・サーベイランスである Sentinel Initiative が実施されることとなった。Sentinel Initiative ではデータインフラと手法の開発を目的として、Harvard Pilgrim Health Care Institute が中心となって医療データベースを所有する数十の医療機関が協力して Mini-Sentinel が実施され、各データストレージは分散したままで、コモンデータモデルに従って共通プログラムで解析するシステムが構



構築されている(図 3)。

図 3 Mini-Sentinel データ連携モデル

- ①運用センターが新規 Query をセキュア・ネットワークポータルに発行する
 - ②ローカルサイトが配布された Query を取り込む
 - ③④ローカルサイトにおいて、自施設データに対して Query を実行する
 - ⑤⑥結果を運用センターに戻す
- (<http://www.mini-sentinel.org> より)

2) Health Information Exchange (HIE)

CDX (Crossflo Data Exchange®) ソフトウェアは、Crossflo 社が開発したデータ交換ソフトウェアで、HL7、GJXDM (Global Justice XML Data Model)、NIEM (National Information Exchange Model)、EDXL、CAP、NCPDP など主要医療データシステムおよび国家的なデータ標準、業界標準に準拠したデータ共有方式を実装している。このうち NIEM は、米国政府が後援する、公共および民間部門の組織間での情報共有を促進するためのイニシアチブである。NIEM の当初の目的は法執行、公安、危機管理に置かれていたが、他の分野にも拡大されている。

本システムの医学領域の応用事例としては、モンタナ州および連邦政府の要請を受けて、国立医療情報科学センターの協力のもと、モンタナ州の 4 つの病院の ED (救急科) と州保健福祉局 (Montana Department of Public Health and Human Services) とをデータ接続するモンタナ州医療情報交換システムプロジェクト (Montana Health Information Exchange Pilot Project)

があげられる。CDX によるシステムで、ほぼリアルタイムでの医療情報の交換を実現し、異種の医療データソースを迅速および効率的に接続する疾病サーベイランスシステムが構築されている。

3. 国際共同疫学研究等での分散型データ共有

1) ViPAR: Virtual Pooling and Analysis of Research Data (Karter KM et al. *Int.J.Epidemiol.* 2015)

ViPAR は、International Collaboration for Autism Registry において、国際共同研究を実施する6つのサイトデータを統合的に解析することを目的に開発された (Karter KM et al. *Int.J.Epidemiol.* 2015)。多くの医学研究では、倫理的、法的な理由から保有するデータの管理と保持を各サイトにおいて行う必要があるが、本研究では、解析の目的で一時的に「Virtual pooling」サイトに移動することは許可されたことから、一か所にデータを持続的に収集することなしに、物理的に離れたリモートサイトのデータを仮想的(一時的)に集約し分析を可能とするシステムが開発された。

ViPAR のリモートサイトとマスターサーバの関連の関連を図 4 に示したが、ViPAR のマスターサーバは、リモートサイトの PC へ接続されており、研究データはそれぞれリモートサイトで管理されている。利用者は Web 上の analytical portal からアクセスし、データをリモート PC からマスターサーバの RAM 上に抽出し、仮想的にデータをプールして解析を行い、終了後は保存することなく消去する。このため外部にデータを収集することなく多施設データを統合した解析が可能となっている。

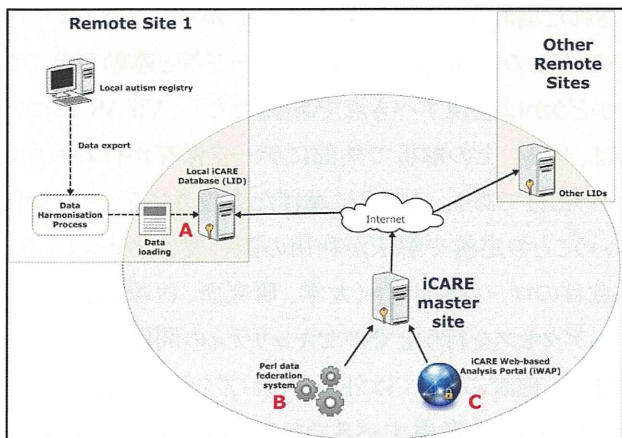


図 4 ViPAR のリモートサイトとマスターサーバの関連

(Karter KM et al. *Int.J.Epidemiol.* 2015)

ViPAR の環境のうち、リモートサイトのデータベースは、データ storage として MySQL サーバを、マスターサイトとの通信のため SSH サーバが導入され構成されている (いずれもオープンソース)。実際に解析を行うマスターサーバは、ViPAR daemon と web-based のポータルで構成されている。ViPAR daemon は、ログ出力、統計パッケージのアクセスとコントロール、ローカルサーバから抽出したデータの統合作業を担う。web-based のポータルは、解析、データ管理を行うインターフェイスであり、新規の解析スタート画面、解析アウトプット画面、コードの管理画面の三つからなる。ViPAR を用いた具体的な解析の流れは、以下の通りである。

- ① 解析用インターフェイスを開き変数やサイトを選択
- ② 統計パッケージの種類 (R, SAS, Stata) を選択し、テキスト入力部分に解析のための syntax を入力しサブミットする
- ③ リモートデータベースにデータのリクエストが送られ、マスターサーバの RAM に読み込まれ、バーチャルプールされる (これらのデータは保存されることはない)。バーチャルプールデータセットは、選択された統計パッケージに読み込まれ、解析が行われる
- ④ ファイルマネージャー画面で、解析の進行状況、完了状況が確認でき、すべての結果とログがダウンロード可能になっている。

その他の追加の機能としては、よく使用する解析プログラムをアップロードでき、リユースやシェアができ、結果を共有できる機能などが存在する。ViPAR のプログラム、マニュアル等は下記で公開されている

<http://bioinformatics.childhealthresearch.org.au/software/vipar/>

2) Data SHIELD (Gaye A et al. *Int. J. Epidemiol.* 2014; 43:1929-1944)

DataSHIELD は、研究の管理部門に設置した分析用 PC からローカルサイトにアクセスすることで、ローカル

データを外部に出す(収集する)ことなく統合的に解析するシステムである。既述の ViPAR と類似したシステムであるが、ViPAR が 3 種類の統計パッケージ (R, STATA, SAS) を使用でき、解析を行う際に柔軟に対応できるのに対し、DataSHIELD は、R 環境で設計されており、使用可能な解析の種類も generalized linear model (GLM) に限られている。ViPAR は、RAM を使用することで解析結果が外部に保存されることはないが、一旦施設外にデータが出ているのに対し、DataSHIELD ではデータがローカルサイトから一度も外に出ることがない点が両者の大きな違いである。

D. 考察

近年整備されつつある分散型ネットワークについて、医療(臨床)データ連携、国際疫学共同研究での先行事例を収集し検討を行ったところ、医療(臨床)と疫学研究の事例では、データ連携(交換)の目的(必要とする背景)、運用方法などが異なっており、本検討課題である疫学追跡終了後コホートデータの共通利用への応用を考える上では、それぞれの特徴を考慮する必要があると思われた。

両者の相違点を整理すると以下のような特徴がみられる。医療(臨床)データ連携と疫学研究データの連携で共通する点は、倫理的規定などからデータ使用の場が制限されており、物理的に施設外に出せないが、連携した利用を積極的に行いたいというニーズがある点である。これに対しては、施設外に出さずにデータ統合と解析を行うさまざまな方法が検討されてきている。医療(臨床)データ連携においては、このようなニーズに加え、データ連携システムの構築を通じて、各種データ形式の標準化を推進していきたいという考えや施設規模の違いや使用しているベンダーの違いを超えた地域医療のデータ連携、災害時のデータバックアップなども目的となっている。さらに、医療(臨床)データ連携では、データの統合方法にも違いがみられ、1) 標準化されたデータと統一様式でのデータ保存(階層化、ストレージ作成)を目指す方向(例:SS-MIX)と、2) データの多種多様性を許容したまま、柔軟にデータ連携(交換)を行える交換様式の開発を目指したものの(例:HIE)などが存在している。これらの違いは、連携

しようとしているデータ量や連携データを即時的に使用する必要があるかといったニーズの違いにも依存していると考えられた。

一方、疫学研究でのデータ連携では、医療(臨床)データ連携でみられるような多くの目的を充足する必要がなく、特に即時的なデータ連携、多種多様なデータをそのまま連携したいというニーズは必ずしも高くはない。疫学研究では、研究間で調査項目の共通性・類似性も高いことから、SS-MIX を例に考えるのであれば、共通性の高い項目は標準化ストレージ用の項目、コホート単位でのオリジナルな項目については、拡張ストレージに保存し、ViPAR のようなオープンソースで構築されたシステムを活用して、ローカルサイトにアクセスして共通利用していくような方法も可能なかもしれない。

逆に医療(臨床)データの連携とは異なる要件としては、大部分の解析をリモートアクセスで行うことを前提にすれば、様々な統計解析を柔軟に行えること、施設外にデータが出るということの定義(法令、指針の解釈)、ローカルサイト(大学、研究所、医療機関を想定)へアクセスを行う上でのセキュリティなどが考えられる。また、人口動態統計の二次利用をローカルサイト単位で各々申請し、死因を付与したデータを統合して用いることや今回紹介した ViPAR での運用のように、施設外部の PC 上の RAM へ一時的な書き出し(保存はされない、電子データの「一時的蓄積」)を行うことの可否など、法令解釈の問題にも留意する必要がある。

一点目の統計解析が柔軟に対応できるかどうかという点については、医療(臨床)データ連携のように、定型化された集計作業(例:Mini sentinel)が中心ではないので、ViPAR のように各種統計パッケージ等と連動可能であるかどうかは重視すべき点である。ただし、ViPAR については、RAM 上の解析で外部にデータ保存が行われないうり、二点目に挙げた施設外にデータが出ることの定義や解釈が利用の際の要件となってくる。三点目のローカルサイト(大学、研究所、医療機関を想定)へアクセスを行う上でのセキュリティの問題については、昨今、施設によっては外部からのアクセスに制限が多くなっている点を考慮すべきである。また、研究によっては、ローカルサイトにセキュリティレベルが低い施設も含む可

能性があることなどもこうすべき点である。四点目の機器利用時・通信過程における一時的なデータ蓄積については、文部科学省文化審議会著作権分科会等でも議論されている。著作権の視点からの議論ではあるが、「一時的固定(複製)」については、次のように整理されている(コンピュータに該当する部分のみ抜粋)。

ア. 瞬時的・過渡的な蓄積であり「複製」ではないもの

- ・処理装置(CPU)の読み込み
- ・ビデオ RAM への書き込み

イ. 一時的固定(複製)のうち、「複製」と判断すべきものではないもの

- ・主記憶(RAM)への蓄積
- ・補助記憶のドライブキャッシュ注釈
- ・CPU における1次キャッシュ2次キャッシュ

ウ. 一時的固定(複製)のうち、「複製」と判断すべきもの

- ・主記憶(RAM)への蓄積(常時蓄積)

このような解釈は、技術動向を見極めて判断されるものでもあり難しい課題であるといえる。

ビッグデータ、IoT、PHR、地域包括ケアにおける地域医療データ連携などが推進されている現代においては、分散型ネットワークによるデータ連携(交換)は、過去の疫学研究データの統合に限定した話ではない。今後、多施設で行う疫学研究を計画する際にも、一施設にデータを集約しないこのような研究方法の利用は検討していく必要があり、これにより従来の疫学(観察)研究のスキームが、より効率性・生産性が高く、信頼性の高い方法へ向上していく可能性も考えられる。

E. 結論

分散型ネットワークの事例を収集し、疫学追跡終了後コホートデータの共通利用に活用できるかどうか検討を行った。医療(臨床)データ連携と疫学研究の事例では、それぞれ、データ連携(交換)の目的(必要とする背景)、運用方法などが異なっており、疫学研究のデータ連携では、様々な統計解析を柔軟に行えること、ローカルサイト(大学、研究所、医療機関を想定)へアクセスを行う上でのセキュリティ、さらには、人口動態統計の二次利用をローカルサイト単位で各々申請し、死因を付与したデータを統合して用いることや、施設外部の PC 上の RAM への一時的な書き出し(保存はされない、電子データの「一時的蓄積」)の可否など、法令解釈の問題にも留意する必要があると思われた。

F. 研究発表

1. 論文発表
 2. 学会発表
- いずれもなし

G. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得
 2. 実用新案登録
 3. その他
- いずれもなし

アーカイブデータ提供のためのガイダンスに含める必要がある項目の整理

研究代表者	玉腰暁子	北海道大学大学院医学研究科
研究分担者	祖父江友孝	大阪大学大学院医学系研究科
研究協力者	辻 一郎	東北大学大学院医学系研究科

研究要旨

特に追跡の終わったコホート研究の利活用を進めることを念頭に、データ提供の際に従うべきガイダンスに含める必要がある項目を 5 点に整理した。その 5 点とは、①対象者の個人情報保護、②インフォームド・コンセントと倫理審査、③知的財産権の帰属、④寄託する項目/しない項目の判断、⑤データ提供先の制限、である。

A. 目的

疾病の発生関連要因の検討を目的としたコホート研究では、生活習慣の詳細や既往歴等の情報をベースに、その後の疾病発生や死亡状況等を追跡し、情報を蓄積する。したがって、追跡を終了したコホート研究データのアーカイブ化を進めるにあたっては、利用を促進すると同時に倫理面や研究参加者の個人情報保護に配慮した適切なデータ提供が求められる。しかし、研究開始時点のデータ取り扱いや対象者への説明同意のあり方は、現在従うことが求められている人を対象とした医学系研究に関する倫理指針で定めるものとは異なっている。そこで、今後利活用を進めることを念頭に、データ提供の際に従うガイダンスに含める必要があると考えられる項目の整理を試みた。

B. 方法

昨年度までの研究成果、ならびに、国内のヒトデータベース(ライフサイエンス分野)の収集・公開を担うバイオサイエンスデータベースセンター(NBDC)のガイドライン、社会科学分野のデータアーカイブの運営

を行っている東京大学社会科学研究所附属社会調査・データアーカイブ研究センターSSJDA のデータ寄託手引きを参考に、特に研究開始時期のデータ取り扱いが現在のものと異なり、かつ多数を取り扱う疫学研究データのアーカイブ化の場合にガイダンスに含めておくべき項目を検討した。

(倫理面への配慮)

特に必要なし。

C. 結果

含める必要があると考えられたのは、以下の項目であった。

① 対象者の個人情報保護:SSJDA では回答者のプライバシー保護に最大限の注意が払われており、調査地域を粗く束ねるなどの手段を用いて、個々の回答者の識別を不可能にしている。我々の実データを用いた過去 2 年間の検討でも調査項目が多岐に渡り、かつ対象者数が多いコホート研究では、データの一意性はあるものと考えた方がよいとしており、この点に関してデータ寄託者、アー

カイク運営者ともに十分に認識する必要がある。

- ② インフォームド・コンセントと倫理審査:NBDC では対象者からデータ寄託に関する同意が得られていることを原則とした上で、寄託者の所属機関で倫理審査を受けることを義務付けている。しかし、既に追跡が終わったコホート研究の場合、その開始は今から 20 年以上前であると考えられ、研究の説明同意プロセスは現在と異なる。また、コホートによっては、地域との関係が現在まで維持されていないことも考えられる。そのため、対象者の同意あるいは再同意の取得を前提とすることは困難であり、所属機関での倫理審査に加え何らかの形で二次利用に関する広報・情報公開を行うことが適切と考えられた。
- ③ 知的財産権の帰属:生体試料を伴わない情報のみのアーカイブでは、知的財産権の発生は稀と思われるが、生体試料を収集・保管するコホート研究では、研究終了時にも試料が残ることが想定される。試料・情報提供者と直接提供を受ける研究者の間であれば、知的財産権は研究者(研究機関)に帰属することとなり、最近では説明事項に含まれることも多い。一方、二次利用者とデータ寄託者の間の分配方法については、一定のルールが定まっていないことから、利用開始時に取り決めを交わしておくことが望ましい。
- ④ 寄託する項目/しない項目の判断:①で述べたような個人同定の可能性が特に高い項目は、丸め処理をするか、提供しないことが望まれる。また、調査方法や設定が特殊である等何らかの理由により、研究内容を十分に理解しない二次利用者の解析・解釈によって誤った結論を引き出すおそれがある項目についても、提供しないことが推奨される。
- ⑤ データ提供先の制限:二次利用者の資格、利用上の制約等については、各研究関係者の意向を反映するような区分けの周知とそれに沿った対応が求められると考えられた。

D. 考察

米国に比べ日本では、未だ疫学研究データの二次料体制が整っているとはいえない。しかし、日本においても社会科学系では既に多くの調査データがアーカイブ化され、利用されていることから、その促進のための教育や環境整備が望まれる。一方で、疫学研究では生活習慣や生活環境のみならず、個人の心身面の健康状態情報を収集することから、その取り扱いには慎重であることが望まれる。今回、特に追跡の終わったコホート研究の利活用を進めることを念頭に、データ提供の際に従うべきガイダンスに含める必要がある項目を5点に整理した。今後そのような研究の寄託を進める際の指標になるものと考えられる。

E. 結論

特に追跡の終わったコホート研究の利活用を進めることを念頭に、データ提供の際に従うべきガイダンスに含める必要がある項目を5点に整理した。その5点とは、①対象者の個人情報保護、②インフォームド・コンセントと倫理審査、③知的財産権の帰属、④寄託する項目/しない項目の判断、⑤データ提供先の制限、である。

F. 健康機器情報

特になし

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

コホート研究で人口動態統計資料から得た死因情報を データアーカイブに付加して提供するための統計法解釈に関する検討

研究代表者	玉腰暁子	北海道大学大学院医学研究科
研究分担者	大橋靖雄	中央大学理工学部人間総合理工学科
研究分担者	祖父江友孝	大阪大学大学院医学系研究科
研究分担者	磯 博康	大阪大学大学院医学系研究科
研究協力者	辻 一郎	東北大学大学院医学系研究科
研究協力者	友岡史仁	日本大学法学部

研究要旨

現在までのところ、多くのコホート研究では統計法の規定に則り人口動態統計資料から死因情報を得ているため、死因をつけた形でのデータアーカイブの公開利用はできないと考えられる。エンドポイントとしての死因情報は非常に価値が高いものであることから、追跡が終了したコホート研究の二次利用を進めるために、死因情報のソースとなる人口動態統計調査の有効活用の方策を模索することが望まれ、その際の論点を2つ提示した。

A. 目的

通常のコホート研究では、追跡情報として重要な死因情報を人口動態統計資料(死亡個票)から、定められた手続きを経て入手する。この人口動態統計資料は、統計法に基づき作成されるところ、同法40条では、同法に特別の定めがある場合を除き行政機関の長等が行った統計調査の目的以外の目的のために、当該統計調査に係る調査票情報を自ら利用したり提供してはならないとされていることから、研究目的での提供は同法第三章に規定される一定の条件の下で行われる。したがって、我々研究者が死因情報の提供を受ける多くの場合は、同法33条2号の規定が適用されるところ、入手にあたっては、公的研究費を受け続けていることなど、以下の様な一定の要件が課せられることになる。

平成23年6月厚生労働省「調査票情報の提供に関する利用申出手引」における

- イ 公的機関から委託を受け、又は公的機関と共同して行う調査研究に係る統計の作成等、
- ロ その実施に要する費用の全部又は一部を公的機関が公募の方法により補助して行う調査研究に係る統計の作成等、
- ハ 行政機関の長又は地方公共団体の長その他の執行機関が、その政策の企画、立案等に有用であると認める統計の作成等。

しかし、このような要件は、追跡が終わったコホート研究にとって決して低いハードルではない。総じて研究者は、複数の研究を実施しているものであること、競争的資金の場合は確実に研究費が得られるとは限らないことなどが、この理由として挙げられる。

さらに、統計法43条2項の規定によれば、同法に基づき提供を受けた情報(コホート研究では死因、死亡の事実と死因年月日は通常別の情報源から入手される)について、「当該調査票情報又は当該匿名データをその提供を受けた目的以外の目的のために自ら利用し、又は提供してはならない」とされているため、対象者の個人情報秘匿した上でもデータの二次利用への提供は認められていないと解される。

以上にあつて、データの二次利用により、公費を投入し多くの人々の協力を得て収集された試料・情報の有効利用、若手研究者の育成、データの検証等幅広い研究利用とその成果に直結し、ひいてはより正確な死因解明に具体的に貢献しうることを考え合わせれば、上記に掲げた統計法の規定による制約を受けない形で、死因も含めたより広い形による該当データの利用可能性が切望される。このため、人口動態統計資料(死因個票)の利用を可能にし得る現実的な考え方について、情報関係法制に詳しい法律学の専門家である友岡史仁日本大学法学部教授を交え、意見交換を行った。

B. 方法

統計法の位置づけと人口動態統計資料の利用価値の観点から、検討を行った。

(倫理面への配慮)

特になし

C. 結果

統計法に基づく調査は機関統計調査として義務的に行われ、その目的は公的統計の作成に限定されているため、当該調査を「研究目的」といった同法の目的外利用に供する場合は厳格に制限しておこうというのが趣旨と解すべきことになる。その場合、コホート研究に不可欠とされる死因情報を人口動態統計資料から得ることも目的外利用を意味するため、その活用に当たり、統計法上設けられた諸手続といった一定の制約が生ずる。しかし、確かに、統計法の限界がある一方、同法以外に個別に死因情報を取得することの困

難さ、今までに行われているほとんど全てのコホート研究で同法に則り死因情報が取得されてきた現状に照らし、当該情報を長期的かつ利用者の範囲を広げてコホート研究に有効活用する道を何らかの形で法制度上担保する必要性を論じておく意義は、十分にある。

この場合の論点として、①統計法の解釈運用を変更することで足りるのか、②統計法とは異なる別立法を行う必要があるのか、という二点が考えられる。

①の場合は、上記の制度趣旨と過去の運用の経緯に照らして乗り越えるべき解釈論を構築する必要があると考えられ、その意味で、死因情報以外の件についても考慮に入れた制度全体の観点から議論を要するものと思われる。対して②の場合は、①とは異なり立法的措置により解決するため、特別法として位置付けられ、法解釈上の齟齬はきたさないとされる一方、既存の仕組みとは異なる新たな制度設計を求められる分、個人情報保護に照らした保護に係る必要なスタンダードを充足する詳細な議論を要する。①または②のいずれを選択するにせよ、新たな立法の煩雑さや既存の仕組みの有効活用という効率性の視点からすれば、現時点では①の選択肢が方向性として現実的と考えられる。

D. 考察

今回の論点は、法の解釈、または立法により、コホート研究において重要な追跡情報となる死因を公的調査情報(人口動態統計資料)から得た上で二次利用を可能にしようとするものである。統計法の従来の解釈経緯などといった従前のコホート研究とは異なる方向性が必要とされており、医学分野の知見だけではない法学分野からの知見に照らし、今後、より視野を広げた説得力ある議論の展開が不可欠である。

E. 結論

現在までのところ、多くのコホート研究では統計法の規定に則り人口動態統計資料から死因情報を得ているため、死因をつけた形でのデータアーカイブの公

開利用はできないと考えられる。エンドポイントとしての死因情報は非常に価値が高いものであることから、追跡が終了したコホート研究の二次利用を進めるために、死因情報のソースとなる人口動態統計調査の有効活用の方策を模索することが望まれ、その際の論点を2つ提示した。

F. 健康機器情報

特になし

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

個人情報保護法改正に伴う疫学研究への影響に関する検討

研究分担者	磯 博康	大阪大学大学院医学系研究科
研究分担者	辻 一郎	東北大学大学院医学系研究科
研究分担者	祖父江友孝	大阪大学大学院医学系研究科
研究代表者	玉腰暁子	北海道大学大学院医学研究科
研究協力者	三浦克之	滋賀医科大学大学院医学系研究科

研究要旨

改正された個人情報保護法において、病歴が要配慮個人情報に位置づけられたことから、疫学研究において生じうる影響を検討した。その結果、追跡情報の入手を医療機関等から受けている場合に、過剰反応により情報が提供されないことが懸念された。

A. 目的

個人情報保護法が 2015 年 9 月に改正され、2 年以内に施行されることとなった。この改正では、個人情報の定義を明確化することによりグレーゾーンを解決し、また誰の情報かわからないように加工された匿名加工情報について自由な利活用を可能とし、経済を活性化することが図られる一方、不正な提供には罰則が設けられ不正な個人情報流通を抑止する措置がとられた。さらに、その取り扱いによって差別や偏見、その他の不利益が生じるおそれがあるため、特に慎重な取り扱いが求められる個人情報として「要配慮個人情報」が定められたが、この中に本人の人種、信条、社会的身分、犯罪の経歴、犯罪により害を被った事実とともに病歴が加えられた。要配慮個人情報では、取得にあたっては原則として本人の同意を得ること、オプトアウト手続きによる第三者提供が認められないこととされている。ただし、関連性を有する範囲内で利用目的を変更したり、匿名加工情報に加工して第三者に提供したりすることは可能である。

喫煙や食事、運動などに代表される生活習慣や病歴等と人々の健康との関連を明らかにし、国民の健康的な生活を守るための対策樹立に資する(実例:健康日本21、喫煙対策、日本人のためのがん予防法など)ことを目的に、多くの疫学研究、特にコホート研究が実施されてきた。研究開始時に収集される情報に病歴が含まれることが少なくなく、また追跡情報では死因のみならず、病院等の協力により疾病罹患情報が用いられる。これから開始される研究はもちろん、2000 年代に入って指針や個人情報保護法の施行以降に開始された研究では、対象者への説明・同意時に取得・提供項目に関し配慮が行き届いていると考えられるが、それ以前の(現在の施策に活かされる成果を輩出してきた)コホート研究では、必ずしも十分な対応がなされていたとは言い難い面がある。そこで、今回、病歴が要配慮個人情報に位置づけられたことにより、以前から行われているコホート研究において生じうる影響、さらに倫理指針の例外規定が正しく認知されていないために疫学研究全般に対し起こりうる事

例につき検討した。

B. 方法

個人情報保護法の条文、特に要配慮個人情報に関する規定を確認するとともに、疫学会の関係者から各々の研究の現在の運用等に関する情報を得るとともに議論を行った。

(倫理面への配慮)

特になし

C. 結果

【以前から行われているコホート研究】

指針等の整備前から開始されたコホート研究であっても、開始以降に倫理審査を受けるなど、研究者は可能な限り対象者の個人情報保護に配慮して研究を行っていた。また、研究開始時に得ている病歴に関しては、本人の自己申告によるものが多く、また医療機関や健診機関から得ている場合でも、その利用に関しては何らかの説明が行われていた。しかし、対象者が多いこと、開始時期が古く対象者が研究に参加していることを記憶していない場合もあること、遡っての同意手続きによって対象者の偏りが懸念されることなどにより、開始時点の説明が現在の基準に照らした場合に不十分であっても改めての同意手続き等を経ることなく、病院から追跡情報として疾病情報を受けている研究も存在した。したがって、今回、病歴が第三者提供に際して対象者からの同意が必須である要配慮個人情報と位置づけられたことから、医療機関等の過剰反応によりコホート研究に重要な追跡情報を得る際に支障を来すことが懸念された。また、仮に、改めて再同意手続きをとり同意が得られた者のみを観察対象とした場合には、対象者の偏りを生じることから研究の価値を著しく損ね、ひいては公衆衛生上の不利益につながるものが危惧された。

一方、本研究班で検討している追跡終了後のコホート研究をアーカイブ化し、二次利用に用いる際には、いずれせよ対象者を遡れないような匿名化措置が必要と考えられることから、既に収集している病歴を

あわせて提供することは差し支えないと考えられた(死因に関しては、統計法の規定によるため、現状では提供不可)。ただし、非常に稀な病気の場合等は、詳細な病名ではなく、ある程度まとめた疾患概念として示すなどの配慮を行うことが必要であろう。

【疾病登録による罹患率等の把握のための疫学研究】

学会や研究者が主体となって地域の医療機関等から医療情報を収集・整理し、循環器疾患(脳血管疾患、心疾患)や腎透析、難治性疾患、新興感染症など、今後予防法や治療法の解明を進めなければならない多くの病気の罹患率、予後、地域差等の実態を明らかにするための疾病登録研究がある(がんはがん登録法が平成28年より施行されるため除く)。このような研究は予防法の解明や医療資源の投入量の判断などに必要である。しかし、同意した患者のみを対象とする研究では正確な罹患率や予後等の把握ができずに誤った解釈につながり、公衆衛生上のリスクを生じることになる。したがって例外規定に該当するものとして研究を実施する必要があるが、医療機関等が要配慮個人情報の「病歴」として本人同意のない情報提供を拒否した場合、このような研究を実施することができなくなる。また、この研究においては、患者の予後の実態を明らかにするため住民基本台帳や死亡小票、人口動態統計データを行政機関で閲覧・照合する必要があるが、この情報提供が拒否されると研究の実施が困難になる。また、新興感染症の流行、化学物質や放射線の曝露を受けるような事故等があった場合、地域別の疾病動向は同意の有無の関わらず全数調査されるべきであり、同意を前提とすると誤判断につながり、社会的な危機につながる。

【特定の疾患の患者の予後及び予後要因を明らかにする疫学研究(臨床研究)】

特定の疾患の患者の予後および予後要因を明らかにするための疫学研究(臨床研究)が、多数の医療機関が共同して実施されることがある。この場合、地域での罹患率を明らかにすることは必ずしも目的にしているが、予後の実態と予後規定因子や治療効果を明らかにする上で必要不可欠な研究である。多くの