

## I. 総括研究報告

厚生労働科学研究費補助金（がん対策推進総合研究事業）  
（総括）研究報告書

全国がん登録を基盤とした長期記述疫学研究用特定匿名化情報の整備に関する研究  
研究代表者 柴田 亜希子 国立がん研究センターがん対策情報センターがん登録センター 室長

**研究要旨**

本研究は、平成 31 年 1 月予定の全国がん登録情報の提供開始の前に、米国の SEER をモデルとして、我が国の法令と照らし、我が国の実情に即して、長期持続可能な、全国がん登録を基盤とする「特定匿名化情報」の提供並びに「がん診療情報」の収集の仕組み（日本版 SEER）の提案を目的とする。平成 30 年度は、日本版 SEER の提案にかかる議論の幅を広げるために、仕組み検討班は「長期持続可能性」、「がん診療情報」の収集、「匿名化処理と易利用性」に関して、米国がん登録関係者、都道府県、病院に対して追加の調査検討を行った。記述疫学研究班は、昨年度に入手した SEER Data を用いた計 4 つの記述疫学研究の分析を進めた。作業部会の結果は計 3 回の研究会議にて報告され、その内容を踏まえて日本版 SEER のあり方について議論した。その成果として、本研究班は、全国がん登録を基盤とした長期記述疫学研究用匿名化情報の 5 要件を提案する。①未来の長期記述疫学研究の継続性を確保するため、がん登録等の推進に関する法律の定める全国がん登録情報等の提供として整備すること。②SEER と同様に、詳細ながん診療情報を含めて、登録の完全性と正確性を備えた住民ベースの記述疫学研究が可能な情報であること。③収集する詳細ながん診療情報は、病院等への届出負担を増大させずに正確な情報となるように、臨床的意義が高く医療者が自ら届け出たくなる内容であること。④長期記述疫学研究用情報の匿名化処理は、米国を参考に、研究に用いる場合の統計学的信頼性の担保と個人識別可能性の低減のバランスを考慮して行うこと。⑤長期記述疫学研究用に匿名化処理された情報は、施設や個人を特定する目的に利用しないこと、特定しようとするしないことを条件として、簡便に利用できる形で整備すること。要件を満たすためには法改正が必要であるが、実現後は、世界でも数少ない大規模の住民ベースのがん記述疫学データベースとなる。

**研究分担者**

伊藤秀美  
愛知県がんセンター・研究所  
がん情報・対策研究分野 分野長  
井上真奈美  
（国研）国立がん研究センター  
社会と健康研究センター  
予防研究部 部長  
大木いずみ  
（地独）栃木県立がんセンター  
がん予防情報相談部 部長  
金村政輝  
宮城県立がんセンター研究所  
がん疫学・予防研究部 部長  
西野善一  
金沢医科大学  
医学部公衆衛生学 教授  
林櫻松  
愛知医科大学  
医学部公衆衛生学 教授（特任）

東尚弘  
（国研）国立がん研究センター  
がん対策情報センター  
がん登録センター センター長  
片野田耕太  
（国研）国立がん研究センター  
がん対策情報センター  
がん統計・総合解析研究部 部長  
堀芽久美  
（国研）国立がん研究センター  
がん対策情報センター  
がん統計・総合解析研究部 研究員  
松田智大  
国立がん研究センター  
がん対策情報センター  
がん登録センター 全国がん登録室長

**研究協力者**

平田公一  
札幌医科大学消化器・総合、乳腺・内分泌  
外科 客員研究員

## A. 研究目的

2016年から施行されているがん登録等の推進に関する法律(がん登録推進法)では、登録情報の十分な活用と国民への結果の還元が求められている。米国では、National Cancer Instituteが、協力州から基本的な州がん登録情報その他のがん診療情報の提供を受けて SEER (The Surveillance, Epidemiology, and End Results) Data を約 40 年以上にわたって整備維持している。そのデータは簡単な申請で研究者に提供されているため、定型の公表集計値以外の部位や組織型の罹患率や生存率の記述疫学研究等に幅広く活用されている。我が国ではがん登録推進法に基づき保存された 2016 年罹患の全国がん登録情報の提供が 2019 年 1 月に開始予定である。本研究は、全国がん登録情報の提供開始の前に、がん登録推進法その他関連法令等と照らし、我が国の実情に即して長期持続可能な、全国がん登録を基盤とする住民ベースのがん罹患の匿名化された情報の提供並びに全国がん登録届出情報以外の「がん診療情報」の収集の仕組み(日本版 SEER)の提案を目的とする。(図 1)。

## B. 研究方法

SEER の制度や組織の仕組みを理解するために、①ウェブサイト等の公開情報の調査及び公開情報で不明な点の現地聞き取り調査、②SEER Data の内容や利用手続きを理解するための SEER Data を用いた記述疫学研究を 2 年計画で実施する計画とする。研究班は、研究分担者の実績に応じて、前者を担当する仕組み検討班と後者を担当する記

述疫学研究班の 2 つの作業部会で構成する。仕組み検討班の研究分担者は、将来的に、SEER 協力州に相当する都道府県の候補となり得る都道府県の、がん登録に見識のある研究者とした。

1 年目の昨年度は、仕組み検討班は SEER のウェブサイト等の公開情報調査と現地訪問による非公開情報の聞き取り調査を行った。記述疫学研究班は、実際に利用申請して入手した米国 SEER Data を用いる記述疫学研究を通して、利用申請の実際を知るとともに、容易な手続きで提供されるデータ内容及び匿名化の程度を確認した。その結果、SEER Data の実利用、ユタ州がん登録と SEER 事務局の調査分析を通して、日本版 SEER 構築においては、SEER の仕組みの基盤である長期に渡る公的研究事業が我が国では維持されにくいこと、研究と個人情報保護のバランスに関する米国と我が国の法体系及び国民意識の違いを考慮する必要性が認識された。

さらに、North American Association of Central Cancer Registries (NAACCR) という組織が、米国における、組織や目的の異なるがん登録が収集する項目の標準化に大きな役割を果たしていることが分かった。また、Centers for Disease Control and Prevention (CDC) の事業である National Program of Cancer Registries (NPCR) が、SEER の公共利用データ提供の仕組みである SEER\*stat というシステムを共同利用して、SEER がカバーしていない州のデータを含む公共利用データを提供していることが分かった。

年度	平成29 (1年目)				平成30 (2年目)						
年次	2017				2018				2019		
月	1-3	4-6	7-9	10-12	1-3	4-6	7-9	10-12	1-3	4-6	7-9
全国がん登録予定			全国がん登録情報の提供に係る厚生科学審議会がん登録部会		全国がん登録情報の提供マニュアル	全国がん登録情報の提供に係る厚生科学審議会がん登録部会	提供マニュアル更新	審議会等による提供審査	情報提供開始予定		
研究班予定	仕組み作りWG 記述疫学WG	SEER公開情報による知見収集 記述疫学研究計画利用手続	班会議 6月 SEER 閣取調査 8月末 提供データ分析	班会議 10月 調査分析知見統合と日本版SEERの検討班会議 1月	A 班会議 5月	B 班会議 7月	米国における住民ベースのがん登録情報の活用 の情報収集追加	日本版SEER 将来構想 記述疫学研究 成果公表	A:「提供のマニュアル」を理解し、現状可能な日本版SEERの提供形式を共有 B:審議会がん登録部会審議会内容の確認と日本版SEERの将来構想継続		

図1 平成29年度からの2年の研究計画（平成30年度一部改変）

以上を踏まえ、平成30年度は、日本版SEERの提案にかかる議論の幅を広げるために、仕組み検討班は「長期持続可能性」、「がん診療情報」の収集、「匿名化と易利用性」に関して、米国がん登録関係者、都道府県、病院に対して調査を行う。

第一に、ウェブサイト等の公開情報の精査後、NAACCRとNPCRの双方に詳しいCDCのNPCR事業担当者を招へいし、聞き取り調査を行う。第二に、その他の詳細ながん診療情報の収集を、SEERと同様に住民ベースのがん登録の拡張の形で行うとすれば、届出病院に今以上の負担をかけることなく、かつ正確に診療情報を登録できる方法でなくてはならない。そのため、分担研究者が所属する病院において、病院に既存の臨床研究用データベースと院内がん登録との連携について、現状調査を行う。最後に、全国がん登録情報を用いた長期記述疫学研究用匿名化情報データベースを整備して、その価値が発揮されるのは少なくとも10年後である。そのため、本研究班では、2015年以前に診

断されたがんが登録されている都道府県単位の地域がん登録情報を用いて、将来の日本版SEERに近い匿名化データベースを国立がん研究センターが整備し、公共利用に提供を行う案の現実性を検討するために、2018年8月、仕組み検討班の分担研究者の関わる都道府県（宮城県、栃木県、愛知県）のがん登録の行政担当者に、共通の依頼文書を用いて、そのようなデータベース構築を目的とする国立がん研究センターへのデータ提供の在り方について、分担研究者を介して意見招請を行う。

記述疫学研究班は、昨年度に入手したSEER Dataを用いた計4つの記述疫学研究の分析を行う。

（倫理面への配慮）

本研究においては人体から採取された試料は用いない。記述疫学研究においては、研究用に整備された匿名化された情報を用いる予定であるが、「人を対象とする医学系研究に関する倫理指針」に準拠し、必要に応じて分担研究者の所属する倫理審査委員会の

審査を受けるものとする。他の公的統計資料を利用する場合は、それぞれの利用手続きに則る。

## C. 研究結果

### 米国における住民ベースのがん登録の公共利用データ提供体制に関する調査

主に、NPCR の公共利用データ提供 (Public Use Research Database Data Use) のウェブサイト (NPCR and SEER Incidence - U.S. Cancer Statistics Public Use Databases [ <https://www.cdc.gov/cancer/uscs/public-use/index.htm> ] (最終検索日 : 2019 年 5 月 28 日) に掲載されている利用規約 (資料 1) と『よくある質問と回答』の内容を調査した。公共利用データ提供において提供されるデータの条件として、郡レベルを含む郡より小さい地理情報や 1 歳年齢階級は利用できないこと等が明示されていた。また、NPCR は秘匿すべき少数集計値を 16 未満と定めており、その理由として統計値としての不安定性を説明していた。

我々は、特定の施設または人物を識別可能としないために、グルーピングという匿名化手法を用いても、州の人口規模やがんの好発年齢の違いによっては厳密な匿名性を得られていない場合はどうするのか、常に集計値 16 未満のために秘匿対象となってしまう希少がんの研究者から不満はないか等の疑問を持ち、聞き取り調査で確認することとした。しかし、招へい予定者の都合が合わず、11 月に開催される国際がん登録協議会年次総会に NAACCR の担当者が参加予定であることを確認できたため、同学会に参加予定の仕組み検討班の分担研究者 (大木、伊藤、松田) との現地での打ち合わ

せを調整した。

打ち合わせ聞き取り結果によると、集計値 16 未満の場合、常に秘匿処理を行うことについて、集計値として統計学的信頼性に乏しいのであるから当然であるとのことであった。利用規約等で定められた匿名化処理を行った上でも残る施設や個人の識別可能性については、そもそも利用規約で施設や個人を特定しようとする行為を禁じているので、それが守られないことは想定しないとのことであった。最後に、SEER、NPCR、米国外科学会 (ACS) が運営する National Clinical Database (NCDB) 等の組織や目的の異なるがん登録が収集する項目の標準化に NAACCR が大きな役割を果たすようになったことについて、10 年以上に渡る関係者の合意形成の話し合いのたゆまぬ努力の結果であり、それぞれの国であらゆる状況が異なるのだから、国内で話し合って決めること、との助言を得た。

### 2015 年以前の地域がん登録情報による長期記述疫学研究用匿名化情報データベース構築に関する都道府県意見招請

3 県とも、がん登録推進法に基づく情報の提供の仕組みに準じて考えるため、あらかじめ匿名化されたデータであっても、国立がん研究センターへの「譲渡」や「データ管理委任」について合理的理由がなく、国立がん研究センターが「研究」として、毎年、必要な手続きを経て構築すべきであるとの意見であった (資料 2)。

### 院内がん登録を基盤とするその他のがん診療情報の収集可能性について

分担研究者の金村は、本人の所属する宮

城県立がんセンターでは、先行研究によると学会等が運営する 14 のがん登録のうちの 11 を含む、合わせて 12 の登録が行われており、すべて各登録が構築した Web ベースへの直接入力のため、電子カルテや院内がん登録データベースとの連携はないことを明らかにした。また、院内がん登録を介して効率よく追加情報収集するためには、データのリレーションが可能となるよう各登録の事前のすり合わせが不可欠であり、そのためには、国立がん研究センターなどがイニシアチブを発揮し、各学会の協力を得て、登録間の協調・連携を進める必要があると考えられるとした。

#### SEER データを用いた記述疫学研究

分担研究者の伊藤は、多発性骨髄腫の死亡に新規治療法が与える影響について検討するため、日米のがん罹患、死亡情報を利用した研究を行った。日米において、新薬の登場により、多発性骨髄腫では死亡率減少率が大幅にアップしていることがわかった。死亡に影響を与える罹患の動向を見ることも本研究には重要なポイントで、罹患の減少と死亡の減少はパラレルではないことを確認した。日本だけでなく米国 SEER データを利用できたことで、日本で観察された現象を米国でも観察でき、妥当性について評価できた点が有益であった。

分担研究者の林は、日米ともに増加傾向にある膵がんについて、罹患率・死亡率の年次推移、診断時ステージ、組織型分布、ステージ別 5 年生存率の日米比較を行った。本研究では、膵がんは日米とも増加傾向にあり、日系アメリカ人はアジア系アメリカ人の中で罹患率が高く、また一貫して顕著な

増加傾向を示すことを明らかにした。膵がんに関する記述疫学は日米それぞれ実施されているものの、両国を比較した詳細な記述疫学は数少ない。両者を比較することで、日本を含む日系における未知の膵がんリスクの可能性を強く想起させる結果を得られた。

分担研究者の堀と片野田は、卵巣がんの年齢別、組織型別罹患率の日米比較を計画した。日本の卵巣癌の生存率は近年、向上し、現在では北米や欧州よりも高いとの報告がある。また、卵巣癌の生存率は、組織型別に大きく違うことも知られている。生存率の差の原因を組織型の分布の違いによると仮定して、同研究を開始し、卵巣癌の組織型罹患率が人種によって差があることを明らかにした。しかし、日本の生存率の高さの原因を裏付ける結果は得られなかった。同じ組織型でも異型度や病期によって生存率が異なることが知られているため、異型度や進行度分布についても比較がさらに必要である。SEER データでは異型度の集計が可能であり、その情報の信頼性についての研究も進められている。組織型、異型度に限らず、詳細な罹患状況を他国と比較する場合、登録情報の質に関する比較が必要である。日本版 SEER では、詳細項目の収集だけではなく、登録情報の信頼性の担保も必要であるとした。

以上の結果は、計 3 回の研究班会議にて報告され、その内容を踏まえて、日本版 SEER のあり方について議論した（資料 3）。

#### その他の成果

分担研究者の井上は、本研究班を通じて SEER データ及び全国がん登録情報に関する

理解を深めた上で、国際的な記述疫学の専門家の立場で日本版 SEER のあり方について検討した。SEER 同様に人口ベースであることは、必ずしも全国がん登録で詳細ながん診療情報の収集が十分にできない項目があるとしても、日本版 SEER では必要不可欠な要素であるとまとめた。人口ベースであれば、単に治療などの生存率への影響を評価できるのみでなく、その集団(国や地域)へのインパクトを評価できるという利点がある。また、もしこの要件が欠ければ、類似の病院施設などで任意に作成されている患者登録データベースでも代用できてしまい、日本版 SEER をあえて作成する必要性がなくなってしまうと考察した。

分担研究者の西野は、2018年9月に国と国立がん研究班が示した「全国がん登録情報の提供マニュアル第2版」別添の「全国がん登録 利用者の安全管理措置」に基づいて、全国がん登録情報の利用者が安全管理措置を実施する上での課題を検討した。個人情報の物理的保存を行っている区画の二重の施錠は利用環境によっては実施が困難な可能性がある。また、安全管理措置によって情報漏えい等のリスクを減らすためには、情報の利用者、利用および保管場所を可能な限り限定することが重要であることを統括利用責任者、利用責任者は認識し環境に応じた適切な対策を作り上げることが求められると結論づけた。

#### D. 考察

本研究班は、全国がん登録を基盤とした長期記述疫学研究用匿名化情報の要件を以下のとおり提案する。

1. 未来の長期記述疫学研究(最低20年以

上)の継続性を確保するため、SEERとは異なり、有期の研究事業ではなく、がん登録推進法の定める全国がん登録情報等の提供として提供する情報の一つとして整備すること。

2. SEERと同様に、詳細ながん診療情報を含めて、登録の完全性と正確性を備えた住民ベースの記述疫学研究が可能な情報であること。
3. 収集する詳細ながん診療情報は、病院等への届出負担を増大させずに正確な情報となるように、臨床的意義が高く医療者が自ら届け出たくなる内容であること。そのため、SEERと同様に、詳細ながん診療情報を収集するがん種や期間を限定する手法も考えられる。
4. 長期記述疫学研究用情報の匿名化処理は、米国を参考に、研究に用いる場合の統計学的信頼性の担保と個人識別可能性の低減のバランスを考慮して行うこと。
5. 長期記述疫学研究用に匿名化処理された情報は、施設や個人を特定する目的に利用しないこと、特定しようとしなかったことを条件として、簡便に利用できる形で整備すること。

上記提案に至った背景と今後の課題について記述する。

米国には、我が国のがん登録推進法に基づく全国がん登録事業のような、合衆国法に基づく全米がん登録事業は存在しない。住民ベースのがん登録は、各州法に基づく州がん登録が最大の運営単位である。SEER及びNPCRは、それぞれの組織であるNCI及びCDC予算による研究事業として、州との

契約に基づいて匿名化が行われたデータの提供を受けて、より大きな住民単位のがん登録データを作っている。

我が国も、がん登録推進法施行以前は米国と類似した構造であって、各都道府県が独自事業として地域がん登録を運営しており、2003年以降は厚生労働省の研究班が都道府県からデータの提供を受けて、Monitoring of Cancer Incidence Japan (MCIJ)として日本全体のがん罹患数・率を報告してきた。SEER及びNPCRは、各州がん登録との契約において、標準化された質のよいがん登録を整備維持するためのシステム、マニュアルや資金を州がん登録に提供し、対価としてデータの提供を受ける関係性である。我が国の厚生労働省研究班も同様の関係性の構築を目指していたが、資金が米国の10分の1に過ぎない、個人情報保護や公衆衛生対策に対する国及び都道府県の法体系や統治体制が異なる、同じ名目の研究事業が何十年も継続された前例がないなどの理由から、都道府県毎の標準化の進捗や質の向上が次第に頭打ちになってきたことも、がん登録推進法に基づく国直轄運営の住民ベースのがん登録が開始される要因の一つであった。

以上の歴史的背景より、我が国では、長期記述疫学研究用情報の提供の仕組みだけ、研究事業で構築することに合理的理由が見当たらず、がん登録推進法の定める全国がん登録情報等の提供の仕組みの中で構築、運営するのが自然であると考えられた。その場合、SEERのように協力する州のみの参加型ではなく、全47都道府県が対象になる。

しかしながら、現行のがん登録推進法では、全国がん登録として収集する項目につ

いては、厚生労働大臣が省令で定めることになっており、詳細ながん診療情報をタイムリーに追加収集することは困難である。この問題を解決する正攻法は、がん登録推進法の改正である。具体的には、がん登録推進法第3条第3項の基本理念「がん診療情報の収集が図られなければならない」の具体的手法を省令等で定めることである。

がん登録推進法に基づき新たに収集することになる、がん対策の充実に寄与する、がんの診療の状況を適確に把握することのできる、診療に関する詳細な情報(がん診療情報)の選択においては、我が国のがん登録の情報収集は病院等からの届出を基本とすることから、病院等に今以上の負担をかけないことや病院等への対価を考慮しなければならない。全国がん登録情報の約8割は院内がん登録由来と考えられることから、少なくとも院内がん登録を介して正確な情報を収集できなければならない。

本研究班の議論の過程では、住民対象ではなく、日本版SEERとして、院内がん登録を実施する病院だけを対象として詳細ながん診療情報を収集し、その活用を行う仕組みの構築も検討された。しかし、それでは、記述疫学研究班が行い、考察で述べているような日米比較とそれによる新たな知見を見出す成果は期待できない。整備が困難な道であっても、本研究班が日本版SEERとして提案するのは、米国SEERと同様に住民ベースの登録である。

米国とは医療提供サービスの異なる我が国の事情を勘案したがん診療情報の収集方法については、我が国では学術研究として実施されている学会等が運営する専門性の高いがん登録(通称「臓器がん登録」)が複



数存在するので、これを利用すべきである。分担研究者の金村の報告のとおり、現状ではそれぞれの臓器がん登録が、それぞれのウェブ登録システムを構築して運営されており、協力する学会員が所属する病院等の院内がん登録との連携はない。データのリレーションが可能となるように臓器がん登録と院内がん登録の事前の前向きの対話が不可欠であり、そのためには、NAACCRのように個々の運営母体に関係なくがん登録でつながるか、いずれかの組織が主導的に動き、協調してデータ連携と質の向上を目指す必要がある。

本研究班の課題名に含む「特定匿名化情報」は、がん登録推進法における「特定匿名化情報」を指す（法律第2条第10項）。がん登録推進法上の匿名化は、「他の情報との照合による識別を含め、当該がん罹患した者の識別ができないように加工すること」と定義されており（法律第2条第9項）、提供の求めを受ける頻度が高い、すなわち利用者の需要が高いと見込まれる匿名化情報を、あらかじめ全国がん登録データベースに記録することができる（法律第21条第5項）とされている。本研究の計画段階では、特定匿名化情報を整備できれば、米国SEERのように容易に利用者に当該情報を提供できると誤解していた。しかし、実際には、匿名化が行われた全国がん登録情報の提供（利用者視点では利用）には、特定匿名化情報であっても、法律第21条第4項の手続きが必要であり、保護規定によって情報の保有制限や安全管理措置が義務づけられている。米国SEERと同等に匿名化された情報の場合は、利用者側の利用規約への誓約を以て提供できる道を法改正によって確保する

必要がある。

以上が、2016年を初年として蓄積されていく全国がん登録情報を用いた長期記述疫学研究用匿名化データベースとして、本研究班が考える要件である。しかし、当該データベースが整備され、その長期記述疫学的価値が発揮されるのは少なくとも10年後である。そのため、本研究班では、2015年以前に診断されたがんが登録されている都道府県単位の地域がん登録情報を用いて、提言に近い匿名化データベースを国立がん研究センターが整備し、提供を行う案を検討した。しかしながら、意見を聴いた3県ともに、地域がん登録情報の提供手続きも今はがん登録推進法に準拠することを理由に、そのようなデータベースの構築と運用は非現実的との意見であった。このことから、全国がん登録情報を用いた長期記述疫学研究用匿名化データベースは、国のがん対策の企画立案のために国が自ら全国がん登録情報を用いて構築し、そのデータベースの利用はがん登録推進法の適用外であることが望ましい。そして、我が国におけるがんに関する記述疫学研究が中断することのないよう、米国SEERに及ばずとも、現行のがん登録推進法の規制の中で可能な限りの2015年以前に診断されたがんの記述疫学研究用匿名化データベースを整備する必要がある。

## E. 結論

本研究班は、米国SEERをモデル、目標とした、全国がん登録を基盤とした長期記述疫学研究用匿名化情報の整備に関して5つの要件提案を行った。以上の要件を満たすためにはがん登録推進法の改正が必要で、2019年1月からの全国がん登録情報の提供

の開始に合わせた整備は間に合わなかったが、実現後は、世界でも数少ない大規模の住民ベースのがん記述疫学データベースとなる。

ベース構築に関する都道府県意見招請（依頼文書及び回答）

資料3 班会議議事録

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

なし

2. 学会発表

1) 柴田亜希子. がん登録推進法と全国がん登録情報の提供について. 第29回日本疫学会学術集会（企画講演）：2019.2：東京都.

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

資料1 National Program of Cancer Registries (NPCR) and Surveillance, Epidemiology and End Results (SEER) Incidence - United States Cancer Surveillance (USCS) Public Use Research Database Data Use Agreement For data submitted in November, 2016

資料2 2015年以前の地域がん登録情報による長期記述疫学研究用匿名化情報データ