

I. 総合研究報告

厚生労働科学研究費補助金（がん対策推進総合研究事業）
総合研究報告書

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 全国がん登録分析室長

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、がん登録等の推進に関する法律（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「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。SEER の制度や組織の仕組みを理解するために、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聞き取り調査と、SEER Data の内容や利用手続きを理解するために SEER Data を用いた記述疫学研究を行った。それらの調査結果を踏まえて、班会議にて日本版 SEER のあるべき姿を段階的に議論した。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、がん登録推進法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、特定しようとしめ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。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あるが、実現すれば世界でも数少ない大規模な住民ベースの公共利用がんデータベースの価値がある。

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 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

がん登録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。この基本理念の推進を目的として、提供要望頻度が高いと見込まれる情報について、あらかじめ、全国がん登録情報の匿名化情報（特定匿名化情報）を整備し、全国がん登録データベースに記録できる、とされている。また、基本理念には、がん対策の充実のためには、全国がん登録の実施のほか、がんの診療の状況を適確に把握することが必要であることに鑑み、院内がん登録により得られる情報その他のがんの診療に関する詳細な情報（がん診療情報）の収集が図られなければならない、とされている。

米国では、National Cancer Institute (NCI) 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。SEER Data の特徴は、最大で協力 18 州のデータに限られるが信頼性の高い罹患率と生存率を算出できること、時代時代で関心の高い情報を追加で収集し、整備されていることである。我が国では 10 数年前から米国を参考に院内がん登録及び地域がん登録の仕組みの標準化と登録精度向上に努め、2013 年にがん登録推進法が成立した。同法に基づき保存される 2016 年罹患の全国がん登録情報の提供は 2019 年 1 月に開始予定である。研究代表者らは、平成

27 年度から、全国がん登録情報の第三者提供に関する国民意識の調査並びにがん診療情報の収集方法の一つとして既存のがん診療情報との連携可能性に関する研究を続けてきた。その結果を踏まえ、我が国で SEER と同等の情報提供が可能であれば、調査研究への速やかな積極的な活用を促し、日米等共同研究に対して比較可能な基本的な情報を提供できると考える。

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、SEER をモデルとして、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「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。

B. 研究方法

SEER の制度や組織の仕組みを理解するために、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聞き取り調査を、SEER Data の内容や利用手続きを理解するために SEER Data を用いた記述疫学研究を 2 年計画で実施する（図 1）。

研究班は、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し、研究分担者の実績に応じて担当した。特に、仕組み検討班の担当者、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。

仕組み検討班（柴田、松田、東、西野、大木、金村、伊藤）

調査によって以下の点を明らかにする。

- ・ SEER のために収集されたデータを第三者

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				全国がん登録情報提供指針		全国がん登録情報提供説明会		厚生科学審議会がん登録部会による提供審議	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	SEER 調査 提供データ分析 班会議	調査分析知見統合と日本版SEERの検討 班会議		日本版SEER 仕組み案まとめ 記述疫学研究まとめ 班会議	日本版SEER 仕組み案まとめ 記述疫学研究まとめ 班会議		研究総括		

図1 平成29年度からの2年の研究計画（平成30年度一部修正）

の研究者に提供可能な理由。特に病院等での情報取得時の本人通知及び本人同意の有無（主担当：柴田、松田、西野）。

- ・希少がん等の個人識別可能性のあるデータの匿名化処理の有無。匿名化処理を行っている場合はその手法（主担当：東、柴田）。
- ・SEER 特別項目を州がん登録及び各州の病院から届出協力してもらう仕組み。必要経費や教育支援の有無と内容（主担当：大木、金村、伊藤、柴田）。

記述疫学研究班（伊藤、井上、林、片野田、堀）

研究を通じて、以下の点を明らかにする。

- ・利用手続きの実際。
- ・通常の利用申請手続きで利用可能なデータの内容や粒度。
- ・通常の利用申請手続きで利用可能なデータを用いて、どの程度の粒度の記述疫学研究が可能か。

2つの作業部会は、調査の効率化を図るため、それぞれの調査で知り得たことを班会議その他で共有する。

（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。仕組み検討班においては、個々のがん登録情報を用いず、運営や仕組みに関する検討を中心に行うため、個人情報保護上、特に問題は発生しない。記述疫学研究班においては、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、必要に応じて国立がん研究センター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

SEER 事業について

研究分担者の伊藤、片野田、林、堀が、記述疫学研究を計画し、SEER Data の利用手続きを行った。SEER linked database を用いた研究も検討されたが、このデータベースの国外からの利用は受け付けていないことが分かった。一方、SEER Data については、

SEER のウェブサイトからオンラインで利用申請し、利用規約に対する誓約書を電子メールで送付することで、国外からの利用申請が可能で、データが提供された。提供されたデータから判断する限り、氏名は含まないものの、年齢は各歳階級、地理情報は日本では郡に相当する County レベル、がんの組織型については国際腫瘍学分類の組織型約 2,000 分類を含む詳細なデータであった。これらの情報を組み合わせると、集計結果が容易に 1 例になる場合があると考えられた。利用規約において、個人を特定できるデータを公表しない、少数例の集計値の公表を避ける、等の少数集計値による個人識別性への公表時の対応を求め、提供に当たり遵守を誓約させていた。

公開情報や実際に提供されたデータからは、前述以上の匿名加工処理が行われているか分からなかったもので、2017 年 8 月 NCI の SEER 担当部門を柴田、井上、林、東、堀が訪問し、聞き取り調査で確認した。聞き取り調査によって、SEER Data は、氏名除去、生年月日の年齢表記と生年処理、診断日の年月処理、生存期間の生存期間表記、地理情報の County までの処理の匿名加工処理がされていることが分かった。氏名は収集段階で別途管理されており、SEER の担当部署では取り扱わない。法的背景には、Health Insurance Portability and Accountability Act (HIPAA) は、その他の公衆衛生目的の情報収集と同様に、がんサーベイランス事業には適用されないこと、また、研究に関する米国連邦法により、個人情報には除かれているが、原情報に連結可能なデータ (de-identified data) を取り扱う研究は、原情報に連結不可能な非特定化

データ (anonymous data) ではなくても「ヒトを対象とする研究」とは見なされないこと、があることが分かった。さらに、HIPAA が非特定化手法を例示しており、SEER 事業ではそのセーフハーバールールを参考にしているとのことだった。データの提供を受けることの叶わなかった SEER linked database のように、がん登録データベースを他のデータベースとリンクすることについては、前述のとおり、個人情報の管理組織を別にする配慮はしているものの、それに関して個人情報の取扱い上問題だという議論は特にないとのことだった。

SEER Data には、各がんの詳細な病期の情報が含まれていることが最大の特徴である。その他、期間限定ではあるが、腫瘍マーカー、治療手技、ホルモンレセプターの情報も含まれている。このような詳細な診療情報を、住民ベースのがん登録で正確に収集するのは、一般的に容易ではない。我々 (柴田、伊藤、井上、大木、金村) は、SEER に協力している州がん登録がどのような手法で詳細な診療情報を収集しているのかを知るために、2017 年 8 月ユタ州がん登録を調査した。前もって、我々の研究上の関心事項を先方の担当者に連絡し、訪問時に以下の 7 項目の議事に分けて詳細な説明を受けた。

- 1) 州がん登録の概要と環境
- 2) データの収集について
- 3) データの編集、集約について
- 4) データの機密保持について
- 5) 追跡について
- 6) データ照合
- 7) データシェア

ユタ州では州法でユタ州健康局が疾病の実態把握をすることと定められており、ユ

ユタ州がん登録は、州の行政規則によって、ユタ州健康局の代理でがんの情報を管理する組織と定められている。現在は、ユタ大学の一部門である。ユタ州がん登録のデータは、ユタ州健康局に帰属し、情報公開される。その他の情報公開については、年2回、ユタ州健康局、病院、研究者の代表から構成される「研究諮問委員会」で検討される。現在、ユタ州がん登録データは、ユタ州健康局、SEER 事業、Centers for Disease Control and Prevention (CDC) の事業である National Program of Cancer Registries (NPCR)、ユタ住民データベース、研究申請、非研究申請に対して提供されている。ユタ州がん登録と SEER 事業の関係は、単純には「契約」とのことだった。ユタ州がん登録は、契約に従って、SEER から資金その他の支援を得て、SEER の求める診療情報も含めて医療機関から情報収集し、SEER に提出する。その他の複数の研究事業とも同様の協力契約を交わしている。州の行政規則は、がん登録について、収集するデータの形式や変数等の具体的な内容に言及していないため、州がん登録は、SEER 事業、NPCR、その他の研究事業が必要とする項目を収集することができるとしている。ユタ州がん登録では、SEER 事業財源で、複数名の実務者が雇用されている。

さらに、ユタ州の行政規則において、すべての病院その他のがんの診断・治療を行う施設はがん登録に報告しなければならないとされている。これによって、病院等は、届出のための情報の抽出にかかる費用を自ら払わなければならない。病院等は、届出のために、腫瘍登録士を雇用したり、同様のサービスを提供する会社と契約している。ユタ

州の病院等は概ね協力的であり、ユタ州がん登録は病院等に適切な時期の報告について勧奨する連絡は絶やさないが、ユタ州がん登録や州が報告に関する行政規則を行使したことはないとのことであった。

さらに、米国では、がん登録のデータの標準化とその適用が進んでいる。North American Association of Central Cancer Registries (NAACCR) が情報の集約のための標準的なレイアウト形式を提供している。SEER や NPCR は標準方式の決定機関であり、NAACCR は標準方式の普及や標準方式の決定機関の協力機関である。ユタ州には、ユタ州がん登録と院内がん登録を支援するユタ州がん登録協議会という地方組織があり、年に一度の会議の開催の他、日頃から電子メールや電話で連絡をとって、助け合っている。

データの収集について、ユタ州がん登録の特筆すべき点は、独立したがんの情報源として病理報告を含めており、その 88% が e-Path という SEER が開発した電子的病理報告情報収集システムで収集されている。本システムはがん患者のケースファインディング用であり、がん登録の報告用ではない。ユタ州では、病院や検査機関に本システムの導入を進め、がん以外の全病理報告を収集し、ユタ州がん登録において一部機械的判定も使いながら新規がん患者の見つけ出しを行っていた。病理報告があつて、がん登録の届出がない場合には、病院に届出を促すさかのぼり調査を行うので、最終的に病理報告のみでがん登録されているケースは約 5% のことであった。

医療機関からがん登録への届出の方法は、電子的な届出については安全性の確立され

たファイル送達システムや暗号化されたウェブポータルサイトも提供されているが、基本的には届出電子情報が暗号化されていれば、一般的なメール添付も可能とのことであった。手書きの届出も一部残っており、郵送のみならずファックスの届出も受け付けていた

住民ベースのがん登録のその他の基本的な業務であるがん届出情報の確認、編集、既登録腫瘍か否かの判定と情報の集約には、SEER や NPCR が開発して提供しているツールその他の独自開発のツールを使って効率的に問題のあるデータの抽出と編集を行っていた。米国独自の認定資格である「腫瘍登録士」が、データの編集、集約を行い、品質管理に責任を持っていた。

同一人物の情報の連結（リンケージ）は、がんの届出情報同士のみならず、死亡その他の追跡情報との連結の作業とまとめ、独立した業務とされていた。個人識別指標は、姓、名、生年月日、社会保障番号で、住所は照合時の確認に用いるとのことであった。照合ソフトは、SEER の開発した SEER Link Plus を利用していた。登録患者の生死状況の追跡には、様々な資料源を用いられていた。SEER 事業に参加することで、SEER の支援の元、国の死亡情報と社会保険庁データとの照合を行うことができる。その他、院内がん登録、公的保険加入名簿、有権者登録名簿、担当医師及び患者への直接連絡で、追跡情報を得ていた。ユタ州全体の追跡率は 99% とのことであった。

ユタ州がん登録情報の SEER や NPCR へのデータ提供は、データの機密保護及び被験者保護の連邦ガイドラインとがん研究促進のための要件を含む契約にて行われていた。

一般的な随時の研究への提供は、研究に対する諮問委員会で審議して決定されるとのことであった。

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 の一般利用データ提供 (Public Use Research Database Data Use) のウェブサイトに掲載されている利用規約と『よくある質問と回答』の内容を調査した。一般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、2018 年 11 月に開催された国際がん登録協議会年次総会にて、NAACCR の担当者との打ち合わせを調整した。

分担研究者ら（伊藤、大木、松田）の聞き取り調査によると、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

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った。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と、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」について合理的説明をできないため、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。

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 Data を用いた記述疫学研究について

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。

研究分担者の林は、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計画した。今年度は、SEER Data の利用申請を行い、SEER が提供している統計解析ソフト SEER*Stat の Client-Server Mode を用いて、米国分の解析を行った。また、日系アメリカ人のがん罹患率や死亡率の情報は、SEER Data には含まれていなかったが、別の Asian Pacific Islander (API) Specialized Population Dataset に含まれており、別途の仕様申請で提供を受けられた。2000 年以降、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。

研究分担者の堀と片野田は、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。今年度は、SEER Data の利用申請を行い、SEER*Stat の Client-Server Mode を用いて、米国分の解析を行った。罹患数は 1 歳刻みの年齢区分が利用可能であるが、人口データは 1 歳年齢階級に対応していないことが分かった。また、地理情報は County まで利用可能で、さらに County に対応した American Community Survey データの利用が可能なので、地域別の社会・経済指標とがん罹患・死亡との関連分析が可能であることが分かった。組織型については、ICD-O-3

分類に基づく組織型／性状情報が含まれ、診断年その他の患者情報（性別、年齢階級、人種等）を研究者自らの任意の組み合わせる集計が可能であった。SEER*Stat から取得したデータを用いて、卵巣がんの年齢階級・人種・組織型別罹患率を計算した。White, Black, American Indian/Alaska Native, API 別では、API とその他の人種の卵巣がん罹患率に、年齢別、組織型別に違いが見られたため、日米比較を行う意義が確認された。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。

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

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

記述疫学研究班の研究結果は、日本版 SEER によってもたらされる我が国のがんによる負荷に関する記述統計の充実と、それによる公平、効率的ながん対策企画立案への貢献の可能性を示している。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。

上記提案に至った背景と今後の課題について記述する。

SEER Data の素晴らしさは、住民ベースのがん登録データであること、診療情報を含むこと、高品質であることである。これらの 3 点を長期にわたり満たすのは容易ではない。一般的に、より正確に数（がん罹患者）

を把握しようとするればその他の項目の正確な取得が、個々の患者に関する診療情報をより多く把握しようとするれば全数登録が難しくなる。

SEER が住民ベースに診療情報を含む高品質のデータを整備し続ける仕組みについては、SEER 事業が、資金提供、教育及び登録システムの環境整備支援を条件として、データの提供元である州がん登録との共同研究事業であることが要であることが分かった。各州のがん登録の存在根拠や基本的な財源は各州の法律と財源によるが、それを越える活動については SEER 事業その他の研究事業との契約によって、自らの可能な範囲で行っている。また、SEER 事業を継続するために、NCI の SEER を管理する部門は、常に事業改善計画を考え、新しい事業を付加し、事業価値を維持、高める工夫をしている。

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米国とは異なるなど

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。また、日本の一つの公的研究事業が SEER 事業のように 40 年も持続することはまれであり、各都道府県がん登録を必要十分に支援できる仕組みでもなかった。現在は、日本の住民ベースのがん登録は、研究事業ではなく、がん登録推進法に基づく国の事業であり、各都道府県には法定受託事務として必要な経費の 2 分の 1 が国から補助されている。

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。

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、すなわち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、日本版 SEER として、住民ベースではなく、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、この方法で収集されるデータはがん診療を専門とする病院情報というバイアスがかかり、それを調整する統計学的手法が見当たらないことから、記述疫学研究班が行い考察で述べた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。また、分母となる集団を想定できないため罹患率を計測できないこと、別途行われている院内がん登録全国集計事業との違いは病院間の重複症例の有無だけであることから、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。現状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、協働してデータ連携と質の向上を目指す必要がある。

SEER Data 提供の仕組みの素晴らしさは、原則として国内外問わず、簡単な手続きで誰でも利用できることである。日本では、2015年の個人情報保護法改正によって匿名加工情報が規定され、匿名加工情報の定義を満たす情報を、事業者は本人の同意なく第三者に提供できるようになった。がん登録推進法における全国がん登録情報の匿名化された情報の定義は、改正個人情報保護法の匿名加工情報とは独立しており、本研究班の発足時に厳密な定義が未定であったが、個人情報保護法の考え方は尊重されるであろうことから、「匿名化」が易利用性において重要になってくる。

匿名化の考え方について、米国では、非特定化データと匿名化データの区別があり、非特定化データを用いる研究は「人を対象とする研究」と見なさないとする法規則が前提となる。日本の場合、個人情報保護法における「匿名化」の定義は、米国の匿名化データの定義（個人情報除かれて、原情報に連結不可能）に近い。一方で、HIPAA の非特定化データに関するセーフハーバー規則は、日本の個人情報保護法ガイドライン（匿名加工情報編）による匿名加工情報の適正な加工に近い。日本の個人情報保護法ガイドラインでは、情報を相互に連結する符号もがあれば必ず削除するか置き換えをしなければならないこと、個人情報データベース等の性質に応じて復元可能性について個別具体的に判断する必要があることが HIPAA のセーフハーバー規則とは異なるが、氏名と個人識別符号の削除又は置き換え、生年

月日等の年月までの加工、地理情報の最小基準までの加工、特異な記述の削除や置き換えについては、ほぼ同様の基準である。

SEER Data は、氏名除去、個人識別符号除去、生年月日の年齢表記と生年処理、診断日の年月処理、生存期間の生存期間表記、地理情報の County までの処理の匿名加工済みの非特定化データであり、原情報への復元可能性は否定されない。その点は、利用者との利用規約において、提供された情報から万が一個人を識別できてしまった場合は事務局に報告すること、その情報は削除すること、少数集計値では公表しないこと、を規定することで配慮している。

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、当該がんに罹患した者の識別ができないように加工すること」と定義されている。個人情報保護法ガイドライン（匿名加工情報編）では、個人情報保護法における「特定の個人を識別することができる」とは、情報単体又は複数の情報を組み合わせて保存されているものから社会通念上そのように判断できるものをいい、一般人の判断力又は理解力をもって生存する具体的な人物と情報の間に同一性を認めるに至ることができるかどうかによるものである。匿名加工情報に求められる「特定の個人を識別することができない」という要件は、あらゆる手法によって特定することができないよう技術的側面から全ての可能性を排除することまでを求めるものではなく、少なくとも、一般人及び一般的な事業者の能力、手法等を基準として当該情報を個人情報取扱事業者又は匿名加工情報取扱事業者が通常の方法により特定できないような状態にすることを求めるものである、

とされていることから、がん登録推進法上の匿名化についても、ほぼ同様に考えてよいであろう。

以上より、本研究班は、全国がん登録情報の匿名化の加工とは、全国がん登録の原情報から、氏名除去、個人識別符号除去、情報を連結する符号の除去、日付項目の置き換え（生年月日の年齢表記、一定年齢以上の高齢年齢の置き換え、死亡日等の生存期間表記）、診断時住所の市区町村より詳細部分の削除を最低限の加工とし、加工後の情報に特定の個人を識別可能な特異な記述が含まれるか否かは、全国がん登録情報等の提供に係る審議会等の意見を聴くことを提案する。

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、当該がんに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と見込まれる、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。2019年1月からの全

国がん登録情報の提供の開始にあたり、将来の日本版SEERのあるべき姿を見据えて、現時点で可能な匿名化の妥当性の判断基準や利用規約の整備に専門的見地から助言を行った。数年以内に予定されているがん登録推進法の改正手続きを前に、同法の理念の達成のために必要な改正事項を具体的に整理できたことは、今後の我が国の科学的根拠に基づくがん対策の一層の充実に資すると考えられた。

(資料1) 公開情報ユタ州がん登録 SEER

(資料2) 訪問ユタ州がん登録_SEER

(資料3) 参考資料(H29第2回班会議から)

(資料4) 参考資料(H29第3回班会議から)

(資料5) 参考資料(H30第1回班会議から)

(資料6) 参考資料(H30第2回班会議から)

(資料7) 記述疫学研究班結果報告

(資料8) 参考資料(H30第3回班会議から)

(資料9) 班会議議事録集

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

(2017年度)

1) Katanoda K, Shibata A, Matsuda T, Hori M, Nakata K, Narita Y, Ogawa C, Munakata W, Kawai A, Nishimoto H. Childhood, adolescent and young adult cancer incidence in Japan in 2009–2011. *Jpn J Clin Oncol*. 2017; 47: 762–771

2) Masaoka H, Ito H, Yokomizo A, Eto M, Matsuo K. Potential overtreatment among men aged 80 years and older with

localized prostate cancer in Japan. *Cancer Sci.* 2017;108(8):1673-1680.

3) 金村政輝：がん転移学(上)Ⅱがんの疫学 2. 臓器別のがん転移に関する疫学情報 (3)乳がん. 日本臨床 75 巻増刊 841-45. 2017 (2018 年度)

1) Inoue, S., H. Ito, S. Hosono, M. Hori, T. Matsuda, M. Mizuno, K. Kato and K. Matsuo. Net Survival of Elderly Patients with Gynecological Cancer Aged Over 75 Years in 2006-2008. *Asian Pac J Cancer Prev.* 2019; 20(2): 437-442.

2) Yoshimura, A., H. Ito, Y. Nishino, M. Hattori, T. Matsuda, I. Miyashiro, T. Nakayama, H. Iwata, K. Matsuo, H. Tanaka and Y. Ito. Recent Improvement in the Long-term Survival of Breast Cancer Patients by Age and Stage in Japan. *J Epidemiol.* 2018; 28(10): 420-427.

3) Matsuda T, Inoue M. Moving towards tailored, region-specific cancer-control measures in China. *Lancet Glob Health.* 2019;7(2):e175-e176.

4) Nakatochi M, Lin Y, Ito H, Hara K, Kinoshita F, Kobayashi Y, Ishii H, Ozaka M, Sasaki T, Sasahira N, Morimoto M, Kobayashi S, Ueno M, Ohkawa S, Egawa N, Kuruma S, Mori M, Nakao H, Wang C, Nishiyama T, Kawaguchi T, Takahashi M, Matsuda F, Kikuchi S, Matsuo K. Prediction model for pancreatic cancer risk in the general Japanese population. *PLoS One.* 2018 Sep 7;13(9):e0203386.

2. 学会発表

(2017 年度)

1) 柴田亜希子. 全国がん登録からみた臓器がん登録. 第 25 回日本乳癌学会学術集会 (シンポジウム): 2017.7: 福岡県.

2) 柴田亜希子, 松田智大. 全国がん登録の匿名化情報の研究利用に関する一般意識調査. 第 76 回日本癌学会学術集会 (ポスター, 英語): 2017.10: 神奈川県.

3) 柴田亜希子. 全国がん登録の現状. 第 55 回日本癌治療学会学術集会 (シンポジウム): 2017.10: 神奈川県.

4) 柴田亜希子. がん登録推進法の 5 つの基本理念: 基本理念の実現に向けて. 第 30 回日本放射線腫瘍学会学術大会 (シンポジウム, 英語): 2017.11: 大阪府.

5) Nakagawa H, Matsuda T, Ito H. Prognostic impact of tumor location in colon cancer: the Monitoring of Cancer Incidence in Japan (MCIJ) Project. 39th IAACR Scientific Conference. (Poster Presentation): Oct, 2017: Utrecht, the Netherlands.

6) 大木いずみ, 西野善一, 松田智大. 地域がん登録データを用いたがん診療実態の把握. 第 76 回日本公衆衛生学会総会: 2017.10: 鹿児島県.

7) Oki I, Nishino Y, Saruki N. Profile of the Japanese Association of Cancer Registries. Asian Cancer Registry Forum 2018, March, 2018. Bangkok, Thailand.

8) 金村政輝. 全国のがん診療連携拠点病院における院内がん登録の現状分析—実施体制と実績との関連. 第 76 回日本公衆衛生学会総会: 2017.10: 鹿児島県.

9) 中柄 昌弘, 林 櫻松, 伊藤 秀美, 木下文恵, 小林 由美子, 石井 浩, 笹平 直樹,

上野 誠、江川 直人、来間 佐和子、中尾 春壽、松田 文彦、菊地 正悟、松尾 恵太郎。一般日本人集団における膀胱がん発症予測モデルの構築。日本人類遺伝学会第 62 回大会。2017. 兵庫県。

(2018 年度)

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。

2) Ito H, Masaoka H, Koyanagi Y, Matsuo K. Potential overtreatment among men aged 80 years and older with localized prostate cancer in Japan. 40th IACR Scientific Conference (Poster Presentation). Nov, 2018. Arequipa, Peru.

3) 碓井喜明、小柳友理子、松田智大、片野田耕太、松尾恵太郎、伊藤秀美。日米における多発性骨髄腫の死亡率の経年変化と新規薬剤導入の影響。第 29 回日本疫学会学術総会(口演): 2019.2: 東京都。

4) 山口通代、小柳友理子、松尾恵太郎、伊藤秀美。愛知県がん登録資料に基づくがん患者 5 年生存率の医療圏間格差-Funnle plot 法を用いた評価。第 29 回日本疫学会学術総会(ポスター発表): 2019.1. 東京都。

5) Inoue M. Epidemiology of esophagogastric junction (EGJ) cancer. 第 91 回日本胃癌学会総会(シンポジウム 1: EGJ tumor update: Topics from epidemiology and diagnosis): 2019.2. 静岡県

6) Izumi Oki, Yoshikazu Nishino, Tomohiro Matsuda. Epidemiologic features of cancer cases diagnosed

and/or treated in nationally designated cancer care hospitals. The 40th Annual Meeting of the IACR (Poster Presentation). Nov, 2018. Arequipa, Peru.

7) 金村政輝, 伊藤秀美, 大木いずみ, 井上真奈美, 柴田亜希子。全国がん登録を基盤とした記述疫学研究「日本版 SEER」を実現するために。第 77 回日本癌学会学術総会(ポスター発表): 2018.9. 大阪府。

8) 金村政輝, 只野尚子。がん登録情報を活用したがん対策の推進(1)市町村別がん統計情報の提供の実態と課題。第 77 回日本公衆衛生学会総会(ポスター発表): 2018.10. 福島県

9) 只野尚子, 金村政輝。がん登録情報を活用したがん対策の推進(2)第 3 期がん対策推進計画の記載状況調査。第 77 回日本公衆衛生学会総会(ポスター発表): 2018.10. 福島県

10) Hori M, Saito E, Matsuda T, Shibata A, Katanoda K. Distribution and incidence of ovarian cancer by histologic subtypes in Japan. The 40th Annual Conference of the International Association of Cancer Registries. Nov, 2018. Arequipa, Peru.

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