

論壇

諸外国でのがん登録データの地理情報の
利用事例とわが国の全国がん登録の諸問題

| | | | |
|---------------------------------|--------------------------------|---|------------------------------------|
| カタノダ ヨウタ 片野田耕太* | イトウ ヒデミ 伊藤 秀美 ^{2*} | イトウ ユリ ^{3*} 伊藤 ゆり ^{3*} | カタヤマ カヨコ 片山佳代子 ^{4*,5*} |
| ニシノ ヨシカズ 西野 善一 ^{6*} | ツツイ アンナ 筒井 杏奈 ^{7*} | トガワ カヨ 十川 佳代* | タナカ ヒロカズ 田中 宏和* |
| オオノ ユウコ 大野ゆう子 ^{7*} | ナカヤ トモキ 中谷 友樹 ^{8*} | | |

諸外国では、がん登録を始めとする公的統計データの地理情報を用いた研究ががん対策および公衆衛生施策に活用されている。日本でも2016年に全国がん登録が開始され、がんの罹患情報のデータ活用が制度的に可能となった。悉皆調査である全国がん登録は、市区町村、町丁字など小地域単位での活用によりその有用性が高まる。一方、小地域単位のデータ活用では個人情報保護とのバランスをとる必要がある。小地域単位の全国がん登録データの利用可否は、国、各都道府県の審議会等で個別に判断されており、利用に制限がかけられることも多い。本稿では、がん登録データの地理情報の研究利用とデータ提供体制について、米国、カナダ、および英国の事例を紹介し、個人情報保護の下でデータが有効に活用されるための方策を検討する。諸外国では、データ提供機関ががん登録データおよび他のデータとのリンケージデータを利用目的に沿って提供する体制が整備され、医療アクセスとアウトカムとの関連が小地域レベルで検討されている。日本では同様の利活用が十分に実施されておらず、利用申請のハードルが高い。全国がん登録の目的である調査研究の推進とがん対策の一層の充実のために、他のデータとのリンケージ、オンサイト利用など、全国がん登録を有効かつ安全に活用できる体制を構築していく必要がある。

Key words : がん登録, 記述疫学, 保健統計, 医療アクセス, 地理空間疫学

日本公衆衛生雑誌 2023; 70(3): 163-170. doi:10.11236/jph.22-093

I はじめに

諸外国では、がん登録を始めとする公的統計データの地理情報を用いた研究が広く行われ、公衆衛生施策に活用されている¹⁻³⁾。わが国でも2013年に成立した全国がん登録等の推進に関する法律（以下、がん登録推進法）に基づき、2016年に全国がん登録が開始された。これにより、人口動態調査に基づく

死亡情報だけでなく、がんの罹患情報についても、全国の地理情報に基づくデータ活用が制度的に可能となった。がん登録の地理情報を活用することにより、がん罹患に地理的な集積や格差がないか、国民・患者ががん医療に均しくアクセスできているかなどの評価が可能となる。諸外国では、このような公的統計データの地理的分析は、学術研究によるデータ利用の仕組みによって実現している。

2022年8月現在、全国がん登録に基づく2016～2019年診断例のデータが政府統計の総合窓口 e-stat で公開され⁴⁾、申請に基づくデータ提供が行われている⁵⁾。しかしながら、患者住所や医療施設住所など、地理情報を含むがん登録情報は、まだ広く利用されているとは言えない。その最も大きな要因として、がん登録データに付随する地理情報の利用が、個人情報保護の観点から制限されていることが挙げられる。諸外国では、がん登録の地理情報が個人情報保護とのバランスをとりながら広く活用され、が

* 国立がん研究センターがん対策研究所予防検診政策研究部

^{2*} 愛知県がんセンター研究所がん情報・対策研究分野

^{3*} 大阪医科薬科大学医学研究支援センター医療統計室

^{4*} 群馬大学情報学部

^{5*} 神奈川県立がんセンター臨床研究所

^{6*} 金沢医科大学医学部公衆衛生学

^{7*} 大阪大学大学院医学系研究科保健学専攻

^{8*} 東北大学大学院環境科学研究科先端環境創成学専攻
責任著者連絡先：〒104-0045 中央区築地 5-1-1

国立がん研究センターがん対策研究所予防検診政策研究部 片野田耕太

ん対策に資する研究が行われている。わが国でも、地理情報を含むがん登録データの利用を推進することで、がん登録推進法の趣旨であるがん対策の一層の充実につなげられる可能性がある。そこで本稿では、がん登録データの地理情報の研究利用とデータ提供体制について、米国、カナダ、および英国の事例を紹介し、わが国の全国がん登録データの地理情報の利用における論点を整理し、個人情報保護の下でデータが有効に活用されるための方策を検討することを目的とする。なお、本稿の「がん登録」は別途記載がない限り住民ベースのがん登録 (population-based cancer registry) を指す。

II 諸外国の事例

米国では州単位で住民ベースのがん登録が実施されており、そのうち一部が National Cancer Institute (NCI) の SEER (Surveillance, Epidemiology, and End Results) プログラムにより運営されている⁶⁾。SEER のデータは、解析用アプリケーション SEER*Stat の形で個票レベルのデータがオンラインでの利用申請と誓約書の送付により利用できる^{7,8)}。誓約書には、目的外に利用しないこと、他のデータベースと個人レベルのリンケージをしないこと、データの適切な保管措置などの項目が含まれている。SEER*Stat データの地理情報はカウンティレベルまでだが、他のさまざまなデータベースと連結することでさらに細かい地理情報や社会経済指標を用いた分析が可能となっている (カウンティ (郡) は米国の州の下位行政区分で、全米で約3,000のカウンティがある)。

その一つ、Census Tract-level SES and Rurality Database は⁹⁾社会経済指標 (SES) と都市化の程度の指標データである。SES は国勢調査統計区ごとに世帯収入、家賃、貧困度、学歴、職業、失業などから算出された五分位データ、都市化の程度は同じく国勢調査統計区ごとに算出された合成指標の2または4カテゴリのデータで、いずれも SEER*Stat のデータに付与される形で使用する。このデータにより、カウンティより細かい地理単位の社会経済指標と、がん罹患などの関連の分析が可能となる。

もう一つの SEER-Medicare Linked Data Resource は¹⁰⁾、SEER のがん登録データと公的医療保険 Medicare/Medicaid データを個人単位で連結したものである。この連結データは米国内のみ利用可能で、研究計画の概要を記した申請書、誓約書、および所属施設の倫理審査委員会の承認書を提出して申請する^{8,11)}。患者および医療施設の詳細な地理情報として、患者住所の国勢調査統計区および郵便番

号、医師および医療施設住所の郵便番号がオプションとしてあり、各州のがん登録から別途承認を得ることで利用可能となる¹²⁾。

SEER-Medicare Linked Data Resource を用いた研究の例として、Donohoe らは米国アパラチア地域 (ペンシルバニア、オハイオ、ケンタッキー、ノースカロライナ州) において、患者住所と施設住所 (いずれも国勢調査統計区) からプライマリーケア施設へのアクセスを評価し、進行乳がん診断 (ステージ III/IV) との関連を調べ、プライマリーケア施設へのアクセスがよいほど進行乳がんの診断が少ないことを示した¹³⁾。同様の研究は米国全体や州単位のがん登録を用いた研究で広く実施されており、医療アクセスのよさと生存率の高さ、非侵襲的治療の利用などとの関連が報告されている^{14~17)}。

カナダでは、連邦政府の統計局ががん登録を含む公的統計、納税、医療情報などを個票レベルで連結したデータの提供を行っている¹⁸⁾。申請者が研究計画を申請し、承認を受け、データ利用に関する誓約書を提出する流れは米国と同様である。地理情報は郵便番号が利用可能である¹⁹⁾。州レベルでも同様の枠組みが整備されており、オンタリオ州では Institute for Clinical Evaluative Sciences (ICES) ががん登録を含む保健医療データの分析と管理を担っている。申請者が研究内容に関する申請書と誓約書を提出し、ICES がそれを審査してデータセットを作成、提供するという流れである²⁰⁾。

オンタリオ州でこの枠組みで行われた研究の例として、Yee らは州がん登録で同定された非治療食道がん患者住所 (郵便番号レベル) と近隣がん治療施設までの距離、腫瘍内科医のコンサルテーションと化学療法の施行、および生存率との関連を調べ、医療施設まで遠い患者ほどコンサルテーションと化学療法を受けない傾向、生存率が低い傾向があることを示した²¹⁾。プリティッシュ・コロンビア州でも、Walker らはがん登録から得た頭頸部がん患者データについて患者住所 (国勢調査統計区) から算出した地理的剥奪指標 (area deprivation index: ADI) と、最寄りの治療施設までの移動時間 (郵便番号レベル) との関連を調べ、ADI が高い (困窮度の高い) 地域ほど治療施設への移動時間が長い傾向があることを報告している²²⁾。

英国では、National Health Service (NHS) に新しく組織された NHS Digital が、保健医療分野の公的統計データの管理プログラム National Disease Registration Service (NDRS) を運営している²³⁾。データ提供の枠組みは米国やカナダと同様、利用者からの申請書と同意書の提出に基づいて行われてい

る²⁴⁾。Murageらはイングランドがん登録のがん患者住所から General Practitioner (GP) への移動時間を算出し、GP への移動時間が長いほど検診ではなく緊急搬送などが診断契機となっていることを示している²⁵⁾。

Ⅲ わが国の状況と論点の整理

わが国でも、全国がん登録の個票データが申請に基づいて提供されている⁵⁾。しかしながら、患者住所や医療施設住所などの地理情報は有効に利用されているとは言えない。実際、がん登録データの利用申請の過程で、地理情報の利用が個人情報保護の観点から制限されている。全国がん登録の利用申請の審査の際には、以下の論点が主に問題となる。

1. 地理的単位

まず、どの程度細かい地理的単位のデータまで提供するか(粒度)という問題がある。全国がん登録データにおける患者住所は、J-LIS 全国町字ファイル11桁コード(字・丁目)で登録されている。したがって、全国がん登録データを利用する際の地理的単位としては、J-LIS 11桁コードが最も粒度が細かく(町丁字に相当)、小学校・中学校区、市区町村、二次医療圏などの順で粒度が粗くなる。どの程度の粒度を必要とするかは研究目的によって異なる。たとえば上記で紹介した諸外国の医療アクセスの研究では、患者の住所そのものまたは郵便番号や国勢調査統計区レベルなどが使われている。移動時間の算出は市区町村の人口重心などでも可能ではあるが²⁶⁾、患者住所の地理的粒度が細かいほど移動時間の算出精度は高くなる。上記 Walker らのように ADI を用いた研究においても²²⁾、同一の市区町村内で地理的条件により居住地別の社会階層構成が異なるため^{27,28)}、より粒度の細かい地理情報を用いることでより正確な居住者の社会階層の分類が可能となる。国内でも全国がん登録実施前は、ADI に基づく社会格差研究で、町丁字など小地域単位のデータが用いられてきた^{29~31)}。アスベストと中皮腫など、環境要因によるがんの地理的集積性を調べる研究でも、より細かい地理的単位のデータが必要になる。実際、大阪府地域がん登録のデータ(全国がん登録実施前)を用いて印刷工場と胆管がん罹患との関連を調べた研究では町字レベルのデータが用いられている³²⁾。地域での保健医療活動は、市区町村全体よりむしろ地域コミュニティで実施されるため、保健医療計画の立案や評価には学校区など社会的まとまりのある単位が適している。一方、希少疾患の受療動向と専門病院の配置の検討など、広い医療圏単位の地理情報が適している場合もある。

地理的単位に関わるもう一つの論点として、汎用性と連続性の問題がある。市区町村、町丁字とそれに基づく J-LIS コードは合併や新設(政令指定都市の区政導入など)による変更がありえる。そのため、町丁字の境界データを正確かつ経時的に保存しておくことが必要である。代替的に、より汎用性が高く時系列比較が容易な地理的単位として、緯度経度に基づくメッシュ(たとえばおおよそ 1 km × 1 km の格子状データ)データがある。現在 J-LIS コードで管理されている全国がん登録情報の町丁字住所情報を、その代表点の緯度経度データと合わせて保有することで、時系列的な分析や他の地理的単位への変換(たとえば 1 km メッシュから 10 km メッシュへ)などの汎用性が高まる。

2. 人口規模と疾患の希少性

次に、人口規模と疾患の希少性の問題がある。小地域のがん統計情報から個人が特定される可能性は、人口規模に大きく左右される。たとえば同じ東京都の市区町村であっても、人口は区部で世田谷区の 948,147 人から千代田区の 66,758 人、市部で八王子市の 579,605 人から羽村市の 54,341 人、町村部で瑞穂町の 31,774 人から青ヶ島村の 169 人までの範囲がある(2020年国勢調査)。同じ 1 例の症例であっても、世田谷区と青ヶ島村では個人の特定可能性が全く異なるため、市区町村レベルまでは可というように一律の基準を設けるのは適切ではない。

個人の特定可能性は当該疾患の希少性にも大きく左右される。胃がんや肺がんなど主要ながん種の場合と希少がんの場合とでは少数例の持つ意味が異なる。Ⅲの論点 1. で述べた中皮腫のように(2019年罹患率人口10万対1.5)、がん種の希少性が実数把握の必要性に関わる場合があるため、目的や必要性を勘案した判断が必要となる。

3. データ利用と公表物の区別

公的調査のデータ利用申請におけるデータの粒度は、申請者に提供されるデータと公表物とで明確に区別して判断される必要がある。提供されるデータは個票レベルであっても、それを公表する際には何らかの属性単位で集計するという形が可能である。個票データや情報粒度の細かいデータは、より細かい分析が可能となる一方、個人の特定可能性が高くなる。公表データの場合、表で数値を掲載するか、地図にするか、散布図にするか、頻度分布にするかなど、匿名性の制御の幅が広い。分析のために細かい地理的単位のデータが必要であっても、論文などで公表する際にはリスク比などの統計値があれば足りる場合もある。集計した結果少数例になった場合は「5 例未満」のように丸めた形で公表することも

可能である(Ⅲの論点5. 秘匿措置参照)。公表物に実測値ではなく空間平滑化などの処理をした推計値を用いることによって匿名性を担保する方法も広く用いられている³³⁾。

上記Ⅱで紹介した諸外国の例では、提供データに個人の特定可能性が高いデータが含まれているが、申請者は統計的な結果のみを公表すること、個人を特定しようとしないうこと、他のデータとリンケージしないことなどを、誓約書で同意する形になっている。つまり、研究目的に沿って、個人の特定可能性が高いデータを提供しながらも、統計的結果の公表のみを認めることと、利用規約の遵守を誓約させることによって、データの有用性と個人情報保護とのバランスを図っている。

4. 他の情報とのリンケージ

全国がん登録などの疾病登録データは、他の情報とリンケージすることで有用性が高まる。上記Ⅱで紹介した諸外国の研究事例でも、がん登録から得られたがん患者の情報に、他の公的調査などから得られた生活習慣、社会経済指標、診断医療施設までの距離や移動時間、受療した医療行為などの情報を付与して分析している研究が多く見られた。これらの付加情報は、研究の主たるアウトカムの場合もあれば、主たるアウトカムの影響を適切に評価するための調整変数や層別変数の場合もある。そういう意味で、他の情報の付与は、全国がん登録から得られたデータから科学的に妥当かつ有用な知見を得るために必要なプロセスでもある。

データリンケージに関わるもう一つの問題として、リンケージ作業をどこで行うかという点がある。Ⅱで紹介した諸外国の事例では、リンケージ利用が可能でデータが豊富である一方、リンケージ作業はデータ管理機関が行い、利用者は提供されたデータのみを使う(他のデータとのリンケージを行わない旨を誓約する)形となっていた。ドイツやオーストラリアのようにデータリンケージを行う第三者機関を設けている国もある。

一方、日本の全国がん登録データの提供では、他の公的データとのリンケージによる提供は行われていない。全国がん登録実施前は、都道府県地域がん登録データとDPCなど臨床情報との連結データを用いて小児・若年成人がんの治療内容と生存率との関連や、合併症のあるがん患者の生存率の検討などが実施されていた^{34,35)}。治療だけでなく、肺がんの遺伝子変異検査の実施状況など、全国がん登録と臨床データを組み合わせた研究は患者団体からのニーズが大きい。しかし、このような研究は現在実施が困難になっている。

患者の医療施設までの移動時間の付与も、現行の運用では利用者側が秘匿措置がされた一時データで移動時間情報を付与し、必要データだけ残して再度利用者側に提供される流れになっている。この流れは、利用者側が精度の高い移動時間情報を算出できる一方で、秘匿措置がされているとはいえ患者と医療施設の組み合わせデータが外部に出るという点でセキュリティ上のリスクを伴う。移動時間の算出は簡易的な無料のものを含めて多くのアプリケーションが存在しており、一定の条件の下でデータ提供機関が実施することも選択肢の一つとして考えられる。

セキュリティを担保した環境で利用者自身がリンケージ作業を行う方法として、オンサイト利用という仕組みも国内ですでに開始されている³⁶⁾。2021年7月21日時点で、厚生労働省所管の公的調査では人口動態調査(2016~2018年)、医療施設調査(2010~2019年)、患者調査(2011, 2014, 2017年)、国民健康・栄養調査(2008~2018年)などとなっており、全国がん登録は含まれていない。

5. 秘匿措置

データの提供および公表においては、個人の特定可能性を低くするためにさまざまな秘匿措置が行われる。公表における秘匿措置の代表的なものは、グルーピングと呼ばれる、特定の値をグループ(「20~24歳」、「島しょ部」など)に変更する方法である。同様に、とくに大きい値や小さい値を「〇〇以上」、「〇〇未満」などのようにまとめるトップ・コーディングもよく用いられる。人口動態統計特殊報告の市区町村別死亡データのように、実測値ではなく推計値のみを公表する方法もある。これらの公表における秘匿措置は、Ⅲの論点3. で述べた通り利用の目的や必要性に応じて粒度の選択が可能である。たとえば中皮腫の死亡データのように精度の高い地理情報が重要である場合、時間的情報である死亡年をプールすることで、データの安定性と匿名性の両方の要請に応えることができる。

一方、データの提供における秘匿措置としては、リサンプリングしたデータ、誤差を付与したデータ、ダミーデータを加えたデータなどを提供する方法がある。しかし、リサンプリングや誤差の付与は、元データに改変を加えることになり、単純な集計値は一致したとしても、元データの持つ数値的特徴すべてが再現されるわけではない。たとえば、地理的データは互いに独立しておらず、地理的に連続した地域間ではさまざまな要因で統計値が類似する傾向を持つ。そうした数値的特性を踏まえたより信頼性の高い空間統計解析には人為的な加工が行われていないデータが不可欠である³⁷⁾。ダミーデータを

加えたデータの提供は、Ⅲの論点4.で述べた移動時間の計算のように、利用者が一時的にデータを受領して追加データを付与するような場合に想定される手法だが、全国規模のデータのように症例数が多いとデータサイズが膨大になってしまう可能性がある。

Ⅳ 全国がん登録の地理情報を有効活用するために

これらの現状を踏まえて、個人情報保護の要請に応えつつ全国がん登録データの地理情報を有効活用する方策について考察する。Ⅲの論点1.地理的単位と必要性、および論点2.人口規模と疾患の希少性については、現状の利用申請においても必要性と個人特定可能性を吟味した審査が行われている。ただ、市区町村レベルより粒度が細かい地理情報や、診断・治療が行われた医療施設の情報については、提供データに制限が加えられることが多く、推計値であっても公表が制限される場合がある。Ⅲの論点3.提供データと公表物の区別、および論点5の秘匿措置で述べた通り、利用者への提供データについては粒度の細かいデータを認め、公表データでの制約や秘匿措置でそのバランスをとるという運用が、データの利用価値と個人情報保護を両立しうる方法である。Ⅱで紹介した諸外国のデータ利用枠組みは、基本的には利用者の誓約という性善説に立っている。全国がん登録情報の利用は「がん登録推進法」に基づいており、情報漏洩には罰則規定がある。申請内容を逸脱した利用をした者に一定期間申請を認めないペナルティを課すなど、運用レベルで適切な利用を担保する方法もある。利活用を前提に作られたデータを過度に防衛的に運用することで、利用者の申請意思をそぐ恐れもある。実際、全国がん登録の匿名データの新規利用件数は年平均10件に満たず（うち約半数が国立がん研究センターからの申請）³⁸⁾、広く活用されているとは言えない。公的データの利用申請全般の問題として、形式的、事務的なやり取りが多く、作業コストの問題も大きい。日本の公的データの利用申請では行政的な観点での審査が中心であるが、カナダのICESのように、研究目的に沿ったデータ利用のコンサルテーションをするような枠組みがあれば、より有効なデータ活用が可能になると考えられる。

Ⅲの論点4.他の情報とのリンケージは、全国がん登録データが利用申請の作業コストに見合う価値を持つかに関わる問題でもある。全国がん登録は基本的ながんの罹患率と生存率を計測するための仕組みであり、それだけでは研究仮説に答えられないこ

とが多い。Ⅱで紹介した諸外国の事例でも、治療内容など他のデータとリンケージすることでがん登録情報の価値を高め、より具体的な研究ニーズに応えている。厚生科学審議会においても、NDB（National Database：レセプト情報・特定健診等情報データベース）、介護DB（要介護認定情報・介護レセプト等情報データベース）、DPC（Diagnosis Procedure Combination：診断群分類別包括評価）などとのリンケージデータの整備が議論されており³⁹⁾、これが実現することで全国がん登録データの有用性が向上することが期待される。がんの詳細なステージ情報がある院内がん登録とDPCデータのリンケージデータについても、現在国立がん研究センターの委託費で限定的に利用されているにとどまるが、臨床ニーズ⁴⁰⁾、患者ニーズに応えられるデータとして広く活用されるべきである。

これらのリンケージデータの提供においては、診療の質の地域差などの検討のために、研究目的に沿った地理情報の利用を可能にすることが望ましい。諸外国では臨床データとのリンケージ作業はデータ提供機関が実施しており、NDBなどとのリンケージでもそのような形が想定される。町丁字など小地域単位のADIや患者住所と医療施設との移動時間などについても、データ提供機関での算出ができる体制、あるいはセキュリティが担保された環境でデータ利用者が作業できるオンサイト利用の体制が求められる。

Ⅴ おわりに

全国がん登録の開始から7年、がん登録推進法の成立からは10年を迎えようとしている。その間、データの利活用が進む諸外国との差は開くばかりである。本稿でみた通り、がん登録の地理情報を活用することによりがんアウトカムの格差を可視化し、その健康行動や医療アクセスとの関連を明らかにすることが可能となり、介入候補の同定や医療資源の配置など、効率のかつ公正な公衆衛生施策の立案が可能となる^り。がん登録の法制化は、長年の関係者の悲願で実現した。がん登録推進法の成立には、がん患者団体および自らががん患者でもある国会議員の尽力が大きく寄与した。がん登録推進法は、その目的ががん医療の質の向上と科学的知見に基づくがん対策の実施を掲げている。これらの目的と経緯に今一度立ち戻り、全国がん登録を有効かつ安全に活用できる体制を構築していく必要がある。

本稿は国立がん研究センター研究開発費（2021-A-20）の一部として執筆された。開示すべきCOI状態はない。

諸外国の状況について貴重な情報をいただいた吉見逸郎先生（国立感染研究所・感染症危機管理研究センター）、浅川恵子先生（前カナダ統計局）に感謝申し上げます。

（ 受付 2022. 9.13
採用 2022.11.18
J-STAGE早期公開 2023. 2.10 ）

文 献

- 1) Wang F. Why public health needs GIS: a methodological overview. *Ann GIS* 2020; 26: 1-12.
- 2) Healthy People 2030. <https://health.gov/healthypeople> (2022年7月7日アクセス可能).
- 3) Gomez SL, Shariff-Marco S, DeRouen M, et al. The impact of neighborhood social and built environment factors across the cancer continuum: current research, methodological considerations, and future directions. *Cancer* 2015; 121: 2314-2330.
- 4) e-Stat 政府統計の総合窓口. <https://www.e-stat.go.jp/> (2022年7月7日アクセス可能).
- 5) 全国がん登録 登録情報の提供. 国立がん研究センターがん情報サービス. https://ganjoho.jp/med_pro/cancer_control/can_reg/national/datause/index.html (2022年7月8日アクセス可能).
- 6) National Cancer Institute Surveillance, Epidemiology, and End Results (SEER). National Cancer Institute, U.S. <https://seer.cancer.gov/> (2022年7月7日アクセス可能).
- 7) How to Request Access to SEER Data. SEER, National Cancer Institute, U.S. <https://seer.cancer.gov/data/access.html> (2022年7月7日アクセス可能).
- 8) 堀芽久美, 片野田耕太. 2017年度厚生労働科学研究費補助金疾病・障害対策研究分野 がん対策推進総合研究「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」分担研究報告書. 2018. <https://mhlw-grants.niph.go.jp/project/26468> (2022年7月7日アクセス可能).
- 9) Census Tract-level SES and Rurality Database (2006-2018). SEER, National Cancer Institute, U.S. <https://seer.cancer.gov/seerstat/databases/census-tract/index.html> (2022年7月7日アクセス可能).
- 10) SEER-Medicare: Brief Description of the SEER-Medicare Database. National Cancer Institute, U.S. <https://healthcaresdelivery.cancer.gov/seermedicare/overview/> (2022年7月7日アクセス可能).
- 11) SEER-Medicare: Required Documents & Instructions for Submitting Applications. National Cancer Institute, U.S. <https://healthcaresdelivery.cancer.gov/seermedicare/obtain/requests.html> (2022年7月7日アクセス可能).
- 12) Encrypted or Restricted Variables. SEER, National Cancer Institute, U.S. <https://healthcaresdelivery.cancer.gov/seermedicare/privacy/variables.html> (2022年7月7日アクセス可能).
- 13) Donohoe J, Marshall V, Tan X, et al. Predicting late-stage breast cancer diagnosis and receipt of adjuvant therapy: applying current spatial access to care methods in Appalachia. *Med Care* 2015; 53: 980-988.
- 14) Lin Y, Wimberly MC, Da Rosa P, et al. Geographic access to radiation therapy facilities and disparities of early-stage breast cancer treatment. *Geospat Health* 2018; 13: 622.
- 15) Wan S, Jubelirer S. Geographic access and age-related variation in chemotherapy use in elderly with metastatic breast cancer. *Breast Cancer Res Treat* 2015; 149: 199-209.
- 16) Freeman VL, Naylor KB, Boylan EE, et al. Spatial access to primary care providers and colorectal cancer-specific survival in Cook County, Illinois. *Cancer Med* 2020; 9: 3211-3223.
- 17) Zahnd WE, Josey MJ, Schootman M, et al. Spatial accessibility to colonoscopy and its role in predicting late-stage colorectal cancer. *Health Serv Res* 2021; 56: 73-83.
- 18) Access to microdata. Statistics Canada. <https://www.statcan.gc.ca/en/microdata> (2022年7月8日アクセス可能).
- 19) Canadian Population Health Survey data (CCHS Annual and Focus Content) integrated with mortality, hospitalization, historical postal codes, cancer, tax data and Census. Statistics Canada. <https://www.statcan.gc.ca/en/microdata/data-centres/data/cchs> (2022年7月8日アクセス可能).
- 20) Access to ICES Data. Institute for Clinical Evaluative Sciences, Canada. <https://www.ices.on.ca/DAS/Public-Sector/Access-to-ICES-Data-Process> (2022年7月8日アクセス可能).
- 21) Yee EK, Coburn NG, Zuk V, et al. Geographic impact on access to care and survival for non-curative esophagogastric cancer: a population-based study. *Gastric Cancer* 2021; 24: 790-799.
- 22) Walker BB, Schuurman N, Auluck A, et al. Socioeconomic disparities in head and neck cancer patients' access to cancer treatment centers. *Rural Remote Health* 2017; 17: 4210.
- 23) National Disease Registration Service (NDRS). NHS Digital, UK. <https://digital.nhs.uk/services/national-disease-registration-service> (2022年7月11日アクセス可能).
- 24) Access to NDRS data. National Disease Registration Service, UK. <https://www.ndrs.nhs.uk/dataaccess/> (2022年7月11日アクセス可能).
- 25) Murage P, Bachmann MO, Crawford SM, et al. Geographical access to GPs and modes of cancer diagnosis in England: a cross-sectional study. *Fam Pract* 2019; 36: 284-290.
- 26) Tanaka H, Ishikawa KB, Katanoda K. Geographic access to cancer treatment in Japan: results from a combined dataset of the Patient Survey and the Survey of Medical Institutions in 2011. *J Epidemiol* 2018; 28: 470-475.

- 27) Nakaya T, Ito Y. Chapter 7: Synthesis, The Atlas of Health Inequalities in Japan. New York: Springer. 2019; 297-305.
- 28) 中谷友樹, 矢野桂司. 社会格差を視る小地域のセンサス指標: 地理的剥奪指標とジオデモグラフィクス. 地域開発 2014; 599: 35-40.
- 29) Ito Y, Nakaya T, Nakayama T, et al. Socioeconomic inequalities in cancer survival: a population-based study of adult patients diagnosed in Osaka, Japan, during the period 1993-2004. Acta Oncol 2014; 53: 1423-1433.
- 30) Kaneko N, Nishino Y, Ito Y, et al. Association of socioeconomic status assessed by areal deprivation with cancer incidence and detection by screening in Miyagi, Japan between 2005 and 2010. J Epidemiol 2022.
- 31) 中谷友樹. 地理疫学とがん登録. JACR Monograph 2012; 17: 11-13.
- 32) Ito Y, Nakaya T, Ioka A, et al. Investigation of spatial clustering of biliary tract cancer incidence in Osaka, Japan: neighborhood effect of a printing factory. J Epidemiol 2016; 26: 459-463.
- 33) 中谷友樹. 階層ベイズモデルを利用した小地域疾病地図. 統計 2014; 65: 22-27.
- 34) Morishima T, Matsumoto Y, Koeda N, et al. Impact of comorbidities on survival in gastric, colorectal, and lung cancer patients. J Epidemiol 2019; 29: 110-115.
- 35) Nakata-Yamada K, Inoue M, Ioka A, et al. Comparison of survival of adolescents and young adults with hematologic malignancies in Osaka, Japan. Leuk Lymphoma 2016; 57: 1342-1348.
- 36) 調査票情報の利用 (オンサイト利用). <https://www.e-stat.go.jp/microdata/data-use/on-site> (2022年7月7日アクセス可能).
- 37) Lawson AB, Banerjee S, Haining RP, et al. (eds.). Handbook of Spatial Epidemiology. Boca Raton: CRC Press. 2016.
- 38) 全国がん登録情報提供等審議委員会. 国立がん研究センター「がん情報サービス」. https://ganjoho.jp/med_pro/cancer_control/can_reg/national/datause/shingi.html (2022年8月3日アクセス可能).
- 39) 第18回厚生科学審議会がん登録部会 (資料). 厚生労働省. 2021. https://www.mhlw.go.jp/stf/newpage_22638.html (2022年8月3日アクセス可能).
- 40) 都道府県がん診療連携拠点病院連絡協議会がん登録部会 Quality Indicator 研究. 国立がん研究センターがん対策研究所医療政策部. <https://www.ncc.go.jp/jp/icc/health-serv/project/010/index.html> (2022年8月12日アクセス可能).
-

Geographic information in National Cancer Registry data: Overseas examples and challenges in Japan

Kota KATANODA^{*}, Hidemi ITO^{2*}, Yuri ITO^{3*}, Kayoko KATAYAMA^{4*,5*},
Yoshikazu NISHINO^{6*}, Anna TSUTSUI^{7*}, Kayo TOGAWA^{*}, Hirokazu TANAKA^{*},
Yuko OHNO^{7*} and Tomoki NAKAYA^{8*}

Key words : access to healthcare, cancer registry, descriptive epidemiology, geospatial epidemiology, health statistics

Worldwide, research based on geographic information of official statistical data, including cancer registries, is utilized for cancer control and public health policies. The National Cancer Registry of Japan was launched in 2016, making it possible to use data on cancer incidence systematically. Given the nature of this comprehensive survey, the usefulness of the National Cancer Registry would be further enhanced when it is utilized at small-regional levels, such as in municipalities or even smaller geographical units. It is essential to maintain a balance between privacy protection and data usability. Currently, the national and prefectural councils determine the availability of the data from the National Cancer Registry at the small-regional level on an individual application basis. Under this framework, use of the data is often restricted or declined. This paper showcases three model countries where geographic information obtained from cancer registry data are widely utilized: the United States, Canada, and the United Kingdom. It further discusses measures to ensure that data are effectively used, without compromising data privacy.

In the three countries, data-providing systems have been established to compile the necessary data from the cancer registry and other linked databases, in accordance with the purpose of use. The relationships between healthcare access and various outcomes are elaborately examined at the small-regional level. In Japan, similar utilization of data has not been fully implemented, and there remain many hurdles to the application of the data use. For the National Cancer Registry to promote research and further enhance cancer control, it is necessary to establish a system that enables effective and safe utilization of the data from the National Cancer Registry, including linkage with other data and on-site use.

^{*} Division of Surveillance and Policy Evaluation, National Cancer Center Institute for Cancer Control

^{2*} Division of Cancer Information and Control, Department of Preventive Medicine, Aichi Cancer Center Research Institute

^{3*} Department of Medical Statistics, Research & Development Center, Osaka Medical and Pharmaceutical University

^{4*} Faculty of Informatics, Gunma University

^{5*} Kanagawa Cancer Center, Research Institute

^{6*} Department of Epidemiology and Public Health, Kanazawa Medical University

^{7*} Graduate School of Medicine Division of Health Sciences, Osaka University

^{8*} Graduate School of Environmental Studies, Tohoku University