

院内がん登録における重複登録割合の推定

ワタナベ タ エ コ ヒガシ タカヒロ サ イ シドウ ヤマシロ カツシゲ
 渡邊 多永子*1 東 尚弘*2 佐井 至道*5 山城 勝重*6
 カイザキ ヤスハル ツ クマ ヒデアキ コ タケ ケン ジ ロウ サル キ ノブヒロ
 海崎 泰治*7 津熊 秀明*8 固武 健二郎*9 猿木 信裕*10
 オカムラ シンイチ シバ タ ア キ コ ニシモト ヒロシ
 岡村 信一*11 柴田 亜希子*3 西本 寛*4

目的 全国のがん診療連携拠点病院から提供される院内がん登録データには、全国のがん罹患の約6割以上についての情報が蓄積されている。一方で匿名化後に収集されており、同一患者が複数の提出施設を受診すると重複して登録されてしまう可能性がある。本解析はデータ内に含まれる重複登録割合の推定を目的とする。

方法 2通りの方法で重複登録割合の推計を行った。1つは、組合管掌健康保険組合8組合から提供された診療報酬請求書（レセプト）データベースを用い、がん診療連携拠点病院を受診した5大がんの患者のうち同一年の間にほかの拠点病院も受診している患者の割合を算出することで重複の割合を推計した。もう1つは、院内がん登録データ2008年症例の中で、多数の性質の詳細が一致する組み合わせ（類似特性症例）を重複と推定して算定した。

結果 重複登録割合は、レセプトデータによる推計では6.6～8.5%、院内がん登録データ内の類似特性症例による推計では8.6%と算定された。

結論 院内がん登録データには8%程度の重複登録が含まれると推定された。

キーワード 院内がん登録、マイクロデータ、重複

I はじめに

2007年に施行されたがん対策基本法（平成18年法98号）においては「がん医療に関する情報の収集提供体制の整備」が定められており、同法に基づくがん対策推進基本計画（平成19年6月閣議決定）では、そのための具体的な仕組みとして「がん登録」が取り上げられている。

「がん登録」とは、一定の対象集団において発生・受療したがん患者の情報を蓄積する仕組みであり、わが国には、実施主体と登録対象の異なる「地域がん登録」「院内がん登録」「臓器がん登録」という3種類のがん登録がある。地

域がん登録は、地方自治体（2012年6月現在45道府県1市）が運営しており、当該地域における全がんの罹患実態を把握して、がん対策やその評価に資することを第1の目標としている。罹患の把握が第1の目的であることから各地域における個人情報保護関連条例などの定める手続きに従って個人情報を使った名寄せを行っており、また実施主体である自治体の保有する死亡情報とのリンクも可能であるため、予後調査に強みがある。一方で、進行度の分類が臨床現場で使われているものと異なることや、平成24年6月現在では医療機関に届出義務がないことから登録精度が低い自治体も存在する。臓器が

*1 東京大学大学院医学系研究科公共健康医学専攻健康医療政策学分野博士課程

*2 独立行政法人国立がん研究センターがん対策情報センターがん政策研究部部長

*3 同がん統計研究部診療実態調査室室長 *4 同部長 *5 岡山商科大学経済学部経済学科教授

*6 独立行政法人国立病院機構北海道がんセンター臨床研究部部長 *7 福井県立病院臨床病理科医長

*8 大阪府立成人病センターがん予防情報センターセンター長 *9 栃木県立がんセンター研究所所長

*10 群馬県立がんセンター院長 *11 独立行政法人医薬品医療機器総合機構主任専門員（調査・臨床医学担当）

ん登録は、学会・研究会が主導して行っている。公的制度としての裏付けがないものの、各がん種に対応した詳細な臨床情報の収集によって、診療ガイドラインや病期分類の改定などに役立てられている¹⁾²⁾。

院内がん登録は、全国で指定されているがん診療連携拠点病院（以下、拠点病院）において、各病院を受診した全がん患者の部位・病理組織などの基礎データが集積されているものである。これらデータは、年ごとに、国立がん研究センターがん対策情報センターに提供される。2007年には287施設から316,089症例、2008年には353施設から420,683症例³⁾が提供され、わが国のがん症例の約6割以上を含んでいると推定されている。これは、施設ごとに自己評価や改善を行う基礎資料として役立つばかりでなく、全体としてわが国のがん診療の現状を示すと考えられ、今後数多くの研究に用いられることが期待される。

この院内がん登録が各病院から収集される際、データからは名前や住所などの個人情報を削除され匿名化された状態となっている。各病院がそれぞれデータを提供するため、同じ患者が同じがんで、複数の拠点病院を受診した場合は重複登録となるが、匿名化後のため整理は行われない。本研究は、院内がん登録データに含まれる重複登録割合の推定を行い、このデータを用いた研究結果を解釈する際の基礎資料に資することを目的とする。

Ⅱ 方 法

重複の推定は、以下の2通りの方法を用いて行った。

（1）診療情報請求書データによる重複率推定

まず、診療報酬請求書（レセプト）データの蓄積・解析・研究機関などへの提供を主たる業務とする日本医療データセンターから研究利用のために供与を受けた、組合管掌健康保険組合のレセプトデータを解析した。利用可能なレセプトデータは、約15万人の加入者（5健康保険

組合）における2005年1月から、約60万人の加入者（3健康保険組合）における2008年1月から、ともに2010年4月診療分までのものであった。これらのうち、当該期間内に5大がん（胃がん、大腸がん、肺がん、肝がん、乳がん）の診断名がついており、2010年4月時点でのがん診療連携拠点病院388施設を受診した集団のうち、同一年に2カ所以上の拠点病院を受診した患者の2カ所目以降の受診を、院内がん登録で重複登録となっている可能性が高いと考え、その数が全体に占める割合を算出した。該当する患者は、病院ごとに算出して足し合わせた。実人数でなく病院ごとに算出して足し合わせたのは、下記の院内がん登録データは各病院初診時に病院ごとに登録されるため、その手順と整合性を持たせるためである。

（2）院内がん登録2008年症例による重複症例推定

次いで、院内がん登録全国データ2008年症例（424,983）³⁾から、診断時居住都道府県が欠損および不明の症例（欠損7,544、不明55）を除外した417,384症例について、以下のすべての条件を満たす組み合わせを重複と推定して数え、その割合を算出した。

- ①性別、生年月、診断時居住都道府県が等しい
- ②部位が等しいまたは矛盾しない
- ③組織が等しいまたは矛盾しない
- ④部位の側性（左右）が等しいまたは矛盾しない
- ⑤データ提出病院が異なる
- ⑥自施設診断同士、自施設初回治療同士ではない

なお、「②部位が等しいまたは矛盾しない」の定義は以下とした。

- a) ICD-O-3の局在コード数字3桁が等しい
「噴門」「胃底部」など、コードされた部位（および亜部位）が一致
- b) ICD-O-3において、詳細不明の分類と、それと矛盾しない分類の組み合わせ

「胃の境界部病巣」および「胃, Not otherwise specified (NOS)」と胃がん全般など

「③組織が等しいまたは矛盾しない」の定義は以下とした。

- a) ICD-O-3の形態コードにおける, 形態の大分類が等しい

「扁平上皮新生物」「基底細胞新生物」など

- b) ICD-O-3の形態コードにおいて, 詳細不明の分類と, それと矛盾しない分類の組み合わせ

「新生物, NOS」とすべての分類など

重複を数える際には, n番目の症例とn+1番目, n+2番目, n+3番目…の症例を, 順次条件を満たすかどうか比較していき, いずれかの相手と合致すれば, このn番目の症例は重複の可能性があるとした(図1)。院内がん登録データには, 性別, 生年月(日は削除済), 診断時居住都道府県, 診断部位コード, 組織コード(国際疾病分類腫瘍学第3版:ICD-O-3

図1 重複症例カウント方法

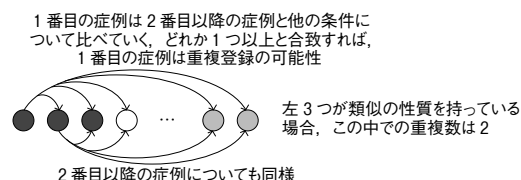


表1 レセプトデータを用いた分析における該当者の特性

	人数 (n=7,126)	%
年齢(歳) (平均±標準偏差)	55.8±10.8	
性別		
男	2 792	39.2
女	4 334	60.8
部位		
胃	1 412	19.8
大腸	1 568	22.0
乳腺	2 656	37.3
肺	1 027	14.4
肝	463	6.5
診療年		
2005	976	13.7
'06	1 055	14.8
'07	1 091	15.3
'08	1 952	27.4
'09	2 052	28.8

に基づく分類コード), 診断施設・治療施設が記載されている。

解析にはStata11 (Stata Corp. College Station, TX, USA) を使用した。

なお, 本研究は, 国立がん研究センターおよび東京大学(審査番号 No.3445, 平成23年6月27日)にて倫理審査を受け, 承認を得ている。

Ⅲ 結 果

レセプトデータを用いた分析において, 該当する患者を病院ごとに算出して足し合わせると, 累計7,126人(レセプト約23万件)であった。その年齢, 性別, 部位, 診療年の内訳を示す(表1)。拠点病院を受診した5大がんの累計患者のうち, 同一年における2カ所目以上の拠点病院受診の割合は, 年次によって異なるものの, 6.6~8.5%であった(表2)。各がん種別で同様の割合を算出すると, 5.6~11.8%であった(表3)。

院内がん登録データ2008年症例を用いた分析において, 多数の性質の詳細が一致する組み合わせ(類似特性症例)を重複と推定したところ, 全がんにおいて条件に合致した症例は35,762例, 8.6%であった。5大がんでは7.8%, 5大がんの各がん種別では5.6~13.0%であった(表3)。

表2 レセプトデータを用いた分析における各年の重複割合

	重複割合 (%)	95%信頼区間
2005年	8.5	6.8-10.4
'06	8.5	6.9-10.4
'07	7.9	6.4- 9.6
'08	6.6	5.5- 7.8
'09	7.8	6.7- 9.0

表3 各がん種の重複割合

	レセプトデータ		院内がん登録データ	
	重複割合(%)	95%信頼区間	重複割合(%)	人数
胃	6.2	5.0- 7.6	7.4	54 966
大腸	6.3	5.2- 7.6	5.6	57 669
乳腺	8.1	7.0- 9.2	11.0	40 927
肺	11.8	9.9-13.9	6.1	48 508
肝	5.6	3.7- 8.1	13.0	20 440

Ⅳ 考 察

レセプトデータを用いた分析では、拠点病院を受診した5大がん（胃・大腸・乳腺・肺・肝）の累計患者のうち、同一年における2カ所目以上の拠点病院受診の割合は6.6～8.5%であった。ただし、これにはがん以外の疾患ではかの拠点病院を受診した場合にも、5大がんの病名がつけられて数えられている可能性がある。そのため、すべてが院内がん登録に重複登録されているわけではない。実際の重複登録割合は今回求めた値より小さくなることが予想される。一方で、同一年での受診に限定したことから、年をまたいで受診については、重複登録割合を過小に推計している可能性がある。また、一部の健康保険組合のデータかつ5大がんのみの分析であり、患者特性の偏りが結果に影響を与えている可能性もある。患者の年齢、居住地やがん種の偏りは、その集団の受療行動に偏りを与えると予想される。ただし、この偏りによって重複受診が増加したか減少したかは、明確にはわからない。

院内がん登録データ2008年症例を用いた分析において、多数の性質の詳細が一致する組み合わせを重複と推定したところ、重複登録は約8.6%と推定された。ただし、これは真の重複登録と、偶然同じような特徴を持つ組み合わせの区別ができていない。そのため、この分析においても、重複登録割合を過大に推計している可能性がある。しかし、今回重複症例と推定するために用いた6つの条件は詳細情報であり取り得る値の組み合わせも非常に多いことから、偶然同じような特徴を持つ確率は高くないと考えられる。

がん種別の重複登録割合の差異は両分析ともにみられ、レセプトを用いた分析では肝がんの5.6%～肺がんの11.8%まで、院内がん登録を用いた分析では大腸がんの5.6%～肝がんの13.0%までの開きがあり、がん種によって受療行動が異なることがうかがわれる。各分析での値のずれ、特に肺がんと肝がんの大きなずれの

原因は不明であるが、レセプトは一部の健康保険組合のものであり上記のとおり患者特性に偏りがあること、また肺がんと肝がんは5大がんの中では患者数が少なく、特にレセプトにおいて標本数が多くなかったことから、推計値自体ある程度の不安定さを含んでいることも影響しているかもしれない。

方法・結果に記載した主な2つの方法により、院内がん登録データには、8%程度の重複登録が含まれると考えられる。今後、院内がん登録データの利用が進むと期待されるが、その際には、データの解釈において、一定の重複登録の程度も考慮に入れる必要がある。また、解析の目的により、この重複登録の割合が許容できる水準よりも高いと考えられるならば、現状の各病院から個人情報を含まない状態でデータが提出されるという形式を改める必要があるかもしれない。

謝辞

本研究を行うにあたり、ご協力下さいましたすべての方々に深謝致します。温かいご指導をいただきました東京大学情報理工研究科の竹村彰通教授、金沢大学経済学部星野伸明准教授、東京大学大学院医学系研究科公共健康医学専攻の小林廉毅教授、豊川智之講師、富尾淳助教と大学院生の皆様、レセプトデータをご提供いただいた日本医療データセンターに深く感謝申し上げます。

なお、本研究は第3次対がん総合戦略事業「院内がん登録の標準化と普及に関する研究」の助成を得て行われた。

文 献

- 1) 東尚弘, 祖父江友孝, 西本寛. 臓器がん登録の現状－臓器がん登録の実態についての調査報告－. 外科治療 2011; 104(2): 169-76.
- 2) 猿木信裕. がん登録の軌跡. 東京: 悠飛社, 2010.
- 3) 独立行政法人国立がん研究センターがん対策情報センターがん統計研究部院内がん登録室, 編. がん診療連携拠点病院 院内がん登録 2008年全国集計報告書. 2011.