

受けた。

ガイドラインは特定の個人を同定しうる情報の提供に慎重であったこと、本研究では、複雑な突合作業を行わないで容易に集計できる結果に関心があったことから、同じ患者の異なる医療機関や、異なる月に請求されたレセプト同士を突合できる情報の提供は受けなかった。そのため、同じ患者の同じ傷病名を複数回計上する場合がある。本研究では、①“疑い”の修飾語コードを持つ、②同一レセプト内にICD-10コードの左3桁（例：C16）が同じレコードの重複分を除外したレコードを計測対象とした。この計測対象の中で、部位別には、ICD-10コードの左3桁単位で集計した。全部位の件数は、何らかの悪性新生物を有する患者の近似として、部位別レコードの合計ではなく、計測対象のレセプト件数とした。それぞれ月単位に集計し、全医科レセプトに占める電子レセプトの割合が安定していた平成22（2010）年8月から平成23（2011）年3月の間の8カ月間の平均を代表値（月平均レセプト件数）とした。

（2） 患者調査に基づく総患者数

平成20年患者調査から作成された、総患者数、性・年齢階級×傷病小分類×都道府県別（患者住所地）の既報の統計表を参照した⁴⁾。最新の平成23年の患者調査は、東日本大震災の影響により、宮城県の一部および福島県の医療施設については調査を実施していないため参考扱いとした⁵⁾。患者調査は統計法に基づく基幹統計調査の1つで、全国の医療施設を利用する患者を対象とし、層化無作為により抽出した医療施設における患者を客体として3年に一度実施される標本調査である。総患者数とは、調査日現在において、継続的に医療を受けている者（調査日には医療施設で受療していない者を含む）の数を、次の算式で推計されたものである。

〔総患者数＝入院患者数＋初診外来患者数＋（再来外来患者数×平均診療間隔×調整係数（6/7）〕

抽出率は、平成23年調査では、病院・入院7.6/10、病院・外来3.9/10、一般診療所6.3/100、

客体数は、病院の入院・外来患者202.5万人、一般診療所の入院・外来患者28.3万人であった。傷病分類別の数値は、主病名についての集計値であり、入院患者においては調査日現在、入院の理由になっている傷病、外来患者においては、主として治療または検査をしている傷病である。

（3） 5年有病数

疫学で用いる“有病数”は、特定の時点において、ある特性または疾病を保有していた者の総数と定義されるので⁶⁾、“患者数”とほぼ同義と考えられる。特定の時点とは、年間、生涯、特定期間、ある一時点でも構わない。計測する人口集団規模が大きい場合、有病状態にある患者を実際に1人ずつ計測することは現実的ではないため、新たに特定の疾患を発症した患者（罹患患者）は、数日から数年の有病期間を経て、治癒もしくは死亡することで有病状態を脱すると仮定して、一般に、有病数は罹患率に有病期間を乗じて推計される。また、有病期間は、一般に、生存率、死亡率で代用される。

本研究では、15府県のデータを元に推計された罹患率と6府県のデータから計測された5年生存率を用いて推計された、日本の2010年から2014年における年平均の5年有病数を利用した¹⁾。この既報では5年有病数を、5年以内に診断され、生存している推計患者数と定義している。がんの罹患数は、人単位ではなく腫瘍単位で数えられている。そのため、全部位の5年有病数には、同一人物の複数の独立した腫瘍が含まれている。5年有病数の場合、5年以内は、生存している限り、治癒している患者も含めて有病状態であるとの前提を置き、5年を越えて生存したがん患者は、治癒していない患者も含めて有病状態ではないとみなす。そのため、5年有病数は、完全治癒が期待できる部位及び進行度のがんについては過大評価になると考えられる。一方、5年を越えても再発の治療を繰り返していたり、寛解状態であったり、完全治癒を宣言するのは難しい部位及び進行度のがんについては、過小評価になると考えられる。この関係は、有病者数×受療割合＝患者数（受療者

図1 提供レセプトレコード総数、レセプトの電子化割合、月別

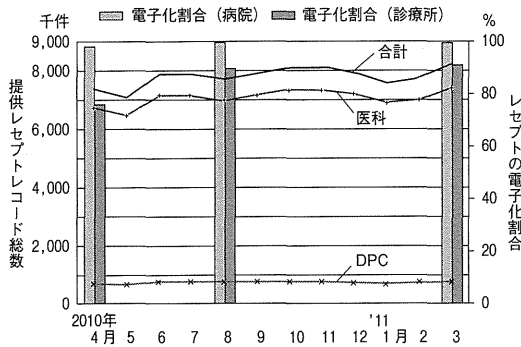


表1 悪性新生物の月平均レセプト件数、総患者数、5年有病数；年齢階級別、性別

	レセプト (A) (千件)	平成23年 総患者数 (千人)	平成20年 総患者数 (千人)	A/B	5年 有病数 (C) (千人)	C/B
全体	2 438	1 526	1 518	1.6	2 310	1.5
0～4歳	2	1	2	1.0	NA	NA
5～14 (0～14)	6 (7)	5 (6)	4 (6)	1.5 1.2	NA 6	NA 0.9
15～24	8	5	5	1.6	NA	NA
25～34	24	15	17	1.4	NA	NA
35～44 (15～44)	78 (111)	62 (82)	62 (79)	1.3 1.4	NA 136	NA 1.7
45～54	167	122	135	1.2	195	1.4
55～64	415	308	313	1.3	474	1.5
65～74	692	450	461	1.5	699	1.5
75～84	789	435	424	1.9	NA	NA
85歳以上 (75歳以上)	256 (1 046)	120 (555)	94 (518)	2.7 2.0	NA 802	NA 1.5
不詳	NA	3	2	NA	NA	NA
男性	1 326	830	837	1.6	1 298	1.6
0～4歳	1	1	1	1.0	NA	NA
5～14 (0～14)	3 (4)	3 (4)	2 (3)	1.5 1.3	NA 3	NA 1.0
15～24	3	2	2	1.5	NA	NA
25～34	6	4	6	1.0	NA	NA
35～44 (15～44)	18 (28)	16 (22)	16 (24)	1.1 1.2	NA 43	NA 1.8
45～54	49	40	46	1.1	73	1.6
55～64	203	149	158	1.3	264	1.7
65～74	416	267	282	1.5	440	1.6
75～84	494	279	274	1.8	NA	NA
85歳以上 (75歳以上)	132 (626)	68 (347)	51 (325)	2.6 1.9	NA 475	NA 1.5
不詳	NA	2	1	NA	NA	NA
女性	1 112	695	680	1.6	1 012	1.5
0～4歳	1	1	1	1.0	NA	NA
5～14 (0～14)	3 (3)	2 (3)	2 (3)	1.5 1.0	NA 3	NA 0.8
15～24	4	3	3	1.3	NA	NA
25～34	18	11	11	1.6	NA	NA
35～44 (15～44)	60 (83)	46 (60)	46 (60)	1.3 1.4	NA 93	NA 1.6
45～54	118	82	89	1.3	122	1.4
55～64	213	158	155	1.4	210	1.4
65～74	276	184	179	1.5	259	1.4
75～84	295	157	150	2.0	NA	NA
85歳以上 (75歳以上)	125 (420)	51 (208)	43 (193)	2.9 2.2	NA 326	NA 1.7
不詳	NA	1	1	NA	NA	NA

注 1) NA: Non-available
2) 悪性新生物: ICD-10 C00-C97

数)で表すことができるだろう。例えば、比較的前後のよい部位では5年有病者数に占める受療中の割合が小さいと考えられるので5年有病数と患者数のかい離が大きく、比較的前後の悪い部位では5年間受療を継続している割合が高いと考えられるので、逆にかい離が小さくなると思われる。

著者らは、総患者数は人単位、5年有病数は腫瘍単位、月平均レセプト件数の全部位は人単位、部位別には腫瘍単位の集計であることやそれぞれの統計値の計測方法の特徴から、以下のような関係を想定し、主に総患者数を基準にして各統計値を比較した。

全部位: 総患者数<<月平均レセプト件数<5年有病数

部位別: 総患者数<<5年有病数<月平均レセプト件数

生存率の高い部位: 総患者数<<5年有病数<月平均レセプト件数

生存率の低い部位: 総患者数<5年有病数<月平均レセプト件数

(4) 倫理的配慮

本研究は、国立がん研究センター倫理審査委員会の承認を得て行われた(研究課題番号2011-091)。厚生労働省から提供を受けるレセプト情報自体は連結可能匿名化状態のデータであるが、連結キーは厚生労働省のみが保持しており、研究者が扱うデータは匿名化状態である。また、少数集計値によって個人が特定されることのないように、集計単位を配慮した。

Ⅲ 結 果

図1に提供を受けた医科領域レセプトレコード総数、医科領域レセプトに占める電子レセプトの割合を月別に示す。提供を受けた全電子医科レセプトレコードは約9300万件で、うち、DPCレセプトレコードの割合は9.4%であった。電子化割合が93%以上になった平成22年8月から平成23年3

月請求分の医科領域レセプトレコードの月件数について、最小値は1月請求分760万件で、最大値は3月請求分820万件であり、最大約60万件の月別変動がみられた。

修飾語コード“疑い”を持つレコードは約40%を占めた。修飾語コード“疑い”を持つレコード、同一部位の重複レコード、上皮内がん及び悪性以外の脳・中枢神経系腫瘍を除外し、悪性新生物（ICD-10：C00-C97）に限定すると約4400万件であった。この中で、同一レセプトで2つ以上のがん傷病名レコードを持つ対象は約34%認められた。1つ以上のがん傷病名レコードを持つレセプト件数でまとめると約2900万件であった。

表1に悪性新生物の月平均レセプト件数、総患者数、5年有病数を、総患者数を基準に比較した。全部位、全年齢では、月平均レセプト件数は

約240万件、総患者数は約150万人、5年有病数は約230万人であった。総患者数を1とした場合、全年齢では5年有病数は1.5倍、月平均レセプト件数は1.6倍であるのに対して、75歳以上では5年有病数は1.5倍、月平均レセプト件数は2倍であり、高年齢層ほど総患者数と月平均レセプト件数の差が大きい傾向がみられた。この傾向は、男女別でも同様であった。

表2に、主な部位について、月平均レセプト件数、総患者数、5年有病数を、性別に示した。胃、大腸、肝臓、肺、前立腺、乳房等の罹患数の多い部位について、総患者数と比較して、月平均レセプト件数は約2倍から2.4倍、5年有病数は約1.5倍から2倍であった。中程度の罹患数を持つ部位については、膵と膀胱では総患者数と5年有病数の差が小さいのに対して、月

表2 悪性新生物の月平均レセプト件数、総患者数、5年有病数；部位別、性別

	ICD-10	レセプト (A) (千件)	平成23年 総患者数 (千人)	平成20年 総患者数 (B) (千人)	A/B	5年 有病数 (C) (千人)	C/B
全体							
胃	C16	476	186	213	2.2	421.3	2.0
大腸	C18-C20	573	233	235	2.4	418.2	1.8
肝臓及び肝内胆管	C22	153	52	66	2.3	94.5	1.4
肺	C33-C34	264	138	131	2.0	194.0	1.5
膵	C25	61	34	30	2.0	21.3	0.7
膀胱	C67	127	58	64	2.0	67.3	1.1
甲状腺	C73	77	29	29	2.7	51.6	1.8
中皮及び軟部組織	C45-C49	14	6	5	2.8	NA	NA
中枢神経系 ¹⁾	C70-C72	11	7	6	1.8	11.0	1.8
白血病	C91-C95	60	30	29	2.1	19.7	0.7
男性							
胃	C16	318	123	141	2.3	293.8	2.1
大腸	C18-C20	327	132	134	2.4	243.8	1.8
肝臓及び肝内胆管	C22	100	36	41	2.4	63.9	1.6
肺	C33-C34	166	88	83	2.0	120.9	1.5
前立腺	C61	385	181	183	2.1	262.1	1.4
膵	C25	33	17	17	1.9	11.5	0.7
膀胱	C67	100	42	47	2.1	53.0	1.1
甲状腺	C73	15	7	6	2.5	11.4	1.9
中皮及び軟部組織	C45-C49	7	3	3	2.3	NA	NA
中枢神経系 ²⁾	C70-C72	6	3	3	2.0	5.6	1.9
白血病	C91-C95	32	16	16	2.0	11.1	0.7
女性							
胃	C16	158	63	72	2.2	127.5	1.8
大腸	C18-C20	246	101	101	2.4	174.4	1.7
肝臓及び肝内胆管	C22	53	17	25	2.1	30.6	1.2
肺	C33-C34	99	50	48	2.1	73.1	1.5
乳房	C50	378	192	177	2.1	251.3	1.4
子宮頸部	C53	39	25	28	1.4	38.4	1.4
膵	C25	28	17	13	2.2	9.8	0.8
膀胱	C67	27	16	17	1.6	14.3	0.8
甲状腺	C73	62	23	23	2.7	40.2	1.7
中皮及び軟部組織	C45-C49	7	3	2	3.5	NA	NA
中枢神経系 ²⁾	C70-C72	5	3	3	1.7	5.4	1.8
白血病	C91-C95	28	14	13	2.2	8.6	0.7

注 1) NA：Non-available

2) 中枢神経系：患者調査にはC75.1-C75.3を含む。

平均レセプト件数は約2倍であった。また、膵では、総患者数よりも5年有病数が少なかった。甲状腺では、5年有病数は総患者数の約2倍、月平均レセプト件数は約3倍であり、罹患数の多い部位と類似の特徴を示した。罹患数の少ない部位については、中枢神経系の腫瘍では、月平均レセプト件数は総患者数の約2倍で、月平均レセプト件数と5年有病数の差は小さかった。白血病では中程度の罹患数の膵と同様に、5年有病数が総患者数より小さかった。中皮及び軟部組織の腫瘍については、5年有病数は利用できず、月平均レセプト件数は総患者数の2倍から3倍であった。

表に示していないが、全部位の悪性新生物について、平成20年総患者数と月平均レセプト件数との比を都道府県別に比較したところ、全国

では1.6で、最小値1.3、最大値2.2であった。人口規模や地域性による明らかな特徴は認められなかった。

Ⅳ 考 察

レセプトは、疾病の治癒の状況に関わらず、診断・診療、検査、管理等の医療・保健サービスについて請求されるので、定義的には患者調査の総患者数に近いと考えられる。しかし、本研究では、同じ患者について、同じ月に請求された複数の医療機関からのレセプトの突合をしていないため、月平均レセプト件数は総患者数と比べて過大評価になる前提で集計を行った。その結果、年齢、性別、部位別、都道府県別で、月平均レセプト件数は総患者数を常に1から2.9倍上回っていた。腫瘍単位ではなく、人単位の集計を意図した全部位でも、月平均レセプト件数が総患者数の1.6倍であったことは、1人の患者について同一月内に、同一傷病名で複数の医療機関でレセプト請求されることがまれではない状況を反映していると考えられる。部位別の月平均レセプト件数が過大評価になる要因としては、最小集計単位をICD-10の主部位とした影響が考えられる。例えば、本研究では、同一レセプト内に行結腸がん、S状結腸がん、直腸がんの記載があれば、それぞれ独立して大腸がんの件数の中に含まれている。この場合、実際に個々の部位に独立したがんが存在した可能性もあれば、傷病名の記載方法での違いの可能性もある。2つ以上のがん傷病名レコードを持つレセプトが34%あったが、これは報告されている多重がんを有する割合5-10%⁷⁾と比べて高いことから、部位によっては、本来は同じ悪性新生物について傷病名の記載の違いによる重複カウントが含まれている可能性がある。

5年有病数について、比較的早期に発見されることの多く生存率の高い、胃、大腸、甲状腺については、総患者数の約2倍であるのに対して、比較的生存率の低い部位である肝臓、肺、膵、白血病については1.5倍以下であり、有病者数×受療割合＝患者数の関係に矛盾しない結

果であった。しかし、膀胱、中枢神経系など、生存率と受療割合の関係だけでは説明できない結果もみられ、有病者数の推計に用いている推計罹患数に由来する影響が疑われた。罹患数の多寡による傾向は明らかではなかったが、胃がんや膀胱がんは最初の病変のみ計測対象とするようながん登録独特の規則が関係している可能性がある。

がん罹患数から推計する5年有病数および患者調査と比べた時のレセプト情報等を利用したがん患者数計測の利点は、ほぼ全数調査であること、毎月計測できる即時性の良い統計指標となることがあげられる。標本調査である患者調査では推計精度の問題から千人未満は表示されないのに対して、レセプト情報では、特定の個人を特定されない範囲で、医療機関所在地の都道府県別、性別、年齢階級別の集計値を求めることができる。5年有病数は、将来推計によって即時性を補える利点はあるが、現時点では、推計に用いるがん罹患率や生存率が全都道府県で把握されていないため、日本全体の5年有病数は推計罹患数に基づく推計値であり、都道府県別の推計値を得ることができない。

レセプト情報等を利用したがん患者数計測において現時点で不可能なことは、第一に患者の居住地別の集計である。レセプトデータベースには、医療機関の所在地情報は持つが、患者の居住地情報は集積されていないためである。その他、レセプト情報等は、個人識別情報はハッシュ関数化されているため、同じ人物でも名前の誤記、結婚等で名前が異なる場合等、ハッシュ値も異なるため突合が困難なこと、保険者が変われば同一人物の突合は困難なことから、完全な同一人物の突合はできない。さらに、一般に保険傷病名と呼ばれる保険請求のための傷病名が一定数含まれている可能性が高いことから、厳密に正確な統計値を求めることは難しいと考えられる。しかし、これらの問題は、現在のレセプトデータベースは、本来、診療報酬請求処理のために構築されたサーバであり、データベースとしての利用を意図して構築されていないために当然であるといえる。レセプト情報

等を研究利用する利用者側が、その利点と限界を十分に理解した上で、存在する大量データを有効に活用する方法を見いだすことが重要であると考え。例えば、今回は提供を受けなかった個人識別情報をハッシュ関数化した「ID1」「ID2」という項目の提供を受ければ、同一人物の重複カウントを大幅に避けることができる。また、医科レセプトにおいて入院と入院外を区別する「レセプト種別」という項目の提供を受ければ、異なる傷病名コードを持つ同一の悪性新生物の判断をする材料として有用である可能性がある。

V 結 語

新たに利用できるようになった電子レセプト情報等について、日本のがん患者数計測資料としての可能性を、患者調査の総患者数と推計5年有病数との比較において記述した。ナショナルレセプトデータベースとして、毎月自動的に、一定の様式で、ほぼ全数調査に近いデータが蓄積されるレセプト情報は、既存資料を利用した日本のがん患者数計測資料として一定の利用可能性があると考えられた。

文 献

- 1) 松田智大, 雑賀久美子. 第6章 日本のがん有病数の推計. 祖父江友孝, 監. がん・統計白書2012. 東京: (株) 篠原出版新社, 2012; 101-16.
- 2) 厚生労働省ホームページ. レセプト情報・特定健診等情報の提供に関するガイドライン(平成23年3月) (http://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryou/iryouhoken/reseputo/dl/guide02_02.pdf) 2014.9.4.
- 3) 疾病, 傷害, 死因および統計分類 (ICD-10 2003年版準拠). 一般財団法人厚生労働統計協会
- 4) 政府統計の総合窓口 (e-Stat), 平成20年患者調査 (厚生労働省). 表番号148-1~6 総患者数 (患者住所地), 性・年齢階級×傷病小分類×都道府県別 (<http://www.e-stat.go.jp>) 2013.6.27.
- 5) 政府統計の総合窓口 (e-Stat), 平成23年患者調査 (厚生労働省). 表番号150-1~6 総患者数 (患者住所地), 性・年齢階級×傷病小分類×都道府県別 (<http://www.e-stat.go.jp>) 2013.6.27.
- 6) Miquel Porta編, 日本疫学会訳. 疫学事典第5版. 東京: 日本公衆衛生協会, 2010.
- 7) Tabuchi T, Ito Y, Ioka A, et al. Incidence of meta-chronous second primary cancers in Osaka, Japan: update of analyses using population-based cancer registry data. Cancer Sci. 2012; 103: 1111-20.

