

レセプトによるがん診断の妥当性研究 —大阪大学附属病院—

研究分担者： 松村 泰志 国立病院機構大阪医療センター 院長

武田 理宏 大阪大学大学院医学系研究科 医療情報学 准教授

要旨

健保組合のレセプトデータから新規がん治療の開始を検出し、がん検診の精度評価に適用することを目的とし、小川分担研究者が開発した「新規がん診断決定ロジック」の精度を、大阪大学医学部附属病院のレセプトデータとがん登録データを突合させて評価した。2014年1月1日から2018年12月31日にレセプト上各5大がん（胃がん、肺がん、大腸がん、乳がん、子宮頸がん）の治療開始が確認され、その内2012年1月1日～2013年12月31日にがん治療された患者を継続治療症例として除いた患者を母数とし、治療開始日365日前～90日後にがん登録で診断日がある症例を真陽性として陽性予測値を算出した。がん登録で2014年1月1日～2017年12月31日の間に各5大がんと診断された患者を母数とし、診断日より90日前～365日後の間にレセプトでそのがんの治療開始が検出された患者を真陽性として感度を算出した。令和2年度は、入院でDPC適用患者では様式1の病名のみを判定に利用したが、DPC適用患者でも病名とがん治療行為のどちらもある場合を陽性したこと、令和2年度はがん治療行為にがん指導管理等の加算を含めていたが、これらの治療行為を除くようにロジックを改訂した。陽性予測値は、胃がん765例/824例(0.93)、肺がん634例/722例(0.88)、大腸がん620例/752例(0.82)、乳がん1061例/1234例(0.86)、子宮頸がん(CINIIIを含む)237例/269例(0.88)であった。感度は、胃がん776例/1060例(0.73)、肺がん641例/712例(0.90)、大腸がん627例/930例(0.67)、乳がん1049例/1087例(0.97)、子宮頸がん241例/378例(0.64)、CINIIIを含めた子宮頸がん241例/615例(0.39)であった。偽陽性例は、病理診断で良性と判定されたもの、他がん、再発がんの症例であり、がん検診としては陽性とすべきものであった。がん検診で陰性で「新規がん診断決定ロジック」で陽性の場合には、がん検診の見逃しと指摘することは妥当であることが確認できた。

A. 研究目的

健保組合では、がん検診データとレセプト情報を保有している。レセプト情報からがんの発症が分かると、がん検診の精度を調べることができる。本研究班では、健診を実施する

5大がん（胃がん、肺がん、大腸がん、乳がん、子宮頸がん）について、がん検診のがんの新規発見についての精度評価を目指している。そのため、レセプト情報から5大がんのそれぞれの新規発生を捉える手法を確立する必要

がある。

レセプト情報は、各医療機関が請求のために作成したデータであり、記載されている病名とその開始日を必ずしも信用できない。がんを疑って検査を実施する場合に、病名として、がん疑いと正しく登録されず、疑いを付けずに登録されているケースがある。また、がんであるにも関わらず、がんの病名を付けずに診療をしている場合もあり得る。このようにレセプトの病名だけでは正しくがんの発症を捉えることができないことが予想されることから、小川分担研究者らは、5大がんについて、レセプトの病名に加えて、レセプトから取得できるがんに対する治療のデータがある時に当該がんを判定する「新規がん診断決定ロジック」を開発した。本研究では、レセプトデータを用いた「新規がん診断決定ロジック」による5大がんの新規がん患者の検出精度を、大阪大学医学部附属病院（阪大病院）のデータを用いて評価することを目的とする。

阪大病院では、レセプトデータに加え、がん登録のデータが利用できる。阪大病院では、がん登録は、がん登録士により高い精度で実施されている。更に、診療録の内容を調査することが可能であり、偽陽性、偽陰性事例の原因の調査が可能である。一方、1病院のデータであるので、阪大病院で診断されながらも、治療は他院で実施されるケースが含まれてくる。健保組合では、複数の医療機関を受診した場合でも、レセプトデータは集約されるので、どこかの医療機関で治療を受けていれば検出できるメリットがある。従って、阪大病院で評価した本ロジックの精度が、そのまま、健保組合のデータを使った場合の精度となるわけではな

い。そこで本研究では、阪大病院のデータでの「新規がん診断決定ロジック」の精度評価をし、次に、偽陽性、偽陰性事例を調査し、「新規がん診断決定ロジック」を健保組合のデータに適用した場合の精度を予測した。

令和2年度に開発した「新規がん診断決定ロジック」では、手術は各5大がんで設定し、がん指導管理等、放射線治療、薬物治療は、臓器非特異的にレセプト上の治療行為コードをリストアップして適用させた。また、入院患者でDPC適用の場合には、様式1の病名でがんの病名がある場合に当該がんを判定させた。しかし、DPC様式1の病名だけでは、その入院でがん治療がされたとは限らないケースがあること、がんに関する指導管理等の行為が、がん診療が行われていることの証拠としては弱いと考えた。そこで、入院患者においても、がんに対する治療（手術、放射線治療、抗がん剤治療）があった症例をがん治療があったこととするようロジックを変更した。また、がんに関する指導管理の行為を、がん治療がされたことの証拠としての行為から外したロジックに変更した。今年度は、この変更した「修正新規がん診断決定ロジック」でデータを処理し結果を再評価した。

B. 研究方法

1. 解析対象データの抽出処理

阪大病院を受診した患者のレセプトのうち2012年1月1日～2018年12月31日のレセプトデータに対して「新規がん診断決定ロジック」を適用して、処理対象となるレコードを生成した。また、阪大病院のがん登録のデータから、起算日（診断日）が2013年1月1日～2017

年 12 月 31 日までのデータを抽出した。個人情報保護担当者が、レセプト処理データとがん登録データに対して、患者 ID を同じルールで別番号に変換して仮名化し、処理対象データを作成した。本研究は、「人を対象とする医学系研究に関する倫理指針」に基づき、阪大病院の倫理審査委員会の承認を得て実施した（承認番号 19125）。

「修正新規がん診断決定ロジック」では、レセプト月単位で 1 患者に対し 1 レコードを生成する。5 大がんのそれぞれについての病名があり、かつ、がん治療（手術、放射線治療、薬物治療）があった場合に 5 大がんのそれぞれにフラグを立てた。阪大病院では、外来患者については医科レセに診療報酬請求情報が記録されるが、入院患者で DPC 適用の患者に対しては DPC レセに診療報酬請求情報が記録される。令和 2 年度に適用したロジックでは DPC レセに対する処理を様式 1 の病名から判定していたが、修正ロジックでは、DPC レセについても、病名があり、かつ、がん治療（手術、放射線治療、薬物治療）があった場合にフラグを付ける方法に変更した。

対象期間の全レコードを集め、患者 ID を第 1 ソートキー、レセプト月を第 2 ソートキーとして並べ替え、各がんについて、各患者でフラグが立った月を抽出した。

2. 陽性予測値の算出

各がんについて、患者 ID を第 1 ソートキー、レセプト月を第 2 ソートキーとして並べ替え、各患者で最初にフラグがたった月を治療開始月とした。レセプト上で 2014 年 1 月から 2018 年 12 月までの間に初めて治療開始が検出され

た患者を対象とした。2014 年 1 月 1 日時点で、その前から継続してがん治療をしている患者を除くために、2 年間でワッシュアウト期間とし、2012 年 1 月 1 日～2013 年 12 月 31 日の間に治療開始月がある患者を除いた。

レセプト上、そのがんの治療が開始された日の 365 日前以後に、がん登録で当該部位のがんと診断された場合を真陽性とした。レセプトでは月単位で検出されるため、治療開始日を各月の 15 日として計算した。DPC 症例については入院月が治療開始月となり、実際の治療開始日が捉えにくい。そこで DPC で捉えた治療開始日より 90 日後までにがん登録の診断日がある場合には、このがん診断による治療と判定した。従って、レセプト上の各がんの治療開始日の 365 日前から 90 日後までの間に、がん登録上、そのがんの診断日がある場合に真陽性として、陽性予測値を算出した。

レセプト上がんの治療が開始されたとした事例でがん登録上そのがんの診断がなかった偽陽性事例について、その原因を調べるために、偽陽性事例について、レセプト上のがん治療開始日を中心に診療録を閲覧し、偽陽性となった原因を調査した。

がん登録では、がんの発生部位は病理学的に厳密に判定される。このため、腹部食道がんは、臨床的には胃がんとして治療がされるが、がん登録上は食道がんに分類される。そこで、がん登録上の腹部食道がんを胃がんを含めて判定した。同様に、がん登録で気管がん、気管支がんを肺がんを含めて解析した。大腸がんについてはがん登録上の盲腸、虫垂、結腸、直腸、肛門管のがんとした。子宮頸がんでは、子宮頸部上皮内腫瘍（CIN）と診断される事例が

多い。CIN I（軽度異形成）とCIN II（中等度異形成）は経過観察とされることが多いが、CIN III（高度異形成・上皮内がん）は、治療の対象となる。そこで、がん登録でCIN IIIと診断されたものは、子宮頸がんを含めて判定した。

3. 感度の算出

がん登録で、2014年1月1日から2017年12月31日の間に、5大がんの各がんと診断された患者を対象とした。ただし、腹部食道がんを胃がん、気管がん、気管支がんを肺がんを含め、盲腸、虫垂、結腸、直腸、肛門管のがんを大腸がんとした。子宮頸がんについては、子宮頸部上皮内腫瘍Ⅲを含める場合と含めない場合について判定した。

同一症例でこの期間に同一部位のがんの診断がある例がある。この場合には、最初のがんと診断された日を採用し、その後の診断については対象から除いた。

がん登録で各5大がんと診断された患者について、当該がんの診断日より90日前から365日後までの間にレセプトでそのがんの治療が検出された患者を正しくレセプトでがん治療の開始が検出できたものとし、これを真陽性として感度を算出した。ただし、レセプトは月単位で判定されるため、レセプト月の15日を治療日とした。

がん登録では、阪大病院で治療がされたか否かが登録されている。阪大病院では、がん治療方針の相談を受けるケースがあり、治療がされない（他院で治療がされる）ケースが比較的多い。これらの症例を含めると感度が落ちることから、がん登録上で、阪大病院で治療を受けていないとされている症例を母数から除

いて感度を算出した。

がん登録上阪大病院で治療を受けているとされる症例で、レセプト上で治療ありとならない偽陰性例について、その理由を調べるために、上記期間内にレセプトにおいて、何らかのがんの病名が登録されているか、手術、放射線治療、薬物治療に加え、がんについての指導管理等の行為があるかについて調べた。

C. 結果

1. 陽性予測値の評価

レセプト上2014年1月から2017年12月までに、各5大がんの治療開始が検出された患者数は、胃がん824例、肺がん722例、大腸がん752例、乳がん1234例、子宮頸がん269例であった。この内、指定期間内のがん登録で当該がんと診断されていた患者数は、胃がん765例、肺がん634例、大腸がん620例、乳がん1061例、子宮頸がん237例であった（表1）。

偽陽性例の数は、胃がん64例、肺がん89例、大腸がん133例、乳がん173例、子宮頸がん32例であった。この内、がん登録の不備で本来がん登録で登録されるべき症例が登録されていなかった症例が胃がん5例、肺がん1例、大腸がん1例あった。これらを修正した場合の陽性予測値は、胃がん $765/824$ (0.93)、肺がん $634/722$ (0.88)、大腸がん $620/752$ (0.82)、乳がん $1061/1234$ (0.86)、子宮頸がん $269/237$ (0.88)であった（表1）。

偽陰性症例の原因の内訳を表2に示す。胃がんでは、胃の周囲臓器のがん等の他臓器がんであった症例が59.3%、病理診断上は良性であったものが11.9%、再発例が28.8%であった。肺がんでは、他臓器がんが60.2%であり、肺に

転移したがんが多く含まれていた。再発がんや他院で治療し当院で継続治療した症例が30.6%、良性腫瘍が9.1%であった。大腸がんでは、良性腫瘍が43.9%と最も多く、他臓器のがんが34.8%、再発および継続治療が20.4%であった。乳がんでは、再発および継続治療が48.0%と多かった。院外処方の場合、レセプト上には内服のがん薬物療法は記録されない。当院で治療の場合は、内服のみの治療はほぼ無かったが、他院で主な治療をしている場合で、当院が継続治療となった場合には、当院での治療が見えないためにワッシュアウトがされず、別目的で入院した際にがん薬物療法が初めてレセプト上で確認され、初発のがんであるかのように見える症例があった。形成外科での乳房再建を目的とした治療でこの間抗癌剤治療が継続された症例が34.7%、他臓器のがんが9.8%であった。子宮頸がんでは、子宮体がんを含む他癌が65.6%と多く、再発および治療の継続が28.1%、良性腫瘍が6.3%であった。

2. 感度の評価

がん登録で2014年1月1日から2017年12月31日までの間に5大がんの診断日がある患者の患者数は、胃がん1291例、肺がん921例、大腸がん1162例、乳がん1297例、子宮頸がん488例、CINⅢを含めた子宮頸がん761例であった。この内、がん登録で、自院で治療ありと登録されている患者数は、胃がん1060例、肺がん712例、大腸がん930例、乳がん1087例、子宮頸がん378例、CINⅢを含めた子宮頸がん615例であった。この内、当該がんの診断日から指定期間の間にレセプト上で当該がんの治

療開始が確認できた患者数と割合(感度)は、胃がん776例(0.73)、肺がん641例(0.90)、大腸がん627例(0.67)、乳がん1049例(0.97)、子宮頸がん241例(0.64)、CINⅢを含めた子宮頸がん241(0.39)であった(表3)。

偽陰性例(がん登録において自院で治療ありと登録されながら、レセプト上、当該がんの治療が見当たらなかった事例)の患者数と、その内、レセプト上、病名、治療行為のいずれかが欠けていた件数を表4に示す。各がんについて、偽陰性件数、病名のみで治療行為がなかった例の件数、治療行為があったが病名ががん疑いや他のがん病名であった例の件数、がん疑いや他のがん病名のみで治療行為がなかった例の件数、がんに関する病名が登録されておらず抽出されなかった例の件数は、それぞれ胃がんで284例/83例/13例/15例/173例、肺がんで71例/16例/2例/4例/49例、大腸がんで303例/168例/34例/10例/91例、乳がんで38例/14例/4例/2例/18例、子宮頸がんで137例/87例/0例/1例/49例であった。

病名が正しく登録されていながら治療行為が検出できなかった事例は、治験等の臨床研究の事例、ダビンチ手術など当時は保険適用がなかった治療が行われた事例、隣接臓器に浸潤しておりレセプト上は他臓器の摘出術として登録されていた事例、悪性腫瘍が付かない鏡視下手術或いは内視鏡手術として術式が登録されたために治療ありと検出できなかった事例、緩和治療がされたために治療が検出できなかった事例であった。

病名が疑いのままで登録されていた事例は胃がん、大腸がんで多かった、レセプト上では

ポリープ等の良性腫瘍での病名が付けられ、病理検査で悪性と判定されたが、病名は修正されなかった事例であった。

D. 考察および結論

5大がんについて、レセプトから治療があったことを検出する「修正新規がん診断決定ロジック」の陽性予測値は、胃がん93%、肺がん88%、大腸がん82%、乳がん86%、子宮頸がん88%であった。

偽陽性症例のうち切除後の病理診断で良性腫瘍のものがあつた。こうした事例は大腸がんで多く、胃がん、肺がん、子宮頸がんの順であつた。これらは、精査して初めて良性か悪性かの鑑別ができた症例であり、健診としては要精査として指摘すべき症例であると言える。

乳がんでは、手術後の抗癌剤治療が長く継続することが多い。こうした場合に、がんそのものの治療のための入院でなくても、抗癌剤投与がされるため、「新規がん診断決定ロジック」でがんフラグが付くことになる。がんの手術はかなり前に実施されているため、或いは他院で実施されているために、ワッシュアウト処理で除かれず、新規のがんと判定されケースがあつた。特に、形成外科で乳房の再建手術を目的に紹介された患者では、この判定ロジックで新規がんとして判定されるケースが多かつた。検診では、既のがんの治療中の患者を陽性と判定することになっても誤った判定とはならない。

それ以外の偽陽性例は、当該のがんの初発事例ではなかつたものの、他癌であつたり、当該がんの再発事例であつたりしたもので、がん検診で陽性とすべき症例であつた。

がん登録がされ、その内阪大病院での治療例として登録された症例に限定した場合の感度は、胃がん73%、肺がん90%、大腸がん67%、乳がん97%、子宮頸がん64%、CINⅢを含めた子宮頸がん39%であつた。阪大病院でがん登録された患者のうち、自院治療とされていない患者は、他院で治療を受けていると考えられ、健保組合のレセプトデータであれば、これも含めて捉えることができるはずである。

自院治療がされているにも関わらずレセプト上で治療開始が検出されなかつた偽陰性例は、多くがレセプトに当該がんの病名が登録されていなかつた事例であつた。病名が正しく登録されていながらレセプト上治療を捉えられなかつた事例には、治験や臨床研究、ダビンチ手術で当時保険適用外の治療、隣接臓器の摘出術、悪性腫瘍が付かない術式で手術が登録されたことが原因のものが多かつた。

子宮頸がんについて、CINⅢまで含めると、感度は39%と低かつた。「新規がん診断決定ロジック」では、子宮頸部上皮内癌までは十分に検出できるロジックではないことに注意する必要がある。

令和2年度の「新規がん診断決定ロジック」では、がん指導管理の算定等をがんの診療行為として抽出していた。これを除くことで陽性予測値は少し改善したが、感度はやや低下する結果となつた。感度分析の偽陰性例の中に、治療行為としてがん指導管理の算定等のみであつた症例が、胃がんで83例、肺がんで16例、大腸がんで165例、乳がんで14例、子宮頸がんで76例あつたことから、これらの分の感度が悪化した。

本研究の最終的な目的は、がん検診の精度

を評価することである。本評価の結果を踏まえて考えると、本「新規がん診断決定ロジック」で陽性と判定された場合、レセプト上の判定が過大であることはほぼ無いと言える。修正した「新規がん診断決定ロジック」では、がんに対する手術か、放射線治療かがんに対する薬物治療がされていることが条件となるので、確実にがん診療を要した患者を検出していると言える。一方、感度は十分に高いとは言えないことから、がん検診で陽性の患者が、レセプトを利用した「新規がん診断決定ロジック」で陰性であったとしても、がん検診が正しい可能性は十分にある。特に、子宮頸がんについては、CINⅢを含めると感度は低いので、検診結果が尊重されるべきである。一方、がん検診で陰性であった受診者がレセプトデータに「新規がん診断決定ロジック」を適用して陽性と判定された場合、がん検診で陽性として判定すべき症例であると言え、検診でがん見逃し

例として判定することは妥当であることが確認できた。

E. 健康危険情報

なし

F. 研究発表

1. 論文発表

なし

2. 学会発表

なし

G. 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

	胃がん	肺がん	大腸がん	乳がん	子宮頸がん
レセ上治療開始検出	824	722	752	1234	269
がん登録 当該がんの診断あり	765	634	620	1061	237
陽性予測値	0.93	0.88	0.82	0.86	0.88

表1 各がんの陽性予測値

	胃がん	肺がん	大腸がん	乳がん	子宮頸がん
他癌(他疾患)	35 (59.3)	53 (60.2)	46 (34.8)	17 (9.8)	21 (65.6)
良性	7 (11.9)	8 (9.1)	58 (43.9)	12 (6.9)	2 (6.3)
再発	17 (28.8)	23 (26.1)	25 (18.9)	69 (39.9)	8 (25.0)
継続治療	0	4 (4.5)	2 (1.5)	14 (8.1)	1 (3.1)
治験	0	0	0	1 (0.6)	0
未確定死亡	0	0	1 (0.8)	0	0
再建術のみ	0	0	0	60 (34.7)	0
計	59	88	132	173	32

()内は%

表2 各がんの偽陽性例の理由

	胃がん	肺がん	大腸がん	乳がん	子宮頸がん	子宮頸がん (CINⅢ含む)
がん登録あり	1291	921	1162	1297	488	761
がん登録あり+自院治療あり	1060	712	930	1087	378	615
自院治療あり&レセ+	776	641	627	1049	241	241
自院治療あり内の感度	0.73	0.90	0.67	0.97	0.64	0.39

表3 各がんの感度

	胃がん	肺がん	大腸がん	乳がん	子宮頸がん
病名のみ、治療行為なし	83(29)	16(23)	168(55)	14(37)	87(64)
疑い病名+治療行為	7(2)	1(1)	31(10)	3(8)	0
他のがん病名+治療行為	6(2)	1(1)	3(1)	1(3)	0
疑い病名のみ	5(2)	1(1)	5(2)	0	0
他のがん病名のみ	10(4)	3(4)	5(2)	2(5)	1(1)
がん病名なし	173(61)	49(69)	91(30)	18(47)	49(36)
計	284	71	303	38	137

()内は%

表4 各がんの偽陰性例の理由