

一般口演

一般口演24

医療データ分析10（NDBの活用）

2018年11月25日(日) 13:40～15:10 G会場 (5F 504+505)

[4-G-3-4] レセプト情報・特定健診等情報データベース（NDB）利用促進に向けた取り組みー1患者1データ化ーレセプト情報・特定健診等情報データベース（NDB）利用促進に向けた取り組みー1患者1データ化ー

○明神 大也^{1,3}, 野田 龍也¹, 久保 慎一郎^{1,4}, 西岡 祐一^{1,5}, 東野 恒之², 今村 知明¹（1.奈良県立医科大学 公衆衛生学講座, 2.（株）三菱総合研究所 経営イノベーション本部, 3.奈良県立医科大学附属病院 病理診断科, 4.奈良県立医科大学附属病院 看護部, 5.奈良県立医科大学附属病院 糖尿病センター）

【背景】レセプト情報・特定健診等情報データベース（NDB）を利用する上で、問題点の1つに、患者の追跡が困難であることが挙げられる。その原因として、レセプトが月別・個人別・医療機関別に発行されるために個人追跡用IDにて名寄せが必要であるが、その個人追跡用IDを用いても1年間に1割以上の患者が脱落する問題があった。発表者は、NDBデータの再データベース化・個人追跡用IDの名寄せ精度向上・1入院中のSIレコードの結合成功といった、患者の追跡にむけての技術開発・発表を進めてきた。【目的】本研究の目的は、医療計画策定に係る評価指標作成に向けての一環として、ほぼ正規化されたデータベース上で患者の追跡を実現することである。【方法】平成25年4月から平成28年3月までのNDBデータ（医科・DPC）を用いた。まず、①入院と②外来に分けて処理した。①入院では、医科入院とDPC出来高部分をCDファイル形式にして、DPC包括部分と結合した。そして医療機関コード・ID0・病棟区分をキーとして、入院日が途切れるまでを1入院として分割・統合した。②外来では医科レセプト（入院外）と調剤レセプトを、ID0・処方箋発行元の医療機関コード・診療日（処方箋発行日）をキーとして結合した。そして①と②をID0で結合した。なおID0は、NDBに備わっている2つの個人追跡用IDと転帰区分、診療年月から生成した、名寄せ精度向上させたIDである。【結果】上記処理後の一例として、平成27年4月から平成28年3月の1年間の性別・年齢階級別入院日数を算出した。その結果は男女ともに20代以降は右肩上がりとなり、女性は伸び続けた100歳以上で80.0日となったが、男性では95～99歳の60.8日が最大となった。【まとめ】本研究で、1患者1データ化がほぼ完了した。本機能を利用して、患者を軸とした長期追跡性が求められる調査、すなわちコホート研究が可能となった。

レセプト情報・特定健診等情報データベース(NDB)利用促進に向けた取り組み

- 1 患者 1 データ化 -

明神 大也^{*1*3}、野田 龍也^{*1}、久保 慎一郎^{*1*4}、西岡 祐一^{*1*5}、東野 恒之^{*2}、今村 知明^{*1}

*1 奈良県立医科大学 公衆衛生学講座、*2 (株)三菱総合研究所 経営イノベーション本部、

*3 奈良医科大学附属病院 病理診断科、*4 奈良医科大学附属病院 看護部、

*5 奈良医科大学附属病院 糖尿病センター

Promoting More Effective Use of the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB)

- trace the patient data over the health insurance claim -

Myojin Tomoya^{*1*3}, Noda Tatsuya^{*1}, Kubo Shinichiro^{*1*4}, Nishioka Yuichi^{*1*5}, Higashino Tsuneyuki^{*2}, Imamura Tomoyaki^{*1}

*1 Department of Public Health, Health Management and Policy, Nara Medical University,

*2 Management Innovation Division, Mitsubishi Research Institute, Inc.,

*3 Department of Diagnostic Pathology, Nara Medical University

*4 Nursing Department, Nara Medical University

*5 Department of Diabetology, Nara Medical University

Background: The National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) is a comprehensive database of health insurance claims data under Japan's National Health Insurance system. One of the NDB problems is the difficulty of tracing the patient data over the month. As the claims were performed by month, medical institution and patient, the NDB users have to use the personal identification variables (ID1/ID2). However, more than 10% of ID1/ID2 change for 1 year. We have reported the new technique to trace the patients' health insurance claims data such as following; to re-construct the relational database of NDB data, make the personal identification variables more precise and a new style of tracing the hospitalization data over the month.

Aim: The aim of this study is to make it possible to trace the patient data over the rezept

Methods: Our study targeted data from health insurance claims filed between April 2013 and March 2016. 1st step was to trace the hospitalization data over the month, 2nd step was to trace the outpatient data over the claim and 3rd step was to combine the 2 steps. 1st step: we converted DPC piecework and medical hospitalization data to CD file style, combined it and DPC comprehension, and converted the consecutive hospitalization days to a hospitalization sequence. The sequence was by ID0, medical institution code and hospital ward classification. 2nd step: we joined dispensing claims and medical outpatient claims by the key of ID0, medical institution code published prescription and the date publishing prescription.

Results: In order to confirm the reliability of the converted data, we calculated the outpatient number and average outpatient visit.

Conclusion: Our study succeeded a new method of tracing the patients' data over the health insurance claim.

Keywords: NDB, rezept, database, health insurance claims in Japan, patient identification

1. 緒論

日本は国民皆保険制度をとっている。日本に合法に居住する人は、90 日以内の短期滞在者と生活保護加入者を除き、外国籍を含め、全員健康保険に加入し、受診の際は保険診療が適用される。保険診療を行う場合、医療機関は保険者に対し保険診療請求(レセプト)を発行する。レセプトは調剤、Diagnosis Procedure Combination(DPC)、医科入院、医科入院外、歯科の 5 種類がある。レセプトは種類ごと、月ごと、医療機関ごと、患者ごとに発行される。

また、DPC は特定の疾患、特定の病院に限って適用される包括医療支払制度のことで、疾患によって対象期間が異なっており、対象期間を超えると出来高算定に移行する。同一个月内に出来高に移行した場合、DPC レセプト内に DPC と出来高の内容が記載されて発行される。

日本では、高齢者医療確保法を根拠法として、2009 年 4 月からレセプト情報・特定健診等情報データベース (NDB) にレセプト情報を蓄積している。生活保護による医療扶助、治験、全額公費負担の治療を除く全診療報酬情報を有し、2017 年 12 月時点で約 148 億 1,000 万件のレセプトデータを格納している。NDB は世界最大級の健康情報データベースであり、RCT や他のコホート集団では扱えないような稀な疾患や稀な状況を扱うことができる。類似の日本のデータベースとして JMDC Claims Database や限られた集団のレセプト分析があるが、NDB は標本サイズが数十倍～数万倍大きい。

レセプトには患者情報、病院情報に加えて、病名(疑いを含む)、行った診療行為、調剤した医薬品などが含まれる。ただし現時点では患者住所は含まれない。歯科レセプトを除く NDB の全データ量は、2010 年代中葉においては、1 年間

で CSV 3TB 程度である。

NDB では匿名化を目的に、個人を特定できる恐れのある健康保険の保険者番号・被保険者証記号・被保険者証番号・氏名・生年月日の情報は消去されている。その代わりに、レセプト同士をつなぐために健康保険の保険者番号・被保険者証記号・被保険者証番号・性別・生年月日をハッシュ化した ID1、および氏名・性別・生年月日をハッシュ化した ID2 が存在する。つまり、「個人を特定」することはできないが、「個人を区別」して紐つけることは可能である。しかし、1 年間の医科入院外レセプトのうちで ID1 は 7.1%、ID2 は 20.4% 変更される脱落すると報告¹⁾されている。そこで久保ら²⁾は ID1、ID2、転帰区分、診療年月をもとに生成した、名寄せ精度向上させた ID0 (ver.1) を提案した。

NDB の問題点の 1 つに、NDB の DPC レセプトが特殊な形であること³⁾が挙げられる。本来 DPC レセプトはレセプト共通 (RE) レコードや保険者 (HO) レコードなどの共通レコードと、包括行為レコードおよび出来高行為レコードに分かれる。診療行為 (SI) レコード、医薬品 (IY) レコード、特定器材 (TO) レコードなどに含まれる出来高行為は一般レセプトと同様のレコードに記録されるが、包括行為の内容についてはコーディング (CD) レコードの中に出来高で算定した場合の内容を記録することになっている。しかし NDB では、包括期間と出来高期間が混在する月のレセプトデータにおいて、包括行為と出来高行為をまとめて包括行為として包括行為レコードに取り込んでいる。また、入院退院日が正確に記録されない。ある月の初回入院日は RE レコードに記録されるが、その同月内に退院・再入院しても退院日や再入院日の情報はレセプトの摘要欄にコメントとして記載されているが、NDB では摘要欄のコメント情報は取り込まない。さらに、月またぎの入院でも 2 か月目以降のレセプトに入院日は記録されない。その結果、入院期間を把握できなかつたり、包括対象の診療行為と出来高の診療行為 (処置・投薬など) を分離することができなかつたりした。これを解決する手法を 1 入院 1 データ化と称した。

この課題に対して、藤森ら⁴⁾は DPC の生データと医科入院データを EF 化した上で、DPC 分析ツール上で処理する方法を開発・運用していた。同様の方法を全国 1 年分の NDB データで実施すると、データ容量が 440GB から 12.3TB と大きく増加する上、2015 年 2 月当時で処理に約 7 ヶ月間を要した。そのため、継続して実施することが現実的に困難であった。

そこで著者らは、まず入院期間を把握できない点に着目した。データ容量の増大を抑えて処理時間を大幅に短縮することを前提にし、1 入院 1 データ化に至る 1 ステップである 1 入院期間の把握手法を開発し、第 37 回日本医療情報学連合大会にて発表した。

そして本研究では上記手法を前提に 1 入院 1 データ化を実現した。加えて外来・調剤レセプトを結合し、外来の患者の追跡を可能とする 1 外来 1 データ化を行った。最後に、入院と外来を結合し、1 患者の入院・外来の情報を追跡できる「1 患者 1 データ化」を行った。

2. 方法

2013 年 4 月から 2016 年 3 月までの NDB データ (医科・DPC・調剤) を用いた。① 1 入院 1 データと② 1 外来 1 データに分けて処理した。

① 1 入院 1 データでは、医科レセプト (入院) と DPC 出来高部分に含まれる SI, TO, IY レコードを CD ファイル形式にして、DPC 包括部分と結合した。そしてレセプト通番・医療機関コー

ド・ID0・病棟区分をキーとして、入院日が途切れるまでを 1 入院として分割・統合した。具体的には、まず医科入院と DPC 出来高部分の SI, TO, IY レコードを CD ファイル形式にして、DPC 包括部分と結合した。その際、DPC と出来高が混在しているレセプトでは包括対象のレコードが重複しているので、その重複を除外した。そして、SI レコードに関しては既報の通り、1 入院期間を把握した。なお、1 入院期間は医療機関及び病棟区分が同一である入院を 1 入院期間と定義した。そして、日付情報 31 列を年月日の 1 列にし、複数行で点数をひとくくりになっている場合は各行の点数に再配分した。次に IY レコードに関しては、回数・日数が同じ薬剤は点数と日数が一つにくられており、各行の点数に再配分した。

② 1 外来 1 データでは、医科レセプト (入院外) の SI レコードに関しては、① 入院と同様、日付情報 31 列を年月日の 1 列にし、複数行で点数をひとくくりになっている場合は各行の点数に再配分した。次に医科レセプト (入院外) の IY レコードに関しても、① 入院と同様、点数と日数が一つにくられている回数・日数が同じ薬剤を、各行の点数に再配分した。調剤レセプトの IY レコードも同様に、点数と日数が一つにくられている回数・日数が同じ薬剤を、各行の点数に再配分した。そして医科レセプト (入院外) と調剤レセプトを、ID0・処方箋発行元の医療機関コード・診療日 (処方箋発行日) をキーとして結合した。

最終的に、① 入院と② 外来を ID0 で結合した。これで 1 患者 1 データ化が完了した。

そして、1 患者 1 データ化後のデータベースをもとに、日本の保険診療の実態を調査した。

なお、本研究の遂行にあたっては、「人を対象とする医学系研究に関する倫理指針」(2014 年 文部科学省、厚生労働省告示) を遵守し、奈良県立医科大学医の倫理審査委員会の承認 (2015 年 10 月 8 日 受付番号 1123) を得た。

3. 結果

上記方法で行った 1 患者 1 データ化に対し、データの信頼性を調べるため、入院および外来の患者数、入院/受診日数を求めた。2015 年 4 月から 2016 年 3 月の、全患者数、入院患者数/平均入院日数/平均入院回数/平均再入院回数、外来患者数/外来総日数/平均外来日数/時間外患者数/時間外日数/平均時間外外来日数を表 1 に示した。時間外とは、時間外加算、休日加算、深夜加算、時間外特例、夜間・早朝等加算を算定された患者数/日数を意味する。平均とは、入院/外来患者数 1 人あたりを意味する。その結果、平均外来日数では年齢が上昇するにしたがって平均外来日数は上昇し、80~84 歳の 31.9 日/年/人をピークに漸減した。

次に、上記方法で行った 1 患者 1 データ化に対し、技術的な 1 例を示すために、急性心筋梗塞に対するステント留置術後の入院外来日数を示した。ここでは、2013 年 4 月から 2014 年 9 月の間に、K549-00 経皮的冠動脈ステント留置術 (急性心筋梗塞) / 診療行為コード 150375210 を算定した患者を対象とした。この処置は、急性心筋梗塞に対し、来院後 90 分以内に経皮的冠動脈ステントを留置し、冠動脈再開通に成功した場合に算定できることとなっている。対象患者を追跡し、上記処置実施 6 ヶ月後から 18 ヶ月後までの入院日数・再入院回数・外来日数を算出した。表 2 に患者数および術後 6 ヶ月~18 ヶ月の 1 患者あたりの入院日数、1 患者あたりの再入院回数、1 患者あたりの外来日数をそれぞれ示した。年齢が上昇するにつれて入院日数は増えるが、再入院回数は 70 歳代

表1 年齢階級別入院/外来の日数・回数(2015年4月～2016年3月)

	全患者数	入院				外来					
		患者数	平均 入院日数	平均 入院回数	平均 再入院回数	患者数	総日数	平均 外来日数	時間外 患者数	時間外 日数	平均時間外 外来日数
0～4歳	5,176,570	588,794	9.8	1.3	0.3	5,142,995	82,325,012	16.0	2,527,545	8,752,106	1.7
5～9歳	5,419,610	166,128	10.0	1.3	0.3	5,420,767	62,555,488	11.5	2,998,929	10,020,245	1.8
10～14歳	5,246,432	115,528	14.9	1.3	0.3	5,246,320	41,526,988	7.9	2,491,852	7,504,535	1.4
15～19歳	5,041,688	142,496	16.6	1.2	0.2	5,039,556	29,974,141	5.9	1,979,717	5,283,742	1.0
20～24歳	5,033,085	176,251	15.9	1.2	0.2	5,028,818	27,430,598	5.5	1,869,232	4,746,977	0.9
25～29歳	5,530,982	272,494	15.5	1.2	0.2	5,526,467	35,787,724	6.5	2,294,450	7,017,813	1.3
30～34歳	6,270,923	382,136	15.9	1.2	0.2	6,264,839	46,061,967	7.4	2,613,006	9,004,109	1.4
35～39歳	6,924,615	378,897	19.1	1.2	0.2	6,917,224	54,722,841	7.9	2,775,041	10,434,818	1.5
40～44歳	7,937,074	364,215	25.3	1.3	0.3	7,926,218	66,841,557	8.4	3,063,219	12,869,665	1.6
45～49歳	7,179,974	346,063	30.7	1.3	0.3	7,166,152	68,835,925	9.6	2,805,131	13,951,227	1.9
50～54歳	6,608,139	377,280	34.7	1.4	0.4	6,590,820	75,145,312	11.4	2,676,983	16,091,856	2.4
55～59歳	6,354,626	443,222	37.1	1.5	0.5	6,333,060	82,157,719	13.0	2,577,777	18,099,092	2.9
60～64歳	7,140,511	626,072	39.1	1.5	0.5	7,108,744	106,817,393	15.0	2,848,034	23,405,742	3.3
65～69歳	9,011,354	994,378	39.5	1.6	0.6	8,963,041	162,685,497	18.2	3,662,554	35,758,338	4.0
70～74歳	6,912,751	993,144	40.0	1.6	0.6	6,865,650	165,030,493	24.0	3,011,766	37,327,041	5.4
75～79歳	6,094,064	1,107,969	45.0	1.6	0.6	6,038,077	174,951,886	29.0	2,822,507	40,929,626	6.8
80～84歳	4,957,061	1,158,685	52.6	1.7	0.7	4,877,127	155,487,192	31.9	2,396,115	37,790,345	7.7
85～89歳	3,252,011	954,534	60.9	1.7	0.7	3,162,054	96,450,175	30.5	1,563,510	23,440,619	7.4
90～94歳	1,536,586	540,752	67.8	1.6	0.6	1,467,166	39,673,558	27.0	699,142	8,982,491	6.1
95～99歳	456,231	184,066	72.5	1.6	0.6	424,977	10,484,854	24.7	191,880	1,983,938	4.7
100歳以上	87,438	36,458	76.7	1.5	0.5	79,102	1,887,199	23.9	34,284	288,791	3.7

平均とは、入院/外来患者数1人あたりを意味する
時間外とは、時間外加算、休日加算、深夜加算、時間外特例、夜間・早朝等加算を算定された患者数/日数を意味する

表2 急性心筋梗塞でステント留置した6～18ヶ月後の患者数および平均の入院日数・再入院回数・外来日数

	患者数	1患者あたり平均		
		入院日数	再入院回数	外来日数
0～4歳	**	**	**	**
5～9歳	**	**	**	**
10～14歳	**	**	**	**
15～19歳	**	**	**	**
20～24歳	**	**	**	**
25～29歳	**	**	**	**
30～34歳	10	2.9	0.8	13.9
35～39歳	75	2.8	0.8	11.7
40～44歳	264	3.6	0.9	14.5
45～49歳	480	4.7	0.9	12.8
50～54歳	663	4.8	0.9	15.1
55～59歳	861	5.0	0.9	16.0
60～64歳	1,210	7.4	1.0	16.8
65～69歳	1,913	8.7	1.1	20.9
70～74歳	1,682	10.6	1.2	25.4
75～79歳	1,521	12.1	1.2	28.9
80～84歳	1,360	15.5	1.1	29.6
85～89歳	811	17.7	1.0	30.4
90～94歳	284	24.1	0.8	25.2
95～99歳	60	24.6	0.7	19.3
100歳以上	**	**	**	**

**：患者数10人未満のためマスキング

でピークになった。一方、外来日数は85～89歳でピークになり、表1に示した全外来患者日数より年齢階級のピークが遅れていた。表1の全患者数の平均外来日数と比較して、表2の75歳未満の平均外来日数は増加したのに対し、75歳以上の平均外来日数は減少した。

4. 考察

本研究では1患者1データ化がほぼ完了した。これによりレセプト単位の横断研究ができるNDBから、患者単位の追跡調査・コホート調査を可能とするデータベースに変化した。

日本の外来患者数の悉皆調査は初めてで、妥当性評価が困難であるが、既存のサンプリング調査としては患者調査が挙げられる。患者調査は層化無作為抽出した全国の医療施設に対し、ある1日を対象とし、外来受診または入院中の患者の性別、出生年月日、患者の住所、入院年月日などを調査し、医療の実態を把握するために2年に1度実施されている。2014年患者調査⁹⁾で示した年齢階級別の外来受療率では80～84歳でピークを迎えている。表1では1患者あたりの外来受診日数は80～84歳でピークになることを示した。受療率の分母は医療機関を受診していない人も含めた推計人口なのに対し、本研究の分母は2015年4月から1年間の医療機関を外来受診した全患者である。そのため算出方法に相違はあるが、基本的に本研究結果を支持する内容であった。

表2ではステント留置実施して、6ヶ月経過後1年間の入院・外来レセプトを抽出した。術後6カ月以内に死亡した場合も含め、6ヶ月～18ヶ月後に何らかの理由で医療機関にかからなかった患者は対象から外れている。そのため、この患者数はK549-00 経皮的冠動脈ステント留置術(急性心筋梗塞)／診療行為コード150375210を算定した患者全員ではない。30歳～74歳までの平均外来日数は、表1に示した全患者平均を上回っているが、75歳以上の平均外来日数は全患者平均を下回ることになった。その原因としては、外来で見れず入院となっているため外来が過小評価されている可能性が考えられる。

本研究では、1患者1データ化によって上記算出が可能となったことを示したものである。著者らが有しているNDBデータは3年分のため、対象期間を1.5年、追跡期間を1.5年とした。より長期間のNDBデータで解析するとより粒度の小さい結果を示すことができると考えられる。また、Kaplan-Meier法などでより詳細の解析をすることで、入院日数や再入院回数

の分布、例えば6ヶ月後～9か月後の再入院率が高い、や同じ再入院回数でも60代と80代では再入院する時期が異なる、などの結果を示すことができるかもしれない。

本研究の限界として、①1入院1データでは、1入院の定義上、退院したが同日中または翌日に再入院した場合でも1入院として処理される。また、連続して同一医療機関に1日入院した場合も1入院として処理される問題がある。加えて、短期滞在手術等基本料を入院期間として考慮できていない。DPC期間の1日ごとの点数を考慮おらず、入院時から1週間に要した点数の合計など、詳細の点数把握ができない。

限界で示した短期滞在手術等基本料および点数把握の問題は改善の余地が残されているものの、本研究によりNDBを用いて医療計画策定に関連するストラクチャー指標およびプロセス指標をほぼ把握できるようになった。残されたのはアウトカム指標であり、現在、死亡転帰未記載の死亡をレセプト上で特定する研究を進めている。

なお、本論文では触れないが、DPC期間内の詳細を除き、点数の把握も可能となったため、医療費分析に応用することは十分可能である。

5. 結論

本研究で行った1患者1データ化により、NDBで患者単位の追跡調査・コホート調査が可能となった。NDBの行政利用・研究利用の普及を進める一端となると考えられる。

今後、アウトカム指標として死亡を含む予後を検討する必要がある。

謝辞

本研究は平成28、29年度厚生労働科学研究費補助金(地域医療基盤開発推進研究事業)「病床機能の分化・連携や病床の効率的利用等ために必要となる実施可能な施策に関する研究」、平成28年度国立研究開発法人日本医療研究開発機構(AMED)地域横断的ICT活用医療推進研究事業「レセプト等の大規模電子診療情報を活用した薬剤疫学研究を含む医療パフォーマンス評価に関する研究」、平成30年度厚生労働科学研究費補助金(地域医療基盤開発推進研究事業)「地域の実情に応じた医療提供体制の構築を推進するための政策研究」、平成30年度文部科学研究費助成事業(科学研究費補助金)基盤研究(A)(一般)「データ科学・疫学・臨床医学の融合による日本の保険診療情報(NDB)の全解析」の一環として実施したものである。

参考文献

- 1) 野田龍也、久保慎一郎、明神大也、西岡祐一、東野恒之、松居宏樹、他. レセプト情報・特定健診等情報データベース(NDB)における患者突合(名寄せ)手法の改良と検証. 厚生指標. 64(12), 7-13, 2017-10.
- 2) Kubo S, Noda T, Myojin T, et al. National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB): outline and patient-matching technique. bioRxiv. [https://www.biorxiv.org/content/early/2018/04/02/280008.full.pdf. (cited 2018-Sep-1)].
- 3) 今村 知明. 地域横断的な医療介護情報のICT化により、世界最先端の臨床研究基盤等の構築を加速するための研究. AMED [http://www.amed.go.jp/content/files/jp/houkoku_h27/0501053/2015lk_chiikiICT_001.pdf (cited 2018-Sep-1)].
- 4) 特許公報. 電子レセプトデータ変換プログラムおよび電子レセプト

データ変換システム. 特開 2011-118538.

- 5) 明神大也、野田龍也、久保慎一郎、西岡祐一、東野恒之、今村知明. レセプト情報・特定健診等情報データベース(NDB)利用促進に向けた取り組みー1入院1データ化ー 第37回日本医療情報学連合大会論文集. 2017.
- 6) 厚生労働省. 平成26年患者調査. [https://www.mhlw.go.jp/toukei/saikin/hw/kanja/14/ (cited 2018-Sep-1)].