

特集 医療データベース活用の現状——ビッグデータを臨床研究に活かす

医療データベースの現状と課題

4) 民間の健康保険データベース

株式会社日本医療データセンター

木村 真也

診療(調剤)報酬明細書(レセプト)は、医療機関または薬局が、患者ごとに外来・入院〔診断群分類包括評価(Diagnosis Procedure Combination: DPC)と医科入院〕・調剤別に月締めで発行し、すべてのレセプトが審査機関を通じて保険者に集積されてくる。2008年度からは、特定健康診査・特定保健指導が制度化されたこともあり、保険者にはレセプトに加えて特定健康診査データや特定保健指導データも集まる。レセプトと健診データを追跡観察可能なレセプト・健診データベース(Database: DB) (図1)にすると、医療の状況や患者受療行動等が把握できる。

レセプトは情報の宝庫であるが、現在の電子レセプトデータは、あくまで診療報酬請求用であり、分析することを目的としていないために、集計解析可能な状態にするためにはかなりのデータ整備を要する。本稿では、本特集趣旨である「主要な医療データベースの整備と活用の現状を提示し、それらを臨床研究に最大限に活用するため、今後を展望する。」に資するために、紙レセプト時代からレセプトDBの開発と普及にかかわった10年以上の経験から、健康保険組合をソースとする民間のレセプト・健診DB¹⁾について述べたい。

民間のレセプト・健診データベースの整備状況

1 観察可能期間と母集団

ひとりの患者は、複数の医療機関・薬局に行き、

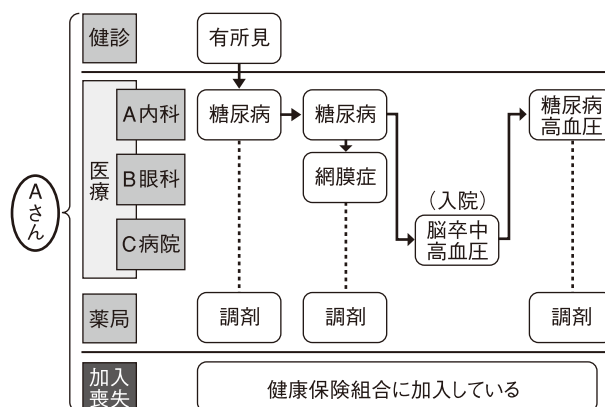


図1 健診時有所見からA内科受診→B眼科受診→C病院で入院し、その後A内科で治療継続していることが、観察できるデータベースイメージ

複数の傷病名をもち、複数の医薬品を処方されることが多い(表1)が、複数のレセプトが同一個人に突合されることにより縦覧観察可能なデータとなる。

日本の代表的な、ある民間のレセプトDBは、2005年1月診療分からDB化され、段階的に増加し、直近では約300万人母集団(累積470万人)に拡大している。長期の観察が必要な場合でも、約22万人は10年近く追跡することができる。約90万人では5年、約180万人では3年、1年間であれば約390万人が観察可能である(図2)。

2 健康診断データ

2008年度から始まった特定健康診査・特定保健指

表 1 患者あたりの傷病数, 薬剤数, 医療機関数, 薬局数

患者あたり傷病数	医科入院	6.39
	DPC	3.73
	医科入院外	8.5
患者あたり薬剤数		11.52
患者あたり受診医療機関数		2.68
患者あたり薬局数		1.88

日本医療データセンターのレセプトデータベース

母集団：実患者 2,386,306 人

(2014 年 4 月～2015 年 3 月)

(データ協力：日本医療データセンター)

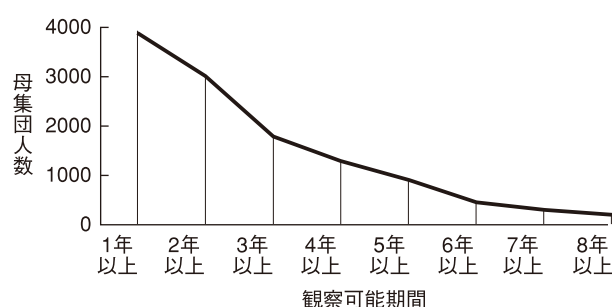


図 2 観察可能年数別の母集団×年齢階層 (データ協力：日本医療データセンター) (単位 1000 人)

導では, 40 歳以上 74 歳までの健診が義務化された。健保組合の平成 25 年度特定健診受診率調査²⁾によると, 受診率は 70.69% である。当該の民間 DB では, 2005～2014 年度で約 470 万件の健診データ (一般健診含む) が管理され, 5 年連続して健診データがある人は 19 万人を超える (表 2)。

3 母集団情報の整備

レセプトは医療機関等を受診したものにのみ発行される。したがって, 各集団における率を知るためには, 分母となる母集団の情報をできるだけリアルタイムに把握しておく必要がある。健康保険加入者の保険加入年月と離脱年月が DB 化されることにより, 期間出現率や罹患率がわかる。

4 分析用マスタの整備

レセプトデータには医薬品の薬効分類や成分名などの分析用情報はない。薬剤であれば, ATC 薬効分類・成分名・剤形など, 傷病は ICD10 コード, 診療

表 2 連続して観察できる年数別の健診受診者実数 (データ協力：日本医療データセンター)

(単位 1000 人)

2014 年度	2 年連続	3 年連続	4 年連続	5 年連続
980	804	389	293	192
6 年連続	7 年連続	8 年連続	9 年連続	10 年連続
117	83	44	8	1

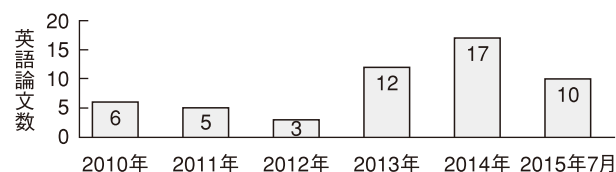


図 3 民間のレセプトデータベースを利用した英語論文数 (データ協力：日本医療データセンター)

行為では診療報酬区分コード (例: A001) などが, レセプトデータに付加されることで分析定義が明確になり, 研究作業効率も格段に良くなる。

5 データベースの英語化

レセプト DB 研究による英語論文は, 2015 年末時点で 50 報を超えている (図 3)。研究者の利便性を増すために, 傷病名や医薬品名など基本的な項目は英語化され, 適宜メンテナンスされている。

6 データ生成の信頼性保証

もともとのレセプトデータの中には, 医療機関における入力エラーデータが入っている。また, データを処理する際に間違いを引き起こすことも想定される。DB 研究の結果が信頼に足るものにするために, 元データと DB の関係は説明できることが求められ, データが適切かつ正確に作成されていることを監視する運用体制が構築されてきた。DB 品質管理の整備は, DB 研究のエビデンス能を引き上げるのみならず, 市販後調査等での活用可能性にもつながる。

レセプトデータベースの限界

1 年齢構成

データソースが社会保険の場合, 65 歳以上が少なく, 75 歳以上は欠損する (表 3)。全数レセプトを集めた厚生労働省のナショナル・データベース (NDB) では全年齢のレセプトが入っているが, デー

表 3 年齢層別実患者数と構成比

性別	年齢階級	年間	
		実患者数	構成比
男性	0～9 歳	199,904	8.4%
	10～19 歳	172,515	7.2%
	20～29 歳	176,613	7.4%
	30～39 歳	209,431	8.8%
	40～49 歳	233,338	9.8%
	50～59 歳	184,123	7.7%
	60～69 歳	90,654	3.8%
	70 歳以上	11,637	0.5%
女性	0～9 歳	188,089	7.9%
	10～19 歳	162,923	6.8%
	20～29 歳	133,591	5.6%
	30～39 歳	190,060	8.0%
	40～49 歳	212,254	8.9%
	50～59 歳	141,078	5.9%
	60～69 歳	67,156	2.8%
	70 歳以上	12,940	0.5%
計		2,386,306	100%

日本医療データセンターのレセプトデータベース

1 年間の年齢階層別の実患者数

(2014 年 4 月～2015 年 3 月)

(データ協力：日本医療データセンター)

タの整備に課題を残す。民間 DB のデータ整備は進んでいるが、データアクセスの関係でどうしても社会保険中心になるため、年齢層に課題がある。

2 傷病名の正確性

診療報酬制度の特性から、保険病名が混在している。たとえば、胃潰瘍に適応のある胃薬を胃炎患者に処方しても、病名は胃炎ではなく、胃潰瘍と書かれることがある。保険病名は、薬剤起因と検査起因が多い。そのため、研究対象症例の定義に工夫をすることにより、対象とすべき症例をより正確に拾い上げる作業を行う。たとえば、レセプト DB から 2 型糖尿病合併症および大血管症の病期進展と医療費の関係を検討した分析³⁾では、糖尿病網膜症の場合、レセプト傷病欄に「糖尿病網膜症」、「2 型糖尿病性網膜症」、「2 型糖尿病・眼合併症あり」のいずれか

があり、診療行為に「網膜光凝固術」や「硝子体切除術」などがあった場合に「増殖網膜症期」と定義している。

3 傷病と診療内容、医療費の関係

レセプトデータには、傷病と診療行為・薬剤を紐つける情報はない。そのため、複数の傷病名が記載されているレセプトでは、医療費（診療行為、医薬品）を傷病別に分解できないので、マクロ的に分析する論理的な按分法が試みられている⁴⁾。

4 レセプトでは把握できない、あるいは把握困難な情報

①臨床検査値がないため、患者の重症度や病態を定義付けることが困難である。

②レセプト転記欄にある死亡情報の完全性には限界がある。人口動態統計年報（死亡）⁵⁾によれば、2010 年の病院または診療所での死亡は全体の 79.3%である。しかしながら、健保組合のレセプトデータを使用した研究では、レセプトから把握できる死亡は 60%との報告がある⁶⁾。

③レセプトには承認されていても、保険償還されない薬剤やワクチンなどは出現しない。

レセプト・健診データベースの活用

民間のレセプト・健診 DB の特徴を整理すると、以下の 3 点が挙げられる。

①追跡性：複数の受診を時系列で観察可能である。

②網羅性：同一保険者に加入しているかぎり、すべてのレセプトが集約される（全数調査）。

③母集団（分母）が把握できる：存在情報があるため、罹患率・有病率などの出現率がわかる。

日本では患者数情報が意外と少ないが、ポピュレーションベースにおける有病割合や罹患率が推計できるのは利点といえる。診療実態の把握もできる。たとえば、診療プロセス、合併症と併発疾患、処方実態（重複、投与量、併用）、検査頻度等の添付文書順守率やガイドライン浸透度（エビデンス診療ギャップ）、イベント発生率、有害事象、医療費、などである。また、複数の医療機関にまたがる受診（治療）継続状況がわかり、病院データではわからなかった未受診状況がわかるのは初めてである。

当該 DB を利用した最近の DB 研究をいくつか紹介する。

1) 関節リウマチ

追跡性が活かされた研究として、関節リウマチ (RA) 臨床疫学データベース構築に関する研究を紹介する。レセプト DB を使用して、2011 年 6 月～2012 年 5 月の期間内に存在確認がとれ、同期間中に RA の診断名 (ICD-10 分類: M05, M060, M062, M063, M068, M069) が 2 回以上、2 カ月以上の間隔をおいて付与された 18 歳以上の 2762 人を RA 患者として解析している。非 RA 対照者として、RA 患者に対し、年齢 (± 5 歳)、性別、合併症の調査期間でマッチングした 27620 名をランダムに選択。虚血性心疾患、脳梗塞、高血圧性疾患、高脂血症、糖尿病、骨粗鬆症の有病率は、RA 群で有意に高かったとし、レセプト DB を用いた RA の合併症研究結果は、これまで全国規模の系統的なデータが存在しなかったわが国の RA 患者の合併症の実態を明らかにして、心血管系疾患などの合併症の有病率を非 RA 対照者と比較検討したアジアで初めての報告であるとしている⁷⁾。

2) 乾癬

母集団情報を活用し、乾癬患者の実態 (患者数、治療内容など) を明らかにする目的で、レセプト DB を使用した例では、2011 年 7 月から 1 年間の約 100 万人母集団を対象に調査している。観察期間中に乾癬と診断された患者数と母集団数から罹患率を年齢・性別で算出し、算出された罹患率を年齢・性別日本人口に乗じて推定患者数 (推定患者数 = 罹患率 $\times$ 日本人口) を算出した。この調査結果では、日本国内で乾癬と診断されている患者は約 56 万人と推定された。また、乾癬全体の罹患率は 0.44% と算出された。考察で、この調査には皮膚科以外の医師も入っており、正しい乾癬の診断がなされたのかという点とレセプトに記載された病名をもとにした推計であることを考慮すると、実際の罹患率はやや少なくなると推察している。しかしながら、日本乾癬学会の報告の 0.1% に比べると高い値であり、この差は、学会の集計が大学病院と基幹病院を受診した患者をもとにしていることと、初診した患者のみを登録していること、さらに今回の調査では約 3 割の患者が皮膚科以外の診療科で診断・治療を受けていることを挙げている⁸⁾。

3) 緑内障

網羅性と追跡性を活かした分析として、緑内障の局所薬剤治療の継続率を検討した研究がある。観察期間中に緑内障と診断された患者 2799 人 (age 47.3 $\pm$ 13.9 歳) の薬剤治療後 3 カ月での継続率は 73.2% で、同様に 6 カ月 68.1%, 12 カ月 60.9%, 3 年では 52.5% であった⁹⁾。

4) そのほか

そのほかにも、統合失調症患者におけるアドヒアランス (medication possession ratio: MPR) を研究した事例では、MPR 不良の再入院率が高いことが報告され¹⁰⁾、また、薬剤の安全性研究では、日本で承認を受けている非定型抗精神病薬 (9 剤) の服用における副作用である脂質異常症の発生を sequence symmetry analysis (SSA) を用いて解析した結果、olanzapine だけが脂質異常症のリスクを高めるとしている¹¹⁾。

さらに、健康診断データを組み合わせた事例では、健康診断で新たに糖尿病と認められた 1393 名の 62% ($N=864$) は治療を開始しておらず、そのなかの 49.2% ($N=425$) は血糖コントロール不良 ($HbA1c \geq 7\%$) であり、糖尿病治療を受けた 38% ($N=529$) と比較した結果、健診で糖尿病を指摘されたら速やかに診療を受け、さらに継続することが血糖コントロールを改善し、臨床的意義があると報告している¹²⁾。

展 望

民間レセプト・健診 DB の量と質の成長とともに、多様な研究に利活用されてきた。現在では、DB を利活用した業務実装へ向けて検討が進んでいる。たとえば、医薬品や医療機器などの費用対効果評価の導入が始まろうとしている。費用対効果評価の費用データはレセプト DB の果たす役割は大きい。

医薬品安全性分野での利活用は社会的価値が大きい、たとえば、医薬品リスク管理計画 (RMP) においては、患者数、complications, comorbidity, 処方実態等を俯瞰的に知ることがリスク最小化策を検討するうえで有用である。さらに、市販後調査における安全性にもデータベースを活用していこうとする動きが世界的に進みつつある。診療の質を評価し、打ち手につなげる PDCA (plan-do-check-act) サイ

クルも発展可能だ。

ICT の発展, 特にウェアラブル・デバイスなどにみられるセンサー類の進化と DB の相乗効果も注目される。たとえば, 手首型血圧計で得られる血圧の日内変動データとレセプト DB を突合することにより, 血圧日内変動の意義の解析検討可能にする。

一方で, レセプト DB の利活用を今以上に進めるためには, データ品質の検証, 保証が求められる。バリデーション研究はさらに積極的に行われる必要がある。

文 献

- 1) Kimura S, Sato T, Ikeda S, Noda M, Nakayama T. Development of a database of health insurance claims: standardization of disease classifications and anonymous record linkage. *J Epidemiol* 2010; 20: 413-9.
- 2) 健康保険組合連合会. 平成 25 年度 特定健診・特定保健指導の実施状況に関する調査分析
http://www.kenporen.com/study/toukei_data/pdf/chosa_h27_4.pdf
- 3) 北里博仁, 池田俊也, 泉和生ほか. 診療報酬明細書 (レセプト) データベースから 2 型糖尿病合併症および大血管症の病期進展と医療費の関係を明らかにするための患者情報を抽出する方法の検討. *日本糖尿病情報学会論文誌* 2010; 9 (肥満と糖尿病別冊 9): 48-64.
- 4) Okamoto E, Hata E. A New Method of Magnitude Estimation for PDM (Proportional Distribution Method) Using an Optimization Technique and Validation by Monte Carlo Simulation. *医療と社会* 2005; 14 (4): 117-26.
- 5) 人口動態統計年報 (死亡) 第 5 表死亡の場所別にみた死亡数・構成割合の年次推移.
<http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/suii10/dl/s03.pdf>
- 6) Ooba N, Setoguchi S, Ando T, et al. Claims-Based Definition of Death in Japanese Claims Database: Validity and Implications. *PLoS one* 2013; 8 (5): 1-6, e66116.
- 7) 厚生労働科学研究費補助金 難治性疾患等克服研究事業 (免疫アレルギー疾患予防・治療研究事業 免疫アレルギー研究分野). 平成 23 年度～平成 25 年度 総合研究報告書 関節リウマチ臨床疫学データベース構築に関する研究 (針谷正祥)
- 8) 照井正, 中川秀己, 江藤隆史ほか. 健康保険組合レセプト情報を利用した乾癬の実態調査. *臨床医薬* 2014; 30: 279-85.
- 9) Kashiwagi K, Furuya T. Persistence with topical glaucoma therapy among newly diagnosed Japanese patients. *Jpn J Ophthalmol* 2014; 58: 68-74.
- 10) Kuwabara H, Saito Y, Mahlich J. Adherence and rehospitalizations in patients with schizophrenia: evidence from Japanese claims data. *Neuropsychiatr Dis Treat* 2015; 11: 935-40.
- 11) Takeuchi Y, Kajimaya K, Ishiguro C, et al. Atypical Antipsychotics and the Risk of Hyperlipidemia: A Sequence Symmetry Analysis. *Drug Saf* 2015; 38: 641-50.
- 12) Heianza Y, Suzuki A, Fujihara K, et al. Impact on short-term glycaemic control of initiating diabetes care versus leaving diabetes untreated among individuals with newly screening-detected diabetes in Japan. *J Epidemiol Community Health* 2014; 68: 1189-95.