

分担研究名 NDB データ研究の実践と人材育成プログラム開発・検証

研究分担者 中山 健夫 京都大学医学研究科 健康情報学分野
研究協力者 酒井 未知 同上
研究協力者 加藤 源太 京都大学医学部附属病院診療報酬センター
研究協力者 岩尾友秀 京都大学医学部附属病院 先端医療研究開発機構
クリニカルトライアルサイエンス部
研究協力者 大寺 祥佑 京都大学医学部附属病院医療情報企画部

要旨

【目的】 NDB サンプルングデータセットを使用した研究を通して、NDB データ研究を実践する人材育成に必要な知識、スキルについて検討した。

【方法】 平成 24 年、25 年、26 年、各年 10 月の、65 歳以上の高齢者の医科入院、DPC レセプトデータを用い、終末期高齢者の診療内容を解析した研究を通し、NDB データを活用する人材に要する知識、スキルを検討し、人材育成の教材の素案を作成した。

【結果】 NDB データを活用する人材には、1)保険診療の仕組み、2)レセプトの記録ルール、3)データベースに関する知識、スキルが必要と考えられた。

【結論】 NDB データを活用できる人材育成プログラムは発展途上にある。今後、データ利用者による利用経験を通して、大規模レセプトデータの活用に必要な知識、スキルを体系化し、NDB データの活用を推進することが課題である。

A. 目的

NDB サンプルングデータセットを使用した研究を通して、NDB データ研究を実践する人材に必要な知識、スキルを検討する。

B. 研究方法

NDB データを用いて終末期高齢者の診療内容を解析した研究を通し、NDB を活用する人材に必要な知識、スキルを検討し、人材育成の教材の素案を作成した。

本研究には、平成 24 年 10 月、25 年 10 月、26 年 10 月、の NDB サンプルングデータを用いた。分析対象は、年齢が 65 歳以上、入院転帰に「死亡」が記録された高齢者の医科入院レセプト、DPC レセプトデータとした。研究目的は、終末期高齢者の診療内容の経年変化を記述することとした。「終末期」の定義は、入院レセプトの転

帰が「死亡」と記録される前の 1 ヶ月間とした。

NDB の第三者提供では、列名が定められていない csv 形式でデータが提供される。本研究では、列名、データ型等が記載されたサンプルングデータレコードフォーマット(厚生労働省より提供)に基づき、データベースマネジメントシステムのソフトウェア PostgreSQL を用いてデータベースを構築した。

研究デザインは repeated cross-sectional study とした。解析用データとして、1 患者 1 行のデータセットを作成した。医科入院、DPC のサンプルングデータには患者個人の識別番号(ID1、ID2)が付与されていないため、レセプトの ID(通番 2)を用いた。解析に用いる変数として、年齢区分、性別、入院レセプトの病棟区分、悪性新生物の傷病名の有無、死亡年度、死亡前 7 日間の生命維持治療、ならびに、集中治療室(ICU)入院

の有無を取得した。医科入院レセプトには死亡日が記録されていないため、診療行為テーブル(SI)に記録された診療行為の実施日の最終の日付を死亡日とした。死亡日を起点として、死亡前7日間の心肺蘇生術、人工呼吸、中心静脈カテーテル挿入、ICU入院の有無を二値変数として取得した。

分析対象として抽出された死亡例の件数2,010,840件を、同年の死亡統計に報告された医療機関内での死亡数2,610,429件と比較し、その妥当性を検証した(1)。調整変数を患者背景、アウトカム変数を死亡前7日間の生命維持治療、集中治療室(ICU)入院として、診療内容の経年変化を、コクランアーミテージ検定、ロジスティック回帰で解析した。

C. 研究結果

NDB サンプリングデータセットを用いることで、終末期高齢者の診療実態の経年変化が明らかになった。平成24年から26年にかけて、死亡前7日間の心肺蘇生術(CPR)、人工呼吸(MV)、中心静脈栄養カテーテル挿入(CVC)、ICU入院の割合は、それぞれCPR:11.0%から8.3%、MV:13.1%から9.8%、CVC:10.6%から7.8%、ICU:9.1%から7.8%に減少していた(2)。

NDB データを利用する研究者には、1)保険診療の仕組み、2)レセプトの記録ルール、3)データベースに関する知識、スキルが必要と考えられた。利用者への教育が必要な事項と、人材育成の教材の素案を以下に記載する。

C.1 保険診療の仕組み

NDBには保険診療下で行われた診療で派生したレセプトの内、電子化されたレセプトのほぼ100%が格納されている。レセプトデータの成り立ちを理解した上でNDBを扱うため、診療報酬制度、DPC制度に関する基礎的な知識が必要である。

1) 診療報酬制度

診療報酬は、保険医療機関、薬局が保険者から受け取る診療の対価(診療報酬)である。診療報酬は公定価格であり、技術・サービスの価格、物の価格で構成される。診療報酬は、診療報酬点数表(医科診療報酬点数表、歯科診療報酬点数表、調剤報酬点数表、診断群分類点数表)に点数化され、定期的に改訂される。医療機関が保険者に保険診療の診療報酬を請求する際、診療報酬明細書(レセプト)が発行される。NDBには、審査支払機関が審査を終えたレセプトデータが格納されている。

2) DPC制度

NDBのDPCレセプトの構造を把握するための背景情報として、DPC制度に関する知識が必要である。

DPCの診療報酬は、DPC(診断群分類)毎に設定される包括評価部分と出来高評価部分の合計である。包括評価部分は、1日当たり点数(3段階の階段設定)に在院日数と医療機関ごとに設定された係数(医療機関別係数)を乗じて算出される。NDBのDPCレセプトを利用する際は、この請求ルールがデータ構造に反映されていることに留意する。DPC対象医療機関で保険診療を受けた患者では、特定入院期間(包括期間)内に入退院が完結した場合、DPCレセプトが発行される。入院期間が包括評価の対象期間を超えた患者では、総括対象DPCレセプトと総括対象医科入院レセプトが発行される。

C.2 レセプトの記録ルール

診療報酬請求を目的としたレセプトデータを二次利用して、研究目的に沿ったデータを取得するため、レセプトの記録ルールを理解し、データの特性、限界を踏まえた解析を行う必要がある。

1) レセプトの発行単位

レセプトは、入院、外来別、医療機関、薬局毎に、1ヶ月単位で発行される。NDBサンプリ

ングデータの医科入院、DPC レセプトには、患者単位の識別子が付与されていないため、レセプトの識別子(通番 2)を患者単位の解析に用いた。同一患者が転院した場合にそれを識別出来ないため、サンプリングデータを縦断研究に用いるには限界が大きかった。

2) レセプトデータの記録条件仕様

2)-1 レセプトのテーブル構造

NDB レセプトは、レセプトの記録条件仕様に沿ったデータである。記録条件仕様は、社会保険診療報酬支払基金のホームページに掲載されており、レセプトのテーブル構造、各テーブルに含まれるデータ項目の定義が記載されている(3)。レセプトデータの仕様は、医科(入院、外来)、DPC、歯科、調剤で異なる(4)(図1)。特に本研究では、診療行為の請求情報から、死亡前の診療内容の情報を取得しており、医科入院レセプトの診療行為テーブル(SI)とDPCレセプトのコーディングレコード(CO)のデータ構造の違いに留意が必要であった。

医療機関情報 患者情報	医療機関情報 患者情報	医療機関情報 患者情報	医療機関情報 患者情報
傷病名	傷病名、書式	医療機関、保険医	診断群分類、傷病名
点数	明細欄	医薬品情報 調剤料 加算料	包括評価分 明細 出来高分 明細
合計点数	摘要欄	合計点数	合計点数
医療レセプト	歯科レセプト	調剤レセプト	DPCレセプト

図1. レセプトのデータ構造

2)-2 入院転帰

本研究では、終末期高齢者を特定するために、入院転帰のデータを用いた。レセプトの記録条件仕様上、転帰の記録は必須とされていないため、入院死亡転記情報の妥当性(感度、特異度)の検証を行った(5)。

2)-3 傷病名

医科レセプトの主傷病は、明確に定義されていない。医科レセプトの傷病名の記録数に制限はない。レセプトの傷病名マスタにない傷病名(対応する傷病名コードをもたない傷病名)が記録されることがある(図2)。本研究の対象とした高齢者は、多くの疾患を併存していると想定され、レセプトに記録された傷病名から病名を特定する限界は大きいと考えられた。本研究では、死亡月のがんの傷病名(疑い病名を除く)の有無を患者背景因子として取得するに留めた。

傷病名	①新生児特発性呼吸窮迫症候群(主) ②先天性骨形成不全症(主) ③慢性呼吸不全 ④皮膚炎の疑い	以下摘要欄	診療開始日	①平成18年1月2日 ②平成18年1月2日 ③平成18年1月2日 ④平成18年1月4日	転帰
-----	----------------------------------------------------------	-------	-------	------------------------------------------------------	----

図2. レセプトの傷病名の記録条件仕様

2)-4 入院日、退院日

レセプトに記録される入院日、退院日は、入院期間の特定に用いるに限界が大きい。医科入院レセプトには、入院日が記録されるが、疾患によっては、3ヶ月以内に再入院した場合、その再入院の入院日は記録されない場合がある。医科レセプトには、退院日は記録されない。医科入院からDPCレセプトに切り替わった場合、該当患者のDPCレセプトの入院日には、DPC算定開始、終了日が記録される(図3)。近年、終末期高齢者の医療費は重要な研究課題であるが、医療費の計算ルールは複雑であり(4)、診療行為単位(基本料、検査、手術単位)の医療費の解析にも限界が大きかった(図4)。

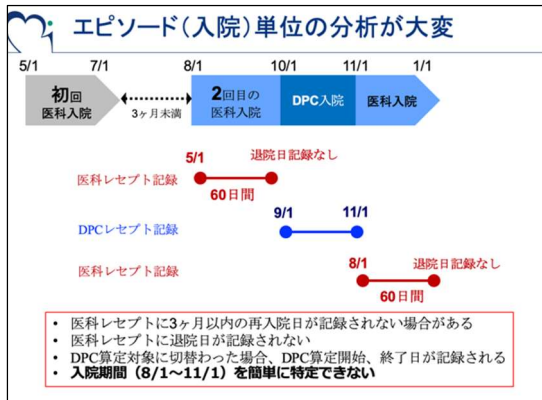


図3. レセプトの入院日の記録条件仕様

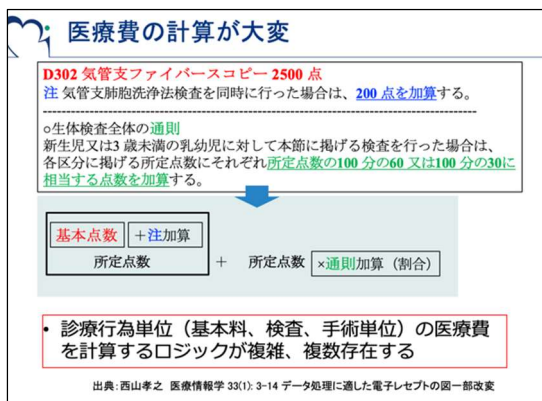


図4. 診療報酬の計算ロジック

C.3 データベース

NDBは、データサイズが多く、第三者提供のデータは、レセプトの記録条件仕様に沿って複数のテーブルに分割されている。ハードウェア、ソフトウェア環境にかけるコストは限られていることから、リレーショナル・データベース(図5)とそのマネジメントに関する知識が必要と考えられた。

本研究では、DELL Precision3510(CPU: Corei7 6820HQ 2.7GHz、メモリ 32GB、HDD 1TB)の端末を用いた。解析目的、データの運用を考慮し、データベースマネジメントシステムを用いてデータベースを構築した。リレーショナル・データベースは、大規模レセプトのデータ量を小さくして格納するに適している。しかし、解析処理の効率の点では限界がある。データベース

マネジメントシステムは、HDD(SSD)へのアクセス回数を最小に抑えることで、メモリに乗りきれない大規模なレセプトデータの高速度処理が可能である(図6、7)。データベースマネジメントシステムのソフトウェアには、PostgreSQLを採用した。PostgreSQLは、分析用のSQLであるOLAP関数に対応している、フリーのソフトウェアである。また、プライマリーキーの代わりとなるレコードIDに対応しているという利点がある。また、データの検索、テーブル間の結合を効率的に行うため、処理を行う必要があった。具体的には、レセプトの識別子(通番2)への索引(インデックス)の付与、データベースの統計情報の収集、更新(VAVUM ANALYSE)を行った。

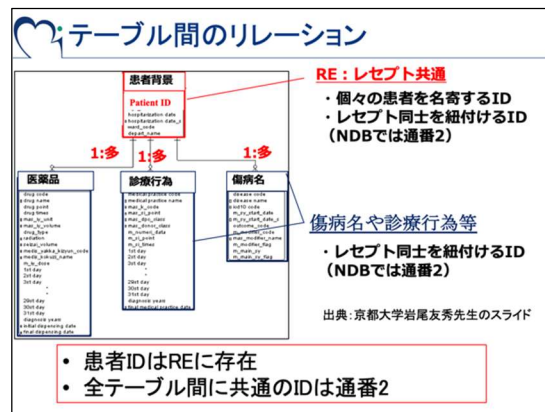


図5. リレーショナル・データベース

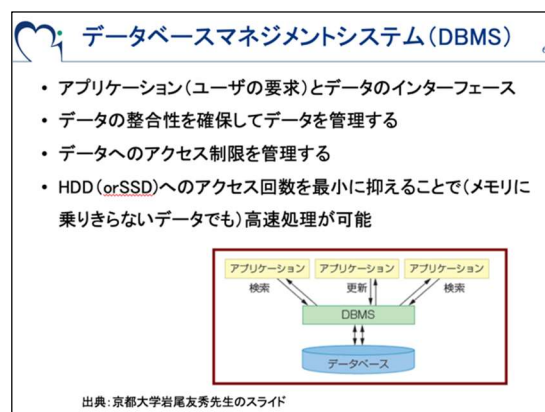


図6. データベースマネジメントシステム

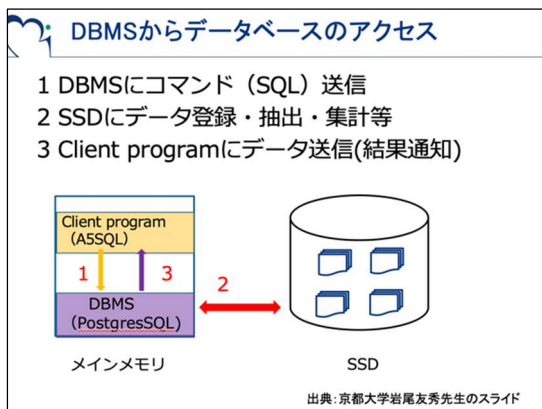


図7. データベースマネジメントシステムからデータベースへのアクセス

D. 考察 & E. 結論

NDB サンプリングデータセットを用いて、終末期高齢者の診療実態を記述する研究を通して、NDBを活用する人材に必要な知識、スキルを検討し、人材育成の教材の素案を作成した。今後、NDBと介護レセプト情報のデータベース等との連結により、データの活用可能性の向上が期待されている。しかし、NDBデータの活用に必要な知識、スキルに関する人材育成は発展途上にある。今後、データ利用者による利用経験を通して、大規模レセプトデータの活用に必要な知識、スキルを体系化し、NDBデータの活用を推進することが課題である。

引用文献

- (1) 酒井未知. 医療経済研究機構 2016 年度 (第 20 回)研究助成 研究要旨. 大規模レセプトデータベースを用いた高齢者終末期医療の実態解明. Monthly IHEP 2017-2018 年 12・1 月号
- (2) Sakai, M.; Ohtera, S.; Iwao, T.; Neff, Y.; Uchida, T.; Takahashi, Y.; Kato, G.; Kuroda, T.; Nishimura, S.; Nakayama, T.; BiDAME (Big Data Analysis of Medical care for the Elderly in Kyoto). Decreased Administration of Life-Sustaining Treatment just before Death among Older Inpatients in Japan: A Time-

Trend Analysis from 2012 through 2014 Based on a Nationally Representative Sample. Int. J. Environ. Res. Public Health 2021, 18, 3135.

- (3) 社会保険診療報酬支払基金. 電子レセプトの作成
https://www.ssk.or.jp/seikyushiharai/rezept/iryokikan/iryokikan_02.html
- (4) 西山孝之. データ処理に適した電子レセプト. 医療情報学 33(1): 3-14
- (5) Sakai M, Ohtera S, Iwao T, Neff Y, Kato G, Takahashi Y, Nakayama T; BiDAME (Big Data Analysis of Medical care for the Elderly in Kyoto). Validation of claims data to identify death among aged persons utilizing enrollment data from health insurance unions. Environ Health Prev Med. 2019 Nov 23;24(1):63.

F. 健康危機情報

なし

G. 研究発表

- (1) Sakai M, Ohtera S, Iwao T, Neff Y, Uchida T, Takahashi Y, Kato G, Kuroda T, Nishimura S, Nakayama T, BiDAME (Big Data Analysis of Medical care for the Elderly in Kyoto). Decreased Administration of Life-Sustaining Treatment just before Death among Older Inpatients in Japan: A Time-Trend Analysis from 2012 through 2014 Based on a Nationally Representative Sample. Int J Environ Res Public Health 2021;18:3135.

H. 知的所有権の取得状況

なし

