

令和 2 年 6 月 9 日現在

機関番号：20101

研究種目：基盤研究(C)（一般）

研究期間：2016～2019

課題番号：16K08910

研究課題名（和文）大規模レセプトデータベースを用いた疫学研究結果の妥当性検証手法の確立

研究課題名（英文）Development of a method for validating epidemiological research results using a large-scale claim database.

研究代表者

樋之津 史郎（Hinotsu, Shiro）

札幌医科大学・医学部・教授

研究者番号：80323567

交付決定額（研究期間全体）：（直接経費） 3,700,000円

研究成果の概要（和文）：レセプトデータをJMDC社から取得した。2005年1月から2017年7月の間に「前立腺がん」と診断された5,767人であった。医療機関数は合計17,001施設で、そのうち85.2%は診療所であった。医薬品処方合計1,426,115件で、前立腺がん治療に用いられるリュープリン4,288件、カソデックス錠4,220件、ゾラデックス3,155件、オダイン錠1,028件であった。検査の総数は5,225,046件で、クレアチニン91,672件、AST85,736件、ALT85,523件、PSA7,877件であった。このデータで腎機能、肝機能の有害事象のリスクを推定することができると考えられた。

研究成果の学術的意義や社会的意義

診療報酬明細書として作成される「レセプト」を研究用に提供する企業から、前立腺癌を対象に大規模で長期間のデータを取得する事ができた。このデータには、医薬品を処方したこと、採血検査を実施したことは記載されているが、たとえば検査値や症状などは含まれていない。そのため、治療のための薬剤を処方した後に、肝機能障害や腎機能障害が発生したかどうかは直接知ることができない。今回行った研究により、検査実施時期を個人ごとにまとめ、その頻度が変化しているかどうかを検討する事により、有害事象の発生リスクを推定することが可能であることが明らかになった。

研究成果の概要（英文）：The claim database was obtained from JMDC. Between January 2005 and July 2017, there were 5,767 patients diagnosed prostate cancer. The total number of hospitals was 17,001, of which 14,487 (85.2%) were clinics with 0-19 beds. The total number of medical prescriptions was 1,426,115. Among them, the drugs used for the treatment of prostate cancer, 4,288 cases of Leuplin, 4,220 cases of Casodex tablets (6,680 generic products), 3,155 cases of Zoradex, and 1,028 cases of Odyne tablets (424 generic products). The total number of laboratory tests was 5,225,046, with Creatinine 91,672, AST 85,736, ALT 85,523, and PSA 7,877. It was considered that this large-data could be used to estimate the risk of adverse events occurring in renal and liver functions.

研究分野：臨床疫学

キーワード：前立腺癌 レセプトデータベース 観察研究

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

診療報酬明細書(通称レセプト)は、医療機関が保険者(健康保険組合や市町村など)に請求する明細書である。近年、複数の保険者と契約してレセプトの情報を入手し、個人情報完全に削除して研究利用のために販売する企業がある。大学に所属する教員は、臨床研究(臨床疫学研究)を行うために、株式会社日本医療データセンター(JMDC)と契約して同社の研究用レセプトデータベースを利用することができる。

JMDCの作成したデータベースは、診断名や処置等の記載を標準化しデータマネジメントを行って表記の揺れなどを統一しているため、質の高い研究用データと考えられる。このデータベースがカバーしている対象は日本国内で作成されるレセプトの一部ではあるが、全国複数の健康保険組合のデータを取得しており最大160万人のデータが使用可能である(2020年4月時点での累積母集団数は730万人)。近年、JMDCを含む複数の提供者が作成したデータベースを用いた、いわゆる「リアルワールドデータ」の解析が多く行われている。

しかしながら、JMDCデータベースに代表されるレセプトを情報源としたデータベースにはレセプトに記載されるデータが病名、処方、処置および施設の情報であり、患者の症状や検査データの値など診療録に記載されている情報は取得できないという限界がある。

### 2. 研究の目的

株式会社日本医療データセンター(JMDC)と契約して取得したデータを、臨床疫学的な研究に用いるデータセットに変換することが可能であるかを検証する。JMDCのデータベースは最大160万人、10年間以上の情報を扱うことができる。疾患を特定して抽出したデータであっても、100万レコードを超えるテキストファイルを扱うことになるため、研究者が一般的に用いるPC環境で取り扱うことが可能かどうかを明らかにすることも目的の一つである。

そこで得られる情報が臨床疫学的研究に用いることができた段階で、大規模なレセプトデータであるビッグデータを用いた分析を行う。その結果を検証する方法として、データのバラツキや単純な分布を解析してデータの妥当性を明らかにする。その結果により病院の診療録データの集計結果と比較検討することが可能であるかを明らかにする。

### 3. 研究の方法

#### (1) データ保存のためのシステム構築

大規模なデータを保存し、研究のためのデータセットを作成するために必要なシステム構成を選択する。また、保存されたデータを抽出し、並べ替えるなどの操作をする際に必要なコンピュータを用いて大規模レセプトデータベースを安定して運用できるシステムを構築する。

#### (2) レセプトデータベースを用いた臨床疫学的研究

処方動向や検査頻度の変化による有害事象リスクなどを解析項目としてレセプトデータベースを用いたデータ抽出を行う。データ解析に用いるソフトウェアも複数(JMP ProとStata)を用いて、いずれのソフトウェアでもデータ抽出が可能であることを確認する。

#### (3) 病院の診療情報データを用いた解析結果との比較可能性の検討

レセプトデータベースを用いた研究で推定される有害事象のリスクが病院の診療情報を用いたデータと比較可能であるか検討する。レセプトデータベースを用いた研究結果から得られる項目を明確にする。

#### 4 . 研究成果

本研究で用いるレセプトデータベースを JMDC から取得した。2005 年 1 月から 2017 年 7 月の間に「前立腺がん」と診断された 5,767 人であった。この間の JMDC データベースに含まれる男女含めた母集団は 4,197,722 人であった。

JMDC のデータベースのデータベースは患者、レセプト、傷病、医薬品、材料、診療行為および施設にわかれてテキストファイルとして提供された。提供された CSV ファイルと JMP に取り込んだ際のデータファイルと Stata15 のデータファイルの大きさを表に示す。

-----  
表：ファイルの大きさ (KB)

	<u>CSV ファイル</u>	<u>JMP ファイル</u>	<u>Stata ファイル</u>
レセプト	45,714	45,517	49,557
医薬品	878,270	809,949	2,176,803
患者	452	382	257
材料	14,572	13,804	19,769
施設	1,306	1,548	1,320
傷病	486,920	469,658	1,092,796
診療行為	1,054,529	1,012,748	2,724,793

-----

これらの結果により、比較的ファイルサイズの大きなデータセットであるが、JMP Pro および Stata で読み込んで各ファイルをデータセットとして解析を行うことが可能であることが明らかになった。

そこで、JMP Pro を用いて分析を行った。

対象患者に発行されたレセプトは合計 512,510 枚で、1 患者あたり月に 1 枚発行されることから、本研究における観察人月がレセプト枚数と同数であると考えられた。患者毎のレセプト数は中央値 70 枚で平均値 89.8 枚であった。

年齢の中央値は 62 歳 (最小値 36 歳、最大値 75 歳)、平均値 61.1 歳 (標準偏差 7.48 歳) であった。今回契約した JMDC のデータベースには後期高齢者を含まないため年齢の最大値は 75 歳であった。

診療していた医療機関は合計 17,001 施設で、そのうち 14,487 施設 (85.2%) は病床数 0-19 の診療所であった。病床数 20-99 の病院が 621 施設 (3.7%)、100-199 症の病因が 716 施設 (4.2%) であった。更に規模の大きな施設は合計 1,177 施設であった。これらのことから、前立腺がん患者の大多数は病床数 0-19 の診療所でケアされていることが明らかになった。

同一レセプト内に複数記載のある「医薬品」について集計すると、合計 1,426,115 件であった。この中で、前立腺がん治療に用いられる薬剤としては、リュープリン 4,288 件、カゾデックス錠 4,220 件、ゾラデックス 3,155 件、オダイン錠 1,028 件、ピカルタミド (カゾデックスの後発品) 31 種類合計 6,680 件、フルタミド (オダインの後発品) 2 種類合計 424 件であった。比較的新しいゴナックス 1354 件、イクスタンジは 894 件、ザイティガ 480 件であった。

検査処置の総数は 5,225,046 件で、クレアチニン 91,672 件、BUN 87,414 件、AST 85,736 件、ALT 85,523 件であった。また前立腺癌の腫瘍マーカーである PSA 検査は 7,877 件であった。こ

のことから、前立腺癌治療中の腎機能、肝機能のチェックを頻繁に行っていることが明らかになった。

これらのデータが解析出来ることが明らかになったことから、レセプトデータベースの解析結果をもとに、診療録に記載された有害事象のリスクと比較可能なデータセットを作成する事が可能であることが明らかになった。今後札幌医科大学泌尿器科で診療された前立腺癌患者の診療録データを取得し、比較検討するための基礎データをまとめることができた。

今回得られたデータは、データベースソフトウェアである MySQL および Microsoft Access で保存することができた。また、JMP Pro および Stata15 を用いて解析用データセットにすることが問題なくできた。今後、大規模レセプトデータベースを用いた解析がこれらのソフトウェアを用いる事で可能であることがわかった。

最後に、レセプトデータベースの信頼性を確認する意味で、今回得られたデータの性別について分析した結果を記載する。前立腺は男性にのみ存在する臓器であるため、前立腺癌は男性に認められる疾患である。今回得られた 5,709 例のうち 11 例の性別が「女性」であった。診断名のリストを見ると、1 例は図 1 のように 1 回のみ「前立腺癌」の診断が記載され、以後「乳腺腫瘍」（乳癌）の診断であった。これは、医療機関における診断名の記載ミスが疑われる。別の症例では図 2 のような診断名が記載されており、一般に前立腺癌の患者に前立腺肥大症が合併することは多く認められることから、こちらは医療機関における性別の記載ミスが疑われる。いずれも単純な記載ミスと考えられ、エラー率も 0.19%と低い事から、このことがレセプトデータベースの信頼性を失うことにはならず、解析の段階で除外することが妥当であると考えられた。大規模データベースを扱う際には、個別に目視で確認作業をする事は困難であるため、単純な集計や分布の確認作業が重要であると考えられた。

以上より、本研究により大規模レセプトデータベースを保管するシステムの構築を行い、得られたデータを分析可能なデータセットにする事ができ、臨床疫学的研究を実施する事ができた。また、得られた結果より、今後実施される診療録を用いた観察研究で得られる結果と比較検討する項目の提案をすることができた。

図 1

匿名化した ID	性別	診断日	診断名
M00017547	女性	2009/09/01	前立腺癌
M00017547	女性	2012/07/01	乳腺腫瘍
M00017547	女性	2016/05/01	乳腺腫瘍
M00017547	女性	2016/05/01	乳腺症

図 2

匿名化した ID	性別	診断日	診断名
M01385480	女性	2009/06/01	前立腺癌
M01385480	女性	2009/06/01	前立腺肥大症
M01385480	女性	2009/04/01	前立腺癌
M01385480	女性	2009/05/01	前立腺癌
M01385480	女性	2009/05/01	前立腺肥大症

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Tatebe Y, Koyama T, Mikami N, Kitamura Y, Sendo T, Hinotsu S	4. 巻 26
2. 論文標題 Hypoglycemia Associated With Pivalate-Conjugated Antibiotics in Young Children: A Retrospective Study Using a Medical and Pharmacy Claims Database in Japan	5. 発行年 2020年
3. 雑誌名 J Infect Chemother .	6. 最初と最後の頁 86-91
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jiac.2019.07.013.	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ohshima A, Koyama T, Ogawa A, Zamami Y, Tanaka HY, Kitamura Y, Sendo T, Hinotsu S, Miller MW, Kano MR.	4. 巻 36
2. 論文標題 Oral anticoagulants usage in Japanese patients aged 18-74 years with non-valvular atrial fibrillation: a retrospective analysis based on insurance claims data	5. 発行年 2019年
3. 雑誌名 Fam Pract	6. 最初と最後の頁 685-692
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/fampra/cmz016	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 大島礼子、小山 敏広、小川 愛子、座間味 義人、樋之津 史郎、狩野 光伸
2. 発表標題 医療ビッグデータに基づく直接作用型経口抗凝固薬のワルファリンに対する有効性・安全性評価に関する臨床疫学研究
3. 学会等名 日本薬学会中四国支部学術大会
4. 発表年 2016年

1. 発表者名 Shiro Hinotsu
2. 発表標題 Experience of Data Convert into CDISC Fromat from Meta-analysis Dataset
3. 学会等名 CDISC Japan Inter change 2019
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	狩野 光伸  (Kano Mitsunobu)  (80447383)	岡山大学・ヘルスシステム統合科学研究科・教授    (15301)	