

厚生労働行政推進調査事業費  
(政策科学総合研究事業)

「患者調査等、各種基幹統計調査における NDB データの利用可能性に関する評価」

「NDB における悉皆性のメリットと機微情報の保護を両立させる方法についての研究」  
分担研究報告書

研究分担者 平木 秀輔

京都大学医学部附属病院 医療情報企画部 助教

研究要旨

NDB に含まれるデータは、その種類によっては希少であること等によって患者個人の特定の可能性が残るなど機微性の高いものもある。そこで研究者が NDB を利用する場合、それらのデータは削除して提供を受けることが通例であるが、そのことは NDB の悉皆性のメリットを損なうことになり得る。

本年度においては、特定性の排除という社会からの要請を踏まえつつ、分析を工夫することによって有益な情報が抽出できるのではないかと問いに取り組んだ。具体的には、薬機法施行規則に規定される「希少疾病用医薬品」の認定水準である年間発症数 5 万人を下回るとある疾患 A について NDB を用いて年間発症数を推定することで、希少疾患の情報が削除されている NDB といえど、希少疾病用医薬品のニーズ調査に有益であることを示した。また、一定の診療行為は医療機関の規模に依存して算定されることに着目し、レセプトを分類することで医療機関の規模ごとの診療実態を明らかにすることができた。

これらにより、NDB の新たな利用方法を提案できると考え、現在は学会発表等の準備をしているところである。

A. 研究目的

レセプト情報・特定検診等情報データベース (NDB) は、いわゆるビッグデータ・リアルワールドデータとしてその二次利用によるエビデンスの創出が期待されている。

反面、国民の健康に関する情報を格納したデータベースとして、その取り扱いには細心の注意が求められる。特に、個人や医療機関の特定につながりうるような機微情報は望ましくない差別や批判につながりかねず、厳重に管理すべきものと考えられ、実際に厚生労働省から研究者に対して提供されるデータにおいては上記のような情報

が削除されるなどの適切な処理が行われている。

本年度においては、研究者に公表されるデータの下でどのような分析が可能であるかについて探索的に検討した。上述のデータ処理は研究者の“手足を縛る”ために行われるものではなく、学術的な興味と社会の求めるプライバシーに配慮した、社会的コンセンサスの産物である。したがって、その範囲の中でどのような知見を抽出することができるか検討することは今後の NDB 研究の可能性と限界を探るにあたって重要と考え、研究を実施した。

## B. 研究方法と結果

### <テーマ A>

NDB サンプルングデータセットを用いて希少疾病の有病者数を推定することができるか？

※下記に示す方法と結果は、共同研究者らとともに2019年7月に開催される第36回日本眼循環学会に発表予定である。そのため、詳細な方法と結果を示すことは差し控え、それらの概要を示す。

#### 【方法】

NDB サンプルングデータセットを用い、2011-2015年・1,4,7,10月の単月分の医科入院外・医科入院およびDPCレセプトを用い、SYファイル（傷病名レコード）から我々の設定したある疾患Aに対応するコードを持つレセプトを抽出し、サンプルング割合で割り戻すことによって年間新規発症患者数を推定した。

#### 【結果・考察】

疾患Aの年間発症患者数は2万人程度と推定され、これは諸外国における既報の有病率と類似した結果であった。共同研究者が本疾患に対する新規薬剤の開発を進めており、希少疾病用医薬品の指定を受けられる可能性を本邦の悉皆データを用いて裏打ちすることができたと考えられた。

### <テーマ B>

NDB サンプルングデータセットを用いて、医療機関の規模別の診療実態を推定することができるか？

※下記に示す方法と結果は、共同研究者らとともに2019年6月に開催される第62回日本腎臓学会学術総会にて発表予定である。そのため、テーマAと同様に方法と結果に

ついては概要を示すにとどめる。

#### 【方法】

2013年10月分のサンプルングデータセット（入院外）におけるSIファイルより、病院の特性を反映する診療報酬請求項目を用いて診療所から請求されたもの・小規模病院から請求されたもの・大規模病院から請求されたものにレセプトを分類した。そのうえで、当該レセプトのIYファイルを用いて降圧薬の処方数をカウントし、病床規模別の降圧薬処方剤数を検討した。また、疾患特異的薬剤の処方有無を検討することで糖尿病あるいは脂質異常症の有無による層別化をしたうえで左記の医療機関規模別処方剤数を検討することを検討している。

#### 【結果・考察】

全体としては、1/100サンプルングで数万人規模の患者が抽出された。これは、既報における本邦の高血圧患者数と合致した結果と考えられる。

降圧薬の処方剤数については、医療機関の規模が大きくなるにしたがって増える傾向にあった。このことより、難治性高血圧が大規模施設に集約されがちであることが示唆された。なお、糖尿病や脂質異常症の有無についての層別化は現在も検討を続けているところである。

## C. 本年度のまとめと考察

本年度は、NDB サンプルングデータセットを用いて、機微情報の保護に配慮されたデータベースの中からどのように有益な情報を抽出できるかという課題に取り組んだ。

個人特定につながる可能性を考慮し、サンプルングデータセットは提供の段階で希少な疾病の情報は削除されている。とはい

うものの、テーマ A を通じて医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律施行規則 251 条に規定される「希少疾病用医薬品」の対象患者数である 5 万人程度の患者数の疾患であれば、サンプリングデータセットを用いて患者数を推定することが可能であることが示唆された。このことは、希少疾患に対する医薬品・機器等の開発にあたり、比較的に入手しやすい NDB であるサンプリングデータセットを用いてニーズ調査が可能になることを示唆している。

また、既存の研究において、規模や教育機能の有無という医療機関の特性が医療の質に影響を与える可能性が示されている。本邦においても他のデータセットを用いた同様の研究は散見されるものの、診療所まで悉皆に含んだデータセットはこれまで得られなかった。

とはいえ NDB において医療機関を特定できる情報を秘匿せずに研究者に公開した場合、医療機関の経営情報を実質的に公開してしまうことになりかねず、少なくともサンプリングデータセットの管理レベルで公表することは困難である。しかしながら、テーマ B を通じて、適切な診療報酬項目を選択することによってレセプトを医療機関の規模別に分類することができることが示唆された。個別医療機関の情報が得られずとも研究者側の創意工夫で情報を抽出できることの実例であるともいえ、今後の研究の知見として有用であると考えられる。

#### D. 結論

研究者に提供される NDB データにおいては、機微情報を含むデータは適切に処理

され、公表されないことがある。そのような制約の中でも、適切な問題設定と分析の工夫により有用な情報が得ることができることが示唆された。

#### E. 健康危険情報

なし

#### G. 研究発表

##### 1. 論文発表

1. Hiragi S, Tamura H, Goto R, Kuroda T. The effect of model selection on cost-effectiveness research: a comparison of kidney function-based microsimulation and disease grade-based microsimulation in chronic kidney disease modeling. *BMC medical informatics and decision making*, 2018, 18.1: 94.
2. Kensuke Morris, Goshiro Yamamoto, Shusuke Hiragi, Shosuke Ohtera, Michi Sakai, Osamu Sugiyama, Kazuya Okamoto, Masayuki Nambu, Tomohiro Kuroda. Designing an Authorization System Based on Patient Privacy Preferences in Japan. *Studies in health technology and informatics*, 2018, 247: 71-75.
3. Helou S, Yamamoto G, Kondoh E, Tamura H, Hiragi S, Sugiyama O, Okamoto K, Nambu M, Kuroda T. Understanding the Roles of EMR Systems in Japanese Antenatal Care Settings. *Studies in health technology and informatics*, 2018, 251: 257-260.
4. Sato M, Kondoh E, Iwao T, Hiragi S, Okamoto K, Tamura H, Mogami H, Chigusa Y, Kuroda T, Mandai M, Konishi I,

Kato G. Nationwide survey of severe postpartum haemorrhage in Japan: an exploratory study using the national database of health insurance claims. *The Journal of Maternal-Fetal & Neonatal Medicine*, 2018, 1-6.

## 2. 学会発表

1. 加藤源太, 田村寛, 平木秀輔, 大寺祥佑, 佐藤大介, 奥村泰之, 酒井未知, 明神大也, 西岡祐一, 久保慎一郎, 野田龍也. 患者調査における NDB データの利用可能性に関する評価:患者一元化および傷病名特定アルゴリズムの観点から. 医療情報学連合大会プログラム・抄録集, 38th, 377 (2018/11/22) 福岡.
2. 藤田健一郎, 杉山治, 平木秀輔, 岡本和也, 竹村匡正, 黒田知宏. 電子カルテ記載量予測モデル創出を目指した電子カルテ年間記載文字数の分析. 医療情報学連合大会プログラム・抄

録集, 38th, 246 (2018/11/22) 福岡.

3. 畠中純, 平木秀輔, 杉山治, 山本豪志郎, 佐々木博史, 岡本和也, 南部雅幸, 黒田知宏: 市場原理に基づくベッドコントロールの妥当性の検証. 日本医療・病院管理学会誌, vol.55, Suppl., p.95 (2018/10/27) 郡山.
4. 市岡光洋, 平木秀輔, 坂井薫, 松原雄, 横井秀基, 柳田素子. 当院の急性持続血液浄化療法における透析離脱に影響を与える因子の検討. 日本腎臓学会誌, 60, 6, 765 (2018/08/20) 新潟.. 当院の急性持続血液浄化療法における透析離脱に影響を与える因子の検討. 日本腎臓学会誌, 60, 6, 765 (2018/08/20) 新潟.

## H. 知的財産権の出願・登録状況

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし