

令和4年度 厚生労働行政推進調査事業費補助金
(医薬品・医療機器等レギュラトリーサイエンス政策研究事業)
分担研究報告書

医療情報データベースの活用推進に関する研究

研究分担者 鈴木 隆弘・国立大学法人千葉大学医学部附属病院 企画情報部 准教授

研究要旨：

医療情報データベースに基づく医薬品の安全性等を評価する場合には、データベースに含まれる情報を組み合わせて対象とする有害事象（アウトカム）を定義する必要があり、効率的なアウトカム定義の作成手法の検討及び実用化可能なアウトカム定義を確立するため、本研究を実施した。

「悪性腫瘍」「心不全」の検討において得られた結果より、当院におけるPPV及び感度、患者背景因子別のPPVについて考察をし報告書に纏めた。また他機関とのPPVの差異の要因を当院において検討し、定義に条件を加えることでさらに良好な結果が得られる可能性が考えられた。

A. 研究目的

製造販売後の医薬品安全性評価は、従来、副作用報告、使用成績調査等の結果に基づくことが主であったが、医療情報データベースの整備等によりリアルワールドでの大規模データに基づく評価が可能となりつつある。

MID-NET[®]は、厚生労働省の医療情報データベース基盤整備事業により構築されたデータベースで、独立行政法人医薬品医療機器総合機構（以下「PMDA」という。）による運用管理の下、平成30年度から運用を開始している。MID-NET[®]は、診療情報明細書（レセプト）、診断群分類（DPC）及び電子カルテの情報が利用可能で、現在では行政による活用の他、臨床研究や製造販売後データベース調査にも活用されており、製造販売後の医薬品安全対策の質の向上が期待されている。

医療情報データベースに基づく医薬品の安全性等の評価では、データベースに含まれる情報から対象とする有害事象（アウトカム）を適切に特定するために、信頼できるアウトカム定義を作成する必要があるが、本邦においてアウトカム定義のバリデーションが実施された例はほとんどない。「MID-NET[®]データの特性解析及びデータ抽出条件・解析手法等に関する研究」（日本医療研究開発機構 医薬品等規制調和・評価研究事業）（以下「先行研究」という。）において、MID-NET[®]を対象にアウトカム定義の作成及びその妥当性の評価を効率的に実施するための検討を実施し、基礎的な検討手法を確立するとともに、バリデーションされた複数のアウトカム定義が作成された。しかしながら、実用化可能なアウトカム定義を増やすためには、更なる検討手法の効率化や具体的なアウトカム定義の

作成を継続的に実施する必要がある。

本研究は、先行研究の成果を踏まえて、研究の流れを見直し All possible cases の定義を決定した上で以降の検討を行う等の検討手法の改善及び実用化可能なアウトカム定義の確立を目指し、医薬品安全性評価における医療情報データベースの活用促進と、より科学的な根拠に基づく安全対策の実現に繋げることを目的とした。

B. 研究方法

安全対策上の必要性や重要性を考慮し、検討の対象とするアウトカムを複数選定した上で、各アウトカムについて、従来法又は機械学習の手法を取り入れて作成したアウトカム定義について、複数医療機関にて妥当性の評価を行った。

研究の流れは以下のとおりである。

- 1) 複数拠点で検討対象とする All possible cases の定義及びアウトカム定義の検討
- 2) 評価基準の作成
- 3) 対象アウトカムについてカルテレビューにより真の症例を特定
- 4) 機械学習及び従来法によるアウトカム定義の作成
- 5) 各アウトカム定義について、陽性的中度（以下「PPV」という。）及び感度の算出並びに評価
- 6) 複数医療機関の PPV の比較及び医療機関間の差異の要因検討

（倫理面への配慮）

医療機関が行う作業は、文部科学省・厚生労働省.人を対象とする医学系研究に関する倫理指針（平成26年文部科学省・厚生労働省告示第3号）を遵守して行った。

C. 研究結果

「悪性腫瘍」についての結果を以下に示す。

全定義の中、包括定義である 1、7 及び 11 の PPV はそれぞれ 45.7%、67.7% 及び 46.1% で、感度は 79.6%、64.3% 及び 79.6% となった。包括定義の中で最も PPV が高かった定義 7 は、DPC データ傷病情報の「医療資源を最も投入した傷病名」に悪性腫瘍に関連する病名が記録されている条件のものであった。それぞれの包括定義を更に条件を絞り込んでいくことで、PPV は大きく改善した。全定義の中で PPV が 82.8%、82.9% と高かったのはそれぞれ定義 6、定義 16 であり、どちらも機械学習に基づく改良型アウトカム定義に関連するものであり、NPV、特異度共に高く良い結果となった。

次に「心不全」についての結果を以下に示す。

全定義の中で PPV が最も高かったのは、87.5% の定義 18 であり、病名を DPC 傷病情報（必須 3 項目のいずれか）に絞り、BNP 高値、利尿剤の使用が条件のものであった。しかし感度は 16.7% と低かった。比較的 Possible case 数を得ることができ PPV が 71.4%、75.7%、77.1% と高かったのは、定義 7、定義 12、定義 17 であった。定義 17 は、定義 18 の病名の条件に SS-MIX 傷病情報（疑いを含まない）を加えたことにより Possible case 数が確保でき、感度も 16.7% から 64.3% と改善された。

D. 考察

「悪性腫瘍」についての考察を以下に示す。

定義 1 及び 11 の PPV が 45~46% と高くなかったのは、SS-MIX2 傷病情報からの抽出の際に疑い病名は含まないように設定はしてあるが、最終的な結果として悪性ではなかったものが含まれていた事が考えられる。また全体的にその他のケースにはがん登録のデータ加工で除外した、腫瘍の性状コードが 0（良性）または 1（良性・悪性の別不詳）であったものや、転移性又は再発性悪性腫瘍のものが含まれていたと考えられた。がん登録は初発が対象だが、当院では初発以外の治療困難な患者が多いことがその他のケースが多くなった理由の一つと考えられる。定義 6、16 において PPV が高い結果となったのは、DPC 患者情報の「がんの初発」に該当するものが条件であるためと考えられる。

次に「心不全」についての考察を以下に示す。

定義 18 の結果より、当院の特徴として外来症例にも真のケースが多く含まれていたということがいえる。真のケース A 又は B の症例のうち、index date 時点で外来であった症例は 14 症例であり、index date 前後 1 週間以内に入院のあった症例を除くと 8 症例が外来でフォローされている症例であった。その全てが判定 1 のフラミンガム診断基準に該当し利尿薬などの服用もあるため真のケース A と判定されており、心不全ではあるもののす

ぐに入院が必要な状況ではなくフォロー中である症例であった。定義 12 に関しては、診療行為として心エコー検査の実施を加えたことにより心機能の評価が必要な症例が集まり PPV を改善することができた。心不全の状態では心機能評価のために心エコーが必須であることを反映していると考えられる。その他のケースの判定理由として多いのは慢性心不全症例であった。当院においては他院からの紹介受診が多く、その多くは他院で心不全症状が落ち着いた頃に手術適応の確認で来院されるケースであり、次に手術後のフォローアップ例が多かった。BNP は有力な判断情報だが、その他のケースとなった中には腎不全による BNP 高値の方が多く含まれていた。また慢性心不全の場合、安定していても軽度高値のケースが多く見られた。BNP だけでは判断することは難しいが、医薬品、診療行為など複合的に条件を組み合わせることで前回の検討時よりも全体的に向上が見られた。

E. 結論

「悪性腫瘍」について感度、PPV が共に高い数値を示した定義は、機械学習による改良型アウトカム定義によるものであった。機械学習の活用の有益性が示された。

今回の研究においてはがん登録のデータを初めて用いて判定を行った。がん登録は定義に沿って収集し、ある程度の時間を置いてから作成され整えられたデータであり、日々の臨床に伴い蓄積されていくリアルワールドデータとは性質が異なることが改めて示されたが、これらを組み合わせることで活用の幅が広がるのではないかと考えられた。「心不全」について当院では、外来に真のケースが多く含まれていたが各種の条件を組み合わせることで改善されることが期待される。

F. 健康危険情報：

（分担研究報告書では記入不要です）

G. 研究発表：

各機関における成果をご記入ください
（発表誌名巻号・頁・発行年等も記入）

1. 論文発表

2. 学会発表

診療データの多次元時系列の構造化を用いたデータベースの作成-脳出血患者データへの適応-石井晃, 齊藤敦子, 鈴木英夫, 他:
第 42 回医療情報学連合大会 e-poster.

H. 知的財産権の出願・登録状況

（予定を含む。）

各機関における状況をご記入ください

1. 特許取得

2. 実用新案登録

3. その他