

医療レセプトデータの傷病名情報に対するvalidation studyの実施

研究代表者 福田治久 九州大学大学院医学研究院・准教授

研究分担者 石黒智恵子 国立健康危機管理研究機構・室長

研究要旨

本研究では、NDB と予防接種台帳データベースの統合によるワクチン有効性・安全性・経済性評価システムの構築を目指し、NDB における病名正確性と電子カルテ情報の有用性を検証した。医療レセプトデータと検査値データを連結した CLEAR Study を立ち上げ、ロタウイルス感染症やネフローゼ症候群についてレセプト情報に基づくアルゴリズムの妥当性を評価するためのデータ収集を行った。また、自治体データと医療機関の電子カルテ情報を連携させ、有害事象例のカルテ情報の検証を行った。その結果、レセプトデータの信頼性を補完するバリデーション研究の必要性が確認され、CLEAR Study の設立が今後のワクチン評価に資する可能性が示された、今後の課題は対象疾患の拡大や多施設データでの検証であるが、本研究を通じてワクチン評価に不可欠な検証研究基盤が構築された。

研究協力者

前田恵 九州大学大学院医学研究院・特任助教

A. 研究目的

我が国では現在、ワクチンデータベース（VDB）が開発中で、ナショナルデータベース（NDB）と統合することで、ワクチンの有効性・安全性・経済性を検証可能なシステム構築が目指されている。本システムを確実に稼働させるためには、VDBとNDBの特徴を踏まえ、何が実施でき、どこにバイアスが入りうるかを把握することが必要である。特にワクチンの評価にあたっては、(1) NDBにおける病名の正確性、(2) VDB・NDBで使用可能なデータに限定して解析した場合のバイアスの程度、(3) 統合が進んでいない電子カルテ等情報を利用することの必要性、を明らかにすることが不可欠である。一方、現時点では、VDBは未稼働で、VDBとNDBは未統合であることから、既に自治体基盤で予防接種台帳、

医療レセプト、HER-SYS等を統合させた実績を有するVENUS Studyにおいて検証することが有用である。研究代表者は既に13自治体128万人分のデータを統合し、多数のワクチン評価論文を報告しており、最も検証可能な立ち位置にある。

そこで本研究では、上記の(1) NDBにおける病名の正確性、(2) VDB・NDBで使用可能なデータに限定して解析した場合のバイアスの程度、(3) 統合が進んでいない電子カルテ等情報を利用することの必要性、を明らかにすることも目的に実施するものである。そのうち、本課題では、(1) NDBにおける病名の正確性評価に向けたvalidation studyと、(3) 統合が進んでいない電子カルテ等情報を利用することの必要性に関する研究を行った。

B. 研究方法

1. NDBにおける病名の正確性評価に向けたvalidation study

我々は医療レセプトデータの validation study を実施するための研究基盤を開発し、この基盤を CLEAR Study (Claims data Learning & Enhancing for Algorithm Refinement Study) と命名した。CLEAR Study ではデータ取得先を日本の保険医療機関と定め、収集データは保険医療機関が支払機関に対して保険請求するために毎月作成している医療レセプトデータを対象にしている。Validation study において必須となる gold standard は当該保険医療機関が有する検査値データを主な対象として、必要に応じてカルテ情報や院内がん登録情報などを取得している。CLEAR Study では、医療レセプトデータと検査値データ等を患者単位でリンケージすることで、医療レセプトデータの validation study を実施した。

(1) データ収集

CLEAR Study の最初のマイルストーンは 10 医療機関の協力体制を構築することである。既に 9 病院から研究協力を得て、そのうち 8 病院からはデータ収集を完了した。

CLEAR Study において収集するデータは医療レセプトデータと検査値データ（および必要に応じてカルテデータや院内がん登録データを対象に加えることができる）である。医療レセプトデータは、医療機関が毎月作成している RECEIPT_C ファイルおよび RECEIPT_D ファイルとした。すなわち、外来診療における医科ファイル、入院診療における医科ファイルおよび DPC ファイルが対象となる。さらに、外来診療における調剤情報を把握するために、DPC データの外来 EF ファイルを対象にした。CLEAR Study では、協力医療機関の承諾の下、保有している医療レセプトデータを最大 2008 年 4 月分まで遡り収集している。2008 年 3 月以前は医療レセプトデータの項目が異なっているため収集対象外とした。本研究では医療機関から提供可能な期間の全疾患領域・全患者分の医療レセプトデータを収集した。一方、検査値データは、validation study の対象疾患に応じて、必要な検査値データの提供を協力

医療機関にその都度、依頼してデータ収集した。2024 年度では、RSV 感染症と腸重積を対象に実施したが、2025 年度ではロタウイルス感染症およびネフローゼ症候群を対象に定めた。ロタウイルス感染症では下記の検査項目を対象にした：ロタウイルス抗原定性（糞便）、ロタウイルス抗原定量（糞便）。一方、ネフローゼ症候群では、下記の検査項目を対象にした：尿蛋白、クレアチニン試験紙法（尿）（蛋白／クレアチニン比）、クレアチニン（尿）、A1b（BCP 改良法・BCG 法）、TP。データ収集は協力医療機関に研究者が訪問し、医療レセプトデータや検査値データに含まれる個人識別情報（氏名、被保険者番号、カルテ番号など）をハッシュ値などに加工した仮名加工情報を作成した。仮名加工情報となった医療レセプトデータや検査値データを研究機関へと持ち帰った。

(2) データリンケージ

医療レセプトデータと検査値データのデータリンケージは、カルテ番号を使用した。医療レセプトデータにはカルテ番号の登録項目が含まれている。医療機関によってはカルテ番号に秘密キーを追加した値として登録されていることもあるが、秘密キーを確認することでカルテ番号に変換した。検査値データはデータ抽出時にカルテ番号を付与することを依頼した。

データ収集時に、医療レセプトデータおよび検査値データの双方のカルテ番号に対して、CLEAR Study が独自に設定するロジックを追加し、カルテ番号を元に研究用 ID を設定した。

(3) アルゴリズム開発

研究機関に持ち帰った医療レセプトデータや検査値データは研究用データへと成形し、validation study を行った。医療レセプトデータにおける傷病名、医薬品名、診療行為名の使用範囲および組み合わせにより、真の判定結果とした検査値で確認した傷病発生との一致状況を評価した。

2. NDBにおける病名の正確性評価に向けた validation study : 死亡アウトカム

日本の自治体が保有する2015年4月～2023年3月の医療・介護レセプトデータを用いてレセプトデータにおける死亡のアルゴリズムを定義し、日本の人口動態統計に基づく死亡票をリファレンススタンダードとして、陽性的中度 (PPV) , 陰性的中度 (NPV) , 感度, 特異度を年齢区分別 (0-39, 40-74, ≥ 75 years) に算出した。さらに、死亡票とリンケージできた症例を対象に、死亡場所別および死因別に感度を算出した。また、レセプト上で定義した死亡年月日と死亡票上の死亡年月日の一致性についてKappa scoresを算出した。

3. 統合が進んでいない電子カルテ等情報を利用することの必要性に関する研究

本研究では、医療機関における電子カルテ等情報を参照することで、ワクチン接種による有害事象発生の因果関係の評価にどの程度有用であるかの評価を行う。そのために、予防接種台帳と医療レセプトデータを連結解析が可能な自治体データにおいてワクチン接種後に有害事象が出現していた対象者を同定し、当該対象者が有害事象の治療のために受診した医療機関の協力を得て、当該医療機関における当該対象者のカルテ情報を検証する。

(倫理面への配慮)

本研究は、人を対象とする医学研究に関する倫理指針に従って実施し、九州大学倫理審査委員会より承認を得て実施した(承認番号:22305-03)。

C. 研究結果

1. NDBにおける病名の正確性評価に向けた validation study

ロタウイルス感染症の評価のために3医療機関からデータを取得し、追加で3医療機関もデータ取得中である。データは2026年度中にはデー

タベース化を完了させ、validation studyに着手する。また、ネフローゼ症候群の評価のために1医療機関からデータを取得し、追加で3医療機関もデータ取得中である。データは2026年度中にはデータベース化を完了させ、validation studyに着手する。

2. NDBにおける病名の正確性評価に向けた validation study : 死亡アウトカム

対象レセプトは、13自治体の対象者1,894,123人 (0-39 years 22.1%; 40-74 years 39.3%; ≥ 75 years 38.7%, 女性55.9%) の68,966,581人月であった。PPV, NPV, 特異度は、どの年齢区分においても高く、感度は年齢区分や死亡場所によって異なることが明らかとなった。病院や診療所での死亡については医療レセプトによって90%程度は把握可能であり、介護施設での死亡も医療レセプトに加えて介護レセプトを使用して定義することで95%以上は把握可能であることが明らかとなった。一方、在宅死亡については、39歳以下は低く年齢によって感度に差が認められたものの、感度を低下させている原因は医療レセプトでは把握が難しい不慮の事故等による死亡であり、がん等の在宅医療における死亡についてはレセプト上で把握できる可能性が示唆された。また、死亡年月日のKappa scoresは0.95以上と高い一致性が認められた。

3. 統合が進んでいない電子カルテ等情報を利用することの必要性に関する研究

予防接種台帳と医療レセプトデータを連結解析が可能な自治体において、1医療機関を対象にカルテレビューを実施した。最初に、予防接種台帳と医療レセプトデータの連結解析から肺炎球菌ワクチン接種後の有害事象の可能性のある症例として3例が同定された。症例A～症例Cの3例を対象にカルテレビューを実施した結果、症例Aでは、ワクチン接種に関する記載はなく、サムスカ導入目的で入院している経過が明らかになった。症例Bでは、ワクチン接種に関する記

載はなく、薬剤アレルギーは"-"との記載しか認められなかった。呼吸困難が発現し救急受診したとの記載から、ワクチン接種後の有害事象の可能性も考えられた。症例Cでは、ワクチン接種に関する記載はなく、膀胱炎の症状があり約2年前に処方された薬を自己判断で内服したところアナフィラキシーショックで受診したとの記載があることから、ワクチン接種によるものではなく、薬剤影響であることが示唆された。その他、肺炎球菌ワクチン接種後57-112日目に出現した有害事象例の25例についてもカルテレビューを行った。しかしながら、いずれも、ワクチン接種との関連性を示唆する記載は認められなかった。

D. 考 察

今後、NDBと予防接種台帳データベースをリンケージしてワクチンの有効性と安全性を検証する上で、感染症と有害事象に関する医療レセプトデータのvalidation studyは必須となる。医療レセプトデータは請求データであるため、その傷病名の信頼性は十分に保証されていないと考えられる。その意味で、本研究において、validation studyを実施可能な研究基盤であるCLEAR Studyを立ち上げることができたことは、今後のワクチン評価の上で、極めて重要な第一歩を進めることができたと考えられる。また、死亡票データと自治体保有レセプトデータを用いた死亡アウトカムのvalidation studyをも実施することができた。

また、医療レセプトデータからは、有害事象の可能性のある症例を判定できるものの、ワクチン接種との因果関係は不明である。この問題に対処するために、1医療機関に限定されるものの、ワクチン接種後の有害事象の可能性のある症例に対するカルテレビューを実施した。その結果、3例中2例は有害事象が他の理由が背景

にあることが確認された。1例は他の理由が不明瞭であり、ワクチン接種後の有害事象である可能性は残る結果であった。施設数・症例数は限定的であるため、今後は施設数を増加させた検証も必要になると考えられる。

E. 結 論

NDBと予防接種台帳データベースをリンケージしてワクチンの有効性と安全性を検証を進める上で不可欠なvalidation studyの研究基盤を用いて、ワクチン接種後の有害事象に関するvalidation研究の成果が生まれつつある。

F. 健康危険情報

該当なし

G. 研究発表

1. 論文発表

1. Fukuda H, Maeda M, Ishiguro C. The Claims data Learning & Enhancing for Algorithm Refinement (CLEAR) Study: Overview of the Study Design and Baseline Profile. Journal of Epidemiology 2026; 18: e70273.

2. 学会発表

該当なし

H. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし