

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、RS ウイルス感染症や腸重積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した。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に向けた環境整備を実施した。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 8 病院から研究協力を得て、そのうち 3 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。特に、検査値データにて鑑別診断が可能になる感染症領域の RSV 感染症と腸重積を対象に実施した。RSV 感染症では下記の検査項目を対象にした：SARS-CoV-2・RS ウイルス核酸同時検出（検査委託）、SARS-CoV-2・RS ウイルス核酸同時検出（検査委託以外）、SARS-CoV-2・インフルエンザ・RS 核酸同時検出（委託）、SARS-CoV-2・インフルエンザ・RS 核酸同時検出（委託外）、RS ウイルス抗体価（定性・半定量・定量）、RS ウイルス抗原定性、SARS-CoV-2・インフルエンザ・RS 抗原同時検出（定性）、SARS-CoV-2・RS ウイルス抗原同時検出（定性）、ウイルス・細菌核酸多項目同時検出（検査委託）、ウイルス・細菌核酸多項目同時検出（検査委託以外）、ウイルス・細菌核酸多項目同時検出。一方、腸重積では、下記の検査項目を対象にした：超音波検査（断層撮影法）（胸腹部）。

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. 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

RSV感染症の評価のために、1医療機関の2011年から2023年8月に少なくとも1回以上、保険診療で受診した1歳未満の患者を対象にした validation studyを実施した。アルゴリズムは、病名と検査を用いて、病名は複数のICDコードをのパターンを設定し、検査有無との組み合わせパターンを作成した。その結果、陽性的中度が最も高い (positive predictive value (PPV) = 93.4)

アルゴリズムは、病名 (J12.1/J20.5/J21.0) および検査 (RSウイルス抗体価・抗原定性検査およびウイルス・細菌核酸多項目同時検査) が出現している場合であった。一方、特異度が最も高い (特異度=98.9) アルゴリズムは、病名 (J12.1/J20.5/J21.0/B34.8のうち傷病名コードが0799008) および検査 (RSウイルス抗体価・抗原定性検査およびウイルス・細菌核酸多項目同時検査) が出現している場合であった。

また、腸重積の評価のために、1医療機関の2017年から2023年に少なくとも1回以上、保険診療で受診した患者を対象にした validation studyを実施した。アルゴリズムは、病名 (ICD-10コード) と検査と診療行為 (レセプト電算処理システムコード) との組み合わせパターンを作成した。その結果、陽性的中度が最も高い (PPV=44.4) アルゴリズムは、病名 (K56.1) および検査 (160072210) および処置 (150181010) が出現している場合であった。特異度が最も高い (特異度=99.9) アルゴリズムも病名 (K56.1) および検査 (160072210) および処置 (150181010, 150181110, 150389510) が出現している場合であった。

2. 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カルテレビューを用いた研究に1医療機関が研究協力可能になった。令和6年度においては倫理審査の承認までを得ることができた。また、当該自治体のレセプトから当該病院の有害事象発生者リストを抽出した。

D. 考 察

今後、NDBと予防接種台帳データベースをリンク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。本研究課題において研究者を対象に実施した予備的なアンケート調査からも、validation studyの必要性が求められていた。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。

既に、RSVや腸重積を対象に陽性的中度や特異度を高めるためのアルゴリズム開発といった成果も出ており、今後、さらに解析対象疾患を拡大していくことが求められる。ただし、これまでの評価は1施設のみを対象にしていたことから、今後は複数施設のデータを統合した検証が求められる。

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の研究基盤が開発された。

F. 健康危険情報

該当なし

G. 研究発表

1. 論文発表

該当なし

2. 学会発表

1. 前田恵, 吉野麻衣, 村田典子, 石黒智恵子, 福田治久. レセプトデータにおける RS ウイルス感染症定義の妥当性評価: CLEAR Study. 第 83 回日本公衆衛生学会総会. 2024 年 10 月 29 日-31 日. 札幌.
2. Ishiguro C, Maeda M, Fukuda H. Validation of claims-based algorithms for intussusception in Japan: The CLEAR Study. 2024 ACPE Annual Meeting. October 12-14, 2024. Tokyo, Japan.

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