

京都大学	博士（薬学）	氏名	山本 浩貴
論文題目	リアルワールドデータを用いた医薬品有害事象の早期検出に関する研究		

（論文内容の要旨）

医薬品有害事象（副作用）は死因の第4位と推計され、その発見が遅れた場合の健康被害は甚大である。承認前の医薬品の安全性に関する情報は限られた条件下で収集されたものであるため、副作用の把握には自発報告システム（SRS：Spontaneous Reporting System）を含めた市販後調査が最も重要であり、ファーマコビジランスでの必須のシステムである。しかし過少報告など既知の限界があるため、SRSを補完できる方法が早急に求められている。近年、大規模ヘルスケアデータを扱うための基盤が構築され、それらを活用した副作用研究への期待が高まっている。レセプトデータには副作用の疑いがあるか否かに関わらず患者の症状の変化や対処薬が登録される。そのため、レセプトに登録される病名を有害事象の発生と考えることで、広範な副作用シグナルを「能動的」に検出できるのではないかと仮説を立てた。そこで本研究では、第1章で広範な医薬品とその副作用を対象に、データマイニングの性能を評価するためのゴールドスタンダードを構築した。その上で第2章では大規模レセプトデータを用いて広範な副作用シグナルを早期に検出する方法を探索した。

### 第1章 実臨床データベースに基づくゴールドスタンダードの作成

データマイニングによる副作用検出法の性能を評価するためには、適切な参照基準（ゴールドスタンダード）が必要である。そこで米国食品医薬品局（FDA）が管理する世界最大の自発報告データベースFAERSと、医薬品投与日と有害事象発現日の詳細な情報に富む国内自発報告データベースJADERを用いて臨床エビデンスに基づくゴールドスタンダードの構築を試みた。各データベースに登録される医薬品と、有害事象グループSMQとの関係の強さについて、2種類の統計解析を網羅的に行うことで最終的に92種類の陽性対照と88種類の陰性対照からなるゴールドスタンダードを構築した。構築の流れとして、「薬とイベント間に因果関係が認められる」と考え得る陽性対照は、各SMQにつき両データベースで共通して強いシグナルを示す1薬剤を選択した。「薬とイベント間の関連性が極めて低い」と考え得る陰性対照は、陽性対照で用いた医薬品のうち、各SMQにつきJADERで被疑薬として1例も報告されず、かつFAERS解析において報告オッズ比 < 1となる医薬品からランダムに1剤選択した。これらの組合せは各医薬品の製品情報および文献検索によりその妥当性を評価した。以上、本検討で構築したゴールドスタンダードを用いることで、広範な医薬品とその副作用を対象にデータ解析手法の性能を評価することを可能にした。

### 第2章 レセプトの相関ルールマイニングによる副作用シグナルの早期検出

相関ルールマイニング（ARM：Association Rule Mining）は膨大なデータから共起確率が高いアイテムを効率的に見出す解析手法であり、近年SRSを用いた副作用研究でその有用性が報告されている。そこで本検討では日本の民間レセプトデータベース

JMDC Claims (2005年～2019年) に登録される医薬品と疾患名に対しARMを適用させ、第1章で構築したゴールドスタンダードを用いて副作用シグナル検出への有用性を評価した。その結果、モデルの性能を表すROC及びPR曲線のAUC値はそれぞれ0.80及び0.83となり、ARMはシグナル検出法として高い性能を示した。次にデータに含まれる各症例を保険組合への加入タイミングにより分割することで、10個のデータセットを準備し同様にARMを適用したところ、ROC-、PR-AUCともに全てのデータセットで0.70を超え、ARMの性能には高い再現性が認められた。次に早期検出のための解析デザインとして、1ヶ月、3ヶ月、6ヶ月と短期間のデータセットを準備し、シグナルを検出した陽性対照の数を既存手法であるSequence Symmetry Analysis (SSA) と比較した。その結果、ARMはどの解析期間においてもSSAより多くの副作用シグナルを検出した。最後に国内において安全性速報が出されたドロスピレノン・エチニルエストラジオールによる血栓症を例に、ARMが新薬の副作用シグナルを早期に検出するか検討を行った。その結果、販売39ヶ月後に発令された安全性速報より25ヶ月前の時点で血栓症のシグナルを検出できた。これらの結果より、レセプトデータにARMを適応させることで既存のファーマコビジランスを強化できる可能性が示された。

以上、申請者は、臨床エビデンスに基づくゴールドスタンダードを構築し、それを用いて副作用シグナルの早期検出にレセプトへのARMが有用であることを示した。本研究成果は、今後利活用が進む大規模臨床データを用いた、副作用予測システムや未知の副作用検出への基盤となることが期待される。

(論文審査の結果の要旨)

医薬品有害事象(副作用)は死因の第4位と推計され、その発見が遅れた場合の健康被害は甚大である。承認前の医薬品の安全性に関する情報は限られた条件下で収集されたものであるため、副作用の把握には自発報告システム(SRS: Spontaneous Reporting System)を含めた市販後調査が最も重要であり、ファーマコビジランスでの必須のシステムである。しかし過少報告など既知の限界があるため、SRSを補完できる方法が早急に求められている。近年、大規模ヘルスケアデータを扱うための基盤が構築され、それらを活用した副作用研究への期待が高まっている。レセプトデータには副作用の疑いがあるか否かに関わらず患者の症状の変化や対処薬が登録される。そのため、レセプトに登録される病名を有害事象の発生と考えることで、広範な副作用シグナルを「能動的」に検出できるのではないかと仮説を立てた。そこで本研究では、第1章で広範な医薬品とその副作用を対象に、データマイニングの性能を評価するためのゴールドスタンダードを構築した。その上で第2章では大規模レセプトデータを用いて広範な副作用シグナルを早期に検出する方法を探索した。

### 第1章 実臨床データベースに基づくゴールドスタンダードの作成

データマイニングによる副作用検出法の性能を評価するためには、適切な参照基準(ゴールドスタンダード)が必要である。そこで米国食品医薬品局(FDA)が管理する世界最大の自発報告データベースFAERSと、医薬品投与日と有害事象発現日の詳細な情報に富む国内自発報告データベースJADERを用いて臨床エビデンスに基づくゴールドスタンダードの構築を試みた。各データベースに登録される医薬品と、有害事象グループSMQとの関係の強さについて、2種類の統計解析を網羅的に行うことで最終的に92種類の陽性対照と88種類の陰性対照からなるゴールドスタンダードを構築した。構築の流れとして、「薬とイベント間に因果関係が認められる」と考え得る陽性対照は、各SMQにつき両データベースで共通して強いシグナルを示す1薬剤を選択した。「薬とイベント間の関連性が極めて低い」と考え得る陰性対照は、陽性対照で用いた医薬品のうち、各SMQにつきJADERで被疑薬として1例も報告されず、かつFAERS解析において報告オッズ比 $< 1$ となる医薬品からランダムに1剤選択した。これらの組合せは各医薬品の製品情報および文献検索によりその妥当性を評価した。以上、本検討で構築したゴールドスタンダードを用いることで、広範な医薬品とその副作用を対象にデータ解析手法の性能を評価することを可能にした。

### 第2章 レセプトの相関ルールマイニングによる副作用シグナルの早期検出

相関ルールマイニング(ARM: Association Rule Mining)は膨大なデータから共起確率が高いアイテムを効率的に見出す解析手法であり、近年SRSを用いた副作用研究でその有用性が報告されている。そこで本検討では日本の民間レセプトデータベースJMDC Claims(2005年~2019年)に登録される医薬品と疾患名に対しARMを適用させ、第1章で構築したゴールドスタンダードを用いて副作用シグナル検出への有用性を評価した。その結果、モデルの性能を表すROC及びPR曲線のAUC値はそれぞれ0.80及び0.83となり、ARMはシグナル検出法として高い性能を示した。次にデータに含まれる各症例を保険組合への加入タイミングにより分割することで、10個のデータセットを準備し同様にARMを適用したところ、ROC-、PR-AUCともに全てのデータセットで0.70を超え、ARMの性能には高い再現性が認められた。次に早期検出のための解析デザインとして、1ヶ月、3ヶ月、6ヶ月と短期間のデータセットを準備し、シグナルを検出した陽性対照の数を既存手法であるSequence Symmetry Analysis(SSA)と比較した。その結果、ARMはどの解析期間においてもSSAより多くの副作用シグナルを検出した。最後に国内において安全性速報が出されたドロスピレノン・エチニルエストラジオールによる血栓症を例に、ARMが新薬の副作用シグナルを早期に検出するか検討を行った。その結果、販売39ヶ月後に発令された安全性速報より25ヶ月前の時点で血栓症のシグナルを検出できた。これらの結果より、レセプトデータにARMを適応させることで既存のファーマコビジランスを強化できる可能性が示された。

以上、申請者は、臨床エビデンスに基づくゴールドスタンダードを構築し、それを用いて副作用シグナルの早期検出にレセプトへの ARM が有用であることを示した。本研究成果は、今後利活用が進む大規模臨床データを用いた、副作用予測システムや未知の副作用検出への基盤となることが期待される。

よって、本論文は博士（薬学）の学位論文として価値あるものと認める。また、令和5年2月14日、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。

要旨公表可能日： \_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日以降