

氏名	うすい みさ 臼井 美紗
学位の種類	博士(薬学)
学位記番号	博士甲 第 5040 号
学位授与の日付	平成 31 年 3 月 23 日
学位授与の要件	学位規則第 4 条第 1 項該当
学位論文題目	副作用に関する患者の表現の収集と活用に関する研究
論文審査委員	(主査) 望月 眞弓 教授 (医学博士) (副査) 山浦 克典 教授 (博士 (臨床薬学)) 漆原 尚巳 教授 (医学博士 (社会健康医学))

論文内容の要旨

【緒言】

近年、患者の医療参画に伴い患者による医薬品の直接的な評価が注目されている。特に、安全性評価においては副作用に関する患者の表現を知ることが患者自身やその家族・介護者による副作用の早期発見や適切なセルフモニタリングにつながるため重要である。患者や医療従事者などから報告された副作用に関する情報は、最終的に MedDRA (Medical Dictionary for Regulatory Activities) に標準化され、統計処理や国際間での情報共有に用いられる。しかしながら、現状では、患者が実際にどのような表現で症状を訴えたのかは共有されていない。また、誤って別の MedDRA に標準化されて共有された場合、医療従事者らが患者の表現を正確に捉えられず、適切に対応できない可能性がある。従って、副作用に関する患者の表現を収集し、その情報を活用可能とする方策が必要である。本邦においては患者の表現を収集・標準化してまとめた研究は報告されていない。そこで本研究では次の 3 点について検討を行った。

検討 1) 副作用に関する患者の表現を収集する方策の検討

検討 2) 自然言語処理による患者の表現の抽出とコーディングの自動化の検討

検討 3) 抽出した表現の辞書化とその有用性の検討

検討 1) 副作用に関する患者の表現を収集する方策の検討

【目的】

副作用に関する患者の表現の収集におけるクラウドソーシングの利用可能性を検討するため、副作用経験談の募集と方言の収集を行った。

【方法】

<副作用経験談の募集> Yahoo!クラウドソーシング登録者を対象にした無記名自記式質問票調査 (調査項目: 回答者の属性、副作用当事者の属性、医薬品名、経験内容) を実施した。得られた回答の経験内容を「MedDRA にある表現のみの回答」または「MedDRA がない表現を含む回答」の何れかに分類した。

<方言の収集と精査>Yahoo!クラウドソーシング登録者を対象にした無記名自記式質問票調査（調査項目：回答者の属性、症状の方言とその使用地域）を行い、予め指定した15の副作用症状について方言を収集した。得られた方言の語彙数を既存辞典（全国方言辞典）の収載語彙数と比較した。さらに、地域薬剤師を対象にした郵送による自記式アンケート調査（調査項目：勤務施設の所在地、クラウドソーシングで収集した方言の使用実態など）により、収集した方言の妥当性を精査した。

【結果】

約2週間で1,098症例の副作用経験談を取得した（有効回答率90.0%）。このうち581症例（52.9%）はMedDRAにない患者の表現を含んでいた。回答者の年齢は20歳～59歳が全体の約97%を占めていた。また、約1週間の調査で619語の方言を収集できた（有効回答率78.0%）。このうち548語は既存の辞典にない方言であった。地域薬剤師への調査では39都道府県254薬局から回答を得た（回収率69.0%）。619語の方言のうち502語（81.1%）について臨床での使用実態が確認された。

【考察】

クラウドソーシングの限界として15歳以下と60歳以上からの情報を得にくい、なりすましを完全には排除できない、アクセントや世代・市区町村レベルでの方言の差異を考慮していないなどが挙げられたが、短期間で大量の情報を得られ、データ処理が容易であることから、患者の表現の情報源として有用であると考えられた。

検討2) 自然言語処理による患者の表現の抽出とコーディングの自動化の検討

【目的】

薬歴のS情報から患者の表現を抽出し既存の用語体系に紐付ける自動処理システムを作成し、その性能を評価した。また、薬歴以外のデータに対するシステムの性能を評価し、複数の情報源への適用可能性を検討した。

【方法】

<ICD-10・薬歴システムの構築>保険薬局1店舗1年分の薬歴のS情報をもとに、自然言語処理を用いた患者の表現の抽出・コーディングの自動処理システムを構築した。システムは予め作成した検索ルールに基づき、「胃+荒れる」「血+止まりにくい」といった体の部位と症状の組み合わせなどにより患者の表現を認識し、対応するコードを同時に付与する。システムの実行結果を作業者のコーディング結果と比較し、適合率・再現率・F値で性能を評価した。

<MedDRA・薬歴システムの構築> <ICD-10・薬歴システムの構築>で作成した検索ルールをICD-10からMedDRAに書き換えたシステムを構築した。14店舗10年分の薬歴のS情報から副作用症例を集め、一文単位に分割したデータセットを作成し、これに対するシステムの性能評価を行った。システムと作業者のコーディング結果が不一致となったものについては、コード毎にその原因を「1. 否定や仮定表現の誤検出」「2.

検査項目/医薬品名の誤検出」「3. 副作用と関係のない語の誤検出」「4. 文脈の解釈が必要な表現」「5. 偽陰性」「6. 不適切なコードの付与」の何れかに分類した。不一致の原因のうち検索ルールの修正で対応可能な 2, 5, 6 に対して改善を行い、検索ルール改訂後のシステムの性能評価および不一致の原因を改訂前の結果と比較した。

＜MedDRA・薬歴システムの応用＞検討 1 のクラウドソーシングから収集した副作用経験談のうち MedDRA にない表現が含まれている 581 症例を一文ごとに分割したデータセットを作成した。これに対して＜MedDRA・薬歴システム＞で作成した改訂前・改訂後の検索ルールをそれぞれに適用し、性能および不一致の原因の比較を行った。

【結果】

＜ICD10・薬歴システムの構築＞では 4,892 症例の S 情報を用いた。検索システムを総合的に評価する指標である F 値は 0.65 であった。＜MedDRA・薬歴システムの構築＞では、14 薬局 10 年分の S 情報から 1,702 文の副作用データセットが作成され、システムの性能は F 値 0.47 となった。不一致の原因は 1,251 件となり、うち検索ルールの修正で対応可能な原因 2, 5, 6 は 958 件（76.6%）となった。これらについて改訂を行った結果、F 値は 0.47 から 0.72 まで改善した。不一致の原因は全体で 1,251 件から 539 件に減少し、原因 2 は 74 件から 7 件（90.5%減少）、原因 5 は 376 件から 31 件（91.6%減少）、原因 6 は 508 件から 13 件（97.4%減少）に減少した。一方で原因 4 は 17 件から 323 件まで増加していた。＜MedDRA・薬歴システムの応用＞では、クラウドソーシングで収集した副作用経験談から 757 文の副作用データセットが得られた。これに＜MedDRA・システムの構築＞で作成した改訂前後の検索ルールを適用すると、改訂前では F 値 0.38 となったのに対し改訂後では F 値 0.66 と性能が改善した。不一致の原因は改訂前の検索ルール適用時は 745 件で、うち原因 2, 5, 6 は 620 件（83.2%）であった。改訂後の検索ルール適用時は不一致が 471 件に減少し、原因 2 は 30 件から 3 件（90.0%減少）、原因 5 は 419 件から 203 件（51.6%減少）、原因 6 は 171 件から 82 件（52.0%減少）に減少した。一方、原因 4 は 11 件から 52 件に増加した。

【考察】

14 薬局の薬歴データを用いた＜MedDRA・薬歴システムの構築＞では、検索ルールの改訂により性能が大幅に改善した。不一致となったもののうち検索ルールの修正で対応可能な原因 2, 5, 6 がいずれも 90%以上減少していたことから、ルールの改訂が正しく行われたと考えられる。一方で、改訂により「4. 文脈の解釈が必要な症状」が増加していた。これは、「5. 偽陰性」を改善するために検索ルールを追加・修正したこと、「6. 不適切なコードの付与」の改善のために候補と考えられる MedDRA を全て出力するようにルールを修正したことにより、誤ったコードが付与される可能性が高くなったためと考えられる。同様の傾向はクラウドソーシングのデータの場合でもみられた。14 薬局の薬歴のデータとクラウドソーシングのデータいずれにおいても、改訂前の検索ルール適用時、不一致の原因で 2, 5, 6 が 7 割～8 割を占めていた。これは、新しいテキス

トデータにシステムを適用した場合、エラーのほとんどは検索ルールの修正で対応可能であることを示唆している。クラウドソーシングのデータにおいては、改訂前・改訂後どちらの検索ルール適用時にも「5. 偽陰性」が不一致の原因として最も多かった。このことから、クラウドソーシングのデータ中には薬歴にない表現が多数存在し、薬歴をベースに作成した検索ルールでは対応できずにとり逃しが多かったことが考えられる。今回、クラウドソーシングのデータを用いた検討では、検索ルールは薬歴のデータに沿って改訂したものを使用しており、クラウドソーシングのデータに対して最適化は行っていない。今後、今回の結果をもとに、クラウドソーシングのデータに対しても適切に MedDRA の付与が行えるように検索ルールを改訂することで、システムの性能をさらに改善することが可能と考えられる。これらの結果を総合すると、作成したシステムは様々な情報源に対して適用可能であると考えた。

検討 3) 抽出した表現の辞書化とその有用性の検討

【目的】

検討 2 までに収集・抽出した患者の表現を整理し『P-MedDRA』として辞書化し、どのような症状に関する患者の表現が集まりやすかったのかを調べる。また、既知の副作用報告例を用いて副作用の早期発見における患者の表現の活用に『P-MedDRA』が有用であるかを検討した。

【方法】

<『P-MedDRA』の作成> 検討 1 で収集した方言および検討 2 で使用した薬歴とクラウドソーシングのデータをもとに、患者の表現と MedDRA の PT が紐付いた辞書データを作成した。方言を除いた語の収載語数と分布 (MedDRA の SOC 毎の語数) について、「Patient-Friendly Term List」(1,440 語) と比較した。

<有用性の検討> 「日本病院薬剤師会雑誌」の「プレアボイド広場」のうち副作用の重篤化を回避したプレアボイド報告について、記載されている患者の症状が MedDRA の LLT に完全一致するか、『P-MedDRA』に同じ語が含まれているかを確認した。

【結果】

クラウドソーシング由来で 593 語、薬歴由来で 496 語が得られ、重複を削除して 1,022 語 (方言を含めると 1,641 語) が『P-MedDRA』となった。このうち MedDRA に対応する語がない患者の表現が 1~2%みられた。『P-MedDRA』の語は 21 種類の SOC に分布し、「Patient-Friendly Term List」と比較すると「神経系障害」、「皮膚および皮下組織障害」が特に集まりやすい傾向にあった。プレアボイド報告 81 症例 172 症状を用いた検討の結果、MedDRA の LLT と完全一致した症状が 107 件 (62.2%)、『P-MedDRA』に同一の語が存在した症状が 9 件 (5.2%)、類似の語が存在した症状は 41 件 (23.8%)、全く存在しなかった症状が 15 件 (8.7%) あった。

【考察】

『P-MedDRA』は「Patient Friendly-Term List」に劣らない語数を収載しており、様々なSOCに分布し、その1~2%はMedDRAに対応する語がなかった。このことから、本研究では多様な患者の表現が得られ、その中には患者独特の表現が含まれていたと考えられた。今回は「神経系障害」、「皮膚および皮下組織障害」に関連する表現が特に多く得られたが、これは患者が自覚しやすい症状の表現が集まりやすかったためと考えられた。プレアボイド報告を用いた検討では重篤な副作用81症例の症状のうちMedDRAと完全一致したのは6割で、『P-MedDRA』に同一または類似の語が存在した症状を合わせると全体の9割の症状はMedDRAと『P-MedDRA』で網羅できた。これは、従来の用語体系ではカバーしきれない患者の表現を『P-MedDRA』を併用することで補完できる可能性を示している。今後、同様の症状が規制当局や製薬企業に報告された場合、『P-MedDRA』を参照することでより簡便に報告内容をMedDRAに標準化して共有することができ、適切な対応につなげられる可能性がある。また、服薬指導や薬剤情報提供書の記述などに『P-MedDRA』を参考とすることで、患者がより理解しやすい言葉で副作用の説明を受けることにつながり、患者やその家族が副作用のセルフモニタリングを行う一助にもなると期待される。今回の検討ではプレアボイド報告のみを用いたため、非重篤な副作用については評価を行えていないものの、副作用の早期発見における患者の表現の活用に『P-MedDRA』が有用である可能性が示唆された。

【結語】

本研究は副作用に関する患者の表現の情報源としてクラウドソーシングや薬歴のS情報の利用可能性を初めて検討し、テキストデータから自動的に患者の表現を抽出し既存の用語体系に紐付ける手法を本邦で初めて検討した。また、日本語の患者の表現をMedDRAに紐づけた用語集を初めて作成した。『P-MedDRA』は多様な症状の表現を含んでおり、副作用の早期発見における患者の表現の活用に有用である可能性が示唆された。本研究で作成した『P-MedDRA』やその過程で検討した方法を様々な場面に応用し、得られた知見を辞書の拡充や検索ルールの改訂などに還元することで、継続的に患者の表現を収集し臨床現場に共有する仕組みの確立が期待できる。

論文審査結果の要旨

論文発表の後、副査2名と研究科委員からの試問ならびに質疑応答を行った。参加型の医療の重要性が叫ばれる中で、医療における患者の発話の重要性が認識されてきている。特に医薬品による副作用について、患者の発話から医療者が気づかない新しい副作用が発見されることもある。発表者は、患者の発話を効率的に集める手法を検討し、集めた言葉を辞書化し、患者の言葉による副作用情報を医療関係者が利用できるようなシステムを開発した。発表者は、研究の背景、研究の目的・意義および研究内容について、専門外の研究科委員に対しても論理的かつ分かりやすく説明した。副査である医療薬学の研究者、薬剤疫学の研究者および副査以外の臨床薬物動態学の研究者等からの質問にも的確に回答し、意義ある研究として高く評価された。

主な質問として、

- 1) 患者の表現の情報源として薬歴のS情報に注目した理由とメリットは。
- 2) 薬歴を用いると薬剤師の解釈が入り患者の表現が加工されている可能性があるが患者の表現の情報源として本当に適しているのか。
- 3) 薬局間でS情報の質的な差があると思うが、今後他の薬局への応用する場合、その差についてはどう考えるか。
- 4) 音声入力と現状のテキスト入力ではどちらのデータが適しているのか。
- 5) 検索ルールは患者ベースで作成しているが、逆側（MedDRA）からのアプローチは考えられないか。患者側からのアプローチを選択した理由とメリットは。
- 7) 今回は限られたデータセットから限られたルールしか作れていないが、実用化を考えた場合の筋道はどのように考えているか。
- 8) 文脈を考慮できないことや患者背景の不足によって今回乗り越えられなかったエラーについては今後どうすれば改善できると考えるか。
- 9) 副作用以外の症例も含めたデータセットに適用すると本来疾患に起因する症状も拾ってしまい性能が非常に落ちると懸念されるが、その点についてはどのように考えるか。

などの質問に対して、発表者は自身の研究結果に基づいて的確に説明・回答し、質問者から概ね了解された。

発表者は、患者の発話の情報源として、クラウドソーシングおよび薬局での薬歴のS情報(Subjective data)に着目し、これらから副作用に関連する患者の言葉を収集した。収集に際してはコンピュータを用いた自然言語処理により、より多くの情報を効率よく処理することを可能とした。さらに副作用を整理するため、医療用語集である MedDRA でのコーディングの自動化を検討し、集めた患者用語を MedDRA の辞書に追加するための P-MedDRA を作成した。ここで作成した P-MedDRA は、日本における患者副作用報告システムを本格稼働する際にも極めて有用なものとなると考えられた。以上より、本論文の研究内容は本学の博士（薬学）の学位に十分値するものと判定された。

論文目録

【主論文に関する原著論文】

Usui M, Aramaki E, Iwao T, Wakamiya S, Sakamoto T, and Mochizuki M. Extracting and Standardizing Patient Complaints from Electronic Medication Histories for Pharmacovigilance: Natural Language Processing Analysis in Japanese. JMIR Med Inform. 2018 Sep 27;6(3):e11021. doi: 10.2196/11021.

【参考論文】

該当なし