

平成29年度 厚生労働科学研究費補助金

政策科学総合研究事業（臨床研究等ICT基盤構築・人工知能実装研究事業）
人工知能を活用した副作用症例報告の評価支援の基盤整備と試行的評価

平成29年度 総括・分担研究報告書

研究代表者 今任 拓也

平成30（2018）年 5月

目 次

I . 総括研究報告

- 人工知能を活用した副作用症例報告の評価支援の基盤整備と試行的評価 1
今任拓也 国立医薬品食品衛生研究所・医薬安全科学部 主任研究官

II . 分担研究報告

- 1 . 副作用症例報告の評価支援のための自動アノテーションおよび辞書作成に関する研究 10
潮田 明 産業総合技術研究所・人工知能研究センター・招聘研究員
- 2 . 医薬品副作用データベース（JADER）を用いた対象副作用に関連する因子の同定に関する研究 16
今任拓也 国立医薬品食品衛生研究所・医薬安全科学部・主任研究官
斎藤嘉朗 国立医薬品食品衛生研究所・医薬安全科学部・部長

III . 研究成果の刊行に関する一覧表 25

平成30年度 厚生労働科学研究費 補助金
(政策科学総合(臨床研究等ICT基盤構築・人工知能実装研究事業))
(総括)研究報告書

研究代表者 今任 拓也 国立医薬品食品衛生研究所・医薬安全科学部・主任研究官

研究要旨:【目的】本研究は、副作用判定における人的作業の効率化を図るため、近年、ディープラーニングという新たな技術が提唱され、医療の分野を含め、様々な分野において注目を集めている人工知能を用いて、副作用評価支援システムの構築および人工知能を用いた副作用判定の試行的な評価を行うことを目的としている。今年度は、副作用評価判定に人工知能を応用するためのアノテーション作業および表記ゆれ解消のための機械学習用辞書の構築を行い、学習評価モデル構築に先立ち、副作用個別症例報告と同様の情報に基づき作成されている医薬品副作用データベース(JADER)を用い、スティーブンス・ジョンソン症候群(SJS)および中毒性表皮壊死融解症(TEN)に特徴的な因子を探索し、SJSおよびTENの判別能について検討を行った。【結果】本研究で使用する副作用報告個別症例の情報は、原則非開示情報とされており、情報の二次利用および外部施設へのデータの提供が考慮されていなかったため、データの精査を行い、使用可能なデータを決定し、専門医による助言に基づき、診断基準が明確であり、検査値情報の必要性の低いSJSおよびTENを対象副作用とした。また、テキストデータを過不足・曖昧なく機械に伝えるためのアノテーション作業を開始した。さらに、表記ゆれ解消のため、webより抽出した標準的病名約2,500語をベースにしたニューラルネットワークによる類義表現抽出実験より類似度評価器を作成し、特定の医療関連用語とほぼ同義かあるいは置換可能なレベルで類似した用語を効率よく収集できることを確認した。また、JADERを用いた解析では、性別、年齢区分(未成人、成人、高齢者)、予後不良、発症までの期間(10日以内または10日以上)および被疑薬(ロキソプロフェンナトリウム水和物および総合感冒薬(一般)の有無)がSJSおよびTENに特徴的な因子と考えられ、SJSおよびTENとそれ以外の重症薬疹との判別では、予後不良、被疑薬が他の因子に比べてやや判別能が高かった。【考察・今後の計画】現段階では、データは限定されているが、可能な限りデータを追加することにより、精度の高いシステムの構築を目指す。今年度実施した使用可能なデータの精査およびアノテーションは、機械が副作用を正確に評価するために最も重要な要素である。精度の高い副作用評価判定モデルの構築に向け、次年度以降もアノテーション作業および機械学習用辞書の作成を継続し、副作用判定モデルの重み付けのためにテキストマイニングの手法を用い、添付文書から吸収・排泄などの体内動態、薬理作用などといったさらなるSJS及びTENに特徴的な因子の検討を行う。

研究代表者:

今任拓也・国立医薬品食品衛生研究所・医薬安全科学部・主任研究官

研究分担者:

潮田 明・産業総合技術研究所・人工知能研究センター・招聘研究員

斎藤嘉朗・国立医薬品食品衛生研究所・医薬安全科学部・部長

渡邊 環・医薬品医療機器総合機構・安全第二部・主任専門員

相原道子・横浜市立大学・耳鼻科・教授

A. 研究目的

医薬品の安全対策に対し、リスクマネジメントプランが導入され、開発段階から市販後までの安全対策の充実強化が図られている。一方で、国際共同治験の推進により臨床試験における日本人症例の割合は減少し、日本人における医薬品の開発段階の安全性に関するエビデンスは十分とは言えない。また予定される条件付き早期承認制度の創設により、これまで限られた患者数での安全性検討により発売される場合もあり、市販後における医薬品安全性監視は、益々重要となってきた。

この市販後医薬品安全対策の骨格を成すものは、医薬品・医療機器等法の第六十八条の十に規定される副作用等の報告である。本報告は、主に製薬企業により症例報告書の形で直接医薬品医療機器総合機構（PMDA）に提出される。企業報告は、年々増加傾向にあり、2015年には、海外症例 35 万件、国内症例 5 万件となっている。そのため、PMDA では膨大な症例報告を効率的かつ適切に評価をすることが必要となっている。

人工知能は、ひとの知的な処理を機械によっ

て行う技術であり、ひとの意思決定に代わるものとなる。人工知能の研究は、1950年代から続いており（Buchanan, Bruce G. A (Very) Brief History of Artificial Intelligence, AI Magazine(Winter) 2005;: 53-60）2006年ごろにディープラーニングという新たな技術が提唱され、現在では、第三次ブームとして、医療分野を含めた様々な分野において大きな期待が寄せられている。

そこで、本研究は、PMDAにおいて管理されている副作用個別症例報告を用いて、人工知能を応用したスティーブンス・ジョンソン症候群（SJS）および中毒性表皮壊死融解症（TEN）の副作用判定システムを構築し、その試行的評価を実施する。今年度は、副作用評価判定に人工知能を応用するためのアノテーション作業および表記ゆれ解消のための機械学習用辞書の構築を行い、学習評価モデル構築に先立ち、副作用個別症例報告と同様の情報に基づき作成されている医薬品副作用データベース（JADER）を用い、SJS および TEN に特徴的な因子を探索し、SJS および TEN の判別能について検討を行うことを目的とした。

B. 研究方法および研究結果

今年度実施した研究の流れを図1に示した。

使用するデータの精査

本研究で使用する副作用個別症例報告は、医薬品医療機器等法に基づき、実施されており、PMDAにてその情報の管理・整理を行っている。データは原則非開示情報（機密性情報3）とされており、これまで情報の二次利用および外部施設へのデータの提供が考慮されていなかった。

たため、個人情報保護の観点を含め、今年度は、今後のデータの取扱いについて十分な検討を行い、研究班内でデータの取扱いに関する規定を作成することとした。また、検討の結果から現時点で使用可能なデータを表1に示した。使用可能なデータは、主に副作用症例報告の経過欄のテキストデータより、被疑薬、日付（ただし、被疑薬投与日を起点とした日数データとする）、傷病名、症状、変化などの情報とした。

対象副作用

使用するデータの精査の検討から検査値データの使用は、個人情報保護法において要配慮個人情報に該当し、PMDAのセキュリティポリシーにおいて、外部施設への提供は困難であるとの見解から、現時点では、検査値情報の使用が困難であり、必ずしも必要な検査値が記載されていない可能性もあるため、診断基準において、検査値情報の必要性が低く、相原道子分担研究者による助言により、診断基準が明確であるSJSおよびTENを対象副作用とすることとした。また、SJSおよびTENの診断基準を表2、3に示した。

アノテーション作業

アノテーションは、専門家がテキストから読み取れる重要な内容を過不足・曖昧性なく機械に伝えるために必須である。アノテーションには、産総研より提供された日本語にも対応したテキストアノテーションツール「brat rapid annotation tool」を用いることとし、副作用症例報告の経過欄のテキストデータを基にアノテーション作業を開始した。現在PMDAにて実施されている副作用評価に必要な情報として、医療用語（医薬品名、病名、症状名、医師によ

る処置名など）投与量、投与期間、投与頻度をはじめ、薬剤の種類と投与量の変化、患者の状態の変化などの「変化」や薬と症状との関係性などを中心にアノテーションを実施している。さらに今後の汎用性も考え、医薬品名は、一般名、傷病名は可能な限り、疾病及び関連保健問題の国際統計分類第10版（ICD10）コードを付与することとした。今後、人工知能を用いたアノテーションの自動化のため、手動でのアノテーションの結果を基に、さらに必要な情報を追加し、アノテーションの精度を上げていく予定である。

機械学習用辞書作成

今年度は、標準的病名約2万5千語（以下「ベース用語」）をベースにしたニューラルネットワークによる類義表現抽出実験を実施した。この類義表現抽出実験は本研究が対象とする皮膚障害に限らず副作用報告書全般に適用可能な表記ゆれ対策を視野に入れて手法の可能性を検証する目的で行った。

2万5千語のベース用語はWebから医療関係のテキストを収集するための検索キーワードとして用いた。まずWebボットを用いてそれぞれのベース用語を含む合計約280万のWebサイトからテキスト情報を抽出し、タグ等を排除して約10億語のテキストを収集した。次にこのテキストを学習データとして、用語を高次元のベクトルで表現（分散意味表現）するためのニューラルネットワークを学習させ、ベクトルの近さで用語の類似度を評価する類似度評価器を作成し、少数サンプルの抽出による調査を行ったところ、本評価器と人手によるチェックを組み合わせることで特定の医療関連用語とほぼ同義かあるいは置換可能なレベルで類

似した用語を効率よく収集することが可能であることが分かった。

対象副作用に関連する因子の同定

データ・ソースは、2017年11月更新のJADERデータセットをPMDAウェブサイト(<https://www.pmda.go.jp/safety/info-services/drugs/adr-info/suspected-adr/0005.html>)よりダウンロードし、解析に用いた。重症薬疹はMedDRAのPTコードに基づき、本研究の対象副作用としたSJS(皮膚粘膜眼症候群を含む)、TENに加え、好酸球増加と全身症状を伴う薬疹(DIHS)、多形紅斑(EM)とした。

今回用いたデータセットのうち、対象とした重症薬疹の報告症例は16,688症例であった(SJS[5,640症例]、TEN[2,683症例]、DIHS[3,647症例]、EM[4,718症例])。報告数が最も多かったのは、SJSであった。年代についてはEMでは、高齢者の報告割合が、50.81%と他の重症薬疹と比べて最も多くなっていた。また、SJS、TEN、DIHSにはほぼ性差は認められなかったが、EMでは女性が60.1%と男性に比べて多くなっていた。TENでは、予後不良者の割合が他の副作用に比べて多くなっていた。発症までの平均日数は、SJS、TEN、EMでおよそ30日となっているのに対し、DIHSはおよそ42日で、他の副作用に比べて遅くなっていた。中央値を比較しても、DIHSが最も遅くなっていた。被疑薬では、SJSは、ラモトリギン、TENは、アロプリノール、DIHSはカルバマゼピン、EMは、ソラフェニブトシル酸塩が最も多くなっていた。ラモトリギン、カルバマゼピン、アロプリノールは4つの副作用に共通して上位被疑薬となっていた。総合感冒剤(一般)、ロキソプロフェンナトリウム水和物は、SJSとTENで上位被疑薬と

なっていた。

そこで、本研究の対象副作用であるSJSおよびTENとそれ以外の重症薬疹(DIHS、EM)との判別について、性別、年齢区分(未成人、成人、高齢者)、予後不良、発症までの期間(10日以内または10日以上)および被疑薬(ロキソプロフェンナトリウム水和物および総合感冒剤(一般)の有無)の判別能の評価を行ったところ、予後不良、被疑薬は、他の因子に比べてやや判別能が高かった。さらにすべての因子をモデルに投入したところ、判別能は最も高かった。

C. 考察

本研究は、PMDAにて整理・管理されている副作用症例報告の経過欄の情報を基に、人工知能を用いて、副作用を精度良く判別できるシステムの構築およびその試行的評価を行うものである。

研究開始に先立ち、本研究で使用する副作用報告症例は、本来、XXを基に使用されており、そのデータの二次利用が考慮されていなかったため、個人情報保護などの観点から、使用可能なデータの精査を行う必要があった。精査の結果、現時点で使用可能であるデータは、日付(投与日を起点とした日数)情報、副作用症例報告の経過欄内の医療用語(医薬品名、病名、症状名、医師による処置名など)、投与量、投与期間、投与頻度をはじめ、薬剤の種類と投与量の変化、患者の状態の変化などの「変化」や薬と症状との関係性などの情報に限られており、検査値などの情報は、要配慮個人情報に当たるため、現時点では使用不可とした。

使用可能なデータと専門医の助言により、まずは対象副作用をSJSおよびTENとし、アノテーション作業を開始した。また、経過欄の情報

は記載者に依存する。そのため、表記ゆれ解消のための辞書の作成も開始し、web サイトより抽出したテキスト情報を学習データとし、ニューラルネットワークを用い、類似度評価器を作成したところ、特定の医療関連用語とほぼ同義かあるいは置換可能なレベルで類似した用語を効率よく収集することが可能であることを確認した。大量のテキストから類似した用語を抽出するクラスタリング技術はこれまで様々な手法が提案されているが、従来のアプローチである統計的手法も最近著しく発展してきている。ニューラルネットワークによる手法も大量のテキスト（コーパス）の存在が前提となっている。しかしながら医療関係のテキストに関しては、研究用に利用できる日本語のリソースはごく限られており、本研究においても、ある程度まとまった量の副作用報告個別症例データの入手を前提に計画を進めているものの、適宜他の言語リソースによる補完も進める必要がある。今年度は使用するデータの精査を行う必要があったため、副作用症例報告データの入手が予定よりも遅れ、その対応策として、Web からテキストを収集するアプローチを試みた。さらに、副作用評価判定の精度を上げるために JADER のデータを用い、重症薬疹に特徴的な因子を検討し、その因子の SJS および TEN の判別能を推定した。

現在、ディープラーニングを始めとした人工知能に関する研究に注目が集まっている。特にディープラーニングは特徴量を自ら作り出すことができるようになった点で画期的な技術ではあるが、大規模かつ質の高いデータが必要であり、単独ではあらゆる分野のさまざまな種類の問題を解決できるような万能な人工知能を生み出すには至っていない。したがって、ま

ずは本研究のように、医療の分野を含め、それぞれの分野における具体的な問題に対応できる人工知能を個々に実用化していくための研究が重要であると考えられる。

本研究は、PMDA において、現在、人的作業により行われている副作用判定に人工知能を応用し、その効率化を図ることを目標としている。現段階では、使用できるデータは限定されているが、可能な限りデータを追加することにより、精度の高いシステムの構築を目指す。今年度実施したデータの精査およびアノテーションは、今後、機械が副作用を正確に評価するために最も重要な要素である。特にアノテーションは、副作用個別症例報告の経過欄の情報を曖昧なく抽出しなければならないため、十分に検討することが必要である。日本語はさまざまな表記ゆれを示すため、使用する用語を統一することも重要である。また、JADER は、reporting bias など様々なバイアスを含んでいるが、より多くの症例データを活用することができるため、対象副作用の特徴を検討するには有用であると考えられる。そのため、JADER の解析によって得られた結果は、今後の人工知能を用いた副作用評価モデルの特徴量およびその重み付けにおいて、重要なエビデンスとなると考えられる。

D. 今後の研究計画

次年度は、今年度開始した機械学習用辞書作成およびアノテーション作業を継続する。アノテーション作業を基に作成したアノテーション機械学習のための学習データおよび機械学習用辞書を基に、自動でアノテーションを行うための機械学習モデルの開発を行う。更に平成 31 年度の人工知能による副作用判定の試行的評価に向けて、副作用評価のための機械学習用

トレーニングデータの作成を開始する。対象副作用に関連する因子の同定では、今年度同様、JADER を用いた更なる解析を実施すると共に、被疑薬候補となる医薬品の電子化された添付文書を用い、テキストマイニングの手法により、吸収・排泄などの体内動態、薬理作用などさらなる因子の検討を行う予定である。

E. 倫理審査状況

本研究で使用するデータは、PMDA が所有する既存の副作用個別症例データの二次利用であり、患者のインフォームド・コンセントは取得していない。しかしながら、医薬品の副作用報告は、「医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律施行規則」第 228 条の 20 第 1 項に基づき実施されるものであり、副作用個別症例情報には、患者氏名、生年月日等などの個人に関する情報は記載されておらず、PMDA にて対応表は保有していないため、特定の個人を識別することは不可能である。また、「人を対象とする医学系研究に関する倫理指針」に則り、PMDA にて承認を受け、国立医薬品食品衛生研究所および産業総合研究所の倫理審査においても承認を受けている。

F. 健康危険情報

該当なし

F. 研究成果

1. 論文発表

該当なし

2. 学会発表

該当なし

表1 副作用個別症例報告データについて

変数	使用可能状況
身長・体重	使用不可
日付データ（投与日、経過欄日付、過去治療歴日付、副作用発症日） （ただし、被疑薬投与日を起点とした日数とする）	使用可
経過欄（テキストデータ） （ただし、患者名や病院名、検査値などはマスキングする）	使用可
検査値	使用不可
被疑薬	使用可

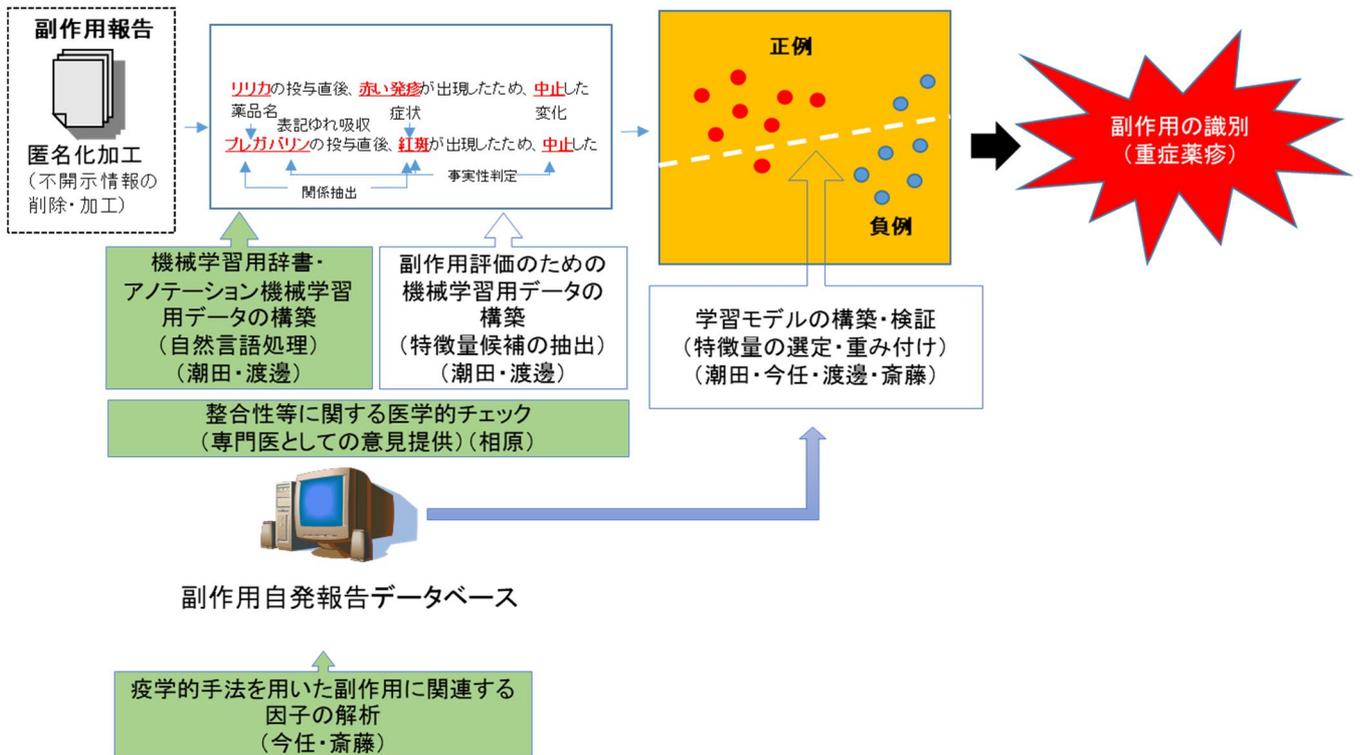


図1 研究の流れ

緑地の吹き出しは今年度実施した課題である。

表 2 SJS の診断基準

主要所見 (必須)	<ol style="list-style-type: none"> 1 皮膚粘膜移行部(眼,口唇,外陰部など)の広範囲で重篤な粘膜病変(出血・血痂を伴うびらん等)がみられる。 2 皮膚の汎発性の紅斑に伴って表皮の壊死性障害に基づくびらん・水疱を認め,軽快後には痂皮,膜様落がみられる。その面積は体表面積の10%未満である。但し,外力を加えると表皮が容易に剥離すると思われる部位はこの面積に含まれる。 3 発熱がある 4 病理組織学的に表皮の壊死性変化を認める* 5 多形紅斑重症型(erythema multiforme [EM] major)**を除外できる。
副所見	<ol style="list-style-type: none"> 1 紅斑は顔面,頸部,体幹優位に全身性に分布する。紅斑は隆起せず,中央が暗紅色の flat atypical targets を示し,融合傾向を認める。 2 皮膚粘膜移行部の粘膜病変を伴う。眼病変では偽膜形成と眼表面上皮欠損のどちらかあるいは両方を伴う両眼性の急性結膜炎がみられる。 3 全身症状として他覚的に重症感,自覚的には倦怠感を伴う。口腔内の疼痛や咽頭痛のため,種々の程度に摂食障害を伴う。 4 自己免疫性水疱症を除外できる。

*病理組織学的に完成した病像では表皮の全層性壊死を呈するが,少なくとも 200 倍視野で 10 個以上の表皮細胞(壊)死を確認することが望ましい。

**多形紅斑重症型(erythema multiforme [EM] major)とは比較的軽度の粘膜病変を伴う多形紅斑をいう。皮疹は四肢優位に分布し,全身症状としてしばしば発熱を伴うが,重症感は乏しい。SJS とは別疾患である。

表 3 TEN の診断基準

<p>主要所見 (必須)</p>	<ol style="list-style-type: none"> 1 広範囲に分布する紅斑に加え体表面積の 10%を超える水疱・びらんがみられる．外力を加えると表皮が容易に剥離すると思われる部位はこの面積に含める (なお，国際基準に準じて体表面積の 10～30%の表皮剥離は，SJS / TEN オーバーラップと診断してもよい) 2 発熱がある． 3 以下の疾患を除外できる． <ul style="list-style-type: none"> ブドウ球菌性熱傷様皮膚症候群 (SSSS) トキシックショック症候群 伝染性膿痂疹 急性汎発性発疹性膿疱症 (AGEP) 自己免疫性水疱症
<p>副所見</p>	<ol style="list-style-type: none"> 1 初期病変は広範囲にみられる斑状紅斑で，その特徴は隆起せず，中央が暗紅色の flat atypical targets もしくはびまん性紅斑である．紅斑は顔面，頸部，体幹優位に分布する． 2 皮膚粘膜移行部の粘膜病変を伴う．眼病変では偽膜形成と眼表面上皮欠損のどちらかあるいは両方を伴う両眼性の急性結膜炎がみられる． 3 全身症状として他覚的に重症感，自覚的には倦怠感を伴う．口腔内の疼痛や咽頭痛のため，種々の程度に摂食障害を伴う． 4 病理組織学的に表皮の壊死性変化を認める．完成した病像では表皮の全層性壊死を呈するが，軽度の病変でも少なくとも 200 倍視野で 10 個以上の表皮細胞 (壊)死を確認することが望ましい．

副作用症例報告の評価支援のための自動アノテーションおよび辞書作成に関する研究

研究分担者 潮田 明 国立研究開発法人産業技術総合研究所・人工知能研究センター・招聘研究員

研究要旨:【目的】本研究は、副作用を迅速かつ客観的に評価するための人工知能を活用した副作用症例報告評価技術の開発に先立ち、副作用症例報告テキスト中の重要情報を過不足・曖昧性なく機械に伝えるための自動アノテーション技術を開発することを目的としている。今年度は自動アノテーションを機械学習を用いて行うための学習用データの作成、および表記ゆれ解消のための機械学習用辞書の構築を目的としている。【方法】アノテーションには、日本語にも対応したテキストアノテーションツール「brat rapid annotation tool」を用いることとした。表記ゆれ解消のための辞書作成に関しては、本研究が対象とする皮膚障害に限らず副作用報告書全般に適用可能な表記ゆれ対策を視野に入れて、Web から収集した大量テキストから類似表現を検出する手法の検討を行った。【結果】現在 PMDA にて実施されている副作用評価に必要な情報として、医療用語(医薬品名、病名、症状名、医師による処置名など)、投与量、投与期間、投与頻度をはじめ、薬剤の種類と投与量の変化、患者の状態の変化などの「変化」や薬と症状との関係性などを中心にアノテーション対象を選定し、それらを configuration ファイルにまとめてアノテーション作業を開始した。辞書作成に関しては、標準的病名約2万5千語をベースにしたニューラルネットワークによる類義表現抽出実験を行った。Web から収集した標準病名を含む大量のテキストをもとに分散意味表現に基づく類似度評価器を作成した結果、評価器と人手によるチェックを組み合わせることで特定の医療関連用語とほぼ同義かあるいは置換可能なレベルで類似した用語を効率よく収集することが可能であることが分かった。【考察】今回試みた分散意味表現を用いる手法は、用語の内部構成ではなく用語のテキスト中の文脈の近さをもとに類似度を評価するものであり、同じ用語のカタカナ表記と漢字表記の類似性を検出できるなど大きな利点を有するものと考えられる。

A. 研究目的

今年度は、副作用を迅速かつ客観的に評価するための人工知能を活用した副作用症例報告評価技術の開発に先立ち、副作用症例報告テキスト中の重要情報を過不足・曖昧性なく機械に伝えるための自動アノテーションを機械学習を用いて行うための学習用データの作成、および表記ゆれ解消のための機械学習用辞書の構築を目的とした。

B. 研究方法

アノテーション作業

副作用症例報告の経過欄のテキストデータを対象にアノテーション作業を開始した。アノテーションには、日本語にも対応したテキストアノテーションツール「brat rapid annotation tool」を用いることとした。アノテーションの作業画面の例を図1に、アノテーション結果を格納したアノテーションファ

イルの例を図2に示した。アノテーションは、専門家がテキストから読み取れる重要な内容を過不足・曖昧性なく機械に伝えるために必須である。現在PMDAにて実施されている副作用評価に必要な情報として、医療用語（医薬品名、病名、症状名、医師による処置名など）投与量、投与期間、投与頻度をはじめ、薬剤の種類と投与量の変化、患者の状態の変化などの「変化」や薬と症状との関係性などを中心にアノテーション対象を選定し、それらを configuration ファイルにまとめてアノテーション作業を開始した。手動でのアノテーション結果は今後2つの用途で使用する。1つ目の用途は、アノテーションの自動化のための学習データである。テキスト中のそれぞれの用語に付与されたタグ（上記アノテーション対象に付与された医薬品名、病名などの名称）の付け方を用語の前後の文脈とともに機械学習させることにより、アノテーション作業の自動化を図る。2つ目の用途は副作用評価のための機械学習用データを作成する際の元データとしてである。本機械学習においては、アノテーション付きのテキストを入力として、副作用評価結果を出力する機能を学習する。

機械学習用辞書作成

機械学習用辞書作成に関しては、「ICD10 対応標準病名マスター」に収載された標準的病名約2万5千語（以下「ベース用語」）をベースにしたニューラルネットワークによる類義表現抽出実験を行った。この類義表現抽出実験は本研究が対象とする皮膚障害に限らず副作用報告書全般に適用可能な表記ゆれ対策を視野に入れて手法の可能性を検証する目的で行ったも

のである。

大量のテキストから類似した用語を抽出するクラスタリング技術に関してはこれまで様々な手法が提案されているが、従来アプローチである統計的手法も最近著しく発展してきているニューラルネットワークによる手法も、大量のテキスト（コーパス）の存在が前提となっている。しかしながら医療関係のテキストに関しては、研究用に利用できる日本語のリソースはごく限られている。本研究においても、ある程度まとまった量の副作用報告症例データの入手を前提に計画を進めており、それらに対するアノテーション作業完了後には、図3に示すようにアノテーション機械学習用データから表記ゆれの吸収を学習する手法を開発する計画である。しかしながら同時に適宜他の言語リソースによる補完も進める必要がある。初年度である今年度は、副作用報告症例の使用において、個人情報保護などの観点から使用可能なデータの精査と匿名化等のデータの加工を行う必要があり、症例データの入手が予定よりも大幅に遅れたため、その対応策として、Web からテキストを収集するアプローチを試みることにした。

C. 研究結果

上記2万5千語のベース用語をWebから医療関係のテキストを収集するための検索キーワードとして使用した。まずWebポットを用いてそれぞれのベース用語を含む合計約280万のWebサイトからテキスト情報を抽出し、タグ等を排除して約32億語のテキストを収集した。次にこのテキストを学習データとして、用語を高次元のベクトルで表現（分散意味表現）するためのニューラルネットワーク（word2vec）を

学習させ、ベクトルの近さで用語の類似度を評価する類似度評価器を作成した。図4に類似度評価器により検出された病名(A)および症状(B)の類義語の例を示す。図4の例は、見出し語との類似度が最も高い300語から人手により類義語/表記ゆれ等を抽出した結果である。少数サンプルの抽出による調査を行ったところ、本評価器と人手によるチェックを組み合わせることで特定の医療関連用語とほぼ同義かあるいは置換可能なレベルで類似した用語を効率よく収集することが可能であることが分かった。

D. 考察

副作用症例報告の経過欄に記載される病名や患者の症状は、診療録中の記載同様非常に多様な表現で記述されることが多い。一方でこれらのテキスト情報を機械で処理する場合、一字でも異なる用語は全く別の用語として処理されるため、表記ゆれの吸収は極めて重要な課題である。同じ病名でも漢字だけの表記、カタカナだけの表記、アルファベットだけの表記、漢字とカタカナを組み合わせた表記、アルファベットと漢字を組み合わせた表記など様々な表記形態が存在し得る。図4(A)に示される例からも推測できる通り、カタカナ表現に対しては非常に多くのバリエーションや誤記表現が用いられる可能性がある。また一部の漢字表現も同様と考えられる。これまで表記ゆれの対策においては、用語中の文字配列の近さ、すなわち文字の置換・挿入・削除といったオペレーションによる用語の相互遷移の容易さを基準に類似度を評価する手法が主流であったが、「アナフィラキシー」の例にも見られるように誤記までを対象に考えた場合、オペレーションの組み合

わせは限りなく存在し、必要となる学習データの量を考えた場合、用語中の文字配列と言った局所的特徴のみに頼るには限界があると考えられる。今回試みた分散意味表現を用いる手法は、用語の内部構成ではなく、用語のテキスト中の文脈の近さをもとに類似度を評価するものであり、同じ用語のカタカナ表記と漢字表記の類似性を検出できるなど大きな利点を有するものと考えられる。

E. 結論

副作用症例報告中の表記ゆれを吸収するためには、アノテーション機械学習用データをなるべく多く作成することが有効であることは間違いないが、人手によるデータ作成のみに頼るのではなく、外部の言語リソースによる補完を行うことも有効である可能性が示された。今回はWebから収集した標準病名を含む大量のテキストをもとに分散意味表現に基づく類似度評価器を作成した結果、評価器と人手によるチェックを組み合わせることで特定の医療関連用語とほぼ同義かあるいは置換可能なレベルで類似した用語を効率よく収集することが可能であることが分かった。

G. 研究発表

1. 論文発表
該当なし
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

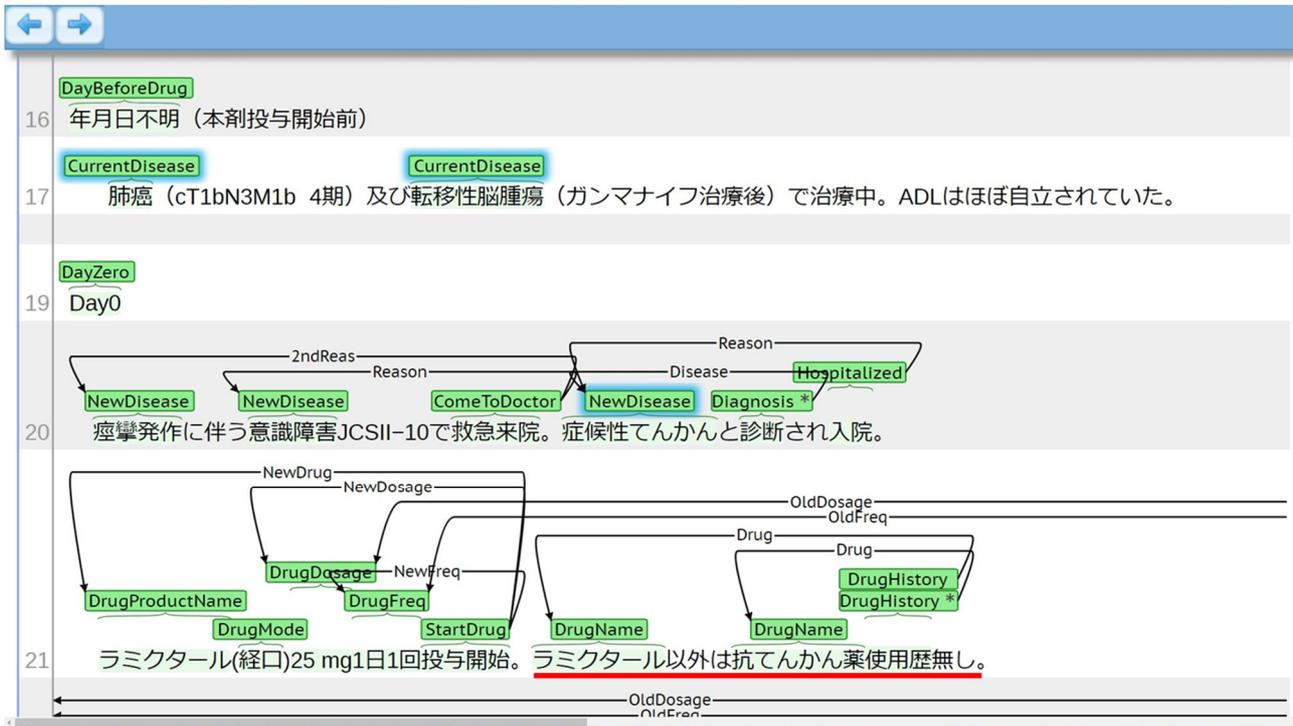


図 1 brat によるアノテーション例

912番目と916番目の文字境界に挟まれた文字列

Entity ID	\$ cat AdverseDrugReaction_14025805.ann	テキスト中文字列
T1	NewDisease 912 916	意識障害
T2	DrugDosage 957 962	25 mg
T3	DrugDosage 1030 1035	50 mg
T5	CurrentDisease 859 865	転移性脳腫瘍
T6	NewDisease 930 937	症候性てんかん
T4	DrugDosageChange 1045 1047	変更
E1	DrugDosageChange:T4 NewDosage:T3 OldDosage:T2 Drug:T13	
T8	DrugFreq 962 966	1日1回
T9	DrugFreq 1035 1044	1日2回（朝、夕）
T10	DrugFreqChange 1045 1047	変更
E2	DrugFreqChange:T10 NewFreq:T9 OldFreq:T8 Drug:T13	
T11	StartDrug 966 970	投与開始
E3	StartDrug:T11 NewDosage:T2 NewFreq:T8 NewDrug:T12	
T12	DrugProductName 947 953	ラミクタール
T13	DrugProductName 1024 1030	ラミクタール
T14	DrugName 1082 1087	ステロイド
T15	StartDrug 1087 1089	投与

図 2 アノテーション情報ファイルの例

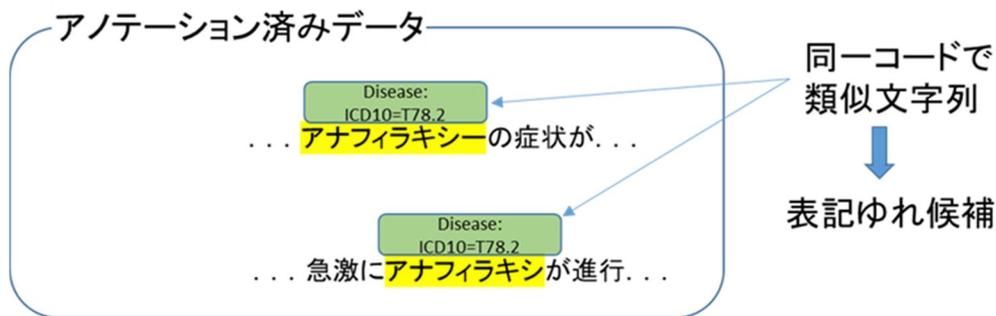


図3 アノテーション機械学習用データからの表記ゆれ吸収学習

(A)

<u>アナフィラキシー</u>	<u>変形性膝関節症</u>	<u>浮腫</u>
Anaphylaxis アナフィラキシ アナフィラキシ- アナフィラキシーー アナフィラシキー アナフィラシーショック アナフィラキシ性ショック アナキラフィシー アナヂラキシーショック アナフェラキシーショック	変形性膝関節炎 変形性膝関節疾患 変形性膝関節足 変形性膝関節大腿 gonarthrosis gonarthrosis膝関節症	浮腫 ふしゆ エデーマ 蓄水症 腫脹 ふしゆ 浮腫み 膨隆 膨瘤

(B)

<u>吐き気</u>	<u>めまい</u>	<u>息苦しい</u>
嘔吐 悪心・嘔気 悪心 はきけ 嘔気 おう吐 おうと 吐き気 むかつき 胸つかえ 吐気 もよおす 嘔気 船酔い感 嘔吐症 胸焼け 嘔気症 げっぷ	眩暈 立ちくらみ感 目眩 身体揺らぎ 動揺感	息切れ 胸部圧迫感 動悸 呼吸難感 胸苦しい 動機息切れ 呼吸困難 息ぐるし 窒息感 息切れ感 呼吸困難感 胸圧迫感 酸欠感

図4 類似度評価器により検出された病名 (A) および症状 (B) の類義

医薬品副作用データベース(JADER)を用いた対象副作用に関連する因子の同定に関する研究

研究代表者 今任 拓也 国立医薬品食品衛生研究所・医薬安全科学部・主任研究官

研究分担者 斎藤 嘉朗 国立医薬品食品衛生研究所・医薬安全科学部・部長

研究要旨:【目的】本研究は、製薬企業などから報告される副作用個別症例報告の情報を用い、人工知能によって、重症薬疹であるスティーブンス・ジョンソン症候群(SJS)および中毒性表皮壊死融解症(TEN)を判別するシステムの構築に先立ち、大規模データベースである医薬品副作用データベース(JADER)の情報を基に副作用判別モデルの最適化のための特徴的な因子を探索し、その因子の判別能を検討することを目的としている。【方法】データソースには、2018年4月版のJADERを用い、ICH国際医薬用語集日本語版のPT用語に基づき、SJS、皮膚粘膜眼症候群、TEN、好酸球増加と全身症状を伴う薬疹(DIHS)および多形紅斑(EM)の報告症例を抽出し、考えられる因子(年齢、性別、予後不良者、発症までの期間、被疑薬)について比較検討し、それら因子のSJSおよびTENの判別能について評価した。【結果】DIHSでは、年齢および発症までの期間、EMでは、年齢および性別、TENでは、予後不良者数が他の重症薬疹と比べて特徴的であった。被疑薬では、ラモトリギン、アロプリノール、カルバマゼピンは、すべての重症薬疹において、上位被疑薬として挙がっていたが、総合感冒薬(一般)およびロキソプロフェンナトリウム水和物は、SJSおよびTENにのみ上位被疑薬となっていた。さらに、これらの因子のSJSおよびTENの判別能の評価では、予後不良者および被疑薬(総合感冒薬(一般)およびロキソプロフェンナトリウム水和物の有無)では、他の因子に比べて、やや判別能は高くなっていた。また、すべての因子をモデルに導入した場合が最も判別能が高かった。【考察】本研究で使用する副作用報告個別症例の経過欄の臨床情報に今回得られた知見を追加することで、重症薬疹の評価判定モデルの精度を上げることが可能と期待され、今後検討していく予定である。

A. 研究目的

医薬品医療機器法の第六十八条の十に規定される副作用等の報告はファーマコビジランスの枠組みにおいて依然として大きな柱となっている。この医薬品医療機器法等に基づく副作用症例報告は、症例報告書の形で医薬品医療機器総合機構(PMDA)に報告され、情報の管理・整理がなされている。近年の国内副作用の年間報告件数は約5万件にもおよび、その膨大な副

作用報告の評価は、現在、人的作業により行われている。このような大量のデータを処理するには、人的・時間的リソースにも限界があり、データ処理を支援する仕組みの構築が必要である。

2006年ごろからディープラーニングなどの新たな技術が開発され、人工知能は現在、医療の分野においても大きな期待が寄せられている。人工知能は、人の知的な処理を機械によっ

て行う技術であり、人の意思決定に代わるものとなる。しかしながら、実務は、複雑であり、100%の条件を機械に与えることは困難であるため、データから様々な傾向や特徴を人手、あるいは、自動で抽出し、最適化することが人工知能の応用のための大きな課題の一つを考えられる。

そこで、本研究は、人工知能を用いて、PMDAにおいて管理されている副作用個別症例報告から、スティーブンス・ジョンソン症候群 (SJS) および中毒性表皮壊死融解症 (TEN) の試行的判定評価システムの構築に先立ち、副作用個別症例報告を同様の情報に基づき作成されている医薬品副作用データベース (JADER) を用い、SJS および TEN に特徴的な因子を探索し、SJS および TEN の判別能について検討を行うことを目的とした。

B. 研究方法

データソース

データソースは、医薬品医療機器総合機構 (PMDA) の医薬品副作用データベース (Japanese Adverse Drug Event Report database; JADER) を用いた。JADER は、PMDA のホームページ (<https://www.pmda.go.jp/safety/info-services/drugs/adr-info/suspected-adr/0003.html>) よりダウンロードした 2018 年 4 月版を用いた。JADER のデータは、図 1 に示すように、症例一覧テーブル (demo)、医薬品情報テーブル (drug)、副作用情報テーブル (reac) および原疾患テーブル (hist) から構成されており、それぞれのテーブルの識別番号を基に統合することが可能である。本研究では、症例一覧データ、医薬品情報データ、副作用情報テーブル

統合し、5,745,017 症例のデータを作成した。

対象副作用

対象副作用は、ICH 国際医薬用語集日本語版 Medical Dictionary for Regulatory Activities Terminology bundled with Japanese Translation (MedDRA/J) に記載されている基本語 (Preferred Term; PT) に基づき、副作用名を重症薬疹として分類した (表 1)。ただし、多形紅斑重症型 (Erythema Multiforme major) は、MedDRA の PT 分類では、抽出することが困難であったため、多形紅斑 (EM) を対象副作用とした。

表 1 対象とした重症薬疹分類とその PT 分類名

重症薬疹分類	MedDRA PT
SJS	スティーブンス・ジョンソン症候群 皮膚粘膜眼症候群
TEN	中毒性表皮壊死融解症
DIHS	好酸球増加と全身症状を伴う薬疹
EM*	多形紅斑

*多形紅斑重症型は PT 分類名に該当なし

統計解析

報告症例の基本属性の比較には、カイ二乗検定を用いた。本研究の対象副作用は、SJS および TEN であるため、SJS/ TEN とそれ以外の重症薬疹との判別を検討するため、ロジスティック回帰分析を用い、それぞれの解析モデルの Receiver operating characteristic curve (ROC) 曲線下面積 (Area under the curve; AUC) を算出した。さらに、解析モデルによる ROC 曲線の図示には、R for Windows の pROC パッケージを用いた。統計学的有意差は、 $P < 0.05$ とした。すべての解析には R for Windows ver.3.4.4 を用いた。

C. 研究結果

本研究で使用したデータベースから、重症薬疹の報告症例を抽出したところ、16,688 症例が抽出された。そのうち、SJS は、5,640 症例 (33.80%)、TEN は、2,683 症例 (16.08%)、DIHS は、3,647 症例 (21.85%)、EM は、4,718 症例 (28.27%) であった。

それぞれの重症薬疹報告症例の基本属性は、表 2 に示した。年齢については、20 歳以下 (未成年)、20 歳代から 50 歳台 (成人)、60 歳台以上 (高齢者) の 3 群に分けると、DIHS 報告症例は、成人群が 49.65% と他の重症薬疹と比べて、多くなっていた。また、EM では、高齢者の報告割合が、50.81% と他の重症薬疹と比べて最も多くなっていた。性別では、SJS、TEN および DIHS では、男女比が 1:1 となっているのに対し、EM では、男性が 40.22%、女性が 59.78% とやや女性の割合が高くなっていた。次に、JADER の転記の項目に基づき、「後遺症あり」、「死亡」および「未回復」を予後不良者と定義した。予後不良者の割合は、EM で 2.31% と他の重症薬疹と比べて、圧倒的に少なく、SJS および DIHS では、およそ 10% 程度となっているのに対し、TEN では、27.36% と圧倒的に予後不良者の割合が高くなっていた。また、発症までの期間については、SJS、TEN および EM では、発症までの平均期間は、およそ 30 日程度であるのに対し、DIHS では、42.49 日と長くなっていた。さらに中央値を比較すると、SJS、TEN および EM は、10 日程度であるのに対し、DIHS は、平均発症期間と同様に 27 日と長くなっていた。

さらに我々は、それぞれの重症薬疹ごとの被疑薬について検討した。SJS では、ラモトリギン、TEN では、アロプリノール、DIHS では、カルバマゼピン、EM では、ソラトフェニブトシル

酸塩が最も多くなった。それぞれの重症薬疹ごとの上位 6 位までの被疑薬を比較すると、カルバマゼピン、アロプリノール、ラモトリギンがすべての重症薬疹の上位被疑薬として挙がっていた。また、ロキソプロフェンナトリウム水和物および総合感冒薬 (一般) は、SJS および TEN のみ上位被疑薬として挙がっていた。

そこで、重症薬疹を本研究の対象としている SJS および TEN とそれ以外の重症薬疹 (DIHS、EM) に分け、性別 (男性、女性)、年齢区分 (未成年、成人、高齢者)、予後不良、発症までの期間 (10 日以内または 10 日以上) および被疑薬 (ロキソプロフェンナトリウム水和物および総合感冒薬 (一般) の有無) の SJS および TEN の判別能を評価することとした (表 4)。性別、年齢区分、発症までの期間の AUC は 0.500 程度と精度は良くなかった。予後不良および被疑薬の AUC は、0.550 程度で精度が高いとは言えないが、他の因子の AUC と比べるとその精度は高かった。次に、やや精度が高かった予後不良と被疑薬をモデルに投入し、その AUC を検討したところ、0.5581 と個々と比べると高くなっていたものの、精度に大きな変化は認められなかった。さらに、すべての因子をモデルに導入したところ、その AUC は、0.6114 と個々の因子の AUC と比べて、精度が高くなった (図 2)。

D. 考察

本研究は、PMDA において管理されている副作用個別症例報告を用いた人工知能による SJS および TEN の試行的評価に先立ち、JADER を用いて、本研究の対象副作用である SJS および TEN と他の重症薬疹 (DIHS および EM) との判別のための特徴的な因子を探索し、SJS および TEN の判別能を評価した。

TEN では、予後不良者が他の重症薬疹と比べて多くなっていた。また、DIHS は、発症までの平均期間が 40 日程度、中央値では、約 30 日とそれ以外の重症薬疹と比べて、長くなっていた。EM 以外の重症薬疹では、性差は認められなかったのに対し、EM では、やや女性が多い傾向が認められた。さらに、それぞれの重症薬疹の上位被疑薬を見てみると、カルバマゼピン、アロプリノール、ラモトリギンは、すべての重症薬疹の上位被疑薬となっていた。総合感冒薬(一般)およびロキソプロフェンナトリウム水和物は、SJS および TEN にのみ上位被疑薬として挙がっていた。さらに、それぞれの重症薬疹で特徴的であった因子を使用し、SJS/TEN とそれ以外との重症薬疹との判別を試みたところ、すべての因子を使用した場合は、やや判別能は高くなり、SJS/TEN 症例の判別に有用な要因となり得ることが示唆された。

SJS の 診 断 基 準 (2016) (日皮会誌 2016;126(9) : 1637-1685) によると、必須となる主要所見は、皮膚粘膜移行部(眼, 口唇, 外陰部など)の広範囲で重篤な粘膜病変(出血・血痂を伴うびらん等)がみられること、皮膚の汎発性の紅斑に伴って表皮の壊死性障害に基づくびらん・水疱を認め、軽快後には痂皮, 膜様落がみられ、その面積は体表面積の 10%未満ではあること、発熱があること、多形紅斑重症型(EM major)を除外できることとなっている。また、TEN では、水疱・びらんなどの表皮剥離は体表面積の 10%以上となっているが、概ね SJS と同様となっている。DIHS の診断基準では、発熱があることは SJS と同様であるが、限られた薬剤投与後に遅発性に生じ、急速に拡大する紅斑が認められること、さらに、肝機能障害や白血球過多などの血液学的異常が認められるこ

とも基準となっている。今回使用した JADER の情報には、検査値のデータはなく、発熱の有無、詳細な皮膚の性状など臨床所見は存在しない。そのため、特徴的な因子は認められたものの、それら因子のみで判別することは非常に困難と考えられる。

本研究は、JADER に含まれている情報の基となっているより詳細な副作用個別症例報告を使用する。まずは、重症薬疹の診断基準に基づく判定を基本とし、今回得られた特徴的な因子は、基本となる判定基準を補完する因子として使用することで、より高精度な判定が可能となるものと考えられる。また、一般的に複数の薬剤を服用している可能性が高く、JADER のような大規模なデータの解析結果により、被疑薬を特定することが可能となり、副作用の判別に寄与できるものと考えられる。今後、副作用症例報告書の詳細情報と本結果で抽出された要因を組合せ、人工知能を用いた SJS/TEN 症例の判別に関する解析を進め、判定の最適化を行う予定である。

E. 結論

それぞれの重症薬疹において、特徴的な因子が認められたものの、今回の検討において、認められた因子だけでは、SJS/TEN のみを判別することは困難であることがわかった。今後、研究を進めるにあたり、主体となる副作用個別症例報告から得られた因子に加え、今回得られた情報を追加することで、より精度の高い副作用評価システムの構築できるかどうか、さらなる検討を継続していく予定である。

G. 研究発表

1. 論文発表

- 1) Imatoh T, Sai K, Fukazawa C, Hinomura Y, Nakamura R, Okamoto-Uchida Y, Segawa K, Saito Y. Association between infection and severe drug adverse reactions: an analysis using data from the Japanese Adverse Drug Event Report database. Eur J Clin Pharmacol. 2017; 73(12):1643-1653.

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

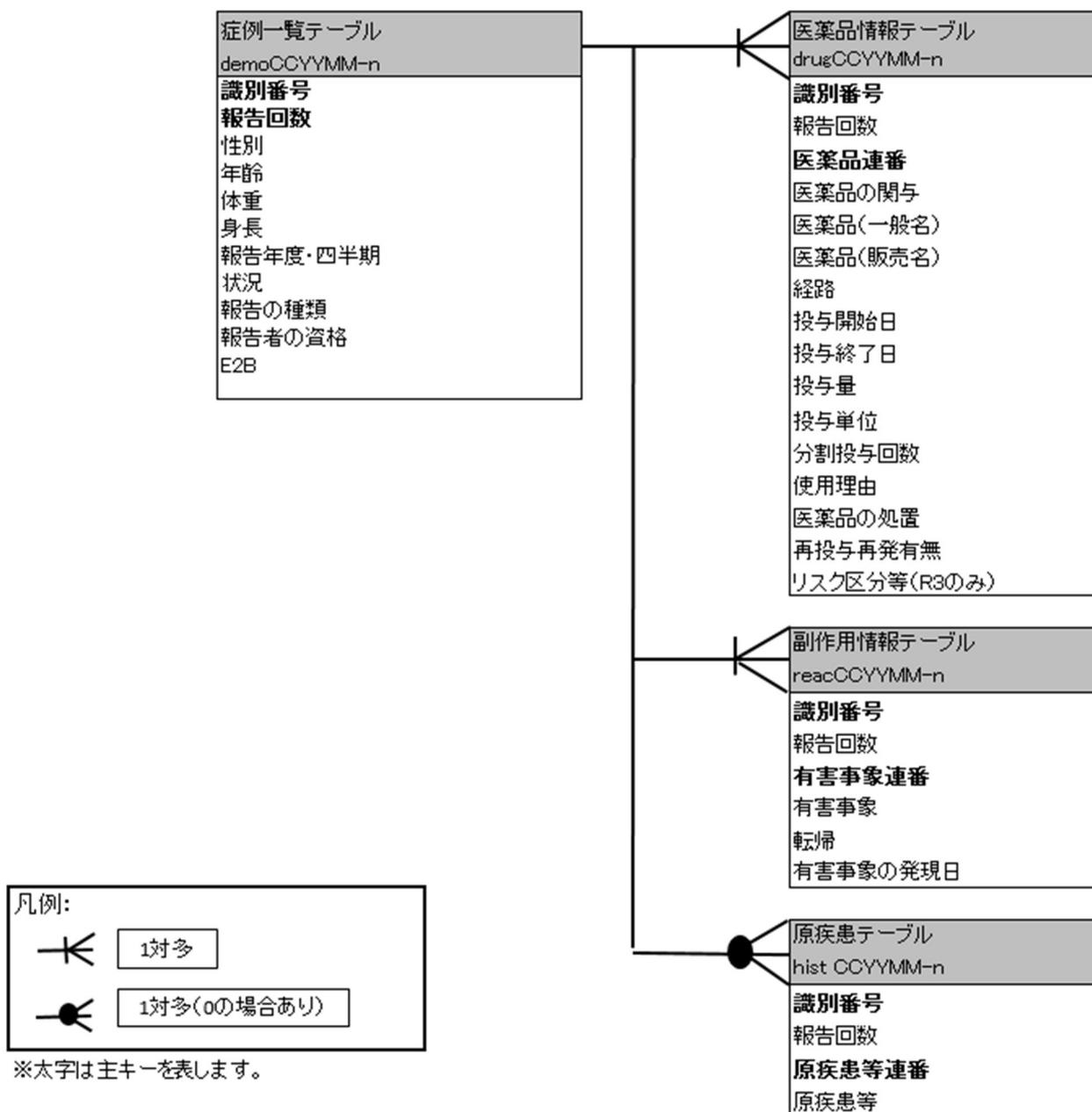


図1 JADERの実体関連図 (PMDAのホームページより抜粋)

表 2. 重症薬疹別の報告症例の基本属性

		SJS		TEN		DIHS		EM	
人数		5640	(33.80)	2683	(16.08)	3647	(21.85)	4718	(28.27)
年齢、 人数 (%)	10 歳未満	331	(5.87)	144	(5.37)	157	(4.30)	297	(6.29)
	10 歳代	246	(4.36)	124	(4.62)	189	(5.18)	135	(2.86)
	20 歳代	456	(8.09)	174	(6.49)	290	(7.95)	254	(5.38)
	30 歳代	628	(11.13)	283	(10.55)	395	(10.83)	466	(9.88)
	40 歳代	539	(9.56)	233	(8.68)	549	(15.05)	456	(9.67)
	50 歳代	745	(13.21)	390	(14.54)	536	(14.70)	665	(14.09)
	60 歳代	999	(17.71)	464	(8.68)	707	(19.39)	1128	(23.91)
	70 歳代	956	(16.95)	501	(18.67)	468	(12.83)	875	(18.55)
	80 歳代	453	(8.03)	294	(10.96)	243	(6.66)	318	(6.74)
	90 歳代	58	(1.03)	26	(0.97)	34	(0.93)	28	(0.59)
	100 歳代	1	(0.02)	1	(0.15)	0	(0.00)	1	(0.02)
	高齢者	14	(0.25)	2	(0.07)	5	(0.14)	10	(0.21)
	小児	10	(0.18)	0	(0.00)	0	(0.00)	4	(0.08)
	成人	25	(0.44)	12	(0.45)	8	(0.22)	6	(0.13)
乳幼児	0	(0.00)	0	(0.00)	0	(0.00)	2	(0.04)	
不明	179	(3.17)	32	(1.19)	66	(1.81)	73	(1.55)	
性別、 人数 (%)	男性	2624	(47.33)	1255	(47.89)	1836	(50.91)	1881	(40.22)
	女性	2920	(52.67)	1409	(52.89)	1770	(49.08)	2796	(59.78)
予後不良者、 人数 (%)		502	(8.90)	734	(27.36)	346	(9.49)	109	(2.31)
発症期間	平均値 (SD)	29.83	(139.19)	30.82	(148.48)	42.49	(140.04)	31.42	(149.13)
	中央値	9		10		27		9	

表 3-1 対象副作用別上位被疑薬

SJS		TEN	
被疑薬	件数	被疑薬	件数
ラモトリギン	438	アロプリノール	185
カルバマゼピン	320	アセトアミノフェン	149
アロプリノール	308	ロキソプロフェンナトリウム水和物	105
総合感冒薬 (一般薬)	190	ラモトリギン	89
ロキソプロフェンナトリウム水和物	177	総合感冒剤 (一般薬)	86
アセトアミノフェン	170	レボフロキサシン水和物	70

表 3-2 対象副作用別上位被疑薬

DIHS		EM	
被疑薬	件数	被疑薬	件数
カルバマゼピン	986	ソラフェニブトシル酸塩	388
アロプリノール	477	ラモトリギン	238
ラモトリギン	312	セレコキシブ	224
サラゾスルファピリジン	195	アモキシシリン水和物	181
メキシレチン塩酸塩	193	カルバマゼピン	158
フェニトイン	147	クラリスロマイシン	117

表 4. SJS/TEN の判別能の評価

	AUC	95%CI	
性別	0.5119	0.5040	0.5195
年齢区分	0.5023	0.4943	0.5103
予後不良	0.5471	0.5425	0.5516
発症までの日数	0.5110	0.5067	0.5153
被疑薬	0.5542	0.5440	0.5644
Model 1	0.5581	0.5522	0.5639
予後不良 被疑薬			
Model 2	0.6114	0.5999	0.6229
性別 年齢区分 予後不良 発症までの日数 被疑薬			

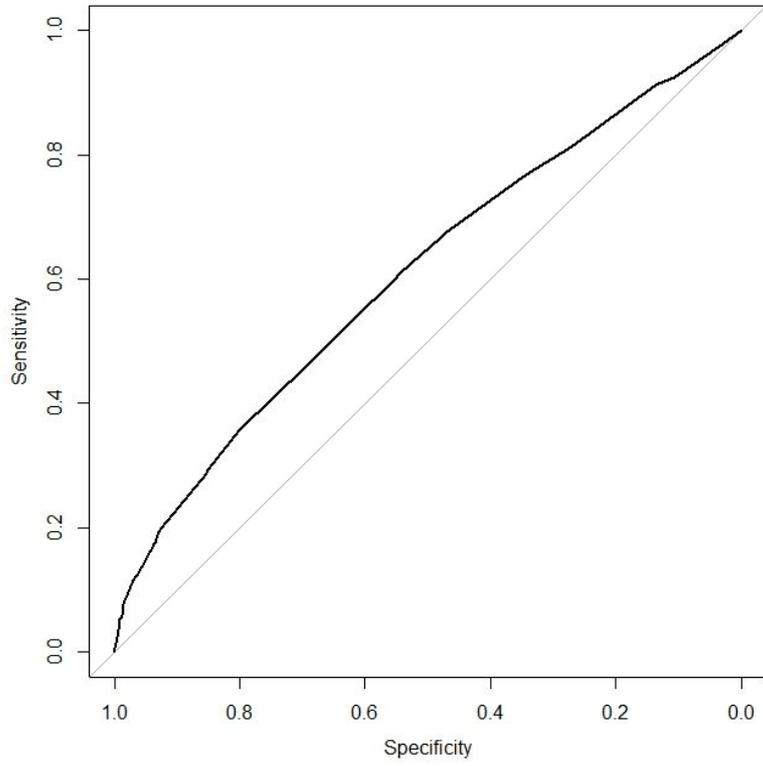


図 2. 表 4 Model12 の ROC 曲線

別添5

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書籍名	出版社名	出版地	出版年	ページ
該当なし							

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Imatoh T, Sai K, Fukazawa C, Hinomura Y, Nakamura R, Okamoto-Uchida Y, Segawa K, Saito Y.	Influence of Japanese Regulatory Action on Denosumab-Related Hypocalcemia Using Japanese Adverse Drug Event Report Database.	Eur J Clin Pharmacol.	73(12)	1643-1653	2017