

# 審査の結果の要旨

氏名 藤原 豊史

本研究は、既存データベースが提供する希少疾患関連症状の不足、およびそれらデータベースを利用する従来の希少疾患ゲノム変異解釈支援システムに対して、PubMed に収載されている症例報告を対象とした希少疾患関連症状の自動取得、およびそれら関連症状を活用する希少疾患ゲノム変異解釈支援システムを提案し、実際に、希少疾患関連症状を自動取得し、基礎技術を統合したシステムを開発したものである。実ユーザとの緊密な共同研究を行うことで、独創性の高いシステムを開発し、複数の希少疾患関連プロジェクトで同システムが採用されている。

論文は、以下の4つの章から構成されている。

第1章では、希少疾患が疑われる未診断患者を対象に行われているゲノム解析を用いた診断、および、ゲノム解析における患者の症状を利用した変異解釈に関する背景を簡潔に整理している。既存データベースが提供する希少疾患関連症状の不足、および、それらデータベースを利用する従来の希少疾患ゲノム変異解釈支援システムの欠陥を指摘し、問題の対処には、症例報告を大規模に用いた希少疾患関連症状の自動取得、およびそれら関連症状を活用する希少疾患ゲノム変異解釈支援システムの構築が有効であることを指摘して、本論文の研究背景とその意義を述べている。

第2章では、症例報告の重要性およびPubMed に収載されている症例報告を概観し、また、希少疾患関連症状を取得する上で、各種資源が整備されてきていること、とくに、遺伝性疾患の症状を表現するために整備されている Human Phenotype Ontology、および希少疾患データベース Orphanet が提供する Orphanet Rare Disease Ontology が本研究の主要な資源となっていることを説明し、それぞれの特質とそれらが本研究で果たす役割とを詳述している。そして、症例報告の書誌情報に対して希少疾患名および症状のアノテーションを付与する手法について、複数手法の精度を比較し、アノテーション結果から希少疾患関連症状を取得する方法を詳述している。また、希少疾患ゲノム変異解釈支援として提供する希少疾患ランキングの生成のために必要な、患者症状セットと疾患症状セットとの類似度計算手法を複数概観し、その中から採用した GeneYenta について詳述している。

第3章は、希少疾患関連症状の自動取得結果、利用者の視点にたった希少疾患ゲノム

変異解釈支援システム **PubCaseFinder** の開発および複数の希少疾患関連プロジェクトとの連携、自動取得した希少疾患関連症状の希少疾患ランキング精度への寄与、および希少疾患関連症状のフィルタリング手法について詳述している章である。自動取得した希少疾患関連症状は、既存データベース **Orphanet** の希少疾患関連症状を約 113% 拡張することを確認した。**PubCaseFinder** は希少疾患ランキング機能と症例報告提示機能を有しており、同システムが複数の希少疾患症例リポジトリに採用されたことは、本研究の大きな成果である。自動取得した希少疾患関連症状の希少疾患ランキング精度への寄与について、本研究で用いた定量的評価の方法を詳述し、実際の希少疾患患者データを用いて評価を行い、その結果、自動取得した希少疾患関連症状が疾患ランキング精度の向上に寄与していることを示した。また、希少疾患関連症状のフィルタリング手法について、効果が高いと報告されている出現頻度を用いた従来手法が、希少疾患においては有効でないことを指摘している。

また、**PubCaseFinder** は **GA4GH (Global Alliance for Genomics & Health)** のドライバープロジェクトである **Matchmaker Exchange** に採用され、新規疾患の定義、新規疾患原因遺伝子の同定に活用されていることは、文献処理と知識処理との融合の一つの形態が、希少疾患分野の研究を推進させるために有効であることを指摘している。

第 4 章は、本研究の貢献を、1) 症例報告を大規模に活用し希少疾患関連症状を自動で収取したこと、2) 自動収取した希少疾患関連症状が希少疾患ランキング精度向上に寄与すること、の 2 点に整理し、アノテーション手法、希少疾患関連症状のフィルタリング手法、症状セット類似度計算手法、**PubCaseFinder** と日本人ゲノム情報統合データベース **TogoVar** との統合化について、今後の研究課題を整理している。

以上のように、本論文は、症例報告から自動取得した希少疾患関連症状の希少疾患ゲノム変異解釈支援システムに対する有効性を示し、また、構築した **PubCaseFinder** は複数の希少疾患関連プロジェクトに採用されており、審査委員会は、その独創性、有効性は、博士号に十分値するものと判断した。

以上 1788 字