

各位

電子カルテに記載された医師の所見を構造化されたデータに変換 ～小児医薬品の安全性監視や開発推進への利活用に期待～

2025年02月17日

株式会社ケアネット

国立成育医療研究センター

株式会社ケアネット（代表取締役社長：藤井 勝博、URL：<https://www.carenet.co.jp>、以下「ケアネット」）と、国立研究開発法人国立成育医療研究センター（理事長：五十嵐隆、URL：<https://www.ncchd.go.jp>、以下「国立成育医療研究センター」）は共同で、小児医療情報を活用した電子カルテ情報の文章データを構造的（定められた形式で整理すること）に分析する研究を行いました。

研究は、気管支喘息やアレルギー性鼻炎などに広く使用されるロイコトリエン受容体拮抗薬（LTRA）¹の処方歴があり、かつアナフィラキシー症状が発現した、国立成育医療研究センターの患者さんの電子カルテ情報を対象に進めました。

H&H CONNECT株式会社（代表取締役社長：林 哲、URL：<https://www.hh-c.co.jp>）が提供する自然言語処理技術²や医療用語の辞書データを活用し研究を行った結果、文章形式で記録されている医師の所見や診断内容などをより分析しやすい構造化されたデータに変換することに成功しました。

研究の背景

近年、臨床現場で日々蓄積される健康医療データは、従来の臨床試験データを補完・代替するデータとして重要性が高まっています。特に小児医療の分野では、医薬品や治療に関するデータが不足していることから、健康医療データの活用が期待されています。国立成育医療研究センターでは「小児と薬情報収集ネットワーク整備事業」（<https://pharmanet.ncchd.go.jp>）において、全国の小児病院やクリニックから電子カルテ情報を集め、小児医療情報収集システム（以下「小児DB」）を構築して、医薬品の使用実態や副作用など幅広い情報を蓄積し、安全性評価や適正使用の推進に取り組んでいます。

¹ ロイコトリエン受容体拮抗薬（LTRA）：気管支を収縮させる作用を持つロイコトリエンの働きを阻害することで、気管支を広げたり、炎症を抑えたりする働きがある薬です。

² 自然言語処理（NLP）技術：コンピューターが人間の言語（日本語や英語など）を理解し、分析できるようにする技術の総称です（NLP：Natural Language Processing）。

一般的に、電子カルテには患者の基本情報や処方履歴といった項目が構造化データとして整理されていますが、医師が自由に記載する所見や経過記録といった文章データは、項目が決まっていない非構造化データ³にあたります。非構造化データには、病状や患者背景など詳細な情報が数多く含まれていて、これを活用できれば健康医療データのより多角的で精密な解析が可能となります。しかし、そのためには文章をコンピューターが理解しやすい形に変換・分析する自然言語処理技術や、文章情報に注釈（タグ付け）を施すアノテーション技術⁴、さらに診断名や症状名などを標準的に整理した構造化用語辞書⁵といった仕組みを整備する必要があり、これまで大きな技術的ハードルとなっていました。

研究の内容

本研究では、自然言語処理技術とアノテーション手法を組み合わせ、さらに構造化用語辞書を活用することで、非構造化データを構造的に分析しました。その結果、電子カルテに記録された医師の所見や診断内容などの文章情報（非構造化データ）を、より分析しやすい形式に効率的に変換することに成功し、この手法の有用性と汎用性も確認することができました。従来の構造化データと統合した解析用データセットを生成することで、より精密な分析が可能となり、小児医療分野における健康医療データ活用のさらなる発展に寄与することが期待されます。

今後の展開について

今回の研究を通じて、非構造化データを利活用することにより医薬品などによる有害事象の原因を一定の割合で特定できる可能性があることが明らかになりました。今後、非構造化データを構造化させ健康医療データの情報をより充実させることにより、健康医療データを活用した小児医薬品の安全性監視や開発がより一層発展することが期待されます。

今回得られた知見を活かし、他の疾患領域への展開や健康医療データのさらなる活用方法を検討することで、医療現場や製薬企業が抱える課題の解決、日本の小児医療を含めた医療全体の発展に貢献していきたいと考えています。

－本リリースに関する問い合わせ先－

株式会社ケアネット カスタマーリレーショングループ

フォームよりお問い合わせください。 <https://www.carenet.co.jp/contact/>

国立成育医療研究センター 企画戦略局 広報企画室 神田・村上

電話：03-3416-0181（代表） E-mail:koho@ncchd.go.jp

³ 非構造化データ：自由形式で記載されたデータの総称です。医療では、カルテの所見や医師のコメントなど、定型化されていない記述が該当します。

⁴ アノテーション技術：文章の中の特定の情報にタグ（注釈）を付ける技術で、自然言語処理（NLP）の精度向上に役立ちます。

⁵ 構造化用語辞書：診断名や症状名などを標準化して整理した「用語の一覧表」のこと。異なる表現で書かれていても同じ意味として扱えるようにします。