令和７年度厚生労働科学研究費補助金（地域医療基盤開発推進研究事業）
分担研究報告書

【分担研究１】LLM を用いて診療記録から不穏状態に関する情報を抽出する研究

川本　章太・宮本　太郎・井汲　沙織・園部真也

【要旨】
看護記録を LLM に入力し、せん妄の有無を判定させた。特定のキーワードを含まず文脈でせん妄を描写しているような文章においては、キーワード検索ではせん妄の検出に偽陰性が多く認められたが、LLM による判定では、これらの偽陰性が抑制され、より高い精度でせん妄が検出された。

【目的】
医療記録に基づくせん妄状態の検出について、キーワード検索と LLM 推論分析の精度を比較した。

【方法】
東北大学病院で 2024 年 3 月 1 日から 2024 年 3 月 31 日までの期間に作成された看護記録を対象とした。キーワード検索においては、せん妄に関連した 44 のキーワードのいずれかを含む看護記録を抽出した。LLM 推論分析においては、パラメーター数の小さい小規模 LLM として Llama3.2:1b、Qwen2:1.5b、Phi3:mini を、パラメーター数の大きい大規模 LLM として Gemini 27B を採用し、それぞれのモデルを用いて看護記録から当該患者がせん妄であるかどうか判定させた。
観察研究であり、オプトアウト方式で実施した。

【結果】
解析の対象となった看護記録は 9,107 件であった。キーワード検索では、209 件（2.3%）の看護記録が抽出された。LLM 推論分析では、Llama3.2:1b が 4,545 件（49.9%）を、Qwen2:1.5b が 0 件（0.0%）を、Phi3:mini が 89 件（1.0%）を、Gemini 27B が 806 件（8.9%）の看護記録において当該患者をせん妄と判定した。せん妄の描写があると Gemini 27B が判定した看護記録の内、キーワード検索において抽出されなかった看護記録の大部分は、せん妄を描写していると臨床的観点から判断される看護記録であった。

【考察】
せん妄の状態にある患者の看護記録は、特定のキーワードを含まず、文脈でせん妄を描写していることがある。このような看護記録は、キーワード検索では抽出されないが、大規模 LLM では検出されており、大規模 LLM は文脈を考慮した評価が出来たと言える。一方で、小規模 LLM は検出結果に過剰や不足が目立ち、文脈を考慮した患者状態判定の性能は限定的と言える。このことから、看護記録のような医学文書の文脈を認識する上で、パラメーター数の多寡が重要であることが示唆される。本研究で使用した大規模 LLM は、システムの規模という観点から現実的にローカルで実装および運用が可能であり、医療において求められるコストおよびセキュリティの要件を満たしやすい点も特筆すべき点である。

【結論】
大規模 LLM を用いた LLM 推論分析により、看護記録からせん妄の有無を判定するような、文脈を考慮した患者状態の判定が、従来手法よりも高い精度で可能となる。これらの手法は現実的に実装可能であり、患者に対する正確な病状把握と適切な治療介入のための支援ツールとして有効と考えられる。

【論文発表】
該当なし

【学会発表】
川本章太，宮本太郎，井汲沙織，園部真也．看護記録に基づくせん妄患者の大規模言語モデルを用いた抽出．第 45 回医療情報学連合大会．

【知的財産権の出願・登録状況】
該当なし