

論文内容の要旨

Machine Learning-Based Prediction of In-Hospital Falls in Adult Inpatients:
Retrospective Observational Multicenter Study

入院成人患者における院内転倒の機械学習による予測：多施設後ろ向き観察研究

日本医科大学大学院医学研究科 医療管理学分野

研究生 西野 拓也

発表誌名 JMIR Medical Informatics

(巻・号・年月日) 13, e75958, 2025 年 12 月 4 日

【背景】

入院患者の転倒は、入院期間の延長や医療費の増加、さらには患者の生活の質の低下を引き起こす重要な医療安全上の問題である。高齢患者では身体機能低下や多剤併用など複数の要因が複雑に関与するため、転倒リスクの評価は容易ではない。多くの医療機関では転倒リスク評価スコアを用いたスクリーニングが行われているが、これらの方法は患者背景や臨床状況の多様性、リスク因子間の相互作用を十分に反映できない可能性がある。近年、電子カルテデータを用いた機械学習モデルは、多数の臨床変数の相互関係を解析できることから、転倒リスク予測への応用が期待されている。しかし、既存研究では識別性能の評価が中心であり、臨床応用に重要なキャリブレーション性能については十分な検討が行われていない。

【目的】

本研究の目的は、電子カルテおよび診療データベースを用いて入院患者の転倒を予測する機械学習モデルを開発し、その識別性能およびキャリブレーション性能を検証することである。さらに、モデル解釈手法を用いて転倒に関連する臨床因子を明らかにし、転倒発生状況の詳細解析を通じて実践的な予防戦略への示唆を得ることを目的とした。

【方法】

本研究は、日本医科大学付属病院および千葉北総病院の2施設において実施した多施設後ろ向き観察研究である。2018年4月から2023年3月までに入院した患者のうち、65歳以上かつ入院期間3日以上を対象とした。診断群分類（Diagnosis Procedure Combination: DPC）データおよび電子カルテから患者背景、併存疾患、薬剤情報、看護必要度、血液検査結果などを取得した。転倒は入院後30日以内に発生した最初の転倒と定義し、インシデントレポートを用いて判定した。

候補変数の中から機械学習による特徴量選択により30変数を抽出し、データを訓練データ80%、検証データ20%に分割した。予測モデルとしてロジスティック回帰、Extreme Gradient Boosting (XGB)、Light Gradient Boosting Machine (LGBM)、Categorical Boosting (CatBoost) の4種類を構築した。クラス不均衡への対応としてSynthetic Minority Oversampling Technique (SMOTE) を適用し、ハイパーパラメータ最適化にはOptunaを用いた。モデル評価はF1スコア、受信者動作特性曲線下面積（ROAUC）、Precision-Recall 曲線下面積（PRAUC）、Brier スコア、キャリブレーションスロープを用いて行った。さらにSHAP（Shapley Additive Explanations）を用いてモデルの解釈を行った。

【結果】

対象患者 159,274 例のうち除外基準を適用した結果、83,917 例が最終解析対象となった。転倒は 2,173 例（2.6%）に発生した。モデル性能の比較では、CatBoost が最も高い F1 スコアおよび PRAUC を示し、転倒高リスク患者の識別に優れていた。一方、LGBM はキャリブレーションスロープが理想値に最も近く、予測確率の信頼性に優れていた。SHAP 解析では、血清アルブミン低値、移乗能力の低下、悪性腫瘍の存在、鎮静薬・睡眠薬の使用、糖尿病治療薬の使用などが転倒リスクに大きく寄与する因子として同定された。さらに転倒インシデント 435 件の詳細解析では、転倒の約半数が排泄行動に関連しており、特に早朝 4～6 時に発生頻度が高かった。また、高リスク群ではベッドサイドでの転倒が多く認められた。

【考察】

本研究では、機械学習を用いた転倒予測モデルが良好な識別性能およびキャリブレーション性能を示すことを明らかにした。特にブースティングアルゴリズムはロジスティック回帰よりも高い予測性能を示し、複雑なリスク因子の相互作用を捉える能力が示唆された。臨床応用の観点では、リスクの高い患者を効率的に抽出する目的には CatBoost が有用であり、介入閾値に基づく確率評価にはキャリブレーション性能に優れる LGBM が適していると考えられる。また、転倒の多くが排泄行動やベッド周囲での移動に関連していたことから、夜間の排泄介助やベッドサイド環境の整備など、個別化された予防介入の重要性が示唆された。

【結論】

本研究では、入院患者の転倒リスクを予測する機械学習モデルを開発し、その有用性を示した。特に CatBoost および LGBM は高い予測性能を示し、転倒リスク評価に有用である可能性が示唆された。電子カルテシステムにこれらのモデルを統合することで、リアルタイムの転倒リスク評価と個別化された予防介入が可能となり、入院患者の安全性向上に寄与することが期待される。