Chirurgia (Bucur). 2026;121(5):489-499
http://dx.doi.org/10.21614/chirurgia.3379

Artificial Intelligence - Driven Risk Stratification in Emergency Surgery: Development and Validation of the PROMETHEUS-AI Model

Cătălin Cosma[1,2], Vlad Olimpiu Butiurca[1,2], Cosmin Nicolescu[1,2], Marian Botoncea[1,2], Călin Molnar[1,2]
[1]George Emil Palade University of Medicine, Pharmacy, Science and Technology of Târgu Mureș, Romania
[2]General Surgery, Emergency County Hospital of Târgu Mureș, Romania

Abstract
Background: Emergency General Surgery is associated with substantial postoperative morbidity, and accurate perioperative risk stratification remains challenging. Artificial intelligence and machine learning may improve risk prediction by integrating multiple demographic, clinical, laboratory, and operative variables. This study aimed to develop and internally validate PROMETHEUS-AI, an artificial intelligence–driven framework for predicting postoperative complications in Emergency General Surgery.
Methods: A retrospective single-center observational cohort study included 250 adult patients who underwent emergency surgery for acute appendicitis, acute cholecystitis, or non-oncological intestinal obstruction between 2020 and 2025. Demographic, clinical, laboratory, and operative variables were incorporated into Logistic Regression, Random Forest, and Extreme Gradient Boosting (XGBoost) models. Internal validation was performed using an 80:20 training–validation split. Model performance was assessed using receiver operating characteristic area under the curve (ROC-AUC), accuracy, sensitivity, specificity, precision, and F1-score. Model interpretability was evaluated using SHapley Additive exPlanations (SHAP).
Results: Postoperative complications occurred in 48 patients (19.2%), while major postoperative complications (Clavien-Dindo grade IIIa–V) occurred in 17 patients (6.8%). XGBoost demonstrated the highest predictive performance, with a ROC-AUC of 0.88, accuracy of 85.0%, sensitivity of 83.0%, specificity of 86.0%, precision of 81.0%, and an F1-score of 0.82. Random Forest achieved a ROC-AUC of 0.84, while Logistic Regression achieved a ROC-AUC of 0.78. SHAP analysis identified serum albumin, age, ASA physical status classification, C-reactive protein, serum lactate, and neutrophil-to-lymphocyte ratio as the six highest-ranked predictors, followed by operative duration, hemodynamic instability, sepsis on admission, and Clinical Frailty Scale.
Conclusions: PROMETHEUS-AI demonstrated promising internally validated performance for postoperative complication risk stratification in Emergency General Surgery, with XGBoost showing the highest discrimination among the evaluated algorithms. These findings support the feasibility of explainable machine learning for perioperative risk assessment. However, given the retrospective single-center design, limited sample size, and absence of external validation, PROMETHEUS-AI should be regarded as an exploratory proof-of-concept model requiring prospective multicenter and external validation before clinical implementation
Keywords: artificial intelligence, emergency surgery, machine learning, risk stratification, postoperative complications, XGBoost, predictive modeling