Machine Learning for Revision Total Knee Arthroplasty: Identifying Patient-Specific Risk Factors for Adverse Outcomes

Jayalakshmi Jain, Jibreel Hussain, Nezar Abunnur, Apurva Swapnil Choubey, Esa Syed, Malyika F Hussain, Yasser Farid, Mark H Gonzalez

INTRODUCTION: Revision total knee arthroplasty (rTKA) presents unique clinical challenges, often associated with higher complication rates, longer hospital stays, and increased likelihood of discharge to non-home settings compared to primary TKA. As these procedures become more common due to the aging population and expanding pool of primary TKA recipients, there is a growing need for robust, individualized risk prediction tools. Traditional risk calculators may not adequately reflect the complex interplay of factors influencing revision outcomes. This study aimed to evaluate the performance of multiple machine learning (ML) models in predicting short-term postoperative outcomes after rTKA and to identify the most influential preoperative variables driving each prediction. The ultimate objective is to develop an interpretable and practical decision-support tool for preoperative risk stratification.

METHODS: Patients who underwent rTKA between 2019 and 2023 were identified in the National Surgical Quality Improvement Program (NSQIP) database using CPT codes 27486 and 27487. After applying standard exclusions, a cohort of 20,869 patients was analyzed. Four supervised ML algorithms—XGBoost, LightGBM, Random Forest, and Elastic Net Logistic Regression—were developed to predict four key 30-day outcomes: readmission, major complications, prolonged length of stay (LOS ≥3 days), and discharge to a non-home setting. A stacked ensemble model combining all four base models was also constructed. The data were split 75% for training and 25% for testing. Model performance was assessed using the area under the receiver operating characteristic curve (AUROC). Feature importance was derived using mean SHapley Additive exPlanations (SHAP) values from the top-performing model for each outcome. For major complications, SHAP dependence plots were generated for the top four features to visualize how changes in these variables influenced model predictions.

RESULTS:

The ensemble model demonstrated superior performance for major complications (AUROC 0.78), prolonged LOS (AUROC 0.74), and non-home discharge (AUROC 0.81). For 30-day readmission, XGBoost achieved an AUROC of 0.81 for readmissions, outperforming the ensemble and other individual models. Predictive variables varied across outcomes. Readmission was most influenced by anesthesia technique, platelet count, white blood cell count, ASA classification, and hematocrit. Prolonged LOS was primarily associated with hematocrit, anesthesia type, race, platelet count, and age. Hematocrit, anesthesia type, sex, and BMI emerged as the strongest predictors for non-home discharge.

The major complications model identified age, ASA classification, sex, and anesthesia type as the top four contributors. SHAP dependence plots revealed that patients over 80 years old had the highest positive contribution to predicted risk. Use of general or epidural anesthesia also conferred elevated risk compared to other techniques. ASA class III was associated with greater predicted complication risk than classes I or II. Additionally, female sex corresponded to modestly increased risk based on SHAP value interpretation.

DISCUSSION AND CONCLUSION: In patients undergoing revision total knee arthroplasty, machine learning models demonstrated strong predictive capability for short-term postoperative outcomes, with XGBoost excelling in predicting readmissions and the ensemble model performing best for other endpoints. Key preoperative features—particularly age, ASA classification, hematologic parameters, and anesthesia technique—varied in importance by outcome, reinforcing the need for outcome-specific modeling. SHAP analysis provided interpretable insights, such as increased complication risk among elderly patients, ASA III classification, and general or epidural anesthesia use. These results support the utility of ML-based tools for personalized risk assessment in the rTKA population. To promote clinical applicability, a risk calculator based on the ensemble model was developed, offering individualized predictions to guide perioperative planning and optimization.