Tailoring Risk in Revision Total Hip Arthroplasty: Machine Learning Models for Patient-Specific Prognosis

Jibreel Hussain MBA[1], Jayalakshmi Jain MS[1], Apurva Choubey MD[2], Nezar Abunnur BS[1], Saif Khan[1], Yasser R. Farid MD PhD[2], Mark H. Gonzalez MD PhD[2], [1]University Illinois College of Medicine, Chicago, IL, [2]University of Illinois at Chicago, Department of Orthopaedics, Chicago IL
mhussa45@uic.edu

Disclosures: Jibreel Hussain MBA (N), Jayalakshmi Jain MS (N), Apurva Choubey MD (N), Nezar Abunnur BS (N), Saif Khan (N), Yasser R. Farid MD PhD (N), Mark H. Gonzalez MD PhD (N)

Introduction: Revision total hip arthroplasty (rTHA) is a technically demanding procedure with higher complication rates, extended hospital stays, and increased discharge to rehabilitation facilities compared to primary THA. As the volume of revision cases grows—driven by implant longevity, increased life expectancy, and rising primary THA rates—there is an urgent need for more accurate tools to stratify patient risk. Traditional risk calculators often lack the sophistication needed for complex surgical populations. In this study, we evaluated the performance of multiple machine learning (ML) models to predict short-term outcomes following rTHA and identified the most influential preoperative variables associated with each outcome. The broader aim is to develop a clinically relevant and interpretable ML-based tool to support perioperative risk assessment.

Methods: Patients who underwent rTHA between 2019 and 2023 were identified in the National Surgical Quality Improvement Program (NSQIP) database using CPT codes 27132, 27134, 27137, and 27138. After standard exclusion criteria were applied, 16,448 cases were included in the final analysis. Four supervised ML algorithms—XGBoost, LightGBM, Random Forest, and Elastic Net Logistic Regression—were trained to predict four 30-day postoperative outcomes: readmission, major complications, prolonged length of stay (LOS ≥3 days), and discharge to a non-home setting. A stacked ensemble model incorporating all four algorithms was constructed to improve predictive performance. The dataset was randomly split into training (75%) and testing (25%) cohorts. Model performance was evaluated using area under the receiver operating characteristic curve (AUROC). Feature importance was derived from the most predictive model for each outcome using mean SHapley Additive exPlanations (SHAP) values. For major complications, SHAP dependence plots were created for the top four features to visualize their relationship with predicted risk.

Results: The ensemble model achieved the highest overall performance across outcomes, with AUROC values of 0.74 for readmission, 0.73 for major complications, 0.78 for prolonged LOS, and 0.79 for non-home discharge. Predictive features varied depending on the outcome. For readmission, the most impactful variables were ASA classification, hematocrit, BMI, and platelet count. Prolonged LOS was primarily influenced by hematocrit, followed by anesthetic technique, race, and platelet levels. Non-home discharge was best predicted by hematocrit, anesthesia method, ASA classification, and platelet levels.

For major complications, the leading predictors were age, sex, ASA classification, and anesthesia type. SHAP dependence plots revealed a positive association between advanced age and complication risk, particularly among patients aged 70 and older. ASA class III also conferred elevated predicted risk compared to lower ASA classifications. General anesthesia was associated with the highest SHAP values among anesthetic techniques, suggesting greater risk than other modalities. Additionally, female patients demonstrated modestly increased predicted risk relative to male patients.

Conclusion: In patients undergoing revision total hip arthroplasty, ensemble machine learning models showed strong performance in predicting short-term postoperative outcomes, particularly for prolonged length of stay and non-home discharge. Risk factor analysis revealed substantial variation across outcomes, with hematocrit, ASA classification, and anesthesia technique appearing consistently across multiple models. For major complications, age, sex, and anesthesia type significantly influenced predicted risk. These findings emphasize the value of outcome-specific modeling and the interpretability offered by SHAP analysis in clinical decision-making. To translate these insights into practice, we developed a clinical risk calculator that provides individualized risk estimates to support preoperative counseling and optimization in the rTHA population.

Significance/Clinical Relevance: Revision hip replacement presents unique clinical challenges due to increased patient complexity and surgical difficulty. This study leverages machine learning to help clinicians identify high-risk patients and tailor interventions accordingly.