Comparing Machine Learning Models to Predict the Development of Prosthetic Joint Infection in Total Joint Arthroplasty

Adrian Lin BS, Kole Joachim BA, Christopher Hamad MD, Brandon Gettleman MD, Othneil Sparks BS, Sumin Jeong BS, Amanda Perrotta BA, Ezekiel Dingle BS, Alexandra Stavrakis MD, Alexander B. Christ MD
[1]David Geffen School of Medicine at UCLA, Department of Orthopaedic Surgery, Los Angeles, CA
kjoachim@mednet.ucla.edu

Disclosures: Alexander B. Christ (3B-Onkos Surgical, Inc.; 3B-Smith & Nephew, Inc.; 3B-Zimmer Biomet Holdings, Inc.; 3B-Stryker; 3B-Daiichi Sankyo; 3B-Globus Medical, Inc.; 9-AAOS Hip & Knee Evaluation Committee; 9-MSTS Fellowship Committee; 9-ORS Implant Section President; 8-Arthroplasty Editorial Board)

INTRODUCTION: Prosthetic joint infection (PJI) remains a rare but severe complication following total joint arthroplasty (TJA). Accurate prediction of PJI risk using electronic health record (EHR) data may improve perioperative risk stratification. Machine learning (ML) methods have demonstrated promise in clinical prediction but must be evaluated against traditional approaches to assess clinical utility. This study seeks to characterize how two machine learning models compare to logistic regression in predicting risk of PJI after TJA.

METHODS: A retrospective cohort of patients from a single center undergoing total hip or knee arthroplasty (2012–2024) was used to train and test three predictive models: (1) logistic regression with LASSO covariate selection, (2) XGBoost, and (3) Random Forest. The dataset included demographic (age, sex, BMI), behavioral (smoking, alcohol use), and comorbidity (CCI-based) variables. Models were trained using a 70/30 train-test split with stratified sampling by outcome. Random Forest covariate importance was determined by mean decrease in Gini. Model performance was evaluated using area under the receiver operating characteristic curve (AUC).

RESULTS SECTION: The cohort included 13,526 patients (mean age 66.3 years, 55% female). The LASSO logistic regression model selected age, BMI, sex, 10 CCI-derived comorbidities, and alcohol use as covariates. Significant covariates included older age (OR=0.97, 95%CI: 0.96–0.98, p<0.001), male sex (OR=1.53, 95%CI: 1.15–2.04, p=0.004), congestive heart failure (OR=2.31, 95%CI: 1.54–3.43, p<0.001), peripheral vascular disease (OR=1.61, 95%CI: 1.12–2.28, p=0.009), dementia (OR=1.97, 95%CI: 1.02–3.52, p=0.031), and rheumatic disease (OR=1.49, 95%CI: 1.00–2.16, p=0.043). Random Forest ranked age, BMI, congestive heart failure, and peripheral vascular disease as the most important covariates. AUCs were: Random Forest (0.70), LASSO (0.69), and XGBoost (0.68)(Figure 1).

DISCUSSION: While ML models can capture non-linear relationships and potential interactions among comorbidities, all three demonstrated similar predictive performance. The LASSO logistic regression model offers the advantage of covariate-level interpretability, enabling inference of individual risk factors.

SIGNIFICANCE/CLINICAL RELEVANCE: These findings underscore the utility of interpretable, parsimonious models in clinical risk prediction, particularly when working with EHR data where self-reported and outcome data can vary substantially in quality and completeness.

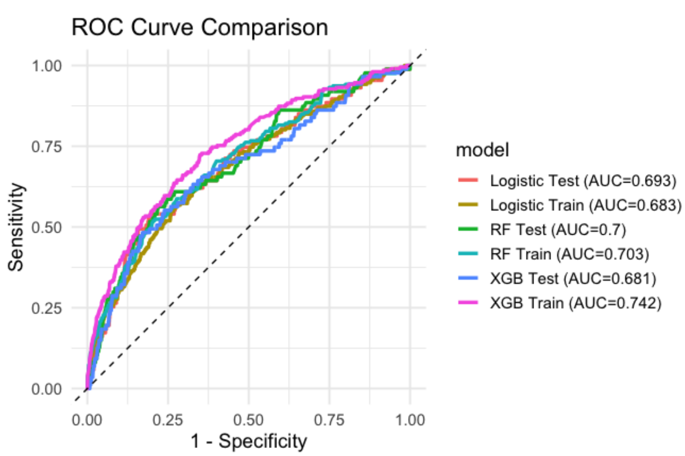

Figure 1. Receiver operating characteristic (ROC) curves comparing the predictive performance of three models for PJI following TJA.