Machine Learning Algorithms Accurately Predict Same-day Discharge and Lengthened Hospital Stay Following Primary Total Knee Arthroplasty

Shane Fei Chen MA[1], Tony Lin-Wei Chen MD, PhD[1], Anirudh Buddhiraju, MD[1], Michelle Shimizu BSc[1], Henry Hojoon Seo BA[1], MohammadAmin RezazadehSaatlou MD[1], Christian A. Pean MD[1], Oh-Jak Kwon[1], Jona Kerluku BSc[1], Ziwei Huang MD, PhD[1], Blake M. Bacevich BSc[1], John G. Esposito MD[1], Young-Min Kwon MD, PhD[1]
[1]Bioengineering Laboratory, Department of Orthopaedic Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA
ymkwon@mgh.harvard.edu

Disclosures: Shane Fei Chen (N), Tony Lin-Wei Chen (N), Anirudh Buddhiraju (N), Michelle Shimizu (N), Henry Hojoon Seo (N), MohammadAmin RezazadehSaatlou (N), Christian A. Pean (N), Oh-Jak Kwon (N), Jona Kerluku (N), Ziwei Huang (N), Blake M. Bacevich (N), John G. Esposito (N), Young-Min Kwon (5- MicroPort; 5- Depuy; 5- Smith & Nephew; 5- Stryker; 5- Zimmer Biomet)

INTRODUCTION: Total knee arthroplasty (TKA) is expected to exceed 1.26 million cases by 2030. Length of stay (LOS) is one substantial driver of the total health care costs. The advancement in treatment protocols has contributed to shortened LOS following TKA without compromising patient outcomes, leading to a growing consideration of ambulatory TKA (with LOS=0) in an increasing number of cases. Currently, there is a scarcity of prediction tools that can stratify patients based on estimates of different LOSs, including zero-day LOS and prolonged LOS. Therefore, the aim of this study was to develop multiclass machine learning models using a national patient cohort and assessed their performance in predicting LOS after primary TKA.

METHODS: The current study analyzed primary TKA (n = 77,444) from the American College of Surgeons National Surgical Quality Improvement Program from 2017 to 2020. LOS was categorized into zero-day LOS (LOS = 0 days), normal LOS (LOS = 1–3 days), and prolonged LOS (LOS > 3 days). Machine learning models including artificial neural network, random forest, k-nearest neighbors, and XGBoost were developed and employed to predict LOS after primary TKA. The method of permutation importance was applied to determine the contribution of each patient factor to LOS. The model performance was evaluated by prediction accuracy and the area under the receiver operating characteristic curve (AUC). We calculated the average AUC to aggregate the model's performance metrics across multiple classes.

RESULTS: In the study cohort, we identified 1,030 patients with zero-day LOS (1.33%), 70,326 patients with normal LOS (90.81%), and 6,088 patients with prolonged LOS (7.86%). Random forest delivered the best performance with a prediction accuracy of 88.8% and micro-averaged AUC of 0.97 (Figure 1). Age and ASA scores were the strongest predictors of LOS. Other contributing factors to the LOS prediction included diabetes, race, BMI, hypertension, and sex (Figure 2).

DISCUSSION: We found supportive evidence that machine learning models were accurate in predicting patients with zero-day (ambulatory), normal, and prolonged LOS following primary TKA. Patients receiving spinal anesthesia and having an ASA lower than 2 are more likely to be discharged on the same day of the surgery while those who received general anesthesia and had an ASA score > 3 are predisposed to prolonged hospital stay.

SIGNIFICANCE/CLINICAL RELEVANCE: Machine learning algorithms possess the potential to predict patients suitable for same-day discharge as well as patients at high risk of extended hospitalization.

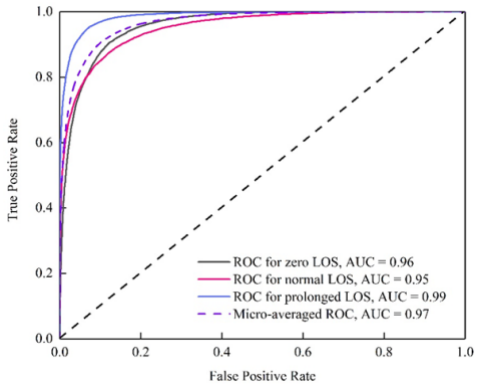

Figure 1. The receiver operating characteristics curves for the trained multiclass model of RF using a one-vs-rest multiclass strategy.

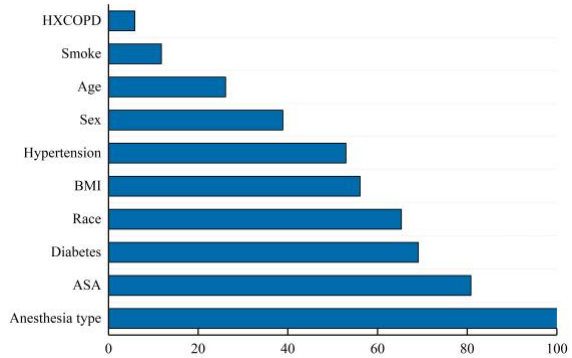

Figure 2. Top 10 important features for the prediction of LOS after primary TKA.