Paper #51

Development of Machine Learning Algorithms for Predicting Complications After Primary THA

Kyle N. Kunze, MD, Aditya V. Karhade, BS, Evan M. Polce, BS, Alex Sadauskas, MD, Joseph H. Schwab, MD, **Brett R. Levine, MD, MS**

Introduction: Complications after primary total hip arthroplasty (THA) which result in readmission or reoperation have the potential to impose a significant cost on the healthcare system with the implementation of alternative payment models. The aims of this study were to 1) develop and internally validate supervised machine learning algorithms to predict complications after total hip arthroplasty; and 2) to develop an open-access clinical decision aid capable of providing patient-level risk explanations.

Methods: This was a retrospective case-control study of institutional registry data. The primary outcome was all-cause complications at two-years after primary THA. Twelve preoperative variables were considered for prediction, including demographics, medical history, opioid use, and preoperative outcome scores. Recursive feature elimination was applied to identify the variables with the greatest predictive value. An 80:20 random sample split was used to stratify patients into training and testing sets. Five supervised machine learning algorithms were developed on the training set using 10-fold cross-validation and internally validated on the independent testing set. All algorithms were assessed by discrimination, calibration, Brier score, and decision curve analysis to quantify performance.

Results: A total of 616 patients were included. The observed complication rate was 16.6%. The stochastic gradient boosting model achieved the best performance in the independent testing set not used for algorithm development, with c-statistic=0.88, calibration intercept=0.1, calibration slope=1.22, and Brier score=0.09. The most important factors for predicting postoperative complications were age, drug allergies, prior hip surgery, smoking, and opioid use. Individual patient-level explanations were provided for the algorithm predictions and the algorithms were incorporated into an open access digital application: *https://sorg-apps.shinyapps.io/tha_complication*.

Conclusions: The stochastic gradient boosting algorithm demonstrated excellent discriminatory capacity for identifying patients at high-risk of experiencing complications following primary THA. The open-access application can be used to determine patient-specific risk of complications based off their medical profile.

Notes