Machine Learning Algorithms Identify At-Risk Patients For Deep Vein Thrombosis After Primary Total Hip Arthroplasty: Analysis of National-scale Patient Cohort

Ziwei Huang MD, PhD[1], Tony Lin-Wei Chen MD, PhD[1], Anirudh Buddhiraju MD[1], Blake M. Bacevich BSc[1], Michelle Shimizu BSc[1], MohammadAmin RezazadehSaatlou MD[1], Henry Hojoon Seo BA[1], Shane Fei Chen MA[1], Christian A. Pean MD[1], Oh-Jak Kwon[1], Jona Kerluku 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: Ziwei Huang (N), Tony Lin-Wei Chen (N), Anirudh Buddhiraju (N), Blake M. Bacevich (N), Michelle Shimizu (N), MohammadAmin RezazadehSaatlou (N), Henry Hojoon Seo (N), Shane Fei Chen (N), Christian A. Pean (N), Oh-Jak Kwon (N), Jona Kerluku (N), John G. Esposito (N), Young-Min Kwon (5- MicroPort; 5- Depuy; 5- Smith & Nephew; 5- Stryker; 5- Zimmer Biomet)

INTRODUCTION: Deep vein thrombosis (DVT) is a complication following primary total hip arthroplasty (THA). However, there is currently a scarcity of robust tools that accurately predict DVT after primary THA due to the variations in patient characteristics and low incidence of DVT. Prior studies have reported promising performance of machine learning models in forecasting various adverse events after arthroplasty. Therefore, this study aimed to develop and assess machine learning models for the prediction of DVT after primary THA using a national-scale patient cohort.

METHODS: A total of 70,733 patients receiving primary THA between 2013 and 2020 were identified from the ACS-NSQIP database. Candidate predictors of DVT were selected based on existing literature and refined using recursive feature elimination. Five machine learning models—artificial neural network (ANN), random forest (RF), histogram-based gradient boosting (HGB), support vector machines (SVM), and K-nearest neighbor (KNN) were constructed and used to predict the risk of DVT following primary THA. The models were assessed by discrimination, calibration, and decision curve analysis.

RESULTS: In the study cohort, 317 patients (0.45%) developed DVT after surgery. The included models demonstrated excellent performance during the training session with a reporting AUC of 0.93–0.97, a calibration slope of 0.88–1.11, a calibration intercept of -0.13–0.21, and a Brier score of 0.009–0.013. HGB was the best performer during the testing session, which returned an AUC of 0.97, a calibration slope of 1.11, a calibration intercept of 0.17, and a Brier score of 0.013 (Table 1). HGB also demonstrated good utility by producing higher net benefits than the default management strategies in the decision curve analysis. According to our model's prediction, length of stay, international normalized ratio, age, partial thromboplastin time, white blood cell, body mass index, hematocrit, platelet count, and operation time were the strongest predictors of DVT after primary THA (Figure 1).

DISCUSSION: Histogram-based gradient boosting machine learning model was the best machine learning model for predicting DVT following primary THA. In line with previous studies, we found that clot-related blood biomarkers, length of stay, and body mass index predominantly determined the probability of DVT. Machine learning models showed the potential to aid surgeons in identifying high-risk patients and optimizing perioperative DVT prophylaxis following primary THA.

SIGNIFICANCE/CLINICAL RELEVANCE: Machine learning algorithms have the potential as a clinical tool to identify patients at risk of DVT following primary THA.

Table 1. Discrimination and calibration of machine learning models in the testing dataset for primary THA.

Metric	ANN	RF	HGB	KNN	SVM
AUC	0.92	0.91	0.93	0.96	0.88
Slope	0.85	1.07	0.92	1.32	0.69
Intercept	-0.07	-0.15	0.18	0.17	0.21
Brier score	0.012	0.009	0.010	0.013	0.015

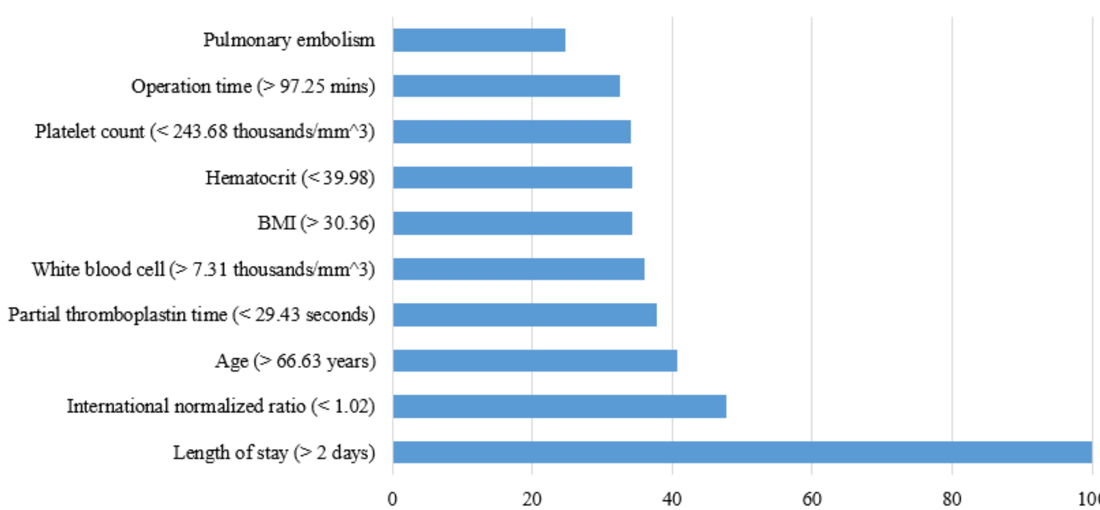

Figure 1. Global feature importance plot for the prediction of DVT following primary THA.