Predicting The Risk of Deep Vein Thrombosis After Primary Total Knee Arthroplasty Using Machine Learning Algorithms: Analysis of 109,478 patients In A National Database

Tony Lin-Wei Chen MD, PhD[1], Ziwei Huang MD, PhD[1], Anirudh Buddhiraju MD[1], Shane Fei Chen MA[1], Michelle Shimizu BSc[1], Blake M. Bacevich BSc[1], Henry Hojoon Seo BA[1], MohammadAmin RezazadehSaatlou MD[1], Christian A. Pean MD[1], Jona Kerluku BSc[1], Oh-Jak Kwon[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

INTRODUCTION: Deep vein thrombosis (DVT) after primary total knee arthroplasty (TKA) and its associated complications remain one of the challenges for orthopedic surgeons. Due to the large variances in patient characteristics and the relatively low incidence rate, DVT is challenging to predict using traditional statistical models. Machine learning models, on the other hand, showed potential utility in analyzing the interconnection between patient variables to predict major adverse events following arthroplasty. This study, therefore, aimed to develop machine learning models based on a national patient cohort and evaluate their performance in predicting DVT after primary TKA.

METHODS: We identified 109,478 patients who underwent primary TKA between 2013 and 2020 from the ACS-NSQIP database. We selected potential predictors of DVT based on existing literature and refined them using recursive feature elimination. Five advanced 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 to predict the risk of DVT following primary TKA. The performance of the models was evaluated in terms of discrimination, calibration, and decision curve analysis.

RESULTS: In the study cohort, 910 patients (0.83%) developed DVT after surgery. All models except for SVM demonstrated great performance during the training session with an AUC of 0.83–0.89, a calibration slope of 0.83–1.23, a calibration intercept of -0.12–0.26, and a Brier score of 0.010–0.015. The models' performance was improved during the testing session. HGB yielded the best prediction accuracy that featured an AUC of 0.87, a calibration slope of 0.96, a calibration intercept of 0.05, and a Brier score of 0.015 (Table 1). HGB also showed clinical utility by producing higher net benefits than the default treatment strategies in the decision curve analysis. The strongest predictors of DVT after primary TKA were the length of stay, body mass index, platelet count, operation time, white blood cell count, partial thromboplastin time, hematocrit, age, and international normalized ratio (Figure 1).

DISCUSSION: Histogram-based gradient boosting machine learning model (HGB) was the most accurate machine learning model in predicting patient-specific DVT after primary TKA. Our study aligned with existing findings that clot-related blood tests (platelet count, partial thromboplastin time, and international normalized ratio) and length of stay were the dominant predictors of DVT probability.

SIGNIFICANCE/CLINICAL RELEVANCE: Machine learning models have the potential to assist surgeons in identifying high-risk patients for DVT and optimizing perioperative prophylactic measures after primary TKA.

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

Metric	ANN	RF	HGB	KNN	SVM
AUC	0.81	0.84	0.87	0.91	0.80
Slope	0.51	0.90	0.96	1.42	0.31
Intercept	0.01	-0.13	0.05	0.19	0.23
Brier score	0.018	0.013	0.015	0.016	0.018
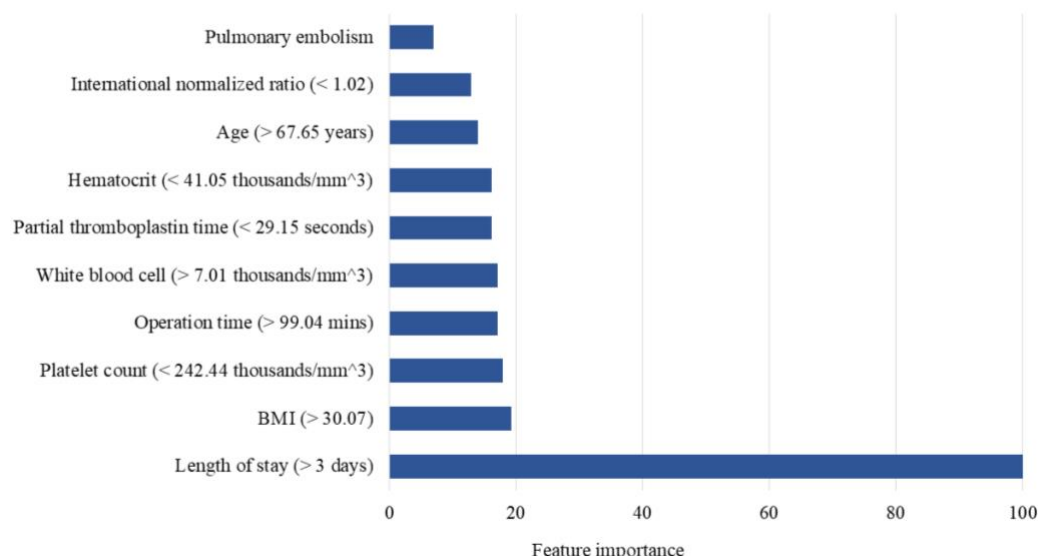
Figure 1. Global feature importance plot for the prediction of DVT following primary TKA.