Machine Learning Approach to Predict Venous Thromboembolism Among Patients Undergoing Multi-Level Lumbar Spinal Fusion with Posterior Instrumentation

Kevin Heo[1], Sameer Khawaja [1], Alexander R. Webb[1], Lauren A. Barber[1], Sangwook Tim Yoon[1]
[1]Emory University, Atlanta, GA
kevin.heo@emory.edu

Disclosures: Kevin Heo (N), Sameer Khawaja (N), Alexander R. Webb (N), Lauren A. Barber (N), Sangwook Tim Yoon (N)

INTRODUCTION: Venous thromboembolism (VTE) after lumbar spinal fusion is associated with significant morbidity and mortality. Currently, there is variability in the timing and usage of chemoprophylactic anticoagulation after spine surgery. The absence of guidelines for thromboprophylaxis in these procedures underscores the importance of identifying patients at risk. Although several studies have identified risk factors associated with VTE, there is a lack of standard scoring systems that accurately stratify patient VTE risk. As a result, this study incorporated machine learning (ML) models to assess key risk factors of VTE in patients who underwent lumbar spinal fusion.

METHODS: Data was collected from the IBM MarketScan Database (2009-2021) for patients ≥18 years old who underwent lumbar spinal fusion with posterior instrumentation (3-6 levels). VTE incidence (including deep vein thrombosis and/or pulmonary embolism) was recorded 90-days post-surgery. Risk factors for VTE were investigated and compared through several machine learning models including logistic regression, linear support vector machine (LSVM), random forest, extreme gradient boosting (XGBoost), and neural networks.

RESULTS: Among the 164,582 patients who underwent lumbar fusion with posterior instrumentation (3-6 levels), the overall rate of 90-day VTE of 5.49%. The LSVM model demonstrated the best prediction with an area under the curve (AUC) value of 0.68 and a diagnostic odds ratio of 13.50 (Table 1). The most important features for prediction of VTE included prior history of VTE, hemiplegia, metastatic cancer, diagnosis of chronic hypercoagulability, and history of cerebrovascular disease. A diagnosis of chronic hypercoagulability included patients with Factor V Leiden, anti-phospholipid syndrome, history of thrombophilia, or protein C or S deficiency. Patients who did not have these five key risk factors had a 90-day VTE rate of 3.88%. Patients who had an increasing number of key risk factors had subsequently higher risks of postoperative VTE (Table 2). For example, patients with any one of these top five risk factors had a VTE rate of 9.71%, which corresponded to being 2.66 times more likely to develop a 90-day VTE compared to patients without any of these five risk factors (P<0.001). Furthermore, patients with any four of the top five risk factors had a VTE rate of 65.85% after surgery, which corresponded to having 47.74 times greater odds of developing a VTE within 90-days compared to patients without the top five risk factors (P<0.001).

DISCUSSION: The analysis of the data with different machine learning models identified 5 key variables that are most closely associated with VTE. Using these variables, we have developed a simple risk model with odds ratio ranging from 2.66 to 47.74 (see table 5) over a period of 90 days after posterior lumbar fusion surgery (3-6 levels). These findings can also help guide surgeons to understand which patients are in need for closer monitoring and detection of early diagnosis, as well as for properly educating patients about their risk levels.

SIGNIFICANCE/CLINICAL RELEVANCE: Our study is the first large data study to incorporate machine learning models to stratify patient risk factors for developing VTE after lumbar fusion with posterior instrumentation (3-6 levels). With current variability in timing and usage of anticoagulation after surgery, clinicians can utilize these findings to identify patients at acute risk in order to implement mitigation efforts such as extending the prophylactic anticoagulation period or utilizing more aggressive chemoprophylactic agents.

IMAGES AND TABLES:

Table 1. Top Five Important Variables for Risk of 90-Day Venous Thromboembolism (VTE) by Model

Model	AUROC	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5
XGBoost Tree	0.61	History of DVT/PE	Chronic Hypercoagulability	Smoking	Diabetes w/ Complications	Solid Cancer
Logistic Regression	0.68	History of DVT/PE	Metastatic Cancer	Hemiplegia	Chronic Hypercoagulability	Cerebrovascular Disease
Random Forest	0.64	History of DVT/PE	Rheumatic Diseases	Chronic Pulmonary Disease	Obesity	Diabetes w/o Complications
Linear Support Vector Machine	0.68	History of DVT/PE	Hemiplegia	Metastatic Cancer	Chronic Hypercoagulability	Cerebrovascular Disease
Neural Networks	0.67	History of DVT/PE	Metastatic Cancer	Chronic Hypercoagulability	Hemiplegia	Dementia

*DVT = Deep vein thrombosis
*PE = Pulmonary embolism

Table 2. Risk of 90-Day Venous Thromboembolism (VTE) Based on Number of Comorbidities Within the Top 5 per the Linear Support Vector Machine (LSVM) model

Number of Top 5 Risk Factors (LSVM)	90-Day VTE* Rate (%)	Odds Ratio	p-value
0	3.88%	(reference)	
1	9.71%	2.66	<0.001
2	21.25%	6.68	<0.001
3	32.18%	11.74	<0.001
4	65.85%	47.74	<0.001

*No patients had all 5 top risk factors
*Top 5 Risk Factors included history of DVT/PE, hemiplegia, metastatic cancer, chronic hypercoagulable state, and history of cerebrovascular disease