Original Article

A Machine Learning Approach to Predict Deep Venous Thrombosis Among Hospitalized Patients

Logan Ryan, BS[1], Samson Mataraso, BS[1], Anna Siefkas, SM[1] ⬤,
Emily Pellegrini, MEng[1], Gina Barnes, MPH[1],
Abigail Green-Saxena, PhD[1], Jana Hoffman, PhD[1],
Jacob Calvert, MSc[1], and Ritankar Das, MSc[1]

Abstract
Deep venous thrombosis (DVT) is associated with significant morbidity, mortality, and increased healthcare costs. Standard scoring systems for DVT risk stratification often provide insufficient stratification of hospitalized patients and are unable to accurately predict which inpatients are most likely to present with DVT. There is a continued need for tools which can predict DVT in hospitalized patients. We performed a retrospective study on a database collected from a large academic hospital, comprised of 99,237 total general ward or ICU patients, 2,378 of whom experienced a DVT during their hospital stay. Gradient boosted machine learning algorithms were developed to predict a patient's risk of developing DVT at 12- and 24-hour windows prior to onset. The primary outcome of interest was diagnosis of in-hospital DVT. The machine learning predictors obtained AUROCs of 0.83 and 0.85 for DVT risk prediction on hospitalized patients at 12- and 24-hour windows, respectively. At both 12 and 24 hours before DVT onset, the most important features for prediction of DVT were cancer history, VTE history, and internal normalized ratio (INR). Improved risk stratification may prevent unnecessary invasive testing in patients for whom DVT cannot be ruled out using existing methods. Improved risk stratification may also allow for more targeted use of prophylactic anticoagulants, as well as earlier diagnosis and treatment, preventing the development of pulmonary emboli and other sequelae of DVT.

Introduction

Venous thromboembolism (VTE), including deep venous thrombosis (DVT), is a leading cause of cardiovascular disease, with significant associated morbidity and mortality in the United States and globally.[1] DVT is also associated with significant healthcare spending. Estimates of total DVT-related costs range from $7 to $35 billion USD per year.[2] Precluding timely and accurate diagnosis, DVT can lead to acute life-threatening complications (i.e. pulmonary embolism) as well as chronic conditions (i.e. post thrombotic syndrome and recurrent DVT).[3] Therefore, research on the identification and prevention of DVT has focused on reducing morbidity, mortality, and medical spending.

Genetic and acquired risk factors affect the course of the disease. Acquired risk factors such as prolonged immobility and bed rest, surgery, and impaired blood flow are associated with hospital stay and acute illness.[4] For reasons likely related to these risk factors, inpatient populations are known to be at a higher risk of DVT as compared to community dwellers.[5] Early identification of DVT in inpatient populations is essential to minimize these risks. Because definitive diagnosis of DVT requires expensive or invasive imaging testing, several methods have been developed to stratify patients by their DVT risk and identify patients for whom testing is appropriate. Clinical

[1] Dascena, Inc., San Francisco, CA, USA

Corresponding Author:
Anna Siefkas, Dascena, Inc., Houston, TX 77080, USA.
Email: anna@dascena.com

decision rules for the stratification of DVT risk include the Wells Criteria for assessing VTE risk in patients suspected of thrombosis[6] and the International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) score for assessing VTE risk among hospitalized patients.[7] In particular, the Wells score may be useful to rule out DVT in low-risk segments of the population,[8] and the use of the 2 level Wells score is recommended by the European Society of Cardiology consensus statement on DVT diagnosis.[9] However, the performance of the Wells score is heterogeneous across age groups, and between patients with and without a history of VTE.[10]

Studies have found that many of the existing DVT scores in clinical use provide insufficient risk stratification of hospitalized patients[11,12] and are not able to accurately predict which patients are likely to present with DVT. To improve clinician ability to identify and monitor patients at a high risk for development of DVT, we have developed and retrospectively validated machine learning algorithms to predict DVT onset 12 and 24 hours in advance.

Materials and Methods

Data Processing

Patient data were extracted from patient electronic medical records (EMR) at a large academic medical center. We included data extracted from inpatients seen between May 2011 and November 2017. These data included information on patient demographics, diagnoses, vital signs and laboratory data, and medication usage. Data were collected passively, and were de-identified in compliance with the Health Insurance Portability and Accountability Act (HIPAA). Studies performed on de-identified data constitute non-human subject studies, and therefore our study did not require Institutional Review Board approval.

Gold Standard and Definition of Onset Time

The primary endpoint was diagnosis of in-hospital DVT. DVT was identified through the presence of an International Classification of Diseases (ICD) 9 and 10 codes for DVT in the inpatient chart (Supplementary Table 1), followed by administration of an anticoagulant (lenovox®, heparin, aspirin, or warfarin). Anticoagulants were identified by performing a search for character strings matching the generic or brand names for each drug. Time of DVT onset was considered to be the time when an anticoagulant was first ordered. All encounters not meeting these criteria were considered to be negative for DVT, and were assigned an onset time at random to assess the accuracy of model predictions. To include all possible DVT patients in our positive class, no further restrictions were placed on the types of DVT that were included in this study. In particular, DVT were not required to be symptomatic for inclusion, and imaging results were not required or reviewed for the purposes of this study.

For the purposes of this study, we included data from all patients with inpatient encounters and with complete chart

Table 1. Patient Inclusion for 12- and 24-Hour Prediction Experiments.

	12 hour	24 hour
All inpatients	131,466	131,466
Inpatients with complete chart data	94,642	90,576
Inpatients with DVT	1,230	999

data. Chart data were considered incomplete if they did not contain raw data over at least a period of 3 hours and, in the case of the positive class, over a sufficient period of time to allow for a prediction 12 or 24 hours in advance of the DVT event as defined by our gold standard. Raw data refers to any of the lab values or clinical measurements in **Supplemental** Table 2. Patient inclusion information for patients included in the 12- and 24-hour prediction experiments is provided in Table 1. A total of 94,642 patients were included in training and testing of the 12-hour DVT prediction model, 1,230 of whom met our DVT gold standard. For training and testing the 24-hour DVT prediction model, a total 90,576 patients were included, 999 of whom met our gold standard.

Clinical measurements, patient diagnoses, and laboratory data were extracted to compute algorithm scores (Supplementary Table 2). All data were collected in the 3 hours preceding prediction time. Prediction time was defined as 12 hours before DVT onset time for 12-hour predictions, and 24 hours before DVT onset time for 24 hour predictions. Clinical and lab measurements were included as hourly measurements and as the change in value between hourly measurements. Lab values were included as boolean values indicating the presence of an abnormal lab result.

Machine Learning Model

Two machine learning models were built: one to predict DVT 12 hours in advance of onset, and one to predict DVT 24 hours in advance of onset. The machine learning classifiers were created using gradient boosting for fitting "boosted" decision trees and were implemented by the XGBoost method. Gradient boosting is an ensemble learning technique that combines results from multiple decision trees to create prediction scores. Each decision tree splits the patient population into successively smaller groups. Each tree branch splits the patients who enter it into 2 groups based on whether their covariate value is above or below a given threshold (e.g. a branch may divide patients according to whether their temperature is above or below 100°F.) If a given value is missing for a patient, a "default" path is taken based on the path that correctly classifies the most patients in the training data. After some number of branches, the decision tree ends in a set of "leaves" with each patient represented in exactly 1 leaf, according to the values of their measurements. All patients in each "leaf" of the decision tree are predicted to have the same risk of DVT. The covariate and threshold value involved in each split are selected by an

Table 2. Demographic Information for the Study Sample.

	Characteristic	12 hr Non-DVT Patients (%) (n = 93,412)	12 hr DVT (%) (n = 1,230)	24 hr Non-DVT Patients (%) (n = 89,577)	24 hr DVT Patients (%) (n = 999)
Age	18-30	10,936 (11.71%)	86 (6.99%)	10,324 (11.53%)	69 (6.91%)
	30-39	14,390 (15.40%)	84 (6.83%)	13,727 (15.32%)	70 (7.01%)
	40-49	12,146 (13.00%)	160 (13.01%)	11,599 (12.95%)	132 (13.21%)
	50-59	17,310 (18.53%)	246 (20.00%)	16,566 (18.49%)	199 (19.92%)
	60-69	19,641 (21.03%)	352 (28.62%)	18,877 (21.07%)	291 (29.13%)
	>70	18,989 (20.33%)	302 (24.55%)	18,484 (20.63%)	238 (23.82%)
Sex	Male	42,189 (45.16%)	648 (52.68%)	40,314 (45.00%)	526 (52.65%)
	Female	51,223 (54.84%)	582 (47.32%)	49,263 (55.00%)	473 (47.35%)
Comorbidities	Renal	6,618 (7.08%)	185 (15.04%)	6,489 (7.24%)	154 (15.42%)
	Organ Transplant	8,396 (8.99%)	176 (14.31%)	8,285 (9.25%)	152 (15.22%)
	Cancer	52,643 (56.36%)	1032 (83.9%)	50,905 (56.83%)	852 (85.29%)
	Diabetes	16,875 (18.07%)	301 (24.47%)	16,498 (18.42%)	249 (24.92%)
	COPD	5,644 (6.04%)	110 (8.94%)	5,544 (6.19%)	91 (9.11%)
	Hepatic	8,006 (8.57%)	181 (14.72%)	7,886 (8.8%)	159 (15.92%)
	Cardiovascular	57,983 (62.07%)	1,230 (100%)	56,297 (62.85%)	999 (100%)

Abbreviations used: COPD Chronic obstructive pulmonary disease. DVT: Deep venous thrombosis.

algorithm designed to minimize loss on the training data and without overfitting by regularizing the loss function.

Prior to model training, data were randomly split in a 80:20 ratio of training to test data, where test data was only evaluated after completion of training process. A cross-validated grid search and 5 fold cross validation were used for the purposes of hyperparameter optimization. For each algorithm, four-fifths of the encounters from within the training set were randomly selected to train a model, and the remaining one-fifth were used as a validation set. This was repeated a total of 5 times, with a cross-validated grid search performed during each iteration, across a sparse parameter grid. The combination of hyperparameters that performed the best across all 5 iterations was selected for incorporation into the final respective model, which was subsequently trained on the entirety of the training data set in XGBoost before being tested on the holdout test set, which had not been seen during any of the model training process. To prevent model overfitting, we included a hyperparameter for the early stopping of the iterative tree-addition procedure in the cross-validated grid search. For 12 hour predictions, we restricted tree depth to a maximum of 4 branching levels, and set the XGBoost learning rate parameter 0.06. For 24 hour predictions, we restricted tree depth to a maximum of 5 branching levels and set the XGBoost learning rate parameter 0.05. For both models, we restricted tree ensembles to 100 trees to limit computational burden.

Statistical Analysis

Model performance was assessed using area under the receiver operating characteristic (AUROC), sensitivity, specificity, diagnostic odds ratio (DOR), and positive and negative likelihood ratios (LR+ and LR-). Model performance was assessed separately for prediction of DVT 12 and 24 hours in advance of onset.

Model performance was compared to that of the IMPROVE score, calculated 12 and 24 hours before DVT onset time. Performance of IMPROVE was assessed using the same metrics as above. For ease of comparison, operating points for both the machine learning and IMPROVE models were chosen to obtain a sensitivity of approximately 0.80.

Feature importance for all inputs was assessed using Shapley values, which take into account the contribution of a feature in producing a more accurate model (that is, the degree to which a feature reduces the error in the loss function). Because many clinical measurements were incorporated as discrete features representing measures as several time points, the importance of a single measurement type at 2 different time points could vary.

Results

In total, 99,412 patient encounters were included in our analysis, 1,230 of whom experienced a DVT during their hospital stay. Those who experienced a DVT were, on average, older, more likely to be male, and were more likely to have comorbidities or a history of organ transplant (Table 2). Median age for the total population was 56 years (interquartile range (IQR): 33, 69), and median age among those who experienced a DVT was 62 (IQR: 47,71). Reason for hospitalization for all patients is included in Supplementary Table 3.

At 12 hours before onset of DVT, the algorithm demonstrated an AUROC of 0.83, along with a sensitivity of 0.80 and a specificity of 0.66 (Table 3). At 24 hours before DVT onset the algorithm achieved an AUROC of 0.85, and a sensitivity and specificity of 0.80 and 0.75, respectively (Table 3). The algorithm outperformed the IMPROVE score at both time points (Figure 1).

Feature importance were similar for 12 and 24 hour prediction models, with cancer status, VTE history, international

Table 3. Performance Metrics for DVT Prediction at 12 and 24 Hours Before Onset.

	12 Hour MLA (95% CI)	12 Hour IMPROVE (95% CI)	24 Hour MLA (95% CI)	24 Hour IMPROVE (95% CI)
AUROC	0.83 (0.81, 0.85)	0.78 (0.75, 0.80)	0.85 (0.83, 0.87)	0.79 (0.77, 0.81)
Sensitivity	0.80 (0.75, 0.85)	0.80 (0.75, 0.85)	0.80 (0.84, 0.86)	0.83 (0.78, 0.88)
Specificity	0.66 (0.65, 0.67)	0.61 (0.60, 0.62)	0.75 (0.74, 0.76)	0.60 (0.59, 0.61)
DOR	7.72 (5.64, 10.57)	6.07 (4.44, 8.29)	11.94 (8.43, 16.92)	7.37 (5.09, 10.67)
LR+	2.34 (2.19, 2.50)	2.03 (1.90, 2.17)	3.19 (2.96, 3.43)	2.08 (1.95, 2.22)
LR-	0.30 (0.24, 0.39)	0.33 (0.26, 0.43)	0.27 (0.20, 0.35)	0.28 (0.21, 0.38)

Abbreviations used: AUROC: Area under the receiver operating characteristic. CI: Confidence interval. DOR: Diagnostic odds ratio. MLA: Machine learning algorithm. LR: Likelihood ratio.
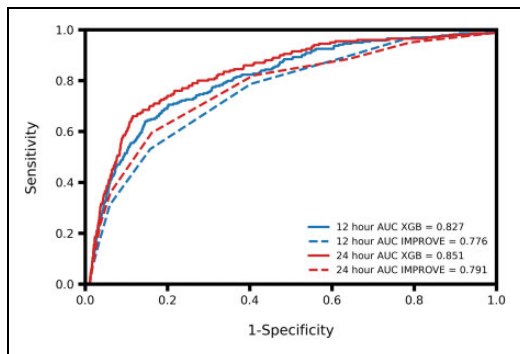
Figure 1. Receiver operating characteristic (ROC) curves and comparison of area under the ROC (AUROC) of the XGBoost (XGB) and IMPROVE models for 12 and 24 hour prediction of deep venous thrombosis.

normalized ratio (INR), change in INR, and heart rate being the top 5 features for both models (Figure 2). Epinephrine, 8 hour average urine output, age, and antibiotic use were also important features of both models as measured by Shapley values.

Discussion

Timely diagnosis of DVT is essential to minimize the risk of complications like pulmonary emboli and post thrombotic syndrome.[13] However, the accuracy of common clinical scoring systems for DVT risk stratification has been shown to be variable in inpatient, elderly, pediatric, and post-operative patient populations.[3] In inpatient settings, commonly used DVT scores such as the Wells score are inflated by comorbidities and non-specific physical findings that are common among hospitalized patients.[14] This leads to a moderate or high characterization of risk among many patients without DVT. Therefore, there is an ongoing need for improved assessment and accurate identification of DVT risk factors specific to hospitalized patients. Our results demonstrate that machine learning algorithms can be used to predict DVT 12 and 24 hours in advance of onset. For both 12- and 24-hour prediction, sensitivity was 80% with specificity exceeding 65%, indicating strong predictive performance while maintaining a low false positive rate. For an in-patient population at increased risk of DVT, these performance

metrics indicate that this algorithm may be useful for an improvement of risk assessment over current tools. An exploration of feature importance revealed that the model is utilizing several known DVT risk factors in generating DVT predictions, in addition to utilizing additional novel clinical and laboratory measures. Cancer is considered a major provoking risk factor for VTE, and was the most important feature in making predictions for both the 12- and 24-hour prediction models.[15,16] As per Virchow's triad, the classic explanatory model for VTE, patients with hypercoagulability are predisposed to developing thromboses. In line with this well-recognized relationship, INR, a measure of coagulability, was another key feature.[17] Obesity has been shown to confer additional risk for VTE.[16,18] Consistent with this prior research, higher weight was associated with greater risk of DVT in our model. This analysis supports that that model is utilizing known risk factors for DVT (such as cancer diagnosis and prior VTE) in generating DVT risk prediction scores.

To avoid referral for testing in patients at low risk of DVT, various clinical decision rules have been developed for the purposes of risk stratification,[19] including the Wells score,[3] Oudega score,[20] Gagne score,[21] LEFT score,[22] and IMPROVE score.[7] The Wells score for DVT is often used by clinicians as a best-available approach, but multiple studies have reported that its utility for risk stratification in the inpatient setting is low.[12,14,19] Silveira et al prospectively evaluated the Wells score for risk stratification among inpatients with suspected DVT, and found that it performed only slightly better than chance.[12] Further, the Wells score was developed for use on patients already suspected of a DVT; the score is therefore not readily interpretable on the general patient population. The IMPROVE score has similarly been assessed for its ability to identify patients at risk of experiencing a DVT. In particular, recent work incorporating D-dimer has been found to increase the accuracy of the score. A study by Spyropoulos et al[23] found that a modified IMPROVE score incorporating D-dimer was capable of identifying patients at a 3-fold higher risk of VTE and who benefited from extended thromboprophylaxis. Another study found that the IMPROVEDD score, which incorporates D-dimer measures, demonstrated a higher AUROC than the IMPROVE score for VTE prediction at 42 days (0.621 vs 0.588).[24] However, the low AUROC of the
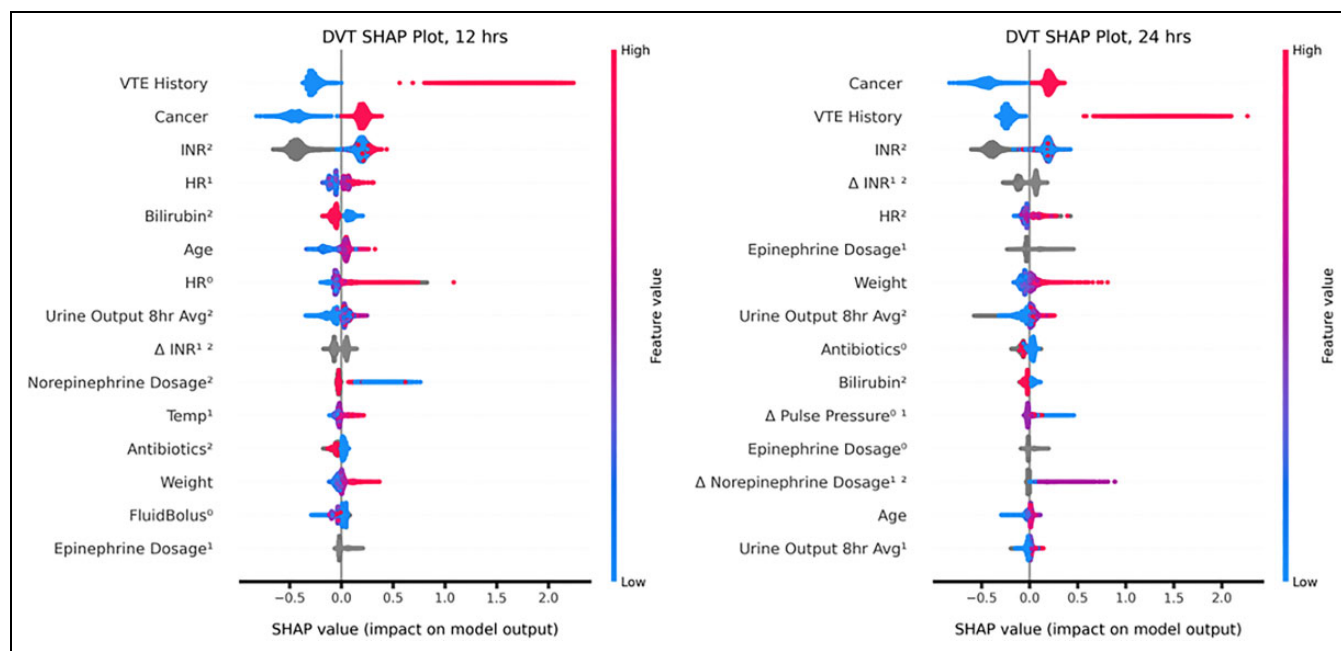
Figure 2. Feature correlations and distribution of feature importance for each patient. Input variables are ranked in descending order of feature importance. Red indicates a high feature value; blue indicates a low feature value. Dots to the right resulted in a higher score; dots to the left resulted in a lower score. The superscript denotes the number of hours prior to the time the algorithm was applied, and Δ denotes change between the measurements at each indicated hour. For example, HR[0]. represents heart rate at the time the algorithm was applied, and Δ Antibiotics[01] represents the change in antibiotic status from the previous hour to the current hour. Abbreviations used: INR: international normalized ratio. HR: heart rate. VTE: venous thromboembolism. PPL: pulse pressure.

IMPROVEDD score supports that there is still substantial unmet need for DVT risk stratification, particularly in the short term.

Use of clinical decision rules in daily practice is often influenced by the expertise of the clinician using them,[3,25] which may affect efficacy of the decision rules in stratifying patients most at risk of disease. These standard scoring systems were generally developed and validated for identifying patients at low risk of developing DVT, so that DVT can be excluded without referring a patient for imaging tests.[26] Most patients suspected of having DVT are not ultimately diagnosed[26] and there is a lack of validated high risk stratification methods for those patients who are particularly likely to develop DVT. Toward this end, our results demonstrate that machine learning may offer an opportunity for more precise risk stratification approaches, enabling improvements in patient monitoring, use of prophylactic anticoagulants, and earlier diagnosis of DVT to prevent its sequelae.

Several retrospective studies have applied ML to the prediction of VTE in patient subpopulations. Previous research has utilized machine learning methods to characterize a patients' risk of developing DVT and VTE.[27] Ferroni et al used kernel machine learning techniques to predict risk of venous thromboembolism in oncology outpatients with high performance, indicating that a machine learning approach may be of clinical value for VTE risk stratification in chemotherapy-treated outpatients. However, this study acknowledged that additional patient complexities, such as obesity and active patient

therapeutics, may impact the performance of ML in DVT prediction.[27] While the model performed well in predicting VTE among oncology outpatients with an AUROC of 0.716, performance remains to be validated in inpatient care settings. A recent study examined the discriminative ability of the IMPROVE ensemble machine learning software to predict VTE onset in a subpopulation of acutely ill patients classified as high risk for VTE.[28] Though the research yielded high discriminative ability with an AUROC of 0.69 for the standard ML algorithm and .68 c-statistic for the reduced ML (rML) model, this study did not present results indicating the length of time prior to onset that VTE can be predicted, and also remains to be validated in a general inpatient population.[28] The results present a strong argument for the capability of ML to predict VTE. Expanding on these findings, our research presents results that may indicate how a ML may be useful in predicting DVT up to 24 hours in advance of onset while being used in more general clinical settings.

A prospective cohort study examined the ability of an ML algorithm to predict the incidence of VTE over the course of 5 years and yielded a 0.75 AUROC.[29] The study sample included patients with estimated DVT risk between 0.5%-10% over 1 year and 5 year periods.[29] Thus, the results are not necessarily representative of how an ML algorithm can perform in a very low-risk population or in a higher-risk patient population. Additionally, while the ability to predict long-term risk of VTE remains important, this tool does not allow for short-term risk prediction of in-patient populations. Our

research fills these gaps by allowing for short term risk predictions in the general inpatient population, rather than a specific population restricted to certain risk levels.

There are several limitations to this study. First, because of the way DVT onset was assessed, it is possible that we assigned an onset time before DVT development if any patients were given prophylactic anticoagulants and later developed an active DVT for which they were given therapeutic anticoagulants. However, given our patient population and the effectiveness of prophylactic anticoagulation to prevent DVTs,[30,31] this is likely a rare occurrence in our data set and is therefore unlikely to significantly impact our findings. It is also possible that our gold standard for identifying DVT did not accurately capture the DVT patient population in our data. There are known difficulties in identifying DVT events in administrative data.[32] Because we did not perform clinician chart adjudication of patient outcomes, it is possible that patients were misclassified by our use of ICD codes. However, we minimized this limitation through the inclusion of anticoagulants as well as ICD codes for identifying DVT. Inclusion of treatment information has been found to increase the PPV of ICD codes for identifying VTE from 72% to 91%.[33] It is worth noting that medications used to define our gold standard and DVT onset time have indications beyond the treatment of acute DVT. It is therefore possible that patients in our positive class were misclassified based on their treatment regimen. Because this is a retrospective study, we are unable to determine the performance of the DVT prediction algorithm in a prospective clinical setting where data availability may differ from our study. Future work assessing algorithm performance in a prospective setting is required to determine how clinicians may respond to DVT risk predictions, as well as to determine whether predictions impact patient outcomes or resource allocation.

Conclusion

The machine learning algorithm presented in this study is a useful predictive tool for anticipating risk of DVT at 12 and 24 hours in advance of onset. Improved prediction and risk stratification enabled by the machine learning algorithm can prevent unnecessary invasive testing in patients for whom DVT cannot be ruled out using existing methods. Improved risk stratification may also allow for more targeted use of prophylactic anticoagulants, as well as earlier diagnosis and treatment, preventing the development of pulmonary emboli and other sequelae of DVT.

Authors' Note

Author Roles: L. Ryan contributed to the data analysis, writing, and revision of this study. S. Mataraso, J. Calvert, and R. Das contributed to the conception and revision of this study. A. Siefkas, E. Pellegrini, G. Barnes, A. Green-Saxena, and J. Hoffman contributed to the writing and revision of this study.

References

1. White R H. The epidemiology of venous thromboembolism. *Circulation*. 2003;107(23_suppl_1):I-4.
2. Mahan CE, Holdsworth MT, Welch SM, Borrego M, Spyropoulos AC. Deep-vein thrombosis: a United States cost model for a preventable and costly adverse event. *Thromb Haemost*. 2011; 106(3):405-415.
3. Kafeza M, Shalhoub J, Salooja N, Bingham L, Spagou K, Davies AH. A systematic review of clinical prediction scores for deep vein thrombosis. *Phlebology*. 2017;32(8):516-531.
4. Heit JA, Spencer FA, White RH. The epidemiology of venous thromboembolism. *J Thromb Thrombolysis*. 2016;41(1):3-14.
5. Heit JA, Melton LJ, Lohse CM, et al. Incidence of venous thromboembolism in hospitalized patients vs community residents. *Mayo Clin Proc*. 2001;76(11):1102-1110.
6. Wells PS, Anderson DR, Rodger M, et al. Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis. *N Engl J Med*. 2003;349(13):1227-1235.
7. Spyropoulos AC, Anderson FA, FitzGerald G, et al. Predictive and associative models to identify hospitalized medical patients at risk for VTE. *Chest*. 2011;140(3):706-714.
8. Modi S, Deisler R, Gozel K, et al. Wells criteria for DVT is a reliable clinical tool to assess the risk of deep venous thrombosis in trauma patients. *World J Emerg Surg WJES*. 2016;11:24.
9. Mazzolai L, Aboyans V, Ageno W, et al. Diagnosis and management of acute deep vein thrombosis: a joint consensus document from the European Society of Cardiology working groups of aorta and peripheral vascular diseases and pulmonary circulation and right ventricular function. *Eur Heart J*. 2018;39(47): 4208-4218.
10. Goodacre S, Sutton AJ, Sampson FC. Meta-analysis: the value of clinical assessment in the diagnosis of deep venous thrombosis. *Ann Intern Med*. 2005;143(2):129-139.
11. Greene MT, Spyropoulos AC, Chopra V, et al. Validation of risk assessment models of venous thromboembolism in hospitalized medical patients. *Am J Med*. 2016;129(9):1001. e9-1001. e18.
12. Silveira PC, Ip IK, Goldhaber SZ, Piazza G, Benson CB, Khorasani R. Performance of wells score for deep vein thrombosis in the inpatient Setting. *JAMA Intern Med*. 2015;175(7): 1112-1117.

13. Hansrani V, Khanbhai M, McCollum C. The prevention of venous thromboembolism in surgical patients. *Adv Exp Med Biol*. 2017; 906:1-8.

14. Price EL, Minichiello T. The Wells deep vein thrombosis score for inpatients: not the right tool for the job. *JAMA Intern Med*. 2015;175(7):1118-1119.

15. Kearon C, Ageno W, Cannegieter SC, Cosmi B, Geersing GJ, Kyrle PA; Subcommittees on control of anticoagulation, and predictive and diagnostic variables in thrombotic disease. Categorization of patients as having provoked or unprovoked venous thromboembolism: guidance from the SSC of ISTH. *J Thromb Haemost*. 2016;14(7):1480-1483.

16. Ortel TL, Neumann I, Ageno W, et al. American Society of Hematology 2020 guidelines for management of venous thromboembolism: treatment of deep vein thrombosis and pulmonary embolism. *Blood Adv*. 2020;4(19):4693-4738.

17. Stone J, Hangge P, Albadawi H, et al. Deep vein thrombosis: pathogenesis, diagnosis, and medical management. *Cardiovasc Diagn Ther*. 2017;7(Suppl 3):S276-S284.

18. Rahmani J, Haghighian Roudsari A, Bawadi H, et al. Relationship between body mass index, risk of venous thromboembolism and pulmonary embolism: a systematic review and dose-response meta-analysis of cohort studies among four million participants. *Thromb Res*. 2020;192:64-72.

19. Ageno W. The Wells rule is not accurate in hospitalized patients. *Nat Rev Cardiol*. 2015;12(8):449-450.

20. Oudega R, Moons KGM, Hoes AW. Ruling out deep venous thrombosis in primary care. A simple diagnostic algorithm including D-dimer testing. *Thromb Haemost*. 2005;94(1):200-205.

21. Gagne P, Simon L, Le Pape F, et al. Clinical prediction rule for diagnosing deep vein thrombosis in primary care. *Presse Med Paris Fr*. 2009;38(4):525-533.

22. Chan WS, Lee A, Spencer FA, et al. Predicting deep venous thrombosis in pregnancy: out in "LEFt" field? *Ann Intern Med*. 2009;151(2):85-92.

23. Spyropoulos AC, Lipardi C, Xu J, et al. Modified improve VTE risk score and elevated d-dimer identify a high venous thromboembolism risk in acutely ill medical population for extended thromboprophylaxis. *TH Open*. 2020;4(1):e59-e65.

24. Gibson CM, Spyropoulos AC, Cohen AT, et al. The improvedd VTE risk score: incorporation of d-dimer into the improve score to improve venous thromboembolism risk stratification. *TH Open*. 2017;1(1):e56-e65.

25. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 2009;338:b606.

26. Geersing GJ, Janssen KJ, Oudega R, et al. Diagnostic classification in patients with suspected deep venous thrombosis: physicians' judgement or a decision rule? *Br J Gen Pract*. 2010; 60(579):742-748.

27. Ferroni P, Zanzotto FM, Scarpato N, Riondino S, Guadagni F, Roselli M. Validation of a machine learning approach for venous thromboembolism risk prediction in oncology. *Dis Markers* [Internet]. 2017;2017:8781379.

28. Nafee T, Gibson CM, Travis R, et al. Machine learning to predict venous thrombosis in acutely ill medical patients. *Res Pract Thromb Haemost*. 2020;4(2):230-237.

29. Hippisley-Cox J, Coupland C. Development and validation of risk prediction algorithm (QThrombosis) to estimate future risk of venous thromboembolism: prospective cohort study. *BMJ* [Internet]. 2011;343:d4656.

30. Mismetti P, Laporte S, Darmon JY, Buchmüller A, Decousus H. Meta-analysis of low molecular weight heparin in the prevention of venous thromboembolism in general surgery. *Br J Surg*. 2001; 88(7):913-930.

31. Laryea J, Champagne B. Venous thromboembolism prophylaxis. *Clin Colon Rectal Surg*. 2013;26(3):153-159. doi:10.1055/s-0033-1351130

32. Fang MC, Fan D, Sung SH, et al. Validity of using inpatient and outpatient administrative codes to identify acute venous thromboembolism: the CVRN VTE Study. *Med Care*. 2017;55(12): e137-e143.

33. Sanfilippo KM, Wang T-F, Gage BF, Liu W, Carson KR. Improving accuracy of international classification of diseases codes for venous thromboembolism in administrative data. *Thromb Res*. 2015;135(4):616-620.