






ORIGINAL ARTICLE

Machine learning to predict venous thrombosis in acutely ill medical patients

Tarek Nafee MD¹   | C. Michael Gibson MS, MD¹  | Ryan Travis MS¹ | Megan K. Yee MPH¹ | Mathieu Kerneis MD¹ | Gerald Chi MD¹ | Fahad AlKhalfan MD¹ | Adrian F. Hernandez MD²  | Russell D. Hull MD³ | Ander T. Cohen MD⁴ | Robert A. Harrington MD⁵  | Samuel Z. Goldhaber MD⁶

¹The Cardiovascular Division, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts

²Duke University, The Duke Clinical Research Institute, Durham, North Carolina

³Foothills Hospital, University of Calgary, Calgary, AB, Canada

⁴Guys and St. Thomas Hospitals, London, UK

⁵Department of Medicine, Stanford University, Stanford, California

⁶Cardiovascular Division, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts

Correspondence

Tarek Nafee, Beth Israel Deaconess Medical Center, 930 Commonwealth Avenue, 3rd Floor, Boston, MA 02215.
Email: tarekomarnafee@gmail.com

Funding information

This study was funded by Portola Pharmaceuticals Inc.

Handling Editor: Dr Cihan Ay.

Abstract

Background: The identification of acutely ill patients at high risk for venous thromboembolism (VTE) may be determined clinically or by use of integer-based scoring systems. These scores demonstrated modest performance in external data sets.

Objectives: To evaluate the performance of machine learning models compared to the IMPROVE score.

Methods: The APEX trial randomized 7513 acutely medically ill patients to extended duration betrixaban vs. enoxaparin. Including 68 variables, a super learner model (ML) was built to predict VTE by combining estimates from 5 families of candidate models. A “reduced” model (rML) was also developed using 16 variables that were thought, a priori, to be associated with VTE. The IMPROVE score was calculated for each patient. Model performance was assessed by discrimination and calibration to predict a composite VTE end point. The frequency of predicted risks of VTE were plotted and divided into tertiles. VTE risks were compared across tertiles.

Results: The ML and rML algorithms outperformed the IMPROVE score in predicting VTE (c-statistic: 0.69, 0.68 and 0.59, respectively). The Hosmer-Lemeshow goodness-of-fit *P*-value was 0.06 for ML, 0.44 for rML, and <0.001 for the IMPROVE score. The observed event rate in the lowest tertile was 2.5%, 4.8% in tertile 2, and 11.4% in the highest tertile. Patients in the highest tertile of VTE risk had a 5-fold increase in odds of VTE compared to the lowest tertile.

Conclusion: The super learner algorithms improved discrimination and calibration compared to the IMPROVE score for predicting VTE in acute medically ill patients.

Abbreviations: ACCP, American College of Chest Physicians; CART, classification and regression tree; CUS, compression ultrasound; DAPT, dual antiplatelet therapy; DOAC, direct oral anticoagulant; DVT, deep vein thrombosis; IMPROVE, International Medical Prevention Registry on Venous Thromboembolism; ML, super learner model; PE, pulmonary embolism; RAM, risk assessment model; rML, reduced super learner model; VTE, venous thromboembolism.

Trial Registration: ClinicalTrials.gov Unique ID: NCT01583218. URL: <https://clinicaltrials.gov/ct2/show/NCT01583218>

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Research and Practice in Thrombosis and Haemostasis* published by Wiley Periodicals, Inc on behalf of International Society on Thrombosis and Haemostasis.

KEYWORDS

acute medically ill, machine learning, personalized medicine, super learner, venous thromboembolism

Essentials

- The identification of acutely ill patients at high risk for VTE is determined clinically or by risk scores.
- Two ensemble learning algorithms were trained to predict venous thrombosis in the APEX trial data set.
- The super learner algorithms improved discrimination and calibration compared to the IMPROVE score.
- This is a proof-of-concept algorithm that requires prospective evaluation in a clinical setting.

1 | INTRODUCTION

Patients with an acute medical illness have an increased risk of venous thromboembolism (VTE) during hospitalization that persists following discharge.^{1,2} Several randomized trials have demonstrated the efficacy of VTE prophylaxis with direct oral anticoagulants (DOACs) compared to low-molecular-weight heparin for 6 to 14 days.³⁻⁵

Based on the results of the APEX trial, the US Food and Drug Administration has licensed betrixaban for first-line thromboprophylaxis in acute medically ill patients at high risk for VTE. The identification of these high-risk patients may be determined clinically or by use of risk assessment models (RAMs) that rely on integer-based scoring systems of known risk factors.^{6,7} Padua and IMPROVE (International Medical Prevention Registry on Venous Thromboembolism) are 2 validated scores for risk stratification of acute medically ill patients.⁸⁻¹⁰ These RAMs demonstrated modest performance in validation data sets.¹¹⁻¹³

Machine learning algorithms are constructed to search for patterns in data that provide maximum predictive ability.^{14,15} These learning methods have demonstrated superiority to traditional diagnostic and prognostic tools in various domains.¹⁶⁻¹⁹ However, the performance of machine learning methods in the prediction of the occurrence of VTE has not been previously explored. The aim of this study is to develop and evaluate the performance of machine learning in the prediction of VTE through 77 days among hospitalized acutely medically ill patients enrolled in the APEX (Acute Medically Ill VTE Prevention With Extended Duration Betrixaban) trial.

2 | METHODS**2.1 | Source of data**

The APEX trial has been described previously.²⁰ In brief, APEX was a randomized, double-blind, double-dummy, active-controlled phase 3 clinical trial that enrolled 7513 patients hospitalized for a specific acute medical illnesses (heart failure, respiratory failure, infectious disease, rheumatic disease, or ischemic stroke). Patients had reduced mobility and specific risk factors for VTE (aged ≥ 75 years, baseline D-dimer

$\geq 2\times$ upper the limit of normal, or 2 additional ancillary risk factors for VTE). Patients were randomized in a 1:1 ratio to receive 80 mg of oral betrixaban for 35 to 42 days, or 40 mg of subcutaneous enoxaparin for a standard duration of 6 to 14 days. Patients were followed for up to 77 days. The 6454 subjects (86%) that had compression ultrasound (CUS) assessment for asymptomatic DVT were included in the data set.

2.2 | Outcome

The primary outcome was a composite of asymptomatic and symptomatic deep vein thrombosis (DVT), pulmonary embolism (PE), and VTE-related mortality through the end of study (77 days). Asymptomatic DVT was assessed by CUS from 32 to 47 days after hospitalization. An independent, blinded, centralized, end-point adjudication committee assessed symptomatic and ultrasound end points.

2.3 | Predictive models

Ensemble learning is a type of machine learning that combines predictions across different candidate models. A total of 39 candidate machine learning algorithms in 5 families of models were built using 10-fold cross validation. Broadly, the families of candidate models built included generalized additive models, elastic net (penalized logistic regression), extreme gradient boosting, random forests, a Bayesian logistic regression with default priors, and a simple classification tree. The super learner ensemble method then used cross validation to select weights that were applied to each candidate model to combine predictions. The weights that the super learner model assigned to each model are demonstrated in Tables S2 and S3. The first model (ML) was built with a total of 68 variables, which comprised all the baseline variables available in the APEX trial data set that had $<1\%$ missing data. A second 'reduced' model (rML) was developed using 16 variables that were thought, a priori, to be risk factors for VTE. These 16 variables were either part of the IMPROVE and/or Padua scores or were identified in a literature search to have associations with VTE. Both super learner ensemble models were compared to each other and to the IMPROVE risk score. The variables included in each of the models are summarized in the Table S1.

2.4 | Performance measures

Predictive performance was compared via model discrimination and calibration. Discrimination was assessed with a cross-validated concordance statistic (c-statistic) and compared using a bootstrapped test of significance. Calibration was assessed via high resolution nonparametric calibration plots and the Hosmer-Lemeshow goodness-of-fit test was used to test for statistical significance. A nonsignificant Hosmer-Lemeshow test ($P > .05$) suggests good calibration, and a significant test ($P < .05$) suggests poor calibration.

2.5 | Risk stratification

The distribution of the predicted risks produced by rML model was plotted and divided into tertiles. Patient characteristics and outcome event rates were described, and median predicted risks were reported in each tertile. Odds ratios and 95% confidence intervals were computed comparing the observed rate of VTE across the

tertiles. An implementation of the rML model in 3 different patient profiles was summarized and marked in relation to the risk distribution in the APEX trial.

Analyses were conducted by an academic research organization, Percutaneous-Pharmacologic Endoluminal Revascularization for Unstable Syndromes Evaluation. All authors drafted and critically revised the manuscript and took responsibility for its content. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

3 | RESULTS

3.1 | Baseline characteristics

Baseline characteristics and outcome summary of the patients enrolled in the APEX trial are shown in Table 1. The mean age was 76 years; 46.1% of patients were hospitalized for heart failure, 11.6% for respiratory failure, 11.2% for ischemic stroke, 27.9% for acute infection, and 3.1% for rheumatic disease. A total of 405 patients

TABLE 1 Baseline characteristics

	Overall (N = 6459)	No VTE event (N = 6052)	VTE event (N = 407)	P value
Age, y, mean (SD)	76.31 (8.28)	76.26 (8.21)	77.05 (9.22)	.06
Male (%)	2924 (45.3)	2732 (45.1)	192 (47.2)	.46
Weight, kg, mean (SD)	80.55 (19.20)	80.62 (19.24)	79.46 (18.66)	.24
BMI, kg/m ² , median (IQR)	28.40 (24.90-33.20)	28.40 (24.90-33.20)	28.40 (24.35-32.40)	.20
Duration of hospitalization, days, median (IQR)	10.00 (8.00-14.00)	10.00 (8.00-14.00)	11.00 (8.00-15.00)	<.001
Creatinine clearance, mL/min, mean (SD)	71.23 (32.92)	71.41 (32.89)	68.56 (33.29)	.09
Race, n (%)				
White	6063 (95.5)	5686 (95.6)	377 (94.5)	.003
Black	116 (1.8)	106 (1.8)	10 (2.5)	
Asian	13 (0.2)	11 (0.2)	2 (0.5)	
American Indian	7 (0.1)	7 (0.1)	0 (0.0)	
Pacific Islander	1 (0.0)	0 (0.0)	1 (0.3)	
Multiple	44 (0.7)	43 (0.7)	1 (0.3)	
Other	104 (1.6)	96 (1.6)	8 (2.0)	
Using strong P-gp inhibitor, n (%)	1161 (18.0)	1091 (18.0)	70 (17.2)	.72
D-dimer, median (IQR)	1.24 (0.65-2.25)	1.20 (0.63-2.15)	2.05 (1.04-3.78)	<.001
History of cancer, n (%)	759 (11.8)	698 (11.5)	61 (15.0)	.04
History of thrombosis, n (%)	512 (7.9)	420 (6.9)	92 (22.6)	<.001
Chronic heart failure, n (%)	1470 (22.8)	1376 (22.7)	94 (23.1)	.91
Acute infectious disease, n (%)	1009 (15.6)	940 (15.5)	69 (17.0)	.49
Severe varicosities, n (%)	1201 (18.6)	1128 (18.6)	73 (17.9)	.77
Hormone replacement, n (%)	59 (0.9)	53 (0.9)	6 (1.5)	.43
Inherited or acquired thrombophilia (%)	7 (0.1)	5 (0.1)	2 (0.5)	.10

BMI, body mass index; IQR, interquartile; P-gp, P-glycoprotein; SD, standard deviation; VTE, venous thromboembolism.

(6.3%) had a VTE through the end of the study. Approximately 76% of VTEs were asymptomatic DVT, 9.6% were symptomatic DVT, 7.1% were PE, and 10% were VTE-related deaths.

3.2 | Performance measures

The ML model had a c-statistic of 0.69 and the rML model had a c-statistic of 0.68, with both demonstrating superior discriminative

ability in this data set compared to the IMPROVE score, which had a c-statistic of 0.59 ($P < .001$). The 2 super learner models had similar discrimination, with a P value of .28 (Figure 1A).

The super learner models appear well calibrated compared to the IMPROVE score, as shown in the calibration plot (Figure 1B). Both the ML and rML models demonstrated good calibration using the Hosmer-Lemeshow test ($P = .06$ and $P = .44$, respectively). In contrast, the IMPROVE score demonstrated poor calibration (Hosmer-Lemeshow, $P < .001$).

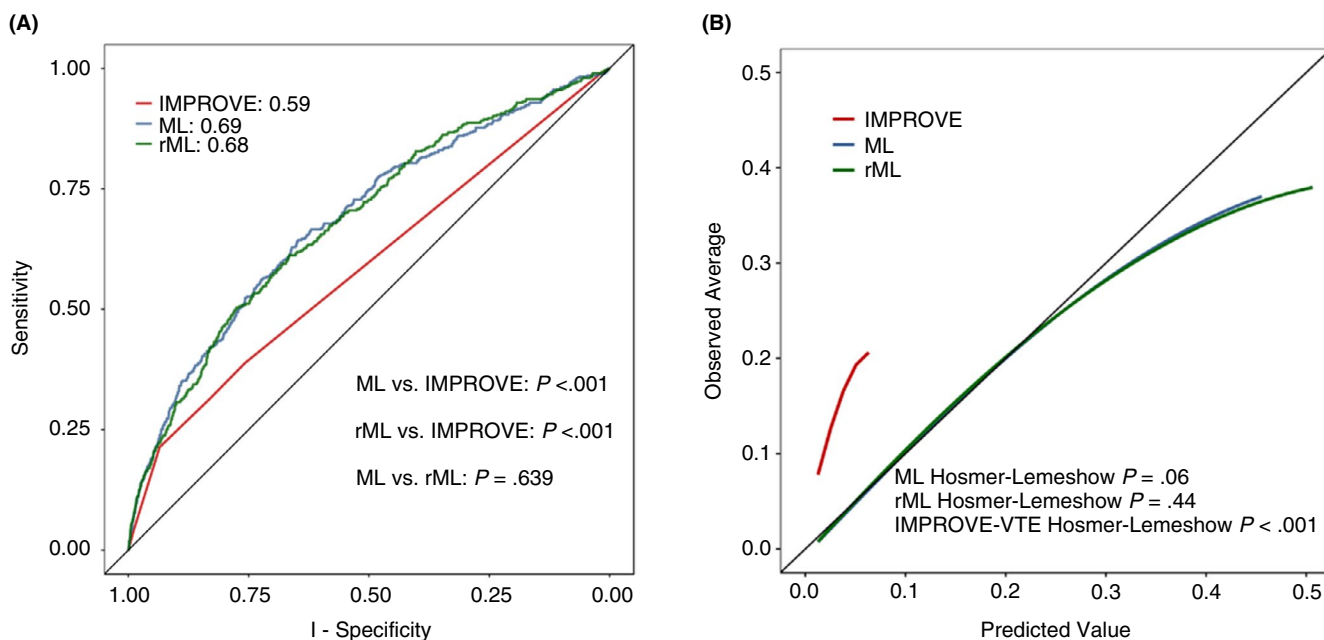


FIGURE 1 Receiver operator characteristics curve and calibration plot. A, the receiver operator characteristics curve for the composite VTE outcome for the complete super learner ensemble (ML), the reduced super learner ensemble (rML), and the IMPROVE risk score. B, the calibration plot for the composite VTE outcome for the complete super learner ensemble (ML), the reduced super learner ensemble (rML), and the IMPROVE risk score. VTE, venous thromboembolism

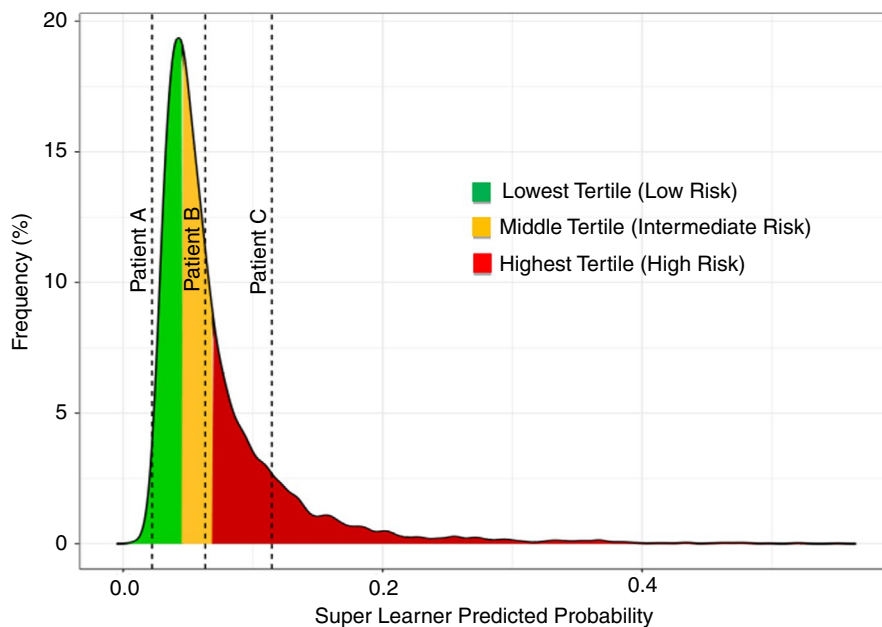


FIGURE 2 Predicted risk distributions. This figure shows the distribution of the predicted risks produced by the rML model. The distribution is divided into tertiles, with the green area representing low risk of VTE, the yellow represents intermediate risk for VTE, and the red represents high risk for VTE. VTE, venous thromboembolism

3.3 | Risk stratification

The frequency of the predicted risks produced by the reduced super learner model (rML) was plotted in Figure 2 and color-coded according to tertiles of risk. Predicted risks were estimated by refitting the rML model to the full data set, with the modification that every patient was assumed to receive the betrixaban treatment. This was done to ensure that the risk stratification is due to additional covariate information, separate from the treatment effects demonstrated in the APEX trial.

Patient characteristics and outcomes in each tertile are described in Table 2. The lowest tertile included predicted risks between 0 and $\leq 3.3\%$, with a median risk of 2.7%. The second tertile represented intermediate risks ranging from $>3.3\%$ - $<5\%$, with a median of 5%. The third tertile predicted risks $\geq 5\%$, with a median of 7.4%. The observed VTE rates in the tertiles were 2.5% in the lowest tertile, 4.8% in the middle tertile, and 11.4% in the third

tertile. The odds of VTE in the second and third tertiles, compared to the lowest tertile, were 1.97 and 5.04, respectively. Table 3 shows VTE prediction risk estimates obtained from the rML model algorithms for 3 different acutely ill patient profiles, and Figure 3 marks these patients' risk of VTE on a plot of the distribution of risks in the APEX trial.

4 | DISCUSSION

This analysis is the first report of a machine learning model to predict VTE in acute medically ill patients. The super learner ensemble method demonstrated improved performance compared to the IMPROVE score, with a marginally higher c-statistic for predicting VTE through 77 days. Importantly, this method produced risk estimates that were well calibrated with observed outcomes. Patients predicted to be in the second and third tertiles of VTE risk were at

TABLE 2 Patient characteristics and outcomes according to predicted risk tertiles

	Lowest tertile (N = 2103)	Middle tertile (N = 2135)	Highest tertile (N = 2221)	P value
Treatment with betrixaban, n (%)	1031 (49.0)	1036 (48.5)	1140 (51.2)	.15
Treatment with enoxaparin, n (%)	1073 (51.0)	1099 (51.5)	1081 (48.8)	.15
Primary outcome event, n (%)	58 (2.8)	109 (5.1)	240 (10.8)	<.001
Age, y, mean (SD)	73.83 (7.24)	77.23 (7.40)	77.78 (9.40)	<.001
Male (%)	988 (47.0)	967 (45.3)	969 (43.6)	.09
Weight, kg, mean (SD)	85.17 (20.56)	78.53 (17.48)	78.11 (18.68)	<.001
BMI, kg/m ² , median (IQR)	29.60 (25.90-35.50)	28.10 (24.60-32.10)	27.50 (24.20-32.00)	<.001
Duration of hospitalization, days, median (IQR)	9.00 (7.00-13.00)	10.00 (7.75-14.00)	11.00 (8.00-15.00)	<.001
Creatinine clearance, mL/min, mean (SD)	75.84 (33.15)	69.93 (31.32)	68.08 (33.71)	<.001
Race (%)				
White	1996 (95.7)	1992 (95.1)	2075 (95.7)	.73
Black	37 (1.8)	35 (1.7)	44 (2.0)	
Asian	4 (0.2)	7 (0.3)	2 (0.1)	
American Indian	2 (0.1)	3 (0.1)	2 (0.1)	
Pacific Islander	0 (0.0)	1 (0.0)	0 (0.0)	
Multiple	14 (0.7)	18 (0.9)	12 (0.6)	
Other	32 (1.5)	39 (1.9)	33 (1.5)	
Use of strong P-gp inhibitor, n (%)	383 (18.2)	352 (16.5)	426 (19.2)	.07
D-dimer, median (IQR)	0.69 (0.39-1.11)	1.24 (0.76-1.83)	2.70 (1.52-4.37)	<.001
History of cancer, n (%)	217 (10.3)	275 (12.9)	267 (12.0)	.031
History of thrombosis, n (%)	2 (0.1)	13 (0.6)	497 (22.4)	<.001
Chronic heart failure, n (%)	500 (23.8)	467 (21.9)	503 (22.6)	.33
Acute infectious disease, n (%)	313 (14.9)	356 (16.7)	340 (15.3)	.24
Severe varicosities, n (%)	475 (22.6)	312 (14.6)	414 (18.6)	<.001
Hormone replacement, n (%)	4 (0.2)	14 (0.7)	41 (1.8)	<.001
Inherited or acquired thrombophilia, n (%)	2 (0.1)	2 (0.1)	3 (0.1)	.89

BMI, body mass index; IQR, interquartile; P-gp, P-glycoprotein; SD, standard deviation.

TABLE 3 Venous thromboembolism risk prediction using the reduced machine learning (rML) model in 3 different patient profiles

Clinical presentation/ description	Patient A	Patient B	Patient C
Variable name	Septic patient	COPD exacerbation patient	Acute decompensated heart failure patient
Age	45	77	86
BMI	35	27.5	18
Any ICU admissions	Yes	No	Yes
Duration of immobility, days	2	3	5
Heart failure at admission	No	No	Yes
Respiratory failure at admission	No	Yes	No
Acute infection at admission	Yes	No	No
Rheumatic disease at admission	No	No	No
NYHA CHF Class III or IV	No	No	Yes
History of thrombosis	No	Yes	Yes
Hormone replacement therapy	Yes	No	No
Active cancer	No	No	Yes
Lower limb paresis	No	Yes	Yes
Chronic respiratory failure	No	Yes	Yes
Protein	109	74	90
D-dimer	2500	750	1500
Predicted risk of VTE, n (%)	2.2	6.3	11.5
Accuracy	0.711 (0.540-0.813)		
Negative predictive value	0.962 (0.956-0.969)		
Positive predictive value	0.120 (0.095-0.158)		
Sensitivity	0.570 (0.430-0.748)		
Specificity	0.720 (0.527-0.837)		

BMI, body mass index; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit; NYHA, New York Heart Association; VTE, venous thromboembolism.

2- and 5-fold increased odds of suffering VTE compared to the lowest tertile.

Traditional risk assessment models have several limitations. They are built primarily using generalized linear models, which require substantial modeling skill and subject matter expertise. Use of poor techniques such as stepwise variable selection, unjustified dichotomization of continuous input variables, unprincipled handling of variable interactions, and other arbitrary procedures reduces generalizability to new patient populations. Additionally, the traditional use of integer scoring systems diminishes the predictive ability of risk assessment models by not allowing models to contain “protective” effects and inappropriately combining estimated variable effects (ie, adding odds ratios). Finally, higher-order interactions are rarely explored and are difficult to interpret or incorporate into integer scoring systems. These limitations can lead to predictions that were poorly calibrated to outcomes.

The use of a modified IMPROVE score to enrich for patients at high risk for VTE in the MARINER (Medically Ill Patient Assessment of Rivaroxaban Versus Placebo in Reducing Post-Discharge Thromboembolism Risk) trial presents a cautionary tale. In APEX, the IMPROVE score did not discriminate on outcomes with a similar score in those that did and did not have a VTE. Similarly, using an integer-based score in the MARINER trial underestimated the event rate.²¹

However, traditional risk scores offer the advantage of ease of use and simple computation by the physician at the point of care, which previously would have outweighed performance limitations. The IMPROVE score and the dual antiplatelet therapy (DAPT) score have a mobile application that has been downloaded over 100 000 times.²²⁻²⁴ Similarly, ML models may be implemented on mobile platforms that would optimize both the computational and convenience challenges that previously necessitated using oversimplified integer-based models.

The super learner ensemble algorithm provides advantages of allowing a priori specification of an algorithm that combines estimates from traditional regression models and machine learning algorithms and determines the weights to apply to each model.²⁵ This allows the data to “decide” the weights to apply to each method in order to achieve the best fit rather than arbitrarily choosing an approach that may not necessarily provide the best performance.

Our reduced model demonstrated similar discrimination compared to the complete model and qualitatively appeared more highly calibrated with patient outcomes through a wider range of risks and had a high *P* value for the Hosmer-Lemeshow test. This finding corroborates a previous study, which used “untransformed” and “reduced” super learner algorithms for the prediction of mortality in intensive care unit patients in which the reduced model performed similarly to the untransformed model.¹⁷

When administering antithrombotic agents to patients, clinicians consider the benefit and harm of the therapies in terms of thrombotic and bleeding risk. In the absence of reliable risk assessment tools, clinicians make subjective judgments based on knowledge of risk factors. The ideal risk assessment tool would simultaneously weigh the risk of thrombosis with the risk of bleeding. For example, increased age, renal insufficiency, aspirin treatment, hypertension, and diabetes are all primary risk factors for both bleeding and thrombosis and are included in both thrombosis (Thrombolysis in Myocardial Infarction, DAPT) and bleeding (HAS-BLED, PRECISE-DAPT). Disentangling the “high bleeding risk” patient from the “high thrombotic risk” patient would require a more dynamic, longitudinal collection of data on nontraditional risk factors. Though this type of data was not available in the present analysis, machine learning has shown promise in predicting treatment responses using longitudinal data.

4.1 | Limitations

Machine learning methods are often described as “black box,” as they do not provide information on the directionality or magnitude of effect for variables on the outcome. The predicted risk distribution in the APEX trial may not apply to other populations and serves only as a validation of progressively increasing risk across classes derived from this data set. Further, the APEX trial included a highly selected population of acute medically ill patients that had risk factors that made them at high risk for VTE. Trial participants were mostly Caucasian, and 70% of them were >75 years of age. Additionally, patients with active cancer or severe renal insufficiency were excluded. Thus, this model may not be generalizable to younger, non-Caucasian populations with severe renal insufficiency or active cancer and is not applicable to surgical patients. The composite end point of the APEX trial included asymptomatic DVT. Although several studies have demonstrated associations between asymptomatic DVT and short-term mortality, the clinical meaningfulness of this asymptomatic event is questionable. Thus, classification of risk into low-, intermediate-, and high-risk tertiles may not correspond to tertiles of risk in other populations. Finally, model

validation was performed within a single data set. External validation in a separate cohort is warranted.

5 | CONCLUSION

This analysis is the first to evaluate the performance of machine learning algorithms, built on randomized clinical trial data, for the prediction of VTE among acute medically ill patients. The super learner produced the highest c-statistic for prediction of VTE compared to the IMPROVE score and produced risk estimates that were well calibrated with observed outcomes. Patients in the lowest tertiles of risk had the lowest observed outcomes, while those in the highest tertile had the highest rate of VTE that corresponds to approximately 5-fold increased odds of VTE compared to the lowest tertile. Therefore, the rML model can be used to identify and group patients according to risk in a clinically meaningful way. Further research is needed to evaluate the effect of the use of this algorithm on outcomes in a clinical setting.

RELATIONSHIP DISCLOSURE

TN, CMG, RT, and MKY have received consulting fees from Portola Pharmaceuticals Inc. All other authors report nothing to disclose.

AUTHOR CONTRIBUTIONS

TN drafted the manuscript; developed the methodology; edited the text, tables, and figures; and critically appraised, consolidated comments, and applied all changes to the drafts of the manuscript. CMG conceived the idea, collected the data (as chairman of the executive committee of the APEX trial), and critically appraised the manuscript. RT and MKY performed statistical analyses and contributed significantly to the methods and results sections of the manuscript. MK, AK, GC, and FA critically appraised the manuscript. AFH, RDH, ATC, RAH, and SZG served on the executive committee of the APEX trial and were involved in data collection and provided critical commentary on this manuscript.

ORCID

Tarek Nafee  <https://orcid.org/0000-0002-8905-7745>

TWITTER

Tarek Nafee  @TarekOmarNafee
 C. Michael Gibson  @CMichaelGibson
 Adrian F. Hernandez  @TexHern
 Robert A. Harrington  @HeartBobH

REFERENCES

1. Cohen AT, Alikhan R, Arcelus JI, Bergmann J-F, Haas S, Merli GJ, et al. Assessment of venous thromboembolism risk and the benefits of thromboprophylaxis in medical patients. *Thromb Haemost.* 2005;94(4):750–9.
2. Amin AN, Varker H, Prinic N, Lin J, Thompson S, Johnston S. Duration of venous thromboembolism risk across a continuum in medically ill hospitalized patients. *J Hosp Med.* 2012;7(3):231–8.

3. Cohen AT, Spiro TE, Büller HR, Haskell L, Hu D, Hull R, et al. Rivaroxaban for thromboprophylaxis in acutely ill medical patients. *N Engl J Med.* 2013;368(6):513–23.
4. Goldhaber SZ, Leizorovicz A, Kakkar AK, Haas SK, Merli G, Knabb RM, et al. Apixaban versus enoxaparin for thromboprophylaxis in medically ill patients. *N Engl J Med.* 2011;365(23):2167–77.
5. Cohen AT, Harrington RA, Goldhaber SZ, Hull RD, Wiens BL, Gold A, et al. Extended thromboprophylaxis with betrixaban in acutely ill medical patients. *N Engl J Med.* 2016;375(6):534–44.
6. Spyropoulos AC, Anderson FA, FitzGerald G, Decousus H, Pini M, Chong BH, et al. Predictive and associative models to identify hospitalized medical patients at risk for VTE. *Chest.* 2011;140(3):706–14.
7. Stuck AK, Spirk D, Schaudt J, Kucher N. Risk assessment models for venous thromboembolism in acutely ill medical patients. A systematic review. *Thromb Haemost.* 2017;117(4):801–8.
8. Barbar S, Noventa F, Rossetto V, Ferrari A, Brandolin B, Perlati M, et al. A risk assessment model for the identification of hospitalized medical patients at risk for venous thromboembolism: the Padua prediction score. *J Thromb Haemost.* 2010;8(11):2450–7.
9. Tapson VF, Decousus H, Pini M, Chong BH, Froehlich JB, Monreal M, et al. Venous thromboembolism prophylaxis in acutely ill hospitalized medical patients: findings from the international medical prevention registry on venous thromboembolism. *Chest.* 2007;132(3):936–45.
10. Rosenberg D, Eichorn A, Alarcon M, McCullagh L, McGinn T, Spyropoulos AC. External validation of the risk assessment model of the International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) for medical patients in a tertiary health system. *J Am Heart Assoc.* 2014;3(6):e001152.
11. Vardi M, Ghanem-Zoubi NO, Zidan R, Yurin V, Bitterman H. Venous thromboembolism and the utility of the Padua prediction score in patients with sepsis admitted to internal medicine departments. *J Thromb Haemost.* 2013;11(3):467–73.
12. Matthew Cerasale SMM, Foraone H, Kaatz S, Ritz J, Watson P, Jordan J. Padua to IMPROVE: matching VTE risk stratification tool to the EHR. *J Hosp Med.* 2017;12(suppl 2): 229.
13. Greene MT, Spyropoulos AC, Chopra V, Grant PJ, Kaatz S, Bernstein SJ, et al. Validation of Risk Assessment Models of Venous Thromboembolism in Hospitalized Medical Patients. *Am J Med.* 2016;129(9):1001.e9–1001.e18.
14. Deo RC. Machine learning in medicine. *Circulation.* 2015;132(20): 1920–30.
15. Mackie P, Sim F, Johnman C. Big data! Big deal? *Public Health.* 2015;129(3):189–90.
16. Golden JA. Deep learning algorithms for detection of lymph node metastases from breast cancer: helping artificial intelligence be seen. *JAMA.* 2017;318(22):2184–6.
17. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med.* 2015;3(1):42–52.
18. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA.* 2016;316(22):2402–10.
19. Sengupta PP, Kulkarni H, Narula J. Prediction of abnormal myocardial relaxation from signal processed surface ECG. *J Am Coll Cardiol.* 2018;71(15):1650–60.
20. Cohen AT, Harrington R, Goldhaber SZ, Hull R, Gibson CM, Hernandez AF, et al. The design and rationale for the acute medically ill venous thromboembolism prevention with extended duration betrixaban (APEX) study. *Am Heart J.* 2014;167(3): 335–41.
21. Spyropoulos AC, Ageno W, Albers GW, Elliott CG, Halperin JL, Hiatt WR, et al. Rivaroxaban for thromboprophylaxis after hospitalization for medical illness. *N Engl J Med.* 2018;379(12):1118–27.
22. Thromboembolism IMPROV. Risk assessment tool for estimating the risk of acute VTE: international medical prevention registry on venous thromboembolism. Available from: https://www.outcomes-umassmed.org/IMPROVE/risk_score/vte/index.html
23. Yeh RW, Secemsky EA, Kereiakes DJ, Normand S-L, Gershlick AH, Cohen DJ, et al. Development and validation of a prediction rule for benefit and harm of dual antiplatelet therapy beyond 1 year after percutaneous coronary intervention. *JAMA.* 2016;315(16):1735–49.
24. Cardiology ACo. DAPT Risk Calculator: American College of Cardiology; 2017. Available from: <http://tools.acc.org/DAPTriskapp/#!/content/calculator/>
25. van der Laan MJ, Eric C, Hubbard AE. Super Learner. *Statistical Applications in Genetics and Molecular Biology.* 2007;6: 25.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Nafee T, Gibson CM, Travis R, et al. Machine learning to predict venous thrombosis in acutely ill medical patients. *Res Pract Thromb Haemost.* 2020;4:230–237. <https://doi.org/10.1002/rth2.12292>