

# Converting IMPROVE bleeding and VTE risk assessment models into a fast-and-frugal decision tree for optimal hospital VTE prophylaxis

Benjamin Djulbegovic,<sup>1</sup> Alice Boylan,<sup>1</sup> Shelby Kolo,<sup>1</sup> Danielle Bowen Scheurer,<sup>1</sup> Sydney Anuskiewicz,<sup>1</sup> Flora Khaledi,<sup>1</sup> Kimberley Youkhana,<sup>1</sup> Sarah Madgwick,<sup>1</sup> Neema Maharjan,<sup>1</sup> and Iztok Hozo<sup>2</sup>

<sup>1</sup>Medical University of South Carolina, Charleston, SC; and <sup>2</sup>Department of Mathematics, Indiana University Northwest, Gary, IN

## Key Points

- American Society of Hematology has endorsed using risk assessment models (RAMs) to individualize VTE prophylaxis.
- Fast-and-frugal decision tree based on IMPROVE RAMs reduces unnecessary VTE prophylaxis by 45% and saves ~\$5 million annually.

Current hospital venous thromboembolism (VTE) prophylaxis for medical patients is characterized by both underuse and overuse. The American Society of Hematology (ASH) has endorsed the use of risk assessment models (RAMs) as an approach to individualize VTE prophylaxis by balancing overuse (excessive risk of bleeding) and underuse (risk of avoidable VTE). ASH has endorsed IMPROVE (International Medical Prevention Registry on Venous Thromboembolism) risk assessment models, the only RAMs to assess short-term bleeding and VTE risk in acutely ill medical inpatients. ASH, however, notes that no RAMs have been thoroughly analyzed for their effect on patient outcomes. We aimed to validate the IMPROVE models and adapt them into a simple, fast-and-frugal (FFT) decision tree to evaluate the impact of VTE prevention on health outcomes and costs. We used 3 methods: the “*best evidence*” from ASH guidelines, a “*learning health system paradigm*” combining guideline and real-world data from the Medical University of South Carolina (MUSC), and a “*real-world data*” approach based solely on MUSC data retrospectively extracted from electronic records. We found that the most effective VTE prevention strategy used the FFT decision tree based on an IMPROVE VTE score of  $\geq 2$  or  $\geq 4$  and a bleeding score of  $< 7$ . This method could prevent 45% of unnecessary treatments, saving ~\$5 million annually for patients such as the MUSC cohort. We recommend integrating IMPROVE models into hospital electronic medical records as a point-of-care tool, thereby enhancing VTE prevention in hospitalized medical patients.

## Introduction

Fifteen years ago, the US Surgeon General issued a call to action to prevent deep venous thrombosis and pulmonary embolism (VTE) as one of the most serious but avoidable hospital complications.<sup>1</sup> Since then, the focus has remained on reducing the underuse of VTE prophylaxis in hospitalized patients, resulting in “mandatory use of a standardized medical order set in which a prescription for prophylaxis is embedded.”<sup>2</sup> Reducing the underuse of VTE prophylaxis has also been promoted as one of the critical quality measures by most national quality organizations in the United States.<sup>3</sup> Increasingly, however, overuse is being recognized as a more significant problem. Studies in the last few years have demonstrated that 78% to 86% of medical hospitalized patients at low risk for VTE have received unnecessary pharmacologic VTE prophylaxis.<sup>4</sup> As a result, there is increasing recognition that some

Submitted 14 March 2024; accepted 4 April 2024; prepublished online on *Blood Advances* First Edition 15 April 2024. <https://doi.org/10.1182/bloodadvances.2024013166>.

Data are available on request from the corresponding author, Benjamin Djulbegovic ([djulbegov@musc.edu](mailto:djulbegov@musc.edu)).

The full-text version of this article contains a data supplement.

© 2024 by The American Society of Hematology. Licensed under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International \(CC BY-NC-ND 4.0\)](https://creativecommons.org/licenses/by-nc-nd/4.0/), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

patients at higher risk of VTE need hospital VTE prophylaxis, whereas others, considered to be at low risk of VTE, do not.<sup>5</sup> Most recently, influential guideline panels, such as the American Society of Hematology (ASH), have endorsed the use of risk assessment models (RAMs) as an approach to individualizing VTE prophylaxis as a way of finding a balance between overuse (excessive risk of bleeding) and underuse (risk of avoidable VTE).<sup>6,7</sup>

ASH has also endorsed IMPROVE (International Medical Prevention Registry on Venous Thromboembolism) VTE RAM to guide VTE prophylaxis.<sup>8</sup> IMPROVE is the only existing RAM for assessing both short-term bleeding (IMPROVE\_bleeding)<sup>9</sup> and VTE risk in acutely ill medical inpatients (IMPROVE\_vte).<sup>8</sup> It has been externally validated in US,<sup>10-12</sup> Canadian,<sup>13</sup> Chinese,<sup>14</sup> and European populations<sup>15</sup> with acceptable discrimination (area under the curve [AUC] = 0.65 to 0.77 [VTE score]; AUC = 0.63 to 0.76 [bleeding score]) and calibration properties.

Although RAMs can (accurately) assess the risk of bleeding and VTE, to date, their use has not been proven to aid decision-making and improve health outcomes. Additionally, no recommendations exist for integrating these 2 predictive models (ie, IMPROVE\_bleeding and IMPROVE\_vte) into 1 coherent management strategy. We previously demonstrated that this can only be accomplished within a decision-theoretical framework.<sup>16-18</sup> One such framework consists of fast-and-frugal decision trees (FFTs), effective implementation of heuristics for problem-solving, and decision-making strategies composed of sequentially ordered cues (tests) and binary (yes/no) decisions formulated via a series of if-then statements reflective of the heuristic-analytic, adaptive theory of human reasoning that dominates clinical decision-making.<sup>17-21</sup> Unlike the popular clinical pathways (algorithms), which are theory-free constructs, FFTs are based on solid theoretical foundations.<sup>20,21</sup> The latter provides a basis for quantitative evaluation of the classification capacity of FFTs. Importantly, clinical algorithms can be easily converted into FFTs to facilitate quantitative assessment of their classification accuracy.<sup>17-21</sup>

In this article, we describe the conversion of IMPROVE bleeding and VTE predictive models to FFT to identify the optimal recommendation for VTE prophylaxis in hospitalized patients.

## Methods

### Practice guidelines for using IMPROVE RAMs

The original IMPROVE VTE risk article<sup>8</sup> and ASH<sup>6,7</sup> guidelines recommend VTE prophylaxis if the IMPROVE VTE risk score is  $\geq 2$ . This converts into the probability of VTE between 1% and 1.5% in the original report.<sup>8</sup> Some authors recommend an IMPROVE score of  $\geq 4$  to indicate thromboprophylaxis in hospitalized patients.<sup>22</sup> The IMPROVE bleeding risk report<sup>9</sup> and ASH<sup>6,7</sup> guidelines recommend IMPROVE bleeding score of  $<7$  and  $\geq 7$  as a low or high bleeding risk to proceed with or avoid VTE prophylaxis, respectively. According to the original publication,<sup>9</sup> IMPROVE bleeding score  $<7$  corresponds to 1.5% of major and clinically relevant bleeding. IMPROVE bleeding score  $\geq 7$  was observed in 7.9% of the patients who experienced major and clinically relevant bleeding.

### Conversion of IMPROVE RAMs into FFT

The cutoffs identified for VTE and bleeding scores lend them naturally to integration into the FFT decision tree as a

decision-support strategy. Figure 1 shows the FFT based on the integration of VTE and bleeding IMPROVE scores in the comprehensive risk-adapted VTE prophylaxis. As shown, we started with the “first, do no harm” decision principle: patients with a high bleeding score ( $\geq 7$ ) should not be given prophylaxis (first cue in Figure 1). Those with bleeding scores  $<7$  are then assessed for the risk of VTE. Then, if the IMPROVE VTE risk score is  $\geq 2$  or  $\geq 4$  (see “Results”), they will be given pharmacological VTE prophylaxis.

By categorizing the disease or making predictions, FFT also functions as a predictive model.<sup>21</sup> This enables personalized forecasting of outcomes and tailored management based on a sequence of responses within the FFT decision tree, while considering the heterogeneity of patients’ risk profiles.<sup>21,23,24</sup>

There are 2 variants of FFT: (a) standard versions of FFT, which aim to classify a condition of interest (ie, whether the patient has VTE or not) but do not take the consequences of treatment into account,<sup>17,19,21</sup> (b) a variant of FFT referred to as FFT with threshold (FFTT),<sup>17,19,21</sup> which incorporates benefits and harms of treatment at each exit of the tree to indicate whether treatment such as VTE prophylaxis is indicated depending on the risk for VTE or bleeding concerning the treatment threshold at a given cue.<sup>16,25,26</sup> Below, we present the impact and classification accuracy of FFT/T-based management strategies (ie, overall accuracy, sensitivity, specificity, and positive and negative predictive values).

### Using real-world data to collect data on IMPROVE RAMs

Patients were eligible for the analysis if they were aged  $\geq 18$  years and had  $\geq 3$  days of hospitalization for an acute medical illness consecutively admitted to the Medical University of South Carolina (MUSC) hospital in Charleston, SC. Patients were excluded if they received anticoagulants or thrombolytics at admission or within 48 hours after admission, if they had bleeding at admission, had major surgery or trauma within 3 months before admission, or were admitted for treatment of VTE. Admissions for major surgeries were also excluded and patients with primary obstetric or mental health diagnoses.

We extracted data on the key features of IMPROVE VTE and bleeding RAMs from MUSC electronic medical (EPIC)/discharge records. Supplemental Appendix 0 shows the ICD10 (International Classification of Diseases) and other codes used to select the variables of interest for the analysis. The outcomes of interest were

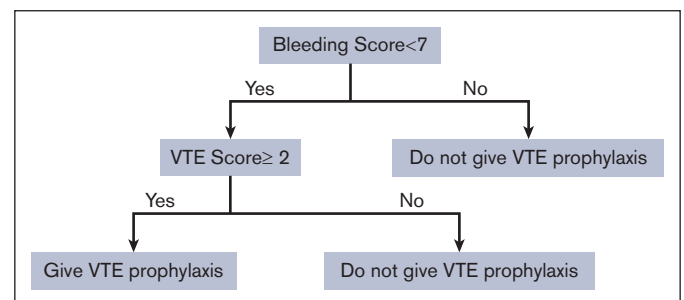


Figure 1. An FFT was constructed using IMPROVE VTE and bleeding RAMs.

VTE at 90 days and bleeding (major and clinically significant bleeding) at 14 days. Two reviewers manually verified all the outcomes (VTE and bleeding). The ascertainment of exposure (whether the patients received VTE prophylaxis) was conducted using 20% of randomly selected records. The Gwet agreement coefficient<sup>27</sup> between the 2 observers was high; the accuracy for classification of VTE outcome was 0.95 (95% CI, 0.94-0.96) for VTE and 0.93 (95% CI, 0.92-0.94) for bleeding outcomes. The overall accuracy of ascertainment of the exposure was 96%.

Enoxaparin at the recommended prophylactic dose (40 mg daily or 30 mg twice a day, further adjusted for renal function as needed)<sup>28</sup> was almost exclusively used in our patient cohort. Patients who received therapeutic/intermediate-intensity anticoagulation<sup>28</sup> (supplemental Appendix 2, Table 1C) were excluded (n = 601). Patients who received aspirin (n = 132) were included in the prophylactic arm in light of the evidence that aspirin has prophylactic effects in a surgical setting.<sup>29</sup> Patients who received antiplatelet combinations without aspirin, heparin flushes to keep the IV lines open, and thrombolytics (alteplase) for the treatment of thrombotically occluded catheters (n = 42) were included in the control arm (supplemental Appendix 2, Table 1B).

### External validation of IMPROVE models

To assess the generalizability of the original IMPROVE models,<sup>8,9</sup> we assessed their performance using the MUSC data. Both the original and external validations of IMPROVE models<sup>10-15</sup> were performed agnostic of prophylactic treatments. The model performance was assessed by calculating the discrimination and calibration properties.<sup>30</sup> Discrimination (capacity of the predictive model to discriminate between patients who truly have outcomes vs those who do not have an outcome of interest) was assessed using c-statistics. Calibration (capacity to determine the agreement between predicted and actually observed outcomes) was assessed using the Hosmer-Lemeshow goodness-of-fit statistic (H-L) and examining intercept, that is, whether the best-fit line of the model crossed the y-axis at zero predicted probability and whether the slope was statistically significantly different from 1. In the case of poor external validation of the model, we assessed the internal validity of IMPROVE models using MUSC data.

### Impact analysis: comparison of various VTE prophylaxis strategies on health outcomes

To estimate the predictive impact of VTE prophylaxis using IMPROVE models, we first used the best evidence approach, relying on the effects of thromboprophylaxis as reported in the ASH guidelines.<sup>6,7</sup> Specifically, we assumed that thromboprophylaxis with medications such as enoxaparin reduces the risk of VTE by 41% ( $RRR = 1 - RR = 1 - 0.59 = 41\%$ ), where RR refers to the relative risk and RRR refers to the relative risk reduction of VTE due to thromboprophylaxis compared with no treatment. The ASH estimated relative risk increase (RRI) of clinically significant bleeding to be 48% higher on thromboprophylaxis than in no treatment arms.<sup>6</sup>

We compared the effects of the following management strategies on the number of bleeding events (major and clinically significant) and VTE events (supplemental Appendix 1): (1) administer no VTE prophylaxis to medically hospitalized patients; (2) administer VTE prophylaxis to all hospitalized patients admitted for treatment

medical disease (current default management strategy in most institutions); (3) administer VTE prophylaxis if VTE risk exceeds 1% (as per the original IMPROVE recommendations<sup>29</sup>); (4) administer VTE prophylaxis if VTE score is  $\geq 2$  (as per ASH guidelines<sup>6</sup> and the original IMPROVE article<sup>9</sup>); (5) administer VTE prophylaxis if VTE score is  $\geq 4$  (as per ref#<sup>22</sup>); (6) apply VTE prophylaxis if the bleeding score is  $< 7$  (as per IMPROVE bleeding score recommendation<sup>9</sup>); (7-10) administer VTE prophylaxis as per FFT/T (Figure 1) by assuming VTE risk  $\geq 2$  or  $\geq 4$ , respectively.

The heart of personalized decision-making consists of contrasting risk assessment against decision thresholds,<sup>17,31</sup> which, in turn, are determined based on the benefits and harms of the treatment. The threshold model that is particularly applicable to the problem of VTE prophylaxis<sup>16,25,26</sup> can be expressed as:

$$T = \frac{RV \cdot (B_{rx} - B_{norx})}{RRR} \quad (\text{equation 1})$$

where RRR represents the treatment effect on VTE expressed as relative risk reduction;  $(B_{rx} - B_{norx})$  refers to the absolute difference between major and clinically significant bleeding in patients who received pharmacological VTE prophylaxis and those who did not at each “cue” of FFT/T decision tree shown in Figure 1; RV (relative value) refers to patients’ values and preferences. When  $RV < 1$ , the patient values avoid the consequences of VTE more than avoiding treatment-related bleeding outcomes; if  $RV > 1$ , the patient places more importance on avoiding bleeding than on the consequences of VTE. When the patient is indifferent between the consequences of bleeding and the consequences of the disease outcome, the RV is set to 1. When  $RV = 1$ , the threshold is solely determined by empirical evidence.<sup>26,32</sup> Our default impact metrics comprised the weighted average of VTE and bleeding events, calculated as  $N_{both} = (N_{vte} + RV \times N_{bleed}) / (1 + RV)$ . We also performed a sensitivity analysis of the impact of RV on the decision whether to recommend prophylaxis by ranging it from 0.75 to 1.33, which is within the estimated empirical assessment of patients’ RV.<sup>33</sup> We also used costs as the impact metrics by projecting the costs of unnecessary VTE prophylaxis from published reports<sup>34,35</sup> to January 2024 dollar costs (supplemental Appendix 8).

According to equation 1, if the probability (risk) of VTE,  $p_{VTE} > T$  we should give treatment; otherwise, we should not provide VTE prophylaxis.<sup>25,26,36,37</sup> Note that according to classic statistical decision theory, when choosing between different options, the most rational choice is the one with the highest expected utility, regardless of statistical significance; the magnitude of differences is irrelevant.<sup>38</sup>

Because patients studied in clinical trials differ from “real-world” patients,<sup>39,40</sup> a fundamental concept behind the “learning health system paradigm”<sup>41</sup> is to evaluate the impact of the best available research evidence in a given population of interest. Therefore, we further extended the modeling of the impact of VTE prophylaxis on health outcomes by using distributions of prognostic scores from our MUSC population while assuming the applicability of the best available evidence as per ASH guidelines.<sup>6,7</sup>

Finally, we performed the same analysis using exclusively MUSC “real-world” data. From the data, we estimated  $RRR_{vte} = 0.058$  and

$RRR_{\text{bleeding}} = 0.197$ ; however, the results for RRR were not significant, potentially indicating a null effect of prophylaxis on VTE (see “Discussion”). All statistical analyses were performed using the statistical package Stata<sup>42</sup> and reported according to Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)<sup>43</sup> and TRIPOD<sup>44</sup> (Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) guidelines.

## Results

Figure 2 shows a STROBE flow diagram displaying the study patients’ enrollment, exclusion criteria, and data availability for the analysis according to anticoagulation treatment. From 1 January 2022 to 31 December 2022, 5051 consecutive patients were admitted to the MUSC hospital in Charleston, SC, of whom 2072 patients were medical patients eligible for the analysis per the criteria in the original IMPROVE reports.<sup>8,9</sup> After further restriction to the observations within 14 days (bleeding) and 90 days (VTE) and exclusion of patients who received therapeutic/intermediate intensity anticoagulation,<sup>28</sup> we ended up with 1429 analyzable patients for the impact analysis. The median length of stay was 5 days (range; 3-85 days). Note that we could not collect data on the “immobility >7 days” due to a lack of adequate ICD10 codes.

Tables 1 and 2 show the distribution of IMPROVE VTE and bleeding risk predictors in our population as per the intention-to-treat population. Not surprisingly, the case-mix differed between

our and the original IMPROVE publications,<sup>8,9</sup> primarily due to a higher number of patients requiring admission to the intensive care unit during the index hospitalization in our vs the original IMPROVE cohorts (30.2% vs 4.89<sup>8</sup> and 8.5%<sup>9</sup>, respectively). However, the risk score proportion distribution was similar despite dissimilarities in the predictors distribution (Table 3).

Figure 3 shows the bleeding and VTE rates as a function of the recommended discriminatory scores: <7 vs ≥7 for bleeding and <2 vs ≥2 VTE scores. Bleeding scores showed a statistically significant absolute risk difference of 2.9% (95% CI, 0.5%-5.3%;  $P = .001$ ), whereas no statistical difference was noted for VTE scores (see “Discussion”). Supplemental Appendix 3 shows the analysis by VTE prophylaxis, which is consistent with what is observed in the original publications<sup>8,9</sup> and other validation studies;<sup>10-15</sup> treatment did not affect the IMPROVE prognostic scores.

Supplemental Appendix 4, Figure 1.1 shows the external validation of IMPROVE bleeding scores with discrimination of AUC = 0.799 (95% CI, 0.731-0.848), similar to that reported in the original IMPROVE report<sup>9</sup> and other external validation studies.<sup>13-15</sup> After updating the intercept,<sup>45,46</sup> the model was well calibrated with a not significant H-L test ( $P = .287$ ), with an intercept close to 0 and slope of 1.016, which was not statistically significantly different from 0 and 1, respectively. However, we could not externally validate the IMPROVE VTE risk model using the original IMPROVE scores. Nevertheless, using our data with the same predictors, we obtained similar discrimination (AUC = 0.704) as was reported in

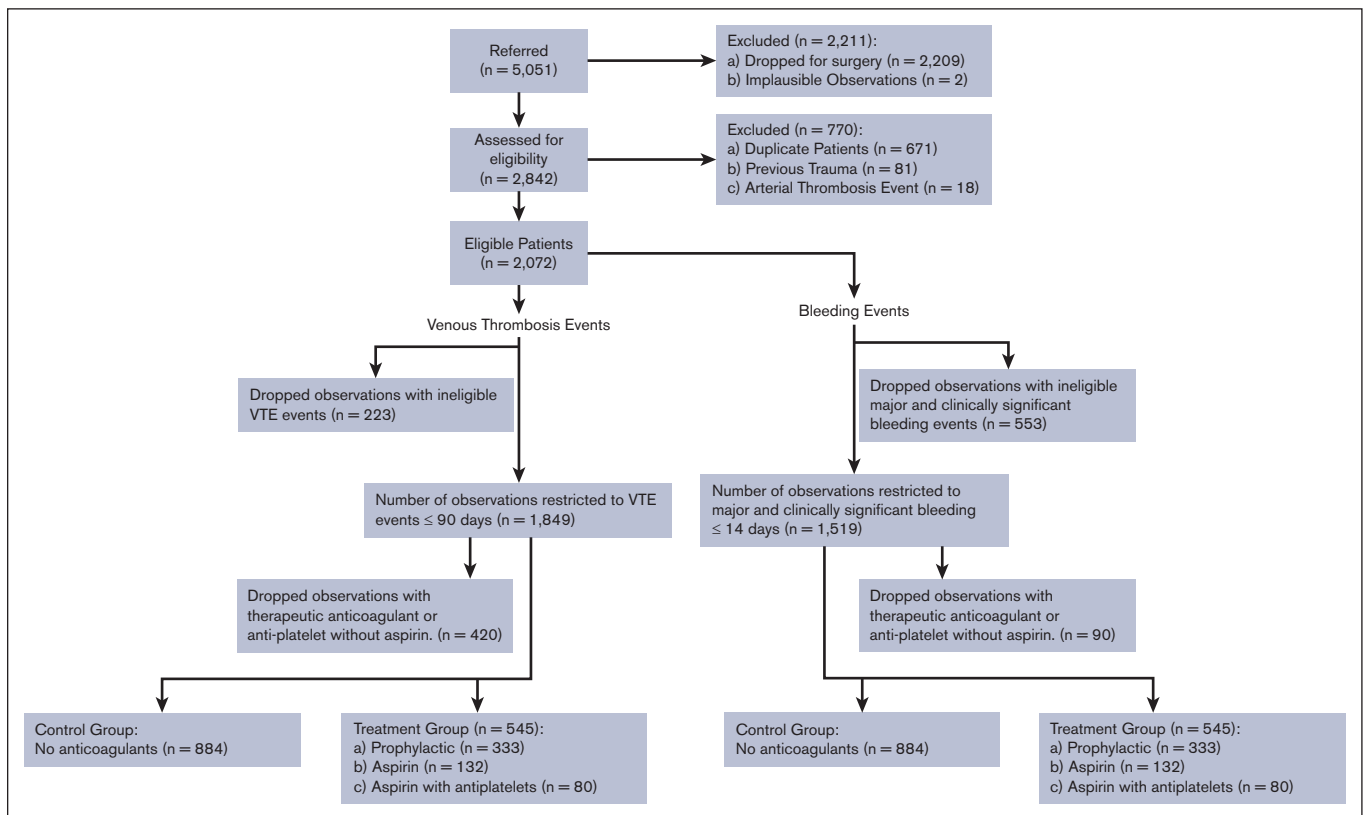


Figure 2. STROBE flow diagram displaying the study patients’ enrollment, exclusion criteria, and data availability for the analysis according to anticoagulation treatment.

**Table 1. VTE baseline characteristics and outcomes**

Predictors N = 2072	%
Age >60	56.2
Current cancer	16.8
Previous VTE	0.7
ICU (yes)	30.2
Lower limb paralysis	0.8
Thrombophilia	1.0
Immobility >7 days	NA
VTE_Outcome (yes)	2.2

NA, not available.

the original IMPROVE score<sup>8</sup> and excellent calibration (H-L test:  $P = .927$ , with an intercept of 0 and slope = 1) (supplemental Appendix 4; Figure 1.2). Importantly, however, both VTE scores and particularly bleeding scores were almost perfectly discriminatory regarding the predicted risk above and below the scores used in FFT (Figure 4; and supplemental Appendix 3; Figure 1.3).

Figures 5A and B) show FFT and FFTT decision trees informed by IMPROVE risk assessment scores  $\geq 2$  (5A) and  $\geq 4$  (5B) and evidence-based ASH guidelines.<sup>6,7</sup> The results were consistent with our hypothesis that the FFT(T)-based decision tree generated by linking VTE and bleeding IMPROVE scores (Figure 1) leads to optimal classification and decision making. We also determined that the tree was fast, reaching the decision in less than 2 steps, and frugal, using only ~5% of all information. In this case, both FFT and FFTT had identical classifications, with an overall accuracy of 93% for VTE scores  $\geq 4$  (Figure 5B) and 72% for VTE scores  $\geq 2$  (Figure 5A), respectively. However, given the low prevalence of outcomes (Figure 3), as expected, the negative predictive value was much higher (>99% for both VTE scores) than the positive

**Table 2. Bleeding baseline characteristics and outcomes**

Predictors N = 2072	%
GFR 30-59	14.4
Male	49.6
Age 40-84	77.6
Current cancer	16.8
Rheumatic diseases	7.4
Central venous catheter	34.5
ICU (yes)	30.2
Severe renal failure GFR <30	11.7
Hepatic failure	2.4
Age >85	8.3
Platelet count <50.10 <sup>2</sup> cells/L	0.9
Bleeding before admission	2.8
Active gastro ulcer	0.0
Bleed_outcome (yes)	1.6

GFR, glomerular filtration rate.

**Table 3. Distribution of IMPROVE risk scores used in ASH guidelines**

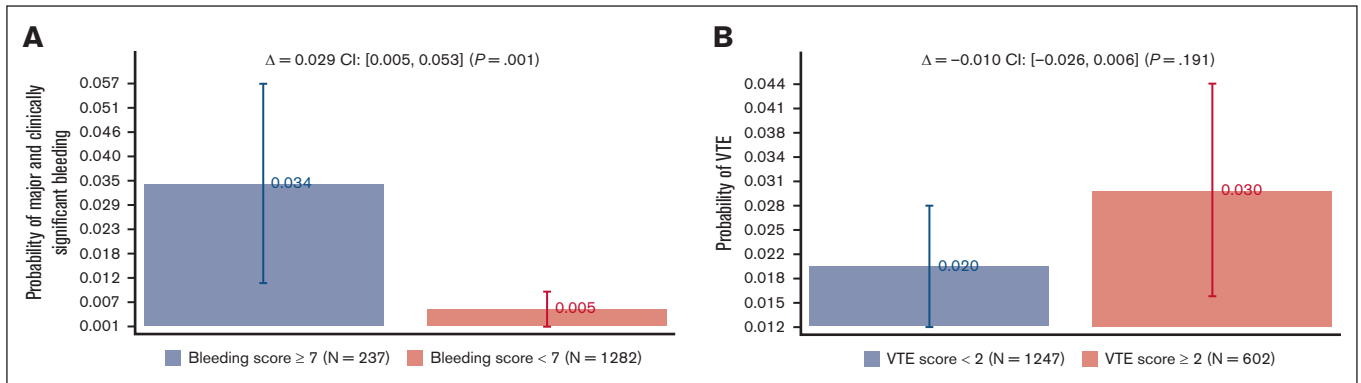
VTE	ASH/original IMPROVE scores (%)	MUSC
0 or 1	69	66
2 or 3	24	30
$\geq 4$	7	4
Bleeding		
<7	90	79
$\geq 7$	10	21

predictive value of 5.6% for FFT(T) using a VTE score  $\geq 4$  and 2.44% for a VTE score  $\geq 2$ . Figure 5C shows an impact analysis comparing 10 different management strategies. The 2 FFT(T)-driven strategies (Rx07-Rx10) were the best, resulting in the lowest weighted average and optimal trade-offs between bleeding and VTE. All FFT(T) strategies outperformed other currently recommended managements. Strategy Rx2 ("treat all") (presently recommended strategy) was the worst, yielding the highest number of weighted average outcomes driven by an increased number of bleedings, even though, as expected, that led to the lowest VTE event rates. The opposite was observed for the Rx1 strategy ("Treat none"), which was the second-worst management option. The changes in RV did not materially affect the results (supplemental Appendix 5).

Figure 6 shows the analysis according to the "learning health system"<sup>41</sup> paradigm, the application of the best research evidence to local data. In our case, this consisted of using MUSC bleeding and VTE risk scores  $\geq 2$  while using ASH guidelines as a source of the most reliable ("best") evidence.<sup>6,7</sup> The results agree with the analysis based on the "best evidence" shown in Figure 5, with a slightly higher overall classification accuracy of 74% and positive predictive value of 6%, while retaining a higher than 99% negative predictive value. Supplemental Appendix 6 shows the results of the sensitivity analysis using the RV.

Supplemental Appendix Figure 7 shows the FFT and FFTT decision tree using real-world MUSC data for VTE, bleeding risks, and treatment data. As shown in Figure 3 and supplemental Appendix 3, in our data set, we detected a significant effect of VTE prophylaxis on bleeding but not VTE outcomes. This resulted in unrealistic estimates of RRR of 5.8% and RRI of 19.7%. The results show that, as expected, when treatment is ineffective but results in more harm, the "no treatment" strategy represents the best option. However, assuming ineffective treatment disagrees with our current understanding of the best evidence (see "Discussion"). Note that if  $RV = 0.75$ , FFT-driven VTE prophylaxis after avoiding prophylaxis in patients at a high risk of bleeding represents the best management option, even if assuming such low effectiveness of treatment (supplemental Appendix 7).

Figure 1 shows that we can define unnecessary VTE prophylaxis as (*overuse*) = VTE prophylaxis given for bleeding score  $\geq 7$  or VTE score <2; failure to administer VTE prophylaxis (*underuse*) = VTE prophylaxis not given if VTE score  $\geq 2$  and bleeding score <7. By applying this FFT(T) within the "learning health system" paradigm, we can calculate that ~44% of inappropriate VTE prophylaxis (of which 13% and 31% consist of underuse and overuse,



**Figure 3.** Major and clinically significant bleeding and VTE rates as a function of the recommended IMPROVE models' discriminatory scores:  $<7$  vs  $\geq 7$  for bleeding and  $<2$  vs  $\geq 2$  VTE scores (see Figure 1).

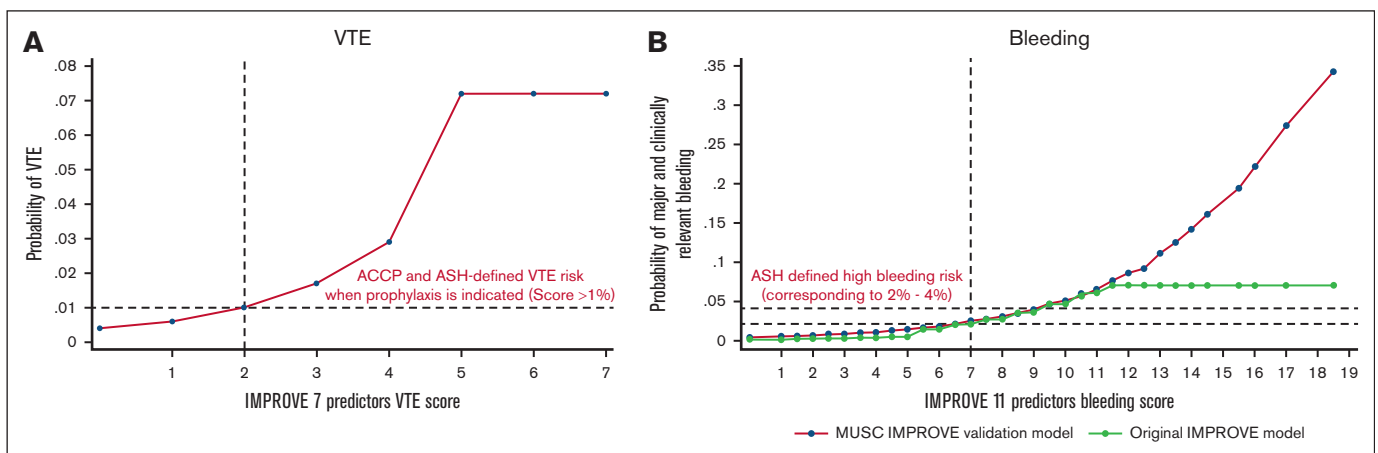
respectively) could be avoided (Figure 7) at annual savings due to avoiding resource waste at ~\$5 million per our patient cohort (supplemental Appendix Figure 8).

## Discussion

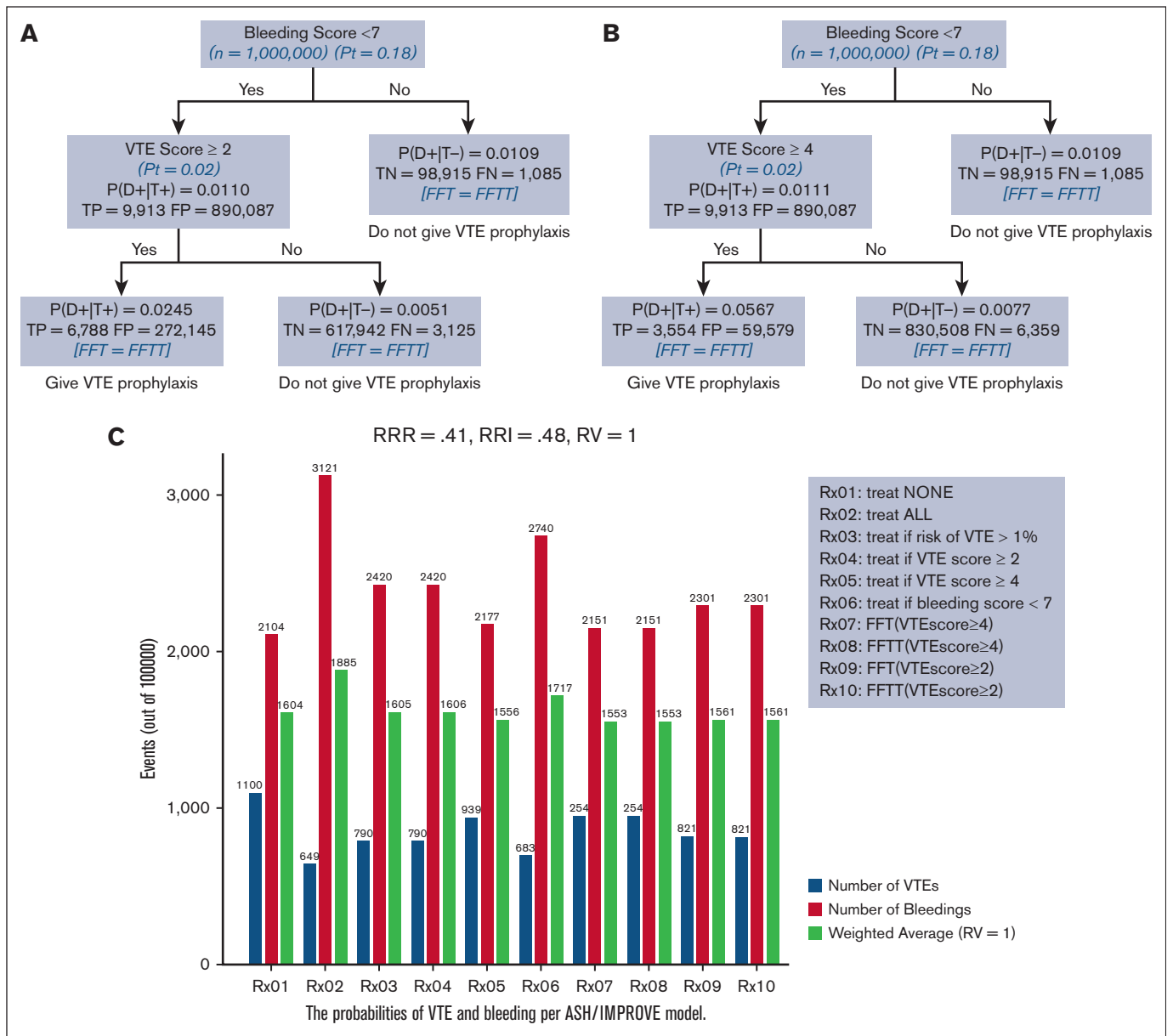
Current hospital VTE prophylaxis for medical patients is suboptimal. It has historically been characterized by underuse but increasingly by overuse, with  $>85\%$  of patients receiving inappropriate VTE prophylaxis in some studies.<sup>4</sup> As a result, calls have been made to “deimplement” VTE prophylaxis at low-risk for VTE and high-bleeding risk patients.<sup>5</sup> To minimize both underuse and overuse, the ASH guidelines on VTE prophylaxis state that when “clinicians and health care systems use ASH VTE guidelines, they should integrate VTE and bleeding risk assessments into clinical decision-making processes.”<sup>6</sup> However, the ASH guidelines also recommend that RAM-based decision-making needs further clinical outcomes impact evaluation.<sup>6</sup> In this article, we showed that the application of RAMs can lead to the development of decision support that is superior to other alternative approaches to VTE prophylaxis. We accomplished this by converting IMPROVE VTE and bleeding risk models<sup>3,9</sup> into FFT(T) decision trees to enable the quantification and impact assessment of competing clinical VTE prophylaxis strategies.<sup>17-21,47</sup> FFT draws on its theoretical

robustness by relating it to signal detection theory, evidence accumulation theory, and the threshold model to help improve decision making.<sup>17-21,47</sup> Importantly, we observed a meaningful impact within a “learning health systems”<sup>41</sup> framework, aiming to bridge the gaps between efficacy data (the best evidence observed in ideal research settings) and real-world case-mix data that often do not reflect controlled research settings.<sup>39,40</sup> In doing so, we demonstrated that  $\sim 45\%$  of inappropriate VTE prophylaxis, amounting to  $\sim \$5$  million in resource waste annually, could be avoided (Figure 7; supplemental Appendix 8).

Our study adds to a series of other studies that have externally validated the IMPROVE bleeding scores in various hospital settings,<sup>13-15</sup> testifying to the utility of the IMPROVE model. Unfortunately, unlike other studies,<sup>10,12,13</sup> we could not externally validate the IMPROVE VTE risk model. This is likely because our sample size was small ( $n = 1849$ , Figure 3B) as opposed to the 15 156 patients enrolled in the original study.<sup>8</sup> Another reason, as explained earlier, is that our population differed from the original IMPROVE cohort in the number of transfers to the intensive care unit during index hospitalization (30.2% vs 4.89%).<sup>13</sup> Finally, we did not collect data on 1 of the IMPROVE predictors (immobility  $>7$  days) (Tables 1 and 2). Despite this, we could internally validate the IMPROVE VTE score and obtain almost identical risk scores, as



**Figure 4.** Relationship between IMPROVE VTE scores (A) and bleeding scores (B) and the probability of VTE and bleeding (major and clinically relevant).



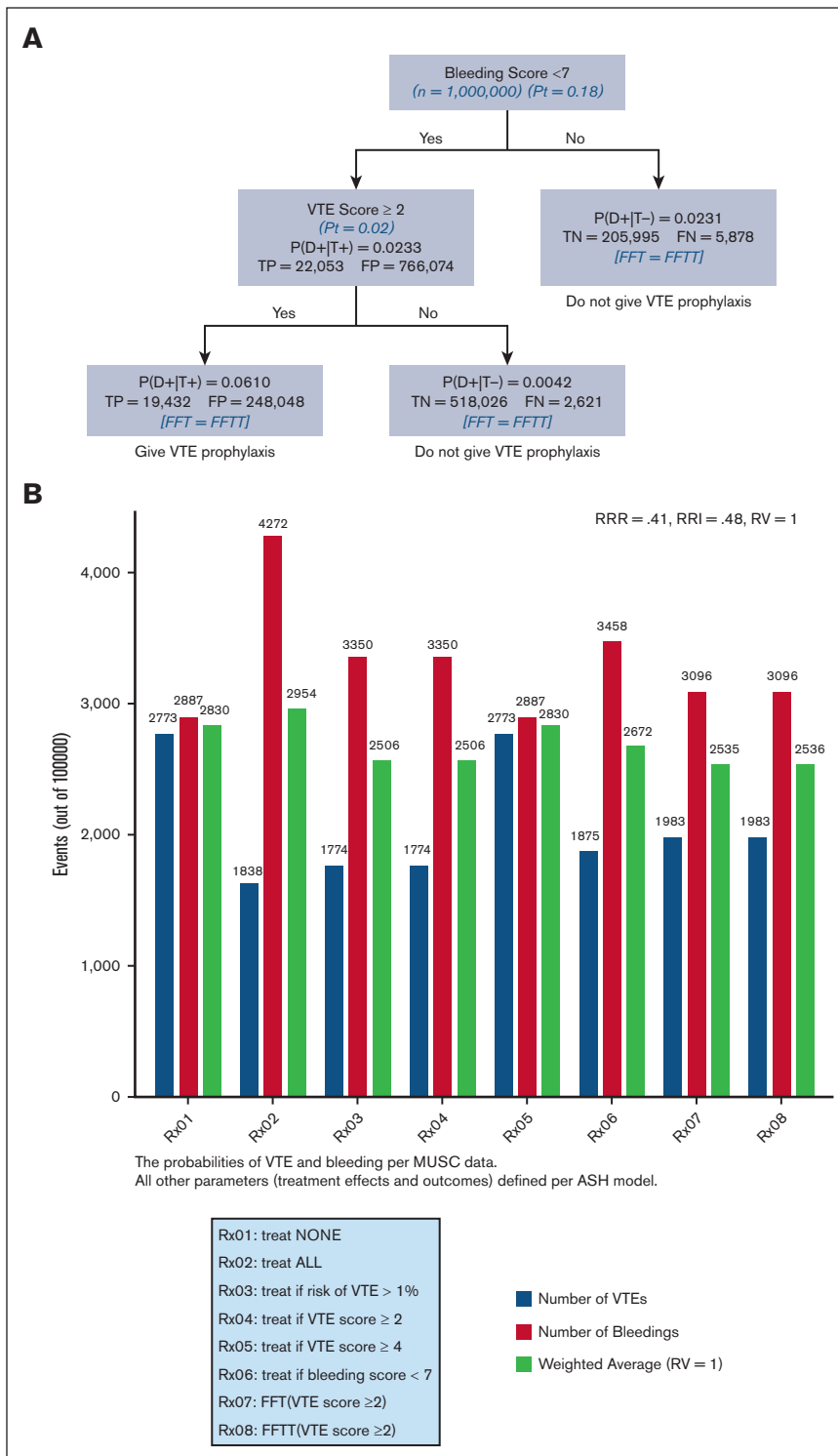
**Figure 5.** An FFT and FFTT decision trees were informed by IMPROVE risk assessment scores  $\geq 2$  (A) and  $\geq 4$  (B) and evidence-based ASH guidelines; (C) impact on bleeding and VTE outcomes. The strategy with the lowest weighted average is considered the best. P(D+|T+)— probability of VTE if the test is “positive” that is, IMPROVE score exceeds given score (= PPV); P(D+|T-)— probability of VTE if the test is “negative” that is, IMPROVE Score <given score; TP, true positives; FP, false positive; TN, true negatives; FN, false negatives.

reported in the original IMPROVE VTE article (Table 3). Thus, IMPROVE predictors appear to have been reproducibly validated for their prognostic value across various settings. In addition, the important evidence regarding the validity of the IMPROVE model came from a recent cluster randomized trial that demonstrated that when a somewhat modified IMPROVE -D-Dimer model (IMPROVE-DD) was embedded in an electronic health record (EHR) clinical decision support system, appropriate thromboprophylaxis significantly increased, yielding a reduction in VTE events with no increase in bleeding rates.<sup>48</sup>

Based on our multidimensional modeling using the “best evidence approach,” “learning health system paradigm,” and “real-world data analysis,” we endorse IMPROVE RAMs as the optimal approach to VTE hospital prophylaxis. The first 2 approaches (Figures 5 and 6) unequivocally demonstrated the superiority of the FFT(T) approach, as shown in Figure 1, for hospital VTE prophylaxis. However, analysis using only MUSC data (supplemental Appendix 7) suggests that VTE prophylaxis should be offered to all patients, regardless of their VTE risk score. This recommendation appears to contradict the impact analysis, which indicates

**Figure 6. FFT and FFFT decision trees analysis according to the “learning health system” paradigm, application of the best research evidence to local data.**

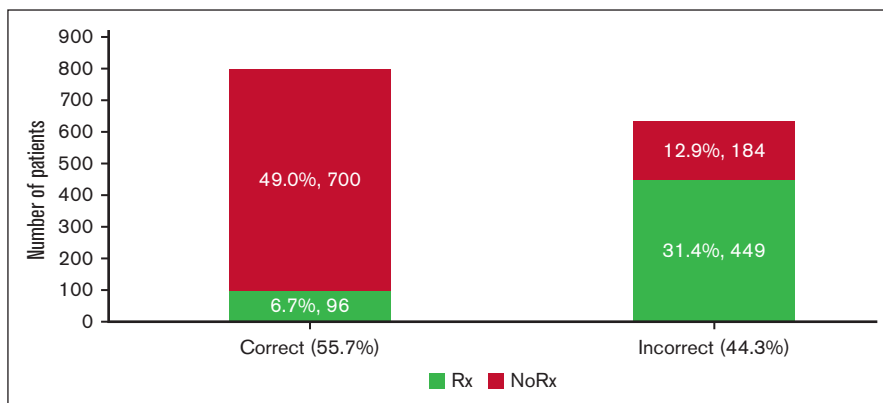
Here, it consisted of using MUSC bleeding and VTE risk scores  $\geq 2$  while using ASH guidelines as a source of the most reliable (“best”) evidence (A). The strategy with the lowest weighted average is considered the best (B). P(D+|T+)—probability of VTE if the test is “positive” that is, IMPROVE score exceeds given score (= PPV); P(D+|T-)—probability of VTE if the test is “negative” that is, IMPROVE score  $<$  given score.



that the optimal strategy might be to “administer prophylaxis to none,” thus questioning the necessity of VTE prophylaxis. We suspect that the discrepancy between the FFT and FFFT classification results stems from the small sample size, yielding unrealistic estimates of the zero-threshold probability at the second FFT

cue. To observe a 40% reduction in VTE from our baseline of 2.2% (as detailed in Tables 1 and 2), following the ASH guidelines,<sup>6</sup> we estimated a requirement to enroll at least 3104 patients. Consequently, this led to pharmacologically implausible results in our calculated treatment effects.





**Figure 7. Estimates of the proportion of correctly and incorrectly administered VTE prophylaxis.** Of the 44% of inappropriate VTE prophylaxis, 13% and 31% consisted of underuse and overuse, respectively (see also supplemental Appendix 7, related to estimates of costs of the inappropriate VTE prophylaxis).

Another limitation is that our study is based on retrospective and modeling analyses. Despite our independent verifications of ICD10 codes, a prospective study implementing the proposed decision support in EHR at the point of care is needed to assess the full impact of the proposed management for VTE prophylaxis.

Our analysis, which has relied on a strong theoretical framework, demonstrates that the FFT(T)-based decision tree (Figure 1) informed by IMPROVE VTE score  $\geq 2$  or  $\geq 4$  and bleeding score  $\leq 7$  offers the best approach to hospital prophylaxis. It is important to place this conclusion in the correct perspective: a claim here is not that IMPROVE models are perfect (none of the RAMs are) but that relying on the RAMs provides the opportunity for much better management than other default strategies (such as to offer VTE prophylaxis to all medically hospitalized patients without contraindication vs offering it to none or even relying on health providers' gestalt for VTE prophylaxis).

In conclusion, our analysis, grounded in a robust theoretical framework, revealed that the decision tree based on FFT(T) (Figure 1), guided by an IMPROVE VTE score of  $\geq 2$  or  $\geq 4$  and a bleeding score of  $\leq 7$ , represents the optimal approach to hospital prophylaxis. Following Spyropoulos et al,<sup>48</sup> we call for embedding the IMPROVE-based FFT(T) into EHR at the point of care to realize the full impact of VTE prophylaxis. Ideally, further implementation of FFT(T) decision support should be accompanied by prospective studies, including multi-institutional

cost-effectiveness pragmatic randomized trials, to best ascertain the broader socioeconomic values of the proposed approach to VTE hospital prophylaxis.

## Acknowledgment

This project is conducted within the purview of the Hematology Stewardship Program at the Medical University of South Carolina, Charleston, SC.

## Authorship

Contribution: B.D. conceived the project, performed analysis, and wrote the first draft; S.A. and F.K. verified the data; N.M. and S.M. identified eligible patients using ICD10 codes; A.B., S.K., and D.B.S. provided organizational and administrative support; K.Y. provided content-related input; I.H. created statistical codes for the analysis; and all authors approved the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profile: I.H., [0000-0003-2349-5707](https://orcid.org/0000-0003-2349-5707).

Correspondence: Benjamin Djulbegovic, Division of Hematology & Oncology, Department of Medicine, Medical University of South Carolina, 39 Sabin St, MSC 635, Charleston, SC 29425; email: [djulbegov@musc.edu](mailto:djulbegov@musc.edu).

## References

1. *The Surgeon General's Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism (Office of the Surgeon General (US)).* National Heart, Lung, and Blood Institute (US); 2008.
2. Gaylis FD, Van SJ, Daneshvar MA, et al. Preprinted standardized orders promote venous thromboembolism prophylaxis compared with traditional handwritten orders: an endorsement of standardized evidence-based practice. *Am J Med Qual.* 2010;25(6):449-456.
3. *Preventing Hospital-Associated Venous Thromboembolism. Appendix A: Tools and Resources.* Agency for Healthcare Research and Quality; 2018.
4. Djulbegovic M, Chen K, Sureshanand S, Chaudhry S. Overuse of primary thromboprophylaxis in medical inpatients at low risk of venous thromboembolism. *J Gen Intern Med.* 2021;36(9):2883-2885.
5. Grant PJ, Conlon A, Chopra V, Flanders SA. Use of venous thromboembolism prophylaxis in hospitalized patients. *JAMA Intern Med.* 2018;178(8):1122-1124.
6. Schunemann HJ, Cushman M, Burnett AE, et al. American Society of Hematology 2018 guidelines for management of venous thromboembolism: prophylaxis for hospitalized and nonhospitalized medical patients. *Blood Adv.* 2018;2(22):3198-3225.

7. Rezende SM, Bauer KA, Zakai NA. Thromboprophylaxis in hospitalized and nonhospitalized medical patients: what's new? *Blood Adv.* 2023;7(18): 5199-5201.
8. Spyropoulos AC, Anderson FA Jr, FitzGerald G, et al. Predictive and associative models to identify hospitalized medical patients at risk for VTE. *Chest.* 2011;140(3):706-714.
9. Decousus H, Tapson VF, Bergmann JF, et al. Factors at admission associated with bleeding risk in medical patients: findings from the IMPROVE investigators. *Chest.* 2011;139(1):69-79.
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. Rosenberg DJ, Press A, Fishbein J, et al. External validation of the IMPROVE bleeding risk assessment model in medical patients. *Thromb Haemost.* 2016;116(09):530-536.
12. Hostler DC, Marx ES, Moores LK, et al. Validation of the international medical prevention registry on venous thromboembolism bleeding risk score. *Chest.* 2016;149(2):372-379.
13. Mahan CE, Liu Y, Turpie AG, et al. External validation of a risk assessment model for venous thromboembolism in the hospitalised acutely-ill medical patient (VTE-VALOURR). *Thromb Haemost.* 2014;112(4):692-699.
14. Zhang Z, Zhai Z, Li W, et al. Validation of the IMPROVE bleeding risk score in Chinese medical patients during hospitalization: findings from the dissolve-2 study. *Lancet Reg Health West Pac.* 2020;4:100054.
15. Villiger R, Juillard P, Darbellay Farhoumand P, et al. Prediction of in-hospital bleeding in acutely ill medical patients: external validation of the IMPROVE bleeding risk score. *Thromb Res.* 2023;230:37-44.
16. Djulbegovic B, Hozo I. *Threshold decision-making in clinical medicine with practical application to hematology and oncology.* Springer Nature Switzerland AG; 2023.
17. Djulbegovic B, Hozo I, Cuker A, Guyatt G. Improving methods of clinical practice guidelines: from guidelines to pathways to fast-and-frugal trees and decision analysis to develop individualised patient care. *J Eval Clin Pract.* 2023.
18. Djulbegovic B, Hozo I, Lizarraga D, et al. Evaluation of a fast-and-frugal clinical decision algorithm ('pathways') on clinical outcomes in hospitalised patients with COVID-19 treated with anticoagulants. *J Eval Clin Pract.* 2023;29(1):3-12.
19. Djulbegovic B, Hozo I. Formulating management strategies using fast-and-frugal trees (a decision tool to transform clinical practice guidelines and clinical pathways into decision support at the point of care). *Cancer Treat Res.* 2023;189:67-75.
20. Djulbegovic B, Hozo I, Dale W. Transforming clinical practice guidelines and clinical pathways into fast-and-frugal decision trees to improve clinical care strategies. *J Eval Clin Pract.* 2018;24(5):1247-1254.
21. Hozo I, Djulbegovic B, Luan S, Tsalatsanis A, Gigerenzer G. Towards theory integration: threshold model as a link between signal detection theory, fast-and-frugal trees and evidence accumulation theory. *J Eval Clin Pract.* 2017;23(1):49-65.
22. Spyropoulos AC, Ageno W, Albers GW, et al. Rivaroxaban for thromboprophylaxis after hospitalization for medical illness. *N Engl J Med.* 2018;379(12): 1118-1127.
23. Luan S, Schooler LJ, Gigerenzer G. A signal-detection analysis of fast-and-frugal trees. *Psychol Rev.* 2011;118(2):316-338.
24. Djulbegovic B, Hozo I, Lizarraga D, Guyatt G. Decomposing clinical practice guidelines panels' deliberation into decision theoretical constructs. *J Eval Clin Pract.* 2023;29(3):459-471.
25. Djulbegovic B, Hozo I, Lyman GH. Linking evidence-based medicine therapeutic summary measures to clinical decision analysis. *MedGenMed.* 2000; 2(1):E6.
26. Djulbegovic B, Hozo I, Mayrhofer T, van den Ende J, Guyatt G. The threshold model revisited. *J Eval Clin Pract.* 2019;25(2):186-195.
27. Klein D. Implementing a general framework for assessing interrater agreement in Stata. *STATA J.* 2018;18(4):871-901.
28. Cuker A, Tseng EK, Nieuwlaar R, et al. American society of hematology living guidelines on the use of anticoagulation for thromboprophylaxis in patients with COVID-19: January 2022 update on the use of therapeutic-intensity anticoagulation in acutely ill patients. *Blood Adv.* 2022;6(17):4915-4923.
29. The ICM-VTE Hip & Knee Delegates. Recommendations from the ICM-VTE: Hip & Knee. *JBJS.* 2022;104(Suppl 1):180-231.
30. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating.* Springer; 2009.
31. Hozo I, Djulbegovic B. Generalized decision curve analysis for explicit comparison of treatment effects. *J Eval Clin Pract.* 2023;29(8):1271-1278.
32. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet.* 2017;390(10092):415-423.
33. Middeldorp S, Nieuwlaar R, Baumann Kreuziger L, et al. American society of hematology 2023 guidelines for management of venous thromboembolism: thrombophilia testing. *Blood Advances.* 2023;7(22):7101-7138.
34. O'Brien JA, Caro JJ. Direct medical cost of managing deep vein thrombosis according to the occurrence of complications. *Pharmacoeconomics.* 2002; 20(9):603-615.
35. Lin J, Lingohr-Smith M, Kwong WJ. Incremental health care resource utilization and economic burden of venous thromboembolism recurrence from a U.S. payer perspective. *J Manag Care Pharm.* 2014;20(2):174-186.
36. Pauker SG. The clinical interpretation of research. *PLoS Med.* 2005;2(11):e395.

37. Pauker SG, Kassirer J. The threshold approach to clinical decision making. *N Engl J Med*. 1980;302:1109-1117.
38. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ*. 1999;18:341-364.
39. Djulbegovic B. A framework to bridge the gaps between evidence-based medicine, health outcomes, and improvement and implementation science. *J Oncol Pract*. 2014;10(3):200-202.
40. Djulbegovic B, Paul A. From efficacy to effectiveness in the face of uncertainty indication creep and prevention creep. *JAMA*. 2011;305(19):2005-2006.
41. Agency for Healthcare Research and Quality. About learning health systems. 2019. Accessed 4 May 2019. <https://www.ahrq.gov/learning-health-systems/about.html>
42. STATA Version 18. 2022.
43. Vandembroucke JP, Elm Ev, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med*. 2007;147(8):W-163-W-194.
44. Debray TPA, Collins GS, Riley RD, et al. Transparent reporting of multivariable prediction models developed or validated using clustered data: TRIPOD-Cluster checklist. *BMJ*. 2023;380:e071018.
45. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230.
46. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. *J Clin Epidemiol*. 2016;74:167-176.
47. Djulbegovic B, Hozo I. *Threshold Decision-making in Clinical Medicine. With Practical Application to Hematology and Oncology*. Springer Nature; 2023.
48. Spyropoulos AC, Goldin M, Koulas I, et al. Universal EHRs clinical decision support for thromboprophylaxis in medical inpatients: a cluster randomized trial. *JACC Adv*. 2023;2(8):100597.