blood advances

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodadvances@hematology.org

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

Tracking no: ADV-2024-013166R1

Benjamin Djulbegovic (Medical University of South Carolina, United States) Alice Boylan (Medical University of South Carolina, United States) Shelby Kolo (Medical University of South Carolina, United States) Danielle Scheurer (Medical University of South Carolina, United States) Sydney Anuskiewicz (Medical University of South Carolina, United States) Flora Khaledi (Medical University of South Carolina, United States) Kimberley Youkhana (Medical University of South Carolina, United States) Sarah Madgwick (Medical University of South Carolina, United States) Neema Maharjan (Medical University of South Carolina, United States) Iztok Hozo (Indiana University Northwest, United States)

Abstract:
Current hospital VTE (venous thromboembolism) prophylaxis for medical patients has been characterized by both underuse and, increasingly, overuse. The American Society of Hematology (ASH) has endorsed the use of risk assessment models (RAM) as an approach to individualize VTE prophylaxis as a way of 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 RAM 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 employed three 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 the most effective VTE prevention strategy utilizes the FFT decision tree, guided by an IMPROVE VTE score of {greater than or equal to} 2 or {greater than or equal to} 4 and a bleeding score of < 7. This method could prevent 45% of unnecessary treatments, saving about $5 million annually for patients like the MUSC cohort. We recommend integrating the IMPROVE models into hospital electronic medical records as a point-of-care tool, thereby enhancing VTE prevention in hospitalized medical patients.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: BD conceived the project, performed analysis and wrote the first draft; SD and FK verified data; NM and SH identified eligible patients using ICD10 codes; AB, SK and DBS provided organizational and administrative support; KY provided content-related input; IH created statistical codes for the analysis. All authors approved the manuscript.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: by writing request to the first author (because data include the patients from MUSC hospitals MUSC legal administrative approval is also required)

Clinical trial registration information (if any):

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

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

[1]Medical University of South Carolina, Charleston, SC, USA

[2]Department of Mathematics, Indiana University Northwest, Gary, Indiana, USA

Word count:

Abstract: 247

Manuscript: 3,993

Correspondence:

Benjamin Djulbegovic, MD, PhD
Professor of Medicine
Medical University of South Carolina
Department of Medicine
Division of Hematology & Oncology
39 Sabin Street
MSC 635
Charleston, SC 29425

e-mail: djulbegov@musc.edu
Phone: 843-792-4879

Data Sharing:

By writing request to the first author (because data include the patients from MUSC hospitals MUSC legal administrative approval is also required)

Key points:

- American Society of Hematology has endorsed using risk assessment models (RAM) to individualize VTE prophylaxis.

- Fast-and-frugal decision tree based on IMPROVE RAMs reduces unnecessary VTE prophylaxis by 45% and saves approximately $5 million annually

Abstract

Current hospital VTE (venous thromboembolism) prophylaxis for medical patients has been characterized by both underuse and, increasingly, overuse. The American Society of Hematology (ASH) has endorsed the use of risk assessment models (RAM) as an approach to individualize VTE prophylaxis as a way of 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 RAM 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 employed three 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 the most effective VTE prevention strategy utilizes the FFT decision tree, guided by an IMPROVE VTE score of ≥ 2 or ≥ 4 and a bleeding score of < 7. This method could prevent 45% of unnecessary treatments, saving about $5 million annually for patients like the MUSC cohort. We recommend integrating the IMPROVE models into hospital electronic medical records as a point-of-care tool, thereby enhancing VTE prevention in hospitalized medical patients.

Keywords: clinical practice guidelines, evidence-based medicine, decision analysis, risk assessment models, fast-and-frugal trees, medical decision-making

Fifteen years ago, the U.S. Surgeon General issued a Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism (VTE) as one the most serious but avoidable hospital complications. [1] Ever since, 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". [2] Reducing the underuse of VTE prophylaxis has also been promoted as one of the critical quality measures by most National Quality Organizations in the U.S. [3] Increasingly, however, overuse is being recognized as a more significant problem. The 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. [4]As a result, there is increasing recognition that some patients at higher risk of VTE need hospital VTE prophylaxis while others, considered to be at low risk of VTE, do not. [5] Most recently, influential guidelines panels such as the American Society of Hematology (ASH) have endorsed the use of risk assessment models (RAM) as an approach to individualizing VTE prophylaxis as a way of finding a balance between overuse (excessive risk of bleeding) vs. underuse (risk of avoidable VTE).[6,7]

ASH has also endorsed IMPROVE (International Medical Prevention Registry on Venous Thromboembolism) VTE risk assessment models (RAMs) to guide VTE prophylaxis.[8] IMPROVE is the only existing RAM for assessing both short-term bleeding (IMPROVE_bleeding) [9] and VTE risk in acutely ill medical inpatients (IMPROVE_vte).[8] It has been externally validated in US[10-12], Canadian[13], Chinese [14], and European population[15] with acceptable discrimination [AUC=0.65-0.77 (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 two predictive models (i.e., IMPROVE_bleeding and IMPROVE_vte) in one coherent management strategy. We have previously demonstrated that this can only be accomplished within a

decision-theoretical framework. [16-18] One such a framework consists of *FFT (fast-and-frugal) decision trees*-an 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. [17-21] Unlike popular clinical pathways (algorithms) that are theory-free constructs, FFTs are based on solid theoretical foundations.[20,21] The latter provides the basis for quantitative evaluation of the FFTs' classification capacity. Importantly, clinical algorithms can easily be converted into FFTs to facilitate the quantitative assessment of their classification accuracy. [17-21]

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

Methods

Practice guidelines for using IMPROVE RAMs

The original IMPROVE VTE risk paper [8] and ASH [6,7] guidelines recommend VTE prophylaxis if the IMPROVE VTE risk score ≥ 2. This converts into the probability of VTE between 1% and 1.5% in the original report. [8] Some authors recommend an IMPROVE score ≥ 4 to indicate thromboprophylaxis in hospital patients.[22] The IMPROVE bleeding risk report [9] and ASH[6,7] guidelines recommend IMPROVE bleeding score of <7 and ≥ 7 as a low or high bleeding risk to proceed or avoid VTE prophylaxis, respectively. According to the original publicaton [9] , IMPROVE bleeding score <7 corresponds to 1.5% of major and clinically relevant bleeding. IMPROVE bleeding score ≥ 7 was observed in 7.9% of 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. Fig 1 shows the FFT based on the integration of VTE and bleeding IMPROVE scores in a comprehensive risk-adapted VTE prophylaxis. As shown, we started with the "first, do no harm" decision principle: the patients with high bleeding score (score ≥ 7) should not be given prophylaxis (first cue in Fig 1). Those with bleeding scores <7 are then assessed for risk for VTE. Then, if the IMPROVE VTE risk score is≥ 2 or ≥ 4 (see results), they will be given pharmacological VTE prophylaxis.

By categorizing disease or making predictions, FFT also functions as a predictive model. [21] This enables the 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.[21,23,24]

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

Using real-world data to collect data on IMPROVE RAMs

The patients were eligible for the analysis if they were aged≥ 18 years and had≥ 3 days of hospitalization for an acute medical illness *consecutively* admitted to the MUSC hospital in Charleston, SC. Patients were excluded if they received anticoagulants or thrombolytics at admission or within 48 hours after admission, if they were bleeding at admission, had major surgery or trauma within 3 months before

admission, 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 key features of IMPROVE VTE and bleeding RAMs from the MUSC electronic medical (EPIC)/discharge records. Appendix 0 shows ICD10 and other codes we used to select the variables of interest for the analysis. Outcomes of interest were VTE at 90 and bleeding (major and clinically significant bleeding) at 14 days. Two reviewers manually verified all outcomes (VTE and bleeding). The ascertainment of exposure (whether the patients received VTE prophylaxis) was conducted on 20% of randomly selected records. Gwet's agreement coefficient[27] among two observers was high- the accuracy for classification of VTE outcome was 0.95 [95%CI: 0.94 to 0.96] for VTE and 0.93 [95%CI: 0.92 to 0.94] for bleeding outcomes, respectively. The overall accuracy of ascertainment of the exposure was 96%.

Enoxaparin at the recommended prophylactic dose (at 40 mg QD or 30 mg bid, further adjusted for renal function as needed) [28] was almost exclusively used in our patients' cohort. The patients who have received therapeutic/intermediate intensity anticoagulation[28] (see Appendix2, Table 1C) were excluded (n=601). The patients who received aspirin (n=132) were included in the prophylactic arm in light of evidence that aspirin has prophylactic effects in a surgical setting.[29] The patients who received anti-platelet combinations without aspirin, heparin flushes to keep intravenous lines open, and thrombolytics (alteplase) for treatment of thrombotically occluded catheters (n=42) were included in the control arm (Appendix2 Table 1B).

External validation of IMPROVE models

To assess the generalizability of original IMPROVE models[8 9], we assessed their performance using MUSC data.. Both original and external validation of IMPROVE models [10-12 13 14 15] have been performed agnostic of prophylactic treatments. Models' performance was assessed by calculating their discrimination and calibration properties.[30] Discrimination (capacity of the predictive model to discriminate between those patients who truly have outcomes vs. those who don't have an outcome of interest) was assessed using c-statistics. Calibration (capacity to determine the agreement between predicted and actually observed outcomes) was assessed by the Hosmer-Lemeshow Goodnes-of-Fit statistic (H-L), and examining intercept, i.e., whether the best-fit line of the model crossed the y-axis at zero predicted probability and the whether the slope was statistically significantly different from 1. In the case of the models' poor external validation, 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.[6,7] 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 relative risk and RRR to relative risk reduction of VTE due to thromboprophylaxis compared to no treatment. The ASH estimated relative risk increase (RRI) of clinically significant bleeding to be 48% higher on thromboprophylaxis than in no treatment arms.[6]

We compared the effects of the following management strategies on the number of bleeding (major and clinically significant) and VTE events (see 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[29]); 4) administer VTE

prophylaxis if VTE score is ≥2 (as per ASH guidelines[6] and the original IMPROVE paper [8]); 5) administer VTE prophylaxis if VTE score is ≥4 (as per ref# [22]]); 6) apply VTE prophylaxis if the bleeding score is < 7 (as per IMPROVE bleeding score recommendation [9]); 7-10) administer VTE prophylaxis as per FFT/T (Fig 1) by assuming VTE risk ≥2 or ≥4, respectively.

The heart of personalized decision-making consists of contrasting the risk assessment against decision thresholds[17,31], 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 [16,25,26] can be expressed as:

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

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 FTT/T decision tree shown in Fig 1; RV $(relative\ value)$ refers to patients' values and preferences (V&P). When RV <1, the patient values avoiding 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, RV is set to 1. When RV=1, the threshold is solely determined by empirical evidence.[26,32] Our default impact metrics comprised the weighted average between VTE and bleeding events calculated as Nboth = (Nvte + RV*Nbld)/(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' R.V.[33] We also used costs as the impact metrics by projecting costs of unnecessary VTE prophylaxis from published reports[34,35] to January 2024 dollar costs (see Appendix 8).

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

Because the patients studied in clinical trials differ from "real-world" patients [39,40], a fundamental concept behind the "*learning health system paradigm*" [41] 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. [6,7]

Finally, we performed the same analysis using exclusively MUSC ("*real-world" data*). From the data, we estimated RRR_{vte}= 0.058 and $RRI_{bleeding}$=0.197; however, the results for RRR were not significant, potentially indicating the null effect of prophylaxis on VTE (see Discussion). All statistical analyses were performed using the statistical package Stata[42] and reported according to STROBE (Strengthening the Reporting of Observational Studies in Epidemiology)[43] and TRIPOD[44] guidelines.

Results

Fig 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 December 31, 2022, 5,051 consecutive patients were admitted to the MUSC hospital in Charleston, SC of whom 2,072 patients were medical patients eligible for the analysis as per the criteria in the original IMPROVE reports.[8 9] After further restriction to the observations within 14 days (bleeding) and 90 days (VTE), and exclusion of the patients who have received therapeutic/intermediate intensity anticoagulation[28] , we ended up with 1,429 analyzable patients for the impact analysis. The median length of stay was 5 days

(range: 3 to 85). Note that we could not collect data on the "immobility >7 days" due to a lack of adequate ICD10 codes.

Table 1 shows a distribution of the IMPROVE VTE and bleeding risk predictors in our population as per intention-to-treat population. Not surprisingly, the case-mix differed between our and original IMPROVE publications[8 9] , primarily due to a higher number of patients requiring admission to ICU during the index hospitalization in our vs. the original IMPROVE cohorts (30.2% vs 4.89[8] and 8.5% [9]). However, the risk score proportion distribution was similar despite dissimilarities in the predictors distribution (Table 2).

Fig 3 shows 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 show a statistically significant absolute risk difference of 2.9% (95%CI: 0.5% to 5.3%, p=0.001), while no statistical difference was noted for VTE scores (see Discussion). App3 shows the analysis by VTE prophylaxis, which is consistent with what is observed in the original publications[8 9] and other validation studies[10-12 13 14 15] - treatment did not affect the IMPROVE prognostic scores.

App 4 Fig 1.1 shows external validation of IMPROVE bleeding scores with discrimination of AUC=0.799 [95%CI: 0.731 to 0.848] , similar to reported in the original IMPROVE report[9] and other external validation studies. [13 14 15] After updating intercept[45,46], the model was well calibrated with not significant H-L test (p=0.287), with an intercept of close to 0 and the slope of 1.016, 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 the original IMPROVE score[8] and excellent calibration [H-L test:p=0.927, with an intercept of 0 and slope=1], (see Appendix 4, Fig1.2). Importantly, however, both VTE scores and particularly bleeding scores were almost perfectly

discriminatory regarding the predicted risk above and below scores employed in FFT (Fig 4; and App3 Fig 1.3).

Figs 5a and b) show FFT and FFTT decision trees informed by IMPROVE risk assessment scores ≥2 (5a) and ≥4 (5b) and evidence-based ASH guidelines. [6,7] The results were consistent with our hypothesis that the FFT(T)-based decision tree generated by linking VTE and bleeding IMPROVE scores (as shown in Fig 1) leads to optimal classification and decision-making. We also determined that the tree was fast, reaching the decision in less than two steps, and frugal, using only about 5% of all information. In this case, both FFT and FFTT had identical classifications at the overall accuracy of 93% for VTE score ≥ 4 (Fig 5b) and 72% for VTE score ≥ 2 (Fig 5a), respectively. However, given the low prevalence of outcomes (Fig 3), as expected, the negative predictive value (NPV) was much higher (>99% for both VTE scores) than the positive predictive value (PPV) of 5.6% for FFT(T) employing VTE score ≥ 4 and 2.44% for VTE score ≥ 2, respectively. Fig 5c shows an impact analysis comparing 10 different management strategies. Two FFT(T) driven strategies (Rx07 to 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 Rx1 strategy ("Treat none")- the second worst management option. The changes in RV did not materially affect the results (App5).

Fig 6 shows the analysis according to the "learning health system" [41] paradigm- application of the best research evidence to local data. In our case, this consisted of using MUSC bleeding and VTE risk scores ≥2 while using ASH guidelines as a source of the most reliable ("best") evidence. [6,7] The results agree with the analysis based on the "best evidence" shown in Fig 5 with slightly higher overall classification

accuracy 74% and PPV=6%, while retaining higher than 99% NPV. App 6 shows the results of sensitivity analysis by RV.

App Fig 7 shows FFT and FFTT decision tree using real-world MUSC data for VTE, bleeding risks, and treatment data. As shown in Fig3 and App3, in our data set, we did detect a significant effect of VTE prophylaxis on bleeding but not on the VTE outcomes. This resulted in the unrealistic estimates of RRR=5.8%, and RRI=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 to patients at high risk of bleeding- represents the best management option even if assuming such low effectiveness of treatment. (see App7)

Fig 1 shows that we can define unnecessary VTE prophylaxis as: (*overuse)*= VTE prophylaxis is given for bleeding score ≥7 or VTE score <2; failure to administer VTE prophylaxis (*underuse*) = VTE prophylaxis is *not* given if VTE score ≥2 and bleeding score <7. By applying this FFT(T) within the "learning health system" paradigm, we can calculate that about 44% of inappropriate VTE prophylaxis (of which 13% and 31% consist of underuse and overuse, respectively) could be avoided (Fig 7) at annual savings due to avoiding resource waste at about $5millions per our patient cohort (AppFig 8)

Discussion

Current hospital VTE prophylaxis for medical patients is suboptimal. It has historically been characterized by underuse but increasingly by overuse- over 85% of patients receiving inappropriate VTE prophylaxis in some studies.[4] As a result, calls have been made to "deimplement" VTE prophylaxis at low-risk for VTE and high-bleeding risk patients.[5] To minimize both underuse and overuse, the ASH guidelines on VTE prophylaxis state that when "clinicians and healthcare systems use ASH VTE guidelines, they should

integrate VTE and bleeding risk assessments into clinical decision-making processes". [6] However, the ASH guidelines also recommended that RAMs-based decision-making needs further clinical outcomes impact evaluation. [6] In this paper, we show 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 risks models[8 9] into FFT(T) decision trees to enable the quantification and impact assessment of competing clinical VTE prophylaxis strategies. [17-21,47] FFT draws its theoretical robustness by relating to signal detection theory, evidence accumulation theory, and the threshold model to help improve decision-making. [17-21,47] Importantly, we observed a meaningful impact within a "learning health systems" [41] - framework aiming to bridge the gaps between efficacy data (best evidence observed in ideal research settings) with the real-world case-mix data that often do not reflect controlled research settings.[39,40] In doing so, we demonstrated that about 45% of inappropriate VTE prophylaxis- amounting to around $ 5 million in resource waste annually- could be avoided. (Fig 7, App 8)

Our study adds to a series of other studies that had externally validated IMPROVE bleeding scores in the various hospital settings[13 14 15] testifying to the utility of the IMPROVE model. Unfortunately, unlike others[10,12 13] we could not externally validate the IMPROVE VTE risk model. This is likely because our sample size was small (n=1849, Fig 3b) as opposed to the 15,156 patients enrolled in the original study. [8] Another reason, as explained earlier, is that our population differed from the original IMPROVE cohort in the number of transfers to the ICU during the index hospitalization (30.2% vs 4.89%).[13] Finally, we did not collect data on one of the IMPROVE predictors (immobility >7 days) (see Table 1). Despite this, we could internally validate the IMPROVE VTE score and obtain almost identical risk scores as reported in the original IMPROVE VTE paper (Table 2). Thus, the IMPROVE predictors appear to have been reproducibly validated for their prognostic values across the various settings. The most important evidence about the validity of the IMPROVE model came from the recent cluster randomized trial that demonstrated that

when a somewhat modified IMPROVE model (IMPROVE-DD) was embedded in an electronic health record (EHR) clinical decision support system, the appropriate thromboprophylaxis significantly increased yielding reduction in VTE events with no increase in bleeding rates.[48]

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 prophylaxis. The first two approaches (Figs 5 and 6) unequivocally demonstrated the superiority of FFT(T) approach shown in Fig 1 to hospital VTE prophylaxis. However, analysis using only MUSC data (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 that the optimal strategy might be to 'administer prophylaxis to none', thus questioning the necessity of VTE prophylaxis. We suspect the discrepancy between FFT and FFTT classification results stems from the small sample size, yielding unrealistic estimates of zero threshold probability at the second FFT cue. To observe a 40% reduction in VTE from our baseline of 2.2% (as detailed in Table 1), following the ASH guidelines report[6], we estimate a requirement to enroll at least 3,104 patients. Consequently, this has led to pharmacologically implausible results in our calculated treatment effects.

Another limitation is that our study is based on retrospective and modeling analyses. Despite our independent verifications of ICD10 codes, a prospective study by 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.

In conclusion, our analysis that has relied on a solid theoretical framework demonstrates that the FFT(T) based decision tree (Fig 1) informed by IMPROVE VTE score ≥ 2 or ≥ 4 and bleeding score ≤ 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. offer it to none or even relying on health providers' gestalt for VTE prophylaxis).

In conclusion, our analysis, grounded in a robust theoretical framework, reveals that the decision tree based on FFT(T) (see Fig 1), guided by an IMPROVE VTE score of ≥ 2 or ≥ 4 and a bleeding score of ≤ 7, represents the optimal approach to hospital prophylaxis. Following Spyropoulos et al[48], 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 socio-economic values of the proposed approach to VTE hospital prophylaxis.

Contributions:

BD conceived the project, performed analysis and wrote the first draft; SD and FK verified data; NM and SH identified eligible patients using ICD10 codes; AB, SK and DBS provided organizational and administrative support; KY provided content-related input; IH created statistical codes for the analysis. All authors approved the manuscript.

None of the authors has a relevant conflict of interest.

Legends

Figure 1. A fast-and-frugal (FFT) tree was constructed using IMPROVE VTE and bleeding risk assessment models. (see manuscript for further information); VTE- venous thromboembolism

Figure 2. STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) flow diagram displaying the study patients' enrolment, exclusion criteria, and data availability for the analysis according to anticoagulation treatment. (see manuscript for further details)

Figure 3 Major and clinically significant bleeding and VTE rates as a function of the recommended IMPROVE models' discriminatory scores: <7 vs.≥ 7 for bleeding and <2 vs.≥ 2 VTE scores (see Fig 1); VTE- venous thromboembolism

Figure 4. A relationship between IMPROVE VTE scores (a) and bleeding scores (b) with the probability of VTE and bleeding (major and clinically relevant), respectively. VTE- venous thromboembolism

Figs 5. FFT (Fast-and-frugal) and FFTT (Fast-and-frugal with threshold) decision trees informed by IMPROVE risk assessment scores ≥2 (a) and ≥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” i.e., IMPROVE Score exceeds given score [=positive predictive value]; P(D+|T-)- probability of VTE if the test is “negative” i.e., IMPROVE Score <given score; TP-true positives;FP-false positive;TN-true negatives;FN-false negatives; VTE- venous thromboembolism

Fig 6. FFT (Fast-and-frugal) and FFTT (Fast-and-frugal with threshold) 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 ≥2 while using ASH guidelines as a source of 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” i.e., IMPROVE Score exceeds given score [=positive predictive value]; P(D+|T-)- probability of VTE if the test is “negative” i.e., IMPROVE Score <given score; TP-true positives; FP-false positive; TN-true negatives; FN-false negatives; VTE- venous thromboembolism

Fig 7. Estimates of the proportion of correctly and incorrectly administered VTE prophylaxis. Note that of 44% of inappropriate VTE prophylaxis, 13% and 31% consist of underuse and overuse, respectively. (see also App 7 related to estimates of costs of the inappropriate VTE prophylaxis.

Acknowledgments

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

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. *American Journal of Medical Quality*. 2010;25(6):449-456. doi:10.1177/1062860610369824
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. *Journal of General Internal Medicine*. 2021/09/01 2021;36(9):2883-2885. doi:10.1007/s11606-020-05993-x
5. Grant PJ, Conlon A, Chopra V, Flanders SA. Use of venous thromboembolism prophylaxis in hospitalized patients. *JAMA Internal Medicine*. 2018;178(8):1122-1124. doi:10.1001/jamainternmed.2018.2022
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*. Nov 27 2018;2(22):3198-3225. doi:10.1182/bloodadvances.2018022954
7. Rezende SM, Bauer KA, Zakai NA. Thromboprophylaxis in hospitalized and nonhospitalized medical patients: what's new? *Blood Advances*. 2023;7(18):5199-5201. doi:10.1182/bloodadvances.2023010067
8. Spyropoulos AC, Anderson FA, Jr., FitzGerald G, et al. Predictive and associative models to identify hospitalized medical patients at risk for VTE. *Chest*. Sep 2011;140(3):706-714. doi:10.1378/chest.10-1944
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*. Jan 2011;139(1):69-79. doi:10.1378/chest.09-3081
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*. Nov 17 2014;3(6):e001152. doi:10.1161/JAHA.114.001152
11. Rosenberg DJ, Press A, Fishbein J, et al. External validation of the IMPROVE Bleeding Risk Assessment Model in medical patients. *Thromb Haemost*. 2017/11/29 2016;116(09):530-536. doi:10.1160/TH16-01-0003
12. Hostler DC, Marx ES, Moores LK, et al. Validation of the International Medical Prevention Registry on Venous Thromboembolism Bleeding Risk Score. *Chest*. Feb 2016;149(2):372-9. doi:10.1378/chest.14-2842
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*. Oct 2014;112(4):692-9. doi:10.1160/TH14-03-0239
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. *The Lancet Regional Health – Western Pacific*. 2020;4doi:10.1016/j.lanwpc.2020.100054

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. *Thrombosis Research*. 2023;230:37-44. doi:10.1016/j.thromres.2023.08.003
16. Djulbegovic B, Hozo I. *Threshold Decision-making in Clinical Medicine with Practical Application to Hematology and Oncology.* Cancer Treatment and Research. Springer Nature Switzerland AG; 2023:XXIII, 138.
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*. Dec 10 2023;doi:10.1111/jep.13953
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. *Journal of Evaluation in Clinical Practice*. 2023;29(1):3-12. doi:https://doi.org/10.1111/jep.13780
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. doi:10.1007/978-3-031-37993-2_5
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 Practice* 2018;24(5):1247-1254. doi:doi:10.1111/jep.12895
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*. Feb 2017;23(1):49-65. doi:10.1111/jep.12490
22. Spyropoulos AC, Ageno W, Albers GW, et al. Rivaroxaban for Thromboprophylaxis after Hospitalization for Medical Illness. *New England Journal of Medicine*. 2018;379(12):1118-1127. doi:10.1056/NEJMoa1805090
23. Luan S, Schooler LJ, Gigerenzer G. A signal-detection analysis of fast-and-frugal trees. *Psychol Rev*. Apr 2011;118(2):316-38. doi:10.1037/a0022684
24. Djulbegovic B, Hozo I, Lizarraga D, Guyatt G. Decomposing clinical practice guidelines panels' deliberation into decision theoretical constructs. *Journal of Evaluation in Clinical Practice*. 2023;29 (3):459-471. doi:https://doi.org/10.1111/jep.13809
25. Djulbegovic B, Hozo I, Lyman GH. Linking evidence-based medicine therapeutic summary measures to clinical decision analysis. *MedGenMed*. Jan 13 2000;2(1):E6.
26. Djulbegovic B, Hozo I, Mayrhofer T, van den Ende J, Guyatt G. The threshold model revisited. *Journal of Evaluation in Clinical Practice*. 2019;25(2):186-195. doi:10.1111/jep.13091
27. Klein D. Implementing a general framework for assessing interrater agreement in Stata. *Stata Journal*. 2018;18(4):871–901.
28. Cuker A, Tseng EK, Nieuwlaat 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 Advances*. 2022;6(17):4915-4923. doi:10.1182/bloodadvances.2022007561
29. Hip TI-V, Delegates* K. Recommendations from the ICM-VTE: Hip & Knee. *JBJS*. 2022;104(Suppl 1):180-231. doi:10.2106/jbjs.21.01529
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 *Journal of Evaluation in Clinical Practice*. 2023;29(8):1271-1278. . doi:doi: 10.1111/jep.13915. Epub 2023 Aug 25.
32. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet*. Jul 22 2017;390(10092):415-423. doi:10.1016/S0140-6736(16)31592-6

33. Middeldorp S, Nieuwlaat 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. doi:10.1182/bloodadvances.2023010177
34. O'Brien JA, Caro JJ. Direct Medical Cost of Managing Deep Vein Thrombosis According to the Occurrence of Complications. *PharmacoEconomics*. 2002/08/01 2002;20(9):603-615. doi:10.2165/00019053-200220090-00004
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*. Feb 2014;20(2):174-86. doi:10.18553/jmcp.2014.20.2.174
36. Pauker SG. The clinical interpretation of research. *PLoS Med*. Nov 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. Claxon K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econom*. 1999;18:341-364.
39. Djulbegovic B. A Framework to Bridge the Gaps Between Evidence-Based Medicine, Health Outcomes, and Improvement and Implementation Science. *Journal of Oncology Practice*. 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. Accessed May 4, 2019, 2019. https://www.ahrq.gov/learning-health-systems/about.html
42. *STATA, ver. 18*. 2022.
43. Vandenbroucke JP, Elm Ev, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *Ann Intern Med*. October 16, 2007 2007;147(8):W-163-194.
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. doi:10.1136/bmj-2022-071018
45. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med*. Dec 16 2019;17(1):230. doi:10.1186/s12916-019-1466-7
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. *Journal of Clinical Epidemiology*. 2016/06/01/ 2016;74:167-176. doi:https://doi.org/10.1016/j.jclinepi.2015.12.005
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: Advances*. 2023/10/01/ 2023;2(8):100597. doi:https://doi.org/10.1016/j.jacadv.2023.100597

Table 1 A) VTE Baseline Characteristics and Outcomes

Predictors N=2,072	%
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

Table1B Bleeding Baseline Characteristics and Outcomes

Predictors N=2,072	%
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 <50.10^2 cells/L	0.9
Bleeding before admission	2.8
Active gastro ulcer	0.0
Bleed_outcome (yes)	1.6

Table 2 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
≥ 4	7	4
Bleeding		
<7	90	79
≥7	10	21

Figure 1

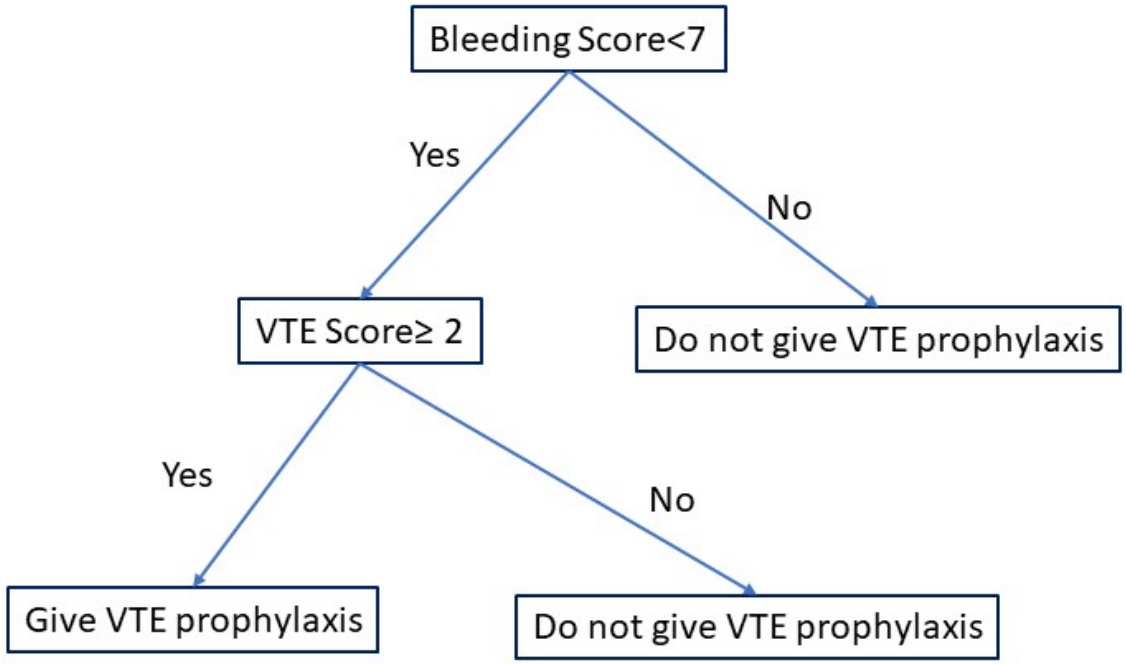


Fig 1

Figure 2

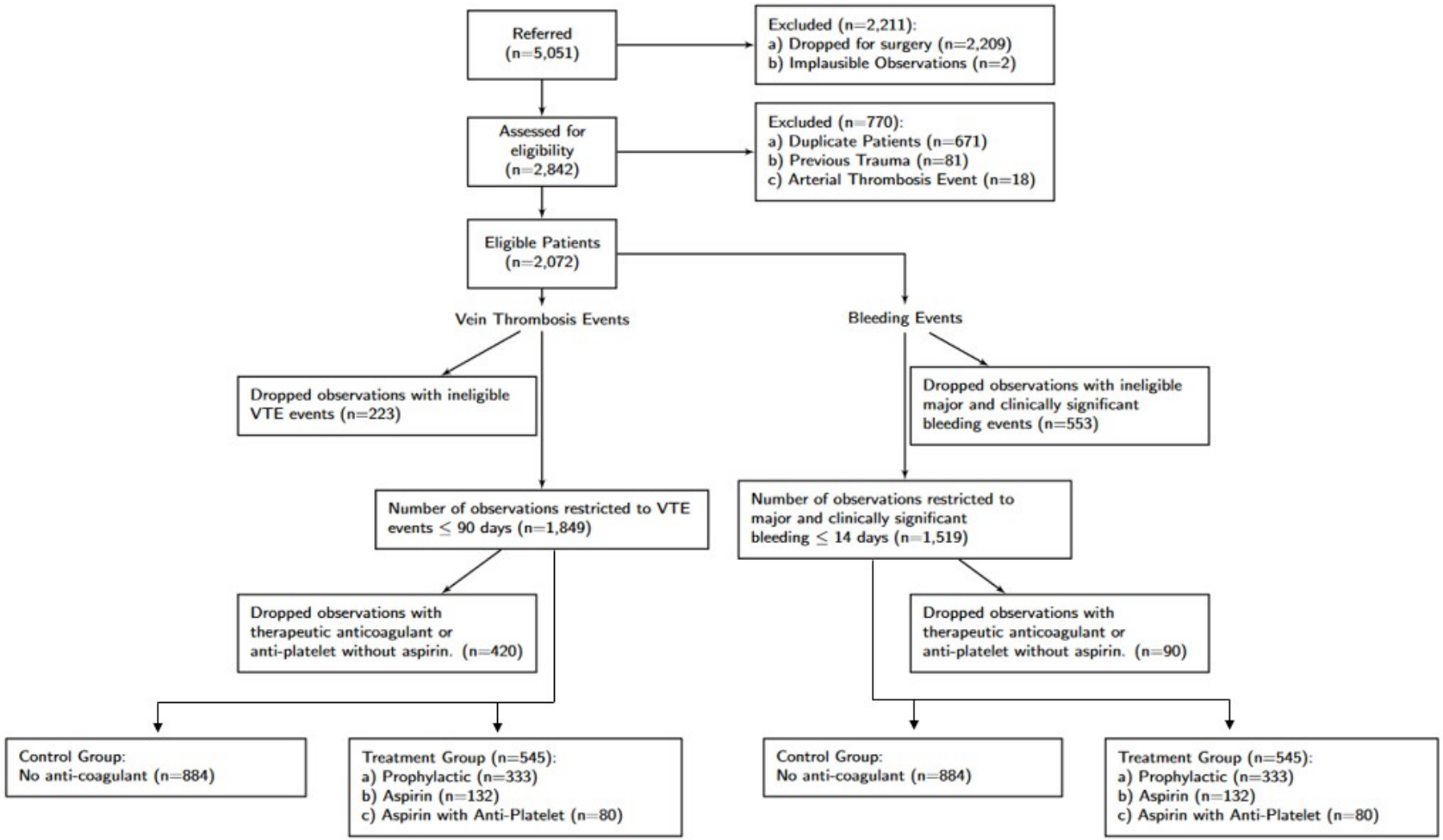


Fig 2

Figure 3

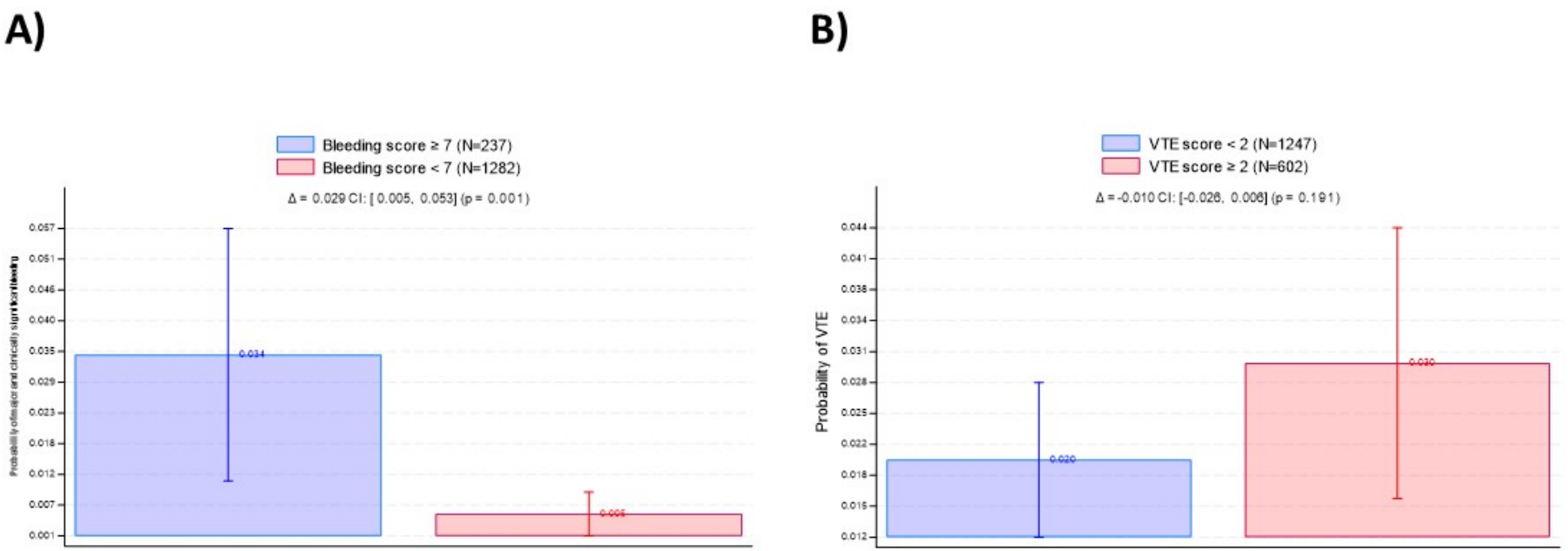


Fig 3

Figure 4

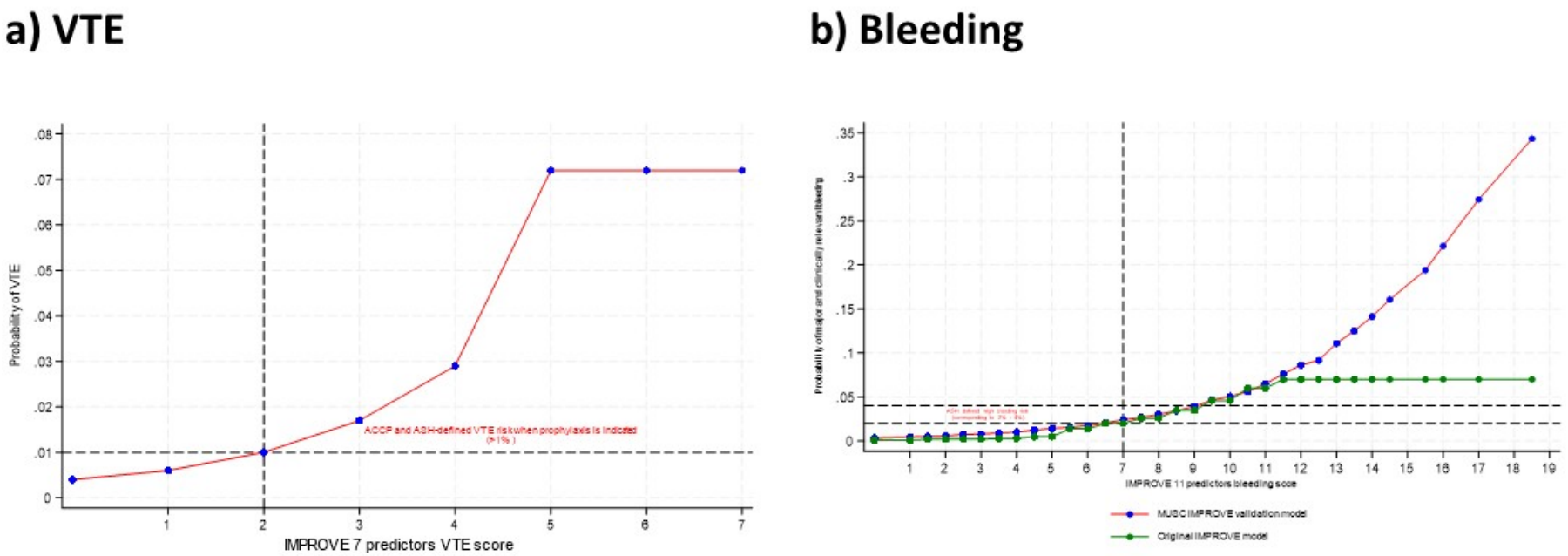
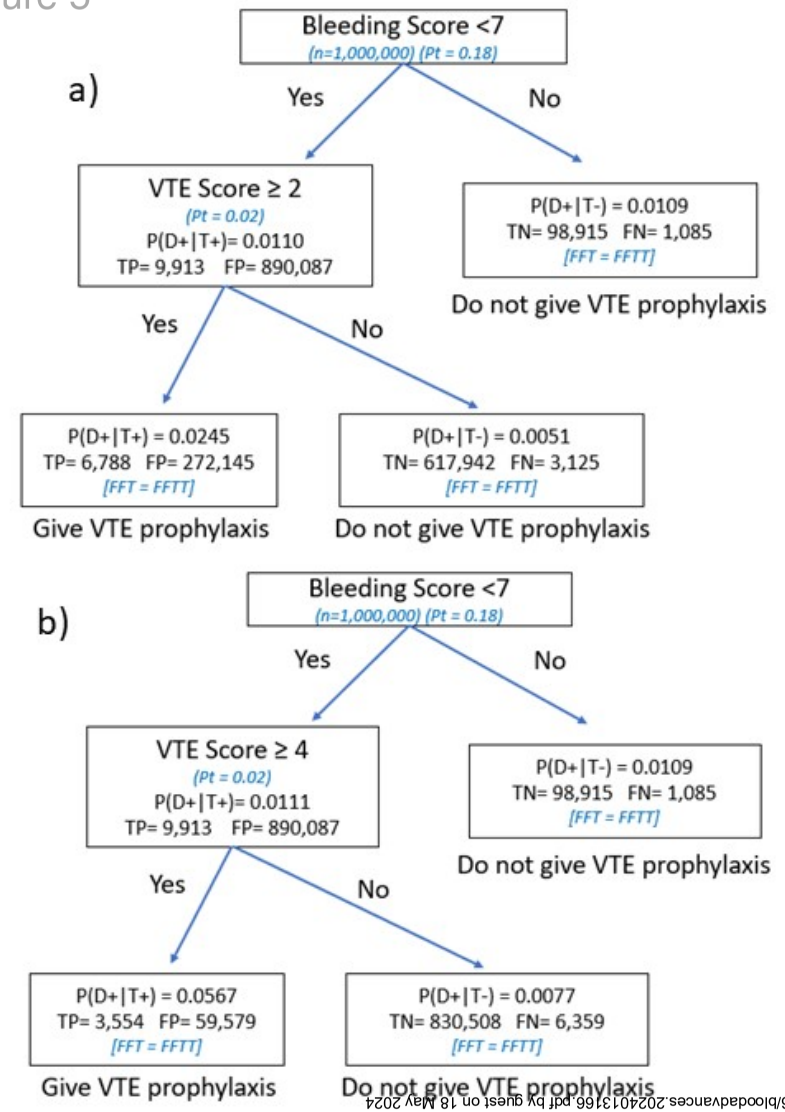


Fig 4

Figure 5



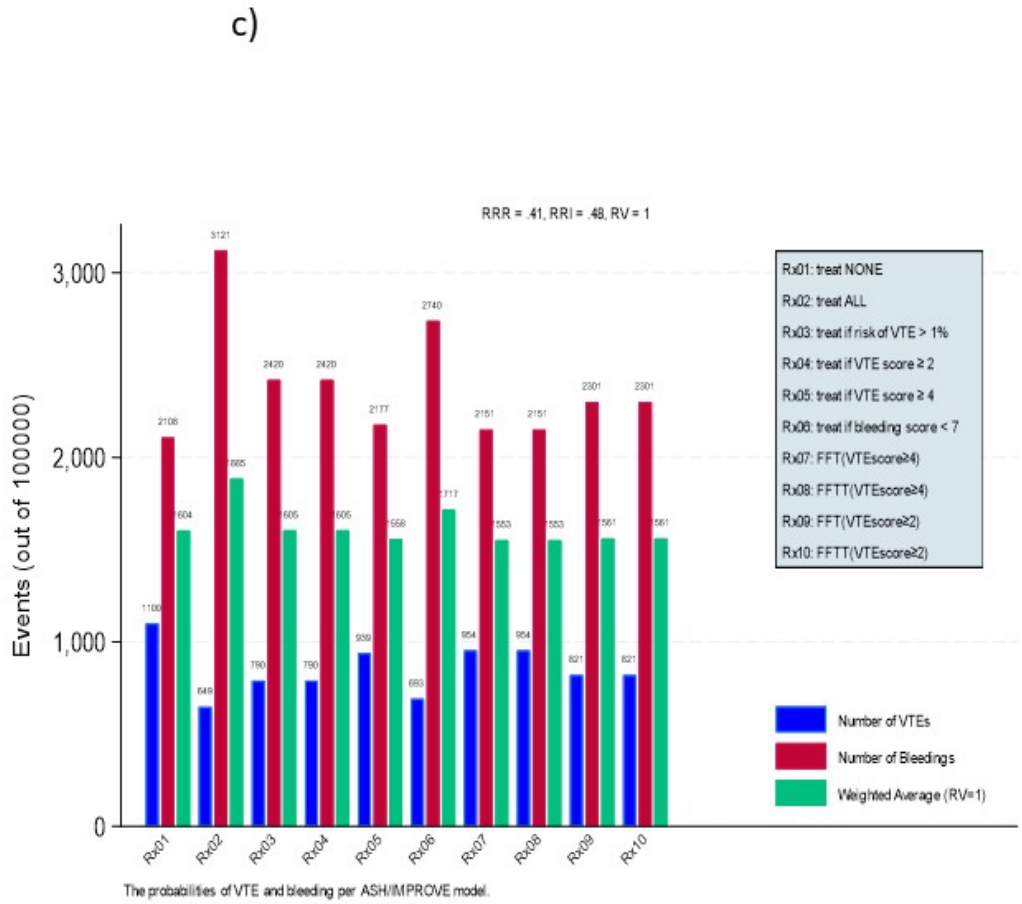


Figure 5

Figure 6

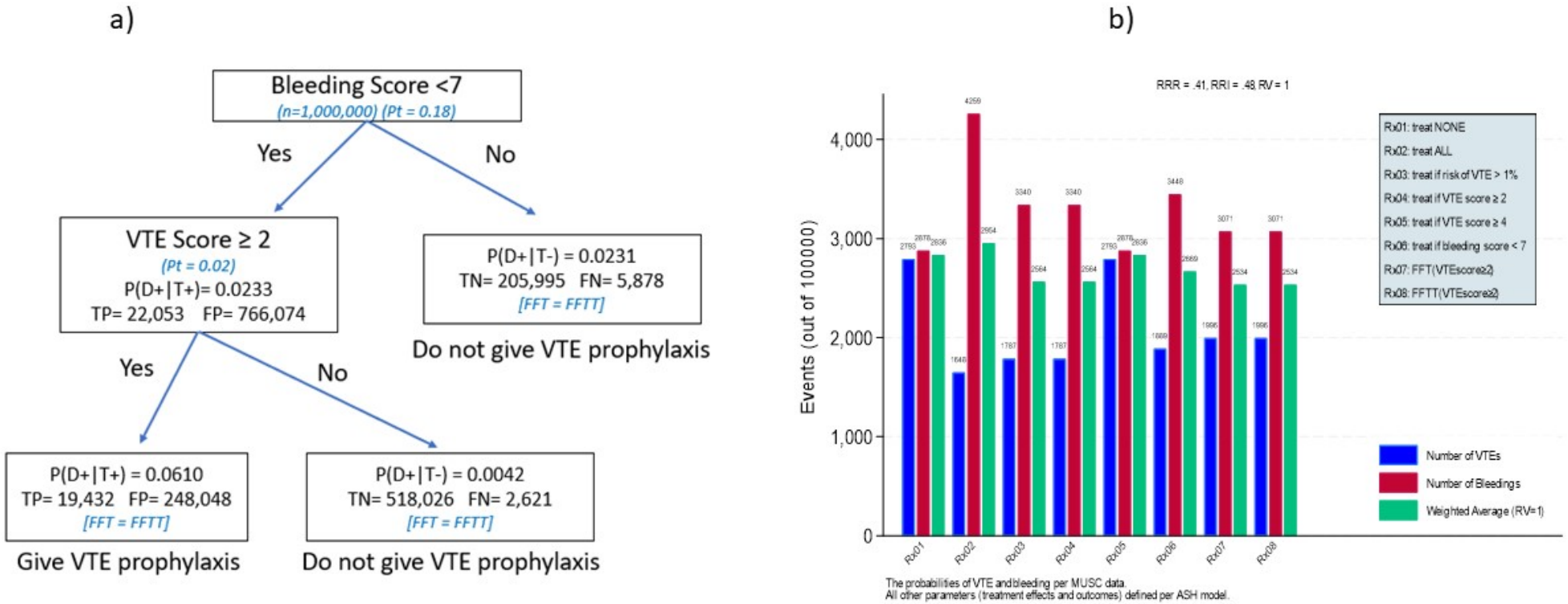


Figure 6

Figure 7

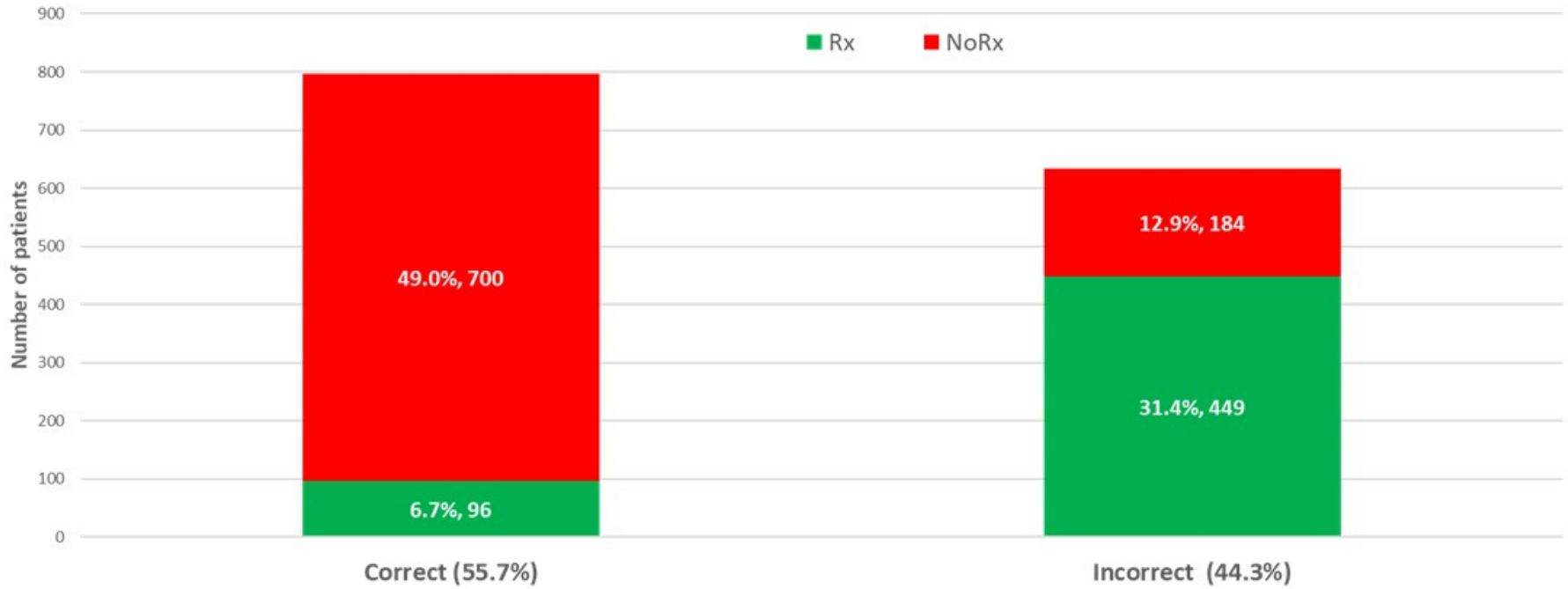


Fig 7