

A Novel Risk Assessment Model Predicts Major Bleeding Risk at Admission in Medical Inpatients

Benjamin G Mittman,^{1,2} Megan Sheehan,¹ Lisa Kojima,¹ Nicholas Cassachia,¹ Oleg Lisheba,³ Bo Hu,⁴

Matthew Pappas,¹ Michael B. Rothberg¹

¹ Center for Value-Based Care Research, Cleveland Clinic, Cleveland, OH

² Department of Population and Quantitative Health Sciences, Case Western Reserve University,
Cleveland, OH

³ Enterprise Analytics eResearch Department, Cleveland Clinic, Cleveland, OH

⁴ Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH

Short title: Predicting bleeding in medical inpatients

Keywords: bleeding; risk assessment model; clinical prediction; machine learning

Abstract word count: 327

Main text word count: 3,418

Corresponding author:

Benjamin G. Mittman, BA, Center for Value-Based Care Research, Cleveland Clinic, Cleveland Clinic, 9500
Euclid Ave, Mail Code G10, Cleveland, OH 44195; Email: MITTMAB2@ccf.org

Abstract

Background: Venous thromboembolism (VTE) is the leading cause of preventable hospital death in the US. Guidelines from the American College of Chest Physicians and American Society for Hematology recommend providing pharmacological VTE prophylaxis to acutely or critically ill medical patients at acceptable bleeding risk, but there is currently only one validated risk assessment model (RAM) for estimating bleeding risk. We developed a RAM using risk factors at admission and compared it with the International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) model.

Methods: A total of 46,314 medical patients admitted to a Cleveland Clinic Health System hospital from 2017-2020 were included. Data were split into training (70%) and validation (30%) sets with equivalent bleeding event rates in each set. Potential risk factors for major bleeding were identified from the IMPROVE model and literature review. Penalized logistic regression using LASSO was performed on the training set to select and regularize important risk factors for the final model. The validation set was used to assess model calibration and discrimination and compare performance with IMPROVE. Bleeding events and risk factors were confirmed through chart review.

Results: The incidence of major in-hospital bleeding was 0.58%. Active peptic ulcer (OR = 5.90), prior bleeding (OR = 4.24), and history of sepsis (OR = 3.29) were the strongest independent risk factors. Other risk factors included age, male sex, decreased platelet count, increased INR, increased PTT, decreased GFR, ICU admission, CVC or PICC placement, active cancer, coagulopathy, and in-hospital antiplatelet drug, steroid, or SSRI use. In the validation set, the Cleveland Clinic Bleeding Model (CCBM) had better discrimination than IMPROVE (0.86 vs. 0.72, $p < .001$) and, at equivalent sensitivity (54%), categorized fewer patients as high-risk (6.8% vs. 12.1%, $p < .001$).

Conclusions: From a large population of medical inpatients, we developed and validated a RAM to accurately predict bleeding risk at admission. The CCBM may be used in conjunction with VTE risk calculators to decide between mechanical and pharmacological prophylaxis for at-risk patients.

Introduction

Venous thromboembolism (VTE) is the leading cause of preventable hospital death in the United States (US).¹ Multiple randomized controlled trials have demonstrated that pharmacological prophylaxis with heparin is highly effective at reducing the incidence of VTE and subsequent mortality.^{2,3} However, heparin also increases the risk of bleeding.³ Therefore, the American College of Chest Physicians (ACCP)⁴ and American Society for Hematology (ASH)⁵ recommend that medical patients be assessed for risk of VTE and those not at low risk receive pharmacological VTE prophylaxis, but only if they are at acceptable bleeding risk.

However, assessment of bleeding risk can be challenging⁶ and the guidelines do not specify how bleeding risk should be determined. Unlike assessment of VTE, for which there are multiple validated risk assessment models (RAMs),⁷⁻¹¹ there is only one validated RAM available for estimating bleeding risk, the International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) model.¹² Originally derived on 9,388 patients from hospitals in 12 countries using multiple logistic regression, IMPROVE includes 11 factors at admission that were independently associated with in-hospital major or non-major but clinically relevant bleeding. A risk score was developed to assign points for each factor, allowing categorization of patients as either high- or low-risk for bleeding based on their total score.

Despite acceptable performance in the original data set, IMPROVE has key limitations. These include insufficient validation of risk factor data extracted from the electronic health records (EHRs), informal variable selection, exclusion of several important potential risk factors, and inclusion in the outcome of non-major bleeds, which although clinically relevant, typically do not warrant modification of VTE prophylaxis. Additionally, although IMPROVE has been externally validated, its performance in validation studies has been poor.^{13,14} In the largest validation study, approximately 20% of patients were categorized as high risk, sensitivity was less than 40%, and the area under the receiver operating characteristics curve (AUROC) was 0.63.¹³

A more accurate and generalizable bleeding RAM is needed to better evaluate bleeding risk at admission in medical inpatients. We identified major in-hospital bleeds and derived a RAM based on more than 30,000 patients from 10 hospitals within the Cleveland Clinic Health System (CCHS). We then validated our RAM on a holdout validation set of nearly 14,000 patients and compared its discrimination and calibration with the IMRPOVE model.

Methods

Setting and Patients

Our cohort consisted of medical inpatients who were admitted to one of 10 CCHS hospitals between October 1, 2017, and January 31, 2020. Hospitals were located in Ohio and Florida and varied in size from a 126-bed community hospital to a 1,400-bed quaternary care academic medical center. We included patients aged 18 years or older who were admitted to a medical service, either directly or from the emergency room (ER). Because we were interested in hospital-acquired bleeding, we excluded patients who had a bleed that was either documented as present-on-admission or occurred within 48 hours of admission. We also excluded patients taking warfarin because they would not be eligible to receive chemoprophylaxis. The study was approved by the Cleveland Clinic Institutional Review Board (IRB #14-240).

Identification of Bleeds

We developed an algorithm to identify major in-hospital bleeds in medical inpatients. We used the International Society on Thrombosis and Hemostasis (ISTH) definition of major bleeding—clinically overt and either fatal or associated with one of the following: (a) fall in hemoglobin of 2 g/dL or more, (b) documented transfusion of at least 2 units of packed red blood cells, or (c) involvement of a critical anatomical site (e.g., intracranial, pericardial, intramuscular with compartment syndrome, retroperitoneal).¹⁵ Additionally, the bleed could not be perioperative and must have resulted in expected blood loss. We used a combination of diagnostic codes and laboratory values to identify

patients with probable bleeds based on these criteria. We reviewed the charts of patients with probable bleeds to confirm all outcomes. We excluded patients with a present-on-admission indicator for a bleeding code, bleed during admission, or major surgical operation without a specific diagnostic code for other bleeding.

Risk Factor Identification

We identified potential risk factors for major bleeding from the IMPROVE¹² model and other published studies.^{16–20} We considered age, sex, race, smoking status, alcohol use, platelet count, international normalized ratio (INR), partial thromboplastin time (PTT), glomerular filtration rate (GFR), prior bleeding within 3 months of admission, rheumatic disease, sepsis, intensive care unit (ICU) admission within 24 hours of hospital admission, central venous catheter (CVC) placement (including peripherally inserted central catheter, PICC) within 24 hours of admission, cancer, peptic ulcer, in-hospital antiplatelet medication use, hemorrhagic or ischemic stroke, transient ischemic attack, hypertension, heart failure, pericarditis, diabetic retinopathy, hospital readmission, and myocardial infarction. All variables were extracted from the Cleveland Clinic EHR system.

Smoking status and alcohol use were categorized as current, former, or never. Platelet count (thousands of platelets per uL), INR, and PTT (seconds) values were not modified. Serum creatinine (umol/L) was used to calculate GFR (mL/min) using the revised chronic kidney disease-epidemiology age and sex (CKD-EPI_{AS}) equation.²¹ Antiplatelet medications (aspirin, clopidogrel, ticagrelor, or prasugrel) were categorized as yes/no. Readmission was defined as having a previous inpatient discharge within 30 days of admission. All other predictors were identified based on a combination of diagnostic and finance codes (see Supplementary Information).

Because many patients received chemoprophylaxis, which is hypothesized to increase the risk of major bleeding,¹⁷ we identified VTE prophylaxis (subcutaneous heparin 5,000 units 2–3 times daily, enoxaparin 40mg daily, dalteparin 5,000 units daily, or fondaparinux 2.5mg daily) from the table of

medications administered. To assess the impact of prophylaxis on model parameter estimates, we performed a sensitivity analysis, refitting a multiple logistic regression model with and without patients who received prophylaxis.

Data extracted from the EHR are often unreliable,²² thus we performed iterative chart review of a random sample of patient charts for each predictor. To ensure that each risk factor preceded the outcome, we required that all predictors be present in the first 24 hours of admission.

Model Development and Training

We first performed single imputation to fill in missing values for INR and PTT based on the distribution of values in patients with normal results. We did not perform multiple imputation because these data were not missing at random. Unlike platelet count and serum creatinine (used to calculate GFR), INR and PTT are not routinely measured in hospitalized patients unless there are specific indicating symptoms or suspected bleeding or coagulation disorders. Thus, we assumed that patients with missing INR and PTT values had values within normal range.

We split the cohort into training (70%) and validation (30%) sets with pseudo-randomization of bleeding events to ensure proportional event rates in each set. We did this to ensure model trainability and accuracy in the presence of a highly unbalanced binary outcome. We used the training set to perform variable selection in two steps. First, we used univariate analyses (t-tests and chi-squared tests for continuous and categorical variables, respectively) and collinearity assessments among the full set of predictors to exclude unimportant variables. We then used a supervised machine learning algorithm, the least absolute shrinkage selection operator (LASSO),²³ to identify remaining unimportant variables in order to minimize model complexity and reduce overfitting. Variables with no importance in predicting major bleeding were penalized and shrunk to zero and removed from the final fitted model. Included variables were regularized based on their amount of shrinkage. Cross-validation was used to determine

the penalty term, lambda, and to compute the regularized coefficient for each variable. Coefficients were exponentiated to yield odds ratios (ORs).

All analyses were performed in R (v4.2.2).²⁴ LASSO regression modeling was performed using the package glmnet (v4.1-6).²⁵ In R, glmnet automatically centers and scales all data prior to fitting the LASSO model, thus we did not standardize any variables. However, we scaled down age and platelet count by a factor of 10 to reduce variability in the range of scales between predictors (since most were binary and coded as 0 or 1) and to facilitate easier interpretation of odds ratios.

Internal Validation

The LASSO model fit on the training set was used to generate a predicted probability of bleeding for each patient in the validation set. Model calibration was assessed using a calibration plot. Model discrimination was assessed using a receiver operating characteristics (ROC) curve, the area under the ROC curve (AUROC), sensitivity, specificity, and the F1 score (harmonic mean). The optimal cut-off for predicting bleeding was chosen as the point on the final ROC curve that maximized the Youden index. Bootstrapping was used to compute the optimism-adjusted AUC and 95% confidence interval (CI) at the optimal cut-off, as well as to plot 95% confidence bands for the ROC curve at a range of sensitivity and corresponding specificity values.

External Validation and Comparison with IMPROVE

We sought to validate the IMPROVE model in an external sample and compare the performance of our LASSO model (hereafter Cleveland Clinic Bleeding Model [CCBM]) with the IMPROVE model. To do this, we fit the IMPROVE model on our validation set using multiple logistic regression¹² and generated predictions in the validation set. We then compared CCBM with IMPROVE on the basis of calibration and discrimination in the validation set. Specifically, we compared models on bleed detection and high- vs. low-risk categorization rates using chi-squared tests, and on AUC using DeLong's test for paired ROC curves.²⁶

In order to compare expected performance of the models in clinical practice, we determined the cut-off on the CCBM that corresponds to a total risk score of ≥ 7 on the IMPROVE model. We then compared F1 scores as well as high- vs. low-risk categorization rates between the models using a chi-square test. Statistical significance was defined as two-sided $p < .05$.

Comparing CCBM with Clinician Assessment

The purpose of developing this RAM was to design a tool that could aid clinician decision-making for VTE and bleeding risk in hospitalized patients. We empirically assessed clinicians' decisions for VTE prophylaxis use in patients stratified by bleeding risk to determine how well clinicians incorporate bleeding risk assessment into prophylaxis decision-making at baseline. We divided patients into deciles based on bleeding risk computed by CCBM and quantified the percentage of eligible patients who received prophylaxis in each decile. Patients ineligible for chemoprophylaxis were those whom clinicians considered to be at high risk of bleeding and were excluded from this analysis.

Results

We identified 48,030 adult medical inpatients with valid bleeding data during the study period. After excluding patients with one or more missing predictors that could not be imputed because the reason for missing values could not be determined, our final cohort contained 46,314 patients. The training set included 32,419 patients (70%) and the validation set included 13,895 patients (30%); each had a major bleed rate of 0.58%. Compared with patients without major bleeding, patients with major bleeds were older (mean age 64.5 vs. 61.4 years); had lower platelet counts (mean count 197,000 vs. 231,000 per uL) and worse renal function (mean GFR 61.5 vs. 80.7 ml/min); and were more likely to have prior bleeding (25.7% vs. 5.9%), sepsis (50% vs. 14.2%), and cancer (28% vs. 15.4%). All patient characteristics and their relationship to major bleeding are shown in Table 1.

The final LASSO model included 16 predictors; see Table 2 for predictors and regularized ORs generated in the training set. The strongest predictors of major bleeding were peptic ulcer (OR = 5.90),

prior bleed (OR = 4.24), and sepsis (OR = 3.29). In the validation set, the optimism-adjusted AUC for CCBM was 0.86 (95% CI, 0.81-0.91). At the optimal cut-off, defined as the Youden index (0.68%), sensitivity was 79%, specificity was 84%, and the F1 score was 0.055. Calibration was good (Figure 1). Predicted risk of major bleeding among all patients in the validation set ranged from 0.017% to 58%, with a mean of 0.58% (SD = 1.75%).

Sensitivity Analysis

In addition to fitting the LASSO model, we fit multiple logistic regression models on the entire cohort with and without patients who received VTE prophylaxis to assess the impact of prophylaxis on model parameter estimates. Peptic ulcer's effect was significantly reduced in the model excluding patients who received prophylaxis (OR = 3.06) compared with the full cohort model (OR = 5.81, 95% CI: 3.99, 8.26), but all other variables did not differ significantly between models (Table 2). Based on these results, we kept all patients with complete data in the final LASSO model, regardless of prophylaxis receipt.

External Validation of IMPROVE

The optimism-adjusted AUC for IMPROVE was 0.72 (95% CI, 0.66-0.78). Using the IMPROVE cut-off of ≥ 7 points, sensitivity was 54%, specificity was 88%, and the F1 score was 0.050. Calibration was generally good but IMPROVE tended to over-predict patients' risk in the top 3 deciles (Figure S1).

Comparing CCBM with IMPROVE

The ROC curves with 95% confidence bands are compared in Figure 2. At most values of sensitivity and corresponding specificity, including at the optimal cut-off point for IMPROVE (defined as ≥ 7 points), CCBM showed better discrimination than IMPROVE. The overall difference was robust; CCBM had a significantly higher area under the curve (AUC) compared with IMPROVE (0.86 vs. 0.72, $Z = 4.20$, $p < .001$; 95% CI for the difference between AUCs = 0.08-0.21).

Next, we matched CCBM and IMPROVE on sensitivity and compared the models on high- vs. low-risk categorization rates. In our validation set, IMPROVE identified 44/81 (54%) of bleeds using the cut-off of ≥ 7 points. This equated to a risk cutoff of 1.45% using CCBM. At a sensitivity of 54%, IMPROVE categorized 1,676/13,895 (12.1%) of patients as high-risk, whereas CCBM categorized 950/13,895 (6.8%) patients as high-risk. The mean predicted probability of developing a major bleed among patients categorized as high-risk was 4.28% for CCBM and 3.35% for IMPROVE. A chi-square test comparing rates of high-risk vs. low-risk categorization rates between models showed that at a sensitivity of 54%, CCBM categorized significantly fewer patients as high-risk compared with IMPROVE (6.8% vs. 12.1%, $\chi^2 = 221.7$, $p < .001$).

We also matched CCBM and IMPROVE on the basis of number categorized as high-risk (12.1%) and compared sensitivity. A chi-square test comparing number of bleeds detected showed that CCBM detected significantly more bleeds than IMPROVE (70% vs. 54%, $\chi^2 = 4.44$, $p = .035$).

Comparing CCBM with Clinician Assessment

A total of 319 patients in the validation set were ineligible for chemoprophylaxis due to clinician assessment of elevated bleeding risk. Among eligible patients, the mean risk for major bleeding by decile, as computed by CCBM, is shown in Figure 3 (left y-axis). The percentage of eligible patients who received VTE prophylaxis in each decile is also shown in Figure 3 (right y-axis). Mean predicted risk increased by decile, with tenth decile patients showing substantially increased predicted risk compared with patients in the first nine deciles. The percentage of patients who received prophylaxis also increased by decile, with the highest risk patients (tenth decile) more likely to have received prophylaxis than the lowest risk patients (first decile).

Discussion

In this retrospective cohort study, we developed and validated a RAM for major bleeding using data from more than 45,000 patients from 10 hospitals within the Cleveland Clinic Health System. The

model, CCBM, contained 17 predictors extracted from the EHR system and demonstrated good calibration and discrimination in the validation set. The CCBM outperformed the IMPROVE¹² model, which has been the only validated bleeding RAM for more than 10 years. When matching CCBM and IMPROVE on sensitivity, CCBM categorized 44% fewer patients as high risk than IMPROVE, and high risk-patients identified by CCBM had a 28% higher mean major bleeding risk. Thus, if we focus on sensitivity, use of CCBM would prevent as many major bleeds as IMPROVE but allow more patients to receive VTE prophylaxis without increasing their risk of bleeding. Overall, CCBM showed better discrimination than IMPROVE throughout the probable operating range of the ROC curve.

Our RAM differs from IMPROVE in several important ways. First, we had a larger number of patients, which allowed us to identify more risk factors that were independently associated with bleeding and thus improve predictive performance. Second, we restricted the outcome to major bleeds whereas IMPROVE included nonmajor bleeds, which decreases the clinical relevance of predictions because patients at elevated risk of nonmajor bleeding should still receive pharmacological VTE prophylaxis.²⁷ Third, we validated both outcomes and all predictors by using both ICD codes and chart review procedures, whereas IMPROVE validated outcomes but relied only on ICD codes to identify risk factors.

Finally, by using a supervised machine learning algorithm, LASSO,²³ we developed a model that achieved better discrimination in the validation set and is less likely to show degraded performance in future validation sets due to the implementation of variable regularization. By using LASSO, we were also able to identify several additional important risk factors for major bleeding that IMPROVE did not include. Moreover, the IMPROVE model utilizes logistic regression, in which each variable is given the same weight in predicting the outcome. However, LASSO penalizes less important variables, which reduces overfitting and produces less biased predictions in the validation set.²⁸ This is important because characteristics vary by region; in our CCHS cohort, peptic ulcers and prior bleeding were the

strongest predictors of major bleeding, but these relationships might be weaker in other health systems or patient populations.

Guidelines from both the ACCP and ASH recommend that bleeding risk assessment be performed in conjunction with VTE risk assessment to support VTE prophylaxis decision-making,^{4,5} because patients at increased bleeding risk should not receive pharmacological prophylaxis. But assessing bleeding risk is challenging and clinicians frequently misclassify patients; in our cohort, patients with the greatest risk for major bleeding were more likely to receive prophylaxis than those with the lowest risk. These findings indicate the need for an accurate RAM for major bleeding. The IMPROVE model has been the only validated RAM since 2011, but we have developed a novel RAM that demonstrates better performance. This RAM may be used in conjunction with readily available VTE risk calculators to identify patients at high bleeding risk and decide between mechanical and pharmacological prophylaxis for patients who are at elevated risk of both VTE and bleeding.

Our study makes important contributions but should be evaluated within the context of its limitations. First, although we validated both the outcomes and predictors via chart review, it is possible that we missed some bleeding events or misclassified some of the risk factors due to the imperfect nature of EHR data. Due to the relatively small number of bleeds, it was possible to manually confirm each individual outcome. However, with over 45,000 patients, we could not manually confirm every predictor for each patient. Instead, for each risk factor, we randomly selected a manageable subset of patients and performed iterative chart review until we were confident all variables were as accurate as possible. Second, we had a small number of outcomes to develop and test our model. Our cohort's major bleeding rate was 0.58%, less than the total bleeding rate in the IMPROVE cohort and external validation studies, which included nonmajor bleeding. Having a highly imbalanced outcome can pose challenges for prediction models and limits the utility of certain evaluation metrics in assessing model performance.²⁹

Third, we included patients who received VTE prophylaxis in our cohort. However, our sensitivity analysis showed that most of the variable estimates were not significantly affected by including patients who received prophylaxis. Only peptic ulcer's effect was meaningfully different, but this risk factor was uncommon in our cohort and thus unlikely to have driven the predictive accuracy of the model. Finally, we did not perform an external validation of our RAM. By splitting our cohort into training and validation sets and adjusting AUC for optimism using bootstrapping, we reduced the bias in our evaluation of model performance. However, the patients in our cohort all came from a single health system, therefore follow-up studies will be needed to validate our RAM in patient populations with different characteristics and distributions of the outcome and risk factors.

Conclusion

The ACCP and ASH guidelines recommend incorporating bleeding risk assessment into decision-making for VTE prophylaxis, but since 2011 only one bleeding RAM—IMPROVE—has been validated. Using more than 45,000 patient records from 10 hospitals within a single health system, we developed a novel RAM that had good calibration and discrimination in a holdout validation set. Compared with IMPROVE, use of our RAM would allow more patients to receive VTE chemoprophylaxis without increasing their bleeding risk. Future studies should validate the model in external data sets and examine the impact of EHR integration on prophylaxis use, bleeding rates, and other patient outcomes.

References

1. Beckman MG, Hooper WC, Critchley SE, Ortel TL. Venous Thromboembolism: A Public Health Concern. *Am J Prev Med*. 2010;38(4):S495-S501. doi:10.1016/J.AMEPRE.2009.12.017
2. Lilly CM, Liu X, Badawi O, Franey CS, Zuckerman IH. Thrombosis Prophylaxis and Mortality Risk Among Critically Ill Adults. *Chest*. 2014;146(1):51-57. doi:10.1378/CHEST.13-2160
3. Kakkar AK, Cimminiello C, Goldhaber SZ, Parakh R, Wang C, Bergmann JF. Low-Molecular-Weight Heparin and Mortality in Acutely Ill Medical Patients. *New England Journal of Medicine*. 2011;365(26):2463-2472. doi:10.1056/NEJMOA1111288/SUPPL_FILE/NEJMOA1111288_DISCLOSURES.PDF
4. Kearon C, Akl EA, Comerota AJ, et al. Antithrombotic therapy for VTE disease: Antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2012;141(2 SUPPL.):e419S-e496S. doi:10.1378/CHEST.11-2301/ATTACHMENT/7547E95F-38B2-4625-A9BB-CE44DDD34D42/MMC1.PDF
5. Schünemann 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. doi:10.1182/BLOODADVANCES.2018022954
6. Klok FA, Huisman M V. How I assess and manage the risk of bleeding in patients treated for venous thromboembolism. *Blood*. 2020;135(10):724-734. doi:10.1182/BLOOD.2019001605
7. Greene MT, Spyropoulos AC, Chopra V, et al. Validation of Risk Assessment Models of Venous Thromboembolism in Hospitalized Medical Patients. *Am J Med*. 2016;129(9):1001.e9-1001.e18. doi:10.1016/J.AMJMED.2016.03.031
8. Barbar S, Noventa F, Rossetto V, et al. A risk assessment model for the identification of hospitalized medical patients at risk for venous thromboembolism: the Padua Prediction Score.

- Journal of Thrombosis and Haemostasis*. 2010;8(11):2450-2457. doi:10.1111/J.1538-7836.2010.04044.X
9. Spyropoulos AC, Anderson FA, FitzGerald G, et al. Predictive and Associative Models to Identify Hospitalized Medical Patients at Risk for VTE. *Chest*. 2011;140(3):706-714. doi:10.1378/CHEST.10-1944
 10. Nendaz M, Spirk D, Kucher N, et al. Multicentre validation of the Geneva Risk Score for hospitalised medical patients at risk of venous thromboembolism: Explicit Assessment of Thromboembolic Risk and Prophylaxis for Medical Patients in Switzerland (ESTIMATE). *Thromb Haemost*. 2014;111(3):531-538. doi:10.1160/TH13-05-0427/ID/JR0427-4
 11. Stuck AK, Spirk D, Schaudt J, Kucher N. Risk assessment models for venous thromboembolism in acutely ill medical patients: A systematic review. *Thromb Haemost*. 2017;117(4):801-808. doi:10.1160/TH16-08-0631/ID/JR0631-6
 12. 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. doi:10.1378/CHEST.09-3081
 13. Rosenberg DJ, Press A, Fishbein J, et al. External validation of the IMPROVE bleeding Risk Assessment Model in medical patients. *Thromb Haemost*. 2016;116(3):530-536. doi:10.1160/TH16-01-0003/ID/JR0003-8
 14. 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. doi:10.1378/CHEST.14-2842
 15. Schulman S, Kearon C. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. *J Thromb Haemost*. 2005;3(4):692-694. doi:10.1111/J.1538-7836.2005.01204.X

16. Kander T, Schött U. Bleeding complications after central line insertions and the relevance of pre-procedure coagulation tests and blood component therapy. *Critical Care* 2013 17:2. 2013;17(2):1-200. doi:10.1186/CC12296
17. Darzi AJ, Karam SG, Charide R, et al. Prognostic factors for VTE and bleeding in hospitalized medical patients: a systematic review and meta-analysis. *Blood*. 2020;135(20):1788-1810. doi:10.1182/BLOOD.2019003603
18. Langsted A, Nordestgaard BG. Smoking is Associated with Increased Risk of Major Bleeding: A Prospective Cohort Study. *Thromb Haemost*. 2019;119(1):39-47. doi:10.1055/S-0038-1675798/ID/JR180444-10
19. Kirchhof P, Haas S, Amarenco P, et al. Impact of modifiable bleeding risk factors on major bleeding in patients with atrial fibrillation anticoagulated with rivaroxaban. *J Am Heart Assoc*. 2020;9(5). doi:10.1161/JAHA.118.009530
20. Liabeuf S, Scaltieux LM, Masmoudi K, et al. Risk factors for bleeding in hospitalized at risk patients with an INR of 5 or more treated with Vitamin K antagonists. *Medicine (United States)*. 2015;94(52). doi:10.1097/MD.0000000000002366
21. Inker LA, Eneanya ND, Coresh J, et al. New Creatinine- and Cystatin C–Based Equations to Estimate GFR without Race. *New England Journal of Medicine*. 2021;385(19):1737-1749. doi:10.1056/NEJMOA2102953/SUPPL_FILE/NEJMOA2102953_DISCLOSURES.PDF
22. Taksler GB, Dalton JE, Perzynski AT, et al. Opportunities, Pitfalls, and Alternatives in Adapting Electronic Health Records for Health Services Research. *Med Decis Making*. 2020;41(2):133-142. doi:10.1177/0272989X20954403
23. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1996;58(1):267-288. doi:10.1111/J.2517-6161.1996.TB02080.X

24. R Core Team. R: A language and environment for statistical computing. Published online 2022.
25. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33(1):1. doi:10.18637/jss.v033.i01
26. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*. 1988;44(3):837. doi:10.2307/2531595
27. Kaatz S, Ahmad D, Spyropoulos AC, Schulman S. Definition of clinically relevant non-major bleeding in studies of anticoagulants in atrial fibrillation and venous thromboembolic disease in non-surgical patients: Communication from the SSC of the ISTH. *Journal of Thrombosis and Haemostasis*. 2015;13(11):2119-2126. doi:10.1111/jth.13140
28. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events. *BMJ*. 2015;351. doi:10.1136/BMJ.H3868
29. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*. 2020;21(1):1-13. doi:10.1186/S12864-019-6413-7/TABLES/5

Acknowledgements

This study was supported in part by NIH grants 5T32GM007250-45 and 5TL1TR002549-04.

Table 1. Patient characteristics and their relationship to major in-hospital bleeding.

Characteristic	Total (n = 46,311)*	No Bleed (n = 46,043)	Major Bleed (n = 268)	p-Value
Age	61.4 (19.0)	61.4 (19.0)	64.5 (14.6)	0.007
Male sex	47.5%	47.5%	53.4%	0.064
Race				0.202
Black	21.1%	21.1%	18.3%	
White	72.3%	72.3%	72.8%	
Other	6.6%	6.6%	9.0%	
Platelets	231.0 (97.4)	231.2 (97.2)	196.6 (123.7)	<0.001
International normalized ratio	1.06 (0.26)	1.06 (0.26)	1.31 (0.50)	<0.001
Partial thromboplastin time	27.7 (5.8)	27.7 (5.7)	31.5 (12.4)	<0.001
Glomerular filtration rate	80.6 (32.2)	80.7 (32.2)	61.5 (36.6)	<0.001
Prior bleed	6.0%	5.9%	25.7%	<0.001
Rheumatic disease	4.3%	4.3%	6.0%	0.223
Sepsis	14.4%	14.2%	50.0%	<0.001
Intensive care unit	19.0%	18.8%	43.7%	<0.001
CVC or PICC	2.7%	2.7%	10.4%	<0.001
Cancer	15.5%	15.4%	28.0%	<0.001
Peptic Ulcer	2.4%	2.4%	15.3%	<0.001
Antiplatelet drug**	38.0%	38.0%	48.5%	<0.001
Smoking (%)				<0.001
Current	37.6%	37.6%	30.2%	
Never	41.5%	41.6%	30.2%	
Former	13.2%	13.2%	13.4%	
Unknown	7.7%	7.6%	26.1%	
Alcohol use (%)				<0.001
Former	45.1%	45.2%	33.2%	
Current	40.4%	40.5%	33.6%	
Unknown	12.7%	12.6%	31.3%	
Never	1.8%	1.8%	1.9%	
Hemorrhagic stroke	1.3%	1.3%	1.5%	>0.99
Ischemic stroke	4.0%	4.05	6.0%	0.138
Transient ischemic attack	3.4%	3.4%	3.0%	0.841
Myocardial infarction	18.7%	18.7%	23.5%	0.052
Hypertension	32.8%	32.85	27.2%	0.06
Heart failure	14.0%	13.9%	25.0%	<0.001
Coagulopathy	1.5%	1.4%	11.9%	<0.001
CNS neoplasm	0.3%	0.35	0.7%	0.506
Pericarditis	0.2%	0.2%	0.0%	0.99
Diabetic retinopathy	1.7%	1.7%	1.9%	0.983
Readmission	5.3%	5.3%	7.1%	0.242
Steroid	24.6%	24.4%	51.9%	<0.001
SSRI	14.8%	14.8%	16.4%	0.51
NSAID	23.1%	23.1%	17.9%	0.053
VTE prophylaxis	58.2%	58.1%	66.4%	0.007

Notes: Platelets are in thousands per uL; partial thromboplastin time is in seconds; glomerular filtration rate is in milliliters per minute (mL/min).

Abbreviations: CVC, central venous catheter; PICC, peripherally inserted central catheter; SSRI, selective serotonin reuptake inhibitor; NSAID, nonsteroidal anti-inflammatory drug; VTE, venous thromboembolism.

*3 patients included in the final LASSO model did not have prophylaxis data.

**Antiplatelet drugs included aspirin, clopidogrel, ticagrelor, and prasugrel

Table 2. Multivariable and LASSO models for major in-hospital bleeding.

Characteristic	OR excluding prophylaxis (n = 19,365)	OR including prophylaxis (n = 46,311)	OR including prophylaxis (95% CI)		OR including prophylaxis, LASSO (n = 46,314)
Age*	0.91	0.87	0.81	0.95	0.91
Male Sex	0.86	1.02	0.80	1.30	1.16
Platelets*	0.97	0.98	0.97	0.99	0.98
International normalized ratio	1.16	1.23	1.00	1.48	1.31
Partial thromboplastin time	1.23	1.23	1.11	1.36	1.33
Glomerular filtration rate	0.87	0.89	0.85	0.92	0.91
Prior Bleed	4.58	3.93	2.94	5.20	4.24
Sepsis	3.53	3.36	2.60	4.33	3.29
Intensive care unit	1.24	1.61	1.23	2.11	1.78
CVC or PICC	1.22	1.01	0.66	1.52	0.71
Cancer	2.03	1.60	1.21	2.11	1.83
Peptic Ulcer	3.06	5.81	3.99	8.26	5.90
Antiplatelet	2.27	1.84	1.42	2.37	1.57
Coagulopathy	3.53	2.46	1.54	3.84	1.87
Steroid	2.94	2.69	2.11	3.43	2.61
SSRI	1.65	1.11	0.79	1.53	1.14
NSAID	-	-	-	-	1.00
VTE prophylaxis	-	1.37	1.05	1.80	-

Abbreviations: OR, odds ratio; LASSO, least absolute shrinkage selection operator; CI, confidence interval; CVC, central venous catheter; PICC, peripherally inserted central catheter; SSRI, selective serotonin reuptake inhibitor; NSAID, nonsteroidal anti-inflammatory drug; VTE, venous thromboembolism.

*Scaled down by a factor of 10 (e.g., from years to tens of years)

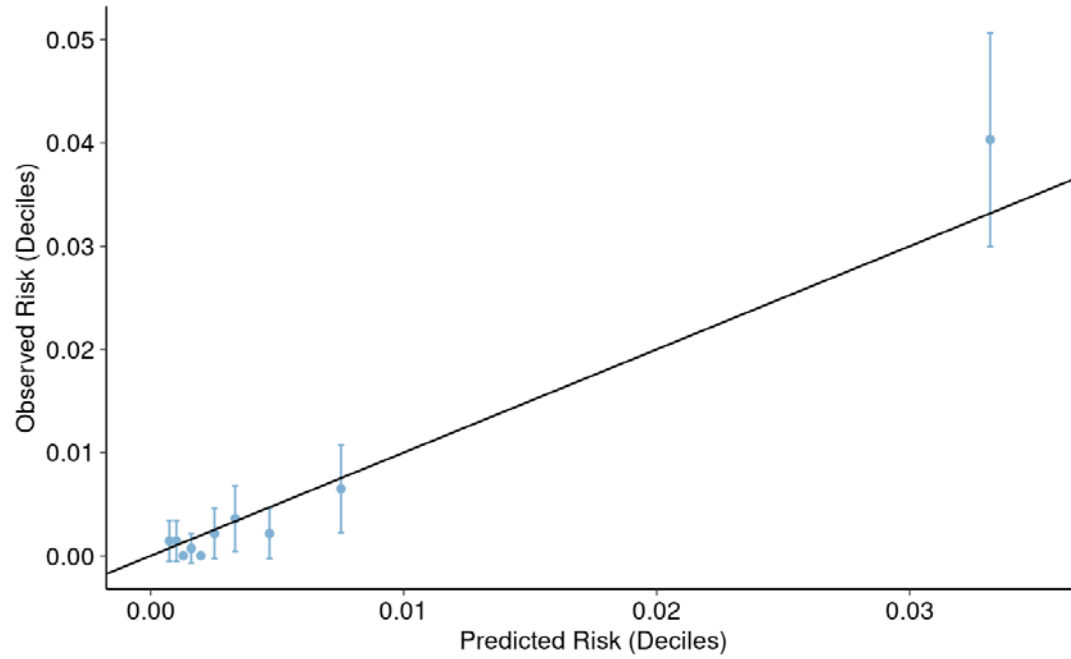


Figure 1. Calibration plot for the Cleveland Clinic Bleeding Model in the validation set.

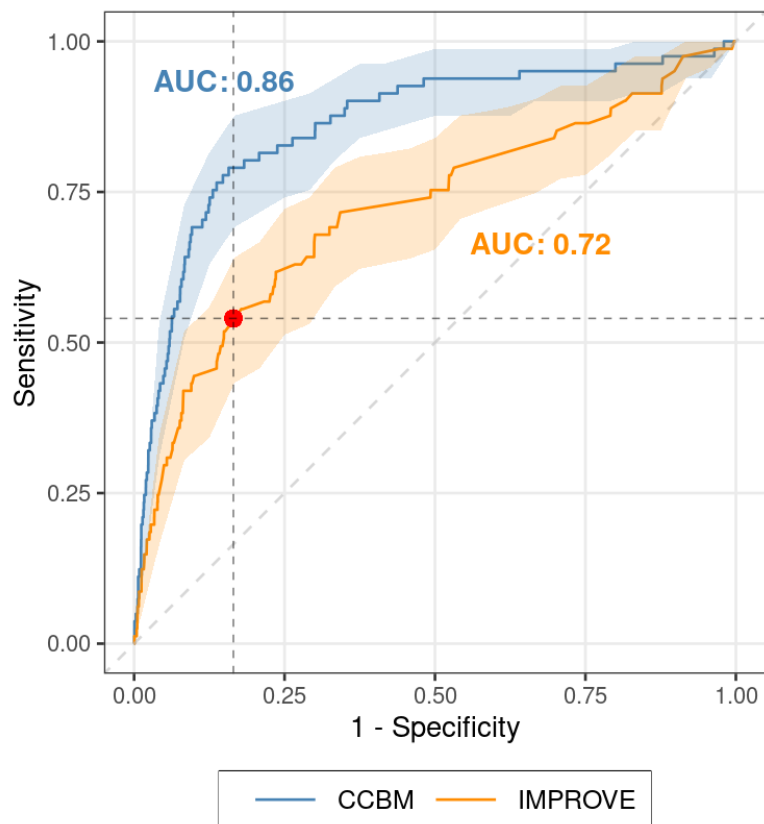


Figure 2. Receiver operating characteristics (ROC) curve comparing the Cleveland Clinic Bleeding Model (CCBM; AUC = 0.86, 95% CI: 0.81-0.91) versus the IMPROVE model (AUC = 0.72, 95% CI: 0.66-0.78) in the validation set (n = 13,895), with bootstrapped 95% confidence bands. The optimal cut-off point for IMPROVE (≥ 7 points) is indicated by the red dot. Horizontal and vertical dashed lines intersect with CCBM at corresponding operating points given equivalent sensitivity or specificity, respectively.

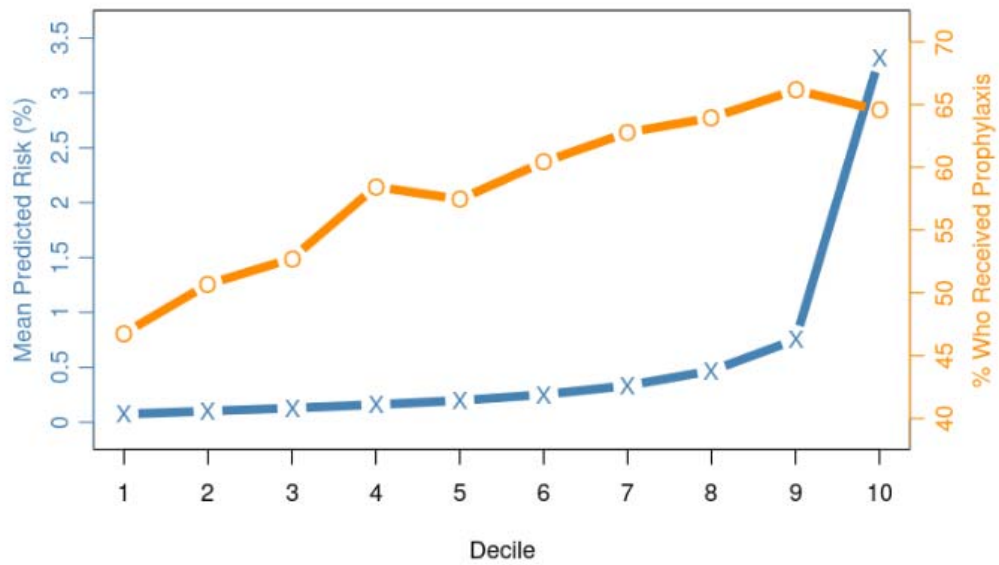


Figure 3. Comparison of mean predicted risk scores computed by CCBM (left y-axis) versus clinician VTE prophylaxis decisions in patients eligible for chemoprophylaxis (right y-axis), by decile of predicted risk.