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

Bo Hu,⁴
ty,

Benjamin G Mittman,*^{,*} Megan Sheehan,*
Matthew Pa_l
Inter for Value-Based Care Research, Cleve
Pratment of Population and Quantitative Lisa Kojima,⁴
ppas,¹ Michae
eland Clinic, C
Health Scien Nicholas Cassachia,⁴
2l B. Rothberg¹
1.
Ceveland, OH
ces, Case Western Re Oleg Lisheba,³
3
serve Universit Matthew Pappas,⁺
search, Cleveland
Quantitative Healt Michael B. Rothberg*
Clinic, Cleveland, OH
h Sciences, Case West \ddot{a} ² Center for Value-Based Care Research, Cleveland Clinic, Cleveland, OH
² Department of Population and Quantitative Health Sciences, Case Wes
Cleveland, OH
³ Enterprise Analytics eResearch Department, Cleveland Clini

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

- 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 Heal $\frac{1}{3}$ Enterprise An
 $\frac{4}{3}$ Department of
Short title: Pre Enterprise Analytics entertains of paramerity extending the chinic, Cleveland, OH
Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH
Short title: Predicting bleeding in medical inpatients
Seywords: 4

Department of Quantitative Health Sciences, Cleveland, OH Clinics, Cleveland, OH
Short title: Predicting bleeding in medical inpatients
Seywords: bleeding; risk assessment model; clinical prediction; machine learni
Nostrac Keywords: bleeding; risk assessment model; clinical p
Abstract word count: 327
Main text word count: 3,418

Reykstands: bleeding; risk assessment model; clinical prediction; machine learning
Abstract word count: 3,418
Corresponding author:

Main text word count: 3,4
Corresponding author:
Benjamin G. Mittman, BA, Corresponding author:
Benjamin G. Mittman, BA, Ce
Euclid Ave, Mail Code G10, Cl (
|}
|} Benjamin G. Mittman, I
Euclid Ave, Mail Code G
. Benjamin Gramman, Bar, Center for Care Research Care Research, Cleveland Clinic, Cleveland Clinic, Cleveland C
Euclid Ave, Mail Code G10, Cleveland, OH 44195; Email: MITTMAB2@ccf.org 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 pharma The Chinamics from the American College of Chest Physician Chinamic College of Chestmanne By
recommend providing pharmacological VTE prophylaxis to acutely or critically ill medical patients at
acceptable bleeding risk, bu recommend providing providing providing providing to 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 ad actimating bleeding risk. We developed a RAM using risk factors at admission and compared it with
International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) model.
Methods: A total of 46,314 medical pa

estimational 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 Internative Methods: A total of 46,314 medical patients admitted to a Cleveland Clinic Health System hot
2017-2020 were included. Data were split into training (70%) and validation (30%) sets with
bleeding event rates in e 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 r 2021-2021 were included. Data were split into training (70%) and validation (30%) sets unit equivalent
bleeding event rates in each set. Potential risk factors for major bleeding were identified from the
IMPROVE model and bleeding emodel and literature review. Penalized logistic regression using LASSO was performed or
training set to select and regularize important risk factors for the final model. The validation set was
used to assess mode Intertualing 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 ris training set to select and regularize important risk for the final model of the final model. Bleed
used to assess model calibration and discrimination and compare performance with IMPROVE. Bleed
events and risk factors wer

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 Results: The incidence of major in-hospital bleeding was 0.589
bleeding (OR = 4.24), and history of sepsis (OR = 3.29) were th
Other risk factors included age, male sex, decreased platelet c
decreased GFR, ICU admission, C Results: The incidence of major in-hospital bleeding was 0.58%. Active peptic ulcer (OR = 5.50), prior
bleeding (OR = 4.24), and history of sepsis (OR = 3.29) were the strongest independent risk factors.
Other 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, steroi Other random and alternative and age, male sex, decreased GFR, ICU admission, CVC or PICC placement, active cancer, coagulopathy, and in-hospit antiplatelet drug, steroid, or SSRI use. In the validation set, the Cleveland antiplatelet drug, steroid, or SSRI use. In the validation set, the Cleveland Clinic Bleeding Model (CCB
had better discrimination than IMPROVE (0.86 vs. 0.72, p < .001) and, at equivalent sensitivity (54%)
categorized fe and 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 i

had better and better distributed to the than μ (1.12, p.m., at equivalent sensitivity (54, p.m.)
categorized fewer patients as high-risk (6.8% vs. 12.1%, p < .001).
Conclusions: From a large population of medical inp **Conclusions**: From a large population of medical inpatients, we de
accurately predict bleeding risk at admission. The CCBM may be u
calculators to decide between mechanical and pharmacological pr Econciduations: From a large population of medical inpatients, we developed and validated a NAM to
accurately predict bleeding risk at admission. The CCBM may be used in conjunction with VTE risk
calculators to decide betw calculators to decide between mechanical and pharmacological prophylaxis for at-risk patients.

calculators to decide between mechanical and pharmacological prophylaxis for at-risk patients. calculators to decide between mechanical and pharmacological prophylaxis for at-risk patients.

Introduction

Venoticial SMS).¹ Multiple randomized controlled trials have demonstrated that pharmacological prophylaxis parin is highly effective at reducing the incidence of VTE and subsequent mortality.^{2,3} However, also increase States (US).⁴
with heparin
heparin also
and America
VTE and tho Multiple randomized controlled trials have demonstrated that pharmacological prophylamic
is highly effective at reducing the incidence of VTE and subsequent mortality.^{2,3} However,
increases the risk of bleeding.³ Ther 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)
and Ameri heparin also increases the risk of bleeding."
and American Society for Hematology (ASH
VTE and those not at low risk receive pharm
bleeding risk.
However, assessment of bleeding ri Therefore, the American College of Chest Physicians (ACCP)*

)⁵ recommend that medical patients be assessed for risk of

nacological VTE prophylaxis, but only if they are at acceptable

sk can be challenging⁶ and the g and American Society for Hematology (ASH)⁵
VTE and those not at low risk receive pharma
bleeding risk.
However, assessment of bleeding risl
bleeding risk should be determined. Unlike a recommend that medical patients be accreted for risk of
icological VTE prophylaxis, but only if they are at acceptab
k can be challenging⁶ and the guidelines do not specify ho
ssessment of VTE, for which there are multip

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 a bleeding

Howe

bleeding risk s

risk assessmer

risk, the Intern However, assessment of bleeding risk can be challenging^o
I risk should be determined. Unlike assessment of VTE, for
Inssment models (RAMs), ^{7–11} there is only one validated RA
International Medical Prevention Registry and the guidelines as *increpent, the the guidated*
M available for estimating bleeding
omboembolism (IMPROVE) model.¹²
is using multiple logistic regression, bleeding
The validated RAM available for estimating bleeding
The determined. The international Medical Prevention Registry on Venous Thromboembolism (IMPROVE) model.¹²
Originally derived on 9,388 patients from hospitals 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 o risk, the International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) model.**
Originally derived on 9,388 patients from hospitals in 12 countries using multiple logistic regression,
IMPROVE includes 11 f IMPROVE includes 11 factors at admission that were independently associated with in-hospital majo
non-major but clinically relevant bleeding. A risk score was developed to assign points for each facto
allowing categorizati IMPLET IMPLET INTERNATION THE TERRAT INTERNATION INTITY INTITY INTITY INTITY INTITY INTITY INTITY INTERNATION
INDITY INTERNATION IN FACTO INTERNATION IN A Allowing categorization of patients as either high- or low-risk for

non-major surfactured in government of the developed to analogic point of claim terms and allowing categorization of patients as either high- or low-risk for bleeding based on their total score.
Despite acceptable performa bespite 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 selecti Insufficient validation of risk factor data extracted from the electronic health records (EHRs),
I variable selection, exclusion of several important potential risk factors, and inclusion in the
e of non-major bleeds, whic 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 modificatic
of VTE prophyla 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 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 validation studies has been poor.^{13,14} In the largest validation study, approximately 20% of patients we
categorized as high risk, sensitivity was less than 40%, and the area under the receiver operating
characteristics validation studies has been poor.^{25,24} 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
characteristic characteristics curve (AUROC) was 0.63 .¹³ characteristics curve (AUROC) was 0.63.13

A more accurate and generalizable accurate and the technical terms accurate and serveral particles.

An in medical inpatients. We identified major in-hospital bleeds and derived a RAM based on

an 30,000 patients from 10 h more than 30,000 patients from 10 hospitals within the Cleveland Clinic Health System (CCHS). We the
validated our RAM on a holdout validation set of nearly 14,000 patients and compared its discriminat
and calibration with more than 30,000 patients from 10 hospitals within the Cleveland China 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 validated our RAM on a hold our RAM on a hold of nearly 14,000 patients and compared its discrimination
and calibration with the IMRPOVE model.
Setting and Patients
Our cohort consisted of medical inpatients who were admit

Methods

Setting and Patients

Methods
Setting and Patients
Our cohort consisted of medical ir
between October 1, 2017, and January 31, October 1, 2017, and January 31, 2020. Hospitals were located in Ohio and Florida and var
om a 126-bed community hospital to a 1,400-bed quaternary care academic medical center
I patients aged 18 years or older who were ad between 2018 by 2019, and January 20, 2020. Hospital were located in 2018 but and January and January in Size from a 126-bed community hospital to a 1,400-bed quaternary care academic medical center. W
included patients ag 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 ble 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 e the emergency room (ER). Because its embedded as present-on-admission or occurred within
pours of admission. We also excluded patients taking warfarin because they would not be eligible t
receive chemoprophylaxis. The stud patients 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).
Iden hours of a discusser of an admission. 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 medi

Identification of Bleeds

receive chemoprophylaxis. The study was approved by the Clinical Clinic Institution of Bleeds
(IRB #14-240).
We developed an algorithm to identify major in-hospital bleeds in medical inpatients. We used
the International S (Internation
Identification
The Internation
Net and either 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 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
anatomic over and either and either face interest and the either the following: (a) fall in the following or a critical
(b) documented transfusion of at least 2 units of packed red blood cells, or (c) involvement of a critical
anat (b) documential 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 use anatomical site (e.g., intracranial, periodicial, intracrement site comparative syndroms,
retroperitoneal).¹⁵ Additionally, the bleed could not be perioperative and must have result
expected blood loss. We used a combina retroperitoneal).²⁵ 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 identi[.]
expected blood loss. We u expected blood loss. We used a combination of diagnostic codes and laboratory values to identify

patients with probable bleeds based on the charts 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 F bleeding code, bleed during admission, or major surgical operation without a specific diagnostic co
other bleeding.
Risk Factor Identification
We identified potential risk factors for major bleeding from the IMPROVE¹²

Risk Factor Identification

bleeding code, bleed during admitted on the grand operation, or major surgical operation, and the promotion without diagnostic code for major surgical operation.
We identified potential risk factors for major bleeding from **Risk Factor Iden
We iden
We iden
published studie
international no** We identified potential risk factors for major bleeding from the IMPROVE²² model and other
published studies.^{16–20} We considered age, sex, race, smoking status, alcohol use, platelet count,
international normalized rat published studies.²⁰ ²⁰ We considered age, sex, race, smoking status, alcohol use, platelet count,
international normalized ratio (INR), partial thromboplastin time (PTT), glomerular filtration rate
prior bleeding with international normalized ratio (INT), partial international ratio (PTT), glocal ratio and antisomer (ICU)
prior bleeding within 3 months of admission, rheumatic disease, sepsis, intensive care unit (ICU)
admission within 2 prior bleeding within 24 hours of hospital admission, central venous catheter (CVC) placement (incluster) admission within 24 hours of hospital admission, central venous catheter (CVC) placement (incluster) peripherally in 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 fai hypertension, heart failure, pericarditis, diabetic retinopathy, hospital readmission, and myocard
infarction. All variables were extracted from the Cleveland Clinic EHR system.
Smoking status and alcohol use were categori

hypertension. All variables were extracted from the Cleveland Clinic EHR system.

Smoking status and alcohol use were categorized as current, former, or never. Platelet coun

(thousands of platelets per uL), INR, and PTT (Smoking status and alcohol use were categorized as current, former, (
(thousands of platelets per uL), INR, and PTT (seconds) values were not modif
(umol/L) was used to calculate GFR (mL/min) using the revised chronic kidn (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 a
and sex (CKD-EPI_{AS}) e (umol/L) was used to calculate GFR (mL/min) using the revised chronic kidney disease-epidemiology
and sex (CKD-EPI_{AS}) equation.²¹ Antiplatelet medications (aspirin, clopidogrel, ticagrelor, or prasrugr
were categorize (umor, graduated chronic calculations) and sex (CKD-EPI_{AS}) equation.²¹ Antiplatelet medications (aspirin, clopidogrel, ticagrelor, or prasrugrel)
were categorized as yes/no. Readmission was defined as having a previou and sex (CKD-EPI_{AS}) equation.²¹ Antiplatelet medications (aspirin, clopidogrel, ticagrelor, or prasrugrel)
were categorized as yes/no. Readmission was defined as having a previous inpatient discharge within 3
days of a 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 incr

days of administration. The promotion.
Codes (see Supplementary Information).
Because many patients received chemoprophylaxis, which is hypothesized to increase the risk of
major bleeding,¹⁷ we identified VTE prophylaxis Because many patients received
major bleeding,¹⁷ we identified VTE prop
enoxaparin 40mg daily, dalteparin 5,000 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 major bleeding,²⁷ 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
noxaparin 40mg daily, enoxaparin 40mg daily, dalteparin 5,000 units daily, or fondaparinux 2.5mg daily) from the table of

mediations administered. To assess the impact of prophyland contract parameter estimates, we
performed a sensitivity analysis, refitting a multiple logistic regression model with and without pati
who received prophylaxis.

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. Data extracted from
a random sample of patien
outcome, we required that
Model Development and 1 Data extracted from the EHR are often unreliable,²² thus we performed iterative chart review of
m sample of patient charts for each predictor. To ensure that each risk factor preceded the
e, we required that all predicto

Model Development and Training

a random sample of the first 24 hours of admission.
A random sample of patient charts for interestion.
A random sample imputation to fill in missing values for INR and PTT based on the
distribution of values in patients wi **Model Development and Training**
We first performed single imputation to fill in missing values for INR and PTT b.
distribution of values in patients with normal results. We did not perform multiple imp
these data were not Mestribution of values in patients with normal results. We did not perform multiple imputation becatives data were not missing at random. Unlike platelet count and serum creatinine (used to calculated GFR), INR and PTT are distribution of values in patients with normal results. The maintenance performanciple imputation states these data were not missing at random. Unlike platelet count and serum creatinine (used to calculate GFR), INR and PT these data were data patients unless there are specific indicat
symptoms or suspected bleeding or coagulation disorders. Thus, we assumed that patients with missin
INR and PTT values had values within normal range.
We spli GER), Interactional exercises the respected 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 vali

symptoms or suspected accuracy or coagulation disorders. Thus, we accurate that patients with missing (NR and PTT values had values within normal range.
We split the cohort into training (70%) and validation (30%) sets wit INR and We split the cohort into training (70%) and v
bleeding events to ensure proportional event rates if
and accuracy in the presence of a highly unbalanced
perform variable selection in two steps. First, we use Explicit the consure proportional event rates in each set. We did this to ensure model trainability
unacy in the presence of a highly unbalanced binary outcome. We used the training set to
variable selection in two steps. bleeding events to ensure proportional event rates in each set. The limitiate model in an interesting set to perform variable selection in two steps. First, we used univariate analyses (t-tests and chi-squared tests for co perform variable selection in two steps. First, we used univariate analyses (t-tests and chi-squared
for continuous and categorical variables, respectively) and collinearity assessments among the ful
predictors to exclude performation variable selection in the stepse integral manufacter analyses (t-tests and an equation is the foll
for continuous and categorical variables, respectively) and collinearity assessments among the full set of
pre 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 minimi predict 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 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 predictir
major bleeding we order to 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
variables we major bleeding were penalized based on their amount of shrinkage. Cross-validation was used to determined was
variables were regularized based on their amount of shrinkage. Cross-validation was used to determined was used variables were regularized based on their amount of shrinkage. Cross-validation was used to determine

the penalty term, lambar, and to compute the regularized community computers of entities the responsibility we
were exponentiated to yield odds ratios (ORs).
All analyses were performed in R (v4.2.2).²⁴ LASSO regression All analyses were performed in R (v4.2.
package glmnet (v4.1-6).²⁵ ln R, glmnet automa
LASSO model, thus we did not standardize any
count by a factor of 10 to reduce variability in t All analyses were performed in R (v4.2.2).²⁴ LASSO regression modeling was performed using the glmnet (v4.1-6).²⁵ ln R, glmnet automatically centers and scales all data prior to fitting the nodel, thus we did not stan 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 platele
count by a factor of LASSO model fit on the training set was used to generate a predictors (since most we

the standard coded as 0 or 1) and to facilitate easier interpretation of odds ratios.
 Internal Validation

The LASSO model fit on the

Internal Validation

country a factor of 10 to reduce variable, in the range of scales a country predictor (since most were between
binary and coded as 0 or 1) and to facilitate easier interpretation of odds ratios.
Internal Validation
The LAS Internal Validation
The LASSO model fit on the training set was used to generate a predicted
for each patient in the validation set. Model calibration was assessed using a cali
discrimination was assessed using a receiver The LASS Decision processed using a 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), sensiti discrimination was assessed using a receiver operating characteristics (ROC) curve, the area under th
ROC curve (AUROC), sensitivity, specificity, and the F1 score (harmonic mean). The optimal cut-off fo
predicting bleedin 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 u Rocketing 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 wel predictionary and the optimism-adjusted AUC and 95% confidence interval (CI) at the potimal cut-off, as well as to plot 95% confidence bands for the ROC curve at a range of sensitivity and corresponding specificity values. Bootstrapping was used to compute 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 IMPRO

External Validation and Comparison with IMPROVE

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 External Validation and Compari
We sought to validate the
of our LASSO model (hereafter Cl
this, we fit the IMPROVE model o 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 t 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 this, we fit the IMPROVE model on our validation set using multiple logistic regression¹⁴ and generated predictions in the validation set. We then compared CCBM with IMPROVE on the basis of calibration and discrimination predistant in the validation set. Specifically, we compared models on bleed detection and high-vs.
Iow-risk categorization rates using chi-squared tests, and on AUC using DeLong's test for paired ROC
curves.²⁶ discrimination in the validation set. Specifically, we compared models on bleed detection and high- valid
low-risk categorization rates using chi-squared tests, and on AUC using DeLong's test for paired ROC
curves.²⁶ $\sum_{i=1}^{\infty}$ curves.²⁶ $curves.²⁶$

COM COM THAT TO COMPTATE EXPERIMMENT OF THE MULTER IN MULTER PRESENT AND THE MORE TO COMPARE THE MORE TO COMP

COMPARE OF THE CONSENT COMPARE compared F1 scores as well as high- vs. low-risk categorization rates between the models using a c
square test. Statistical significance was defined as two-sided $p < .05$.
Comparing CCBM with Clinician Assessment
The pur

Comparing CCBM with Clinician Assessment

Comparing CCBM with Clinician Assessment
The purpose of developing this RAM was to design a tool tha
for VTE and bleeding risk in hospitalized patients. We empirically asse
prophylaxis use in patients stratified by bleedin The purpose of developing the basis of developing and an extended and anti-material meaning.
The prophylaxis use in patients stratified by bleeding risk to determine how well clinicians incorporate
bleeding risk assessment 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 prophylaxis use in patients stratified the stratified the percentage. We divided patients into de
based on bleeding risk computed by CCBM and quantified the percentage of eligible patients who
received prophylaxis in each 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 clinician
considered to be at hig based on bleeding and the perceived prophylaxis in each decile. Patients ineligible for chemoprophylaxis were those whom clime
considered to be at high risk of bleeding and were excluded from this analysis.
Results
We iden

Results

received 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 o **Results**
We identified 48,030 adult medical inpatients with valid bleeding data
After excluding patients with one or more missing predictors that could not be
reason for missing values could not be determined, our final c cluding patients with one or more missing predictors that could not be imputed because the
or missing values could not be determined, our final cohort contained 46,314 patients. The
set included 32,419 patients (70%) and t Franching patients with one or 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 raining set included 32,419 patients (70%) and the validation set included 13,895 patients (30%); each and a major bleed rate of 0.58%. Compared with patients without major bleeding, patients with ma
bleeds were older (mea training set included 22,424 patient (82,4) minimized and the validation set included 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. heeds 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 bleeds were older were older were older (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 the 232,000 per uL) and worse renal function (mean CFR 61.5 cm m.), and cancer (28% vs. 15.4%). All patient
have prior bleeding (25.7% vs. 5.9%), sepsis (50% vs. 14.2%), and cancer (28% vs. 15.4%). All patient
characteristics

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 OR:
generated in the training set. The strongest predi The final LASSO model included 16 predictors; see Table 2 for predictions are shown in Table 1.
generated in the training set. The strongest predictors of major bleeding were Example 16 produced in the training set. The strongest predictors of major bleeding were peptic ulcer (OR = 5.90)
and the training set. The strongest predictors of major bleeding were peptic ulcer (OR = 5.90) generated in the strongest predictors of major bleeding were perturbated were perturbated were perturbated were perturbated were perturbated with α prior bleed (OR = 4.24), and separation since contribute the september (OR = 4.24). In the contributed as the Youden index (0.68%),
CCBM was 0.86 (95% Cl, 0.81-0.91). At the optimal cut-off, defined as the Youden index (0. Sensitivity was 79%, specificity was 84%, and the F1 score was 0.055. Calibration was good (Figure
Predicted risk of major bleeding among all patients in the validation set ranged from 0.017% to 5
with a mean of 0.58% (SD sensitivity was 1949 producted 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 f

Sensitivity Analysis

Predicted risk of major bleeding anong and the validation to the validation to the validation of 0.58% (SD = 1.75%).

Sensitivity Analysis

In addition to fitting the LASSO model, we fit multiple logistic regression models Sensitivity Analysis

In addition to fitting the LAS

cohort with and without patients wh

model parameter estimates. Peptic 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 condition and with a meta-patition with the interpretent of prophylamic of the impact of prophylaxis (model parameter estimates. Peptic ulcer's effect was significantly reduced in the model excluding
patients who received model parameter estimates. Perfect was significantly reduced in the model (OR = 5.81, 95
patients who received prophylaxis (OR = 3.06) compared with the full cohort model (OR = 5.81, 95
3.99, 8.26), but all other variables patients who received the received prophylaxis
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 3.99, 3.99, 8.26), but all patients with complete data in the final LASSO model, regardless of prophylaxis
1.99, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8.26, 8
1.99,

External Validation of IMRPOVE

receipt.
Figures with complete the final patients with complete data in the final LASSO model, regardless of prophylaxis
The optimism-adjusted AUC for IMPROVE was 0.72 (95% CI, 0.66-0.78). Using the IMPROVE
off of ≥7 poin External
off of ≥7
generall

Comparing CCBM with IMPROVE

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

sensi 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 (definerally 27 ty and corresponding specificity, including at the optimal cut-off point for IMPROVE (def
s), CCBM showed better discrimination than IMPROVE. The overall difference was robus
gnificantly higher area under the curve (AUC) c \ge 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 And a significantly higher area under the curve (AUC) compared with IMPROVE (0.86 vs. 0.72, Z = 4.20, p
 \le .001; 95% CI for the difference between AUCs = 0.08-0.21). $k = 0.001$; 95% CI for the difference between AUCs = 0.08-0.21). < .001; 95% CI for the difference between AUCs = 0.08-0.21).

Next, we matched CCBM and IMPROVE identified 44/81 (54%) of bleeds using the
next categorization rates. In our validation set, IMPROVE identified 44/81 (54%) of bleeds using the
cut-off of \geq 7 points. This equated to a cut-off of ≥7 points. This equated to a risk cutoff of 1.45% using CCBM. At a sensitivity of 54%, IMPROVI
categorized 1,676/13,895 (12.1%) of patients as high-risk, whereas CCBM categorized 950/13,895 (6.8%)
patients as hi 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-ris 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-ri categorized as high-risk was 4.28% for CCBM and 3.35% for IMPROVE. A chi-square test comparing
of high-risk vs. low-risk categorization rates between models showed that at a sensitivity of 54%, C
categorized significantly 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%, χ^2 = 221.7, p < .0 categorized significantly fewer patients as high-risk compared with IMPROVE (6.8% vs. 12.1%, χ^2 = 221.7
p < .001).
We also matched CCBM and IMPROVE on the basis of number categorized as high-risk (12.1%)
and compared

categorized significantly fewer patients as high-risk compared with IMPROVE (6.8% vs. 12.1%, χ^2
p < .001).
We also matched CCBM and IMPROVE on the basis of number categorized as high-risk (2
and compared sensitivity. = 221.7
|
|
|
|
|
|
|
|
|
|
| p

W

and comp

detected

Comparin pared sensitivity. A chi-square test comparing number of bleeds detected showed that CCBM
d significantly more bleeds than IMPROVE (70% vs. 54%, χ^2 = 4.44, p = .035).
ing CCBM with Clinician Assessment
A total of 31

detected significantly more bleeds than IMPROVE (70% vs. 54%, χ^2 = 4.44, p = .035).
 Comparing CCBM with Clinician Assessment

A total of 319 patients in the validation set were ineligible for chemoprophylaxis due t detected significantly more bleeds than IMPROVE (70% vs. 54%, χ²
 Comparing CCBM with Clinician Assessment

A total of 319 patients in the validation set were ineligible

assessment of elevated bleeding risk. Among el for chemoprophyl
mean risk for major
percentage of eli A total of 319 patients in the validational sessessment of elevated bleeding risk. Among
decile, as computed by CCBM, is shown in Fig
received VTE prophylaxis in each decile is also 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 proph assessment of electromation and any one process, the mean risk for myanisation of patients.
The percentage of eligible patients
received VTE prophylaxis in each decile is also shown in Figure 3 (right y-axis). Mean predict 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 received by decile, with tenth decile patients showing substantially increased predicted risk comp
with patients in the first nine deciles. The percentage of patients who received prophylaxis also
increased by decile, with 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 Find the first nine decile, with the highest risk patients (tenth decile) more likely to have received pro
than the lowest risk patients (first decile).
Discussion
In this retrospective cohort study, we developed and valid

Discussion

increased by decide, with the highest risk patients (tends decide) more likely to the highest risk patients (first decile).

Discussion

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

model, CCBM, CCBM, CCBM, CCBM, CCBM, CONTERT AND METATRON CONTEQ 17 produces calibration and discrimination in the validation set. The CCBM outperformed the IMPROVE¹² model.
Which has been the only validated bleeding RAM calibration and discrimination in the validation set. The CCBM outperformed the IMPROVE⁴⁴ model,
which has been the only validated bleeding RAM for more than 10 years. When matching CCBM and
IMPROVE on sensitivity, CCBM IMPROVE on sensitivity, CCBM categorized 44% fewer patients as high risk than IMPROVE, and high r
patients identified by CCBM had a 28% higher mean major bleeding risk. Thus, if we focus on sensitiv
use of CCBM would preve patients identify the control 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 t use of CCBM showed better discrimination than
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

prophylaxis MPROVE 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 facto IMPROVE in several important ways. First
Dur RAM differs from IMPROVE in several important ways. First
patients, which allowed us to identify more risk factors that were inde
bleeding and thus improve predictive performanc 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 ble
whereas IMPROVE included non bleeding and thus improve predictive performance. Second, we restricted the outcome to major
whereas IMPROVE included nonmajor bleeds, which decreases the clinical relevance of prediction
because patients at elevated risk 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 val whereas IMPROVE validated outcomes and all predictors by using both ICD codes and charge prophylaxis.²⁷ Third, we validated both outcomes and all predictors by using both ICD codes and charge prophylaxis.²⁷ Third, we v prophylaxis.²⁷ Third, we validated both outcomes and all predictors by using both ICD codes and
review procedures, whereas IMPROVE validated outcomes but relied only on ICD codes to ident
factors.
Finally, by using a sup 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,

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 valida achieved
future v:
also able 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 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 also able to identify several additional important risk factors for major bleeding that IMPROVE did no
include. Moreover, the IMPROVE model utilizes logistic regression, in which each variable is given the
same weight in p and and to identify several additional important risk factors of major additional interval additional include.

Same weight in predicting the outcome. However, LASSO penalizes less important variables, which

reduces overf 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 same weight in predicting the chromatic transformation of permission in permission in prediction set.²⁸ This is important because characteristics vary by region; in our CCHS cohort, peptic ulcers and prior bleeding were reduces overfitting and produces less biased predictions in the validation set.²⁹ I his is important
because characteristics vary by region; in our CCHS cohort, peptic ulcers and prior bleeding were
reduces overfacterist because characteristics vary by region; in our CCHS cohort, peptic ulcers and prior bleeding were the

stronger predictors of major and 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 Guidelines from
performed in conjunctic
because patients at incr
assessing bleeding risk i performed in conjunction with VTE risk assessment to support VTE prophylaxis decision-making
because patients at increased bleeding risk should not receive pharmacological prophylaxis. Bu
assessing bleeding risk is challen performed in conjunction with VTE risk assessment to support VTE prophylaxis decision-making,^{4,3}
because patients at increased bleeding risk should not receive pharmacological prophylaxis. But
assessing bleeding risk is because patients at increases patients at the greatest risk for major bleeding were more likely to receive prophylaxis than the prophylaxis findings indicate the need for an accurate RAM for major bleeding. The with the lo assessing bleeding state state in an emission trequently misclassify patently, at site stately,
patients with the greatest risk for major bleeding were more likely to receive prophylaxis than th
with the lowest risk. These patit 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 those
demonstrates better perf WIMPROVE model has been the only validated RAM since 2011, but we have developed a novel RAM demonstrates better performance. This RAM may be used in conjunction with readily available VTE calculators to identify patients IMPROVED HAS BEEN THE ONLY VALIDATE MULTIME IN EXTEMPTED A NOTE TISK
demonstrates better performance. This RAM may be used in conjunction with readily available VTE risk
calculators to identify patients at high bleeding ri 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 b

pharmacological prophylaxis for patients who are at elevated risk of both VTE and bleedi
Our study makes important contributions but should be evaluated within the cor
limitations. First, although we validated both the out pharmacological prophylaxis in patients who are at electronical number of and are an electronics.
I imitations. First, although we validated both the outcomes and predictors via chart review, if
that we missed some bleedin Iimitations. First, although we validated both the outcomes and predictors via chart review, it is pos
that we missed some bleeding events or misclassified some of the risk factors due to the imperfect
nature of EHR data. In that we missed some bleeding events or misclassified some of the risk factors due to the imperfect
Inature of EHR data. Due to the relatively small number of bleeds, it was possible to manually confirm
Individual outcom that we missed some bleeding some blunch some bleeds, it was possible to manually confireed
hature of EHR data. Due to the relatively small number of bleeds, it was possible to manually confir
each individual outcome. Howe 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 each individual outcome. However, with outcomes the patients, we confident all variables subset contients and performed iterative chart review until we were confident all variables were as accurate possible. Second, we had 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 wa 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 possible. Second, we had a small number of outcomes to develop and test our model. The small number of our coho
major bleeding rate was 0.58%, less than the total bleeding rate in the IMPROVE cohort and externa
validation 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 mode
performance.²⁹ validation studies, which included nonline you be coming a highly imbalanced outcome can pose
challenges for prediction models and limits the utility of certain evaluation metrics in assessing mode
performance.²⁹ performance.²⁹
prediction models and limits the utility of certain evaluation metrics in assessment metric in a
discrete in a sessing model with the utility of certain evaluation metric in a sessing model with the utilit performance.29

This way included patients who received prophylaxis. Only peptic ulcer's effect was meaningfully different, but this risk factor was
eived prophylaxis. Only peptic ulcer's effect was meaningfully different, but this risk f analysis 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 who received prophylaxis. Only peptic and the condensation grant, and the correct was uncommon in our cohort and thus unlikely to have driven the predictive accuracy of the model. Finally, we did not perform an external va uncommon in our cohort and thus unlikely to the enterprise accuracy of the models analy,
we did not perform an external validation of our RAM. By splitting our cohort into training and
validation sets and adjusting AUC for 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 h
system, therefore follow-up studi evaluation of model performance. However, the patients in our cohort all came from a single he
system, therefore follow-up studies will be needed to validate our RAM in patient populations w
different characteristics and d extem, 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 recomm

Conclusion

system, the state follow-up studies will be needed to validate such a main patient population.

Similar different characteristics and distributions of the outcome and risk factors.

Similar to validate our RAM in patient p Conclusion
The ACCP and ASH guidelines recommend incorporating bleeding r
making for VTE prophylaxis, but since 2011 only one bleeding RAM—IMPR
Using more than 45,000 patient records from 10 hospitals within a single h making for VTE prophylaxis, but since 2012 only one are sample from that the the section and also Using more than 45,000 patient records from 10 hospitals within a single health system, we develope
novel RAM that had good Internal 45,000 patients from 10 hospital movel 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 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 integra IMPROVER IN THE CONDITIONS INTO A PARTICLE TO PERFORM PROVIDENT MARKED INTO THE USE INTO A DISTURBAND INTO THE
INCREASING THE INTERNATION OF THE MARKED MOVED IN A SERVING THE USE OF THE USE OF THE MARKED OF EXAMINE THE US
 examine the impact of EHR integration on prophylaxis use, bleeding rates, and other patient outo
examine the impact of EHR integration on prophylaxis use, bleeding rates, and other patient outons. examine the impact of EHR integration on prophylaxis use, bleeding rates, and other patient outcomes.

It is made available under a CC-BY-ND 4.0 International license. **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: [https://doi.org/10.1101/2023.04.29.23289304;](https://doi.org/10.1101/2023.04.29.23289304) this version posted May 2, 2023. The copyright holder for this preprint

- 1. Beck
2. Lilly
2. Lilly
- 1. Beckman Man Man Man 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 Rist
3. Kakkar AK, Cimminiel
- Concern. *Am J Prev Med*. 2010;38(4):S495-S501. doi:10.1016/J.AMEPRE.2009.12.017
Lilly CM, Liu X, Badawi O, Franey CS, Zuckerman IH. Thrombosis Prophylaxis and Mort
Among Critically III Adults. *Chest*. 2014;146(1):51-57. 2. Lim, 2014, 2014, 2014, 2014, 146(1): 51-57. doi: 10.1378/CHEST. 13-2160

2. Kakkar AK, Cimminiello C, Goldhaber SZ, Parakh R, Wang C, Bergmann JF. Low-Molecular-Weight

2011;365(26): 2463-2472. Among Critically III Addits. Chest. 2014,146(1):51-57. doi:10.1378/CHEST.13-2166
Kakkar AK, Cimminiello C, Goldhaber SZ, Parakh R, Wang C, Bergmann JF. Low-Mo
Heparin and Mortality in Acutely III Medical Patients. *New Eng* 3. Kakkar Ak, Cimministry, Cimministry, Ak, Cimmin, Aka, Cimministry, Aka, Cimministry, Aka, Cimministry, Ak,
2011;365(26):2463-2472.
doi:10.1056/NEJMOA1111288/SUPPL_FILE/NEJMOA1111288_DISCLOSURES.PDF
4. Kearon C, Ak! EA,

- Heparin and Mortality in Acutely in Micultary attents. New England Soundry J Medicine.
2011;365(26):2463-2472.
doi:10.1056/NEJMOA1111288/SUPPL_FILE/NEJMOA1111288_DISCLOSURES.PDF
Kearon C, Akl EA, Comerota AJ, et al. Antith 2011;365(26):2463-2472. Rearon C, Akl EA, Comerota AJ, et al. Antithrombotic therapy for VTE disease: Ant
therapy and prevention of thrombosis, 9th ed: American College of Chest Physicia
based clinical practice guidelines. *Chest*. 2012;141(2 SUP 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-38B2-46 based clinical practice guidelines. Chest. 2012,141(2 SUPPL.).e4155-e4505.
doi:10.1378/CHEST.11-2301/ATTACHMENT/7547E95F-38B2-4625-A9BB-
CE44DDD34D42/MMC1.PDF
Schünemann HJ, Cushman M, Burnett AE, et al. American Society o
- Schünemann HJ, Cushman M
for management of venous tl
medical patients. *Blood Adv*.
Klok FA, Huisman M V. How I 5. Schwammann, Journalism, Journalism, Journalism Level, Thematology 2014 guidelines

for management of venous thromboembolism: prophylaxis for hospitalized and nonhospitalized

medical patients. *Blood Adv.* 2018;2(22):31
- medical patients. *Blood Adv.* 2018;2(22):3198-3225. doi:10.1182/BLOODADVANCES.2018022954
Klok FA, Huisman M V. How I assess and manage the risk of bleeding in patients treated for
venous thromboembolism. *Blood*. 2020;135
- medical patients. *Blood Adv. 2018*;2(22):3158-5225. doi:10.1182/BLOODADVANCES.20180225344
Klok FA, Huisman M V. How I assess and manage the risk of bleeding in patients treated for
venous thromboembolism. *Blood*. 2020;13 6. Kload Fa, Huisman M V. How I and the risk Fastess and the risk of the risk of tensors.

6. Klok Fa, Spyropoulos AC, Chopra V, et al. Validation of Risk Assessment Models of Venors

6. Thromboembolism in Hospitalized Med venous thromboembolism. Blood. 2020,135(10):724-734. doi:10.1102/BLOOD.2015001005
Greene MT, Spyropoulos AC, Chopra V, et al. Validation of Risk Assessment Models of Venot
Thromboembolism in Hospitalized Medical Patients. 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
hospita
- Thromboembolism in Hospitalized Medical Patients. Am J Med. 2010;129(9):1001.e9-1001.e18.
doi:10.1016/J.AMJMED.2016.03.031
Barbar S, Noventa F, Rossetto V, et al. A risk assessment model for the identification of
hospitali Barbar S, Noventa F, Rossetto V, et al
hospitalized medical patients at risk f hospitalized medical patients at risk for venous thromboembolism: the Padua Prediction of the identities of the interval for the interval 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):1538-
7836.2010.04044.X
Spyropoulos AC, Anderson FA, FitzGerald G, et al. Predictive and Associative Models
Hospitalized Medical Patients at Risk fo Spyropoulos AC, An
Hospitalized Medica
doi:10.1378/CHEST
Nendaz M, Spirk D, I 9. 9. Spyropoulos Action Fa, Anderson Fa, FitzGerald G, 2011;140(3):706-714.
Hospitalized Medical Patients at Risk for VTE. *Chest*. 2011;140(3):706-714.
10. Nendaz M, Spirk D, Kucher N, et al. Multicentre validation of the G
- Hospitalized Medical Patients at Risk for VTE. Chest. 2011,140(3):706-714.
doi:10.1378/CHEST.10-1944
Nendaz M, Spirk D, Kucher N, et al. Multicentre validation of the Geneva Ri
hospitalised medical patients at risk of veno Nendaz M, Spirk D, Kucher N
hospitalised medical patients
Thromboembolic RIsk and Pr
Haemost. 2014;111(3):531-5 11. Nendal M, Spirk D, Spirk and Prophylaxis for Medical PATients in SwitzErland (ESTIMATE).

11. Stuck AK, Spirk D, Schaudt J, Kucher N. Risk assessment models for venous thromboem

11. Stuck AK, Spirk D, Schaudt J, Kuche Thromboembolic Rlsk and Prophylaxis for Medical PATients in SwitzErland (ESTIMATE). The
Haemost. 2014;111(3):531-538. doi:10.1160/TH13-05-0427/ID/JR0427-4
Stuck AK, Spirk D, Schaudt J, Kucher N. Risk assessment models for
- 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
Stuck AK, Spirk D, Schaudt J, Kucher N. Risk assessment models f Haemost. 2014,111(3):331-338. doi:10.1160/1113-63-0427/ID/JR0427-4
Stuck AK, Spirk D, Schaudt J, Kucher N. Risk assessment models for venou
acutely ill medical patients: A systematic review. Thromb Haemost. 2017;
doi:10.11 11. Stuck Action Actions: A Systematic review. Thromb Haemost. 2017;117(4):801-808.
12. Decousus H, Tapson VF, Bergmann JF, et al. Factors at Admission Associated With Bleeding Risk in Medical Patients: Findings From the I
- acutely in medical patients: A systematic review. Thromb Haemost. 2017;117(4):801-808.
doi:10.1160/TH16-08-0631/ID/JR0631-6
Decousus H, Tapson VF, Bergmann JF, et al. Factors at Admission Associated With Bleeding
Medical P Decousus H, Tapson VF, Bergmann JF, et
Medical Patients: Findings From the IMPI
doi:10.1378/CHEST.09-3081
Rosenberg DJ, Press A, Fishbein J, et al. E. 12. Medical Patients: Findings From the IMPROVE Investigators. Chest. 2011;139(1):69-79.

13. Rosenberg DJ, Press A, Fishbein J, et al. External validation of the IMPROVE bleeding Risk

13. Rosenberg DJ, Press A, Fishbein
- Medical Patients: Findings From the IMPROVE Investigators. Chest. 2011,139(1):69-79.

Rosenberg DJ, Press A, Fishbein J, et al. External validation of the IMPROVE bleeding Ris

Assessment Model in medical patients. *Thromb* Rosenberg DJ, Press A, Fishber
Assessment Model in medica
doi:10.1160/TH16-01-0003/I
Hostler DC, Marx ES, Moores 13. Assessment Model in medical patients. Thromb Haemost. 2016;116(3):530-536.
13. doi:10.1160/TH16-01-0003/ID/JR0003-8
14. Hostler DC, Marx ES, Moores LK, et al. Validation of the International Medical Prevention
14. On V
- Assessment Model in medical patients. Thromb Haemost. 2010,110(3):530-536.
doi:10.1160/TH16-01-0003/ID/JR0003-8
Hostler DC, Marx ES, Moores LK, et al. Validation of the International Medical Pre
on Venous Thromboembolism B Hostler DC, Marx ES, Moores LK, et al. Va
on Venous Thromboembolism Bleeding F
doi:10.1378/CHEST.14-2842
Schulman S, Kearon C. Definition of majo 14. Thom Venous Thromboembolism Bleeding Risk Score. Chest. 2016;149(2):372-379.
14. Hostler D. Marx Essay: Moore L. Definition of major bleeding in clinical investigations of antihemostatic
15. Schulman S, Kearon C. Defin
- on venous Thromboembolism Bleeding Risk Score. Chest. 2010;149(2):372-379.
doi:10.1378/CHEST.14-2842
Schulman S, Kearon C. Definition of major bleeding in clinical investigations of ar
medicinal products in non-surgical pa schulman S, Kearon C. Defini
medicinal products in non-su
doi:10.1111/J.1538-7836.200 11. Schwalter States Communition of major bleeding in clinical investigations of antihemost

medicinal products in non-surgical patients. J Thromb Haemost. 2005;3(4):692-694.

doi:10.1111/J.1538-7836.2005.01204.X medicinal products in non-surgical patients. J Thromb Haemost. 2005;3(4):692-694.
doi:10.1111/J.1538-7836.2005.01204.X 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):1-2.
200. doi:10.1186/CC12296
Darzi AJ, Karam SG, Charide R, et al. Prognostic factors for VTE and bleeding in hospitalized
medic 200. Darzi AJ, Karam SG, Charide
medical patients: a systema
doi:10.1182/BLOOD.201900
Langsted A, Nordestgaard E 17. Darzina, minimize, Emanuela, Charlie R, et al. Prognostic Frederic Frederic metering in hospitalized
18. doi:10.1182/BLOOD.2019003603
18. Langsted A, Nordestgaard BG. Smoking is Associated with Increased Risk of Major
- medical patients: a systematic review and meta-analysis. Blood. 2020;135(20):1788-1810.
doi:10.1182/BLOOD.2019003603
Langsted A, Nordestgaard BG. Smoking is Associated with Increased Risk of Major Bleeding
Prospective Coho MARREADED, 2000 DENDED DED DES
Langsted A, Nordestgaard BG. Sm
Prospective Cohort Study. *Thromb*
1675798/ID/JR180444-10
Kirchhof P, Haas S, Amarenco P, e 2. 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
- Prospective Cohort Study. *Thromb Haemost. 2019*,119(1):39-47. doi:10:1035/3-0038-
1675798/ID/JR180444-10
Kirchhof P, Haas S, Amarenco P, et al. Impact of modifiable bleeding risk factors on mi
bleeding in patients with at Experience, the Extreme Kirchhof P, Haas S, Amarer
bleeding in patients with a
2020;9(5). doi:10.1161/JAI
Liabeuf S, Scaltieux LM, Mi 19. bleeding in patients with atrial fibrillation anticoagulated with rivaroxaban. *J Am Heart As.*
2020;9(5). doi:10.1161/JAHA.118.009530
20. Liabeuf S, Scaltieux LM, Masmoudi K, et al. Risk factors for bleeding in hospit
- bleeding in patients with atrial fibrillation anticoagulated with rivaroxaban. J Am Heart Assoc.
2020;9(5). doi:10.1161/JAHA.118.009530
Liabeuf S, Scaltieux LM, Masmoudi K, et al. Risk factors for bleeding in hospitalized 2020;9, 20202020;0, MARSHOLD 2020;
Liabeuf S, Scaltieux LM, Masmoudi K, et al
patients with an INR of 5 or more treated
2015;94(52). doi:10.1097/MD.000000000
Inker LA, Eneanya ND, Coresh J, et al. New 2015;94(52). doi:10.1097/MD.000000000002366
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
21. Inker LA, Eneanya ND, Coresh J, et a
- patients with an INR of 5 or more treated with Vitamin K antagonists. Medicine (*Ollice States)*.
2015;94(52). doi:10.1097/MD.0000000000002366
Inker LA, Eneanya ND, Coresh J, et al. New Creatinine- and Cystatin C–Based Equ 2023; 2023; 2020
Inker LA, Eneanya ND, Coresh J, et al. New Creatinin
Estimate GFR without Race. *New England Journal o*
doi:10.1056/NEJMOA2102953/SUPPL_FILE/NEJMO
Taksler GB, Dalton JE, Perzynski AT, et al. Opportun 22. Estimate GFR without Race. New England Journal of Medicine. 2021;385(19):1737-1749.
22. Taksler GB, Dalton JE, Perzynski AT, et al. Opportunities, Pitfalls, and Alternatives in Adapt
22. Taksler GB, Dalton JE, Perzynsk
- Estimate GFR without Race. New England Journal by Medicine. 2021,385(15).1737-1745.
doi:10.1056/NEJMOA2102953/SUPPL_FILE/NEJMOA2102953_DISCLOSURES.PDF
Taksler GB, Dalton JE, Perzynski AT, et al. Opportunities, Pitfalls, an Taksler GB, Dalton JE, Perzynski AT, et al. Opportunities, Pitfalls, and Alternatives
Electronic Health Records for Health Services Research. *Med Decis Making*. 2020;
doi:10.1177/0272989X20954403
Tibshirani R. Regression 2220;41(2):133-14 Electronic Health Records for Health Services Research. *Med Decis Making*. 2020;41(2):133-14
doi:10.1177/0272989X20954403
23. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of t*
- Electronic Health Records for Health Services Research. Med Decis Midking. 2020,41(2):133-142.
doi:10.1177/0272989X20954403
Tibshirani R. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical
S Tibshirani R. Regression Shrinkage
S*ociety: Series B (Methodological,*
6161.1996.TB02080.X 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 Society: Series B (Methodological). 1996;58(1):267-288. doi:10.1111/J.2517-
6161.1996.TB02080.X
-
-
- 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
DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing th 26. 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 Cor
Receiver Operating Characteristic Curves: A Nonparametr Coordinate Descent. J Stat Softw. 2010;33(1):1. doi:10.186377jss.v033.i01
DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Tv
Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biol*
198 20. Deceiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*.
26. Deceiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*.
27. Kaatz S, Ahmad D, Spyropoulos AC, Schulman S.
- Receiver Operating Characteristic Carves: A Nonparametric Approach. Biometries:
1988;44(3):837. doi:10.2307/2531595
Kaatz S, Ahmad D, Spyropoulos AC, Schulman S. Definition of clinically relevant nor
bleeding in studies of Example 2017, 2018
Kaatz S, Ahmad D, Spyropoulos AC, Sch
bleeding in studies of anticoagulants in
non-surgical patients: Communication
Haemostasis. 2015;13(11):2119-2126. 27. Bleeding in studies of anticoagulants in atrial fibrillation and venous thromboembolic disea

27. Kaanmare S, Ahmad D, Spyropoulos Ac, Schulman S. Definition from the SSC of the ISTH. Journal of Thrombosis and

28. Pav bleeding in statict is anticoagularity in attribution and ventor anticomputed in attribution and
haemostasis. 2015;13(11):2119-2126. doi:10.1111/jth.13140
Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurat
- non-surgical patients: Communication from the SSC of the ISTH. Sournar by Thrombosis and
Haemostasis. 2015;13(11):2119-2126. doi:10.1111/jth.13140
Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk p
- Pavlou M, Ambler G, Seaman SR, et al. How to develop a more
when there are few events. *BMJ*. 2015;351. doi:10.1136/BMJ.
Chicco D, Jurman G. The advantages of the Matthews correlat
and accuracy in binary classification eva 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 sco
and accuracy in binary classification evaluation. *BMC Geno* When there are few events. BMJ. 2015,351. doi:10.1136/BMJ.15868
Chicco D, Jurman G. The advantages of the Matthews correlation coe
and accuracy in binary classification evaluation. *BMC Genomics*. 2020
doi:10.1186/S12864-0 29. and accuracy in binary classification evaluation. *BMC Genomics*. 2020;21(1):1-13.
doi:10.1186/S12864-019-6413-7/TABLES/5
Acknowledgements and accuracy in binary classification evaluation. *BMC Genomics*. 2020,21(1):1-13.
doi:10.1186/S12864-019-6413-7/TABLES/5
wledgements

$\frac{1}{1}$

doi:10.1186/S12864-019-6413-7/TABLES/5 Acknowledgements
This study was support This study was supported in part by NIH grants 5T32GM007250-45 and 5TL1TR002549-04.

	1. Patient characteristics and their relationship to major in hospital biccums.			
	Total	No Bleed	Major Bleed	
Characteristic	$(n = 46,311)^*$	$(n = 46,043)$	$(n = 268)$	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%	183%	
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%	25.7% 5.9%		< 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	187%	187%	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
	0.3%	0.35	0.7%	0.506
CNS neoplasm				
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

rate is in milliliters per minute (mL/min).
Nobreviations: CVC, central venous catheter; PICC, peripherally inserted central catheter; SSRI, selective
serotonin reuptake inhibitor; NSAID, nonsteroidal anti-inflammatory dru Abbreviations: CVC, central venous cathe
serotonin reuptake inhibitor; NSAID, non:
thromboembolism.
*3 patients included in the final LASSO m
**Antiplatelet drugs included aspirin, clo 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, clo served in the final LASSO model did not have prophylaxis data.

Internal and the final LASSO model did not have prophylaxis data.

Internal anti-inflame included aspirin, clopidogrel, ticagrelor, and prasrugrel

The same p

thromboembolism.

**Antiplatelet drugs included aspirin, clopidogrel, ticagrelor, and prasrugrel
**Antiplatelet drugs included aspirin, clopidogrel, ticagrelor, and prasrugrel $*_{\alpha}$. Antiplatelet drugs included aspirin, α , practiced as piring α , and practice α

TODIC 2. INTOITIVATIONE AND LASSO INVOLS TOT INQUITITITIOSPILAT DICEDITIS.								
	OR excluding	OR including	OR including		OR including			
	prophylaxis	prophylaxis	prophylaxis		prophylaxis, LASSO			
Characteristic	$(n = 19, 365)$	$(n = 46, 311)$	(95% CI)		$(n = 46, 314)$			
$Age*$	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 2 4			
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		\blacksquare			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)								

SSRI 1.65 1.11 0.79 1.53 1.14

NSAID - 1.53 1.14

VTE prophylaxis - 1.37 1.05 1.80 - 1.00

Abbreviations: OR, odds ratio; LASSO, least absolute shrinkage selection operator; CI, cc

interval; CVC, central venous catheter; NE prophylaxis

VTE prophylaxis

Abbreviations: OR, odds ratio; LASSO, least absolute shrinkage selection operator; CI, conterval; CVC, central venous catheter; PICC, peripherally inserted central catheter; SSRI

Serotonin VTE prophylaxis

Abbreviations: OR, odds ratio; LASSO, least absolute shrinkage selection operator; Cl, conterval; CVC, central venous catheter; PICC, peripherally inserted central catheter; SSRI

serotonin reuptake inhibi Abbreviations: OR, odds ratio; LASSO, least absolute shrinkage selection operator; C
interval; CVC, central venous catheter; PICC, peripherally inserted central catheter; S
serotonin reuptake inhibitor; NSAID, nonsteroidal 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 interval; CVCC, central ventral ventra
serotonin reuptake inhibitor; NSAID, nonsteroidal anti-inflammatory drug; VTE, venous
*Scaled down by a factor serotonin reuptake inhibitor; NEW (e.g., from years to tens of years)
*Scaled down by a factor of 10 (e.g., from years to tens of years)
*Scaled down by a factor of 10 (e.g., from years to tens of years)

thrombolism.
*Scaled down by a f $\mathcal{L}=\left\{ \begin{array}{ll} 0.6\,c & \text{if }c\in\mathcal{C}, \end{array}\right.$

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

rigure 2. Receiver operating characteristics (ROC) carve 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 validation set (n = 13,895), with bootstrapped 95% confidence bands. The optimal cut-off point for
Validation set (n = 13,895), with bootstrapped 95% confidence bands. The optimal cut-off point for
IMPROVE (≥7 points) is IMPROVE (27 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. IMPROVER CORDIG (27 points) is the red dot. Horizontal and vertical data corresponding operating points given equivalent sensitivity or specificity, respectively.
The red of the red dot. However, when the red of the red of CCBM at corresponding operating points given equivalent sensitivity or specificity, respectively.

prophylaxis decisions in patients eligible for chemoprophylaxis (right y-axis), by decile of predicted risk.

prophylaxis decisions in patients eligible for chemoprophylaxis (right y-axis), by decile of predicted risk.

Fi prophylaxis decisions in patients eligible for chemoprophylaxis (right y-axis), by decile of predicted risk.