Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Performance of the Padua Prediction Score in Critically Ill Patients With Gastrointestinal Bleeding: A 24-Hour Landmark Cohort Study Using MIMIC-IV

Yun Weng
Huai'an First People's Hospital
Miaomiao Cheng
Huai'an First People's Hospital
Mengmeng Jiang
Huai'an First People's Hospital
Ru Jiao
Huai'an First People's Hospital
Caixia Liu
Huai'an First People's Hospital
Tingting Meng
55573675@qq.com

Huai'an First People's Hospital

Research Article

Keywords: gastrointestinal bleeding, intensive care unit, Padua Prediction Score, thromboprophylaxis, venous thromboembolism, MIMIC-IV

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11244545/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Performance of the Padua Prediction Score in Critically Ill Patients With Gastrointestinal Bleeding: A 24-Hour Landmark Cohort Study Using MIMIC-IV

Yun Weng[1], Miaomiao Cheng[1], Mengmeng Jiang[1], Ru Jiao[1], Caixia Liu[1]

1 Department of Gastroenterology, Huai'an First People's Hospital

Corresponding author: Tingting Meng, Department of Gastroenterology, Huai'an First People's Hospital. Email: 905595154@qq.com

Abstract

Background

Critically ill patients with gastrointestinal bleeding may have substantial venous thromboembolism risk, but pharmacologic prophylaxis can complicate hemostatic management and raise concern for recurrent bleeding. The usefulness of the Padua Prediction Score in this setting remains uncertain.

Methods

We performed a retrospective 24-hour landmark cohort study using MIMIC-IV version 3.1. Adults with gastrointestinal bleeding among the first three discharge diagnoses, an intensive care unit admission, and a hospital stay of at least 72 hours were included. The primary endpoint was a non-primary-position venous thromboembolism diagnosis code during the index hospitalization. A stricter sensitivity endpoint additionally required treatment-dose anticoagulation initiated after the landmark. We evaluated discrimination, threshold performance, apparent calibration, multivariable associations, and exploratory overlap-weighted associations of early prophylaxis with venous thromboembolism and severe rebleeding proxies.

Results

Among 2,758 patients, 101 (3.7%) met the primary endpoint. The median age was 69 years, and 60.2% were male. The Padua score had an area under the receiver operating characteristic curve of 0.610 (95% confidence interval, 0.555–0.665). At the conventional threshold of ≥4, sensitivity was 86.1%, specificity 25.1%, positive predictive value 4.2%, and negative predictive value 97.9%. The maximum-Youden threshold was ≥6, with sensitivity of 57.4% and specificity of 62.5%. Results were similar for the strict endpoint, with 52 events and an area under the curve of 0.641 (95% confidence interval, 0.569–0.713), and among patients with complete core Padua data, with an area under the curve of 0.644 (95% confidence interval, 0.575–0.713). Exploratory weighted associations between prophylaxis and thrombotic or rebleeding proxies were imprecise and not statistically significant.

Conclusions

In critically ill patients with gastrointestinal bleeding, the conventional Padua threshold captured most venous thromboembolism-coded cases but classified three quarters of the cohort as high risk and had poor specificity. The score should not be used alone to guide prophylaxis; decisions should incorporate bleeding activity, hemostatic stability, and repeated clinical reassessment.

Keywords: gastrointestinal bleeding; intensive care unit; Padua Prediction Score; thromboprophylaxis; venous thromboembolism; MIMIC-IV

Introduction

Venous thromboembolism (VTE) is a frequent and clinically important complication of critical illness. Prolonged immobility, systemic inflammation, advanced age, malignancy, organ dysfunction, and invasive treatment all contribute to thrombotic risk in the intensive care unit (ICU). A meta-analysis of ICU studies estimated that approximately one in ten critically ill patients develops VTE, although reported rates vary according to case mix, diagnostic surveillance, and the use of thromboprophylaxis [1]. Current guidelines therefore recommend pharmacologic prophylaxis for acutely or critically ill medical patients when the risk of bleeding is acceptable, with mechanical methods used when anticoagulant prophylaxis is contraindicated [2]. In practice, however, this recommendation is difficult to apply to patients admitted with acute gastrointestinal bleeding. These patients may have several strong VTE risk factors while

simultaneously facing active hemorrhage, uncertain hemostasis, transfusion requirements, and a substantial risk of recurrent bleeding.

The balance between thrombosis and bleeding is particularly unstable during the early phase of hospitalization. Decisions regarding prophylaxis are often revisited as bleeding control, hemoglobin trends, endoscopic findings, mobility, and overall clinical status change. Gastroenterology guidelines provide detailed recommendations on the interruption and resumption of antithrombotic therapy after gastrointestinal bleeding [3,4], but they do not offer a validated VTE risk model specifically developed for critically ill patients with active bleeding. Clinicians must therefore rely on general medical risk-assessment tools together with bedside judgment. This approach may be reasonable, but its accuracy in a selected ICU gastrointestinal bleeding population has not been well established.

The Padua Prediction Score is among the most widely used tools for estimating VTE risk in hospitalized medical patients. The score assigns weighted points to 11 factors, including active cancer, previous VTE, reduced mobility, known thrombophilia, recent surgery or trauma, advanced age, cardiorespiratory failure, acute myocardial infarction or ischemic stroke, acute infection or rheumatic disease, obesity, and hormonal therapy [5]. In the original study, a score of 4 or higher identified patients considered at high risk and has since been incorporated into clinical guidance and institutional decision-support pathways [2,6]. Nevertheless, the performance of the Padua score is not uniform across settings. A large contemporary study of medical inpatients reported moderate discrimination, but accuracy differed across clinical subgroups [7]. The score may be less informative in the ICU, where reduced mobility, older age, infection, and major comorbidity are common. When many patients accumulate similar risk factors, the score distribution shifts upward and its ability to distinguish between those who will and will not experience VTE may decline.

Additional concerns arise when the Padua score is reconstructed from electronic health records. Some components, such as mobility, obesity, recent surgery, or thrombophilia, may be incompletely documented or difficult to define using structured data. Absence of a code or nursing entry does not always indicate true absence of the condition. These limitations are especially relevant in database studies and may affect both risk classification and calibration. At

the same time, electronic health records provide an opportunity to examine how the score performs in large, clinically complex populations that are underrepresented in prospective validation cohorts.

Evidence focused on gastrointestinal bleeding remains limited. In a retrospective study of 504 patients admitted with gastrointestinal bleeding, thromboembolic events occurred in 4%, while pharmacologic prophylaxis was used infrequently [8]. More recent observational data suggest that a history of gastrointestinal bleeding and thrombocytopenia reduce the likelihood of receiving pharmacologic prophylaxis, even among patients with substantial thrombotic risk [9]. These findings reflect the caution required in this population, but they do not clarify whether the Padua score can identify a clinically useful high-risk subgroup or support prophylaxis decisions after early stabilization.

We therefore evaluated the performance of the Padua score in critically ill adults with gastrointestinal bleeding using a 24-hour landmark cohort from MIMIC-IV. The primary objective was to assess discrimination, threshold characteristics, and apparent calibration for an in-hospital VTE coding endpoint. We also examined a stricter VTE proxy requiring post-landmark treatment-dose anticoagulation, assessed prespecified clinical and nursing-related associations, and described exploratory associations between early prophylaxis strategies, VTE-related outcomes, and severe rebleeding. The study was designed to determine whether the conventional Padua threshold provides meaningful risk separation in this high-bleeding-risk ICU population and to define the limits of its use in routine clinical decision-making.

Methods

Study design and data source

This retrospective cohort study was conducted using version 3.1 of the Medical Information Mart for Intensive Care IV database. MIMIC-IV contains deidentified clinical information from patients treated at Beth Israel Deaconess Medical Center between 2008 and 2022, including

hospital diagnoses, procedures, medication records, laboratory measurements, and detailed intensive care unit documentation [10,11].

The study was prepared in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology statement [12]. The MIMIC-IV database was developed with approval from the institutional review boards of Beth Israel Deaconess Medical Center and the Massachusetts Institute of Technology. Requirement for individual informed consent was waived because the database contains only deidentified information. Access to the data was obtained after completion of the required training and approval under the PhysioNet data-use agreement.

Study population and landmark definition

Hospital admissions were screened for eligibility if the patient was 18 years of age or older and had a diagnosis of acute gastrointestinal bleeding listed among the first three discharge diagnoses. Admissions were further required to include an ICU stay and a total hospital length of stay of at least 72 hours. When a patient had more than one eligible admission, only the first qualifying hospitalization was included. The first ICU stay during that hospitalization was used for analysis.

Admissions with a primary diagnosis of trauma were excluded. We also excluded patients with a documented VTE during a hospitalization that ended within 90 days before the index admission. This restriction was intended to reduce the inclusion of patients with recently diagnosed or actively treated thromboembolic disease.

The time of first ICU admission was defined as the study time origin. Baseline characteristics and prophylaxis exposure were assessed during the subsequent 24 hours. Patients who were discharged before completion of this period or who had received treatment-dose anticoagulation by the 24-hour landmark were excluded. Follow-up for study outcomes began after the landmark and continued until hospital discharge.

Reconstruction of the Padua Prediction Score

The Padua Prediction Score was reconstructed from diagnosis and procedure codes, previous hospital records, medication data, anthropometric measurements, and structured ICU nursing

assessments. The score comprised 11 established components: active cancer, previous VTE, reduced mobility, known thrombophilia, recent surgery or trauma, age of at least 70 years, heart or respiratory failure, acute myocardial infarction or ischemic stroke, acute infection or rheumatic disease, obesity, and ongoing hormonal treatment [5].

Active cancer, previous VTE, known thrombophilia, heart or respiratory failure, acute myocardial infarction or ischemic stroke, and acute infection or rheumatic disease were identified using prespecified diagnosis-code groups. Recent surgery was defined conservatively as a qualifying procedure performed during a hospitalization that ended within 30 days before the index admission. Hormonal treatment was identified from medication records available for the index hospitalization.

Obesity was defined as a body mass index of at least 30 kg/m². Body mass index was derived from the nearest available hospital measurement or calculated from height and weight recorded during the ICU stay. Reduced mobility was identified from the first Braden activity assessment recorded within 24 hours of ICU admission. Patients classified as bedfast or chairfast were considered to have reduced mobility.

The original Padua weighting scheme was retained. Active cancer, previous VTE, reduced mobility, and known thrombophilia were assigned 3 points each; recent surgery or trauma was assigned 2 points; and each remaining component was assigned 1 point. A total score of at least 4 was considered high risk according to the conventional definition [5].

A Padua component was coded as absent when no corresponding structured diagnosis, procedure, medication, or nursing record was found. Because lack of documentation may not always indicate true absence, a sensitivity analysis was performed in patients with complete information for the two core components most dependent on bedside documentation: mobility status and body mass index.

Baseline variables

Baseline laboratory measurements were collected from 6 hours before ICU admission through the end of the 24-hour landmark period. When more than one measurement was available, the

value selected according to the prespecified extraction rule was used in the analysis. Variables of interest included hemoglobin, platelet count, international normalized ratio, serum creatinine, blood urea nitrogen, and white blood cell count.

Physiological and nursing variables were obtained from records entered during the first 24 hours of ICU care. The first available Braden activity, mobility, and nutrition assessments were used. The Glasgow Coma Scale and other clinical measurements were derived from structured ICU records. Comorbidity burden was summarized using the Charlson Comorbidity Index, reconstructed from diagnosis codes recorded for the index hospitalization. Gastrointestinal bleeding was categorized as upper, lower, mixed, or unspecified according to the relevant diagnosis codes.

Thromboprophylaxis exposure

Pharmacologic and physical thromboprophylaxis administered or documented during the first 24 hours after ICU admission was assessed separately.

Pharmacologic prophylaxis included subcutaneous unfractionated heparin at doses of 5,000 or 7,500 units, enoxaparin at doses of 20–40 mg, fondaparinux at doses no greater than 2.5 mg, and dalteparin at doses no greater than 5,000 units. Medication administrations used for catheter flushing, line maintenance, dialysis, or other non-prophylactic purposes were excluded. Regimens consistent with treatment-dose anticoagulation were not classified as prophylaxis.

Physical prophylaxis was identified from structured documentation of sequential compression devices, compression sleeves, elastic stockings, or compression wraps. Documentation of a physical device indicated that the intervention had been ordered or recorded during the landmark period; it was not taken as evidence of uninterrupted use or adherence.

For descriptive analyses, patients were assigned to one of four mutually exclusive groups: physical prophylaxis only, pharmacologic prophylaxis only, combined physical and pharmacologic prophylaxis, or no documented prophylaxis.

Study outcomes

The primary outcome was the presence of a VTE diagnosis code in a non-primary discharge diagnosis position during the index hospitalization. VTE included diagnostic codes for deep vein thrombosis and pulmonary embolism. Restriction to non-primary diagnosis positions was used to reduce the likelihood of including admissions primarily caused by VTE.

Diagnosis records in MIMIC-IV do not contain the exact time at which each diagnosis was established. The primary outcome therefore represented an in-hospital VTE coding endpoint rather than a confirmed incident event occurring after the 24-hour landmark. It was not considered equivalent to a clinically adjudicated VTE diagnosis.

A stricter VTE endpoint was examined in sensitivity analysis. Patients met this definition when they had both a non-primary-position VTE diagnosis code and initiation of treatment-dose anticoagulation after the landmark. This definition was intended to improve outcome specificity, although it could not fully establish the timing or clinical certainty of the thromboembolic event.

Severe rebleeding was evaluated using a reproducible laboratory and transfusion-based proxy. The endpoint required a decrease in hemoglobin of at least 2 g/dL within a 24-hour period together with administration of at least 600 mL of packed red blood cells during a temporally adjacent 24-hour interval after the landmark. This outcome was not clinically adjudicated and was used only as an indicator of potentially important recurrent or ongoing bleeding.

Statistical analysis

Continuous variables are presented as medians with interquartile ranges, and categorical variables as frequencies with percentages. Baseline characteristics were compared between patients with and without the primary VTE endpoint using nonparametric tests for continuous variables and categorical tests as appropriate. Standardized mean differences were also calculated to describe the magnitude of between-group differences. Missingness was reported for each baseline variable.

The discriminative performance of the Padua score was quantified using the area under the receiver operating characteristic curve with 95% confidence intervals. The primary analysis evaluated the score as a continuous predictor of the VTE coding endpoint. Additional analyses

examined discrimination for the stricter VTE endpoint and in the subgroup with complete mobility and body mass index information.

Sensitivity, specificity, positive predictive value, and negative predictive value were calculated at each integer Padua threshold. Confidence intervals for these measures were estimated using the Wilson method. The conventional threshold of at least 4 points was specified before analysis. The threshold with the highest Youden index was also reported to describe the best numerical balance between sensitivity and specificity in the study cohort; it was not treated as an externally validated clinical cut-off.

Apparent calibration was examined using a logistic regression model in which the Padua score was entered as the sole predictor. Model-predicted probabilities were grouped and compared with observed endpoint proportions. Calibration was further summarized using the Brier score and the scaled Brier score. Because both model fitting and evaluation were performed in the same cohort, these estimates describe apparent rather than externally validated calibration.

A complete-case multivariable logistic regression model was fitted to assess whether the Padua score remained associated with the primary endpoint after adjustment for selected baseline characteristics. Covariates were chosen before model fitting and included sex, Charlson Comorbidity Index, hemoglobin, platelet count, serum creatinine, Braden nutrition category, physical prophylaxis, and pharmacologic prophylaxis. Creatinine was $\log_2$-transformed so that the corresponding odds ratio represented the association per doubling of the measured value.

The form of each continuous variable was assessed by comparing models with linear terms against models containing restricted cubic spline terms using likelihood-ratio tests. Variance inflation factors were examined to identify potentially important collinearity. Results are reported as adjusted odds ratios with 95% confidence intervals. The Benjamini–Hochberg procedure was used to control the false-discovery rate across coefficients included in the multivariable model.

A separate analysis examined prespecified clinical and nursing variables, including individual Padua components and bedside measures. Known thrombophilia was uncommon but showed a strong relationship with the endpoint, raising concern about sparse-data bias and quasi-separation.

Firth penalized logistic regression was therefore used for this model [13]. Multiplicity adjustment was again performed using the Benjamini–Hochberg procedure.

Associations between thromboprophylaxis and study outcomes were explored among patients with Padua scores of at least 4. Pharmacologic and physical prophylaxis were evaluated in separate propensity score models. Each model included age, sex, Charlson Comorbidity Index, Padua score, hemoglobin, platelet count, creatinine, Braden activity category, and use of the alternative prophylaxis modality.

Overlap weights were calculated from the estimated propensity scores to emphasize patients with comparable probabilities of receiving or not receiving each prophylaxis strategy [14]. Covariate balance before and after weighting was assessed using absolute standardized mean differences. Weighted logistic regression with robust standard errors was used to estimate odds ratios for the primary VTE endpoint, the stricter VTE endpoint, and the severe rebleeding proxy. These analyses were exploratory. Because treatment decisions were not randomized and relevant clinical factors such as active bleeding status and contraindications could not be fully measured, weighted estimates were interpreted as associations rather than causal effects.

All analyses were performed in R version 4.6.0. A fixed random seed of 42 was used for procedures involving random computation. Statistical tests were two-sided, and nominal P values below 0.05 were considered statistically significant. Where multiple coefficients were examined, emphasis was placed on Benjamini–Hochberg-adjusted P values. No formal sample-size calculation was conducted because all admissions meeting the eligibility criteria were included.

Results

Study cohort and baseline characteristics

Of 546,028 hospital admissions available in MIMIC-IV, 13,517 included a qualifying diagnosis of gastrointestinal bleeding. After restriction to adult patients with a hospital stay of at least 72 hours and an ICU admission, and retention of the first eligible hospitalization for each patient,

3,040 admissions remained. We subsequently excluded 37 admissions with a primary trauma diagnosis and 66 involving a recent VTE-related hospitalization. Among the 2,937 patients who reached the pre-landmark stage, 179 were excluded because they were discharged or had received treatment-dose anticoagulation before 24 hours. The final landmark cohort therefore comprised 2,758 patients (Figure 1).

The median age was 69 years (interquartile range [IQR], 57–81), and 1,660 patients (60.2%) were male. Upper gastrointestinal bleeding was the most common presentation, accounting for 1,616 cases (58.6%). The median Padua score was 5 (IQR, 4–6), and 2,077 patients (75.3%) met the conventional high-risk threshold of at least 4 points.

The primary VTE coding endpoint was present in 101 patients (3.7%). Compared with patients without the endpoint, those with the endpoint had a higher median Padua score (6 [IQR, 4–7] vs 5 [IQR, 3–6]; P<0.001) and lower Braden mobility categories, indicating greater mobility impairment (P=0.005). The distribution of demographic, laboratory, comorbidity, nursing, and prophylaxis variables is presented in Table 1.

Discrimination and threshold performance of the Padua score

The Padua score showed limited discrimination for the primary VTE coding endpoint, with an area under the receiver operating characteristic curve of 0.610 (95% confidence interval [CI], 0.555–0.665; Figure 2A and Table 2).

At the conventional threshold of at least 4 points, 87 of the 101 endpoint-positive patients were classified as high risk. This corresponded to a sensitivity of 86.1% (95% CI, 78.1%–91.6%) and a specificity of 25.1% (95% CI, 23.5%–26.8%). The positive predictive value was 4.2% (95% CI, 3.4%–5.1%), whereas the negative predictive value was 97.9% (95% CI, 96.6%–98.8%).

The threshold that maximized the Youden index was at least 6 points. At this cut-off, sensitivity decreased to 57.4% (95% CI, 47.7%–66.6%), while specificity increased to 62.5% (95% CI, 60.6%–64.3%). The corresponding positive and negative predictive values were 5.5% and 97.5%, respectively. Thus, the improvement in specificity at the higher threshold was accompanied by a marked loss of sensitivity.

The proportion of patients meeting the primary endpoint increased across Padua score categories. The endpoint occurred in 14 of 681 patients with scores of 0–3 (2.1%), 29 of 1,022 patients with scores of 4–5 (2.8%), and 58 of 1,055 patients with scores of at least 6 (5.5%) (Figure 2D). Although the event proportion was highest in the upper score category, considerable overlap remained between patients with and without the endpoint.

The apparent Brier score was 0.0351, and the scaled Brier score relative to the intercept-only model was 0.006 (Figure 2B). These findings indicated little improvement in overall prediction beyond the baseline event rate.

Sensitivity and subgroup analyses

The stricter VTE proxy, which required both a non-primary-position VTE code and initiation of treatment-dose anticoagulation after the landmark, was met by 52 patients (1.9%). Discrimination remained modest, with an area under the curve of 0.641 (95% CI, 0.569–0.713).

Among the 1,741 patients with complete information for mobility and body mass index, 61 met the primary endpoint. The area under the curve in this complete-core-item subgroup was 0.644 (95% CI, 0.575–0.713), showing no substantial improvement over the main analysis.

Performance was also examined according to the recorded site of gastrointestinal bleeding. The area under the curve was 0.621 among patients with upper gastrointestinal bleeding and 0.643 among those with lower gastrointestinal bleeding. Confidence intervals were wide, and the mixed-site subgroup contained only one endpoint-positive patient; an estimate was therefore not calculated for that subgroup.

Multivariable associations with the primary endpoint

The complete-case multivariable logistic regression analysis included 2,707 patients and 100 primary endpoints. The adjusted model had an area under the curve of 0.616 (95% CI, 0.562–0.670), which was only marginally higher than that of the Padua score alone.

Each 1-point increase in the Padua score was associated with a 17% increase in the odds of meeting the primary endpoint after adjustment for sex, Charlson Comorbidity Index, hemoglobin,

platelet count, creatinine, Braden nutrition category, physical prophylaxis, and pharmacologic prophylaxis (adjusted odds ratio [OR], 1.17; 95% CI, 1.06–1.28; Benjamini–Hochberg-adjusted P=0.008). None of the remaining covariates retained statistical significance after correction for multiple testing (Table 3 and Figure 3A).

In the prespecified Firth penalized logistic regression model, known thrombophilia was strongly associated with the primary endpoint (adjusted OR, 11.56; 95% CI, 3.73–31.46; adjusted P=0.002). The estimate was based on a small number of patients with documented thrombophilia and should therefore be interpreted with appropriate caution. Heart or respiratory failure showed a nominal association of borderline significance, but this was not retained after multiplicity adjustment. Confidence intervals for the remaining clinical and nursing variables included the null value after correction (Figure 3B).

Reduced mobility was the most prevalent Padua component, occurring in 76.0% of the cohort, followed by acute infection or rheumatic disease (61.3%) and age of at least 70 years (49.5%) (Figure 3C). Endpoint proportions were higher in patients with lower Braden mobility categories, although uncertainty was greatest in the least mobile group because of the smaller number of patients (Figure 3D).

Thromboprophylaxis patterns

Physical prophylaxis without pharmacologic prophylaxis was the most frequently documented strategy during the first 24 hours and was used in 1,795 patients (65.1%). Combined physical and pharmacologic prophylaxis was documented in 376 patients (13.6%), pharmacologic prophylaxis alone in 129 (4.7%), and neither modality in 458 (16.6%) (Figure 4A).

Unadjusted proportions of the primary VTE coding endpoint were 3.5% in the physical-only group, 3.3% among patients with no documented prophylaxis, 4.5% in the combined-prophylaxis group, and 4.7% in the pharmacologic-only group. The strict VTE proxy occurred in 1.8%, 1.1%, 2.4%, and 3.9% of these groups, respectively. The corresponding proportions for the severe rebleeding proxy were 1.3%, 1.1%, 2.4%, and 2.3% (Figure 4B and Table 4). These unadjusted comparisons were descriptive and reflected substantial differences in treatment selection.

Exploratory overlap-weighted analyses

Exploratory analyses of prophylaxis were restricted to patients with Padua scores of at least 4. After overlap weighting, the measured baseline covariates were well balanced in both the pharmacologic and physical prophylaxis comparisons, with absolute standardized mean differences below 0.001 (Figure 4C).

Pharmacologic prophylaxis was not significantly associated with the primary VTE coding endpoint (weighted OR, 1.22; 95% CI, 0.72–2.08), the strict VTE proxy (weighted OR, 1.57; 95% CI, 0.79–3.10), or the severe rebleeding proxy (weighted OR, 1.45; 95% CI, 0.64–3.29).

Similarly, physical prophylaxis was not significantly associated with the primary endpoint (weighted OR, 1.43; 95% CI, 0.51–4.00), the strict VTE proxy (weighted OR, 3.08; 95% CI, 0.42–22.73), or the severe rebleeding proxy (weighted OR, 0.73; 95% CI, 0.21–2.49) (Figure 4D and Table 4). All multiplicity-adjusted P values exceeded 0.59. The wide confidence intervals, particularly in the physical prophylaxis analysis, indicated substantial statistical uncertainty and did not exclude clinically important benefit or harm.

Discussion

In this cohort of critically ill patients with gastrointestinal bleeding, the Padua Prediction Score had limited ability to distinguish patients with and without an in-hospital VTE coding endpoint. Although the conventional threshold of ≥4 identified most endpoint-positive patients, it classified approximately three quarters of the cohort as high risk and yielded a specificity of only 25.1%. Increasing the threshold to ≥6 improved specificity, but sensitivity declined substantially. Comparable results were obtained with the stricter VTE proxy and in patients with complete mobility and body mass index data. Taken together, these findings suggest that the Padua score retains some association with thrombotic risk in this population but provides insufficient separation for use as a stand-alone prophylaxis decision rule.

The observed discrimination was lower than that reported in broader medical inpatient populations. In a recent cohort of more than 42,000 hospitalized medical patients, the Padua score achieved an area under the curve of 0.735 [7], compared with 0.610 in the present study.

This difference likely reflects the more restricted and clinically complex case mix of ICU patients with gastrointestinal bleeding. Reduced mobility, acute infection, advanced age, cardiorespiratory failure, malignancy, and other comorbidities are common in critically ill patients[15]. In our cohort, reduced mobility was present in 76.0% of patients, and 75.3% met the conventional high-risk threshold. The resulting concentration of scores in the higher range reduced the score's ability to separate patients who met the VTE endpoint from those who did not. The Padua model was developed in a heterogeneous general medical population [16], and its original threshold may be less informative when applied to a group in which multiple risk factors are already highly prevalent.

The Padua score nevertheless remained independently associated with the primary endpoint, and the proportion of endpoint-positive patients increased across score categories[17]. These findings indicate that the score preserved an overall risk gradient. However, an association at the population level does not necessarily translate into accurate prediction for individual patients. This distinction is particularly relevant in gastrointestinal bleeding, where classification as high VTE risk may influence the use of anticoagulant prophylaxis and therefore must be weighed against the possibility of recurrent hemorrhage.

The conventional threshold showed high sensitivity but poor specificity. Its positive predictive value was only 4.2%, meaning that most patients classified as high risk did not meet the primary endpoint. The negative predictive value was high, but this result should be interpreted in light of the low endpoint prevalence. When an outcome is uncommon, negative predictive values are generally high even when model discrimination is modest. Similarly, the low apparent Brier score was largely influenced by the low frequency of the endpoint, whereas the scaled Brier score showed little improvement over prediction based on the cohort event rate alone. These findings caution against interpreting apparently favorable negative predictive values or unscaled calibration measures as evidence of strong clinical performance.

A threshold of ≥6 provided a more balanced combination of sensitivity and specificity, but it missed more than 40% of endpoint-positive patients. Although this threshold maximized the Youden index in the present cohort, it should not be proposed as a replacement for the established cut-off. The estimate was derived from the same dataset in which performance was

assessed, and the outcome was a coding-based proxy rather than an adjudicated incident VTE event. The threshold analysis instead illustrates the limited room for improvement through simple cut-off modification: gains in specificity were accompanied by a clinically important loss of sensitivity.

Outcome definition is an important consideration when interpreting these results. The primary endpoint was based on a non-primary-position VTE discharge code because diagnosis timestamps were unavailable in MIMIC-IV. Consequently, it was not possible to determine with certainty whether VTE occurred after ICU admission or after the 24-hour landmark. The stricter endpoint additionally required initiation of treatment-dose anticoagulation after the landmark and produced a slightly higher area under the curve. This may indicate that outcome misclassification contributed to the weak performance of the primary analysis. Nevertheless, treatment-dose anticoagulation is also an imperfect marker because it may be withheld in patients with active bleeding or prescribed for other indications. Both outcomes should therefore be understood as clinically relevant proxies rather than confirmed incident events.

Early prophylaxis patterns reflected the difficult balance between thrombosis and bleeding in this population. Physical prophylaxis alone was the predominant strategy, whereas pharmacologic prophylaxis was documented in fewer than one fifth of patients. After overlap weighting, neither pharmacologic nor physical prophylaxis was significantly associated with the VTE or severe rebleeding proxies. These results do not demonstrate equivalence between prophylaxis strategies or absence of treatment effect. Confidence intervals were wide, especially for physical prophylaxis, and treatment selection was likely influenced by factors not fully captured in the database, including active bleeding, endoscopic hemostasis, transfusion requirements, planned procedures, and clinician assessment of bleeding stability. Moreover, documentation of a compression device does not confirm consistent application or adherence. The weighted analyses should therefore remain exploratory and should not be interpreted causally.

The findings have practical implications. The Padua score may help clinicians organize established thrombotic risk factors, but it does not account for the severity or activity of gastrointestinal bleeding. A high score cannot by itself establish that pharmacologic prophylaxis is appropriate, particularly before hemostasis has been achieved. Conversely, a score below 4

does not exclude clinically important thrombotic risk in a patient with persistent immobility or prolonged critical illness. Prophylaxis decisions in this setting are better approached through parallel assessment of thrombotic risk and bleeding status, with reassessment after endoscopic treatment, hemodynamic stabilization, and changes in mobility. A dynamic approach is likely to be more clinically relevant than a single score calculated during the first day of ICU admission.

This study included a relatively large cohort of ICU patients with gastrointestinal bleeding and evaluated the Padua score across conventional and alternative thresholds, calibration measures, sensitivity analyses, and clinically relevant subgroups. The 24-hour landmark also provided a defined baseline period for measuring patient characteristics and early prophylaxis. Several limitations should be acknowledged. Discharge diagnoses lacked timestamps, and some VTE codes may have represented conditions present before the landmark. Padua components reconstructed from discharge codes may have incorporated information documented later in the hospitalization, whereas absent documentation may have led to under-recognition of certain conditions. The single-center design limits generalizability, and the small number of endpoints reduced precision in subgroup and weighted analyses. Residual confounding was unavoidable in the prophylaxis analyses, despite good balance in measured covariates. Finally, the severe rebleeding endpoint was based on hemoglobin decline and transfusion rather than clinical adjudication.

In conclusion, the Padua Prediction Score showed limited discrimination in critically ill patients with gastrointestinal bleeding. The conventional threshold retained high sensitivity but had poor specificity and classified most patients as high risk. The score may contribute to thrombotic risk assessment, but it should be interpreted alongside current bleeding activity, hemostatic stability, and repeated clinical evaluation rather than used alone to guide thromboprophylaxis.

Declarations

Ethics approval and consent to participate

The MIMIC-IV data resource was approved by the institutional review board of Beth Israel Deaconess Medical Center with a waiver of individual informed consent. This study used deidentified data under the PhysioNet data-use framework.

Consent for publication

Not applicable.

Availability of data and materials

MIMIC-IV version 3.1 is available to credentialed users through PhysioNet under its data-use agreement. Derived patient-level data should remain under equivalent access controls. Analysis code is available in the study project directory and can be prepared for a controlled repository after data-use compliance review.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

References

1. Gao X, Zeng L, Wang H, et al. Prevalence of venous thromboembolism in intensive care units: a meta-analysis. J Clin Med. 2022;11(22):6691. doi:10.3390/jcm11226691.
2. Schunemann HJ, Cushman M, Burnett AE, et al. American Society of Hematology 2018 guidelines for management of venous thromboembolism: prophylaxis for hospitalized and nonhospitalized medical patients. Blood Adv. 2018;2(22):3198-3225. doi:10.1182/bloodadvances.2018022954.
3. Sengupta N, Feuerstein JD, Jairath V, et al. Management of patients with acute lower gastrointestinal bleeding: an updated ACG guideline. Am J Gastroenterol. 2023;118(2):208-231. doi:10.14309/ajg.0000000000002130.
4. Triantafyllou K, Gkolfakis P, Gralnek IM, et al. Diagnosis and management of acute lower gastrointestinal bleeding: European Society of Gastrointestinal Endoscopy guideline. Endoscopy. 2021;53(8):850-868. doi:10.1055/a-1496-8969.
5. 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. J Thromb Haemost. 2010;8(11):2450-2457. doi:10.1111/j.1538-7836.2010.04044.x.

6. Gerotziafas GT, Papageorgiou L, Salta S, Nikolopoulou K, Elalamy I. Updated clinical models for VTE prediction in hospitalized medical patients. Thromb Res. 2018;164(suppl 1):S62-S69. doi:10.1016/j.thromres.2018.02.004.

7. Yifang H, Jun D, Jingting Y, Ying S, Ping Z, Xiaomei D. Comparison of the PADUA and IMPROVE scores in assessing venous thromboembolism risk in 42,257 medical inpatients in China. J Thromb Thrombolysis. 2024;57(5):775-783. doi:10.1007/s11239-024-02979-y.

8. Malhotra N, Chande N. Venous thromboprophylaxis in gastrointestinal bleeding. Can J Gastroenterol Hepatol. 2015;29(3):145-148. doi:10.1155/2015/525623.

9. Repp AB, Sparks AD, Wilkinson K, et al. Factors associated with venous thromboembolism pharmacoprophylaxis initiation in hospitalized medical patients: the Medical Inpatients Thrombosis and Hemostasis study. J Thromb Haemost. 2024;22(12):3521-3531. doi:10.1016/j.jtha.2024.08.016.

10. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data. 2023;10(1):1. doi:10.1038/s41597-022-01899-x.

11. Johnson A, Bulgarelli L, Pollard T, et al. MIMIC-IV (version 3.1). PhysioNet. 2024. doi:10.13026/kpb9-mt58.

12. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology statement: guidelines for reporting observational studies. Bull World Health Organ. 2007;85(11):867-872. doi:10.2471/BLT.07.045120.

13. Heinze G, Schemper M. A solution to the problem of separation in logistic regression. Stat Med. 2002;21(16):2409-2419. doi:10.1002/sim.1047.

14. Li F, Thomas LE, Li F. Addressing extreme propensity scores via the overlap weights. Am J Epidemiol. 2019;188(1):250-257. doi:10.1093/aje/kwy201.

15. Fernando SM, Tran A, Cheng W, et al. VTE prophylaxis in critically ill adults: a systematic review and network meta-analysis. Chest. 2022;161(2):418-428. doi:10.1016/j.chest.2021.08.050.

16. Peng D, Zhai H. Application of antithrombotic drugs and risk factor analysis in ICU patients with lower gastrointestinal bleeding from MIMIC-IV. BMC Gastroenterol. 2024;24(1):319. doi:10.1186/s12876-024-03380-y.

17. Villiger R, Julliard P, Darbellay Farhoumand P, et al. Prediction of in-hospital bleeding in acutely ill medical patients: external validation of the IMPROVE bleeding risk score. Thromb Res. 2023;230:37-44. doi:10.1016/j.thromres.2023.08.003.

Figure legends

Figure 1. Study cohort flow

Sequential construction of the 24-hour landmark cohort from all MIMIC-IV version 3.1 hospital admissions. GIB indicates gastrointestinal bleeding; ICU, intensive care unit; VTE, venous thromboembolism.

Figure 2. Padua score performance for the primary venous thromboembolism coding endpoint

(A) Receiver operating characteristic curve with area under the curve and 95% confidence interval. (B) Apparent calibration using grouped observed and model-predicted proportions; error bars are Wilson 95% confidence intervals and point size reflects group size. (C) Sensitivity and specificity across integer Padua thresholds, highlighting the conventional threshold of at least 4 and the maximum-Youden threshold of at least 6. (D) Primary-endpoint proportions across descriptive Padua score tiers. Because discharge diagnoses lack timestamps, the endpoint is a coding proxy rather than an adjudicated incident event.

Figure 3. Clinical and nursing-related associations with the primary endpoint

(A) Adjusted odds ratios from the complete-case multivariable logistic model. (B) Adjusted odds ratios from the Firth penalized logistic model of prespecified risk factors. (C) Prevalence of individual Padua components in the landmark cohort. (D) Primary-endpoint proportions across first-24-hour Braden mobility categories. Error bars show 95% confidence intervals where applicable. CI indicates confidence interval; GCS, Glasgow Coma Scale.

Figure 4. Prophylaxis patterns, covariate balance, and exploratory weighted associations

(A) Distribution of physical and pharmacologic prophylaxis patterns during the first 24 hours. (B)

Unadjusted proportions of the primary VTE code, strict VTE proxy, and severe rebleeding proxy by prophylaxis group. (C) Absolute standardized mean differences before and after overlap weighting for separate pharmacologic and physical prophylaxis analyses; the dashed line marks 0.10. (D) Overlap-weighted odds ratios with robust 95% confidence intervals. The weighted analyses are exploratory and do not establish causal effects. VTE indicates venous thromboembolism.

Main Tables

Table 1. Baseline characteristics by the primary venous thromboembolism coding endpoint

Characteristic	Overall (N=2,758)	No VTE code (n=2,657)	VTE code (n=101)	P value	SMD	Missing, %
Patients, n	2758	2657	101			
Age, years, median [IQR]	69.00 [57.00, 81.00]	69.00 [57.00, 81.00]	73.00 [62.00, 81.00]	0.117	0.166	0.0
Male sex, n (%)	1660 (60.2)	1597 (60.1)	63 (62.4)	0.723	0.047	0.0
Race/ethnicity, n (%)				0.079	0.293	0.0
Asian	99 (3.6)	92 (3.5)	7 (6.9)			
Black	357 (12.9)	350 (13.2)	7 (6.9)			
Hispanic/Latino	97 (3.5)	95 (3.6)	2 (2.0)			
Other/unknown	302 (10.9)	287 (10.8)	15 (14.9)			
White	1903 (69.0)	1833 (69.0)	70 (69.3)			
Charlson score, median [IQR]	6.00 [4.00, 8.00]	6.00 [4.00, 8.00]	6.00 [4.00, 9.00]	0.044	0.192	0.0
Padua score, median [IQR]	5.00 [4.00, 6.00]	5.00 [3.00, 6.00]	6.00 [4.00, 7.00]	<0.001	0.387	0.0
Padua score >=4, n (%)	2077 (75.3)	1990 (74.9)	87 (86.1)	0.014	0.287	0.0
Gastrointestinal bleeding site, n (%)				0.786	0.106	0.0
Lower	467 (16.9)	451 (17.0)	16 (15.8)			
Mixed	34 (1.2)	33 (1.2)	1 (1.0)			
Unspecified	641 (23.2)	621 (23.4)	20 (19.8)			
Upper	1616 (58.6)	1552 (58.4)	64 (63.4)			
Hemoglobin, g/dL, median [IQR]	9.00 [7.70, 10.40]	9.00 [7.70, 10.40]	8.75 [7.68, 10.10]	0.426	0.072	0.6
Platelet count, x10^9/L, median [IQR]	172.00 [115.00, 240.00]	172.00 [115.00, 239.00]	179.50 [118.25, 261.75]	0.277	0.118	0.7

Characteristic	Overall (N=2,758)	No VTE code (n=2,657)	VTE code (n=101)	P value	SMD	Missing, %
International normalized ratio, median [IQR]	1.30 [1.10, 1.60]	1.30 [1.10, 1.60]	1.30 [1.20, 1.50]	0.786	0.132	9.3
Creatinine, mg/dL, median [IQR]	1.00 [0.70, 1.60]	1.00 [0.70, 1.58]	1.00 [0.70, 1.92]	0.984	0.054	0.9
Blood urea nitrogen, mg/dL, median [IQR]	28.00 [17.00, 49.00]	28.00 [17.00, 48.00]	31.00 [19.00, 58.00]	0.078	0.233	1.0
White blood cell count, x10^9/L, median [IQR]	9.10 [6.50, 12.70]	9.10 [6.47, 12.60]	9.55 [7.20, 14.25]	0.109	0.212	0.7
Braden activity category, median [IQR]	1.00 [1.00, 2.00]	1.00 [1.00, 2.00]	1.00 [1.00, 2.00]	0.611	0.078	0.8
Braden mobility category, median [IQR]	3.00 [3.00, 3.00]	3.00 [3.00, 3.00]	3.00 [2.00, 3.00]	0.005	0.283	0.8
Braden nutrition category, median [IQR]	2.00 [2.00, 3.00]	2.00 [2.00, 3.00]	2.00 [2.00, 3.00]	0.234	0.126	0.8
Physical prophylaxis in first 24 h, n (%)	2171 (78.7)	2091 (78.7)	80 (79.2)	1.000	0.013	0.0
Pharmacologic prophylaxis in first 24 h, n (%)	505 (18.3)	482 (18.1)	23 (22.8)	0.294	0.115	0.0
Severe rebleeding proxy, n (%)	41 (1.5)	36 (1.4)	5 (5.0)	0.016	0.207	0.0
In-hospital death, n (%)	172 (6.2)	152 (5.7)	20 (19.8)	<0.001	0.432	0.0

Data are median [interquartile range] or n (%). P values are descriptive. SMD indicates standardized mean difference; VTE, venous thromboembolism.

Table 2. Padua score discrimination and selected threshold performance

A. Discrimination analyses

Analysis	N	Events	AUC	95% CI
Primary non-primary-position VTE code	2758	101	0.610	0.555-0.665
Strict VTE proxy	2758	52	0.641	0.569-0.713
Complete Padua core items	1741	61	0.644	0.575-0.713

B. Selected thresholds for the primary endpoint

Threshold	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Conventional threshold >=4	86.1 (78.1-91.6)	25.1 (23.5-26.8)	4.2 (3.4-5.1)	97.9 (96.6-98.8)
Maximum-Youden threshold >=6	57.4 (47.7-66.6)	62.5 (60.6-64.3)	5.5 (4.3-7.0)	97.5 (96.6-98.1)

AUC indicates area under the receiver operating characteristic curve; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; VTE, venous thromboembolism. Apparent Brier score=0.0351; scaled Brier score=0.006.

Table 3. Adjusted associations with the primary venous thromboembolism coding endpoint

A. Complete-case multivariable logistic model (N=2,707; events=100)

Variable	Adjusted OR	95% CI	P value	BH-adjusted P
Braden nutrition, per category	0.91	0.66-1.24	0.546	0.951
Charlson comorbidity score, per point	0.99	0.90-1.09	0.856	0.951
Hemoglobin, per g/dL	0.98	0.88-1.08	0.681	0.951
Creatinine, per doubling	0.99	0.79-1.23	0.951	0.951
Padua score, per point	1.17	1.06-1.28	<0.001	0.008
Pharmacologic prophylaxis in first 24 h	1.32	0.80-2.11	0.260	0.951
Physical prophylaxis in first 24 h	0.98	0.61-1.65	0.938	0.951
Platelet count, per 50 x10^9/L	1.03	0.95-1.11	0.441	0.951
Male vs female	1.11	0.73-1.70	0.637	0.951

B. Prespecified Firth penalized logistic model

Variable	Adjusted OR	95% CI	P value	BH-adjusted P
Known thrombophilia	11.56	3.73-31.46	<0.001	0.002
Heart or respiratory failure	1.55	1.00-2.40	0.051	0.185
Acute infection or rheumatic disease	1.34	0.87-2.14	0.190	0.299
Active cancer proxy	1.60	0.94-2.63	0.082	0.226
Age >=70 y	1.36	0.89-2.10	0.157	0.288
GCS, per point	0.94	0.90-0.99	0.025	0.135
Braden mobility, per category	0.86	0.62-1.17	0.335	0.460
Male vs female	1.12	0.73-1.73	0.601	0.721
Hemoglobin, per g/dL	0.98	0.88-1.08	0.656	0.721
Platelet count, per 50 x10^9/L	1.06	0.98-1.13	0.150	0.288

Variable	Adjusted OR	95% CI	P value	BH-adjusted P
Creatinine, per doubling	0.97	0.77-1.20	0.789	0.789

BH indicates Benjamini-Hochberg; CI, confidence interval; OR, odds ratio. Adjusted P values control the false-discovery rate within each model.

Table 4. Prophylaxis patterns and exploratory overlap-weighted associations

A. Unadjusted outcomes by prophylaxis group

Prophylaxis group	N (%)	Primary VTE code, n (%)	Strict VTE proxy, n (%)	Severe rebleeding proxy, n (%)
Physical only	1,795 (65.1)	63 (3.5)	33 (1.8)	24 (1.3)
None documented	458 (16.6)	15 (3.3)	5 (1.1)	5 (1.1)
Combined	376 (13.6)	17 (4.5)	9 (2.4)	9 (2.4)
Pharmacologic only	129 (4.7)	6 (4.7)	5 (3.9)	3 (2.3)

B. Overlap-weighted associations

Analysis	Outcome	N	Exposed	Weighted OR	95% CI	BH-adjusted P
Physical prophylaxis	Primary VTE code	1772	1635	1.43	0.51-4.00	0.595
Physical prophylaxis	Strict VTE proxy	1772	1635	3.08	0.42-22.73	0.595
Physical prophylaxis	Severe rebleeding proxy	1772	1635	0.73	0.21-2.49	0.614
Pharmacologic prophylaxis	Primary VTE code	2047	377	1.22	0.72-2.08	0.595
Pharmacologic prophylaxis	Strict VTE proxy	2047	377	1.57	0.79-3.10	0.595
Pharmacologic prophylaxis	Severe rebleeding proxy	2047	377	1.45	0.64-3.29	0.595

BH indicates Benjamini-Hochberg; CI, confidence interval; OR, odds ratio; VTE, venous thromboembolism. All measured covariates had absolute standardized mean differences below 0.001 after weighting. Estimates are exploratory and should not be interpreted causally.

Main Figures

Figure 1. Study cohort flow.

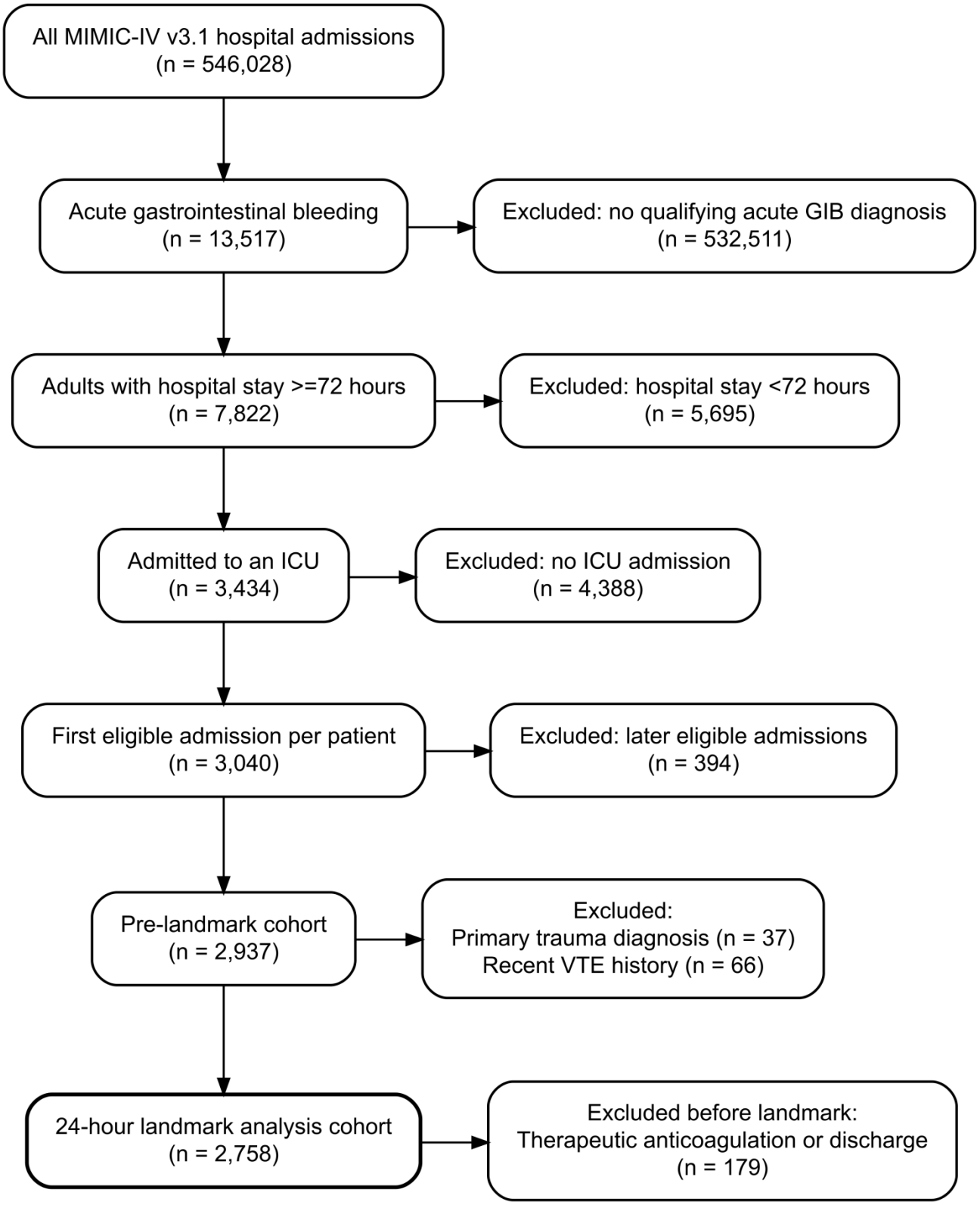


Figure 2. Padua score performance for the primary venous thromboembolism coding endpoint.

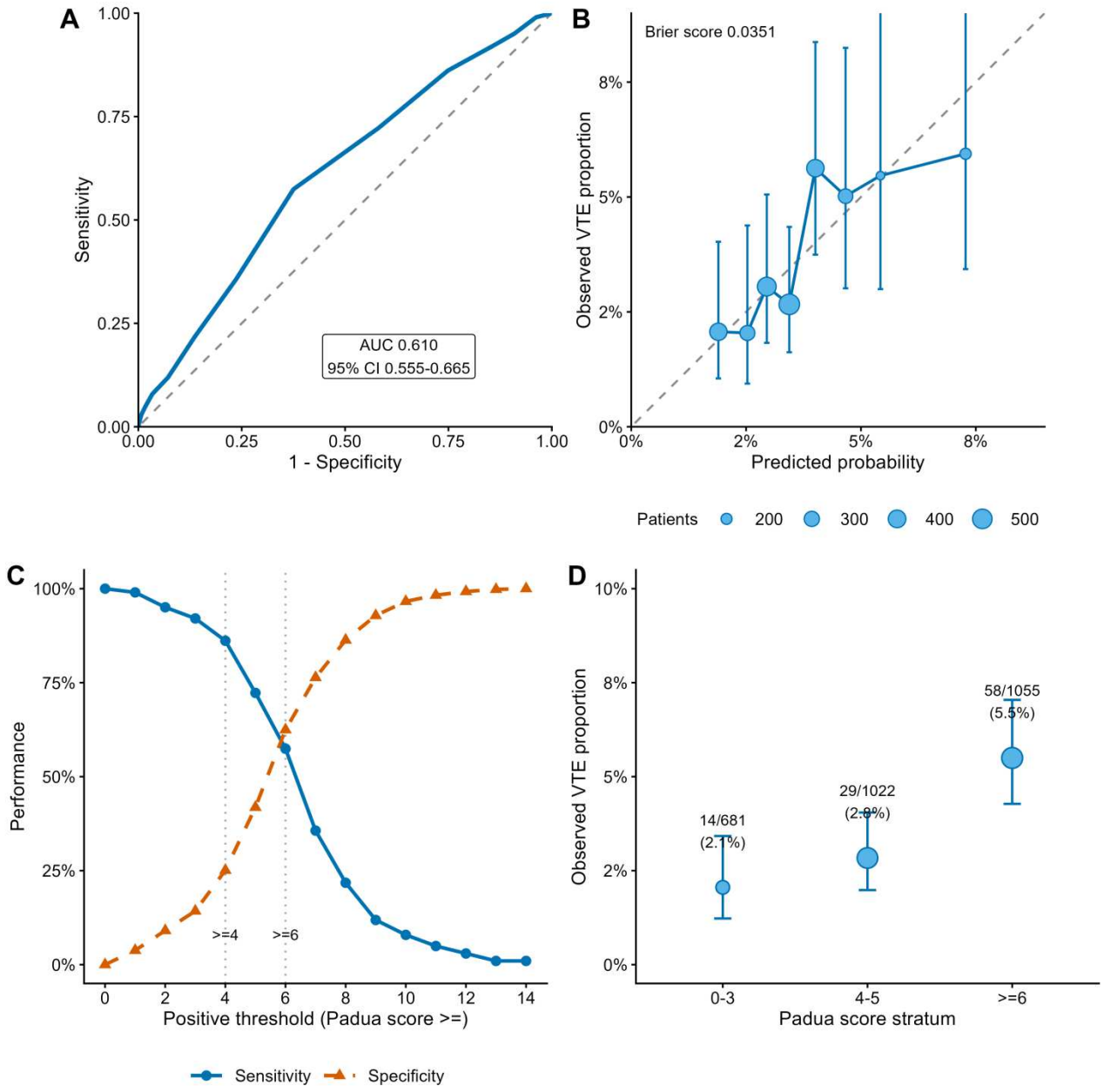


Figure 3. Clinical and nursing-related associations with the primary endpoint.

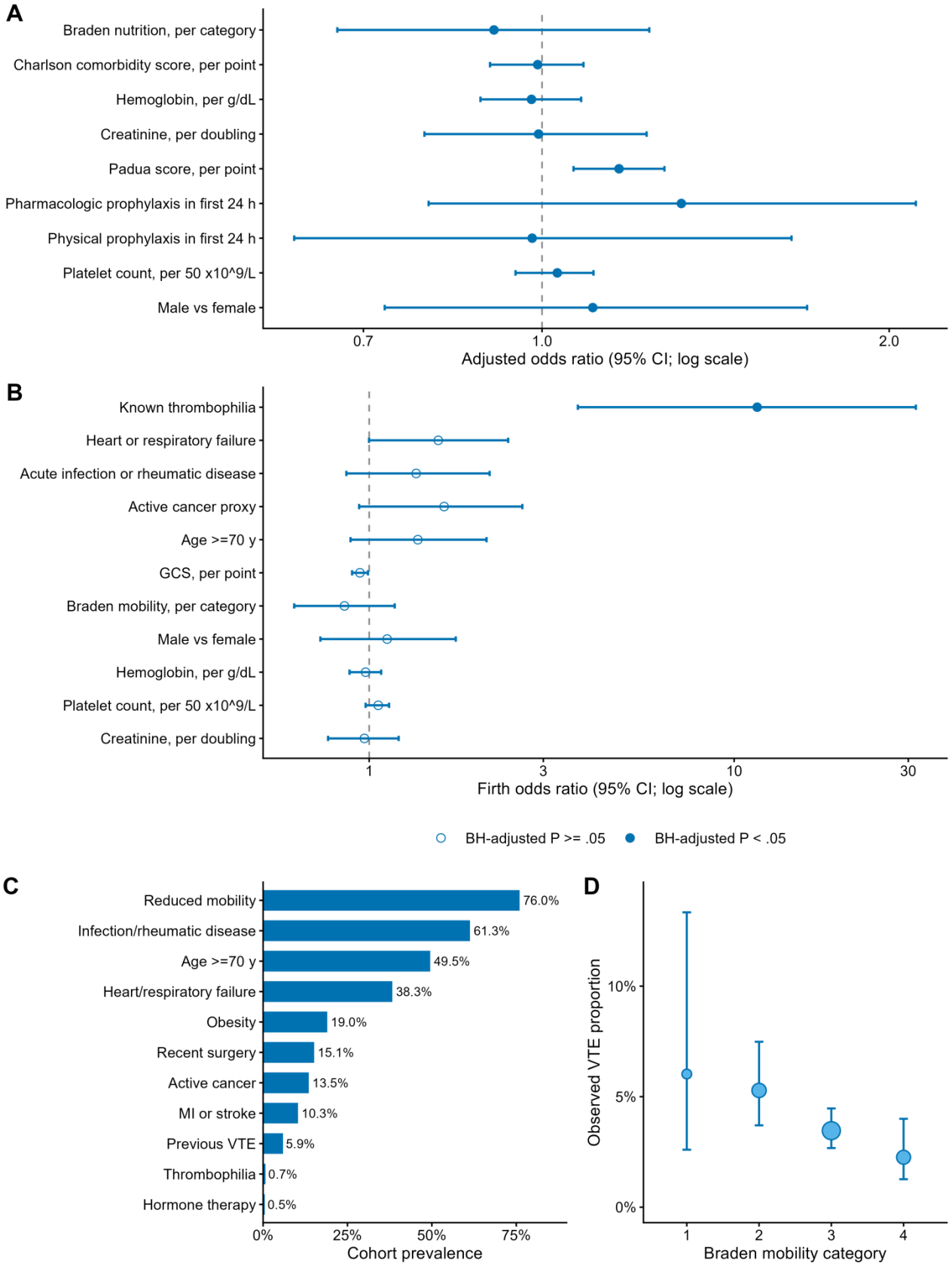


Figure 4. Prophylaxis patterns, covariate balance, and exploratory weighted associations.

