Received: 1 July 2024 | Revised: 25 September 2024 | Accepted: 29 October 2024

https://doi.org/10.1016/j.rpth.2024.102650

ORIGINAL ARTICLE

External validation of a novel cancer-associated venous thromboembolism risk assessment score in a safety-net hospital

Karlynn N. Dulberger[1] | Jennifer La[1,2,3] | Ang Li[4] | Saran Lotfollahzadeh[5] | Asha Jose[5] | Nhan V. Do[1,2] | Mary T. Brophy[1,2] | J. Michael Gaziano[1,2,3] | Katya Ravid[6,7] | Vipul C. Chitalia[2,5] | Nathanael R. Fillmore[1,2,3]

[1]Section of Hematology & Medical Oncology, Boston University School of Medicine, Boston, Massachusetts, USA

[2]Massachusetts Veterans Epidemiology Research and Information Center, Veterans Affairs Boston Healthcare System, Boston, Massachusetts, USA

[3]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA

[4]Section of Hematology-Oncology, Baylor College of Medicine, Houston, Texas, USA

[5]Renal Section, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

[6]Department of Medicine, Whitaker Cardiovascular Institute, Boston University School of Medicine, Boston, Massachusetts, USA

[7]Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts, USA

Correspondence
Nathanael R. Fillmore, VA Boston Healthcare System, Harvard Medical School, 150 S Huntington Ave, Boston, MA 02130, USA.
Email: Nathanael.Fillmore@va.gov

Handling Editor: Nick van Es

Abstract

Background: Cancer-associated thrombosis (CAT) is a leading cause of death in patients diagnosed with cancer. However, pharmacologic thromboprophylaxis use in cancer patients must be carefully evaluated due to a 2-fold increased risk of experiencing a major bleeding event within this population. The electronic health record CAT (EHR-CAT) risk assessment model (RAM) was recently developed, and reports improved performance over the widely used Khorana score. Extensive RAM external validation is crucial to determine accuracy across diverse patient populations prior to clinical utilization.

Objectives: To externally validate EHR-CAT using data from 2103 patients with cancer at the Boston Medical Center (BMC), New England's largest safety-net hospital, and to compare this RAM with the Khorana score.

Methods: We conducted a retrospective study of BMC cancer patients diagnosed between January 2014 and December 2022 using data from the BMC tumor registry and EHR system. We validated the RAM using measures of discrimination and calibration.

Results: The EHR-CAT score exhibited a strong ability to discriminate the risk of CAT (C statistic, 0.67), which was substantially higher than the classic Khorana score (C statistic, 0.58). This increased discrimination power reflects the 20% of patients that were reclassified into high or low risk by the expanded score. Model calibration was also strong in this dataset.

Conclusion: In our external validation, the recently published EHR-CAT score showed clear and improved separation of patients at high and low risk for CAT. The utilization of this expanded CAT score could facilitate improved targeting of at-risk cancer patients for prophylactic therapy.

Karlynn N. Dulberger, Jennifer La, and Ang Li contributed equally to this study as first authors.

KEYWORDS
cardiology, electronic health records, oncology, risk assessment, venous thromboembolism

Essentials

- Identifying cancer patients at increased risk of venous thromboembolism is of clinical importance.
- We performed an external validation of the EHR-CAT score in BMC patients.
- The EHR-CAT score showed clear and improved separation of patients into high- and low-risk groups.
- Using this score could help target cancer patients for prophylactic therapy.

1 | INTRODUCTION

Venous thromboembolism (VTE) is a potentially preventable condition that is responsible for significant patient morbidity and mortality [1–5]. Cancer-associated thrombosis (CAT) accounts for as much as 30% of all primary VTE events [6–12] and is a leading cause of non–cancer-related death in ambulatory patients diagnosed with cancer [3]. Both cancer-specific and treatment-related factors are responsible for the 4- to 7-fold increased risk of VTE, which accompanies a cancer diagnosis [3,4,13–16]. Antithrombotics can reduce VTE risk by 50% to 80% [17–20]. However, universal pharmacologic thromboprophylaxis use in cancer patients is not justified due to a 2-fold increased risk of a major bleeding event within this population [21]. Therefore, proper identification of which cancer patients are at a dangerously elevated risk of CAT is crucial.

A novel risk assessment model (RAM) recently developed by Li et al. [22,23], the electronic health record CAT (EHR-CAT) score, reports improved discrimination of VTE events among cancer patients compared with the widely used Khorana score. The Khorana score is a clinical tool used to identify patients with an increased risk of CAT [24]. This score incorporates patient-specific predictors (body mass index, white blood cell count, hemoglobin, and platelet levels) as well as cancer type to assign points that indicate patient risk level. The EHR-CAT score refines the cancer type variable and incorporates additional predictors pertaining to treatment type, VTE history, recent paralysis or immobility, and Asian race.

EHR-CAT was developed using retrospective data from a diverse patient population from the Harris Health System, the largest safety-net hospital system in Texas. The model has been validated using data from the National Veteran Affairs healthcare system [22] and, separately, data from the MD Anderson cancer center [23]. Rigorous external validation helps determine the reproducibility and generalizability of models in diverse clinical settings [25–28]. Accordingly, in the current study, we performed external validation of the risk score in a cancer cohort from New England's largest safety-net hospital, the Boston Medical Center (BMC). This cohort is regionally and ethnically distinct from all prior applications of this score. The findings from this study could improve confidence in this model and support its application as a clinical tool across diverse patient populations.

2 | METHODS

This longitudinal cohort study was approved by the Boston Medical Center Institutional Review Board prior to data collection and analysis with a waiver of informed consent. Retrospective analysis was performed on patients diagnosed with cancer at BMC between January 1, 2014, and December 31, 2022. Cancer type, diagnosis date, and American Joint Committee on Cancer staging were obtained from the BMC Tumor Registry. Other data, including patients' visit details, medication history, laboratory values, and imaging notes, were obtained from EHR data in the BMC Clinical Data Warehouse.

The inclusion criteria were as follows: patients were required to have a confirmed first-ever cancer diagnosis that was not benign, *in situ*, or stage 0; patients were also required to have systemic therapy within 1 year of diagnosis; patients were excluded if they received an anticoagulation prescription 30 days on or before therapy initiation or if they experienced acute VTE within the 6 months prior to treatment; patients were also excluded if they were missing key medical information at the time of therapy initiation, such as cancer group, height, weight, white blood cell count, and hemoglobin or platelet numbers; patients with undisclosed race or American Joint Committee on Cancer stage (except in cases of brain, leukemia, or myeloma) were also excluded (Figure 1). Variable definitions are described in detail in the data dictionary (Supplementary Table S1).

The primary outcome was overall VTE. This included radiologically confirmed symptomatic or incidental pulmonary embolism (PE), lower extremity (LE) deep vein thrombosis (LE-DVT), and upper extremity DVT. The secondary outcome was PE and LE-DVT (PE/LE-DVT). The index date was defined as the date of systemic therapy initiation. Patients were followed from the index date until the first outcome event, death, loss of follow-up (defined as a 90-day gap without any clinical encounters), or administrative censoring on December 31, 2022. Outcomes were identified by Internal Classification of Disease codes (Supplementary Table S2) and a natural language processing algorithm applied to imaging notes [22], followed by a manual chart review to adjudicate all putative VTE events. A cumulative incidence competing risk model was used to calculate VTE incidence accounting for the competing risk of death [29].

To assess performance, EHR-CAT was applied, and patients were separated either by their tallied numeric risk score or into high- and

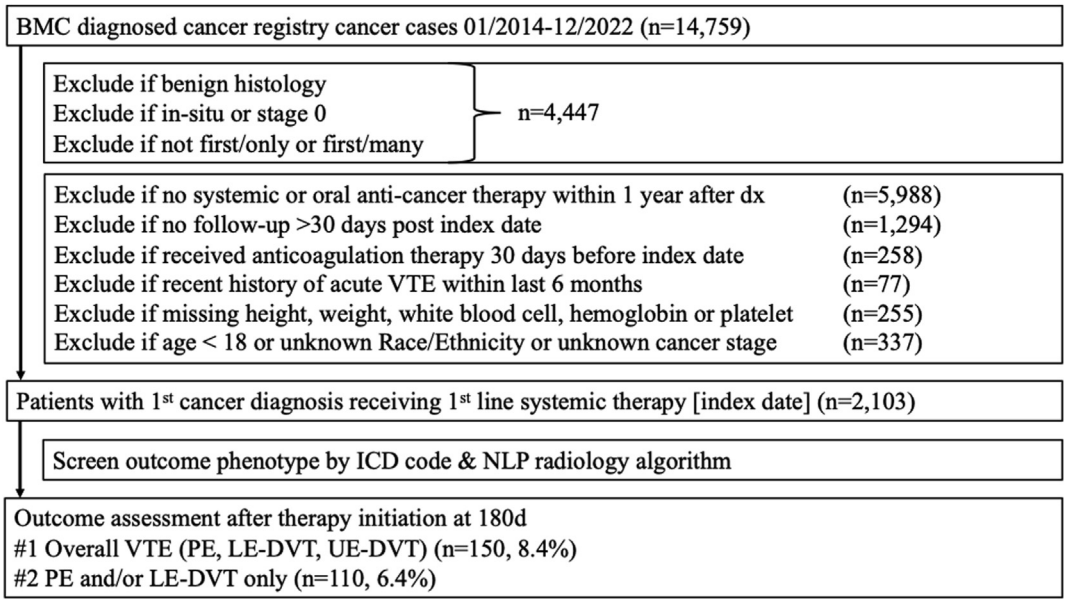

FIGURE 1 Cohort and outcome definition. Cohort selection required stringent inclusion and exclusion criteria. The primary and secondary outcomes were assessed at 180 days posttreatment initiation (index date) by Internal Classification of Disease (ICD) codes and natural language processing (NLP) radiology algorithm. BMC, Boston Medical Center; LE-DVT, lower extremity deep vein thrombosis; PE, pulmonary embolism; UE-DVT, upper extremity deep vein thrombosis; VTE, venous thromboembolism.

low-risk groups. We used previously published cutoffs to distinguish low- vs high-risk groups [22,23]. Patients with a score of 2 or lower were considered low risk, and those with a score of 3 or higher were considered high risk. Discrimination was measured using the time-dependent C statistic with bootstrapped CIs, and calibration was evaluated by plotting predicted vs empirical risk in numeric risk score groups. Subgroup analyses by age, sex, and race/ethnicity were performed to ensure generalizability across appropriate populations. As an additional point of reference, this analysis was repeated with the Khorana score, and reclassification tables were used for comparison between scores. For the Khorana score, patients with a score of 1 or lower were considered low risk, and those with a score of 2 or higher were considered high risk. A complete case analysis was performed, and patients with missing covariates were excluded from the analysis.

3 | RESULTS

Between 2014 and 2022, 10,312 patients received a new invasive cancer diagnosis at BMC, of whom 2103 received first-line systemic therapy at BMC within 1 year of their first cancer diagnosis and met other inclusion/exclusion criteria (Figure 1). There was an even distribution of patients by sex (50.9% female), with a median patient age of 61 years. Most patients were non-Hispanic Black (45.6%), followed by non-Hispanic White (30.2%), Hispanic (18.2%), and then non-Hispanic Asian Pacific Islander (6.0%). The median time to treatment was 55 days, with 67.4% of patients receiving chemotherapy, 21.2% endocrine therapy, 9.5% targeted therapy, and 1.9% immune checkpoint inhibitors as first-line treatment. The most prevalent cancer diagnosis was breast (21.9%), followed by lung (11.7%) and prostate (10.6%). A total of 54.3% of patients had an advanced cancer stage of III/IV, 11.8% had a pretherapy body mass index $\geq$ 35 kg/m^2, 13.6% had a pretherapy white blood cell count >11,000 /mm^3, 18.7% had pretherapy hemoglobin levels <10 mg/dL, and 21.6% had a pretherapy platelet count $\geq$ 350,000/mL. As for other risk predictors, 1.7% of patients had a lifetime history of VTE, 0.2% had a history of paralysis or immobility in the 12 months prior to the index date, and 33.3% of patients had a hospitalization lasting >3 days in the 3 months prior to the index date (Table 1) [30, 31]. The median follow-up time was 222 days (IQR, 179-289), and 29 patients died during follow-up.

The EHR-CAT score [22] was calculated for each patient (Supplementary Table S3), and patients were grouped together into low- and high-risk categories. The cumulative incidence of overall VTE for the low-risk group (n = 1271) was 5.0%, and for the high-risk group (n = 832) was 13.3%, with a C statistic of 0.67 (95% CI, 0.63-0.71; Table 2). The cumulative incidence for PE/LE-DVT was 3.4% and 10.8% for the low- and high-risk groups, respectively, with a C statistic of 0.68 (95% CI, 0.63-0.72). To facilitate the comparison of this newly expanded score to a current clinical tool, we then calculated the Khorana score [24] for our patients. The cumulative incidence of overall VTE for the Khorona et al. [24] stratified low-risk group was 7.4% (n = 1484), and the high-risk group (n = 619) was 10.6%, with a C statistic of 0.58 (95% CI, 0.54-0.63). The Khorana et al. [24] group incidences for PE/LE-DVT were 5.8% and 7.8% for low and high risk, respectively, with a C statistic of 0.59 (95% CI, 0.54-0.64; Table 2).

The 2 scores concordantly classified 80% of the patients, while the remaining 20% were differentially classified (Table 3). In contrast to results obtained with the Khorana score, EHR-CAT placed 108

TABLE 1 Cohort characteristics with risk assessment model predictor prevalence and point values.

Patient characteristics	BMC cohort, *n* (%)	EHR-CAT	Khorana score
Age, y, median (IQR)	61 (53-69)		
Sex			
Male	1033 (49.1)		
Female	1070 (50.9)		
Race/ethnicity[a]			
Non-Hispanic White	636 (30.2)		
Non-Hispanic Black	958 (45.6)		
Hispanic	382 (18.2)		
Non-Hispanic Asian Pacific Islander	127 (6.0)	−1	
Cancer type[b]			
Breast	461 (21.9)		
Prostate	222 (10.6)		
Lung	247 (11.7)	+2	+1
Lower GI/colorectal	166 (7.9)	+1	
Upper GI/gastric and esophageal	85 (4.0)	+3	+2
Liver	69 (3.3)		
Biliary and gallbladder	28 (1.3)	+3	
Pancreas	61 (2.9)	+3	+2
Head and neck	163 (7.8)		
Cervical	43 (2.0)		+1
Ovarian	20 (1.0)	+2	+1
Uterine	35 (1.6)	+2	+1
Other gynecologic	13 (0.7)		+1
Bladder	25 (1.2)	+2	+1
Kidney	21 (1.0)	+2	+1
Testicular	10 (0.5)	+2	+1
Sarcoma	19 (0.9)	+2	
Brain (CNS)	20 (1.0)	+2	
Acute lymphocytic leukemia	2 (0.1)	+2	
Acute myeloid leukemia	28 (1.3)		
Hodgkin's and indolent non-Hodgkin lymphoma[c]	76 (3.6)		
Aggressive non-Hodgkin lymphoma[d]	71 (3.4)	+2	+1
Leukemia[e]	63 (3)		
Multiple myeloma	88 (4.2)	+2	
Other solid tumors	67 (3.2)		
Pretherapy BMI			
BMI, median (IQR)	27 (23-31)		
BMI ≥ 35	249 (11.8)	+1	+1
Pretherapy WBC count			
WBC, median (IQR)	7.3 (5.7-9.4)		
WBC > 11	285 (13.6)	+1	+1

(Continues)

TABLE 1 (Continued)

Patient characteristics	BMC cohort, *n* (%)	EHR-CAT	Khorana score
Pretherapy Hb			
Hb, median (IQR)	12.3 (10.6-13.5)		
Hb < 10	394 (18.7)	+1	+1
Pretherapy Plt			
Plt, median (IQR)	265 (207-337)		
Plt ≥ 350	454 (21.6)	+1	+1
No. of days labs preceded treatment, median (IQR)	14 (34-6)		
Cancer stage			
1	353 (16.8)		
2	445 (21.2)		
3	468 (22.3)	+1	
4	673 (32.0)	+1	
Not applicable[f]	164 (7.8)		
Days to treatment initiation, median (IQR)	55 (26-103)		
Days of patient follow-up, median (IQR)	222 (179-289)		
First treatment regimen received[g]			
Chemotherapy	1418 (67.4)		
Immune checkpoint inhibitors	40 (1.9)		
Targeted therapy	200 (9.5)	−1	
Endocrine therapy	445 (21.2)	−1	
Comorbidities and laboratory values			
History of VTE[h]	36 (1.7)	+1	+1
Hospitalization >3 d in the past 3 mo	700 (33.3)	+1	+1
Congestive heart failure	96 (4.6)		
Cardiac arrhythmia	669 (31.8)		
Cardiac vascular disease	5 (0.2)		
Peripheral vascular disease	54 (2.6)		
Chronic obstructive pulmonary disease	64 (3.0)		
Paralysis or immobility	4 (0.2)	+1	+1
Diabetes	234 (11.1)		
Renal disease	29 (1.4)		
Liver disease	118 (5.6)		
Rheumatologic disease	6 (0.3)		
History of myocardial infarction	72 (3.4)		

BMC, Boston Medical Center; BMI, body mass index; CNS, central nervous system; EHR-CAT, electronic health record cancer-associated thrombosis; GI, gastrointestinal; Hb, hemoglobin; Plt, platelet; VTE, venous thromboembolism; WBC, white blood cell.

[a]Non-Hispanic Asian Pacific Islander category includes Native Americans, Alaskan Natives, Asian Pacific Islanders, and other people of Asian descent.

[b]Cancer type was defined by Internal Classification of Disease-O-3 site and histology using the World Health Organization 2008 Cancer Classifications for common solid tumors [30] and Rare Cancer Classifications for sarcoma, neuroendocrine cancers, and hematologic malignancies [31].

[c]Indolent non-Hodgkin lymphoma (NHL) includes chronic lymphatic leukemia/small lymphocytic lymphoma and follicular and mantel cell lymphoma.

[d]Aggressive non-Hodgkin lymphoma includes diffuse large B cell lymphoma, Burkitt lymphoma, and T and natural killer cell lymphomas.

[e]Leukemia includes acute/chronic leukemia and myelodysplastic syndrome.

[f]Not applicable cancer stage was considered acceptable only if the patient's cancer type was brain/CNS, leukemia, or myeloma.

[g]Refer to Li et al. [22] for a list of systemic cancer therapy classifications.

[h]History of VTE and paralysis were defined by Internal Classification of Disease codes listed in Supplementary Table S2.

TABLE 2 Patient high-low score groups.

Risk score	Classification	*n*	Overall VTE incidence at 6 mo, % (*n*)	C statistic (95% CI)	PE/LE-DVT incidence at 6 mo, % (*n*)	C statistic (95% CI)
EHR-CAT (0-5+)	Low risk (2–)	1271	5 (52)	0.67 (0.63-0.71)	3.4 (34)	0.68 (0.63-0.72)
	High risk (3+)	832	13.3 (98)		10.8 (76)	
Khorana score (0-3+)	Low risk (1–)	1484	7.4 (92)	0.58 (0.54-0.63)	5.8 (69)	0.59 (0.54-0.64)
	High risk (2+)	619	10.6 (58)		7.8 (41)	

EHR-CAT, electronic health record cancer-associated thrombosis; LE-DVT, lower extremity deep vein thrombosis; PE, pulmonary embolism; VTE, venous thromboembolism.

patients with a cumulative VTE incidence of 3.2% into the low-risk group and 321 patients with a cumulative VTE incidence of 15.3% into the high-risk group. EHR-CAT performed well across subgroups of age, sex, and race/ethnicity (Supplementary Table S4; C statistics 0.63-0.74 vs 0.51-0.65 for the Khorana score), and a calibration plot shows a good model fit in the BMC data (Supplementary Figure S1).

4 | DISCUSSION

Several RAMs and accompanying scores have been developed to identify cancer patients who are at an increased risk for VTE [14,22,24,32–35]. Unfortunately, many of these models perform poorly in external validation and remain limited for solid malignancies and patients receiving chemotherapy [14,32,33,35–37]. The novel EHR-CAT score developed by Li et al. [22,23] overcomes these limitations and has shown success in stratifying VTE risk in various diverse cohorts now encompassing a total of 112,531 patients. The characteristics of the patients within the 4 cohorts utilized to date encompass significant cancer, medical history, regional, and racial/ethnic breadth [22,23].

In this external validation, the EHR-CAT score accurately stratified 2103 patients from a regionally distinct safety-net hospital into high- and low-risk groups, with a C statistic of 0.67 for VTE and 0.68 for PE/LE-DVT. This C statistic was greater than the C statistic of 0.58 for VTE and 0.59 for PE/LE-DVT obtained with the Khorana score. This reflects the improved discrimination power of this score, visually represented in Figure 2. These values are very similar to the VTE and PE/LE-DVT C statistics obtained in the derivation cohort (0.71 and 0.72), the external validation in the Veterans Affairs healthcare system cohort (0.68 and 0.68), and the external validation in the MD Anderson cancer center cohort (0.71 and NA). The incorporation of additional predictors beyond those utilized by the current clinical standard facilitated the score's improved performance, reported here and in prior applications [22,23].

Our study has limitations. First, our study's cohort sample size is relatively small compared with the total number of patients diagnosed with cancer at BMC (Figure 1). This is primarily due to the fact that a substantial number of patients were excluded because they did not begin systemic anticancer treatment or lost to follow-up shortly after cancer diagnosis within 1 year of diagnosis at BMC. These excluded patients include many whose diagnosis did not require immediate systemic therapy, such as those with early-stage cancer or cancer types that do not require immediate systemic therapy (Supplementary Table S5), and possibly some that were referred elsewhere for treatment. Note, the percentage of BMC patients who received systemic treatment within 1 year of diagnosis is consistent with data on similar cohorts in other studies reporting frequency of systemic treatment initiation at 27% within 90 days of diagnosis [38] or 59% within 1 year of diagnosis [39]. We acknowledge that the exclusion of patients who do not meet the criteria to initiate systemic treatment for cancer limits our understanding of how these scores would perform in other settings (eg, patients receiving later lines of treatment and patients not initially treated with systemic therapy), and risk-stratification in these patients is of interest in future work. Furthermore, the sample size of the final cohort was still amply adequate to demonstrate that the EHR-CAT score offers improved stratification of patients into high- and low-risk groups compared with the Khorana score in this independent setting with a highly diverse patient population from a

TABLE 3 Patient reclassification table.

Category	Khorana score	EHR-CAT	*n*	Overall VTE incidence at 6 mo, % (*n*)	C statistic comparison	PE/LE-DVT incidence at 6 mo, % (*n*)	C statistic comparison
Concordant (80%)	Low risk	Low risk	1163	5.1 (49)	0.58 (Khorana score) vs 0.67 (EHR-CAT)	3.6 (33)	0.59 (Khorana score) vs 0.68 (EHR-CAT)
	High risk	High risk	511	12.1 (55)		9.1 (40)	
Reclassified (20%)	Low risk	High risk	321	15.3 (43)		13.3 (36)	
	High risk	Low risk	108	3.2 (3)		1.1 (1)	
Total	All	All	2103	8.4 (150)		6.4 (110)	

EHR-CAT, electronic health record cancer-associated thrombosis; LE-DVT, lower extremity deep vein thrombosis; PE, pulmonary embolism; VTE, venous thromboembolism.

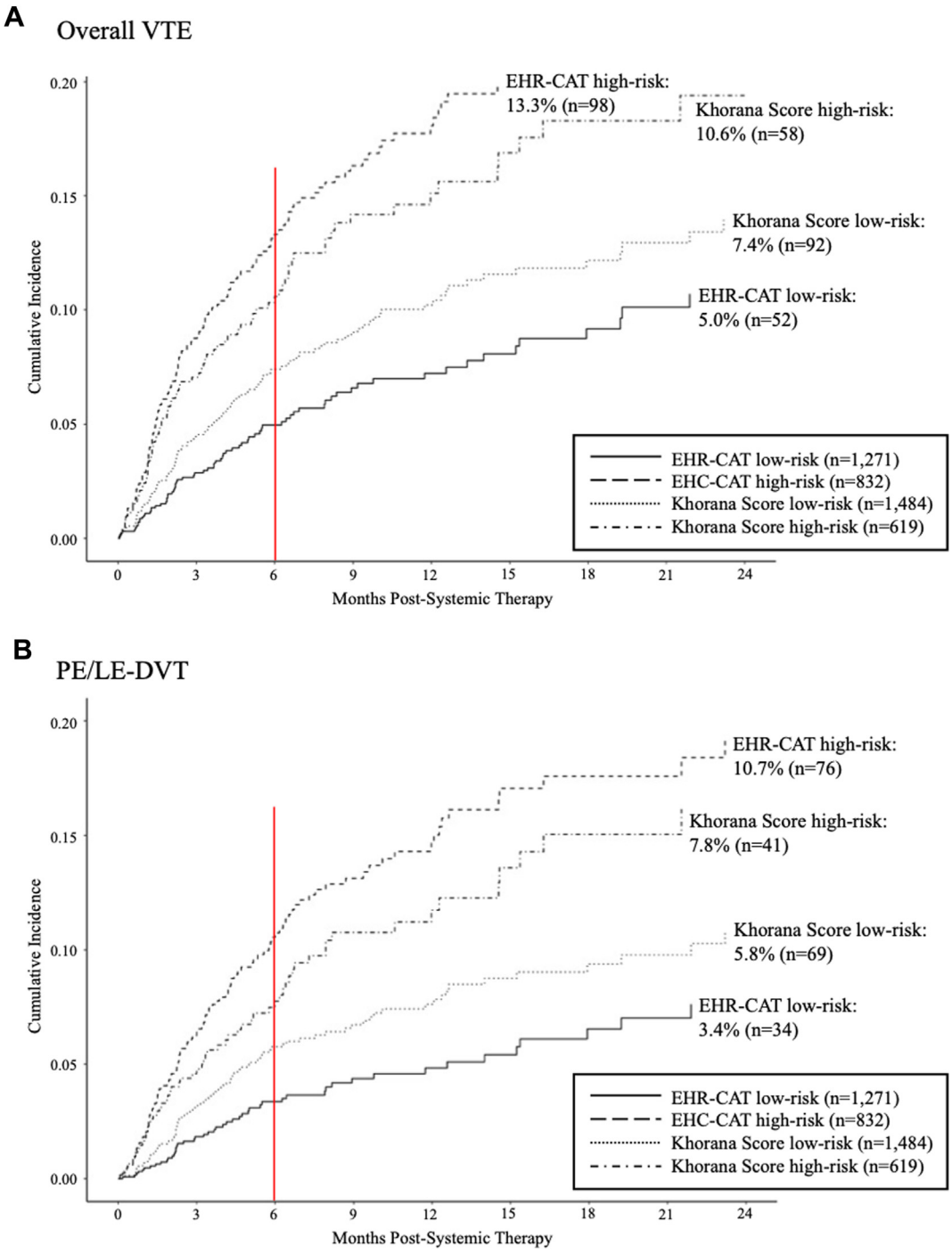

FIGURE 2 Outcome cumulative incidence plots by group. (A) Cumulative incidence competing risk curves show venous thromboembolism (VTE) incidence over time following the initiation of systemic therapy for score-defined patient groups. The vertical line draws attention to the 6-month values reported in Table 2, with an electronic health record cancer-associated thrombosis (EHR-CAT) C statistic of 0.68 and a Khorana score C statistic of 0.59. (B) This analysis was repeated for the secondary outcome, pulmonary embolism/lower extremity deep vein thrombosis (PE/LE-DVT). The EHR-CAT C statistic for PE/LE-DVT at 6 months was 0.69 vs 0.59 for this outcome with the Khorana score.

safety-net hospital. Second, our cohort has a lower proportion of patients receiving immune checkpoint inhibitors and targeted therapy compared with national trends [40,41], likely related to the challenge of therapy access in an underserved patient population, and further work could be necessary to definitively evaluate performance in cohorts receiving a higher proportion of contemporary anticancer therapies. Third, since the VTE outcome was ascertained from EHR data, it is possible we misidentified some events. However, we addressed this limitation by using a sensitive algorithm [22] that looks for outcomes both using medical codes and by parsing text in clinical notes, followed by full manual adjudication of all putative events in order to ensure specificity. In addition, we censored at the end of continuous follow-up in BMC. Nevertheless, it is

still possible that some events occurring outside BMC may be missed, which is an issue inherent to any analysis based on a single institution's EHRs. Fourth, mortality data are incomplete at BMC. Death is not the outcome of this study, but the incompleteness of mortality data will result in some patients being coded as lost to follow-up rather than as having experienced the competing risk of death. Finally, the study cohort includes patients with lymphoma who have unique risk predictors that are not fully accounted for in the EHR-CAT score [42]. However, we saw no significant difference when our validation of EHR-CAT was repeated after the exclusion of patients with acute or chronic lymphocytic leukemia as well as those with acute or chronic myeloid leukemia (n = 2019; VTE C statistic, 0.67; PE-DVT C statistic, 0.68; data not shown).

Future work is possible in several directions. First, although EHR-CAT represents a substantial improvement in the state-of-the-art, the absolute percentage of high-risk patients who experience VTE is still relatively low, and future work is of interest to gain further improvements in accuracy, especially in later lines of therapy settings. Second, future work can incorporate patient bleeding risk into the risk calculation, as bleeding risk is a critical consideration for CAT prophylaxis. Finally, the construction and incorporation of these risk decision support tools in EHRs is an important next step for ease of clinical use. Toward that end, a Shiny application for EHR-CAT is available for clinician use [43].

Cancer-associated risk assessment scores are widely used in the clinic to identify patients at an increased risk of experiencing a VTE event. Our work, along with other recently published studies [22,23], support the enhanced performance of the EHR-CAT score and further endorse it for clinical consideration.

ACKNOWLEDGMENTS

We would like to thank Linda Rosen at Boston Medical Center for her assistance in pulling electronic health record data.

FUNDING

This research was, in part, funded by the American Heart Association (870726 & 857078; K.N.D., J.L., N.R.F., K.R., V.C.C., S.L., and A.J.), by the Veterans Affairs cooperative studies program for providing core funding for the Massachusettes Veterans Epidemiology Research and Informatics Center work (J.L., N.R.F., N.V.D., M.T.B., and J.M.G.), and by the Aim-Ahead National Institutes of Health (NIH) grant (1OT2OD032581-02-239; J.L. and A.L.). S.L. was also funded through the Multidiciplinary Training Program in Cardiovascular Epidemiology (T32HL125232). A.L. was also funded by the NIH (NHLBI K23 HL159271), NIH Agreement (3OT2OD032581-01S1), and the Cancer Prevention and Research Institute of Texas (CPRIT; RR190104).

AUTHOR CONTRIBUTIONS

K.N.D. and J.L. performed analysis. N.R.F., A.L., K.R., V.C.C., N.V.D., M.T.B., and J.M.G. assisted in the conceptual study design. V.C.C., K.R., N.V.D., M.T.B., J.M.G., and N.R.F. provided funding for this work. J.L., A.L., and N.R.F. provided essential knowledge, direction, and feedback for this project. S.L. and V.C.C. provided patient chart review. K.N.D. wrote the manuscript. All authors provided manuscript feedback.

RELATIONSHIP DISCLOSURE

The authors K.N.D., S.L., A.J., V.C.C., K.R., and A.L. have no conflict of interest to declare. J.L., J.M.G., and N.R.F. have research funding from Merck.

ORCID

Karlynn N. Dulberger https://orcid.org/0000-0002-3481-813X
Nathanael R. Fillmore https://orcid.org/0000-0002-8058-3423

REFERENCES

[1] Sørensen HT, Mellemkjær L, Olsen JH, Baron JA. Prognosis of cancers associated with venous thromboembolism. *N Engl J Med*. 2000;343:1846–50.
[2] Ay C, Pabinger I, Cohen AT. Cancer-associated venous thromboembolism: burden, mechanisms, and management. *Thromb Haemost*. 2017;117:219–30.
[3] Khorana AA, Francis CW, Culakova E, Kuderer NM, Lyman GH. Thromboembolism is a leading cause of death in cancer patients receiving outpatient chemotherapy. *J Thromb Haemost*. 2007;5:632–4.
[4] Lyman GH, Eckert L, Wang Y, Wang H, Cohen A. Venous thromboembolism risk in patients with cancer receiving chemotherapy: a real-world analysis. *Oncologist*. 2013;18:1321–9.
[5] Tick LW, Kramer MH, Rosendaal FR, Faber WR, Doggen CJ. Risk factors for post-thrombotic syndrome in patients with a first deep venous thrombosis. *J Thromb Haemost*. 2008;6:2075–81.
[6] Naess IA, Christiansen SC, Romundstad P, Cannegieter SC, Rosendaal FR, Hammerstrøm J. Incidence and mortality of venous thrombosis: a population-based study. *J Thromb Haemost*. 2007;5:692–9.
[7] Spencer FA, Lessard D, Emery C, Reed G, Goldberg RJ. Venous thromboembolism in the outpatient setting. *Arch Intern Med*. 2007;167:1471–5.
[8] Braekkan SK, Borch KH, Mathiesen EB, Njølstad I, Wilsgaard T, Hansen JB. Body height and risk of venous thromboembolism: the Tromsø study. *Am J Epidemiol*. 2010;171:1109–15.
[9] Gussoni G, Frasson S, La Regina M, Di Micco P, Monreal M, RIETE Investigators. Three-month mortality rate and clinical predictors in patients with venous thromboembolism and cancer. Findings from the RIETE registry. *Thromb Res*. 2013;131:24–30.
[10] Heit JA. Epidemiology of venous thromboembolism. *Nat Rev Cardiol*. 2015;12:464–74.
[11] Imberti D, Agnelli G, Ageno W, Moia M, Palareti G, Pistelli R, et al. Clinical characteristics and management of cancer-associated acute venous thromboembolism: findings from the MASTER Registry. *Haematologica*. 2008;93:273–8.
[12] Timp JF, Braekkan SK, Versteeg HH, Cannegieter SC. Epidemiology of cancer-associated venous thrombosis. *Blood*. 2013;122:1712–23.
[13] Blom JW, Doggen CJM, Osanto S, Rosendaal FR. Malignancies, prothrombotic mutations, and the risk of venous thrombosis. *JAMA*. 2005;293:715–22.
[14] Ay C, Dunkler D, Marosi C, Chiriac AL, Vormittag R, Simanek R, et al. Prediction of venous thromboembolism in cancer patients. *Blood*. 2010;116:5377–82.
[15] Cronin-Fenton DP, Søndergaard F, Pedersen LA, Fryzek JP, Cetin K, Acquavella J, et al. Hospitalisation for venous thromboembolism in cancer patients and the general population: a population-based

cohort study in Denmark, 1997-2006. *Br J Cancer*. 2010;103:947–53.

[16] Sørensen HT, Mellemkjaer L, Steffensen FH, Olsen JH, Nielsen GL. The risk of a diagnosis of cancer after primary deep venous thrombosis or pulmonary embolism. *N Engl J Med*. 1998;338:1169–73.

[17] Collins R, Scrimgeour A, Yusuf S, Peto R. Reduction in fatal pulmonary embolism and venous thrombosis by perioperative administration of subcutaneous heparin. Overview of results of randomized trials in general, orthopedic, and urologic surgery. *N Engl J Med*. 1988;318:1162–73.

[18] Lau BD, Haut ER. Practices to prevent venous thromboembolism: a brief review. *BMJ Qual Saf*. 2014;23:187–95.

[19] Schünemann HJ, Cushman M, Burnett AE, Kahn SR, Beyer-Westendorf J, Spencer FA, et al. American Society of Hematology 2018 guidelines for management of venous thromboembolism: prophylaxis for hospitalized and nonhospitalized medical patients. *Blood Adv*. 2018;2:3198–225.

[20] Nicholson M, Chan N, Bhagirath V, Ginsberg J. Prevention of venous thromboembolism in 2020 and beyond. *J Clin Med*. 2020;9:2467. https://doi.org/10.3390/jcm9082467

[21] Prandoni P, Lensing AW, Piccioli A, Bernardi E, Simioni P, Girolami B, et al. Recurrent venous thromboembolism and bleeding complications during anticoagulant treatment in patients with cancer and venous thrombosis. *Blood*. 2002;100:3484–8.

[22] Li A, La J, May SB, Guffey D, da Costa Jr WL, Amos CI, et al. Derivation and validation of a clinical risk assessment model for cancer-associated thrombosis in two unique US health care systems. *J Clin Oncol*. 2023;41:2926–38.

[23] Li A, De Las Pozas G, Andersen CR, Nze CC, Toale KM, Milner EM, et al. External validation of a novel electronic risk score for cancer-associated thrombosis in a comprehensive cancer center. *Am J Hematol*. 2023;98:1052–7.

[24] Khorana AA, Kuderer NM, Culakova E, Lyman GH, Francis CW. Development and validation of a predictive model for chemotherapy-associated thrombosis. *Blood*. 2008;111:4902–7.

[25] Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014;14:40. https://doi.org/10.1186/1471-2288-14-40

[26] Mahar AL, Compton C, McShane LM, Halabi S, Asamura H, Rami-Porta R, et al. Refining prognosis in lung cancer: a report on the quality and relevance of clinical prognostic tools. *J Thorac Oncol*. 2015;10:1576–89.

[27] Mallett S, Royston P, Waters R, Dutton S, Altman DG. Reporting performance of prognostic models in cancer: a review. *BMC Med*. 2010;8:21. https://doi.org/10.1186/1741-7015-8-21

[28] Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*. 2012;98:691–8.

[29] Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol*. 2009;170:244–56.

[30] National Cancer Institute. Site recode ICD-O-3/WHO 2008 definition. https://seer.cancer.gov/.siterecode/icdo3_dwhoheme/index.html. [accessed June 14, 2024].

[31] National Cancer Institute. Rare cancer classification. https://seer.cancer.gov/seerstat/variables/seer/raresiterecode/. [accessed June 14, 2024].

[32] Cella CA, Di Minno G, Carlomagno C, Arcopinto M, Cerbone AM, Matano E, et al. Preventing venous thromboembolism in ambulatory cancer patients: the ONKOTEV study. *Oncologist*. 2017;22:601–8.

[33] Gerotziafas GT, Taher A, Abdel-Razeq H, AboElnazar E, Spyropoulos AC, El Shemmari S, et al. A predictive score for thrombosis associated with breast, colorectal, lung, or ovarian cancer: the prospective COMPASS–cancer-associated thrombosis study. *Oncologist*. 2017;22:1222–31.

[34] Pabinger I, van Es N, Heinze G, Posch F, Riedl J, Reitter EM, et al. A clinical prediction model for cancer-associated venous thromboembolism: a development and validation study in two independent prospective cohorts. *Lancet Haematol*. 2018;5:e289–98. https://doi.org/10.1016/S2352-3026(18)30063-2

[35] Verso M, Agnelli G, Barni S, Gasparini G, LaBianca R. A modified Khorana risk assessment score for venous thromboembolism in cancer patients receiving chemotherapy: the Protecht score. *Intern Emerg Med*. 2012;7:291–2.

[36] van Es N, Di Nisio M, Cesarman G, Kleinjan A, Otten HM, Mahé I, et al. Comparison of risk prediction scores for venous thromboembolism in cancer patients: a prospective cohort study. *Haematologica*. 2017;102:1494–501.

[37] Mulder FI, Horváth-Puhó E, van Es N, van Laarhoven HWM, Pedersen L, Moik F, et al. Venous thromboembolism in cancer patients: a population-based cohort study. *Blood*. 2021;137:1959–69.

[38] Martens KL, Li A, La J, May SB, Swinnerton KN, Tosi H, et al. Epidemiology of cancer-associated venous thromboembolism in patients with solid and hematologic neoplasms in the Veterans Affairs Health Care System. *JAMA Netw Open*. 2023;6:e2317945. https://doi.org/10.1001/jamanetworkopen.2023.17945

[39] da Costa WL Jr, Guffey D, Oluyomi A, Bandyo R, Rosales O, Wallace CD, et al. Patterns of venous thromboembolism risk, treatment, and outcomes among patients with cancer from uninsured and vulnerable populations. *Am J Hematol*. 2022;97:1044–54.

[40] Prasad V, Haslam A, Olivier T. Updated estimates of eligibility and response: immune checkpoint inhibitors [abstract]. *J Clin Oncol*. 2024;42:e14613. https://doi.org/10.1200/JCO.2024.42.16_suppl.e14613

[41] Haslam A, Kim MS, Prasad V. Updated estimates of eligibility for and response to genome-targeted oncology drugs among US cancer patients, 2006-2020. *Ann Oncol*. 2021;32:926–32.

[42] Ma S, Swinnerton KN, La J, Guffey D, Bandyo R, Amos CI, et al. Development and external validation of a lymphoma-specific venous thromboembolism risk assessment model. *Blood*. 2023;142:565. https://doi.org/10.1182/blood-2023-184770

[43] Baylor College of Medicine. EHR-CAT risk assessment model for venous thromboembolism (VTE) in treated cancer patients. https://li-lab-bcm.shinyapps.io/EHR-CAT/; 2023 [accessed September 5, 2024].

SUPPLEMENTARY MATERIAL

The online version contains supplementary material available at https://doi.org/10.1016/j.rpth.2024.102650