

Published in final edited form as:

Lancet Haematol. 2018 July 01; 5(7): e289–e298. doi:10.1016/S2352-3026(18)30063-2.

Development and validation of a clinical prediction model for cancer-associated venous thromboembolism in two independent prospective cohorts

Ingrid Pabinger, MD¹, Nick van Es, MD², Georg Heinze, PhD³, Florian Posch, MD MSc^{1,4}, Julia Riedl, MD PhD¹, Eva-Maria Reitter, MD PhD¹, Marcello Di Nisio, MD PhD⁵, Gabriela Cesarman-Maus, MD PhD⁶, Noémie Kraaijpoel, MD², Christoph C. Zielinski, MD⁷, Harry R. Büller, MD PhD², Cihan Ay, MD¹

¹Clinical Division of Haematology and Haemostaseology; Department of Medicine I; Medical University of Vienna; Vienna, Austria ²Department of Vascular Medicine; Academic Medical Center Amsterdam; University of Amsterdam; Amsterdam, The Netherlands ³Section for Clinical Biometrics; Center for Medical Statistics, Informatics and Intelligent Systems; Medical University of Vienna; Vienna, Austria ⁴Division of Oncology; Department of Internal Medicine; Medical University of Graz; Graz, Austria ⁵Department of Medicine and Ageing Sciences, G. D'Annunzio University, Chieti, Italy ⁶Department of Hematology, National Cancer Institute Mexico, Mexico City, Mexico ⁷Clinical Division of Oncology; Department of Medicine I; Medical University of Vienna; Vienna, Austria

Abstract

Background—Venous thromboembolism (VTE) is a frequent complication of cancer, however the risk is highly variable among individuals depending on various factors, including types of cancer. To enable a personalized risk prediction of VTE we developed and externally validated a clinical prediction model for cancer-associated VTE.

Methods—The prospective Vienna Cancer and Thrombosis Study (CATS, n=1,423) was used for model development, and the prospective Multinational cohort study to Identify Cancer patients at risk of VTE (MICA, n=832) was used for external validation. Primary outcome was objectively confirmed VTE at 6 months. The cumulative 6-month VTE risk was 5.7% in CATS (95% CI: 4.5-6.9), and 6.3% (95%CI: 4.7-8.2) in MICA. Tumor sites were categorized into low/intermediate, high, and very high VTE categories. Predictive variables were selected from a broad set of clinical and laboratory factors.

Address for correspondence: Ingrid Pabinger, MD, Clinical Division of Haematology and Haemostaseology, Department of Medicine I, Medical University of Vienna, Währinger Gürtel 18-20, 1090 Vienna – Austria, Tel: +43-1-40400-44480, ingrid.pabinger@meduniwien.ac.at.

The authors have no conflicting interests to declare.

Ethics: Both the CATS and the MICA studies were approved by the local ethics committees of the participating hospitals.

Author contributions according to ICMJE standards: Design of the study: IP (CATS), NvE & HRB (MICA); Contributed patients: FP & JR & EMR & CA (CATS), NvE & MdN & GCM & NK (MICA); Inter-study project coordination: FP; Statistical analysis: NvE (MICA), GH FP (CATS); First draft of the manuscript: IP NvE FP; Critical revision of the draft for important intellectual content: All authors; Agree with the manuscript's conclusions: All authors; Final approval of the version to be published: All authors.

Findings—The final prediction model included two variables: tumor site category (hazard ratio for “high” vs. “low/intermediate”, and “very high” versus “high” VTE-risk tumor site=1.96 (95% CI: 1.41-2.72, p=0.0001). and D-Dimer (hazard ratio per doubling=1.32, 95% CI: 1.12-1.56, p=0.001). The C-Indices of the model were 0.66 (95%: 0.63-0.67) in internal validation (CATS), and 0.68 (95%: 0.62-0.74) in external validation (MICA), respectively. The clinical prediction model was adequately calibrated in both cohorts.

Interpretation—An externally-validated clinical prediction model incorporating only one clinical factor (tumor site category) and one biomarker (D-Dimer) predicts the risk of VTE in ambulatory patients with solid cancers. This simple model considerably improves on previous models for predicting cancer-associated VTE, and can aid physicians in selecting patients who will likely benefit from thromboprophylaxis.

Funding—Austrian Science Fund, Austrian National Bank Memorial Fund, Unrestricted grants from participating hospitals

Introduction

Venous thromboembolism (VTE), including both, deep vein thrombosis or/and pulmonary embolism, is a burdensome complication of malignancy with an incidence varying from 1% to 20%.¹ Pharmacological thromboprophylaxis has the potential to reduce the burden of this illness in the cancer population.¹

Several randomized trials have consistently demonstrated that prophylactic anticoagulation with low-molecular-weight heparin approximately halves the relative risk of VTE in cancer patients.² However, the absolute risk reduction of this intervention appears to be modest for the majority of unselected ambulatory cancer patients, in whom the risk of VTE is about 3% to 5% in the first months of chemotherapy.² Moreover, anticoagulant-related bleeding complications are frequent in cancer patients.³ Thus, the decision to provide anticoagulation for prevention of cancer-associated VTE should ideally be informed by a valid risk stratification strategy.⁴ With such a personalized approach, thromboprophylaxis could be provided to cancer patients with the highest risk of VTE, while avoiding the burden and risks of anticoagulation in low risk patients.

The most widely used clinical prediction model for this purpose is a score proposed by Khorana and colleagues, which aims to identify ambulatory cancer patients at increased risk of VTE during chemotherapy using two clinical (tumour site and body mass index) and three laboratory (platelets, haemoglobin and leucocytes) variables.^{5,6} Other scores have also been proposed, including the Vienna modification of the Khorana score (addition of biomarkers D-Dimer and soluble (s) P-selectin).⁷ the PROTECHT score (addition of gemcitabine and platinum based chemotherapy),⁸ and the CONKO score (addition of World Health Organisation performance status).⁹ However, in a recent prospective validation study, which included 70% of patients during chemotherapy and only 30% at diagnosis of cancer, only two scoring approaches (the Vienna modification and the PROTECHT score) were found to be predictive for VTE occurrences.¹⁰ A more recently described score investigated only breast, colon, lung and ovary cancer and included cancer and treatment related factors,

showing good discriminatory capacity. However, this score is restricted to the specific cancer types and remains to be externally validated.¹¹

It is widely accepted that tumour sites can be highly associated with VTE.¹ In addition, biomarkers reflecting activation of the hemostatic system, such as D-Dimer, thrombin generation, and soluble P-Selectin, have been shown to be independent prognostic factors for VTE in cancer patients.^{7,12} These markers can facilitate and potentially enhance the clinical prediction of cancer-associated VTE.⁷ However, given the limited availability of most of these biomarkers in clinical practice, their adoption for clinical validation has been limited.¹²

In this study, we aimed to address some of these issues by developing and externally validating a clinical prediction model for VTE in ambulatory patients with active solid cancers. We designed a simple VTE prediction model that provides valuable estimates of thrombosis risk over a time frame of 6 months, and furthermore, might be easily applied to a routine clinical setting. Having such a model available would allow targeted thromboprophylaxis in high risk patients.

Methods

Study Design and participants

Two independent prospective cohorts were used for the development and external validation of the clinical VTE prediction model in this study, namely, the Vienna Cancer and Thrombosis Study (CATS),¹³ and the Multinational Cohort Study to Identify Cancer Patients at High Risk of Venous Thromboembolism (MICA,¹⁰). Both cohorts were specifically designed for the purpose of VTE risk factor identification in cancer patients. This report adheres to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines (see Appendix **page 2-3-2**).¹⁴

CATS is an ongoing, prospective, single-center, observational cohort study with a baseline biobank (EK 126/2003). Patients with newly-diagnosed active malignancy or disease progression after complete or partial remission were enrolled at a single tertiary academic center in Vienna, Austria. Detailed inclusion and exclusion criteria have been presented in previous reports.^{7,15} Inclusion criteria were histologic confirmation of diagnosis, age more than 18 years, willingness to participate and written informed consent. Exclusion criteria were overt bacterial or viral infection within the last 2 weeks, venous or arterial thromboembolism within the last 3 months, and continuous anticoagulation with vitamin K antagonists or low molecular weight heparin (LMWH) or direct oral anticoagulants. Patients were allowed to take aspirin, ticlopidine, or clopidogrel, and immobilized patients were treated with LMWH as thrombosis prophylaxis during their hospital stay. Further exclusion criteria were surgery or radiotherapy within the last 2 weeks and chemotherapy within the last 3 months to exclude a transient influence of these interventions on the hemostatic system. Patients were followed until VTE, death, or censoring within an observation period of 24 months. Data from 1,737 patients included in the CATS study between October 2003 and March 2014 were considered for the development of the VTE clinical prediction model. For the model we included patients with solid cancer (excluding primary brain tumors) and

lymphoma, if classifiable by Ann Arbor. Of the initially 1,737 patients, 305 were excluded due to the following reasons: n=240 because of primary brain tumors, n=16 because D-Dimer was not available, n=5 because of lymphoma not stageable by Ann Arbor, n=39 because of loss of follow up, and n=5 because data review revealed a secondary primary malignancy making it unable to assign the patient to a single tumor site category. Cancer types of the patients are shown in detail in table 1 (patient characteristics).

Patients were followed until VTE, death, or censoring within an observation period of 24 months.

Demographic, laboratory, and outcome data obtained in MICA (enrolment between July 2008 and February 2016) were used for external validation of the prediction model. MICA (#2010.259) is a completed prospective, multinational, observational cohort study with baseline blood sampling in which patients with advanced solid cancer were enrolled from seven centers in the Netherlands, France, Italy, and Mexico. Ambulatory patients with lung, esophageal, colorectal, pancreatic, breast, prostate, gastric, ovarian, or bladder cancer were eligible if they were scheduled for chemotherapy within 7 days or had started chemotherapy in the previous 3 months. Exclusion criteria included prophylactic or therapeutic anticoagulation or adjuvant therapy. Patients in MICA were included before and during chemotherapy. Patients were followed for a maximum of 6 months until the occurrence of VTE, death, censoring due to curative surgery (only for patients receiving neoadjuvant therapy), start of anticoagulation for other reasons, or loss-to-follow-up. There were 1,027 patients eligible of whom 12 were excluded because of prophylactic or therapeutic anticoagulation, 14 because they had a hematological malignancy (other than lymphoma), 56 because they were receiving adjuvant chemotherapy, and 123 because D-dimer testing was not performed (in most of these patients due to a faulty shipment from one of the centers to Amsterdam during which samples probably were thawed).

Both studies were approved by the local ethics committees of the participating hospitals, for CATS the Vienna Medical University (EK 126/2003), for MICA 2 hospitals in Amsterdam, one in Slotervaart hospital had their EC at VU University Medical Center (#2010.259). The coordinating investigators for CATS were in Vienna and for MICA in Amsterdam. All patients provided written informed consent prior to any patient-related activities. There was no randomization, however, laboratory technicians were blinded to the study outcomes.

Procedures

At baseline, blood samples were collected from participants from both CATS and MICA studies in citrated tubes by sterile, antecubital vein puncture or through a peripheral catheter directly after placement. Platelet-poor plasma was prepared and aliquots were stored at -80 °C until further analysis. D-Dimer levels were measured using the STA-Liatest assay (Diagnostica-Stago, Asnières, France) in CATS and the INNOVANCE assay in MICA (Siemens Healthcare, Marburg, Germany). Other laboratory assays from these studies are reported in Appendix **page 4-5**.

Outcomes

The primary outcome of CATS was a composite of symptomatic, objectively confirmed and independently-adjudicated VTE, defined as distal or proximal deep vein thrombosis (DVT) of the leg, upper extremity DVT, symptomatic splanchnic DVT and/or pulmonary embolism (PE) occurring during a 2-year observation period. So-called incidental PEs were counted as events, given that the adjudication committee deemed them to be of clinical significance with a requirement for anticoagulation. Upper extremity DVT related to indwelling venous catheters and incidental splanchnic vein thrombosis were not considered as events.

The primary outcome in MICA was a composite of objectively confirmed symptomatic or incidental PE, distal or proximal DVT, non-catheter-related upper extremity DVT, or symptomatic catheter-related upper extremity DVT occurring during 6 months of follow-up. All diagnoses were centrally verified based on imaging reports. Asymptomatic upper extremity DVT related to indwelling venous catheters and splanchnic vein thrombosis were not considered as events.

Routine screening for VTE was not performed in CATS or MICA.

Analytical Approach—For the development of the clinical prediction model, four principles were prespecified based on clinical grounds. The target population was defined as ambulatory cancer patients, since approximately 75% of all cancer-associated VTEs can occur within this population.⁴ Patients with high-grade gliomas, multiple myelomas, were not included, because specific clinical prediction models have already been developed for those distinct tumor entities.^{16,17}

The primary endpoint was considered to be the cumulative incidence of VTE during the first 6 months, and was defined as the time horizon for the prediction model. A shorter interval would only provide guidance during the first chemotherapy cycles, while the VTE risk remains elevated thereafter. A longer interval was deemed less relevant, because few physicians will consider primary thromboprophylaxis for more than 6 months given the current evidence that the probability is highest in the first 6 months after diagnosis of cancer.² Tumor sites were categorized as harbouring “low/intermediate”, “high”, and “very-high” risk of VTE according to the Khorana et al. tumour site criteria, with previously proposed modifications.^{5,7} As the risk of VTE in patients with colorectal cancer was substantial in CATS (8%),⁷ this type of cancer was assigned to the “high” VTE risk group. All assignments were predefined before the development of this model.

Statistical analysis

All statistical analyses were performed using R (Version 3.3.3, R Core Development Team). Distributional differences of baseline variables between CATS and MICA were evaluated using measures of standardized mean differences (SMDs). SMDs of more than 0.2 were considered to indicate a potentially relevant difference between the two cohorts.¹⁸ The cumulative incidence of VTE was estimated with cause-specific cumulative incidence estimators, treating death not related to VTE as the competing event.¹⁹ All-cause mortality was estimated using a Kaplan-Meier estimator. A penalized regression approach (LASSO, R

package `glmnet`²⁰ with cause-specific VTE hazards was used to select prognostic variables for the clinical prediction model from a large pool of clinical and laboratory candidate variables in CATS (Appendix page 2-3).²¹ Continuous variables were log₂-transformed prior to variable selection to avoid disproportional impact of high values. Continuous variables with standardized hazard ratios between 0.80 and 1.25 were omitted to prevent the inclusion of variables with a small magnitude of association. The resulting model was further reduced by fitting a Fine & Gray competing risk regression (R package `cmprsk`) with a backward selection algorithm with $p > 0.05$ for exclusion. Missing data in selected variables were multiply imputed using the predictive mean matching method in a chained equations algorithm. The five imputed datasets were analyzed separately and results were pooled using Rubin's rules. The resulting model was simplified into a nomogram (R package `Design`). External validation was performed in MICA using complete case analysis.

Discrimination (a measure of a model's ability to distinguish between patients who developed VTE vs those who didn't, as indicated by a modification of Harrell's c-index to accommodate censoring and competing risks (R package `conreg`)) and calibration (a measure of agreement between observed proportions and predicted VTE probabilities, as indicated by calibration plots) were used to assess the performance of the model.^{22,23} The c-index was cross-validated with 1000 bootstrap samples to account for potential over-optimism. External validation was performed in a semi-blinded fashion without data pooling, i.e. the MICA investigators had no access to CATS data and vice versa. The developed clinical prediction model was compared to the Khorana score by computing a population-weighted net reclassification improvement (NRI) statistic.²⁴ CATS was started already in 2003 and registered at the Medical University Vienna (EK 126/2003), MICA was registered at ClinicalTrials.gov (NCT02095925).

Role of the funding source—The funding agencies had no role in the design, analysis, interpretation, or writing of the manuscript.

Results

Description of study cohorts

The development cohort (CATS) included 1,423 patients and the validation cohort (MICA) 832 patients. As indicated by standardized mean differences, important baseline variables such as age, sex, and D-Dimer were similarly distributed in CATS and MICA (Table 1). Patients with esophageal cancer were more prevalent in MICA than in CATS, whereas patients with lymphoma were more prevalent in CATS than in MICA. The percentage of newly diagnosed patients was 70.8% in CATS and 30.2% in MICA. In CATS none of the patients was under active chemotherapy. The average Khorana score was slightly higher in MICA (1.4 points) than in CATS (1.1 points). During a median follow-up of 180 days (IQR 180-180), 80 (5.6%) and 48 (5.8%) patients developed VTE in CATS. during a median follow-up of 180 days (IQR, 109-180), 48 (5.8%) developed VTE in MICA. In competing risk analysis, this corresponded to cumulative 6-month VTE risks of 5.7% (95% confidence interval [CI]: 4.5-6.9) in CATS and 6.3% (4.7-8.2) in MICA, respectively (Figure 1, **Panels A and B**). The most frequent types of VTE in CATS and MICA were lower extremity DVT

and PE (Appendix **page 6**). The estimated 6-month mortality was 13.7% (12.0-15.6) in CATS, and 18.0% (15.4-20.6) in MICA (Figure 1, **Panels A and B**).

Clinical prediction model development and internal validation in CATS

Univariable modeling of cause-specific VTE hazards identified a number of clinical prognostic factors and biomarkers Appendix **page 2-3**). Among the risk factors for cancer-associated VTE, the pre-specified variable selection process selected two variables for the clinical prediction model, namely tumor site category (“low/intermediate”, “high”, and “very high” risk of VTE) and continuous D-Dimer levels. In this model, the multivariable subdistribution hazard ratios (SHR) were 1.96 for “very-high” vs. “high”, and “high” vs. “low/intermediate” VTE-risk tumor site categories (95% CI: 1.41-2.72; $p < 0.001$), and 1.32 (1.12-1.56; $p = 0.001$) per doubling of D-Dimer.

The cross-validated c-index of this model was 0.66 (95% CI: 0.63-0.67). The model was adequately calibrated (Figure 2, **Panel A**), with no indication of systematic under- or overestimation of VTE. A corresponding nomogram was constructed (Figure 3). No statistical interaction between tumor site category and VTE risk was observed ($p = 0.18$), suggesting that D-Dimer may be useful for further VTE risk stratification within individual tumor site categories. Sensitivity, specificity, positive and negative predictive value in the CATS cohort were, at a cut-off of a predicted 6-month cumulative VTE risk of 10%, 33% (95% CI: 23-47%), 84% (83-87%), 12% (8-16%), and 95% (94-96%), respectively. At a cut-off of a risk of 15% the respective rates were 15% (8-24%), 96% (95-97%), 18% (9-29%), and 95% (94-96%), respectively.

External validation of the clinical prediction model in MICA

The c-index of the model developed in CATS and then applied unmodified to MICA was 0.68 (0.62-0.74). Predicted VTE risks in MICA based on the clinical prediction model were in agreement with the observed VTE incidences (Figure 2, **Panel B**). At a cut-off of a predicted 6-month cumulative VTE risk of 10%, sensitivity, specificity, positive and negative predictive value in the MICA validation cohort were 21% (95% CI: 10-35%), 87% (85-90%), 9% (4-16%), and 95% (93-96%), respectively. At a cut-off of a risk of 15% the respective rates were 8% (2-20%), 99% (98-99%), 29% (8-58%), and 95% (93-96%), respectively.

Reclassification statistics and decision curve analysis

Among the five items of the Khorana score, only tumor site category was significantly associated with VTE risk in both CATS and MICA (Table 2). The c-indices of the Khorana score for prediction of the 6-month VTE risk were 0.61 (95% CI: 0.51-0.70) in CATS and 0.56 (0.50-0.63) in MICA, which were lower than the corresponding c-indices of the current clinical prediction model. The current clinical prediction model also had a comparable c-index to the more complex “Vienna Model”, in which the Khorana score is extended with D-Dimer and soluble P-Selectin levels, in both CATS (0.66, 95% CI: 0.58-0.73) and MICA (0.63, 95% CI: 0.55-0.70). Applying the current model instead of the Khorana score reclassified 31% of patients in CATS according to their correct VTE outcome (population-weighted net reclassification improvement (NRI)=0.31).

To gauge the clinical utility of the model for indicating thromboprophylaxis, a decision curve analysis was performed. Here, using the model for thromboprophylaxis indication resulted in a higher clinical utility than strategies of "treat all" and "treat none". This was particularly the case for physicians who have their personal thresholds for indicating thromboprophylaxis at 6-month VTE risks between 6% and 11% (Figure 4).

Discussion

In this study, a novel clinical prediction model for VTE in ambulatory patients with various types of solid cancer was developed in a large prospective cohort study and subsequently externally validated in an independent prospective cohort. Evaluating a large number of clinical and laboratory parameters our analysis yielded a simple model including only one clinical item (tumor site category) and one widely available biomarker (D-Dimer). The resulting nomogram was able to discriminate between patients who developed VTE during 6-month follow-up versus those who did not, and was appropriately calibrated. Decision curve analysis confirmed that applying the model for thromboprophylaxis indication resulted in higher clinical utility than a "treat-all" and a "treat-none" approach. The model was subsequently implemented as a paper-based nomogram as well as an online risk calculator.

This novel and simple tool might enable clinicians to identify ambulatory patients harbouring solid cancers with a 6-month VTE risk of 10-15% or more, and who might benefit from thromboprophylaxis. Moreover, this clinical prediction model can identify a large group of patients with a very low risk of VTE, in which the potential harms of thromboprophylaxis might likely exceed its benefits because of a potentially increased risk of bleeding.² The novelty of the present model is that first, we followed a certain methodological pathway to identify an optimal biomarker— amongst many biomarkers, including parameters from the Khorana score – and we identified this to be the D-Dimer measurement. Second, we used the MICA study as an independent cohort for external validation which revealed a similar statistical power than CATS. Third, we provide a tool, the nomogram, which allows the simple assessment of an individual's risk of VTE, which can be easily acquired via connection to the internet or used in a paper-version.

A limited number of VTE clinical prediction models in the oncological setting have previously been developed,²⁵ however, improvement is desired in order to better allocate patients to high or low risk groups of VTE. At present, the Khorana score is most widely used due to its successful external validation in several studies and its endorsement by pertinent guidelines of the American Society of Clinical Oncology.^{5,6,8} The most important item of the Khorana score is the tumor site category²⁶. Thus, in the present model we have 'borrowed strength' from the Khorana study by adopting the tumour site categories. The 6-month prediction time window chosen in the present model covers the period of highest VTE incidence.⁷ Compared to the model from our group published in 2010⁷ the most important improvement is that the present clinical prediction rule is not dependent on blood count parameters, which can vary strongly for various reasons, but especially based on chemotherapy regimens, whereas D-Dimer measurements are not influenced by chemotherapy treatments.²⁷ In addition, the determination of sP-selectin concentrations is not required, as this biomarker did not reach our predefined cutoff for a meaningful

predictive parameter as outlined in the methods section (less than standardized hazard ratio of 1.25. Most relevant, sP-selectin determination is not available in routine laboratory settings and thus, broad acceptance of this parameter in daily clinical care cannot be expected.

A personalized approach for using thromboprophylaxis in high-VTE-risk cancer patients selected by the Khorana score has been shown to be feasible, leading to a reduction in VTE rates, yet with increases in clinically-relevant bleeding being reported.²⁸ Thus, further improvement of clinical prediction models is necessary in order to determine the benefit of thromboprophylaxis, most appropriately in risk-adapted trials of cancer patients with the highest risk of VTE. Until results from such trials become available, the present data suggest useful thresholds that might be considered for primary thromboprophylaxis, although further research is warranted. For this purpose, our clinical prediction model was evaluated for clinical utility with decision curve analysis.²⁹ This analysis suggested that the clinical prediction model could be useful in a realistic clinical setting, in particular if thromboprophylaxis is considered for patients with a VTE risk greater than 5-15%.

The performance of various biomarkers for cancer-associated VTE has been assessed in previous studies, including some with limited availability outside research environments, such as soluble vascular endothelial growth factor and thrombin generation.^{12,30} Among these, our LASSO variable selection approach identified D-Dimer as the strongest prognostic biomarker in our study. D-Dimer has been validated across multiple cohorts for exclusion of VTE in diagnostic settings and as an independent VTE risk factor in prognostic settings for both cancer and non-cancer patients.³¹ This test is widely available in healthcare facilities. Despite the fact that an assay from another manufacturer was used in the validation cohort than in the development cohort, the prognostic performance of D-dimer testing was consistent, which is reassuring for the putative use of other D-Dimer assays with this clinical prediction model. Indeed, the successful external validation of the model in a cohort with a different D-Dimer assay is considered a major strength. However, it is evident that other D-Dimer assays need to be validated before they can be proposed to be used in this specific setting.

VTE risk thresholds for considering prophylactic anticoagulation in cancer patients are subjective from both a physician's and a patient's perspective. By providing both a paper-based nomogram as well as its implementation as an online risk calculator (catscore.meduniwien.ac.at (will be implemented after acceptance of the manuscript)), doctors and patients may use this tool for individualized VTE risk assessment and shared decision making on thromboprophylactic strategies.

Irrespective of subjective thresholds for indicating thromboprophylaxis, there is reasonable evidence that thromboprophylaxis halves the absolute risk of VTE in patients with cancer.² Based on this assumed absolute risk reduction, the numbers-needed-to-treat (NNT) to prevent one cancer-associated VTE would be 40 or more in the 2-5% VTE risk range, between 20 and 40 in the 5-10% risk range, between 14 and 19 in the 10-15% risk range, and less than 14 in the 15% or more risk range.² Although arbitrary, we posit that thromboprophylaxis is justified at least for cancer patients with a predicted 6-month VTE

risk of 15% or more, and perhaps also in those with a risk of 10% to 15%, because in non-cancer patients with VTE, an approximate 10% VTE recurrence rate at 12 months is seen as enough justification for long-term anticoagulation by the authors of the ACCP guidelines.³² Furthermore, in non-cancer patients undergoing orthopedic surgery, a risk of symptomatic VTE even less than 5% is considered relevant for initiation of thromboprophylaxis.³³

There are some specific limitations of this study that undermine its generalizability. One limitation is that D-dimer assays not used in the CATS or MICA cohort could reveal different results and might thus need to be evaluated separately. As deep vein thrombosis and pulmonary embolism were both seen as a composite for the primary outcome, we cannot comment on the validity of the risk assessment model separately for deep vein thrombosis or pulmonary embolism. The derivation and validation cohort, CATS and MICA, were performed mainly at academic centers, which likely does not reflect the full spectrum of cancer patients. A considerable proportion of patients with esophageal cancer were enrolled in MICA, while these patients were underrepresented in the derivation study and in the Khorana score. By contrast, only patients with lymphoma were enrolled in CATS. Furthermore, 30% of the patients in CATS were not newly diagnosed but had a history of cancer, albeit no recent chemotherapy and in MICA 70% of patients were enrolled after the start of chemotherapy, which limits the evaluation of the Khorana score in such a population. However, one might also consider that a strength of our model is that it can also be used in patients who have in fact already started chemotherapy. This might in turn, potentially broaden its application in daily practice. In conclusion, we present a novel clinical prediction model for VTE in ambulatory patients with solid cancers which includes one clinical factor (tumor entity) and one biomarker (D-Dimer), and which is able to discriminate between patients at low and high risk of VTE. This tool also has the potential of being used for the selection of cancer patients that may require thromboprophylaxis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful to Ankie Kleinjan, Pieter W. Kamphuisen, Hans-Martin Otten, and Isabelle Mahé for enrolment of patients in MICA.

Funding for CATS

Austrian National Bank Memorial Fund (Grants No. 10935 & 12739), Austrian Science Fund (Special Research Program FWF-SFB54 “InThro”)

Funding for MICA

Unrestricted grants of participating hospitals.

References

1. Ay C, Pabinger I, Cohen AT. Cancer-associated venous thromboembolism: Burden, mechanisms, and management. *Thromb Haemost.* 2017; 117(2):219–230. [PubMed: 27882374]

2. Di Nisio M, Porreca E, Candeloro M, et al. Primary prophylaxis for venous thromboembolism in ambulatory cancer patients receiving chemotherapy. *Cochrane Database Syst Rev.* 2016; 12
3. Prandoni P, Lensing AWA, Piccioli A, et al. Recurrent venous thromboembolism and bleeding complications during anticoagulant treatment in patients with cancer and venous thrombosis. *Blood.* 2002; 100(10):3484–3488. [PubMed: 12393647]
4. Ay C, Pabinger I. VTE risk assessment in cancer. Who needs prophylaxis and who does not? *Hamostaseologie.* 2015; 35(4):319–324. [PubMed: 25740182]
5. Khorana AA, Kuderer NM, Culakova E, Lyman GH, Francis CW. Development and validation of a predictive model for chemotherapy-associated thrombosis. *Blood.* 2008; 111(10):4902–4907. [PubMed: 18216292]
6. Lyman GH, Khorana AA, Kuderer NM, et al. Venous thromboembolism prophylaxis and treatment in patients with cancer: American Society of Clinical Oncology clinical practice guideline update. *J Clin Oncol Off J Am Soc Clin Oncol.* 2013; 31(17):2189–2204.
7. Ay C, Dunkler D, Marosi C, et al. Prediction of venous thromboembolism in cancer patients. *Blood.* 2010; 116(24):5377–5382. [PubMed: 20829374]
8. 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(3):291–292. [PubMed: 22547369]
9. Pelzer U, Sinn M, Stieler J, Riess H. Primary pharmacological prevention of thromboembolic events in ambulatory patients with advanced pancreatic cancer treated with chemotherapy? *Dtsch Med Wochenschr* 1946. 2013; 138(41):2084–2088.
10. van Es N, Di Nisio M, Cesarman G, et al. Comparison of risk prediction scores for venous thromboembolism in cancer patients: a prospective cohort study. *Haematologica.* 2017; 102(9):1494–1501. [PubMed: 28550192]
11. Gerotziafas GT, Taher A, Abdel-Razek H, et al. A Predictive Score for Thrombosis Associated with Breast, Colorectal, Lung, or Ovarian Cancer: The Prospective COMPASS-Cancer-Associated Thrombosis Study. *The Oncologist.* 2017; 22(10):1222–1231. [PubMed: 28550032]
12. Pabinger I, Thaler J, Ay C. Biomarkers for prediction of venous thromboembolism in cancer. *Blood.* 2013; 122(12):2011–2018. [PubMed: 23908470]
13. Ay C, Simanek R, Vormittag R, et al. High plasma levels of soluble P-selectin are predictive of venous thromboembolism in cancer patients: results from the Vienna Cancer and Thrombosis Study (CATS). *Blood.* 2008; 112(7):2703–2708. [PubMed: 18539899]
14. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med.* 2015; 13:1. [PubMed: 25563062]
15. Ay C, Vormittag R, Dunkler D, et al. D-dimer and prothrombin fragment 1 + 2 predict venous thromboembolism in patients with cancer: results from the Vienna Cancer and Thrombosis Study. *J Clin Oncol Off J Am Soc Clin Oncol.* 2009; 27(25):4124–4129.
16. Riedl J, Preusser M, Nazari PMS, et al. Podoplanin expression in primary brain tumors induces platelet aggregation and increases risk of venous thromboembolism. *Blood.* 2017; 129(13):1831–1839. [PubMed: 28073783]
17. Thaler J, Ay C, Kaider A, et al. Biomarkers predictive of venous thromboembolism in patients with newly diagnosed high-grade gliomas. *Neuro-Oncol.* 2014
18. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med.* 2015; 34(28):3661–3679. [PubMed: 26238958]
19. Ay C, Posch F, Kaider A, Zielinski C, Pabinger I. Estimating risk of venous thromboembolism in patients with cancer in the presence of competing mortality. *J Thromb Haemost JTH.* 2015; 13(3):390–397. [PubMed: 25529107]
20. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J Stat Softw.* 2011; 39(5):1–13.
21. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* 1997; 16(4):385–395. [PubMed: 9044528]

22. Dunkler D, Schemper M, Heinze G. Gene selection in microarray survival studies under possibly non-proportional hazards. *Bioinforma Oxf Engl*. 2010; 26(6):784–790.
23. Collins GS, Reitsma JB, Altman DG, Moons KGM, TRIPOD Group. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD Group. *Circulation*. 2015; 131(2):211–219. [PubMed: 25561516]
24. Kerr KF, Wang Z, Janes H, et al. Net reclassification indices for evaluating risk prediction instruments: a critical review. *Epidemiol Camb Mass*. 2014; 25(1):114–121.
25. Angelini D, Khorana AA. Risk Assessment Scores for Cancer-Associated Venous Thromboembolic Disease. *Semin Thromb Hemost*. 2017; 43(5):469–478. [PubMed: 28264198]
26. Pabinger I, Posch F. Flamethrowers: blood cells and cancer thrombosis risk. *Hematol Am Soc Hematol Educ Program*. 2014; 2014(1):410–417.
27. Reitter E-M, Kaider A, Ay C, et al. Longitudinal analysis of hemostasis biomarkers in cancer patients during antitumor treatment. *J Thromb Haemost JTH*. 2016; 14(2):294–305. [PubMed: 26662117]
28. Khorana AA, Francis CW, Kuderer NM, et al. Dalteparin thromboprophylaxis in cancer patients at high risk for venous thromboembolism: A randomized trial. *Thromb Res*. 2017; 151:89–95. [PubMed: 28139259]
29. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak Int J Soc Med Decis Mak*. 2006; 26(6):565–574.
30. Posch F, Thaler J, Zlabinger G-J, et al. Soluble Vascular Endothelial Growth Factor (sVEGF) and the Risk of Venous Thromboembolism in Patients with Cancer: Results from the Vienna Cancer and Thrombosis Study (CATS). *Clin Cancer Res Off J Am Assoc Cancer Res*. 2015
31. Ay C, Dunkler D, Pirker R, et al. High D-dimer levels are associated with poor prognosis in cancer patients. *Haematologica*. 2012; 97(8):1158–1164. [PubMed: 22371182]
32. Kearon C, Akl EA, Comerota AJ, et al. Antithrombotic therapy for VTE disease: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012; 141(2 Suppl):e419S–94S. [PubMed: 22315268]
33. Falck-Ytter Y, Francis CW, Johanson NA, et al. Prevention of VTE in orthopedic surgery patients: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012; 141(2 Suppl):e278S–325S. [PubMed: 22315265]

Research in context

Evidence before this study

Venous thromboembolism (VTE) is a frequent complication in patients with cancer. Although pharmacologic thromboprophylaxis significantly reduces the relative risk of cancer-associated VTE, this intervention has not been routinely adopted in clinical practice because the absolute risk reduction of VTE is low for most patients with cancer. A personalized approach to VTE risk assessment in the oncologic setting using clinical prediction models holds great promise to help clinicians in identifying those cancer patients with a high VTE risk justifying pharmacologic thromboprophylaxis. Clinical parameters and biomarkers of haemostatic activation, such as D-Dimer, thrombin generation or soluble P-selectin have been shown to harbor important information on VTE risk in patients with cancer. Several risk scores for cancer-associated VTE have been developed in the past, including the Khorana score, the PROTECHT score, the Vienna update of the Khorana score, and the CONKO score, but their performance should be improved, in order to further increase the predictive power and to provide easy to use tools. Therefore, an advanced clinical prediction model is needed to allow improvement of an individualized approach towards primary thromboprophylaxis in patients with cancer.

Added value of this study

A clinical prediction model for cancer-associated VTE in ambulatory patients with solid tumors was developed and externally validated in an independent second prospective cohort study. A high number of clinical variables and biomarkers were considered during model development. The new model includes only one clinical factor (tumor site category) and one biomarker (D-Dimer), and features characteristics that might outperform previous clinical prediction scores. For eased clinical applicability, the model can, in addition to the printed nomogram, be available as an online prediction tool.

Implications of all the available evidence

A simple clinical prediction model considerably improves predicting cancer-associated VTE, and can aid physicians in selecting ambulatory patients with solid tumors who will likely benefit from pharmacologic thromboprophylaxis.

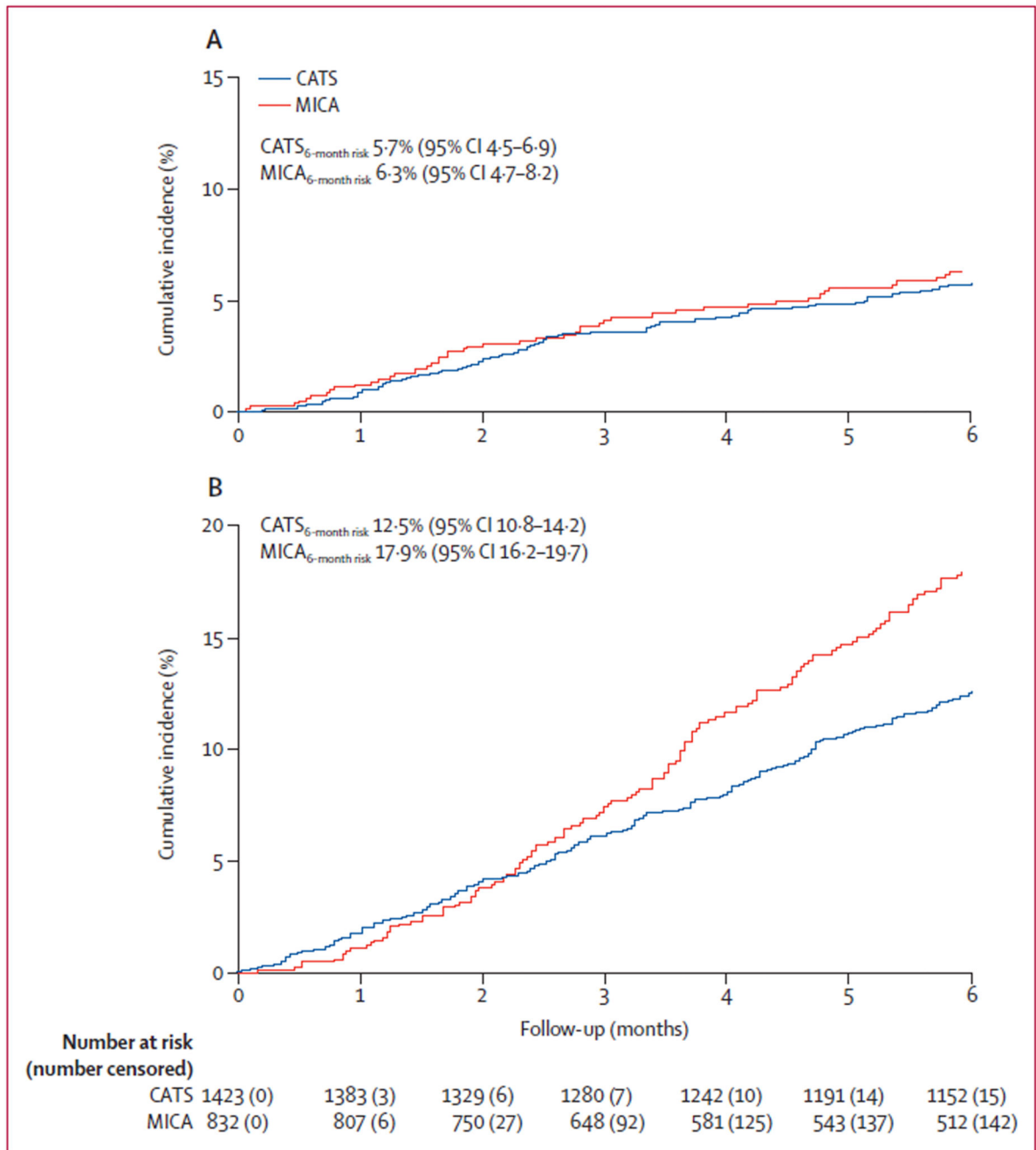


Figure 1. Cumulative incidence of VTE and death-from-any-cause in CATS and MICA. Panel 1A depicts results for CATS, and Panel 1B for MICA. Risks of VTE and death were similar in CATS and MICA. The cumulative incidence of VTE was estimated with competing risk estimators treating death-from-any-cause-except-fatal-VTE as the competing event of interest. Risk of death-from-any-cause was estimated with a 1-Kaplan-Meier estimator.

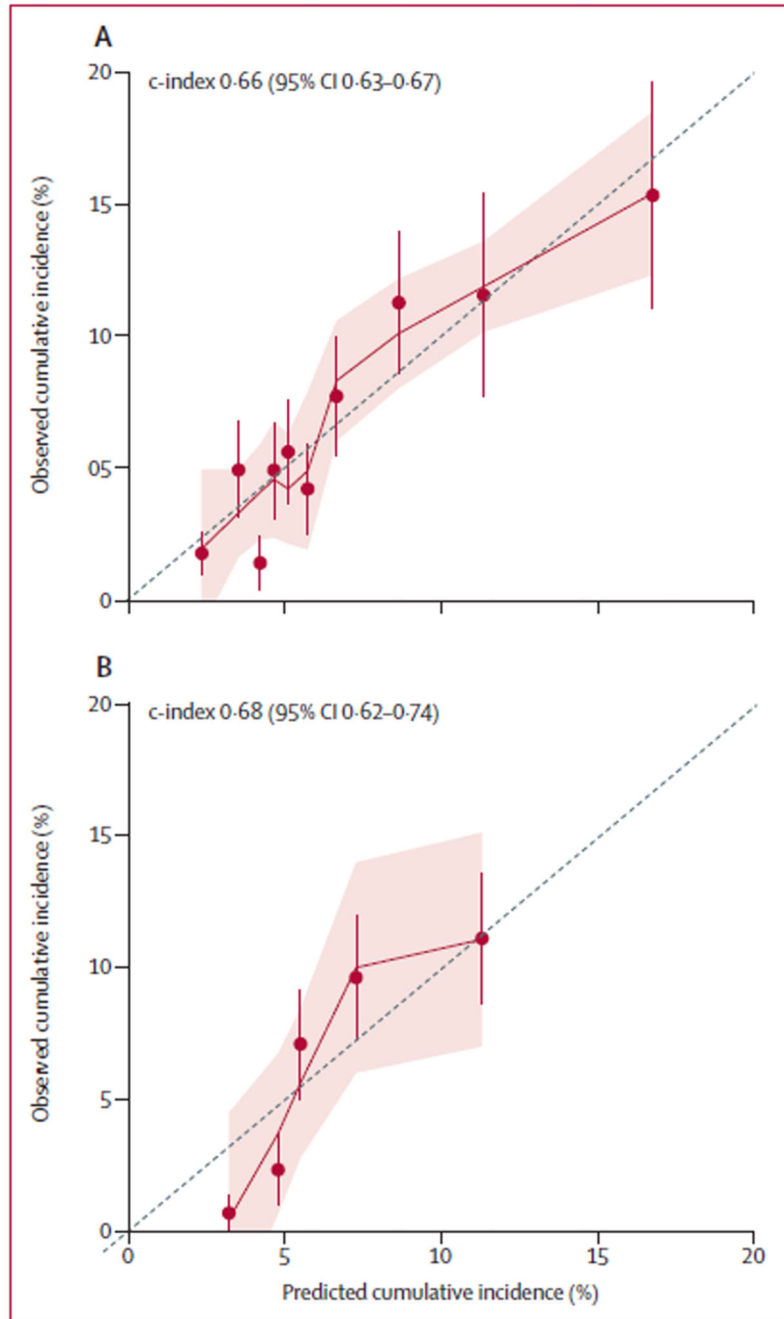


Figure 2. Cross-validated calibration plots of the clinical prediction model in CATS and MICA. Panel 2A depicts results for CATS, and Panel 2B for MICA. These graphs plot observed against predicted VTE risks within deciles (CATS) and quintiles (MICA) of the score’s linear predictor. A smaller distance of the scatter points to the 45° line indicates better calibration. Error bars represent 95% CIs of the prediction. Some underprediction of VTE risk according to the clinical prediction model can be observed in MICA patients at very low risk of VTE.

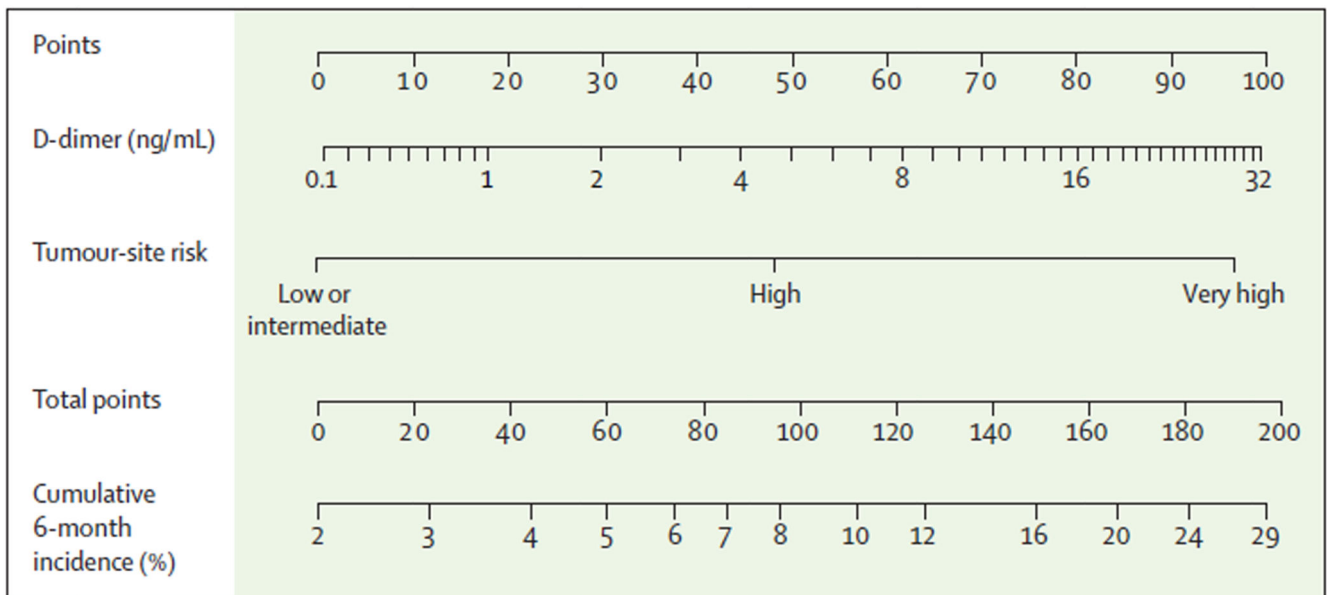


Figure 3. Externally-validated nomogram for predicting the risk of cancer-associated VTE.

This nomogram can be used to obtain 6-month VTE risk predictions for outpatients with solid cancers. The unit of D-Dimer is $\mu\text{g/mL}$. To compute the 6-month VTE risk for an individual patient, D-Dimer (second row) and tumor site category (third row) need to be referred to the point caliper (first row). The obtained points then need to be added, the resulting sum (fourth row) can then be referred to the fifth row to obtain the predicted 6-month risk of VTE.

$$6\text{month VTE risk}(\%) = 100 * (1 - (1 - 0.02137053)^{e^{(0.6709158) \times \text{cancersite} + 0.2793001 \times \log_2(d\text{ dimer} + 1)}}))$$

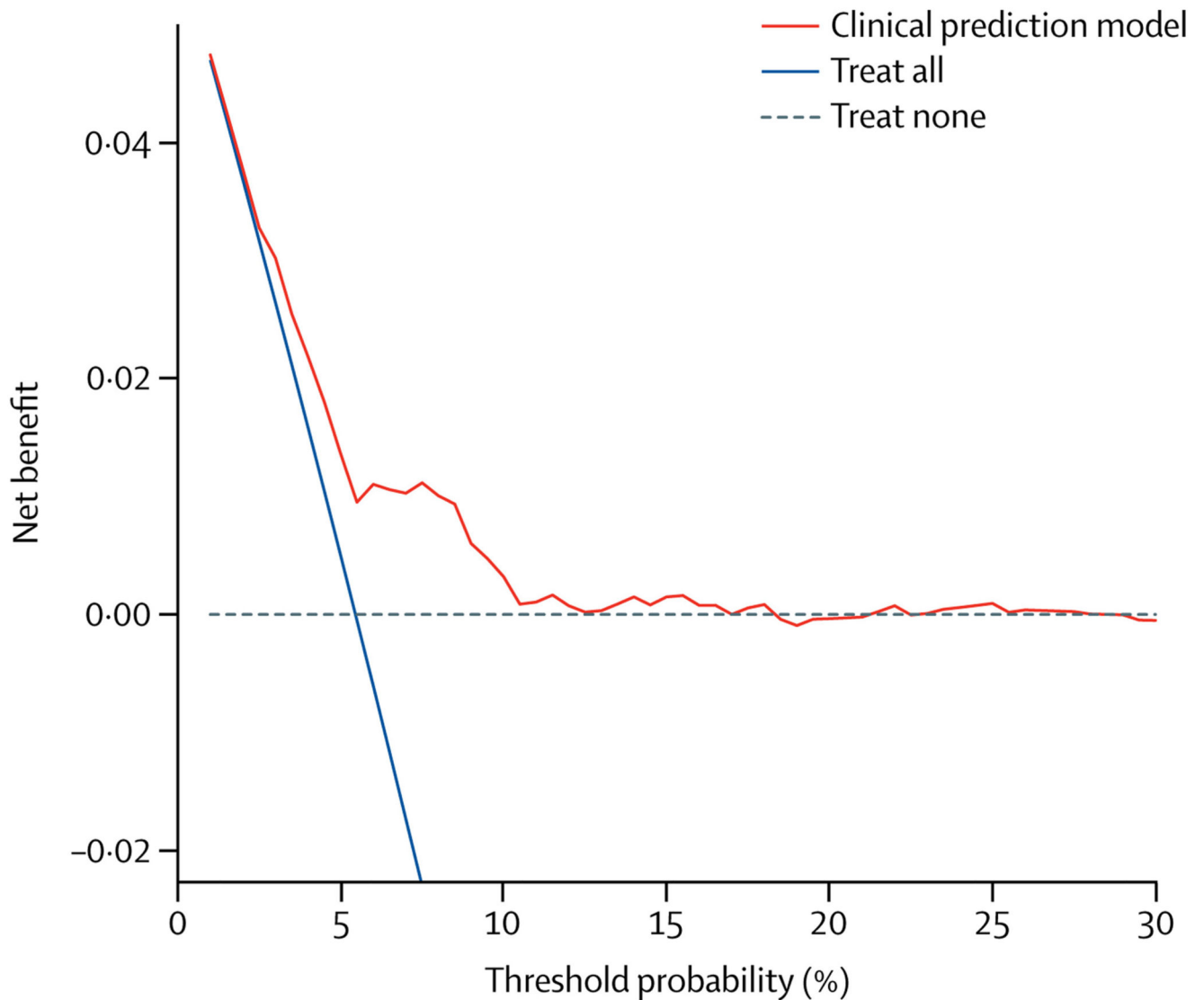


Figure 4. Decision curve analysis for primary thromboprophylaxis in patients with cancer.

This decision curve analysis was performed in CATS, and evaluates the net benefit (avoiding VTE vs. unnecessary thromboprophylaxis) of using the clinical prediction model for targeted thromboprophylaxis compared to providing thromboprophylaxis to all (“treat all”) or to no patients (“treat none”). The threshold probability (x-axis) represents the predicted 6-month VTE risk in CATS for recommending a primary thromboprophylaxis, and the y-axis represents the net clinical benefit (true positive rate minus weighted false positive rate). The threshold probability expresses the relative weight of the assumed potential harms by unnecessary thromboprophylaxis compared to the risk of VTE. For example, a threshold probability of 10% implicitly considers unnecessary treatment of 9 patients $(=(100-10)/10\%)$ equally harmful as missing one successfully prevented VTE event.

Table 1

Baseline characteristics of the study populations

Distribution in CATS and MICA studies. The column “n (% miss.)” indicates the number of patients with missing values. Results represent median [25th percentile (Q1) – 75th percentile (Q3)] for continuous variables, and absolute frequencies (%) for categorical variables. SMDs are standardized mean differences. An SMD of 1 indicates that the center of the data would be one standard deviation higher in CATS than in MICA or vice versa. By convention (Austin PC et al. Stat Med. 2015) we considered SMDs >0.2 indicating a potentially relevant difference between CATS and MICA. Abbreviations: BMI – body mass index, UICC – Union Internationale pour le Lutte Contre le Cancer, N/A – not available *”Other sites” mainly included patients with sarcomas and testicular germ cell tumors

Variable	CATS (n=1,423)		MICA (n=832)		SMD
	n (%missing)	Summary measure	n (%missing)	Summary measure	
Clinical Variables					
Age at entry (years)	1,423 (0.0%)	62.9 [54.0-68.9]	832 (0.0%)	63.7 [55.9-70.3]	0.16
BMI (kg/m ²)	1,418 (0.4%)	25.0 [22.1-28.3]	825 (0.8%)	24.7 [22.5-27.4]	0.09
Female	1,423 (0.0%)	651 (45.8%)	832 (0.0%)	354 (42.7%)	0.06
Use of erythropoiesis-stimulating agents (ESAs)	1,423 (0.0%)	50 (3.5%)	823 (0.7%)	20 (2.4%)	0.06
Tumor site	1,423 (0.0%)	/	832 (0.0%)	/	/
Low/Intermediate Risk of VTE					
Breast	/	379 (26.6%)	/	144 (17.3%)	0.23
Prostate	/	226 (15.9%)	/	89 (10.7%)	0.15
	/	153 (10.8%)	/	39 (4.7%)	0.3
High Risk of VTE	/	863 (60.7%)	/	535 (64.0%)	0.08
Lung	/	292 (20.5%)	/	183 (22.0%)	0.04
Colorectal	/	173 (12.2%)	/	127 (15.3%)	0.09
Esophagus	/	13 (0.9%)	/	177 (21.3%)	0.69
Kidney	/	43 (3.0%)	/	0 (0%)	0.25
Lymphoma (“stageable” by AnnArbor)	/	249 (17.5%)	/	0 (0%)	0.65
Bladder/Urothelial	/	7 (0.5%)	/	11 (1.3%)	0.09
Uterus	/	8 (0.6%)	/	2 (0.2%)	0.05
Cervical	/	16 (1.1%)	/	2 (0.2%)	0.11
Ovary	/	5 (0.4%)	/	33 (4.0%)	0.25
Other site*	/	57 (4.0%)	/	16 (1.9%)	0.12

Variable	CATS (n=1,423)		MICA (n=832)		SMD
	n (%omissing)	Summary measure	n (%omissing)	Summary measure	
Very High Risk of VTE	/	181 (12.7%)	/	153 (18.4%)	0.16
Pancreas	/	118 (8.3%)	/	116 (13.9%)	0.18
Stomach	/	63 (4.4%)	/	37 (4.4%)	0.00
Tumor characteristics					
Newly-diagnosed cancer	1,423 (0.0%)	1,008 (70.8%)	581 (30.2%)	454 (78.1%)	0.17
Tumor grade G3/G4	1,392 (2.2%)	524 (37.6%)	0 (100%)	N/A	N/A
Tumor stage (UICC/AnnArbor)	1,354 (4.8%)	/	826 (%)	/	/
Stage I	/	139 (10.3%)	/	10 (1.2%)	0.40
Stage II	/	311 (23.0%)	/	48 (5.8%)	0.50
Stage III	/	222 (16.4%)	/	258 (31.2%)	0.35
Stage IV	/	682 (50.4%)	/	510 (61.7%)	0.23
Laboratory Parameters & Biomarkers					
Haemoglobin (g/dL)	1,416 (0.5%)	12.9 [11.6-14.0]	820 (1.4%)	13.1 [11.7-14.2]	0.11
Leukocyte count (G/L)	1,416 (0.5%)	7.2 [5.7-9.4]	820 (1.4%)	7.7 [6.0-9.9]	0.10
Neutrophil count (G/L)	1,112 (21.9%)	4.7 [3.5-6.3]	0 (100%)	N/A	N/A
Platelet count (G/L)	1,416 (0.5%)	248 [197-309]	819 (1.6%)	280 [224-353]	0.37
D-Dimer (µg/mL)	1,423 (0.0%)	0.7 [0.4-1.5]	832 (0.0%)	0.94 [0.46-2.08]	0.10
sP-Selectin (ng/mL)	1,410 (0.9%)	40.3 [30.9-50.6]	828 (0.4%)	34.0 [26.0-43.0]	0.40
Fibrinogen (mg/dL)	1,416 (0.5%)	392 [324-489]	0 (100%)	N/A	N/A
Factor VIII activity (%)	1,417 (0.4%)	183 [143-234]	0 (100%)	N/A	N/A
Prothrombin fragment 1.2 (pM/L)	1,406 (1.2%)	232 [169-330]	0 (100%)	N/A	N/A
Peak of Thrombin Generation (nM)	1,415 (0.6%)	392 [207-542]	0 (100%)	N/A	N/A
Velocity Index of Thrombin Generation (nM/min)	1,415 (0.6%)	81.3 [31.8-147.4]	0 (100%)	N/A	N/A
Current VTE prediction models					
Khorana score	1,411 (0.8%)	1 [0-2]	814 (2.2%)	1 [1-2]	0.28
Khorana score 0 points	/	441 (31.3%)	/	163 (19.6%)	0.26
Khorana score 1 point	/	512 (36.3%)	/	292 (35.1%)	0.01
Khorana score 2 points	/	322 (22.8%)	/	242 (29.1%)	0.16

Variable	CATS (n=1,423)		MICA (n=832)		SMD
	n (%missing)	Summary measure	n (%missing)	Summary measure	
Khorana score 3 points	/	136 (9.6%)	/	117 (14.1%)	0.15
Mean Khorana Score (\pm Standard Deviation)	1,411 (0.8%)	1.1 (\pm 1.0)	814 (2.2%)	1.4 (\pm 1.0)	0.28
Vienna update to the Khorana score	1,402 (1.5%)	/	810 (2.6%)	/	0.21
Vienna score 0 points	/	317 (22.6%)	/	107 (12.9%)	0.25
Vienna score 1 point	/	406 (29.0%)	/	232 (27.9%)	0.01
Vienna score 2 points	/	359 (25.6%)	/	240 (28.8%)	0.09
Vienna score 3 points	/	203 (14.5%)	/	145 (17.4%)	0.09
Vienna score 4 points	/	117 (8.3%)	/	86 (10.3%)	0.08

Table 2
Associations between individual Khorana score items and 6-month VTE risk in CATS and MICA.

Among the 5 items, only the tumor site category was associated with venous thromboembolic risk in both cohorts. Results were estimated with multivariable Fine & Gray competing risk regression models, considering death-from-any-cause-except-fatal-VTE as the competing event of interest. Abbreviations: SHR – Subdistribution hazard ratio, 95% CI: 95% Confidence interval, p – Wald-test p-value, Ref. – Reference category, N/A – not applicable (The SHR for BMI could not be estimated since no events were observed in the 26 MICA patients with a BMI $\geq 35\text{kg/m}^2$).

Variable	Multivariable SHR (95% CI, p) in CATS	Multivariable SHR (95% CI, p) in MICA
Tumor site category	/	/
---Low/Intermediate risk	Ref.	Ref.
---High risk	1.99 (1.00-3.94, p=0.05)	2.29 (1.09-4.81; p=0.028)
---Very high risk	4.54 (2.15-9.62, p<0.0001)	2.00 (0.82-4.87; p=0.13)
Body mass index $\geq 35\text{ kg/m}^2$	1.85 (0.66-5.15, p=0.24)	N/A
Platelet count $\geq 350 \times 10^9/\text{L}$	1.15 (0.65-2.03, p=0.63)	1.25 (0.69-2.28; p=0.46)
Hemoglobin level $< 10\text{ g/dL}$ or ESA use	1.47 (0.81-2.68, p=0.21)	1.45 (0.59-3.59; p=0.42)
White blood cell count $> 11 \times 10^9/\text{L}$	1.03 (0.54-1.96, p=0.93)	0.80 (0.37-1.73; p=0.56)