Multivariable clinical-genetic risk model for predicting venous thromboembolic events in patients with cancer

Andrés J. Muñoz Martín1,2, Israel Ortega3, Carme Font2,4, Vanesa Pachón5,5, Victoria Castellón2,6, Virginia Martínez-Marín2,7, Mercedes Salgado2,8, Eva Martínez2,9, Julia Calzas2,10, Ana Rupérez2,11, Juan C. Souto12, Miguel Martín12, Eduardo Salas11 and Jose M. Soria13

BACKGROUND: Venous thromboembolism (VTE) is a leading cause of death among patients with cancer. Outpatients with cancer should be periodically assessed for VTE risk, for which the validated risk assessment tool developed by Khorana—the Khorana score7 is commonly recommended. However, in recent years, a number of studies have questioned whether this tool is sufficiently accurate at identifying patients who should receive thromboprophylaxis.

METHODS: We included 391 outpatients with a recent diagnosis of cancer and candidates for systemic outpatient chemotherapy. All were treated according to standard guidelines. The study population was monitored for 6 months, and VTEs were recorded. The Khorana and the TiC-Onco scores were calculated for each patient and their VTE predictive accuracy VTEs was compared.

RESULTS: We recorded 71 VTEs. The TiC-Onco risk score was significantly better at predicting VTE than the Khorana score (AUC 0.73 vs. 0.58, sensitivity 49 vs. 22%, specificity 81 vs. 82%, PPV 37 vs. 22%, and NPV 88 vs. 82%).

CONCLUSIONS: TiC-Onco risk score performed significantly better than Khorana score at identifying cancer patients at high risk of VTE who would benefit from personalised thromboprophylaxis.

REFERENCES:

1. British Journal of Cancer (2018) 118:1056–1061; https://doi.org/10.1038/s41416-018-0027-8

INTRODUCTION

Patients with cancer (taking all types together) are at 7 times the risk of developing a venous thromboembolism (VTE); in some malignancies the risk increases to 28 times.1 The incidence of cancer-associated VTE is particularly high during the first few months after diagnosis, when distant metastases are present, and after initiating chemotherapy.1, 2 VTE is a leading cause of death among patients with cancer,3 and the survival of an episode may have clinical and economic implications, including hospitalisation, potential delays in cancer therapy, recurrent VTE, post-thrombotic syndrome, and chronic thromboembolic pulmonary hypertension. Indeed, these problems are common, costly, and have a profound impact on the patient’s quality of life.3

Different guidelines cover the identification of patients with cancer at risk of VTE, VTE prevention strategies, and treatment.4–6 These documents indicate that most hospitalised patients with active cancer require thromboprophylaxis throughout hospitalisation. However, in the outpatient setting, it is indicated only for high-risk patients. Outpatients with cancer should be periodically assessed for VTE risk, for which the validated risk assessment tool developed by Khorana—the Khorana score7 is commonly recommended. However, in recent years, a number of studies have questioned whether this tool is sufficiently accurate at identifying patients who should receive thromboprophylaxis.8–10

In the present work, we hypothesised that a genetic component is involved in the appearance of VTE, a factor that the Khorana score does not take into account. Recently, the Vienna Group11 showed that the Leiden (rs6025) variant of the gene coding for factor V in the coagulation pathway doubles the risk of a VTE event occurring in patients with cancer. This variant might, therefore, provide a promising biomarker of venous thrombosis in such patients, and could be used in individual risk prediction.12 Other genes are known to increase the risk of VTE in the general population, and a tool (the Thrombo inCode or TiC tool) involving them as markers has been developed to predict VTE—but only for
non-oncological populations.13 The use of such tools ought to allow more tailored thromboprophylaxis strategies to be followed.

The present work proposes a TiC-derived risk score—the TiC-Onco risk score—which takes into account both genetic and clinical risk factors, and which can be used to identify patients with cancer in the outpatient setting who are at high risk of VTE. Its capacity to identify such patient was compared with that of the Khorana score. The results suggest that the TiC-Onco risk score can better predict which patients should receive thromboprophylaxis.

## PATIENTS AND METHODS

### Study design and participants

The study protocol was approved by the participant hospitals’ institutional review boards. Signed informed consent was obtained from each patient.

This study—the ONCOTHROMB12-01 study (Clinicaltrial.gov identifier: NCT03114618)—is an observational cohort study involving an 18 month monitoring period with analysis at 6, 12, and 18 months. This paper presents the results for the first 6 months.

Selection criteria were as follows:

- Over 18 years of age
- Recent diagnosis of cancer of the following types: colorectal, oesophago-gastric, lung, or pancreatic.
- ECOG/WHO/Zubrod score of 0–2
- Candidates for systemic outpatient chemotherapy according to standard guidelines.
- No outpatient thromboprophylactic therapy deemed mandatory by the treating oncologist.

The Khorana score (reference tool) and the proposed TiC-Onco score (index tool) were calculated for each patient at the moment of initial diagnosis and their accuracy in terms of predicting the observed VTE events of the two tools was compared.

### Diagnosis of VTE events

Deep vein thrombosis in the lower limbs was diagnosed by ultrasound or ascending venography. Pulmonary embolism was diagnosed by ventilation–perfusion lung scanning, pulmonary angiography, or spiral computed tomography. Intracranial venous thrombosis was diagnosed by magnetic resonance imaging.

### Development of the TiC-Onco risk score

The TiC-Onco risk score tool was developed in three steps:

1. Development of a genetic risk score.
2. Selection of clinical variables associated with the development of VTE.
3. Combination of the genetic and clinical variables to create the TiC-Onco risk score.

Data were collected from all patients on the clinical risk factors cited in the literature14 as being associated with VTE and that could be known at the time of diagnosis: primary tumour site, tumour node metastasis stage, and body mass index (BMI), use of tobacco, age, sex, family (first degree) history of VTE, the presence of diabetes, hypertension, and high blood cholesterol level, the Khorana score, previous surgery, number of platelets, number of leukocytes, and immobilisation. The risk of VTE associated with the

---

### Table 1. Study population characteristics

|                        | VTE    | No-VTE | \(p\)-value |
|------------------------|--------|--------|-------------|
| N                      | 71     | 320    |             |
| Sex (female), n (%)    | 27 (38.0) | 108 (33.8) | 0.584       |
| Age, mean (sd)         | 64.1 (11.0) | 64.3 (10.5) | 0.903       |
| Diabetes, n (%)        | 12 (16.9) | 62 (19.4) | 0.754       |
| Smoking, n (%)         | 21 (29.6) | 66 (20.6) | 0.138       |
| Family history (%)     | 6 (8.4) | 12 (3.7) | 0.112       |
| BMI >25, n (%)         | 36 (50.7) | 144 (45.0) | 0.459       |
| Hypercholesterolemia, n (%) | 29 (40.8) | 106 (33.1) | 0.271       |
| Hypertension (%)       | 33 (46.5) | 141 (44.1) | 0.932       |
| Khorana \(\geq 3\)     | 16 (22.5) | 58 (18.1) | 0.505       |

| Primary site of tumour: |        |        |        |
|------------------------|--------|--------|--------|
| Colon                  | 22 (31.0) | 141 (44.1) | 0.059       |
| Pancreas               | 29 (40.8) | 13 (3.7) | 0.011       |
| Lung                   | 11 (15.5) | 76 (23.8) | 0.175       |
| Oesophagus             | 2 (2.8) | 12 (3.7) | 0.976       |
| Stomach                | 7 (9.9) | 48 (15.0) | 0.348       |

| Tumour Stage:          |        |        |        |
|------------------------|--------|--------|--------|
| I + II                 | 5 (7.0) | 66 (20.6) | 0.012       |
| III                    | 18 (25.4) | 121 (37.8) | 0.065       |
| IV                     | 48 (67.6) | 133 (41.6) | <0.001       |
| Haemoglobin <100 g/L, n (%) | 4 (5.6) | 18 (5.6) | >0.999       |
| Platelet >350 \(\times 10^9\)/L, n (%) | 13 (18.3) | 74 (23.1) | 0.469       |
| Leukocyte >11 \(\times 10^9\)/L, n (%) | 15 (21.1) | 58 (18.1) | 0.675       |

| SNPs, risk alleles (%) |        |        |        |
|------------------------|--------|--------|--------|
| F5 rs6025               |        |        |        |
| 0 Risk Alleles          | 68 (95.8) | 314 (98.1) | 0.213       |
| 1 Risk Allele           | 3 (4.2) | 6 (1.9) |        |
| F5 rs4524               |        |        |        |
| 0 Risk Alleles          | 1 (1.4) | 22 (6.9) | 0.108       |
| 1 Risk Allele           | 22 (31.0) | 115 (35.9) |        |
| 2 Risk Alleles          | 48 (67.6) | 183 (57.2) |        |
| F2 rs1799963            |        |        |        |
| 0 Risk Alleles          | 69 (97.2) | 307 (95.9) | >0.999       |
| 1 Risk Allele           | 2 (2.8) | 12 (3.7) |        |
| 2 Risk Alleles          | 0       | 1       | 0.3       |
| F12 rs1801020           |        |        |        |
| 0 Risk Alleles          | 46 (64.8) | 204 (63.7) | >0.999       |
| 1 Risk Allele           | 23 (32.4) | 103 (32.2) |        |
| 2 Risk Alleles          | 2 (2.8) | 13 (4.1) |        |
| F13 rs5985              |        |        |        |
| 0 Risk Alleles          | 36 (50.7) | 184 (57.5) | 0.514       |
| 1 Risk Allele           | 30 (42.3) | 119 (37.2) |        |
| 2 Risk Alleles          | 5 (7.0) | 17 (5.3) |        |
| SERPINC1 rs121909548    |        |        |        |
| 0 Risk Alleles          | 71 (100.0) | 319 (99.7) | >0.999       |
| 1 Risk Allele           | 0       | 1       | 0.3       |
| SERPINA10 rs2232698     |        |        |        |
| 0 Risk Alleles          | 68 (95.8) | 314 (98.1) | 0.213       |
| 1 Risk Allele           | 3 (4.2) | 6 (1.9) |        |
| A1 blood group          |        |        |        |
| 0 A1 Allele             | 41 (57.7) | 194 (60.6) | 0.776       |
| 1 A1 Allele             | 24 (33.8) | 105 (32.8) |        |
| 2 A1 Alleles            | 6 (8.4) | 21 (6.6) |        |
primary tumour site (low, high, and very high) was categorised as when determining the Khorana score. The risks associated with platelet and leukocyte numbers were categorised using the same cut-offs as for the Khorana score.

At 6 months, univariate analysis was performed to determine which of these variables were associated with the appearance of a VTE event. Those associated with an increased risk of VTE ($p \leq 0.25$) were selected.

### Development of the clinical-genetic model.

The genetic risk score and the clinical variables selected were subjected to multivariate logistic regression analysis using an AIC-based backward selection process.

#### Internal validation

Internal validation to obtain the degree of optimism in the area under the receiver operating characteristic (ROC) curve (AUC) estimation was done using the bootstrap approach, considering 100 resamples from the original data.

#### Comparing the Khorana and TiC-Onco risk scores

The risk prediction capacity of the Khorana and TiC-Onco risk scores was evaluated using the area under the receiver operating characteristic (ROC) curve (AUC). Standard measures of sensitivity, specificity, positive, and negative predictive value (PPV and NPV), and positive and negative likelihood ratios (PLR and NLR) were determined for specific cut-off points.

For Khorana score the cut-off defining high risk was set at $\geq 3$ (the normal cut-off value), and 0 for the low risk category definition. We contemplate two scenarios when determining a cut-off for the TiC-Onco score. In the first one (the main scenario), the cut-off is selected as the point on the ROC curve giving the same specificity as provided by the Khorana score (around 80%). It defines those individuals who are in high risk. This will be the default cut-off for TiC-Onco when not specified in the text. In this scenario, we also determine a second cut-off to classify the non-high risk individuals into either intermediate or low risk. Thus, we allow the TiC-Onco score to provide three risk categories—high, intermediate, and low risk. This second cut-off to discriminate between low and intermediate categories is selected as the point giving a sensitivity of 90%. The second scenario is presented just for informative purposes. In this case, we only consider a TiC-Onco with two categories, high risk and non-high risk.

### Statistical analysis

Continuous variables were recorded as median [1st-3rd Quartiles], and categorical variables as proportions. Univariate association between clinical/genetic variables and events was determined using either t-test or Wilcoxon rank sum tests for continuous variables, and $\chi^2$ or Fisher tests for categorical variables. All calculations were performed using R statistical software (version 3.1.3).

### RESULTS

Patient characteristics

Tables 1 and 2 show the clinical and demographic characteristics of the 391 patients at the start of the study. For each variable, the
number and percentage of patients who experienced a VTE, or not, at some point in the 6-month study period, are shown. The overall incidence of VTE was 18%. Patients suffering from pancreatic cancer experienced VTE at a significantly higher frequency (40%) than patients with other type of cancers (p = 0.001) (Table 2).

Development of the TiC-Onco risk model

Table 3 shows the genetic and clinical markers that were significantly associated by multivariate analysis with a VTE event, and thus selected for inclusion in the TiC-Onco risk score model.

Table 4. Predictive capability of TiC-Onco and Khorana scores

|                  | TiC-Onco (1) | TiC-Onco (2) | Khorana | p (TiC-Onco (1) vs Khorana) | p (TiC-Onco (2) vs Khorana) |
|------------------|--------------|--------------|---------|-----------------------------|----------------------------|
| AUC (95% CI)     | 0.734 (0.67–0.79) | 0.734 (0.67–0.79) | 0.580 (0.51–0.65) | <0.001 | <0.001 |
| Sensitivity, %   | 49.30 (37.7–60.9) | 85.92 (77.8–94.0) | 22.54 (12.8–32.3) | <0.001 | <0.001 |
| Specificity, %   | 81.25 (77.0–85.5) | 49.06 (43.6–54.5) | 81.76 (77.5–86.0) | 0.823 | <0.001 |
| PPV, % (95% CI)  | 36.84 (27.1–46.5) | 27.23 (21.4–33.1) | 21.62 (12.2–31.0) | 0.004 | 0.218 |
| NPV, % (95% CI)  | 87.84 (84.1–91.6) | 94.01 (90.4–97.6) | 82.54 (78.3–86.7) | <0.001 | <0.001 |
| PLR (95% CI)     | 2.63 (1.89–3.65) | 1.69 (1.46–1.95) | 1.24 (0.76–2.02) | 0.005 | 0.244 |
| NLR (95% CI)     | 0.62 (0.49–0.79) | 0.29 (0.16–0.52) | 0.95 (0.83–1.09) | 0.001 | <0.001 |

TiC-Onco (1) shows the predictive capabilities for the default cut-off (see Methods). TiC-Onco (2) shows the predictive capabilities for the cut-off providing the best Youden’s Index. AUC Area Under the Roc Curve, PPV Positive Predictive Value, NPV Negative Predictive Value, PLR Positive Likelihood Ratio, NLR Negative Likelihood Ratio

The TiC-Onco score showed an AUC of 0.73 (0.67–0.79), a sensitivity of 49%, and a specificity of 81%. Its PPV was 37%, NPV 88%, PLR 2.6, and NLR 0.6% (Table 4). The TiC-Onco score showed a significantly lower capacity to distinguish between patients who experienced/did not experience a VTE event (AUC 0.73 vs. 0.58; p = 0.001). The sensitivity of the TiC-Onco score was significantly higher than that of the Khorana (49 vs. 22%; p < 0.001), while the specificities of both scores were similar (81 vs. 82%; p = 0.823). The PPV and NPV of the TiC-Onco score were significantly higher than those of the Khorana score (37 vs. 22%; p = 0.004 for PPV and 88 vs. 82%; p < 0.001 for NPV). The LRs of the TiC-Onco score were also significantly better (Table 4).

Table 5 shows the distribution of patients with or without VTE according to the Khorana score. The great majority of patients who suffered a VTE event (77%) were identified by the Khorana score as being at low or moderate risk (values 0, 1, and 2). Among these 55 patients, however, 17 (31%) were detected as high-risk patients by the TiC-Onco score. When the cut-off for high risk was taken as the best Youden Index, the TiC-Onco score returned significantly better predictions of risk than the Khorana score, especially in terms of sensitivity (86 vs. 22%, p < 0.001) (Table 4). In this scenario, of the 55 patients who experienced a VTE event (but who were classified as not being at high-risk by the Khorana score), 40 (73%) were detected as high risk patients by the TiC-Onco score. Table 6 shows rates of VTE according to prespecified risk categories for both TiC-Onco and Khorana.

The NNT values for: (a) if all patients included in the study had been treated (NNT = 12); (b) if only the patients with a Khorana score of ≥3 had been treated (NNT = 10), or (c) if only patients with a high risk TiC-Onco score (with the cut-off set at the same specificity as the Khorana score) had been treated (NNT = 6).

DISCUSSION

When deciding whether to use primary antithrombotic prophylaxis in outpatients with cancer who are candidates for chemotherapy, a clinician needs to determine the risk of VTE and weigh the likely benefit against the risk of bleeding. Despite the awareness of scientific societies regarding cancer-associated VTE, thromboprophylaxis is limited among outpatients, probably due to the sub-optimal predictive capacity of the existing tools used to predict the risk of experiencing a VTE event. The present work presents a new predictive score, the TiC-Onco score, which shows significantly better predictive power in this regard than the Khorana score.

The incidence of VTE in the present population at 6 months of follow-up was 18% (occurring in 71/391 patients); this is within the range of figures cited in a previous publication for the same follow-up period. It is also in agreement with the observation that the incidence of VTE is highest among patients with pancreatic cancer. However, no other clear differences between tumour types were seen with respect to the Khorana score, probably because a large proportion (46%) of the present patients had stage IV tumours.

The statistical analysis of the present data detected four genetic variants that were independently associated with VTE in outpatients with cancer (Table 2). These were combined into the algorithm for the TiC-Onco score, which initially allowed the patients to be classified as either at high or low risk of VTE. Among the patients in the TiC-Onco high risk group, 37% eventually suffered a VTE event, while 12% of those in the low risk group experienced the same (Table 4). However, when the three-tier Tic-Onco risk category system was contemplated (explained in Methods), 37% of the high risk, 18% of the moderate risk, and 6% of the low risk patients experienced a VTE event. In comparison, 22, 20, and 13% of the patients in the equivalent
Khorana score categories experienced an event. The result obtained for the high risk Khorana score ≥3 (22%) is similar to that reported for high risk group at 6 months by other authors (18%) (p = 0.6).26, 27

The majority of genetic studies have excluded individuals with cancer-related thrombosis, and the relatively few studies that have been performed (which have mainly focused on the factor V Leiden and prothrombin G20210A genetic variants), have reported conflicting results.1, 11, 12, 28-30 These discrepancies are most likely due to the use of a single-marker, and inherent problems of low statistical power and poor reproducibility. The present work overcomes these problems by using several markers based on previous knowledge, a strategy that provided good results in previous work performed with non-oncology patients.13

Although the presently noted distribution of patients in the different Khorana risk categories is similar to that previously reported for the same follow-up time by other authors,26, 27 the predictive power of the TiC-Onco was found to be significantly greater than that of the Khorana Score. This superiority is demonstrated by a better AUC, better likelihood ratios, a higher PPV and NPV, and, importantly, a much higher sensitivity (49%).

In summary, this paper reports a clinical-genetic risk score that is significantly better than the Khorana score at identifying outpatients with cancer at high risk for experiencing a VTE event. Patients identified as being at high risk by the TiC-Onco risk score (using a specificity equal to that provided by the Khorana index as the cut-off) would likely benefit from thromboprophylaxis despite the risk of haemorrhage; they should, therefore, be seen as candidates for prophylactic treatment for VTE. The lower NNT of the TiC-Onco score reveals it can identify those patients most likely to benefit from prophylaxis. It is important that an accurate predictive tool like the TiC-Onco score be available if the morbidity associated with VTE is to be reduced, and because patients with cancer who experience a VTE are more likely to die than those who do not (31 vs. 11% in the present study; p < 0.001).

The reduced number of cancer types included might be seen as a weakness of the present study. However, it has the strengths of a multi-site setting, and a large proportion of patients in the advanced stages of tumour node metastasis classification.

As the risk of cancer-associated VTE is high even 6 months before cancer diagnosis, and the peak incidence is from 0 to 6 months1, 31 post-diagnosis, it is recommended that the TiC-Onco score be calculated at the moment cancer is suspected.32, 33

ACKNOWLEDGEMENTS
The authors thank the Sociedad Española de Oncología Médica, the European Society of Cardiology, the Fundación Josep Carreras contra la Leucemia, the Sociedad Española de Trombosis y Hemostasia, CERC Programme/Generalitat, Ferrer inCode, and Leo-Pharma for funding and institutional support. The authors thank Adrian Burton (www.physicaledvidence.es) for editorial assistance. The study has been presented in part in: ISTH congress (Toronto-2015), ICTHIC congress (Bergamo-2016), SEOM congress (Madrid-2016), WHITH congress (Barcelona-2017), and ISTH (Berlin-2017).

### Table 6. Percentage of study population among risk categories, and percentage of patients with VTE. See Methods for details about the definition of two and three categories

|                | TiC-Onco |                 | Khorana |                 |
|----------------|----------|-----------------|---------|-----------------|
|                | % of cancer population | % of cancer patient with VTE | % of cancer population | % of cancer patient with VTE |
| High risk      | 24.30    | 36.84           | 18.93   | 21.62           |
| Non-high risk  | 75.70    | 12.16           | —       | —               |
| Moderate risk  | 39.13    | 18.30           | 52.94   | 19.81           |
| Low risk       | 36.57    | 5.59            | 27.62   | 12.96           |

### ADDITIONAL INFORMATION

#### Competing interests:
A.M., J.M.S., and E.S. are inventors in a patent application claiming the use of TiC-Onco. A.M. and Jose M.S will receive royalties for giving the right to Gendiag of being the assignee of that patent application. Israel Ortega and Eduardo Salas are employees of Gendiag. C.F., V.P., V.C., M.S., E.M., J.C., A.R., M.M., and J.C.S. have no conflict of interest.

**Funding:** This work was supported by Leo Pharma; Ferrer inCode; Fundación Josep Carreras; European Society of Cardiology; and the Spanish Society of Thrombosis and Haemostasis.

#### REFERENCES

1. Blom, J. W., Doggen, C. J. M., Osanto, S. & Rosendaal, F. R. Malignancies, pro-thrombotic mutations, and the risk of venous thrombosis. JAMA 293, 715–722, (2005).
2. Khorana, A. A., Dalal, M., Lin, J. & Connolly, G. C. Incidence and predictors of venous thromboembolism (VTE) among ambulatory high-risk cancer patients undergoing chemotherapy in the United States. Cancer 119, 648–655, (2013).
3. Noble, S. & Pasi, J. Epidemiology and pathophysiology of cancer-associated thrombosis. Br. J. Cancer 102, 52–59, (2010).
4. Lyman, G. H. et al. Venous thromboembolism prophylaxis and treatment in patients with cancer: American Society Of Clinical Oncology Clinical Practice Guideline Update 2014. J. Clin. Oncol. 33, 654–656, (2015).
5. Muñoz Martín, A. J., Font Puig, C., Navarro Martín, L. M., Borrega García, P. & Martín Jiménez, M. Clinical guide SEOM on venous thromboembolism in cancer patients. Clinical Transl. Oncol. 16, 1079–1090, (2014).
6. Falanga, A., Roila, F. & Mandala, M. Management of venous thromboembolism (VTE) in cancer patients: ESMO Clinical Practice Guidelines clinical practice guidelines. Ann. Oncol. 22, 85–92, (2011).
7. Khorana, A. A., Kuderer, N. M., Culakova, E., Lyman, G. H. & Francis, C. W. Development and validation of a predictive model for chemotherapy-associated thrombosis.Blood 111, 4902–4908 (2015).
8. Muñoz Martín, A. J. et al. Incidence of venous thromboembolism (VTE) in ambulatory pancreatic cancer patients receiving chemotherapy and analysis of Khorana’s predictive model. Clinical & Transl. Oncol. 16, 927–930, (2014).
9. Srikanthan, A. et al. Large retroperitoneal lymphadenopathy as a predictor of venous thromboembolism in patients with disseminated germ cell tumors treated with chemotherapy. Journal. clinical Oncol. 33, 582–587, (2015).
10. Ugarte Fornell, G. et al. Predictive Khorana’s model in patients with venous thromboembolic disease and cancer. Med. clinica 141, 479–481, (2013).
11. Pabinger, I. et al. Factor V Leiden mutation increases the risk for venous thromboembolism in cancer patients-results from the Vienna Cancer And Thrombosis Study (CATS). J. Thromb. Haemost. 13, 17–22, (2015).
12. Tinholt, M., Sandset, P. M. and Iversen, N. (2016) Polymorphisms of the coagulation system and risk of cancer, Thromb. Res. 149, S49-S54. https://doi.org/10.1016/S0049-3848(16)30098-6.
13. Soria, J. M. et al. Multilocus genetic risk scores for venous thromboembolism risk assessment. J. Am. Heart Assoc. 3, e001060, (2014).
14. Falanga, A. & Russo, L. Epidemiology, risk and outcomes of venous thromboembolism in cancer. Harnostaseologie 32, 115–125, (2012).
15. Khorana, A. A., Kuderer, N. M., Culakova, E., Lyman, G. H. & Francis, C. W. Development and validation of a predictive model for chemotherapy-associated thrombosis. Blood 111, 4902–4907, (2008).
16. Venables, W. N. and Ripley, B. D. (2002) Modern Applied Statistics with S: Springer New York (Statistics and Computing). https://doi.org/10.1007/978-0-387-21706-2.
17. Smith, G. C., Seaman, S. R., Wood, A. M., Royston, P. & White, I. R. Correcting for optimistic prediction in small data sets. Am. J. Epidemiol. 180, 318–324, (2014).
18. Hanley, J.A. & Hajian-Tilaki, K.O. Sampling variability of nonparametric estimates of the areas under receiver operating characteristic curves: an update. *Acad. Radiol.* **4**, 49–58, (1997).

19. Attia, J. Moving beyond sensitivity and specificity: using likelihood ratios. *Aust. Prescr.* **26**, 111–113 (2003).

20. Schisterman, E. F., Perkins, N. J., Liu, A. & Bondell, H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology* **16**, 73–81 (2005).

21. R Development Core Team (2015) *R: a Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, 2015).

22. Sackett, D. L. & Haynes, R. B. Summarizing the effects of therapy: a new table and some more terms. *Acp. J. Club.* **127**, A15–A16 (1997).

23. Di Nisio, M., Porreca, E., Otten, H.-M. & Rutjes, A. W. S. Primary prophylaxis for venous thromboembolism in ambulatory cancer patients receiving chemotherapy. *Cochrane Database Syst. Rev.* **8**, CD008500, (2014).

24. Streiff, M. B. et al Cancer-associated venous thromboembolic disease, version 1.2015. *Journal. Natl. Compr. Cancer Netw.* **13**, 1079–1095 (2015).

25. Qureshi, W. et al. Venous thromboembolism in cancer: an update of treatment and prevention in the era of newer anticoagulants. *Front. Cardiovasc. Med. Front.* **3**, 24, (2016).

26. Ay, C. et al. Prediction of venous thromboembolism in cancer patients. *Blood* **116**, 5377–5382, (2010).

27. Thaler, J., Ay, C. & Pabinger, I. Venous thromboembolism in cancer patients- Risk scores and recent randomised controlled trials. *Thromb. Haemost.* **108**, 1042–1048, (2012).

28. Phiusch, R. et al. Impact of thrombophilic gene mutations on thrombosis risk in patients with gastrointestinal carcinoma. *Cancer* **94**, 3120–3126, (2002).

29. Gran, O. V. et al. Joint effects of cancer and variants in the factor 5 gene on the risk of venous thromboembolism. *Haematologica* **101**, 1046–1053, (2016).

30. Kennedy, M. et al. Factor V Leiden, prothrombin 20210A and the risk of venous thrombosis among cancer patients. *Br. J. Haematol.* **128**, 386–388, (2005).

31. Khorana, A. A., Dalal, M., Lin, J. & Connolly, G. C. Incidence and predictors of venous thromboembolism (VTE) among ambulatory high-risk cancer patients undergoing chemotherapy in the United States. *Cancer* **119**, 648–655, (2013b).

32. Gran, O. V. et al. Joint effects of cancer and variants in the factor 5 gene on the risk of venous thromboembolism. *Haematol. Haematol.* **101**, 1046–1053, (2016).

33. Lyman, G. H., Bohle, K. & Falanga, A. American Society of Clinical Oncology: Venous thromboembolism prophylaxis and treatment in patients with cancer: American Society of Clinical Oncology clinical practice guideline update. *Journal. Oncol. Pract.* **11**, e442–e444, (2015).