Medicine®

OPEN

Derivation and validation of a clinical prediction model for risks of venous thromboembolism in diabetic and general populations

Cheng-Chieh Lin, MD, PhD[a,b,c], Ching-Chu Chen, MD[a,d], Chia-Ing Li, PhD[a,c], Chiu-Shong Liu, MPH, MD[a,b,c], Wen-Yuan Lin, MD, PhD[a,b], Chih-Hsueh Lin, MD, PhD[a,b], Shing-Yu Yang, BS[e], Tsai-Chung Li, PhD[e,f,*] (iD)

Abstract

Most studies on the prediction of venous thromboembolism (VTE) focused on hospitalized, surgery, and cancer patients or women receiving hormonal contraceptives or menopausal hormone therapy. No study considered diabetic and general populations to establish a VTE prediction model, especially in Asia. We developed a predictive model for VTE among type 2 diabetic patients and the general population.

This study considered 2 nationwide retrospective cohort studies consisting of 52,427 diabetic participants and 508,664 participants from the general population aged 30 to 85 years during 2001 to 2004 in Taiwan. All participants were followed up until VTE event, death, or December 2011. The outcome event was VTE, including deep venous thrombosis and pulmonary embolism. Candidate predictors consisted of socio-demographic factors, diabetes-related factors and biomarkers, comorbidities, and medicine use. Our study followed the procedures proposed by the Framingham Heart Study to develop prediction models by using a Cox regression model. The predictive accuracy and performance characteristics were assessed using the area under curve of receiver operating characteristics curve and calibration of a risk score were performed by Hosmer–Lemeshow goodness-of-fit test.

The common factors for persons with type 2 diabetes and general population included age, hospitalization status 1 year before the baseline, hypertension, chronic kidney disease, chronic obstructive pulmonary disease, and anti-diabetes medications; the specific factors for persons with type 2 diabetes consisted of body mass index, glycosylated hemoglobin A1C, and creatinine; and the factors for general population included gender, peripheral vascular disease, cancer, hypertension medication, cardiovascular medication, and non-steroidal anti-inflammatory drug. The area under curve of 3-, 5-, and 8-year VTE prediction models were 0.74, 0.71, and 0.69 in the diabetic population and 0.77, 0.76, and 0.75 in the general population, respectively.

The new clinical prediction models can help identify a high risk of VTE and provide medical intervention in diabetic and general populations.

Abbreviations: AUC = area under curve, BMI = body mass index, CI = confidence interval, COPD = chronic obstructive pulmonary disease, CVD = cardiovascular disease, DVT = deep vein thrombosis, HbA1c = glycosylated hemoglobin A1C, ICD-9-CM = International Classification Disease, Ninth Revision, Clinical Modification, LHID2000 = Longitudinal Health Insurance Database 2000, NDCMP = National Diabetes Care Management Program, NHI = National Health Insurance, NHIRD = National Health Insurance Research Database, NSAIDs = non-steroidal anti-inflammatory drugs, PE = pulmonary embolism, ROC = receiver operating characteristic, SDs = standard deviations, TDS = Taiwan Diabetes Study, VTE = venous thromboembolism.

Keywords: cohort studies, diabetes mellitus, type 2, venous thromboembolism

1. Introduction

Cardiovascular disease (CVD) is the leading global cause of death. Venous thromboembolism (VTE), a common CVD with complex and multifactorial aspects and a leading cause of morbidity and mortality, has an annual incidence of 1 to 2 new cases per 1000 persons.[1,2] VTE is the third most common vascular disorder after myocardial infarction and stroke in the world.[3] VTE is a blood clot that starts from a vein that includes deep vein thrombosis (DVT) in the lower legs, which is caused by disorders, such as interrupted blood flow, damage of the vessel wall, and hypercoagulability.[4] Pulmonary embolism (PE) is aroused by a blood clot secluded from a venous thrombus in the lower legs that moves into the pulmonary artery through the right heart.[4] In Taiwan, the crude incidence of VTE ranges from 15.9 to 16.5 per 100,000 person-years and increases with age.[5,6] The overall 1-month mortality rates for VTE, DVT, and PE are 8.8%, 7.1%, and 12.9%, respectively.[5] VTE is a major challenge to public health and healthcare systems due to frequent morbidity and a high mortality rate.

VTE prediction models can guide health professionals to provide appropriate thromboprophylaxis and thus reduce morbidity and health care costs. Recent studies have developed VTE prediction models in high-risk groups, such as patients with Cushing syndrome,[7] women receiving hormonal contraceptives or menopausal hormone therapy,[8,9] patients with surgery therapy,[10–12] patients with cancer,[13,14] and hospitalized patients.[15–20] Several risk assessment models, including Caprini and Padua score and Kucher, Padua, Improve, and Intermountain score, have been developed to predict the risk of an individual for VTE in clinical practice. These existing risk assessment models for VTE have been applied to divide hospitalized patients into VTE risk subgroups of low, moderate, and high. DVT is approximately 10% to 40% among medical or general surgical patients or hospitalized patients.[21] PE is approximately 5% to 10% of all deaths among hospitalized patients.[22] On the basis of the findings of a retrospective cohort study, approximately 25% of patients with diabetes are estimated to experience at least 1 hospital admission.[23] Acute changes in blood glucose and insulin resistance acutely intensify the hypercoagulable state and then increase the risk of VTE.[24] Although diabetes is a risk factor of VTE, the determinants of VTE in patients with type 2 diabetes remain unknown. Glucose control has been associated with CVDs, including stroke, coronary heart disease, and end-stage renal disease.[25–27] Thus, glucose control may be associated with the development of VTE. However, this association has never been examined by prior studies.

The Framingham risk score is the most famous point system based on complex statistical models.[28] The system can help clinicians and patients estimate disease risk and monitor this risk over time. The Framingham Heart Study has already developed prediction models for many heart diseases, such as coronary heart disease,[29] CVD,[30] atrial fibrillation,[31] hypertension,[32] and stroke,[33] but not for VTE. Furthermore, many previous studies have developed and validated risk prediction algorithms for VTE in patients with Cushing syndrome,[7] surgery,[10–12] cancer,[13,14] and hospitalization,[15–20] but no study has included diabetic and general populations. Thus, development and validation of score systems to predict VTE in patients with type 2 diabetes and in the general population are needed. The prediction models included risk factors that are generally accepted and available in clinical practice for the general population and additional variables of combined glycemic control and management of diabetes for patients with type 2 diabetes.

The studies for VTE prediction model were limited in Asian patients, mainly based on Western populations. Previous studies validated a risk prediction model for VTE in hospitalized Chinese patients[16] and the population of United States,[10,13,15,17–19] Italy,[7] Sweden,[8] Switzerland,[9] Canada,[12] Austria,[14] Spanish,[20] and 3 European countries.[11] However, to the best of our knowledge, no study has reported a risk score system for VTE in diabetic or general populations. Therefore, we created 2 VTE risk tools in patients with type 2 diabetes by using a nationwide cohort, Taiwan Diabetes Study (TDS), and in the general population by using a large national cohort, Longitudinal Health Insurance Database 2000 (LHID2000).

2. Methods

2.1. Data sources and study subjects

2.1.1. TDS cohort.
A retrospective cohort study was conducted including all enrollees in the National Diabetes Care Management Program (NDCMP) in Taiwan from 2002 to 2004. The NDCMP is a case management program for diabetic patients. Patients diagnosed with diabetes based on the criteria established by the American Diabetes Association (International Classification Disease, Ninth Revision, Clinical Modification, abbreviated as ICD-9-CM Code of 250) were included as study subjects. Index date was defined as the entry date into the NDCMP. Each individual was followed up from the date of entry until December 31, 2011. These patients were monitored for withdrawal from the National Health Insurance (NHI) program, death, or development of VTE events. All patients with type 2 diabetes in the NDCMP from 2002 to 2004 were included. The study consisted of 63,084 enrolled diagnosed patients with type 1, type 2, or gestational diabetes. We included 59,721 patients who were diagnosed as having type 2 diabetes, experienced no VTE event, and aged 30 to 85 years old at baseline. We further excluded 7297 patients who showed missing information regarding baseline characteristics, diabetes-related variables, comorbidities, and medication use. As shown in Figure S1, Supplemental Digital Content, http://links.lww.com/MD2/A499, 52,427 patients with type 2 diabetes were included in the study.

2.1.2. LHID2000 cohort.
Taiwan's National Health Insurance Research Database (NHIRD), one of the administrative health care databases, is derived from the registration files and original claims data. LHID2000 is a data set released by the NHIRD that includes 1 million people who were randomly selected from the entire 23 million Taiwan residents in 2000. The LHID2000 was combined to obtain information on demographic data, date and institution for diagnosis, out-patient visit, in-patient admission, out-patient and in-patient treatment, and the comprehensive health assessment. The datasets were included from January 2000 to December 2011. We identified 508,664 individuals who showed no VTE event and aged 30 to 85 years old. We handled 19 individuals who showed missing baseline characteristics using multiple imputations. As shown in Figure S1, Supplemental Digital Content, http://links.lww.com/MD2/A499, 508,664 individuals of the general population were included in the study analysis.

All participants of both general and diabetic populations for analysis were randomly allocated into the derivation and

validation sets in a 2:1 ratio. Ethics approval was obtained from the Ethical Review Board of China Medical University Hospital (CMUH107-REC2-158).

2.2. Measurements

2.2.1. Sociodemographic factors and biomarkers. In the TDS cohort, we retrieved socio-demographic factors, lifestyle behaviors, and anthropometric measurements in the NDCMP database. These variables included age, gender, duration of diabetes, smoking habits, and alcohol drinking. Smoking habits and alcohol drinking were categorized into 2 levels: yes and no. The number of days between dates for NDCMP enrollment and diabetes onset was calculated as the duration of diabetes. The number of days was then divided by 365 days. Measured body mass index (BMI), glycosylated hemoglobin A1C (HbA1c), fasting plasma glucose, diastolic blood pressure, systolic blood pressure, total cholesterol, high density lipoprotein, low density lipoprotein, and triglycerides, serum glutamate-pyruvate transaminase, and creatinine at the index date were also retrieved. In the LHID2000 cohort, we retrieved socio-demographic factors in the database of NHIRD, including age and gender.

2.2.2. Comorbidities, hospitalization, and surgery for both TDS and LHID2000. Baseline comorbidities, hospitalization status, and surgery status were derived using outpatient and inpatient data within 1 year before enrollment. The criteria to define status of comorbidities at baseline were for those who experienced at least 3 service claims for ambulatory care or 1 service claim for inpatient care. In Taiwan, the ICD-9-CM was converted to ICD-10-CM/PCS by the NHI Agency in 2016 for reimbursement systems of medical expenses of outpatient, inpatients, and emergency services, thus ICD-10-CM was not available in the present study and ICD-9-CM codes were used for comorbidity ascertainment. Comorbidity measures included hypertension (ICD-9-CM codes 404-405), CVD (ICD-9 codes 410–413, 414.01–414.05, 414.8, and 414.9), stroke (ICD-9-CM codes 431–438), atrial fibrillation (ICD-9-CM code 427.3), peripheral neuropathy (ICD-9-CM code 356), hypoglycemia (ICD-9-CM codes 250.3, 250.8, 251.0–251.2, 270.3, 775.0, 775.6, and 962.3), diabetes retinopathy (ICD-9-CM code 362.0), chronic kidney disease (ICD-9 code 585), lower limb amputation (ICD-9-CM codes 895–897), ketoacidosis (ICD-9-CM code 250.1), postural hypotension (ICD-9-CM code 458), peripheral vascular disease (ICD-9-CM code 443), hyperlipidemia (ICD-9-CM code 272), chronic obstructive pulmonary disease (COPD) (ICD-9-CM code 490–496), asthma (ICD-9-CM code 493), and cancer (ICD-9-CM codes 140–165, 170–175, 179–200, 202, 203, 210–213, 215–229, 235–239, 654.1, 654.10, 654.11, 654.12, 654.13, and 654.14). Hospitalization status and surgery status were defined using procedure codes in inpatient claims.

2.2.3. Medication use for both TDS and LHID2000. Data on medication use prescribed for the treatment of disease were derived for a 1-year period preceding the cohort entry. When a patient used drugs for more than 3 months, he/she was defined as a user of this specific drug. Medication use included anti-diabetes medications, hypertension medications, hyperlipidemia medications, cardiovascular medications, non-steroidal anti-inflammatory drugs (NSAIDs), and aspirin. The medications for diabetes were classified into no medication, oral anti-diabetes drug, insulin only, and insulin plus oral anti-diabetes drug. The medications for hypertension included angiotensin converting enzyme inhibitors, angiotensin II receptor blockers, β-blockers, calcium channel blockers and diuretics. The medications for CVD included anti-arrhythmic, anti-coagulants, anti-platelet, digoxin, and nitrates. The medications for hyperlipidemia consisted of fibrates and statins.

2.2.4. Outcome ascertainment for both TDS and LHID2000. The major outcome was the first diagnosis of VTE, which was measured through record linkage with inpatient care dataset or enrollment dataset in the NHIRD. We identified VTE events using ICD-9 codes (451 and 453 for DVT and 415 for PE)[34] and ICD-10 codes (I80 and I82 for DVT and I26 for PE). The outcome event was defined as patients showed at least 3 ambulatory visits or at least 1 inpatient claim of VTE event after the index date to increase the true positive rate. The time of follow-up began with recruitment (index date) and ended with death or a new VTE event, withdrawal from the insurance program, or the end of follow-up.

2.3. Statistical analysis

We presented proportions for categorical variables and means and standard deviations (SDs) for continuous variables. We calculated standardized effect sizes to assess the comparability of the baseline factors between the validation and derivation sets. Standardized effect sizes of less than 0.1 indicate that the differences between derivation and validation sets are trivial. We applied Cox proportional hazards models to estimate the hazards ratios of predictor variables for the development of a VTE prediction model in the derivation set and to evaluate the predictive ability of the model in the validation set. A sensitivity analysis was further performed by considering the competing risks of death with extended Cox proportional hazards models and by using multiple imputation approach to handle missing data.

The approach to choose and include the variables into the predictive model is through the process of reviewing literature to make sure the variables in our dataset were of clinical relevance and biological plausibility. If the variables fulfil these criteria, they were included into models for assessing their statistical significance. The steps in the development of the predictive model are based on the Framingham heart study,[28] guiding us in the assignment of the VTE risk score (Supplemental A, Supplemental Digital Content, http://links.lww.com/MD2/A498). We determined the receiver operating characteristic (ROC) curves for the predictive accuracy evaluation and area under curves (AUCs) of ROC curves for the discriminatory ability assessment of the predictive model. Harrell estimator of the c index was also applied to time-to-event data. The Kaplan–Meier curves stratified for the 3 risk categories were estimated for assessing the discriminatory capability of the risk model for the full cohort. Goodness-of-fit by Hosmer–Lemeshow x^2 tests was applied by comparing the observed and predicted VTE events. Bootstrap resampling approach based on 1000 samples was carried out to assess the internal validity of model performance in term of the potential for over fitting or "optimism." All statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC). Two-tailed $P < .05$ denotes statistical significance.

3. Results

We identified 731 (84.3% DVT and 15.7% PE) and 4904 (81.5% DVT and 18.5% PE) individuals with VTE events from

the TDS and LHID2000 cohorts at the end of follow-up in 2011, respectively. The 52,427 patients with type 2 diabetes from the TDS cohort showed a mean age of 60.9 years (SD of 11.0 years), and 47.1% of these patients were male. The 508,664 individuals from the LHID2000 cohort showed a mean age of 48.1 years (SD of 13.6 years), and 50.6% of these individuals were male. The baseline sociodemographic and clinical characteristics of the 2 cohorts are displayed in Table 1. In both cohorts, the derivation and validation sets showed similar distributions of observed baseline variables, and all variables showed standardized differences of less than 0.05.

Table 2 presents the significant baseline predictors of VTE risk in the univariate and multivariate proportional hazards regression analyses from the TDS and LHID2000 cohorts. In the 2 cohorts, age, hospitalization status 1 year before the baseline, comorbidities of hypertension, chronic kidney disease, COPD, anti-diabetes medications, and cardiovascular medications were both associated with an increased risk of VTE. In addition, BMI ≥ 30 kg/m^2, HbA1c $\geq 9\%$, and creatinine ≥ 2 mg/dL were associated with increased risks of VTE in patients with type 2 diabetes, whereas peripheral vascular disease, cancer, hypertension medication, and NSAID were significantly associated with increased VTE risks. Point assignment based on the final Cox proportional hazards model is shown in Table 3. For patients with type 2 diabetes, individual VTE risk scores were assigned based on a –4 to 48 point scoring system of risk factors, including age (–2 to 8 points), gender (1 point), BMI ≥ 30 kg/m^2 (4 points), hospitalization status 1 year before the baseline (4 points), HbA1c (0–6 points), creatinine (0–9 points), hypertension (2 points), chronic kidney disease (6 points), COPD (3 points), anti-diabetes medications (–2 to 3 points), and cardiovascular medications (2 points). For the general population, individual VTE risk scores were assigned based on a –2 to 30 point scoring system of risk factors, including age (–2 to 8 points), gender (1 point), hospitalization status 1 year before the baseline (2 points), hypertension (1 point), chronic kidney disease (6 points), peripheral vascular disease (2 points), COPD (1 point), cancer (2 points), anti-diabetes medications (0–3 points), hypertension medications (1 point), cardiovascular medications (1 point), and NSAID (2 points). The 3-, 5-, and 8-year estimated VTE risks were determined by total points and calculated with the equations for the Framingham Heart Study (Table S1, Supplemental Digital Content, http://links.lww.com/MD2/A503). The baseline predictors of VTE risk based on the final multivariate Cox proportional hazards model with competing risks death were presented in Table S2, Supplemental Digital Content, http://links.lww.com/MD2/A504. The results are similar to those in the original analysis, indicating the robustness of our study findings.

The ROCs of both VTE risk models in predicting the 3-, 5- and 8-year risks are shown in Figure S2, Supplemental Digital Content, http://links.lww.com/MD2/A500. For patients with type 2 diabetes, the results showed AUCs of 0.74 (95% confidence interval [CI], 0.68–0.79), 0.71 (0.66–0.75), and 0.69 (0.65–0.72) for predicting the 3-, 5- and 8-year risks of VTE in the validation set, respectively. For the general population, it showed AUCs of 0.77 (0.75–0.79), 0.76 (0.74–0.78), and 0.75 (0.74–0.77) for the 3-, 5-, and 8-year VTE risks in the validation set, respectively. Both VTE risk models exhibited good ability for discriminating VTE event. No significant differences between observed and predicted events of VTE were found according to deciles of the 5- and 8-year risk for diabetic population and 3-year risk for the general population and ($P > .05$) (Figure S3,

Supplemental Digital Content, http://links.lww.com/MD2/A501). Multivariable cubic spline plots for VTE risk according to continuous predictors of age, BMI, HbA1c, creatinine from TDS cohort, and age from LHID2000 cohort in validation set were shown in Figure S4, Supplemental Digital Content, http://links.lww.com/MD2/A502.

Figure 1 presents the Kaplan Meier survival curves for VTE risk according to risk tertiles among patients with type 2 diabetes and the general population (both log-rank test, $P < .001$). The medium- and high-risk groups showed increased risks of VTE compared with the low-risk group for patients with type 2 diabetes (hazard ratio, 1.92 [1.54–2.41]; 4.12 [3.41–4.99]) and the general population (2.31 [2.08–2.57]; 9.24 [8.41–10.14]), indicating the predictive abilities of our prediction models. Based on 1000 bootstrap samples, the slope and intercept of the 3-year calibration curve are 0.91, 0.0056 for patients with type 2 diabetes and 1.002, 0.0024 for general population in a 3-year period.

4. Discussion

This study established 2 risk score systems for predicting VTE among patients with type 2 diabetes and the general population based on 2 national population-based databases. Due to the fundamental difference in these 2 populations, 11 candidate predictors were considered for the type 2 diabetic patients and 12 variables for the general population. We found risk factors have a different baseline impact in people with or without diabetes. Eight common predictors were identified to be associated with VTE for both 2 cohorts. The other associated predictors included BMI, HbA1c, and creatinine in type 2 diabetic patients and peripheral vascular disease, cancer, hypertension medications, and NSAID in the general population. In general population, the significant effects of glucose lowering drugs presented the effect of diabetes while the effects of glucose lowering drugs were graduated and diabetes-specific factor such as HbA1c was a strong predictor in persons with diabetes. We identified the 2 best combinations of factors predicting VTE according to the 2 cohorts of diabetes and general populations separately. Moreover, 2 models of diabetes and general populations exhibited good discrimination and calibration with a Harrell C-index of 0.76 (95% CI: 0.75–0.78) and 0.71 (95% CI: 0.68–0.75) in the validation sets, respectively.

Our prediction model for the general population identified that type 2 diabetes is an important predictor of VTE. Patients with type 2 diabetes mellitus encounter a high rate of complications, including CVD. VTE and atherosclerosis may be simultaneously triggered by biological stimuli responsible for activating inflammatory and coagulation pathways in the arterial and venous systems, leading to inflammation, hypercoagulability, and endothelial injury.[35] Therefore, VTE and cardiovascular disorders share common risk factors, including age, obesity, hypertension, dyslipidemia, smoking, and diabetes, and VTE may occur as the first symptomatic cardiovascular event.[36,37] Diabetes is a risk factor for atherosclerosis and VTE.[35] Two systematic review studies showed that persons with diabetes mellitus exhibit a 1.4-fold risk of VTE compared with those without diabetes mellitus,[36,38] which is consistent with our study that considered anti-diabetes medication as well as glucose control status instead of diabetes status. Diabetes is a risk factor for atherosclerosis and VTE show an imbalance of pro- versus anti-coagulation, resulting in hypercoagulability.[39] Hypercoag-

Table 1

Baseline clinical characteristics of 2 study cohort in derivation and validations sets.

Variables	TDS cohort			LHID2000 cohort		
	Derivation set (n = 34,951) Mean ± SD or n (%)	Validation set (n = 17,476) Mean ± SD or n (%)	Standardized effect size	Derivation set (n = 339,109) Mean ± SD or n (%)	Validation set (n = 169,555) Mean ± SD or n (%)	Standardized effect size
Socio-demographic factors						
Age (yrs)	60.88 ± 10.95	60.99 ± 10.94	−0.01	48.12 ± 13.54	48.07 ± 13.56	0.00
Gender						
Female	18,568 (53.13)	9149 (52.35)	0.02	167,480 (49.39)	83,645 (49.33)	0.00
Male	16,383 (46.87)	8327 (47.65)	−0.02	171,629 (50.61)	85,910 (50.67)	0.00
Smoking habit	5319 (15.22)	2700 (15.45)	−0.01			
Alcohol drinking	2913 (8.33)	1508 (8.63)	−0.01			
Age of diabetes onset (yrs)	54.51 ± 11.01	54.57 ± 10.97	−0.01			
Duration of type 2 diabetes (yrs)	6.36 ± 6.34	6.42 ± 6.38	−0.01			
Systolic blood pressure (mm Hg)	134.60 ± 17.88	134.77 ± 18.03	−0.01			
Diastolic blood pressure (mm Hg)	80.00 ± 10.72	80.00 ± 10.64	0.00			
Body mass index (kg/m^2)	25.59 ± 3.77	25.62 ± 3.81	−0.01			
Obesity	12,458 (35.64)	6271 (35.88)	−0.01			
Hospitalization status 1 year before the baseline	7417 (21.22)	3765 (21.54)	−0.01	29,249 (8.63)	14,310 (8.44)	0.01
Surgery 1 year before the baseline	4293 (12.28)	227 (12.69)	−0.01	21,144 (6.24)	10,453 (6.16)	0.00
Diabetes-related factors and biomarkers						
HbA1c level (%)	8.29 ± 2.00	8.26 ± 2.01	0.01			
Fasting blood glucose (mg/dL)	174.81 ± 67.89	174.17 ± 66.86	0.01			
Low-density lipoprotein (mg/dL)	118.15 ± 31.28	117.78 ± 31.25	0.01			
High-density lipoprotein (mg/dL)	46 ± 13.45	46.08 ± 13.78	−0.01			
Creatinine (mg/dL)	1.06 ± 0.63	1.06 ± 0.63	0.00			
Serum glutamate-pyruvate transaminase (u/l)	32.16 ± 32.69	32.52 ± 34.51	−0.01			
Total cholesterol (mg/dL)	195.99 ± 41.15	195.94 ± 41.42	0.00			
Triglyceride (mg/dL)	173.36 ± 132.61	172.46 ± 130.11	0.01			
eGFR (mg/dL)	74.04 ± 22.51	74.07 ± 22.51	0.00			
Comorbidities						
DM				16,648 (4.91)	8209 (4.84)	0.00
Hypertension	15,316 (43.82)	7794 (44.6)	−0.02	37,123 (10.95)	18,387 (10.84)	0.00
Stroke	1873 (5.36)	962 (5.50)	−0.01	7180 (2.12)	3624 (2.14)	0.00
Cardiovascular disease	3086 (8.83)	1512 (8.65)	0.01	7997 (2.36)	4005 (2.36)	0.00
Atrial fibrillation	185 (0.53)	96 (0.55)	0.00	876 (0.26)	441 (0.26)	0.00
Peripheral neuropathy	223 (0.64)	128 (0.73)	−0.01	195 (0.06)	90 (0.05)	0.00
Diabetes retinopathy	481 (1.38)	256 (1.46)	−0.01	858 (0.25)	430 (0.25)	0.00
Hypoglycemia	1319 (3.77)	696 (3.98)	−0.01	563 (0.17)	298 (0.18)	0.00
Chronic kidney disease	369 (1.06)	184 (1.05)	0.00	1605 (0.47)	750 (0.44)	0.00
Lower limb amputation	69 (0.20)	34 (0.19)	0.00	86 (0.03)	40 (0.02)	0.01
Ketoacidosis	295 (0.84)	158 (0.90)	−0.01	110 (0.03)	67 (0.04)	−0.01
Postural hypotension	24 (0.07)	16 (0.09)	−0.01	114 (0.03)	53 (0.03)	0.00
Peripheral vascular disease	247 (0.71)	121 (0.69)	0.00	514 (0.15)	291 (0.17)	−0.01
Hyperlipidemia	8493 (24.30)	4135 (23.66)	0.01	9618 (2.84)	4776 (2.82)	0.00
COPD	1590 (4.55)	771 (4.41)	0.01	13,477 (3.97)	6703 (3.95)	0.00
Asthma	545 (1.56)	259 (1.48)	0.01	5686 (1.68)	2922 (1.72)	0.00
Cancer	770 (2.20)	371 (2.12)	0.01	11,043 (3.26)	5346 (3.15)	0.01
Medication use						
Anti-diabetes medications						
No medication	403 (1.15)	194 (1.11)	0.00	327,346 (96.53)	163,745 (96.57)	0.00
Oral only						
Metformin only	2112 (6.04)	1081 (6.19)	−0.01	675 (0.2)	335 (0.20)	0.00
Sulphonylureas only	5071 (14.51)	2429 (13.9)	0.02	3951 (1.17)	1946 (1.15)	0.00
Other	24278 (69.46)	12181 (69.7)	−0.01	6164 (1.82)	2992 (1.76)	0.00
Insulin	1142 (3.27)	591 (3.38)	−0.01	525 (0.15)	283 (0.17)	−0.01
Insulin + oral agent	1945 (5.56)	1000 (5.72)	−0.01	448 (0.13)	254 (0.15)	−0.01
Hypertension medications	15,460 (44.23)	7694 (44.03)	0.00	27,235 (8.03)	13,551 (7.99)	0.00
Hyperlipidemia medications	11,392 (32.59)	5629 (32.21)	0.01	8101 (2.39)	3959 (2.33)	0.00
Cardiovascular medications	10,109 (28.92)	4865 (27.84)	0.02	15,526 (4.58)	7566 (4.46)	0.01
NSAID	285 (0.82)	119 (0.68)	0.02	435 (0.13)	271 (0.16)	−0.01
Aspirin	5450 (15.59)	2630 (15.05)	0.01	5777 (1.70)	2870 (1.69)	0.00
Outcome						
VTE	487 (1.39)	244 (1.40)	0.00	3314 (0.98)	1590 (0.94)	0.00

COPD = chronic obstructive pulmonary disease, DM = diabetes mellitus, eGFR = estimated glomerular filtration rate, HbA1c = glycosylated hemoglobin A1C, LHID2000 = Longitudinal Health Insurance Database 2000, NSAID = non-steroidal anti-inflammatory drug, SD = standard deviation, TDS = Taiwan Diabetes Study, VTE = venous thromboembolism.

Table 2

Cox model measured hazard ratio and 95% confidence intervals of VTE predictors.

Variables	TDS cohort		LHID2000 cohort	
	Crude HR (95% CI)	Adjusted HR (95% CI)	Crude HR (95% CI)	Adjusted HR (95% CI)
Socio-demographic factors				
Age (yrs)				
30–34	1.96 (0.55, 7.01)	1.80 (0.50, 6.46)	0.44 (0.35, 0.54)‡	0.44 (0.36, 0.55)‡
35–39	2.03 (0.86, 4.77)	1.99 (0.85, 4.70)	0.71 (0.59, 0.86)‡	0.72 (0.60, 0.87)‡
40–44	1.00	1.00	1.00	1.00
45–49	1.26 (0.62, 2.56)	1.21 (0.6, 2.47)	1.39 (1.18, 1.64)‡	1.35 (1.14, 1.59)‡
50–54	1.71 (0.89, 3.27)	1.55 (0.81, 2.97)	1.91 (1.62, 2.26)‡	1.78 (1.51, 2.10)‡
55–59	1.86 (0.97, 3.55)	1.67 (0.87, 3.20)	3.11 (2.64, 3.66)‡	2.74 (2.33, 3.22)‡
60–64	2.17 (1.15, 4.08)*	1.85 (0.98, 3.50)	4.07 (3.49, 4.75)‡	3.41 (2.91, 3.99)‡
65–69	2.63 (1.40, 4.93)†	2.15 (1.14, 4.05)*	5.39 (4.63, 6.28)‡	4.15 (3.54, 4.85)‡
70–74	3.44 (1.84, 6.43)‡	2.89 (1.53, 5.44)†	6.23 (5.34, 7.28)‡	4.52 (3.85, 5.31)‡
75–79	3.48 (1.82, 6.68)‡	2.85 (1.47, 5.51)†	8.36 (7.11, 9.84)‡	5.86 (4.95, 6.93)‡
80–84	6.06 (3.00, 12.22)‡	4.94 (2.42, 10.07)‡	8.84 (7.23, 10.80)‡	6.10 (4.96, 7.49)‡
Gender				
Male	1.00	1.00	1.00	1.00
Female	1.17 (0.98, 1.41)	1.11 (0.93, 1.34)	1.18 (1.10, 1.26)‡	1.17 (1.10, 1.26)‡
Body mass index (kg/m^2)				
<25		1.00		
25–30	1.00 (0.82, 1.22)	1.03 (0.84, 1.26)		
≥30	1.67 (1.31, 2.13)‡	1.73 (1.35, 2.22)‡		
Hospitalization status 1 year before the baseline				
No	1.00	1.00	1.00	1.00
Yes	2.36 (1.96, 2.84)‡	1.81 (1.49, 2.20)‡	2.67 (2.44, 2.92)‡	1.58 (1.44, 1.74)‡
HbA1c (%)				
<6	1.00	1.00		
6–7	1.45 (0.94, 2.22)	1.48 (0.97, 2.28)		
7–8	1.38 (0.90, 2.12)	1.43 (0.93, 2.19)		
8–9	1.43 (0.92, 2.21)	1.51 (0.97, 2.34)		
9–10	1.69 (1.08, 2.64)*	1.78 (1.13, 2.80)*		
≥10	1.90 (1.25, 2.88)†	2.27 (1.48, 3.47)‡		
Creatinine (mg/dL) (ref: <2.0)				
<2.0	1.00	1.00		
2.0–4.0	4.59 (3.23, 6.50)‡	2.74 (1.87, 4.01)‡		
>4.0	5.71 (3.05, 10.70)‡	3.44 (1.76, 6.71)‡		
Comorbidities				
Hypertension				
No	1.00	1.00	1.00	1.00
Yes	1.86 (1.55, 2.22)‡	1.30 (1.07, 1.58)†	3.54 (3.28, 3.82)‡	1.11 (1.00, 1.23)*
Chronic kidney disease				
No	1.00	1.00	1.00	1.00
Yes	5.57 (3.59, 8.62)‡	2.31 (1.42, 3.77)‡	10.83 (8.93, 13.13)‡	4.38 (3.58, 5.35)‡
Peripheral vascular disease				
No			1.00	1.00
Yes			4.32 (2.76, 6.79)‡	1.76 (1.12, 2.76)*
COPD				
No	1.00	1.00	1.00	1.00
Yes	2.24 (1.63, 3.08)‡	1.50 (1.08, 2.09)*	3.26 (2.92, 3.65)‡	1.44 (1.28, 1.62)‡
Cancer				
No			1.00	1.00
Yes			2.40 (2.10, 2.76)‡	1.57 (1.36, 1.80)‡
Medication use				
Anti-diabetes medications				
No medication	0.64 (0.20, 2.03)	0.77 (0.24, 2.48)	1.00	1.00
Oral only				
Metformin only	1.04 (0.65, 1.64)	1.12 (0.70, 1.78)	3.28 (2.06, 5.21)‡	1.20 (0.75, 1.91)
Sulphonylureas only	1.00	1.00	3.74 (3.10, 4.52)‡	1.37 (1.13, 1.66)†
Other	1.14 (0.86, 1.51)	1.12 (0.85, 1.48)	3.17 (2.69, 3.74)‡	1.19 (1.01, 1.41)*
Insulin	2.47 (1.58, 3.87)‡	1.37 (0.86, 2.18)	7.04 (4.67, 10.60)‡	1.99 (1.31, 3.02)†
Insulin + oral agent	2.36 (1.62, 3.44)‡	1.65 (1.12, 2.43)*	4.66 (2.75, 7.87)‡	1.51 (0.89, 2.56)
Hypertension medications				
No			1.00	1.00

Table 2

(continued).

Variables	TDS cohort		LHID2000 cohort	
	Crude HR (95% CI)	Adjusted HR (95% CI)	Crude HR (95% CI)	Adjusted HR (95% CI)
Yes			3.84 (3.54, 4.17)‡	1.21 (1.07, 1.36)†
Cardiovascular medications				
No	1.00	1.00	1.00	1.00
Yes	1.75 (1.46, 2.10)‡	1.32 (1.09, 1.59)†	4.44 (4.03, 4.88)‡	1.32 (1.17, 1.50)‡
NSAID				
No			1.00	1.00
Yes			5.33 (3.44, 8.27)‡	1.81 (1.16, 2.81)†

CI = confidence intervals, COPD = chronic obstructive pulmonary disease, HbA1c = glycosylated hemoglobin A1C, HR = hazard ratio, LHID2000 = Longitudinal Health Insurance Database 2000, NSAID = non-steroidal anti-inflammatory drug, TDS = Taiwan Diabetes Study, VTE = venous thromboembolism.
* $P < .05$.
† $P < .01$.
‡ $P < .001$.

ulability may play an important pathogenic role in the increased risk of VTE.[39]

The known risk factors for VTE reported in previous studies[35,36,38,40] include age, BMI, cancer, surgery, hospitaliza-

tion, hypertension, diabetes, and COPD. Similar to previous studies, our models involve these important risk indicators and additionally consider glycemic control for patients with type 2 diabetes and medication use for both diabetic and general

Table 3

VTE risk score based on the final multivariate Cox proportional hazards model.

Predictive factors	TDS cohort				LHID2000 cohort			
	β (SE)	Mean or proportion	P value	Risk score	β (SE)	Mean or proportion	P value	Risk score
Socio-demographic factors								
Age	0.03 (0.01)	60.88	<.001	−2 to 8	0.05 (0.001)	48.12	<.001	−2 to 8
Gender (female)	0.09 (0.09)	0.53	.32	1	0.16 (0.04)	0.49	<.001	1
Body mass index (kg/m²) (ref: <25)								
25–30	0.02 (0.10)	0.45	.84	0				
≥30	0.55 (0.13)	0.13	<.001	4				
Hospitalization status 1 year before the baseline	0.60 (0.10)	0.21	<.001	4	0.44 (0.05)	0.09	<.001	2
HbA1c (%) (ref: <6)								
6–7	0.39 (0.22)	0.20	.08	3				
7–8	0.34 (0.22)	0.22	.12	3				
8–9	0.40 (0.22)	0.17	.08	3				
9–10	0.56 (0.23)	0.13	.01	4				
≥10	0.80 (0.22)	0.20	<.001	6				
Creatinine (mg/dL) (ref: <2.0)								
2.0–4.0	1.00 (0.20)	0.02	<.001	7				
>4.0	1.23 (0.34)	0.01	<.001	9				
Comorbidities								
Hypertension	0.26 (0.10)	0.44	.009	2	0.11 (0.05)	0.11	.04	1
Chronic kidney disease	0.82 (0.25)	0.01	.001	6	1.48 (0.10)	0.005	<.001	6
Peripheral vascular disease					0.57 (0.23)	0.002	.01	2
COPD	0.42 (0.17)	0.05	.01	3	0.35 (0.06)	0.04	<.001	1
Cancer					0.45 (0.07)	0.03	<.001	2
Medication use								
Anti-diabetes medications								
No medication	−0.22 (0.59)	0.01	.71	−2	1.00			0
Oral only								
Metformin only	0.12 (0.24)	0.06	.60	1	0.18 (0.24)	0.002	.45	1
Sulphonylureas only	1.00			0	0.33 (0.10)	0.012	<.001	1
Other	0.11 (0.14)	0.69	.44	1	0.21 (0.09)	0.018	.01	1
Insulin	0.31 (0.24)	0.03	.19	2	0.73 (0.21)	0.002	<.001	3
Insulin + oral agent	0.50 (0.20)	0.06	.01	3	0.45 (0.27)	0.001	.10	2
Hypertension medications					0.18 (0.06)	0.08	.004	1
Cardiovascular medications	0.27 (0.10)	0.29	.004	2	0.25 (0.06)	0.05	<.001	1
NSAID					0.56 (0.23)	0.001	.01	2

COPD = chronic obstructive pulmonary disease, HbA1c = glycosylated hemoglobin A1C, LHID2000 = Longitudinal Health Insurance Database 2000, NSAID = non-steroidal anti-inflammatory drug, TDS = Taiwan Diabetes Study, VTE = venous thromboembolism.
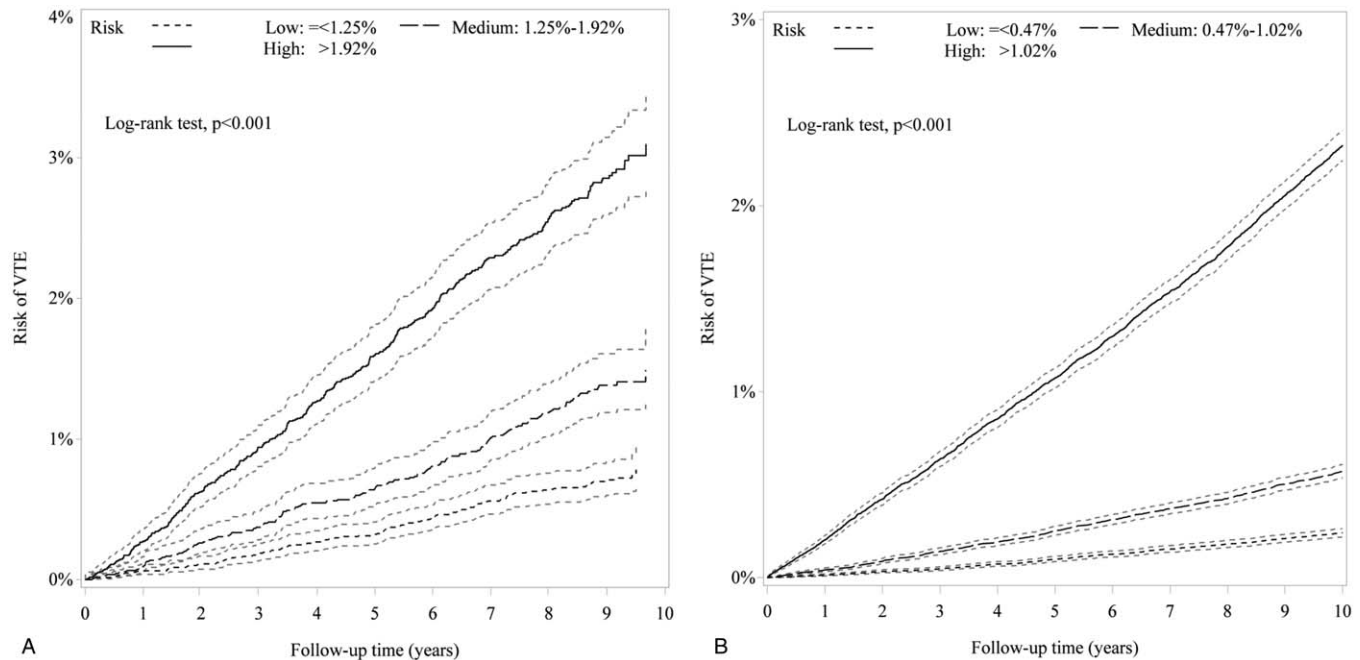
Figure 1. Kaplan-Meier survival curves for VTE risk according to the low, medium, and high-risk groups in (A) the patients with type 2 diabetes and (B) the general population. VTE = venous thromboembolism.

populations. HbA1c is an independent predictor for CAD in patients with type 2 diabetes,[25–27] and a systematic review demonstrated that the use of NSAIDs increases the risk of VTE,[41] which is in line with our study findings that HbA1c and NSAID are significant predictors for VTE risk in diabetic and general populations, respectively.

Most existing prediction models were used to predict the VTE risk in patients with surgery therapy,[10–12] cancer,[13,14] and hospitalization.[15–20] Theses current prediction models for VTE are not applicable in the diabetic and general populations.[10–20] Therefore, the development of VTE prediction in the diabetic and general populations is necessary, which could help healthcare providers to appropriately risk-stratify patients and systematically prevent VTE occurrence.[28] Our 2 predictive models for 3- and 5-year periods showed AUC values greater than 0.70, which indicated their good discrimination ability as a screening tool to identify diabetic patients and general population at high risk for VTE.

Our study showed several strengths. First, this study included 2 nationwide population-based cohorts to develop and validate 2 prediction models for the diabetic and general populations separately. Second, NHIRD is a health care administrative database that is a highly reliable source of information. In addition to traditional factors, we identified the novel predictors of glycemic control and medication use to establish a predictive model for patients with type 2 diabetes. We considered traditional risk factor for VTE risk estimation for easy use in clinical practice. All the risk factors in our score system are accessible to any clinicians and biomarkers available at every endocrinology clinic. Finally, to our knowledge, these 2 prediction models are the first ones to be developed and validated in large unselected patients with type 2 diabetes and a representative sample of general population with retrospective cohort studies. Our results demonstrated that our prediction model stratifies patients who are at risk of VTE and who would benefit the most from treatment.

Some limitations of our study merit comments. First, this study used the baseline measures as predictors and ignored the possibility of time-varying effects of these predictors. Second, although good performance of models was shown in the internal validation, external validation could not be reported due to unavailability of an external dataset. Given that predictors considered in this study can be collected in clinical routine practice, the external validation can be validated by any researcher who owns clinical data for general population and patients with type 2 diabetes. Third, all of these clinical variables were defined using the claims database for the general population. However, the claims data could not provide information on mobility or confined to bed, smoking, family history of VTE, and genetic factors. Finally, the inherent limitations of a retrospective cohort study with information involving claim data dependent on ICD-9-CM codes are present, which may thus result in potential under coding or over coding of disease status. Nevertheless, charts are randomly reviewed by experts of the NHI Bureau, and patients are interviewed for verification of the diagnosis accuracy. Audits were performed for hospitals with outlier chargers or practice. If malpractice or discrepancies were found, heavy penalties were given. These procedures taken by NHI enhanced the accuracy of NHIRD datasets. In addition, the misclassification of disease status may be non-differential, that is, random misclassification, resulting in an effect toward the null, which is a decreased threat to validity.

To the best of our knowledge, our models are the first risk scoring systems developed for diabetes and general populations to predict VTE risk. Despite the low overall risk of VTE, it is a common preventable cause of death, which is easy to be ignored. Our risk scoring systems can alert the attention of physicians and may be used to screen patients with the highest risk of VTE and assist clinicians in making decisions for prevention intervention and treatments. In addition, our novel and simple tools may be used to identify people with type 2 diabetes who may benefit from

thromboprophylaxis to reduce the risk of VTE and further prevent bleeding and immature death.

Acknowledgments

Author contributions

Conceptualization: Ching-Chu Chen.
Data curation: Cheng-Chieh Lin, Chia-Ing Li, Chiu-Shong Liu, Wen-Yuan Lin, Chih-Hsueh Lin, Shing-Yu Yang, Tsai-Chung Li.
Formal analysis: Chia-Ing Li, Chiu-Shong Liu, Wen-Yuan Lin, Chih-Hsueh Lin, Shing-Yu Yang.
Methodology: Cheng-Chieh Lin, Ching-Chu Chen, Chia-Ing Li, Tsai-Chung Li.
Project administration: Cheng-Chieh Lin, Ching-Chu Chen.
Validation: Tsai-Chung Li.
Writing – original draft: Tsai-Chung Li.
Writing – review & editing: Cheng-Chieh Lin, Ching-Chu Chen, Chiu-Shong Liu, Wen-Yuan Lin, Chih-Hsueh Lin, Shing-Yu Yang.

References

[1] Cushman M, Tsai AW, White RH, et al. Deep vein thrombosis and pulmonary embolism in two cohorts: the longitudinal investigation of thromboembolism etiology. Am J Med 2004;117:19–25.

[2] Naess IA, Christiansen SC, Romundstad P, Cannegieter SC, Rosendaal FR, Hammerstrom J. Incidence and mortality of venous thrombosis: a population-based study. J Thromb Haemost 2007;5:692–9.

[3] Raskob GE, Angchaisuksiri P, Blanco AN, et al. Thrombosis: a major contributor to global disease burden. Arterioscler Thromb Vasc Biol 2014;34:2363–71.

[4] Nakamura M, Yamada N, Oda E, et al. Predictors of venous thromboembolism recurrence and the bleeding events identified using a Japanese healthcare database. J Cardiol 2017;70:155–62.

[5] Lee CH, Cheng CL, Lin LJ, Tsai LM, Yang YH. Epidemiology and predictors of short-term mortality in symptomatic venous thromboembolism. Circ J 2011;75:1998–2004.

[6] Lee CH, Lin LJ, Cheng CL, Kao Yang YH, Chen JY, Tsai LM. Incidence and cumulative recurrence rates of venous thromboembolism in the Taiwanese population. J Thromb Haemost 2010;8:1515–23.

[7] Zilio M, Mazzai L, Sartori MT, et al. A venous thromboembolism risk assessment model for patients with Cushing's syndrome. Endocrine 2016;52:322–32.

[8] Bruzelius M, Bottai M, Sabater-Lleal M, et al. Predicting venous thrombosis in women using a combination of genetic markers and clinical risk factors. J Thromb Haemost 2015;13:219–27.

[9] McDaid A, Logette E, Buchillier V, et al. Risk prediction of developing venous thrombosis in combined oral contraceptive users. PLoS One 2017;12:e0182041.

[10] Hewes PD, Hachey KJ, Zhang XW, et al. Evaluation of the Caprini model for venothromboembolism in esophagectomy patients. Ann Thorac Surg 2015;100:2072–8.

[11] Nemeth B, van Adrichem RA, van Hylckama Vlieg A, et al. Venous thrombosis risk after cast immobilization of the lower extremity: derivation and validation of a clinical prediction score, L-TRiP(cast), in three population-based case-control studies. PLoS Med 2015;12:e1001899discussion e1001899.

[12] Dang JT, Switzer N, Delisle M, et al. Predicting venous thromboembolism following laparoscopic bariatric surgery: development of the BariClot tool using the MBSAQIP database. Surg Endosc 2019;33:821–31.

[13] Matsuo K, Yessaian AA, Lin YG, et al. Predictive model of venous thromboembolism in endometrial cancer. Gynecol Oncol 2013;128:544–51.

[14] Pabinger I, van Es N, Heinze G, 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.

[15] Greene MT, Spyropoulos AC, Chopra V, et al. Validation of risk assessment models of venous thromboembolism in hospitalized medical patients. The American journal of medicine 2016;129:1001.e1009–18.

[16] Liu X, Liu C, Chen X, Wu W, Lu G. Comparison between Caprini and Padua risk assessment models for hospitalized medical patients at risk for venous thromboembolism: a retrospective study. Interact Cardiovasc Thorac Surg 2016;23:538–43.

[17] Mahan CE, Liu Y, Turpie AG, et al. External validation of a risk assessment model for venous thromboembolism in the hospitalised acutely-ill medical patient (VTE-VALOURR). Thromb Haemost 2014;112:692–9.

[18] Park MS, Perkins SE, Spears GM, et al. Risk factors for venous thromboembolism after acute trauma: a population-based case-cohort study. Thromb Res 2016;144:40–5.

[19] Rosenberg D, Eichorn A, Alarcon M, McCullagh L, McGinn T, Spyropoulos AC. External validation of the risk assessment model of the International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) for medical patients in a tertiary health system. J Am Heart Assoc 2014;3:e001152.

[20] Soria JM, Morange PE, Vila J, et al. Multilocus genetic risk scores for venous thromboembolism risk assessment. J Am Heart Assoc 2014;3:e001060.

[21] Geerts WH, Bergqvist D, Pineo GF, et al. Prevention of venous thromboembolism: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). Chest 2008;133:381s–453s.

[22] Huang W, Anderson FA, Spencer FA, Gallus A, Goldberg RJ. Risk-assessment models for predicting venous thromboembolism among hospitalized non-surgical patients: a systematic review. J Thromb Thrombolysis 2013;35:67–80.

[23] Donnan PT, Leese GP, Morris AD. Hospitalizations for people with type 1 and type 2 diabetes compared with the nondiabetic population of Tayside, Scotland: a retrospective cohort study of resource use. Diabetes care 2000;23:1774–9.

[24] Carr ME. Diabetes mellitus: a hypercoagulable state. J Diabetes Complications 2001;15:44–54.

[25] Liao LN, Li CI, Liu CS, et al. Extreme levels of HbA1c increase incident ESRD risk in Chinese patients with type 2 diabetes: competing risk analysis in National Cohort of Taiwan Diabetes Study. PLoS One 2015;10:e0130828.

[26] Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Hu G. HbA1c and coronary heart disease risk among diabetic patients. Diabetes care 2014;37:428–35.

[27] Yang YF, Li TC, Li CI, et al. Visit-to-visit glucose variability predicts the development of end-stage renal disease in type 2 diabetes: 10-year follow-up of Taiwan diabetes study. Medicine 2015;94:e1804.

[28] Sullivan LM, Massaro JM, D'Agostino RBSr. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. Stat Med 2004;23:1631–60.

[29] Merry AH, Boer JM, Schouten LJ, et al. Risk prediction of incident coronary heart disease in The Netherlands: re-estimation and improvement of the SCORE risk function. Eur J Prev Cardiol 2012;19:840–8.

[30] Kengne AP, Masconi K, Mbanya VN, Lekoubou A, Echouffo-Tcheugui JB, Matsha TE. Risk predictive modelling for diabetes and cardiovascular disease. Crit Rev Clin Lab Sci 2014;51:1–12.

[31] Alonso A, Krijthe BP, Aspelund T, et al. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc 2013;2:e000102.

[32] Zheng L, Sun Z, Zhang X, et al. Predictive value for the rural Chinese population of the Framingham hypertension risk model: results from Liaoning Province. Am J Hypertens 2014;27:409–14.

[33] Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: a risk profile from the Framingham Study. Stroke 1991;22:312–8.

[34] Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data. Pharmacoepidemiol Drug Saf 2012;21(Suppl 1):154–62.

[35] Mi Y, Yan S, Lu Y, Liang Y, Li C. Venous thromboembolism has the same risk factors as atherosclerosis: a PRISMA-compliant systemic review and meta-analysis. Medicine 2016;95:e4495.

[36] Ageno W, Becattini C, Brighton T, Selby R, Kamphuisen PW. Cardiovascular risk factors and venous thromboembolism: a meta-analysis. Circulation 2008;117:93–102.

[37] Becattini C, Vedovati MC, Ageno W, Dentali F, Agnelli G. Incidence of arterial cardiovascular events after venous thromboembolism: a systematic review and a meta-analysis. J Thromb Haemost 2010;8:891–7.

[38] Bai J, Ding X, Du X, Zhao X, Wang Z, Ma Z. Diabetes is associated with increased risk of venous thromboembolism: a systematic review and meta-analysis. Thromb Res 2015;135:90–5.

[39] Tripodi A, Branchi A, Chantarangkul V, et al. Hypercoagulability in patients with type 2 diabetes mellitus detected by a thrombin generation assay. J Thromb Thrombolysis 2011;31:165–72.

[40] Heit JA. Epidemiology of venous thromboembolism. Nat Rev Cardiol 2015;12:464–74.

[41] Ungprasert P, Srivali N, Wijarnpreecha K, Charoenpong P, Knight EL. Non-steroidal anti-inflammatory drugs and risk of venous thromboembolism: a systematic review and meta-analysis. Rheumatology (Oxford, England) 2015;54:736–42.