External validation of bleeding risk models for the prediction of long-term bleeding risk in patients with established cardiovascular disease

Maria C. Castelijns, MD [a], Steven H.J. Hageman, PhD [a], Martin Teraa, PhD [b], Manon G. van der Meer, PhD [c], Jan Westerink, PhD [a,d], Francesco Costa, PhD [e], Jurriën M. ten Berg, PhD [f,g], and Frank L.J. Visseren, PhD [a], on behalf of the UCC-SMART study group ***Utrecht, The Netherlands; Zwolle, The Netherlands; Messina, Italy; Nieuwegein, The Netherlands; Maastricht, The Netherlands***

Abstract

Objective The long-term predictive performance of existing bleeding risk models in patients with various manifestations of cardiovascular disease (CVD) is not well known. This study aims to assess and compare the performance of relevant existing bleeding risk models in estimating the long-term risk of major bleeding in a cohort of patients with established CVD.

Methods Seven existing bleeding risk models (PRECISE-DAPT, DAPT, Ducrocq et al, de Vries et al, S_2TOP-BLEED, Intracranial B_2LEED$_3$S and HAS-BLED) were identified and externally validated in 7,249 patients with established CVD included in the Utrecht Cardiovascular Cohort–second manifestations of arterial disease study. Predictive performance was assessed in terms of discrimination and calibration, both at 10 years and the original prediction horizon of the models. Major bleeding was defined as Bleeding Academic Research Consortium type 3 or 5.

Results After a median follow-up of 8.4 years (interquartile range 4.5-12.5), a total of 233 (3.2%) major bleeding events occurred. C-statistics for discrimination at 10 years ranged from 0.53 (95%CI 0.49-0.57) to 0.64 (95%CI 0.60-0.68). Calibration plots after recalibration to 10 years showed best agreement between predicted and observed bleeding risk for De Vries et al, S_2TOP-BLEED, DAPT and PRECISE-DAPT.

Conclusions The performance of existing bleeding risk models to predict long-term bleeding in patients with CVD varied. Discrimination and calibration were best for the models of de Vries et al, S_2TOP-BLEED, DAPT and PRECISE-DAPT. Of these, recalibrated models requiring the least predictors may be preferred for use to personalize prevention with antithrombotic therapy. (Am Heart J 2023;260:72–81.)

In the prevention of cardiovascular disease (CVD), risk estimation of cardiovascular events plays an important role in clinical decision making.[1,2] This concerns not only the risk of ischemic events, but also the risk of bleeding due to antithrombotic treatment, which is even more important in a population which is increasingly older and more frail. The question arises how bleeding risk can be estimated accurately in patients with CVD.

Several bleeding risk scores are available for different patient populations. The HAS-BLED score is available to assess bleeding risk in atrial fibrillation (AF) patients.[3,4] DAPT and PRECISE-DAPT scores are advised to guide decision making on duration of dual antiplatelet therapy in patients undergoing percutaneous coronary intervention with stent implantation.[5-8] However, the use of bleeding risk scores in patients with chronic coronary syndromes without recent intervention and patients with cerebrovascular and peripheral artery disease is less well defined, while the scores mentioned have not been validated in a population with different manifestations of CVD.[9,10] Moreover, since atherosclerosis is a progressive disease often involving multiple vascular beds, it is desir-

From the [a]Department of Vascular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands, [b]Department of Vascular Surgery, University Medical Center Utrecht, Utrecht, The Netherlands, [c]Department of Cardiology, University Medical Center Utrecht, Utrecht, The Netherlands, [d]Department of Internal Medicine, Isala Clinics Zwolle, Zwolle, The Netherlands, [e]Department of Cardiology, G. Martino University Hospital Messina, Messina, Italy, [f]Department of Cardiology and Platelet Function Research, St. Antonius Hospital Nieuwegein, Nieuwegein, The Netherlands, [g]Cardiovascular Research Institute Maastricht, Maastricht, The Netherlands

Submitted October 18, 2022; accepted February 18, 2023

Reprint requests: *Frank L.J. Visseren, Department of Vascular Medicine, University Medical Center Utrecht, PO Box 85500, 3508 GA, Utrecht, The Netherlands*

E-mail address: F.L.J.Visseren@umcutrecht.nl.

0002-8703

https://doi.org/10.1016/j.ahj.2023.02.011

able to be able to predict bleeding in a population with various manifestations of CVD.[11]

Ultimately, arriving at optimal treatment of individual patients calls for weighing benefit and harm of preventive strategies, thus contrasting CVD risk reduction with increased bleeding risk. Individual treatment benefit can already be estimated by predictions of 10-year recurrent CVD risk, for example using the SMART risk score,[12] or lifetime risk with SMART-REACH,[13] in combination with hazard ratios from trials and meta-analyses. Which scores should be used to estimate bleeding risk is not well known. Recently, dual antiplatelet therapy and dual pathway inhibition have become available as guideline supported treatment options, making it even more important to be able to predict bleeding risks from intensified antithrombotic therapy.[2] Therefore, the aim of this study is to evaluate the performance of relevant existing bleeding risk scores in predicting long-term major bleeding in a cohort of patients with various manifestations of CVD.

Methods

Identification of bleeding risk models

Existing bleeding risk models which could potentially be relevant to individuals with established CVD in the long term were identified. Based on the European Society of Cardiology (ESC) guidelines on AF, chronic coronary syndromes, dual antiplatelet therapy in coronary artery disease and peripheral artery disease,[4,9,10,14] the HAS-BLED score, DAPT bleeding prediction model and PRECISE-DAPT score were selected.[3,5,6] Other models are also commented on in these guidelines, but only the models recommended for use, eg, based on meta-analyses, were included.[14,4] The 2021 ESC consensus document on antithrombotic therapies in aortic and peripheral arterial diseases identified bleeding prediction models of de Vries et al and Ducrocq et al for patients with established CVD.[15-17] The score of Spiliopoulos et al was not included because it was designed to predict short-term bleeding complications of endovascular interventions, and the studies of Cea Soriano et al and Ward et al identified risk factors without reporting a bleeding risk model or score.[18-20] To identify possible additional bleeding models published in the 2 years before or after the guidelines were issued, a literature search (up to March 2016) was performed (Supplementary Figure 1). This yielded the S_2TOP-BLEED and Intracranial-B_2LEED$_3$S score.[21,22] Table I provides an overview of the predictors used in each model. A short description of the studies in which the bleeding risk models were derived is presented in Supplementary Table 1.

Study population

External validation of the bleeding risk models was performed in patients originating from the Utrecht Cardiovascular Cohort–Second manifestations of arterial disease (UCC-SMART) study, of which the design and methods have been described previously.[23] Briefly, UCC-SMART is an ongoing cohort study in both inclusion and follow-up of patients aged 18 to 90 years referred to a teaching hospital in the Netherlands for management of atherosclerotic CVD and marked cardiovascular risk factors. Patients with a short life expectancy, pregnant women and those insufficiently fluent in Dutch are excluded. The study was approved by the ethics committee and institutional board of the University Medical Center Utrecht (reference number 22-088) and all participants gave their written informed consent. For the present study, 7249 patients with a history of CVD enrolled between July 2001 and January 2020 were included, because information on specific antithrombotic agents use is only available from July 2001 onward. Established CVD was defined as cerebrovascular (transient ischemic attack, cerebral infarction, ischemic retinal syndrome), coronary artery (angina, myocardial infarction), peripheral artery disease or abdominal aortic aneurysm (Supplementary Table 2). In case of a recent cardiovascular event or intervention, baseline measurements are recorded more than 30 days after discharge and as such, the UCC-SMART study can be regarded as a cohort of patients with stable CVD. Information on the predictors was recorded by a questionnaire, physical examination and laboratory measurements upon enrollment in UCC-SMART. No extramural funding was used to support this work.

Outcome

Definitions of bleeding end points used in the bleeding risk model derivation studies differ (Supplementary Table 1). In the present study, major bleeding was defined as type 3 or 5 bleeding according to the Bleeding Academic Research Consortium definition, developed for standardized bleeding end point definitions in cardiovascular research for patients receiving antithrombotic therapy. It comprises bleeding leading to a drop in hemoglobin of >3 g/dL (1.9 mmol/L), cardiac tamponade, bleeding requiring transfusion, surgical intervention or intravenous vasoactive agents, intracranial, intraocular and fatal bleeding.[24] In the UCC-SMART study, information on outcomes is collected through annual questionnaires on hospitalizations and outpatient clinic visits. When patients reports a possible event, hospital discharge letters and results of laboratory and radiology examinations are collected and audited by 3 physicians from the end point adjudication committee. The first occurring bleeding event was used for analyses in case of multiple bleeding events. The competing event was defined as death due to other causes than bleeding.

Data analyses

Use of predictors

Predictors of the models were primarily handled according to their definitions in the original studies. If spe-

Table I. Predictors included in bleeding risk models

	DAPT	PRECISE-DAPT	Ducrocq et al	De Vries et al	S_2TOP-BLEED	Intracranial B_2LEED$_3$S	HAS-BLED	Available in UCC-SMART
Age	x	x	x	x	x	x	x	Yes
Sex				x	x	x		Yes
Ethnicity				o	o	o		No
Geographical region				x				Yes
Alcohol							x	Yes
Smoking			x	x	x			Yes
Clinical performance					o			No
BMI					x	x		Yes
Hypertension/blood pressure	x		x	x	x	x	x	Yes
Hypercholesterolemia/cholesterol level			x	x				Yes
Kidney function/creatinine level	x	x		x			x	yes
Abnormal liver function							o	No
Hemoglobin level		x						Yes
Diabetes mellitus			x	x	x			Yes
History of bleeding		x		o			o	No
History of stroke					x	x	x	Yes
Lacunar stroke						o		No
History of myocardial infarction						x		Yes
Peripheral artery disease	x		x					Yes
Number of cardiovascular beds affected				x				yes
Heart failure			o	o				No
Antiplatelet therapy	x		x		x	x	x	Yes
Anticoagulant therapy			x			x		Yes
Labile INR							o	No

BMI, body mass index; *INR*, international normalized ratio.
An "x" indicates the predictor was part of the risk score and also available in the UCC-SMART study, an "o" indicates the predictor was part of the risk score but not available in the UCC-SMART study.

cific variables were not available in UCC-SMART, they were replaced by a proxy (Supplementary Table 3). Missing predictor values for alcohol use (1%), smoking (0.5%), hypertension history (2.9%), hypercholesterolemia history (1.5%), stroke history (0.2%), myocardial infarction history (0.2%), BMI (0.2%), blood pressure (0.1%), kidney function (0.4%), hemoglobin (0.4%), and cholesterol (0.4%) were imputed using single imputation by predictive mean matching if the variables were indeed collected in UCC-SMART. Information on history of bleeding (other than intracerebral hemorrhage), heart failure, liver failure, labile INR, leukocytes, continued (ie, >12 months) use of dual antiplatelet therapy, ethnicity, lacunar type stroke and modified Rankin scale was not available in the UCC-SMART cohort. The 4-item PRECISE-DAPT model was used because leukocytes were not measured in UCC-SMART. The model of Ducrocq et al and HAS-BLED only report point scores based on the variables and corresponding bleeding risks, where PRECISE-DAPT, DAPT, de Vries et al S_2TOP-BLEED and Intracranial-B_2LEED$_3$S provide predicted risks estimated by regression model equations. For bleeding models only reporting scores, all patients were assigned a value of zero for liver failure, bleeding history, labile INR and heart failure, assuming that the majority of patients did not have these risk factors. For the bleeding models reporting an equation for risk prediction, the prevalence of previous bleeding, heart failure and lacunar type stroke from the derivation cohorts were used in the linear predictors. Prevalence of history of bleeding of the COMPASS trial (the derivation cohort from de Vries et al) was used, because SMART population resembles this study population more than the HAS-BLED derivation study including previous bleeding as a predictor. For heart failure, the study of Ducrocq et al was used, because it is based on a cohort including patients similar to UCC-SMART rather than trials. In addition, sensitivity analyses were performed assuming patients using a combination of ACE-inhibitors, betablockers, and mineralocorticoid receptor antagonists had heart failure.[25] White ethnicity was chosen by default because of the geographical region of the UCC-SMART cohort. Since continued use of dual antiplatelet therapy beyond 12 months after percutaneous coronary intervention or coronary bypass is not standard care according to ESC guidelines used in the Netherlands, none of the patients were assumed to be treated with dual antiplatelet therapy beyond 12 months.[8,9,26] Nevertheless, observational studies showed 43% to 57% of acute coronary syndrome patients continued dual antiplatelet therapy after 2 years, and one of these studies found continuation in 29% of patients at 4 years.[27,28] Therefore, sensitivity analyses

were performed presuming 29% of coronary artery disease patients continued dual antiplatelet therapy >12 months by adding this prevalence to the linear predictor of the DAPT model to calculate predicted bleeding risks, given the longer prediction horizon upon validation. Lastly, none of the patients were assumed to have a clinical performance of ≥3 on the Rankin scale in the S_2TOP-BLEED model, since a modified Rankin scale ≤3 is an inclusion criterion in UCC-SMART.

Performance and predicted risks

Performance of the bleeding models was assessed in terms of discrimination and calibration. Discrimination is depicted by C-statistics for time-to-event data adjusted for competing risk, using the time-dependent Receiver Operating Characteristic curve estimation function in R. The C-statistic equals the area under the curve in which the sensitivity is plotted against 1—specificity for consecutive cut-offs for the probability of a major bleeding, adjusted for competing risk. Calibration is displayed by plotting adjusted predicted risks against observed risks. Although bleeding risk scores are often developed for short-term prediction horizons, in clinical practice preventive antithrombotic therapy is generally prescribed long-term and in many cases lifelong. Therefore, the bleeding models were primarily validated and compared at the long-term prediction horizon of 10 years, in line with how ischemic event risks are often considered. Additional analyses were performed using the original prediction horizons of the bleeding risk models. The model of de Vries et al used a Fine and Gray model for lifetime predictions and was externally validated at 10 year follow-up, so an original time span of 10 years was chosen for this model. Since most of the risk scores were derived in different populations than a population with various manifestations of established CVD, the risk scores were using a constant multiplicative term (recalibration of the intercept).[29] Expected-observed ratios were used to recalibrate all individual predicted risks in the validation population. The original intercepts and predictors coefficients of the models were applied to calculate predicted risks of major bleeding according to the PRECISE-DAPT, DAPT, S_2TOP-BLEED, and Intracranial-B_2LEED$_3$S scores and the model from de Vries et al This information was not provided for the model of Ducrocq et al and HAS-BLED, hence the scores were calculated and the bleeding risks corresponding to the scores were used. Predicted risks of the HAS-BLED score could only be studied up until a score of 6, because patients from the derivation cohort did not score higher than 6 and in our study, patients could not score higher than 6 because of 3 missing predictors. UCC-SMART patients with a score of 5 or 6 were grouped together, because in the original study only 2 patients had a score of 6 and no bleeding. For the model of Ducrocq et al, patients were divided into 4 groups based on their score (0-6, 7-8, 9-10, or 11-21), because the derivation study only reported predicted risks for score quartiles. We contacted the corresponding authors of these risk scores to request the original regression equations, but unfortunately the authors were unable to provide these. Statistical analyses were performed using R 4.0.3 for Windows.

The authors are solely responsible for the design and conduct of this study, all study analyses, the drafting and editing of the paper and its final contents.

Results

Baseline characteristics of the 7,249 patients with established CVD included in this study are shown in Table II. After a median follow up of 8.4 years (interquartile range 4.5-12.5), 233 first major bleeding events occurred (incidence rate 3.8 major bleeding events per 1,000 person years). Of the 233 major bleeding events, 43 occurred in the first year, 63 in the first 2 years, 73 in the first 3 years and 196 in the first 10 years of follow-up (Figure I).

Performance of bleeding risk models

Discrimination

Discrimination of the bleeding risk models ranged from 0.53 (95% CI 0.49-0.57) for Intracranial-B_2LEED$_3$S to 0.64 (95%CI 0.60-0.68) for the model of de Vries et al and DAPT in patients with CVD at a 10-year prediction horizon (Table III). The DAPT, PRECISE-DAPT, de Vries et al and S_2TOP-BLEED model had comparable C-statistics, where Ducrocq et al, Intracranial-B_2LEED$_3$S and HAS-BLED showed poorer discrimination. The discriminative ability at the original prediction horizons of the bleeding risk models was approximately similar (Supplementary Table 4). Time-dependent discrimination was fairly stable across the first 12 years of follow-up for most of the risk models, except for the Intracranial-B_2LEED$_3$S model which showed slightly lower C-statistics with longer follow-up duration more clearly (Figure II). Presuming that patients using a combination of ACE-inhibitors, betablockers and mineralocorticoid receptor antagonists had heart failure, 2 percent of the study population had heart failure. This sensitivity analysis showed no change in C-statistics for the models of Ducrocq et al and de Vries et al (Supplementary Table 5).

Calibration

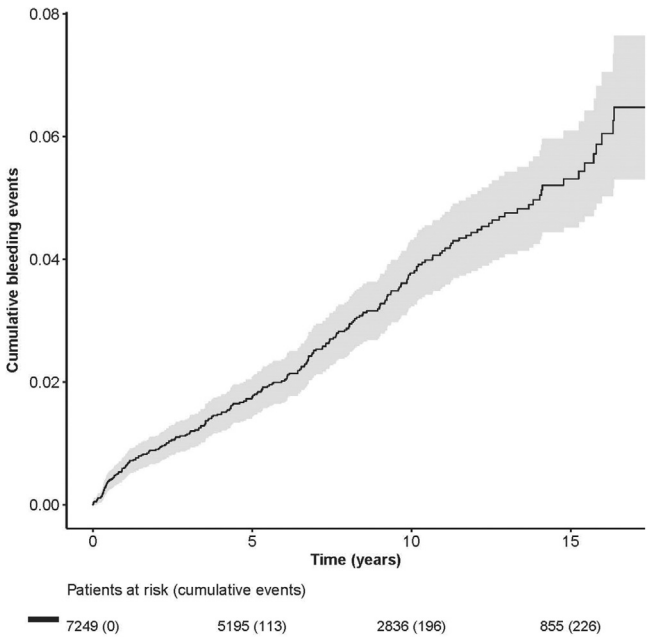
Figure III shows the calibration plots after recalibration at a 10-year prediction horizon. The recalibration factors for every model are shown in Supplementary Table 6. For the HAS-BLED score, a substantial overestimation of bleeding risk is seen for the highest score group. In the present study population, only 23 patients had a HAS-BLED score of 5 or 6 of whom no patients experienced a major bleeding during follow-up, leading to a downward deflect in the last group in the calibration plot (Figure III). The predicted risks of PRECISE-DAPT, DAPT, de Vries et al and S_2TOP-BLEED correspond adequately

Table II. Baseline characteristics of the UCC-SMART validation cohort

	Total population *n* = 7,249	No major bleeding *n* = 7,016	Major bleeding *n* = 233
Age (years)	60.5 ± 10.3	60.4 ± 10.3	64.4 ± 9.3
Male sex	5,310 (73)	5,134 (73)	176 (76)
Current alcohol use	5,084 (70)	4,929 (70)	155 (67)
Current smoking	2,023 (28)	1,943 (28)	80 (34)
Diabetes mellitus	1,252 (17)	1,212 (17)	40 (17)
History of CeVD	2,104 (29)	2,039 (29)	65 (28)
Stroke	1,405 (19)	1,367 (19)	38 (16)
History of CAD	4,647 (64)	4,516 (64)	131 (56)
Myocardial infarction	2,465 (34)	2,398 (34)	67 (29)
History of PAD	1,056 (15)	1,003 (14)	53 (23)
Polyvascular disease	913 (13)	875 (12)	38 (16)
Antiplatelet therapy use	5,845 (81)	5,669 (81)	176 (76)
Aspirin	5,293 (73)	5,133 (73)	160 (69)
Clopidogrel	2,170 (30)	2,105 (30)	65 (28)
Dual antiplatelet	1,774 (24)	1,721 (25)	53 (23)
Oral anticoagulant use	774 (11)	744 (11)	30 (13)
BMI (kg/m^2)	27 ± 4	27 ± 4	27 ± 4
Systolic blood pressure (mmHg)	138 ± 20	138 ± 20	146 ± 23
Diastolic blood pressure (mmHg)	81 ± 11	81 ± 11	83 ± 13
Total cholesterol (mmol/L)	4.6 ± 1.1	4.6 ± 1.1	4.8 ± 1.1
LDL-cholesterol (mmol/L)	2.6 ± 1.0	2.6 ± 1.0	2.8 ± 1.0
HDL-cholesterol (mmol/L)	1.3 ± 0.4	1.3 ± 0.4	1.3 ± 0.4
eGFR (ml/min/1.73m^2)	78 ± 18	78 ± 18	72 ± 19

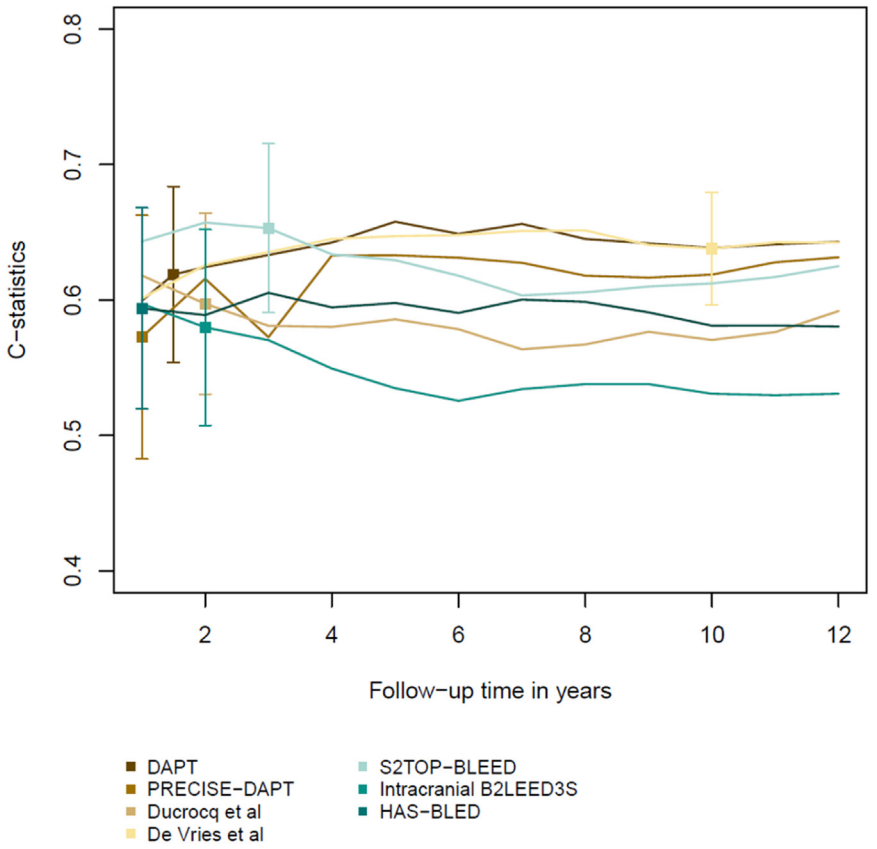
BMI, body mass index; *CAD*, coronary artery disease; *CeVD*, cerebrovascular disease; *eGFR*, estimated glomerular filtration rate; *HDL*, high-density lipoprotein; *LDL*, low-density lipoprotein; *PAD*, peripheral artery disease.
Data reported as mean ± standard deviation or *n* (%).
Polyvascular disease meaning 2 or more locations of vascular disease: cerebrovascular disease, coronary artery disease and abdominal aortic aneurysm and/or lower extremity disease. Estimated glomerular filtration rate calculated using the Chronic Kidney Disease Epidemiology Collaboration equation.

Figure I



Cumulative incidence of first major bleeding events in the UCC-SMART validation cohort. *UCC-SMART*, Utrecht Cardiovascular cohort–second manifestations of arterial disease.

Figure II



Discrimination of the bleeding risk models over time. The points indicate the C-statistics with their 95% confidence intervals of the bleeding risk models at their original prediction horizon, as can be found in Supplementary Table 4.

Table III. Discrimination of the bleeding risk models in patients with CVD at a 10-year prediction horizon

Bleeding risk model	C-statistic (95%CI)	*P*-value*
DAPT	0.64 (0.60-0.68)	NA (reference)
PRECISE-DAPT	0.62 (0.58-0.66)	.17
Ducrocq et al	0.57 (0.53-0.61)	<.01
De Vries et al	0.64 (0.60-0.68)	.77
S_2TOP-BLEED	0.61 (0.57-0.65)	.21
Intracranial B_2LEED_3S	0.53 (0.49-0.57)	<.01
HAS-BLED	0.58 (0.54-0.62)	.01

95%CI, 95% confidence interval; *CVD*, cardiovascular disease.
*Comparison of 2 time-dependent C-statistics, using the DAPT model as a reference.

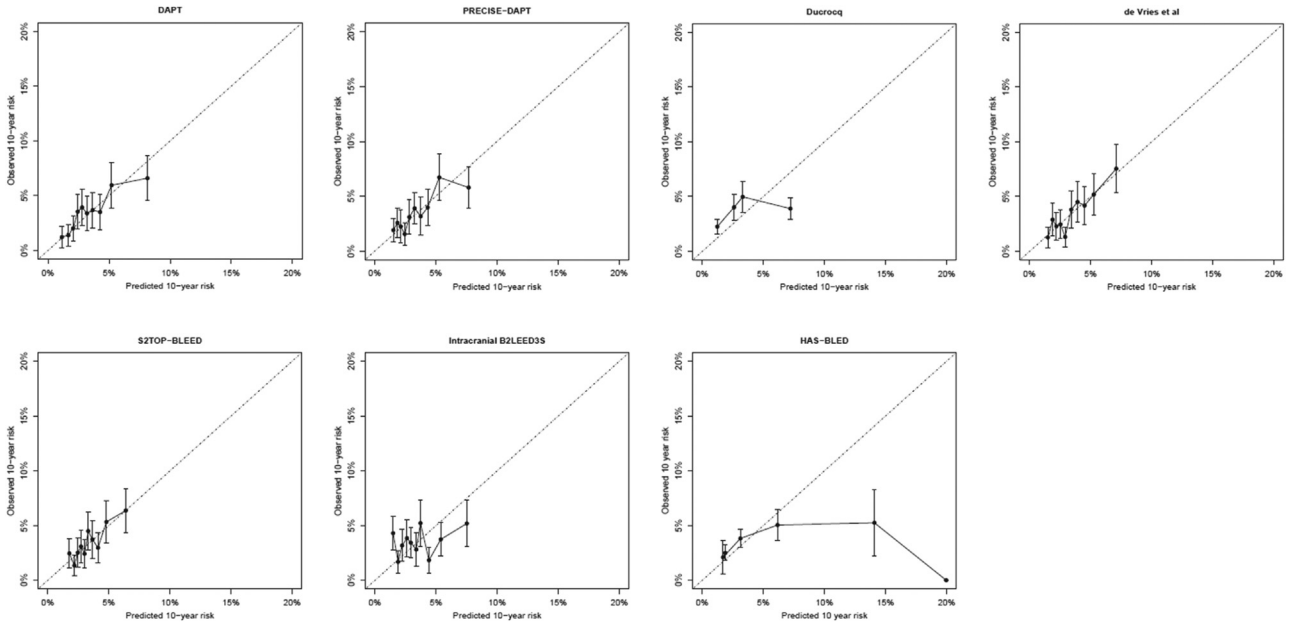
with the observed incidences of major bleeding. Sensitivity analyses presuming a proportion of coronary artery disease patients continued dual antiplatelet therapy beyond 12 months showed similar discrimination and calibration for the DAPT model at 10 years (Supplementary Figure 2). The analyses using a combination of medication for the presence of heart failure showed similar calibration plots (Supplementary Figure 3). Recalibrated plots using the original prediction horizons of the bleeding risk models are shown in Supplementary Figure 4. Calibration of the bleeding risk models before recalibration is shown in Supplementary Figure 5, displaying the over- and underestimation of bleedings risks due to different incidences between populations.

Discussion

The long-term performance of relevant existing bleeding risk models in patients with established CVD was evaluated in terms of discrimination and calibration. The bleeding risk model of de Vries et al, DAPT, S_2TOP-BLEED and PRECISE-DAPT showed higher discriminative ability and good calibration.

Whereas the models were developed in specific populations (ie, AF, postpercutaneous coronary intervention, cerebrovascular disease patients), the performance of some of the validated models appears to be adequate

Figure III



Calibration plots after recalibration of the bleeding risk models in UCC-SMART at a 10-year prediction horizon. The plots display the observed major bleeding risks plotted against the average predicted major bleeding risks within groups defined by deciles of predicted risk or by scores for HAS-BLED (0 up to 5 or 6) and Ducrocq et al (score of 0-6, 7-8, 9-10 and 11-21). *UCC-SMART*, Utrecht Cardiovascular cohort–second manifestations of arterial disease.

in accurately predicting long-term major bleeding risk in a population with various manifestations of CVD. This may be due to similarities between derivation populations and the UCC-SMART population both comprising patients at high CVD risk, and the associated selected predictors. The superiority of the models of de Vries et al, DAPT, S_2TOP-BLEED, and PRECISE-DAPT could be explained by the fact that these models provided equations to calculate predicted risks based on patient characteristics and that they were developed to predict severe bleeding events, similar to BARC type 3 and 5 bleedings used in this study. In contrast, the HAS-BLED score and model of Ducrocq et al only provided scores and risks per score, and the Intracranial-B_2LEED$_3$S model was designed to predict intracranial bleeding only. In clinical practice, however, the model of Ducrocq et al and HAS-BLED also use scores with corresponding risks.

Previous external validation of the HAS-BLED score in an AF cohort treated with various antithrombotic agents showed moderate C-statistics (0.57-0.66) which could be explained by incomplete data on predictors.[30] However, discrimination in patients from AF registries with only information on labile INR missing was moderate too.[31] External validation of the other bleeding risk models in their target population showed similar discrimination, with C-statistics ranging from 0.59 for Intracranial-B_2LEED$_3$S to 0.64 for the PRECISE-DAPT score.[5,16,17,21,22,32] Remarkably, validation of PRECISE-DAPT in PCI-registry patients showed similar C-statistics despite 1 of the 4 predictors missing in UCC-SMART.[6] This could indicate that applying bleeding models to a population with various manifestations of CVD with some predictors missing does not fully explain the limited performance of some of the models. The authors state that availability of candidate predictors from trial populations used for derivation could be a possible explanation for the moderate performance in external validation.[5,21,22] The lower model performance of Ducrocq et al in the present study compared to previous external validation could be explained by using the 4 groups based on score quartiles reported in the original publication.[17] De Vries et al previously used the UCC-SMART population for external validation and reported higher C-statistics (0.69 [95%CI 0.67-0.70]), which could be explained by using a composite outcome of bleeding and nonbleeding mortality.[16]

Accurate risk prediction can support individualized antithrombotic therapy. Calibration, indicating the accuracy of risk estimates (ie, agreement with observed risks), is a more relevant measure to assess model performance for use in clinical practice, when treatment recommendations are based on the absolute values of

predicted risks. Appraisal of absolute numbers of C-statistics can be difficult, since they are based on ranks and are dependent of distribution of risk in the population.[33] The C-statistics in external validation of guideline-recommended models for vascular events in CVD patients for example, are not considerably higher.[34,35] The models of de Vries et al, DAPT, S_2TOP-BLEED, and PRECISE-DAPT showed the best performance to predict major bleeding in a population with various CVD manifestations, hence these may be suitable for use in individuals with established CVD. The recalibrated PRECISE-DAPT or DAPT bleeding model could be a pragmatic choice because of the fewest predictors needed. However, patients from the derivation study populations come from clinical trials and may therefore not be representative of CVD patients in the consulting room. Patients with characteristics leading to higher bleeding risks are often excluded from trials, whereas accurate risk estimates are desired in these patients. Only predictors available in the trials could be selected for model derivation, and a lower variety of predictor values in derivation populations may reduce model performance.[29] Besides, DAPT, S_2TOP-BLEED, and PRECISE-DAPT have not been adjusted for competing risk of nonbleeding mortality. Ideally, a large, less selected cohort population containing numerous possible predictors should be used in future research to optimize competing risk adjusted bleeding risk prediction in patients with CVD. Until such models are available, the recalibration factors as presented in the current study may be used to tailor the predictions of existing models to clinical practice for individuals with established CVD.

Strengths of this study include the evaluation and comparison of potential relevant bleeding risk models, providing a valuable overview in a large population with various manifestations of CVD. Moreover, recalibration to a 10-year prediction horizon allows for contrasting predicted bleeding risks against ischemic event risks. This long-term performance is of importance because in clinical practice treatment decisions are usually based on 10-year or even lifetime predicted risks. Lastly, competing-risk adjustment was performed to reduce overestimation of bleeding risk in a population at high risk of nonbleeding mortality with a long follow-up. Some study limitations should also be considered. First, the most important limitation is that not all predictors of the bleeding risk models were available in the UCC-SMART study. This is mainly a limitation of the cohort study, since the predictors history of bleeding, heart failure and lacunar type stroke are relevant predictors that are usually well recorded in clinical practice. In contrast, information on liver function and ethnicity are not typically collected in cardiovascular care,[36] and the Rankin clinical performance score is only collected in patients with cerebrovascular disease rather than in all vascular patients. Lastly, labile INR defined as therapeutic time in range <60% might not be very practical to assess for clinicians and liver failure has a low prevalence, so missing these predictors is close to clinical practice. Assigning the same value to all patients could have resulted in loss of discriminative ability as is seen by lower C-statistics even for models showing reasonable calibration. Similarly, incorporating a 29% prevalence of continued dual antiplatelet therapy in the linear predictor of the DAPT model does not improve discrimination but only affects calibration. Additional analyses using a proxy for heart failure based on medication use showed similar discrimination and calibration, although only 2 percent of patients had heart failure using this proxy. The prevalence of heart failure is possibly underestimated, because not all patients with heart failure may yet have been clinically diagnosed and thus treated, and the medication used in the proxy is a cornerstone treatment for heart failure with reduced ejection fraction in particular.[25] The prevalence in the derivation study of Ducrocq et al was higher (14%) in CVD patients ≥45 years. Next, comparison of the models is challenging given the different study populations in which models were derived, different bleeding definitions and different prediction horizons. However, by comparing all models in the same population, with the same relevant end point and at the same prediction horizon of 10 years, we have been able to compare relevant models as good as possible with current data. Lastly, assessing the performance of bleeding models using predictor values upon inclusion in the UCC-SMART study could be considered a limitation, since possible changes of predictors throughout life are not accounted for. However, in clinical practice risk estimates are also made based on information known at that time and would be repeated regularly, particularly in case of clinical changes.

In conclusion, the long-term predictive performance of all relevant existing bleeding risk models has been shown in a large population of individuals with established CVD. The performance in terms of discrimination and calibration varied and was best for the model of Vries et al, DAPT, S_2TOP-BLEED, and PRECISE-DAPT. The DAPT bleeding model or PRECISE-DAPT may be preferred for use, given the readily available and least predictors needed. With the use of long-term data in less selected, observational studies, prediction of bleeding may be improved in future studies, thereby further contributing to personalized prevention in individuals with established CVD.

Funding

The UCC-SMART study was financially supported by a grant of the University Medical Center Utrecht, the Netherlands. The supporting sources had no involvement in study design, analysis, interpretation, writing of the results, or the decision to submit for publication.

Conflicts of Interest

Dr Costa reports receiving payment or honoraria for lectures, presentations, speaker bureaus, manuscript writing from Chiesi Farmaceutici and Astra Zeneca, outside the submitted work. Ms Castelijns, Dr Hageman, Dr Teraa, Dr van der Meer, Dr Westerink, Prof ten Berg and Prof Visseren have no competing interests to declare.

Acknowledgments

We gratefully acknowledge the contribution of the research nurses; R. van Petersen (data-manager); A. Vandersteen (study manager) and the members of the Utrecht Cardiovascular Cohort-Second Manifestations of ARTerial disease-Studygroup (UCC-SMART-Studygroup): M.J. Cramer, M.G. van der Meer and H.M. Nathoe, Department of Cardiology; G.J. de Borst, Department of Vascular Surgery; M.L. Bots and M.I. Geerlings, Julius Center for Health Sciences and Primary Care; M.H. Emmelot-Vonk, Department of Geriatrics; P.A. de Jong, Department of Radiology; A.T. Lely, Department of Gynecology and Obstetrics; N.P. van der Kaaij, Department of Cardiothoracic Surgery; L.J. Kappelle and Y.M. Ruigrok, Department of Neurology & Hypertension; M.C. Verhaar, Department of Nephrology & Hypertension, J.A.N. Dorresteijn and F.L.J. Visseren (chair), Department of Vascular Medicine, University Medical Center Utrecht and Utrecht University.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ahj.2023.02.011.

CRediT authorship contribution statement

Maria C. Castelijns: Conceptualization, Methodology, Formal analysis, Writing – original draft. **Steven H.J. Hageman:** Conceptualization, Methodology, Formal analysis, Validation, Writing – review & editing. **Martin Teraa:** Writing – review & editing. **Manon G. van der Meer:** Writing – review & editing. **Jan Westerink:** Conceptualization, Methodology, Writing – review & editing. **Francesco Costa:** Software, Writing – review & editing. **Jurriën M. ten Berg:** Writing – review & editing. **Frank L.J. Visseren:** Conceptualization, Methodology, Writing – review & editing, Supervision.

References

1. Rossello X, Dorresteijn JA, Janssen A, et al. Risk prediction tools in cardiovascular disease prevention: a report from the ESC Prevention of CVD Programme led by the European Association of Preventive Cardiology (EAPC) in collaboration with the Acute Cardiovascular Care Association (ACCA) and the As. Eur J Prev Cardiol 2019;26:1534–44.
2. Visseren FLJ, MacH F, Smulders YM, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur Heart J 2021;42:3227–337.
3. Pisters R, Lane DA, Nieuwlaat R, et al. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation. Chest 2010;138:1093–100. [Internet] https://linkinghub.elsevier.com/retrieve/pii/S0012369210605855.
4. Hindricks G, Potpara T, Dagres N, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): the Task Force for the diagnosis and management of atrial fibrillation of the Europe. Eur Heart J 2021;42:373–498.
5. Yeh RW, Secemsky EA, Kereiakes DJ, et al. Development and validation of a prediction rule for benefit and harm of dual antiplatelet therapy beyond 1 year after percutaneous coronary intervention. JAMA 2016;315:1735.
6. Costa F, van Klaveren D, James S, et al. Derivation and validation of the predicting bleeding complications in patients undergoing stent implantation and subsequent dual antiplatelet therapy (PRECISE-DAPT) score: a pooled analysis of individual-patient datasets from clinical trials. Lancet 2017;389:1025–34.
7. Collet JP, Thiele H, Barbato E, et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. Eur Heart J 2021;42:1289–367.
8. Valgimigli M, Bueno H, Byrne RA, et al. 2017 ESC focused update on dual antiplatelet therapy in coronary artery disease developed in collaboration with EACTS. Eur Heart J 2018;39:213–60.
9. Knuuti J, Wijns W, Saraste A, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. Eur Heart J 2020;41:407–77.
10. Aboyans V, Ricco JB, Bartelink MLEL, et al. 2017 ESC guidelines on the diagnosis and treatment of peripheral arterial diseases, in collaboration with the European Society for Vascular Surgery (ESVS). Eur Heart J 2018;39:763–816.
11. Weissler EH, Jones WS, Desormais I, et al. Polyvascular disease: a narrative review of current evidence and a consideration of the role of antithrombotic therapy. Atherosclerosis 2020;315:10–17.
12. Dorresteijn JAN, Visseren FLJ, Wassink AMJ, et al. Development and validation of a prediction rule for recurrent vascular events based on a cohort study of patients with arterial disease: the SMART risk score. Heart 2013;99:866–72. [Internet] https://heart.bmj.com/lookup/doi/10.1136/heartjnl-2013-303640.
13. Kaasenbrood L, Bhatt DL, Dorresteijn JAN, et al. Estimated life expectancy without recurrent cardiovascular events in patients with vascular disease: the SMART-REACH model. J Am Heart Assoc 2018;7. [Internet] https://www.ahajournals.org/doi/10.1161/JAHA.118.009217.
14. Valgimigli M, Bueno H, Byrne RA, et al. 2017 ESC focused update on dual antiplatelet therapy in coronary artery disease developed in collaboration with EACTS: the Task Force for dual antiplatelet therapy in coronary artery disease of the European Society of Cardiology (ESC) and of the European. Eur Heart J 2018;39:213–60.
15. Aboyans V, Bauersachs R, Mazzolai L, et al. Antithrombotic therapies in aortic and peripheral arterial diseases in 2021: a consensus document from the ESC working group on aorta and peripheral vascular diseases, the ESC working group on

thrombosis, and the ESC working group on cardiovascular pharm. Eur Heart J 2021;42:4013–24.
16. de Vries TI, Eikelboom JW, Bosch J, et al. Estimating individual lifetime benefit and bleeding risk of adding rivaroxaban to aspirin for patients with stable cardiovascular disease: results from the COMPASS trial. Eur Heart J 2019;40:3771–3778a.
17. Ducrocq G, Wallace JS, Baron G, et al. Risk score to predict serious bleeding in stable outpatients with or at risk of atherothrombosis. Eur Heart J 2010;31:1257–65. [Internet] https://academic.oup.com/eurheartj/article-lookup/doi/10.1093/eurheartj/ehq021.
18. Spiliopoulos S, Tsochatzis A, Festas G, et al. A new preprocedural score to predict bleeding complications of endovascular interventions for peripheral artery disease. J Endovasc Ther 2019;26:816–25.
19. Cea Soriano L, Fowkes FG, Allum A, et al. Predictors of bleeding in patients with symptomatic peripheral artery disease: a cohort study using the health improvement network in the United Kingdom. Thromb Haemost 2018;118:1101–12.
20. Ward R, Huang Z, Rockhold FW, et al. Major bleeding in patients with peripheral artery disease: insights from the EUCLID trial. Am Heart J 2020;220:51–8.
21. Hilkens NA, Algra A, Diener HC, et al. Predicting major bleeding in patients with noncardioembolic stroke on antiplatelets. Neurology. 2017;89:936–43.
22. Amarenco P, Sissani L, Labreuche J, et al. The intracranial-B2LEED3S score and the risk of intracranial hemorrhage in ischemic stroke patients under antiplatelet treatment. Cerebrovasc Dis 2017;43:145–51.
23 Castelijns MC, Helmink MAG, Hageman SHJ, et al. Cohort profile: the Utrecht Cardiovascular Cohort-Second Manifestations of Arterial Disease (UCC-SMART) Study-an ongoing prospective cohort study of patients at high cardiovascular risk in the Netherlands. BMJ Open 2023;13:e066952.
24. Mehran R, Rao SV, Bhatt DL, et al. Standardized bleeding definitions for cardiovascular clinical trials: a consensus report from the bleeding academic research consortium. Circulation 2011;123:2736–47.
25. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur J Heart Fail 2022;24:4–131.
26. Ibánez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. Rev Esp Cardiol (Engl Ed) 2017;70:1082.
27. Bueno H, Pocock S, Danchin N, et al. International patterns of dual antiplatelet therapy duration after acute coronary syndromes. Heart 2017;103:132–8.
28. Schiele F, Puymirat E, Bonello L, et al. Impact of prolonged dual antiplatelet therapy after acute myocardial infarction on 5-year mortality in the FAST-MI 2005 registry. Int J Cardiol 2015;187:354–60.
29. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.
30. Friberg L, Rosenqvist M, Lip GYH. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. Eur Heart J 2012;33:1500–10.
31. Fox KAA, Virdone S, Pieper KS, et al. GARFIELD-AF risk score for mortality, stroke, and bleeding within 2 years in patients with atrial fibrillation. Eur Heart J Qual Care Clin Outcomes 2022;8:214–27.
32. Wester A, Mohammad MA, Olivecrona G, et al. Validation of the 4-item PRECISE-DAPT score: A SWEDEHEART study. J Am Heart Assoc 2021;10.
33. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928–35.
34. De Bacquer D, Ueda P, Reiner Ž, et al. Prediction of recurrent event in patients with coronary heart disease: the EUROASPIRE risk model. Eur J Prev Cardiol 2022;29:328–39.
35. Kaasenbrood L, Boekholdt SM, van der Graaf Y, et al. Distribution of estimated 10-year risk of recurrent vascular events and residual risk in a secondary prevention population. Circulation 2016;134:1419–29.
36. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight — reconsidering the use of race correction in clinical algorithms. N Engl J Med 2020;383:874–82.