

Risk of bleeding with oral anticoagulants: an updated systematic review and performance analysis of clinical prediction rules

Peter Loewen · Karen Dahri

Received: 24 February 2011 / Accepted: 24 May 2011 / Published online: 14 June 2011
© Springer-Verlag 2011

Abstract This study aimed to determine whether progress in developing bleeding risk estimation tools for patients on oral anticoagulant therapy has been made since 2006 when we last systematically reviewed this topic, and to refresh previously published quantitative evaluations of the clinical prediction rules (CPRs) available for estimating bleeding risk in patients on oral anticoagulant therapy. A systematic review of English language literature published since December 2006 was conducted when our previous systematic review ended. Studies were analyzed for predictive performance using likelihood ratios. Six studies detailed CPRs used to assess risk of bleeding prior to commencing warfarin therapy, four of which were included in the analysis. Three studies evaluated new CPRs (“RIETE” and “HAS-BLED”). One of the studies was a further validation of the modified outpatient bleeding risk index (mOBRI) in patients with atrial fibrillation. Individual trial and pooled analyses using likelihood ratios for mOBRI and HAS-BLED showed they have weak predictive accuracy. A RIETE score of 0 point was moderately predictive of the absence of major bleeding. None of the CPRs exhibited sufficient predictive accuracy or had sufficient validation to be recommended for routine use in practice. None of the available CPRs exhibit sufficient predictive accuracy or have trials evaluating the impact of their use on patient outcomes.

P. Loewen
Education & Research, Lower Mainland Pharmacy Services,
University of British Columbia,
Vancouver, British Columbia, Canada

K. Dahri
Vancouver Coastal Health, University of British Columbia,
Vancouver, British Columbia, Canada

P. Loewen (✉)
Faculty of Pharmaceutical Sciences,
University of British Columbia,
2146 East Mall, Vancouver, BC, Canada V6T 1Z3
e-mail: ploewen@interchange.ubc.ca

Hence, no existing oral anticoagulation major bleeding CPR can be recommended for routine use in practice at present.

Keywords Bleeding · Warfarin · Anticoagulants · Clinical prediction rules

Background

Warfarin is commonly used for primary and secondary prevention of numerous thrombotic conditions [1]. New oral anticoagulants (OAC) including dabigatran and rivaroxaban have recently become available in several countries, and randomized controlled trials to support their use in many of these indications now exists [2–6]. Bleeding is the main serious adverse effect when using any of these agents. Estimates of bleeding rates for warfarin vary widely depending on the study design with annual incidences of 0.6% for fatal bleeding, 3.0% for major bleeding, and 9.6% for major or minor bleeding being reported [7]. Newer OACs have been shown to have similar [2, 3, 8] or lower rates of major bleeding [2], depending on the doses used. Many different factors such as concurrent use of interacting medications, labile international normalized ratio (INR), previous bleeding episodes, and age can increase a patient’s risk for bleeding [9, 10]. Risk assessment tools [clinical prediction rules (CPR)] that are able to quantify an individual patient’s risk of bleeding while on OAC would be a valuable addition to clinical practice by aiding clinicians in evaluating the benefits versus risks of initiating OAC.

When we previously systematically reviewed this topic [11], no such tools had been sufficiently evaluated and/or had performance characteristics [defined by likelihood ratios (LR)] sufficient to be recommended for use in routine practice. Given the emergence of new OACs and time for better tools to be developed, the purpose of this review was to systematically review and evaluate the performance of

CPRs for estimating bleeding risk in patients on OAC therapy based on research published since 2006.

Methods

A standardized systematic review methodology was employed [12, 13] with the required adaptations for studies of diagnostic tests or clinical prediction rules.

Criteria for considering studies for this review

Types of studies

We included studies that prospectively or retrospectively evaluated the ability of a CPR using only readily available clinical parameters to distinguish between patients at high and low risk of experiencing major bleeding on warfarin, ximelagatran, rivaroxaban, apixaban, or dabigatran therapy for any indication. We excluded studies of CPRs that were published in abstract format only, were conference proceedings, did not report observed bleeding rates in the CPR risk strata, or were published in languages other than English.

Types of participants

Participants included adults with any condition for which OAC (warfarin, ximelagatran, rivaroxaban, apixaban, or dabigatran) was prescribed. The latter three of these OAC agents were included in our search strategy because they are either recently marketed in several countries or are expected to be soon. Ximelagatran was included because although its development was halted due to non-bleeding-related safety concerns, it has been subjected to large trials that could inform our research questions [14, 15].

Types of intervention

To be included, study participants must have undergone risk stratification via a CPR that used only clinical variables readily available when the decision to commence OAC therapy is being made.

Types of outcome measures

The primary outcome of interest was the ability to distinguish between patients at high and low risk of experiencing major bleeding on warfarin therapy. Predictive ability was defined in terms of LR. Authors' definitions of "major bleeding" were accepted if they were similar to the International Society on Thrombosis and Haemostasis' definition, which includes fatal bleeding, and/or symptomatic bleeding in a critical area or organ such as intracranial, intraspinal, intraocular,

retroperitoneal, intra-articular or pericardial, or intramuscular with compartment syndrome, and/or bleeding causing a fall in hemoglobin level of 20 g/L (1.24 mmol/L) or more, or leading to transfusion of two or more units of whole blood or red cells [16].

Secondary outcome measures included the ability of the CPRs to predict mortality, minor bleeding, and erratic INR (for warfarin only).

Search strategy for identification of studies

A systematic search of the following databases was performed: PubMed, MEDLINE, EMBASE, Cochrane Database of Systematic Reviews, and International Pharmaceutical Abstracts. The search timeframe was 2006 to January 2011 for each given our previously published systematic review covering the literature up to 2006 [11]. The following search terms were used: risk assessment, assessment tools, clinical prediction tools, bleeding risk, bleeding, and warfarin, ximelagatran, dabigatran, apixaban, or rivaroxaban. In MEDLINE, a search strategy previously identified as having 98% sensitivity for identifying CPRs was employed [17]. A manual review of reference lists from retrieved articles was performed to identify any additional studies.

Methods of the review

Initially, two reviewers independently evaluated the titles, abstracts, and citations of all identified articles to select those potentially meeting the inclusion criteria. Articles so selected were independently subjected to full text review to establish whether they met the inclusion criteria and did not meet any exclusion criteria. Concordance between reviewers was measured by simple agreement. Any disagreements were resolved by discussion and consensus. Reviewers were not blinded to author or journal names.

Quality assessment

Both reviewers independently subjected the included studies to quality assessment. Three methods of quality assessment were carried out according to the methodologic standards published by Laupacis et al. [18], McGinn et al. [19], and the hierarchy of evidence for CPRs published by McGinn [19]. For the methodologic quality assessment, each question was assigned a score of 0 (criteria not met), 1 (criteria partially met), and 2 (criteria fully met). The score was summed to give an overall quality score out of 34 points for each study.

Data extraction

Both reviewers independently read each article and abstracted data using a standardized content review form.

Data extracted included: the aim of the study, design, number, type, sample size, duration, inclusion and exclusion criteria, patient characteristics, elements comprising the CPR, and results. Attempts were made to acquire additional information from investigators as required. Discrepancies of data extraction were resolved by consensus through review of the published report.

Analysis of CPR performance characteristics

The LR for each stratum of the CPRs was calculated using published data and the method of Peirce and Cornell [20]. In this case, LRs represent the ratio of the probability of each test result (e.g., low, intermediate, high risk using the CPR) in people who end up experiencing bleeding to the probability in those who did not bleed. Unlike the *c* statistic commonly used to evaluate CPRs and which only provides an overall representation of the percentage of cases in which the CPR correctly predicts the outcome, LR is the most directly clinically applicable measure of diagnostic test performance, particularly when the tests produce more than two strata of results (e.g., low, intermediate, or high risk of bleeding) [21–24]. Likelihood ratios can be used clinically to estimate an individual patient's risk of an outcome when combined with the clinician's baseline level of suspicion of that risk. The cutoffs for performance based on LR proposed by Jaeschke et al. were used in interpreting these estimates [25, 26]. Performance was considered moderate when the LR was greater than 5.0 or less than 0.20. Performance was considered strong when the LR was greater than 10.0 or less than 0.10. Statistically significant LRs, less extreme than these values were considered to have weak predictive performance.

For CPRs where multiple studies (>2) assessed their performance or where more than one study had a validation cohort containing more than 2,000 patients, we attempted to quantitatively combine the results. Our aim was thereby to provide a synthesis of evidence for CPRs that have been the most extensively studied.

Planned subgroup analysis

No subgroup analyses were prospectively planned.

Results

Description of studies

Six studies meeting the inclusion criteria were identified [27–32]. Two of these [27, 32] were excluded due to the authors not reporting the actual bleeding risks identified

within the CPR strata. The four included studies are described in Table 1.

Patient characteristics

Patient characteristics varied within the included studies. Mean age of subjects ranged from 58.4 years [19] to 80.2 years [22]. Three of the studies specifically evaluated patients with atrial fibrillation [28–30] and one focused on patients with VTE/PE [31]. Three of the studies involved only patients receiving warfarin [28, 30, 31], and one involved patients receiving warfarin or ximelagatran [29]. No studies involving dabigatran, apixaban, or rivaroxaban were identified.

Intervention characteristics

One of the studies evaluated a CPR which existed at the time of our previous review [11], the modified outpatient bleeding risk index (mOBRI) [30]. The other three studies described novel CPRs, which we refer to as “RIETE” [31] and “hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile INR, elderly, drugs/alcohol concomitantly score (HAS-BLED)” [28, 29]. Lip et al. [29] are unique in that they simultaneously evaluated the performance of three other CPRs (all of which we have previously evaluated) in a new retrospective cohort of patients.

mOBRI

The outpatient bleeding risk index was first developed from a retrospective review by Landefeld et al. in 1989 [33]. It was subsequently modified by Beyth et al. [34] in a reanalysis of Landefeld's cohort resulting in the mOBRI which includes age ≥ 65 years, history of stroke, history of gastrointestinal bleeding, and serious comorbidity (renal insufficiency, recent myocardial infarction, severe anemia, or diabetes). Others have evaluated the mOBRI in different settings [35, 36].

More recently, Airaksinen et al. [29] examined the mOBRI's performance in patients undergoing percutaneous coronary intervention, 72% of whom were receiving anticoagulation for atrial fibrillation. Only one prior study evaluated this CPR in atrial fibrillation patients [35]. That study evaluated the mOBRI in an anticoagulation clinic and failed to demonstrate moderate or better performance [11]. In Airaksinen et al.'s retrospective study, 9% of the population was deemed to have a high bleeding risk (mOBRI score 3–4). Of these, ten patients (26%) had a major bleed; three of whom were on oral anticoagulation alone and the rest on additional antiplatelet therapy.

Table 1 Included studies

Study (CPR name)	Sample size	Mean age	Length of follow-up	Indication for anticoagulation	Major bleeding definition	Methodology (level of evidence)	Comments and potential sources of bias
Lip 2011 [29] [HAS-BLED, mOBRI, Shireman, HEMORR2HAGES, SBRPS]	7,329 [retrospective], 3,665 received warfarin only. The balance received ximalagatran in the SPORTIF III or V trials	70.9 (SD 8.9) years in "no bleed" population, 73.9 (SD 8.6) years in bleeding event population	Mean 499 days (SD 196)	Atrial fibrillation	Fatal or clinically overt bleeding associated with either transfusion of ≥ 2 U of blood or ≥ 20 g/l decrease in hemoglobin or bleeding involving a critical anatomic site other than the brain parenchyma	Validated in a retrospective database (level 4)	Analysis of the SPORTIF III and V trial databases [14, 15]
Pisters 2010 [28] [HAS-BLED]	Total=3,456 [retrospective], 2,242 received oral anticoagulation	66.8 \pm 12.8 years	1 year	Atrial fibrillation without mitral valve stenosis or valvular surgery	Any bleeding requiring hospitalization and/or causing a decrease in hemoglobin level of >2 g/L and/or requiring blood transfusion that was not hemorrhagic stroke	Derived and validated in a retrospective database (level 4)	Used database from the Euro Heart Survey on AF (data collected between 2003–2004). Includes "labile INR", a post-initiation risk factor. Possible detection bias due to use of retrospective database designed for another purpose
Airaksinen 2010 [30] [mOBRI]	Validation=421 [retrospective]	mOBRI 0, 58 years	Median 1,276 days (3.5 years)	Atrial fibrillation (72%) + undergoing percutaneous coronary intervention (PCI) with stenting	Decrease in blood hemoglobin level ≥ 4.0 g/dl, need for transfusion of ≥ 2 U of blood, need for corrective surgery, occurrence of intracranial or retroperitoneal hemorrhage, or any combination of these	Validated in a retrospective database (level 4)	Possibly confounded by three distinct populations (triple therapy, OAC only, OAC + 1 antiplatelet) with no inclusion of this in the CPR
Ruiz-Gimenez 2008 [31] [RIETE]	Test=13,057 Validation=6,572 [retrospective]	mOBRI 1–2, 71 years mOBRI 3–4, 74 years 66 +/- 17 years	First 3 months of therapy	Venous thromboembolism	Overt and required a transfusion of 2 units of blood or more, or were retroperitoneal, spinal, or intracranial, or when they were fatal.	Derived and validated in a split sample or a retrospective database (level 4)	Possible detection bias due to use of retrospective database designed for another purpose

mOBRI modified outpatient bleeding risk index, *RIETE* Spanish venous thromboembolism registry, *HAS-BLED* hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile INR, elderly, drugs/alcohol concomitantly score, *HEMORR2HAGES* hepatic or renal disease, ethanol abuse, malignancy, older (age >75 years), reduced platelet count or function, hypertension (uncontrolled), anemia, genetic factors, excessive fall risk, and stroke, *OAC* oral anticoagulants, *INR* international normalized ratio, *CPR* clinical prediction rules

Lip et al. also retrospectively evaluated the mOBRI in a cohort of 3,665 warfarin recipients with atrial fibrillation participating in the SPORTIF III and V trials [29]. One hundred thirty-six major bleeds were observed during 1.4 years of observation.

RIETE

A novel six-point major bleeding risk CPR was developed and evaluated using a derivation sample and a validation sample (both retrospective) and logistic regression methods to determine optimal cutoffs for low, intermediate, and high risk of major bleeding [31]. The study data were derived from patients in a large Spanish venous thromboembolism registry (“RIETE”). The CPR involves four clinical variables and two laboratory-based ones (serum creatinine and presence/absence of anemia). Major bleeding rates over 3 months of exposure were 0.1%, 2.8%, and 6.2%, respectively. Based on LRs, the authors concluded that the CPR was capable of distinguishing between high and low bleeding risk (Table 2).

HAS-BLED

This most recent bleeding risk CPR was retrospectively evaluated patients in the Euro Heart Survey on Atrial Fibrillation cohort using data derived in this population along with a literature review to develop their seven-characteristic scale. HAS-BLED includes hypertension, abnormal renal and liver function, stroke, bleeding history, labile INRs, age ≥ 75 , aspirin/NSAID use, and alcohol use (1 point each). Based on their findings, the authors assert that the risk of bleeding for a particular patient will outweigh the potential benefit of oral anticoagulation if the HAS-BLED score exceeds the CHADS₂ score for atrial fibrillation-associated stroke risk [37]. HAS-BLED was also retrospectively evaluated in a cohort of 7,329 warfarin or ximelagatran recipients participating in the SPORTIF atrial fibrillation trials [29].

None of the included studies addressed our secondary outcome measures of mortality, minor bleeding, or erratic INR (for warfarin only).

Methodological quality of included studies

Table 3 lists the scores for the quality assessment completed by the reviewers for the included studies. According to the predefined criteria, the studies included in this update of our previous review were of similar or slightly higher quality than those previously evaluated. The new studies all lacked prospective validation or prospective evaluation of their effects on clinical use. The HAS-BLED CPR is the only one for which its proponents propose a course of

action based on a score [28], although the effects of doing so have not been studied.

Levels of evidence

Table 1 depicts the design and population characteristics of the included studies and their level of evidence. The included studies were all derived and/or validated in either a split sample or a retrospective database.

Based on the evidence hierarchy by McGinn et al. [9] and our previous analysis [11], the mOBRI would still be classified as having level 2 evidence as it has been validated in different populations prospectively and has demonstrated reproducibility. Evidence that its application has changed clinician-prescribing behavior with favorable outcomes is not yet available. The two new CPRs evaluated would be classified as having level 4 evidence “Rules that need further evaluation before they can be applied clinically” [9].

Predictive ability of included studies

The performance characteristics of the included studies are described in Table 2. Based on the finding of five studies evaluating the mOBRI and two studies involving more than 2,000 subjects each for HAS-BLED, we chose to perform a pooled analyses of those studies, also depicted in Table 2.

Our analysis on the mOBRI data of Airaksinen et al. [30] revealed performance characteristics somewhat worse than in our previous pooled analysis of studies involving this CPR [11]. Only for “high risk” did the LR achieve statistical significance and its magnitude was in the “weak” category. Airaksinen et al. used survival analysis to assess differences in major adverse cardiac events or mortality between the low/intermediate/high bleeding risk categories and found an association only with the latter. They did not report other measures of CPR predictive ability. Our pooled analysis of all five studies involving the mOBRI (Table 2) shows that the “low” and “high” risk strata have weak predictive ability and classifying a patient as intermediate risk has no predictive utility at all.

The RIETE CPR [31] showed statistical significance for the low- and high-risk strata, and the LRs were considered strong for the low-risk stratum (LR 0.03). These authors used LRs to assess predictive ability in the validation cohort and our analysis closely matched theirs. In other words, a “low-risk” classification on the RIETE CPR may be clinically useful for identifying low bleeding risk venous thromboembolism patients. This requires prospective validation.

The authors of the HAS-BLED analyses did not recommend a low- or high-risk stratification scheme, and instead recommend that the decision to initiate oral

Table 2 Performance characteristics of CPRs in validation groups

Study (mean follow-up duration) (CPR name)	Number, major + minor bleeds (%)	Number, major bleeds (%)	Total in group	LR (95% CI) major + minor bleeding	LR (95% CI) major bleeding	Predictive power ^a
Lip [29] (1.4 years) (HAS-BLED) full cohort (warfarin and ximelagatran patients combined)						
0	N/A	21 (1.2)	1,757	N/A	0.36 (0.24–0.55)	Weak
1	N/A	75 (2.8)	2,717	N/A	0.86 (0.71–1.04)	NS
2	N/A	63 (3.6)	1,752	N/A	1.13 (0.91–1.40)	NS
3	N/A	50 (6.0)	834	N/A	1.93 (1.50–2.49)	Weak
4	N/A	23 (9.5)	241	N/A	3.12 (2.12–4.81)	Weak
5	N/A	2 (7.4)	27	N/A	2.42 (0.57–10.2)	NS
6	N/A	0	1	N/A	N/A	
≥2 vs. <2	N/A	138 (2.6) vs. 96 (2.1)	2,855 vs. 4,474	N/A	1.54 (1.38–1.72)	Weak
≥3 vs. <3	N/A	75 (6.8) vs. 159 (2.6)	1,103 vs. 6,226	N/A	2.21 (1.82–2.69)	Weak
≥4 vs. <4	N/A	25 (9.3) vs. 209 (3.0)	269 vs. 7,060	N/A	3.11 (2.1–4.59)	Weak
Lip [29] (1.4 years) (Shireman) warfarin patients only						
Low	N/A	99 (3.3)	3,013	N/A	0.88 (0.79–0.98)	Weak
Moderate	N/A	37 (5.7)	649	N/A	1.57 (1.18–2.08)	Weak
High	N/A	0	3	N/A	N/A	
Lip [29] (1.4 years) (HEMORR2HAGES) warfarin patients only						
Low	N/A	81 (3.0)	2,694	N/A	0.80 (0.70–0.93)	Weak
Moderate	N/A	53 (6.1)	872	N/A	1.68 (1.35–2.09)	Weak
High	N/A	2 (2.0)	99	N/A	0.54 (0.13–2.15)	NS
Lip [29] (1.4 years) (mOBRI) warfarin patients only						
Low	N/A	8 (2.1)	374	N/A	0.57 (0.29–1.12)	NS
Moderate	N/A	113 (3.9)	2,917	N/A	1.05 (0.97–1.13)	NS
High	N/A	15 (4.0)	374	N/A	1.08 (0.67–1.77)	NS
Lip [29] (1.4 years) (SBRPS) warfarin patients only						
Low	N/A	11 (3.0)	330	N/A	0.89 (0.50–1.59)	NS
Moderate	N/A	120 (3.8)	3,141	N/A	1.03 (0.97–1.10)	NS
High	N/A	5 (2.6)	194	N/A	0.69 (0.29–1.64)	NS
Pisters [28] (1 year) (HAS-BLED)						
0	N/A	9 (1.1)	798	N/A	0.71 (0.40–1.3)	NS
1	N/A	13 (1.0)	1,286	N/A	0.64 (0.40–1.03)	NS
2	N/A	14 (1.9)	744	N/A	1.21 (0.77–1.89)	NS
3	N/A	7 (3.7)	187	N/A	2.45 (1.21–4.93)	Weak
4	N/A	4 (8.7)	46	N/A	6.0 (2.23–16.0)	Moderate
5	N/A	1 (12.5)	8	N/A	9.0 (1.12–71.7)	Moderate
6 ^d	N/A	0 (0)	2	N/A	N/A	
≥2 vs. <2 ^c	N/A	26 (2.6) vs. 22 (1.0)	987 vs. 2,084	N/A	1.70 (1.31–2.22)	Weak
≥3 vs. <3	N/A	12 (4.9) vs. 48 (1.6)	243 vs. 3,071	N/A	3.27 (1.97–5.42)	Weak
≥4 vs. <4	N/A	5 (9.2) vs. 43 (1.4)	54 vs. 3,017	N/A	6.42 (2.68–15.4)	Moderate
Lip and Pisters pooled [28, 29] (HAS-BLED)						
0	N/A	30 (1.2)	2,555	N/A	0.43 (0.30–0.60)	Weak
1	N/A	88 (2.2)	4,003	N/A	0.81 (0.68–0.96)	Weak
2	N/A	77 (3.1)	2,496	N/A	1.14 (0.94–1.39)	NS
3	N/A	57 (5.6)	1,021	N/A	2.12 (1.67–2.70)	Weak
4	N/A	27 (9.4)	287	N/A	3.73 (2.55–5.44)	Weak
5	N/A	3 (8.6)	35	N/A	3.36 (1.03–10.9)	Weak
6	N/A	0 (0)	3	N/A	N/A	
≥2 vs. <2	N/A	164 (4.2) vs. 118 (1.8)	3,842 vs. 6,558	N/A	1.60 (1.44–1.77)	Weak
≥3 vs. <3	N/A	87 (6.5) vs. 195 (2.2)	1,346 vs. 9,054	N/A	2.48 (2.07–2.98)	Weak
≥4 vs. <4	N/A	30 (8.5) vs. 252 (2.5)	352 vs. 10,048	N/A	3.34 (2.3–4.77)	Weak

Table 2 (continued)

Study (mean follow-up duration) (CPR name)	Number, major + minor bleeds (%)	Number, major bleeds (%)	Total in group	LR (95% CI) major + minor bleeding	LR (95% CI) major bleeding	Predictive power ^a
Ruiz-Gimenez [31] (3 months) (RIETE)						
Low	N/A	1 (0.1)	1,340		0.03 (0.01–0.20) (reported in publication) 0.03 (0.004–0.21) (authors' calculation)	Strong
Intermediate	N/A	137 (2.8)	4,891		1.16 (0.92–1.48) (reported in publication) 1.16 (1.09–1.23) (authors' calculation)	NS
High	N/A	21 (6.2)	341		2.65 (1.61–4.32) (reported in publication) 2.65 (1.75–4.00) (authors' calculation)	Weak
Airaksinen [30] (median 3.5 years) (mOBRI)						
Low	N/A	4 [6]	64	N/A	0.40 (0.15–1.083)	NS
Intermediate	N/A	45 [14]	319	N/A	1.0 (0.86–1.18)	NS
High	N/A	10 [26]	38	N/A	2.19 (1.12–4.27)	Weak
Pooled Lip, Airaksinen Byeth, Aspinall, Wells [29, 30, 34–36] (mOBRI)						
Low	N/A ^b	15 (1.9)	776	N/A ^b	0.40 (0.24–0.65)	Weak
Intermediate	N/A ^b	208 (4.7)	4,437	N/A ^b	0.99 (0.91–1.05)	NS
High	N/A ^b	54 (8.6)	628	N/A ^b	1.89 (1.47–2.43)	Weak

Calculations performed using <http://spph.ubc.ca/sites/healthcare/files/calc/bayes.html>

mOBRI modified outpatient bleeding risk index, *RIETE* Spanish venous thromboembolism registry, *HAS-BLED* hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile INR, elderly, drugs/alcohol concomitantly score, *HEMORR2HAGES* hepatic or renal disease, ethanol abuse, malignancy, older (age >75 years), reduced platelet count or function, hypertension (uncontrolled), anemia, genetic factors, excessive fall risk, and stroke (1 point each), *CPR* clinical prediction rules, *CI* confidence interval, *NS* not significant, *LR* likelihood ratio

^a Predictive power, LR <0.2 or >5 = “moderate”; LR <0.1 or >10 = “strong”; LR less extreme than these but statistically different from 1.0 = “weak”. Non-statistically significant LR = “NS”

^b Not pooled due to Byeth, Airaksinen, and Lip not reporting minor bleeding events

^c The authors do not specify risk groupings but suggest that HAS-BLED score should be compared to CHADS₂ score, whose critical threshold for recommending anticoagulant therapy is 2

^d Although a nine-point scale is presented in the publication, no patients were in any strata above 6

anticoagulation be based on comparing the HAS-BLED score to the CHADS₂ score for stroke risk. According to the authors, “the risk of bleeding outweighs the potential benefit of OAC if the HAS-BLED bleed score exceeds the individual CHADS₂ index” [28]. In a study of Pisters et al. [28] this was based on the *c* statistic, which was 0.72 (0.65–0.79) in the validation cohort, corresponding to having “modest value” and below the threshold of 0.8 for “genuine clinical utility” [38, 39]. In addition to analyzing the LRs for each of the six risk scores, we performed the following additional analyses: To explore for threshold effects, we chose to analyze HAS-BLED scores of ≥ 2 vs. <2, ≥ 3 vs. <3, and ≥ 4 vs. <4 in our analysis of the individual studies and the pooled analysis. In the study of Pisters et al. [28], scores of 4 or 5 showed moderate predictive ability [LR 6.0 (2.23–16.0) and 9.0 (1.12–71.7), respectively] and a score of ≥ 4 was associated with moderate predictive ability [LR 6.42 (2.68–15.4)]. In the analysis of Lip et al. [29], much larger cohort did not reveal

any score stratum with more than weak predictive ability, and no threshold effects were detected. Their analysis was based on *c* statistics, which were in the 0.6–0.7 range for all the reported analyses, corresponding to “limited clinical value” [38, 39]. For the pooled HAS-BLED analysis, none of the risk strata or threshold groupings yielded LRs any better than “weak” and the threshold effect detected in the study of Pisters et al. [28] was not preserved. This may represent a more accurate estimate of the effect or an artifact of pooling statistically or clinically heterogeneous data.

Comments

Since our previous evaluation of performance of CPRs for bleeding [11], two new CPRs have been developed (RIETE and HAS-BLED) and the mOBRI has been further studied.

Table 3 Quality characteristics and scores for the included studies, based on the method of Laupacis [18]

Study	Airaksinen [30] (mOBRI)	Ruiz-Gimenez [31] (RIETE)	Pisters [28] (HAS-BLED)	Lip [29] (HAS-BLED)
Outcome				
Definition	2	2	2	2
Clinical importance	2	2	2	2
Blind assessment	0	0	0	0
Predictive variables				
Identification and definition	2	2	2	2
Blind assessment	0	0	0	0
Important patient characteristics described	2	2	2	2
Study site described	2	2	2	2
Mathematical techniques described	2	2	2	2
Results of the rule described	2	2	1	2
Reproducibility				
Of predictive variables	2	2	1	2
Of the rule	0	0	0	0
Sensibility				
Clinically sensible	2	2	2	2
Easy to use	2	2	2	2
Probability of disease described	2	2	2	2
Course of action described	0	0	2	0
Prospective validation	0	0	0	0
Effects of clinical use prospectively measured	0	0	0	0
Total (/34)	22	22	22	22

0=criteria not met, 1=criteria partially met, 2=criteria fully met

mOBRI modified outpatient bleeding risk index, *RIETE* Spanish venous thromboembolism registry, *HAS-BLED* hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile INR, elderly, drugs/alcohol concomitantly score

Quality

Based on the comprehensive rating system used, we consider the overall quality of the available studies to be poor. For any of the existing CPRs to be recommended for routine use in clinical practice, prospective validation of their use and, ultimately, evidence that using them improves patient outcomes would be required [19]. At this point, no such evidence exists and no progress toward it is evident since our last review.

Performance

Using pre-specified thresholds for clinical usefulness based on LR estimates (LR >10, >5 or <0.2, <0.1) in only one case was moderate or better performance detected: a RIETE score of 0 points, which is strongly predictive of the absence of major bleeding [LR 0.03 (0.01–0.20)]. The finding in one study [28] that a HAS-BLED score ≥ 4 was moderately predictive did not persist when pooled with data from another study involving similar patients [29].

Based on our pooled analysis of all trials involving the mOBRI (Table 2), we believe that sufficient evidence is now available that this CPR lacks the ability to make clinically useful distinctions and it is doubtful that more studies are warranted. A limitation of our approach is that because the mOBRI has been studied in a variety of patient populations, pooling could obscure clinically useful LRs. We believe this is unlikely, however, since none of the individual studies of the mOBRI demonstrated such LRs either.

HAS-BLED has recently been endorsed by the European Society of Cardiology and the Canadian Cardiovascular Society for routine use in clinical practice [40, 41]. This seems premature given the CPR authors' conclusion that it requires prospective validation prior to routine use [28], *c* statistics indicating only limited clinical value in a large cohort study [29], and our finding that it lacks power to predict major bleeding in any of the evaluable strata. Hence, the authors' assertion that bleeding risk exceeds the prospect of benefit when the HAS-BLED score is greater than the CHADS₂ score [28] requires prospective evalua-

tion before being adopted as a decision-making rule. Also, adoption of such a rule in practice would somehow have to integrate the evidence that many patients place significantly more disutility on stroke than on major bleeding [42, 43].

Several explanations for the lack of moderate to strong predictive ability of the studied CPRs are possible. In studies with short follow-up periods and/or subjects at low inherent risk of bleeding, the incidence of major bleeding may be too low to result in sufficiently large LRs or results in excessively wide confidence intervals. The predictive variables used in existing CPRs simply may not contain enough useful information to distinguish between different risks of bleeding. Many reasons for this are possible, prominently including that in CPR derivation populations exposure to bleeding risk factors like aspirin, non-steroidal anti-inflammatory drugs, oral anticoagulants, or combinations thereof are so heterogeneous (by dose, drug, or duration) that these do not emerge in regression models as precisely as necessary to be predictive. This may be especially true of the mOBRI, for which the LR estimates are fairly precise and not very predictive. The list of patient factors that can increase bleeding risk is extensive and includes genetic factors, age, sex, personal characteristics such as compliance and dietary intake, comorbid conditions such as cancer and liver disease, and concurrent medications [44]. The metabolism and action of warfarin itself have been associated with 30 different genes with polymorphisms leading to large interindividual variations in dosing requirements and higher bleeding risks [44]. Perhaps this complexity makes creation of a clinical prediction rule that can be easily understood by clinicians and widely used in different populations impossible.

Applicability to practice

The mOBRI, RIETE, and HAS-BLED CPRs involve a small number of easily ascertained variables. HAS-BLED includes labile INR, a factor that can only be determined after initial warfarin exposure. Presumably this does not disqualify warfarin-naïve patients, however, who could simply not be given a point for that parameter.

Using the dual criteria of levels of evidence and quantitative performance characteristics (LR), none of the identified CPRs are supported by sufficient evidence to recommend their adoption in clinical practice. The mOBRI and HAS-BLED are the most developed in terms of quantity of evidence; however, they both suffer from poor predictive performance and their widespread uptake into clinical practice could only be recommended once a positive impact on clinical decision making had been demonstrated. A low score on the RIETE CPR appears to be useful for identifying venous thromboembolism patients

at negligible risk of bleeding, but this requires further validation.

Conclusion

Bleeding risk CPRs could provide clinicians with valuable individualized information to aid in decision making prior to starting oral anticoagulation therapy. Unfortunately, none of the available CPRs exhibit sufficient predictive accuracy or have trials evaluating the impact of their use on patient outcomes. Hence, it remains the case that no existing CPR can be recommended for widespread use in practice at present. A low RIETE score is promising as a means to identify patients at extremely low risk of bleeding. The current evidence does not support the use of HAS-BLED or mOBRI in routine practice. While the goal of accurately estimating a patient's risk of major bleeding remains elusive, risk factors that are known to increase an individual patient's risk of bleeding should be evaluated and minimized if possible. Further prospective trials are required to develop a CPR that can be reliably employed in clinical practice.

Conflicts of interest No funding was provided to assist in the preparation of the manuscript. The authors have no potential conflicts of interest to declare.

References

1. Ansell J, Hirsh J, Hylek E et al (2008) The pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest* 133:160S–198S
2. Connolly SJ, Ezekowitz MD, Yusuf S, et al. (2009) Dabigatran versus warfarin in patients with atrial fibrillation (RE-LY). *N Engl J Med* 361. [10.1056/NEJMoa0905561](https://doi.org/10.1056/NEJMoa0905561)
3. Schulman S, Kearon C, Kakkar AK, et al. (2009) Dabigatran versus warfarin in the treatment of acute venous thromboembolism (RE-COVER). *N Engl J Med* 361. [10.1056/NEJMoa0906598](https://doi.org/10.1056/NEJMoa0906598)
4. Eriksson BI, Dahl OE, Rosencher N et al (2007) Dabigatran etexilate versus enoxaparin for prevention of venous thromboembolism after total hip replacement: a randomised, double-blind, non-inferiority trial (RE-NOVATE). *Lancet* 370:949–956
5. Turpie AGG, Lassen MR, Davidson BL et al (2009) Rivaroxaban versus enoxaparin for thromboprophylaxis after total knee arthroplasty (RECORD4): a randomised trial. *Lancet* 373:1673–1680
6. Lassen MR, Raskob GE, Gallus A et al (2010) Apixaban versus enoxaparin for thromboprophylaxis after knee replacement (ADVANCE-2): a randomised double-blind trial. *Lancet* 375:807–815
7. Landefeld CS, Beyth RJ (1993) Anticoagulant-related bleeding: clinical epidemiology, prediction, and prevention. *Am J Med* 95 (3):315–328
8. ROCKET AF Study Investigators (2010) Rivaroxaban-once daily, oral, direct factor XA inhibition compared with vitamin K antagonism for prevention of stroke and embolism trial in atrial fibrillation rationale and design of the ROCKET AF study. *Am Heart J* 159(3):340–347.e.1

9. Gasse C, Hollowell J, Meier CR et al (2005) Drug interactions and risk of acute bleeding leading to hospitalisation or death in patients with chronic atrial fibrillation treated with warfarin. *Thromb Haemost* 94(3):537–543
10. Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S (2007) Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. *Circulation* 115:2689–2696
11. Dahri K, Loewen P (2007) The risk of bleeding with warfarin: a systematic review and performance analysis of clinical prediction rules. *Thromb Haemost* 98:980–987
12. The Cochrane Library. Cochrane reviewers' handbook 5.0.2. 2009; available at <http://www.cochrane-handbook.org>. Accessed 12 December 2010
13. Liberati A, Altman DG, Tetzlaff D et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 151:W-65–W-94
14. Sportif III (2003) Investigators. Stroke prevention with the oral direct thrombin inhibitor ximelagatran compared with warfarin in patients with non-valvular atrial fibrillation (SPORTIF III): randomised controlled trial. *Lancet* 362:1691–1698
15. Sportif V (2005) Investigators. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation. *JAMA* 293:690–698
16. Schulman S, Kearon C (2005) Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. *J Thromb Haemost* 3:692–694
17. Ingui BJ, Rogers MAM (2001) Searching for clinical prediction rules in Medline. *J Am Med Inform Assoc* 8(4):391–397
18. Laupacis A, Sekar N, Stiell IG (1997) Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA* 277(6):488–494
19. McGinn TG, Guyatt GH, Wyer PC et al (2000) Users' guides to the medical literature XXII: how to use articles about clinical decision rules. *JAMA* 284(1):79–84
20. Peirce JC, Cornell RG (1993) Integrating stratum-specific likelihood ratios with the analysis of ROC curves. *Med Decis Making* 13(2):141–151
21. Deeks JJ, Altman DG (2004) Diagnostic tests 4: likelihood ratios. *BMJ* 329(7458):168–169
22. Greenhalgh T (1997) How to read a paper: papers that report diagnostic or screening tests. *BMJ* 315:540–543
23. Sackett DL, Haynes RB, Guyatt GH, Tugwell P (1991) *Clinical epidemiology—a basic science for clinical medicine*. London, Little Brown 51–68
24. Cook NR (2007) Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 115:928–935
25. Jaeschke R, Guyatt G, Sackett DL (1994) Users' guides to the medical literature: III. How to use an article about a diagnostic test: A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 271(5):389–391
26. Jaeschke R, Guyatt GH, Sackett DL (1994) User's guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA* 271(9):703–707
27. van der Meer FJM, Rosendaal FR, Vandenbroucke JP et al (1996) Assessment of a bleeding risk index in two cohorts of patients treated with oral anticoagulants. *Thromb Haemost* 76(1):12–16
28. Pisters R, Lane DA, Nieuwlaat R et al (2010) A novel user-friendly score (HAS-BLED) to assess one-year risk of major bleeding in atrial fibrillation patients: the Euro Heart Survey. *Chest* 138(5):1093–1100
29. Lip GYH, Frison L, Halperin JL, Lane DA (2011) Comparative validation of a novel risk score for predicting bleeding risk in anticoagulated patients with atrial fibrillation: the HAS-BLED (hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile INR, elderly, drugs/alcohol concomitantly) score. *J Am Coll Cardiol* 57:173–180
30. Airaksinen KEJ, Suurmunne H, Porela P et al (2010) Usefulness of outpatient bleeding risk index to predict bleeding complications in patients with long-term oral anticoagulation undergoing coronary stenting. *Am J Cardiol* 106:175–179
31. Ruiz-Gimenez N et al (2008) Predictive variables for major bleeding events in patients presenting with document acute venous thromboembolism: findings from the RIETE registry. *Thromb Haemost* 100:26–31
32. Wess ML, Schauer DP, Johnston JA et al (2007) Application of a decision support tool for anticoagulation in patients with non-valvular atrial fibrillation. *J Gen Intern Med* 23(4):411–417
33. Landefeld CS, Goldman L (1989) Major bleeding in outpatients treated with warfarin: incidence and prediction by factors known at the start of outpatient therapy. *Am J Med* 87(2):144–152
34. Beyth RJ, Quinn LM, Landefeld CS (1998) Prospective evaluation of an index for predicting the risk of major bleeding in outpatients treated with warfarin. *Am J Med* 105(2):91–99
35. Aspinall SL, DeSanzo BE, Trilli LE et al (2005) Bleeding risk index in an anticoagulation clinic: assessment by indication and implications for care. *J Gen Intern Med* 20(11):1008–1013
36. Wells PS, Forgie MA, Simms M et al (2003) The outpatient bleeding risk index: validation of a tool for predicting bleeding rates in patients treated for deep venous thrombosis and pulmonary embolism. *Arch Intern Med* 163(8):917–920
37. Gage BF, Waterman AD, Shannon W et al (2001) Validation of clinical classification schemes for predicting stroke. *JAMA* 285:2864–2870
38. Ohman EM, Granger CB, Harrington RA, Lee KL (2000) Risk stratification and therapeutic decision making in acute coronary syndromes. *JAMA* 284:876–878
39. Harrell E, Califf R, Pryor DB, Lee KL, Rosati RA (1982) Evaluating the yield of medical tests. *JAMA* 247:2543–2546
40. Canadian Cardiovascular Society (2010) Canadian Cardiovascular Society 2010 Atrial Fibrillation Guidelines. Available at http://www.ccs guidelineprograms.ca/index.php?option=com_content&view=article&id=111&Itemid=66. Accessed 12 December 2010
41. The Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC) (2010) Guidelines for the management of atrial fibrillation. *Eur Heart J* 31:2369–2429. doi:10.1093/eurheartj/ehq278
42. Solomon NA, Glick HA, Russo CJ, Lee J, Schulman KA (1994) Patient preferences for stroke outcomes. *Stroke* 25:1721–1725
43. Devereaux PJ, Anderson DR, Gardner MJ et al (2001) Differences between perspectives of physicians and patients on anticoagulation in patients with atrial fibrillation: observational study. *Br Med J* 323:1218–1222
44. Palareti G, Cosmi B (2009) Bleeding with anticoagulation therapy—who is at risk, and how best to identify such patients. *Thromb Haemost* 102:268–278