



Published in final edited form as:

*Circulation*. 2023 September 19; 148(12): 936–946. doi:10.1161/CIRCULATIONAHA.123.064556.

## Development and Validation of the DOAC Score: A Novel Bleeding Risk Prediction Tool for Patients with Atrial Fibrillation on Direct-Acting Oral Anticoagulants

Rahul Aggarwal, MD<sup>1,2</sup>, Christian T. Ruff, MD, MPH<sup>3</sup>, Saverio Viridone, MSc<sup>4</sup>, Sylvie Perreault, BPharm PHD<sup>5</sup>, Ajay K. Kakkar, MBBS, PhD<sup>4,6</sup>, Michael G. Palazzolo, MPH<sup>3</sup>, Marc Dorais, MSc<sup>7</sup>, Gloria Kayani, BSc<sup>4</sup>, Daniel E. Singer, MD<sup>8</sup>, Eric Secemsky, MD<sup>1</sup>, Jonathan Piccini, MD, MHS<sup>9</sup>, Usman A. Tahir, MD, MBI<sup>1</sup>, Changyu Shen, PhD<sup>10</sup>, Robert W. Yeh, MD, MSc<sup>1</sup>

<sup>1</sup>Richard A. and Susan F. Smith Center for Outcomes Research in Cardiology, Division of Cardiology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts

<sup>2</sup>Heart and Vascular Center, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts

<sup>3</sup>TIMI Study Group, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States

<sup>4</sup>Thrombosis Research Institute, London, United Kingdom

<sup>5</sup>Faculty of Pharmacy, University of Montreal, Montreal, Quebec, Canada

<sup>6</sup>University College London, London, United Kingdom

<sup>7</sup>StatSciences Inc., Notre-Dame-de-l'Île-Perrot, Quebec, Canada

<sup>8</sup>Division of General Internal Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts

<sup>9</sup>Duke Clinical Research Institute, Duke University, Durham, North Carolina

<sup>10</sup>Advanced Analytics, Biogen Digital Health and Worldwide Medical, Cambridge, Massachusetts

### Abstract

**Background:** Current clinical decision tools for assessing bleeding risk in individuals with atrial fibrillation (AF) have limited performance and were developed for individuals treated with warfarin. This study develops and validates a clinical risk score to personalize estimates of bleeding risk for individuals with AF taking direct-acting oral anticoagulants (DOACs).

**Methods:** Among individuals taking dabigatran 150mg twice per day from 44 countries and 951 centers in this secondary analysis of the RE-LY trial, a risk score was developed to

---

**Corresponding Author:** Robert W. Yeh, MD, MSc, Richard A. and Susan F. Smith Center for Outcomes Research in Cardiology, Beth Israel Deaconess Medical Center, Harvard Medical School, 375 Longwood Avenue, 4th Floor, Boston, MA 02215, ryeh@bidmc.harvard.edu.

determine the comparative risk for bleeding, based on covariates derived in a Cox proportional hazards model. The risk prediction model was internally validated with bootstrapping. The model was then further developed in the GARFIELD-AF registry, with individuals taking dabigatran, edoxaban, rivaroxaban, and apixaban. To determine generalizability in external cohorts and among individuals on different DOACS, the risk prediction model was validated in the COMBINE-AF pooled clinical trial cohort and the RAMQ administrative database. The primary outcome was major bleeding. The risk score, termed the DOAC Score, was compared to the HAS-BLED score.

**Results:** Of the 5684 patients in RE-LY, 386 (6.8%) experienced a major bleeding event, within a median follow-up of 1.74-years. The prediction model had an optimism-corrected C statistic of 0.73 after internal validation with bootstrapping and was well-calibrated based on visual inspection of calibration plots (goodness-of-fit  $P = 0.57$ ). The DOAC Score assigned points for age, creatinine clearance/glomerular filtration rate, underweight status, stroke/transient ischemic attack/embolism history, diabetes, hypertension, antiplatelet use, non-steroidal anti-inflammatory use, liver disease, and bleeding history, with each additional point scored associated with a 48.7% (95% CI: 38.9%-59.3%,  $P < 0.001$ ) increase in major bleeding in RE-LY. The score had superior performance to the HAS-BLED score in RE-LY (C Statistic: 0.73 vs 0.60,  $P$  for difference  $< 0.001$ ) and among 12,296 individuals in GARFIELD-AF (C statistic: 0.71 vs 0.66,  $P$  for Difference = 0.025). The DOAC Score had stronger predictive performance than the HAS-BLED score in both validation cohorts, including 25,586 individuals in COMBINE-AF (C statistic: 0.67 vs 0.63,  $P$  for Difference  $< 0.001$ ) and 11,945 individuals in RAMQ (C statistic: 0.65 vs 0.58,  $P$  for Difference  $< 0.001$ ).

**Conclusion:** In individuals with AF potentially eligible for DOAC therapy, the DOAC Score can help stratify patients based on expected bleeding risk.

### Keywords

Atrial fibrillation; Bleeding; Prediction; Apixaban; Rivaroxaban; Dabigatran; Edoxaban. DOAC Score

## INTRODUCTION

Patients with atrial fibrillation (AF) are at an increased risk for antithrombotic events (ischemic stroke/systemic emboli), a risk that is reduced with oral anticoagulation.[1] When deciding to initiate anticoagulation, clinicians must balance the tradeoffs of decreasing the risk of antithrombotic events against increasing the risk of bleeding.[1] While the decision to prescribe oral anticoagulants is straight-forward for certain patients, assessing this tradeoff can be particularly challenging in some patients, especially in the elderly, those with multiple comorbidities, and those at increased risk for bleeding.[1,2]

The CHA<sub>2</sub>DS<sub>2</sub>-VASc risk score was developed to assist clinicians in estimating the risk of ischemic stroke/systemic emboli risk in individuals with AF[3] and has been endorsed for clinical use by major cardiovascular societies.[1,4,5] Assessing bleeding risk, however, has been more challenging, as bleeding risk scores have exhibited limited discrimination, are difficult to use in clinical settings, have received fewer endorsements by professional societies, and are consequently less clinically useful.[1,6–9] The most popular and predictive

clinical tool for determining bleeding risk in patients with AF is the HAS-BLED score. [10,11] This risk score, however, has demonstrated limited accuracy in multiple studies and was developed in cohorts of few bleeding events.[6,8–10] Additionally, HAS-BLED was developed for patients taking warfarin, whereas many patients are now treated with direct-acting oral anticoagulants (DOACs).[1] The risk score requires specifying warfarin anticoagulation quality (i.e., labile international normalized ratio [INR]), a metric that does not apply to patients taking DOAC, making incorporation of this score for patients on DOAC therapy, challenging. Having an improved decision tool to assess bleeding risk for individuals on more contemporary therapies would inform shared-decision making conversations with patients who have AF.[7]

In order to address the current clinical need for a bleeding risk prediction tool relevant to contemporary populations of AF patients, the “DOAC Score” was developed and validated.

## METHODS

Because of the sensitive nature of the data collected for this study, requests to access the dataset from qualified researchers trained in human subject confidentiality protocols may be sent to Vivli (<https://vivli.org>), GARFIELD-AF investigators (<https://af.garfieldregistry.org>), TIMI Study Group (<https://timi.org>), or to RAMQ (<https://www.ramq.gouv.qc.ca/en>).

## DEVELOPMENT COHORT

The bleeding risk prediction tool was initially developed in the RE-LY trial (N=18,113) dabigatran 150mg group. [12] The RE-LY trial has previously been described.[12] In brief, RE-LY was a multicenter clinical trial evaluating antithrombotic and bleeding outcomes for patients with AF on dabigatran 150 mg, dabigatran 110 mg, or warfarin. Individuals were enrolled from 951 clinical centers and 44 different countries, with enrollment occurring from 2005-2007 (follow-up through 2009).[12] Initial model development was based on the 150mg twice daily arm of RE-LY. The model was then further developed, prior to validation, among GARFIELD-AF individuals,[13] as GARFIELD-AF included a large proportion of patients on apixaban and rivaroxaban, so development could occur among individuals on a variety of DOACs. This second stage of development was termed as model refinement.

For model development, individuals were included in the RE-LY development cohort if they were treated with a standard dose DOAC (dabigatran 150 mg twice daily) and were required to be on treatment. Individuals with missing follow-up data (N=48) or missing predictor variable data (N=344) were excluded (final cohort: N=5684). While standard dose DOAC was used for the development cohort, the performance of the algorithm for individuals prescribed dabigatran 110 mg twice daily was also evaluated, as this dose is approved for use in certain non-US countries and has different efficacy and safety outcomes.[12] The same inclusion/exclusion criteria were applied for this cohort (final low-dose cohort: N = 5594). Outcomes in RE-LY were adjudicated by experts blinded to the treatment received by the individual.

## PREDICTORS

A total of 15 baseline variables were selected as potential predictors of bleeding outcomes. Predictor variables were selected based on prior evidence for association with bleeding risk or clinical judgement,[8,10,14] and included: age, sex, body mass index, creatinine clearance/glomerular filtration rate (calculated by the Cockcroft–Gault equation), smoking history, alcohol use, stroke/transient ischemic attack/systemic embolism history, prior bleeding history (minor or major), liver disease, coronary artery disease, heart failure, diabetes, hypertension, aspirin or dual antiplatelet therapy use, and non-steroidal anti-inflammatory use.[6,10,14] Variables were determined prior to model development. Recognizing that creatinine clearance was used in RE-LY for development, and glomerular filtration rate is more clinically accessible, cohorts enriched for categorization based on chronic kidney disease stages (GARFIELD-AF, and RAMQ) were included, as these classifications are based on glomerular filtration rate.[15] Definitions and categorizations of the variables are presented in detail in Supplemental Table 1.

## OUTCOMES

The primary outcome was major bleeding at one-year. Major bleeding in RE-LY was defined per the International Society on Thrombosis and Haemostasis (ISTH) criteria: a 20 gram per liter reduction in hemoglobin, 2-unit transfusion need, fatal bleeding, or symptomatic bleeding in a critical area or organ.[12,16] The secondary outcome was life-threatening bleeding at one-year, a subset of major bleeding. Life-threatening bleeding was defined as bleeding causing death, bleeding resulting in a 50 gram per liter reduction in hemoglobin, bleeding requiring a 4 unit or greater transfusion, bleeding requiring inotropic agents, or bleeding necessitating surgical intervention.[12]

## DEVELOPMENT OF A BLEEDING PREDICTION MODEL

To develop the bleeding prediction model, statistical methods were employed similar to those used for other prediction scores.[17–19]. The association of major bleeding outcome with predictor variables were assessed with univariate Cox regression analysis, retaining all variables with a significance threshold of less than 0.30. A stepwise selection approach was then used with the remaining candidate variables to determine the final variables for the risk prediction score, optimizing the Aikike Information Criterion (AIC) and using a 0.10 significance level. Individuals without events were censored at the time of last anticoagulation dose. Variables with likely high-predictive ability based on clinical rationale but low event counts in the development cohort (bleeding history and liver disease) were retained in the final model based on clinical validity, similar to prior approaches for developing other risk scores.[10] The competing risk of mortality and thrombotic events were accounted for using the Fine-Grey approach.[20]

The initial model was a Cox proportional hazard model, fit for the entire study period of RE-LY. Model discrimination was evaluated with the Harrell's C statistic, determining this estimate for the overall-study period and at one-year. One-year was chosen for ease of clinical interpretation. For the one-year analyses, individuals were censored at this time of follow-up. Calibration was evaluated by visual examination of calibration plots and by the use of the corrected Nam and D'Agostino goodness-of-fit-test.[21] Internal bias-correction

was performed (for internal validation) by bootstrapping, which involved resampling with replacement of 200 iterations of the development cohort. The C statistic was penalized for the optimism from the bootstrapped resamples.[22] As missing data was <5% per variable in the development cohort, no correction techniques were applied in the development cohort.

**DEVELOPMENT OF A SIMPLIFIED CLINICAL RISK PREDICTON TOOL**—To facilitate clinical implementation, the final multivariable model was converted into a simplified points-based scoring system, termed the DOAC Score. Points were assigned proportional to model coefficients.[18] The highest scoring individuals were combined into one maximum score category, such that the highest scoring category had at least 50 major bleeding events. This was done to prevent overestimation of risk in the highest risk group. [19] A Cox proportional hazard model was then fit with the DOAC Score. Patients were then censored at one-year or time-to-first event, whichever occurred first. Discrimination and calibration were assessed for the points-based model at one-year and at the end of follow-up. Performance of the DOAC Score was compared to the HAS-BLED score, with statistical differences determined using DeLong’s method.[23] For the labile INR variable, which is not applicable to DOACs, individuals were assumed to not have labile INRs as dosing adjustments are not necessary on DOACs. Performance of the algorithm was also assessed in the low-dose DOAC arm.

Individuals’ scores were categorized into clinical risk groups based on 1-year major bleeding rates in RE-LY, to facilitate clinical interpretation. Rates were determined by Kaplan-Meier estimates. Risk groups included: very low risk (<1.00% major bleeding per year), low risk (1.00%-1.99%), moderate risk (2%-4.99% per year), high risk (5.00%-10% per year), and very high risk ( 10% per year).

**MODEL REFINEMENT**—The model was refined in GARFIELD-AF. GARFIELD-AF is a multi-national registry of adults with recently diagnosed AF and at least one risk factor for stroke (N=34,903).[13] GARFIELD-AF was used for refinement due its non-clinical trial population, broad inclusion criteria, enrollment of individuals from many countries, stringent quality checks for outcome ascertainment, and large cohort size. Individuals were included if they were on any DOAC, irrespective of dose. Recruitment occurred from 2013 to 2016, with follow-up through 2019 (GARFIELD-AF cohort 3-5). Individuals with missing follow-up information were excluded. For any missing data, multiple imputation was applied. The final cohort consisted of N=12,296. The primary outcome for validation was major bleeding, defined similarly to RE-LY (per ISTH criteria, see Supplemental Methods in the Supplement). Life-threatening bleeding could not be assessed in this cohort.

The DOAC Score’s predictive performance for the primary outcome was evaluated in GARFIELD-AF. The variables with poor association for the primary outcome were identified. These variables were then modified based on evidence review or excluded. Then a Cox regression model was re-fit to the development cohort RE-LY and coefficients were re-estimated. Performance of the updated DOAC Score was then compared with the HAS-BLED score in GARFIELD-AF using DeLong’s method.[23]

**EXTERNAL VALIDATION**—External validation was conducted in two different cohorts: the COMBINE-AF clinical trials, and the RAMQ administrative database. Different external cohorts were used to capture different populations and maximize generalizability of the risk score.

COMBINE-AF has previously been described.[24] In brief, this cohort consists of five randomized control trials (RE-LY, ROCKET AF [Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared With Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in AF][25], ARISTOTLE [Apixaban for Reduction in Stroke and Other Thromboembolic Events in AF][26], ENGAGE AF-TIMI 48 [Effective Anticoagulation With Factor Xa Next Generation in AF–Thrombolysis in Myocardial Infarction 4][27], and AVERROES [Apixaban Versus Acetylsalicylic Acid [ASA] to Prevent Stroke in AF Patients Who Have Failed or Are Unsuitable for Vitamin K Antagonist Treatment][28]) (N=77,282). ROCKET AF, ARISTOTLE, and ENGAGE AF-TIMI 48 were clinical trials comparing DOACs (rivaroxaban, apixaban, edoxaban respectively) with warfarin in patients with AF. AVERROES was a clinical trial comparing apixaban to aspirin in patients unsuitable for vitamin K antagonist therapy.

RE-LY was used as the development cohort and was thus excluded from COMBINE-AF for external validation. Only adults on DOAC therapy with a factor Xa inhibitor were included (apixaban, rivaroxaban, or edoxaban). Adults on apixaban and rivaroxaban meeting eligibility for dose-reduced anticoagulation were dose reduced. Among adults who were enrolled in ENGAGE-AF-TIMI 48 in COMBINE-AF, only those in the high-dose edoxaban group were included, based on regulatory approval of this dosing scheme.[29] The high-dose edoxaban group consisted of 60mg as standard dosing, with dosing halved if patients had an estimated creatinine clearance of 30-50 ml/min, a body weight of 60kg or less, or use of verapamil or quinidine. The final cohort consisted of N=25,586. One-year performance of the DOAC Score was then determined and compared performance to the HAS-BLED score. The primary outcome included major bleeding defined similarly to RE-LY (per ISTH criteria, see Supplemental Methods in the Supplement). As COMBINE-AF recorded fatal bleeding instead of life-threatening bleeding, this outcome was used as the secondary outcome in this cohort. Fatal bleeding was defined in Supplemental Methods in the Supplement. The DOAC Score and the HAS-BLED score were compared with DeLong p-values.

The second validation was conducted in the RAMQ administrative database. This cohort has previously been described.[30] In brief, RAMQ includes data on patients receiving care in Canada. Adults were identified by hospitalization between January 1, 2011 to December 31, 2017. Individuals with a primary or secondary diagnosis of AF were identified using ICD-9 (427.3, 427.31 or 427.32) and ICD-10 (I-48) codes at time of hospital admission (2011-2017). Patients who then filled at least one prescription for rivaroxaban 20mg daily or apixaban 5mg twice per day were included, and then longitudinally followed after hospital stay. The final cohort consisted of N=11,945 individuals. Definitions for predictors were either directly recorded in the database or for those not directly recorded, were based on ICD-9 and ICD-10 codes (in Supplemental Methods in the Supplement). The primary outcome was major bleeding and defined using ICD-9 and ICD-10 codes (Supplemental



Methods in the Supplement). Life-threatening bleeding could not be assessed. The DOAC Score and HAS-BLED were compared using pairwise differences of the C statistic at one-year.[31]

Analyses adhered strictly to TRIPOD recommendations for development and validation of clinical prediction models (Supplement).[32] The study was determined to be not human subjects research by the Beth Israel Deaconess Medical Center institutional review board due to use of deidentified data. RE-LY, GARFIELD-AF, and COMBINE-AF were performed in accordance with local data protection regulations that were in place at the time of study conduct and all study participants provided written informed consent. RAMQ was approved for informed consent waiver. Model development and internal validation was done at the Smith Center for Outcomes Research in Cardiology, Beth Israel Deaconess Medical Center in Boston, MA, using R version 3.4.1 and the rms, survival, and MASS packages (RE-LY).[33–36] Data for RE-LY was obtained from the secure research sharing platform Vivli.[37] Model refinement was conducted at the Thrombosis Research Institute in London, United Kingdom, using SAS 9.4 (SAS Institute, Cary, NC) (GARFIELD-AF).[38] External validation was conducted at the Thrombolysis in Myocardial Infarction (TIMI) Study Group, Brigham and Women’s Hospital in Boston, MA using SAS 9.4 (SAS Institute, Cary, NC) (COMBINE-AF), and at the University of Montreal at Montreal, Quebec, Canada using SAS 9.4 (SAS Institute, Cary, NC) (RAMQ). [38]

## RESULTS

### MODEL DEVELOPMENT POPULATION

Of the 5684 individuals included in the development cohort (RE-LY), 386 (6.8%) experienced a major bleeding event during follow-up. Median follow-up time was 1.74 years (range 0.01 to 3.06 years). Individuals who experienced a major bleeding event were on average older (75.8 years vs 71.1 years), more likely to take aspirin (64.2% vs 44.1%) or dual antiplatelet therapy (9.3% vs 4.4%) along with a DOAC, and more likely to have diabetes (83.9% vs 78.5%). Mean creatinine clearance was lower in those who experienced a major bleeding event (63.6 mL/min vs 73.6 mL/min). Complete characteristics of individuals with and without a major bleeding event are presented in Table 1.

GARFIELD-AF, the refinement cohort, consisted of N=12,296. 4847 patients were prescribed rivaroxaban (39.4%), 3567 apixaban (29.0%), 2435 dabigatran (19.8%), 339 edoxaban (2.8%), and 1108 were on an unknown DOAC (9.0%). 131 (1.1%) individuals experienced a major bleeding outcome. Complete baseline characteristics are presented in Supplemental Table 9.

### BLEEDING RISK PREDICTION MODEL

Stepwise selection resulted in 11 final predictors, including age, creatinine clearance/ glomerular filtration rate, body mass index, smoking history, stroke/transient ischemic attack/embolism history, diabetes, hypertension, antiplatelet use, and non-steroidal anti-inflammatory use, in addition to bleeding history and liver disease which were required to be retained in the model.

After model refinement was conducted in GARFIELD-AF, smoking history was excluded from the model for poor association with major bleeding and body mass index was reclassified into the binary categories of underweight and not underweight (Supplemental Table 2).

Final model coefficients are presented in Supplemental Table 3. The final model had moderate discrimination in RE-LY over the study period (C statistic: 0.75, [95% CI: 0.72-0.75]) and at one-year (C statistic = 0.75 [95% CI: 0.73-0.77]). The model was well-calibrated based on visual inspection of calibration plots (Supplemental Figure 1–2; goodness-of-fit P = 0.57). After internal validation with bootstrapping, discrimination remained similar (C statistic = 0.73 [95% CI: 0.70-0.75]).

### CLINICAL RISK PREDICTION TOOL – THE DOAC SCORE

Points were assigned to each variable in the model proportional to the model coefficients, with a maximum allocation of five points per risk factor (Supplemental Table 3). Point values included age (2 points: 65-69 years, 3 points 70-74 years, 4 points 75-79 years, 5 points 80 years), creatinine clearance/glomerular filtration rate (1 point: 30-59 mL/min, 2 points: 30 mL/min), underweight (1 point), stroke/transient ischemic attack/embolism history (1 point), diabetes (1 point), hypertension (1 point), antiplatelet use (2 point: aspirin, 3 points: aspirin + P2Y12 inhibitor), and non-steroidal anti-inflammatory use (1 point). Due to low major bleeding event counts for individuals with liver disease, and bleeding history, and thus high likelihood for imprecision of coefficients from the small event counts, these characteristics were assigned points of 2 and 3 respectively to be consistent with point assignments of other risk factors.[6,10,14] Individuals were assigned a maximum of 10 points to allow for at least fifty events in each points assignment (this was done to prevent over-prediction in the high scoring groups as detailed in the methods) (Table 2).

Discrimination of the DOAC Score at 1-year in RE-LY was similar to the full model (C statistic: 0.72 [95% CI: 0.71-0.74]) and was superior to that of the HAS-BLED score (C statistic: 0.60 [95% CI: 0.58-0.62], P for Difference <0.001) (Figure 1A, Table 3). Each additional point scored by the clinical prediction tool was associated with a 48.7% (95% CI: 38.9%-59.3%) increased risk of major bleeding (P<0.001) (Supplemental Table 4, Figure 2).

Discrimination of the DOAC Score in GARFIELD-AF for major bleeding prior to refinement was modest (C Statistic: 0.69 [95% CI: 0.65-0.73]) at one year, as shown in Supplemental Table 2. After model updating, discrimination improved (C Statistic: 0.71 [95% CI: 0.67-0.75]) and the final DOAC Score's performance was superior to that observed for the HAS-BLED score (C statistic: 0.66 [95% CI: 0.62-0.71], P for difference = 0.025) at one year (Figure 1B, Table 3). Major bleeding event rates in GARFIELD-AF were overall lower (Table 4) than those observed in RE-LY.

In secondary analyses, the DOAC Score had similar discrimination (C statistic 0.74 [95% CI: 0.72-0.77]) for life-threatening bleeding in the RE-LY trial at one-year, and had superior performance to the HAS-BLED score (C statistic: 0.61 [95% CI: 0.59-0.61], P for Difference <0.001) (Supplemental Table 5, Supplemental Figure 3). Each additional point



was associated with 55.4% (95% CI: 39.1%-73.5%) increased risk of life-threatening bleed ( $P<0.001$ ).

In the low-dose DOAC population of RE-LY (dabigatran 110mg twice per day), the risk score had similar performance to the full dose group for major bleeding at one-year (C statistic: 0.71 [95% CI: 0.70-0.73]), and was similarly superior in prediction to the HAS-BLED score (C statistic: 0.61 [95% CI: 0.59-0.63], P for Difference  $<0.001$ ).

## CLINICAL RISK GROUPS

Risk scores determined by the DOAC Score were categorized into clinical risk groups based on observed 1-year major bleeding rates (Supplemental Table 4). For each higher clinical risk group, there was an incremental increase in risk of major bleeding (HR: 2.02 [95% CI: 1.84-2.22],  $P<0.001$ ) and an increase in risk of life-threatening bleeding (HR 2.16 [95% CI: 1.87-2.50],  $p<0.001$ ) (Figure 1, Supplemental Table 3, Supplemental Table 5).

## EXTERNAL VALIDATION:

25,586 participants from COMBINE-AF were included, of which 11,712 (45.8%) were prescribed apixaban, 6,819 (26.7%) were prescribed edoxaban, and 7,055 (27.6%) were prescribed (rivaroxaban). Baseline characteristics of this cohort are presented in Supplemental Table 10. 692 (2.7%) individuals experienced a major bleeding event. 99 individuals experienced an intracranial bleeding event (0.4%) (Supplemental Table 12). The DOAC Score had superior prediction of major bleeding risk (C Statistic: 0.67 [95% CI: 0.64-0.69]) than the HAS-BLED score (C Statistic: 0.63 [0.61-0.65], P for Difference  $<0.001$ ) (Figure 2A). Predictive ability of the DOAC Score for fatal bleeding is shown in Supplemental Table 13. Predictive ability of the DOAC Score among the different trials in COMBINE-AF is also demonstrated in Supplemental Table 13. Individuals in the very low risk group (scores 0-3) had the lowest bleeding rates (1.5% per year) compared to those in very high risk group (score 10) (7.7% per year).

The second validation cohort, RAMQ, consisted of 11,945 participants. 4876 participants were prescribed rivaroxaban (40.8%) and 7069 (59.2%) were prescribed apixaban. 258 (2.2%) experienced a major bleeding event. 49 (0.4%) experienced an intracranial bleeding event. Full baseline characteristics are shown in Supplemental Table 11. The DOAC score had superior major bleeding risk prediction (C Statistic: 0.65 [95% CI: 0.61-0.68]) than the HAS-BLED score (C Statistic: 0.58 [0.55-0.62], P for Difference  $<0.001$ ) (Figure 2B). Individuals in the very low risk group (scores 0-3) had the lowest bleeding rates (0.6% per year) compared to those in very high risk group (score 10) (3.7% per year).

## DISCUSSION

In this study, the DOAC Score was developed, a novel bleeding risk score to stratify bleeding risk among patients with AF who are prescribed a DOAC. Unlike many commonly used bleeding prediction tools, the DOAC risk prediction score was developed exclusively for individuals on a variety of DOACs, the most common form of anticoagulation for patients with AF. The prediction tool had moderate discrimination in the development cohort, refinement cohort, and in two large-scale validation cohorts. The DOAC Score

consistently outperformed the HAS-BLED score in every cohort and was able to stratify patients by levels of bleeding risk across both randomized trials and observational populations (Figure 1–2).

The increased predictive performance of the DOAC Score is likely due to multiple factors. First, the risk score distinguishes differences in bleeding risk among more age groups than prior scores, allowing for improved estimates of personalized risk.[10] Second, the scoring system accounts for variation in bleeding risk among different levels of renal function, a strong risk factor for bleeding.[39–41] Third, many individuals with AF are on multiple medications associated with a high-risk for bleeding, and the DOAC Score accounts for the higher cumulative risk for those on multiple medications, as well as differences in risk conferred by different combinations of therapy. Prior risk scores often considered those on any combination of aspirin, dual-antiplatelet therapy, and a nonsteroidal anti-inflammatory agents as similar risk—though evidence strongly suggests that the risk from multiple agents is higher.[10,42] Fourth, the risk prediction tool assigns risk differently for each risk factor, in contrast to several prior risk scores which assign equivalent risk to each risk factor despite these risk factors having very different magnitudes of association with bleeding.[10,43]

While the decision to treat patients with AF with anticoagulation requires considering multiple clinical factors, the DOAC score could inform shared-decision making conversations for anticoagulation. By quantifying bleeding risk, physicians can better communicate the risks and benefits of anticoagulation, especially as non-pharmacological options for embolic stroke/systemic emboli prevention (left atrial appendage closures) become available.[44] The DOAC score could be easily implemented in clinical settings, as it contains has a similar number of inputs to prior bleeding risk scoring systems, and only includes commonly obtained clinical risk factors.[10,43] Further, electronic medical record implementation could further automate risk classification with the tool and provide an easier avenue to inform clinical decisions.

Higher bleeding rates were observed in the RELY and COMBINE-AF populations as compared with those in the GARFIELD-AF and RAMQ registries. These differences may reflect higher sensitivity in ascertaining bleeding endpoints in prospective randomized pivotal trials compared with observational cohorts. GARFIELD-AF also included a meaningful proportion of individuals who had reduced dosing of anticoagulation, [45] though this may reflect real-world prescribing patterns. RAMQ individuals may have had lower bleeding rates due to the limitations of billing code outcome ascertainment. Using both types of data in each phase of risk score creation (model development and external validation) allowed us to leverage the unique strengths of each.

This study has several limitations. First, the risk score was primarily validated for major bleeding outcomes, but decisions to treat patients with anticoagulation often require weighing the severity of different types of bleeding. The secondary validation in life-threatening or fatal bleeding was performed to help inform prediction of the DOAC Score in bleeding outcomes of higher severity. We were unable to evaluate the DOAC Score's ability to predict intracranial bleeding events due to low outcome rate. We do report event counts by DOAC risk category in COMBINE-AF, the cohort with the largest count of intracranial

bleeding events, to help inform this limitation (Supplemental Table 12). Second, validation in life-threatening or fatal bleeding was limited by low event counts, though we did observe that the DOAC Score had similar predictive ability for these outcomes as for major bleeding. Third, while the DOAC Score had superior prediction to the HAS-BLED score, discrimination in the validation cohorts was modest (C statistic range 0.65-0.67). Fourth, when comparing the HAS-BLED score, patients on DOACs were assumed not to have labile INRs as DOACs do not require regular dosage monitoring. Fifth, the DOAC Score was only compared to the HAS-BLED score and not other bleeding risk scores [10,11]. HAS-BLED was preferentially chosen, however, because it is the most used bleeding risk score.

In conclusion, the DOAC Score was developed, a tool for assessing personalized bleeding risk in patients with AF who are prescribed a DOAC. The DOAC Score had stronger predictive performance than HAS-BLED, and was able to predict bleeding in separate, large multinational cohorts—supporting generalizability of the risk score.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements:

This publication is based on research using data from data contributor Boehringer Ingelheim that has been made available through Vivli, Inc. Vivli has not contributed to or approved, and is not in any way responsible for, the contents of this publication. We would like to thank the RE-LY investigators for sharing individual patient level data from the trial, as well as Vivli for providing a platform to securely obtain and analyze this data.

We would also like to thank Robert P. Giugliano, MD, SM and Sabina Murphy, MPH for assistance in accessing and using the COMBINE-AF cohort for validation.

## Funding:

Rahul Aggarwal was supported by the Weener Family Grant, which was not involved in the conduction of this study but supports trainees for conducting original research of their choice. Daniel E. Singer was supported in part by the Eliot B. & Edith C. Shoolman Fund for Research of Cerebrovascular Disease, Massachusetts General Hospital, Boston, MA.

## Disclosures:

Dr. Christian Ruff is a member of the TIMI Study Group, which has received institutional research grant support through Brigham and Women's Hospital from: Abbott, Amgen, Anthos Therapeutics, ARCA Biopharma, Inc., AstraZeneca, Bayer HealthCare Pharmaceuticals, Inc., Daiichi-Sankyo, Eisai, Intarcia, Ionis Pharmaceuticals, Inc., Janssen Research and Development, LLC, MedImmune, Merck, Novartis, Pfizer, Quark Pharmaceuticals, Regeneron Pharmaceuticals, Inc., Roche, Siemens Healthcare Diagnostics, Inc., Softcell Medical Limited, The Medicines Company, Zora Biosciences. He endorses research grants through institution: Anthos, AstraZeneca, Daiichi Sankyo, Janssen and Novartis; Honoraria for scientific advisory boards and consulting: Anthos, Bayer, Bristol Myers Squibb, Daiichi Sankyo, Janssen, Pfizer.

Professor Ajay K. Kakkar reports research grants from Bayer Pharma AG and Sanofi. Personal fees from Anthos Therapeutics, Bayer Pharma AG, Sanofi S.A.

Mr. Michael Palazzolo is a member of the TIMI Study Group which has received institutional research grant support through Brigham and Woman's Hospital from: Abbott, Amgen, Anthos Therapeutics, ARCA Biopharma, Inc., AstraZeneca, Daiichi-Sankyo, Ionis Pharmaceuticals, Inc., Janssen Research and Development, LLC, MedImmune, Merck, Novartis, Pfizer, Regeneron Pharmaceuticals, Inc., Roche, Softcell Medical Limited, The Medicines Company, Zora Biosciences

Dr. Daniel E. Singer has research grants from Bristol Myers Squibb and consulting agreements with Bristol Myers Squibb, Fitbit, Medtronic, and Pfizer.

Dr. Eric A Secemsky reports the following research grants to Beth Israel Deaconess Medical Center: NIH/NHLBI K23HL150290, Food & Drug Administration, BD, Boston Scientific, Cook, CSI, Laminar Medical, Medtronic and Philips. He also endorses consulting/speaking with Abbott, Bayer, BD, Boston Scientific, Cook, Cordis, CSI, Inari, Medtronic, Philips, Shockwave and VentureMed.

Dr. Jonathan P. Piccini receives grants for clinical research from Abbott, American Heart Association, Association for the Advancement of Medical Instrumentation, Bayer, Boston Scientific, and Philips; and serves as a consultant to Abbott, Abbvie, Abilene, Altathera, Biotronik, Boston Scientific, Bristol Myers Squibb, LivaNova, Medtronic, Milestone, ElectroPhysiology Frontiers, Pfizer, Sanofi, Philips, and Up-to-Date.

Dr. Changyu Shen reports that he is an employee of Biogen and owns Biogen stocks.

Dr. Robert W. Yeh reports grants and consulting fees from Abbott Vascular, Boston Scientific and Medtronic

## Non-standard Abbreviations and Acronyms:

<b>(AF)</b>	Atrial Fibrillation
<b>(DOAC)</b>	Direct-acting Oral Anticoagulants
<b>RE-LY</b>	Randomized Evaluation of Long-Term Anticoagulation Therapy
<b>GARFIELD-AF</b>	Global Anticoagulant Registry in the Field-Atrial Fibrillation
<b>COMBINE-AF</b>	A Collaboration between Multiple Institutions to Better Investigate Non-vitamin K Antagonist Oral Anticoagulant use in Atrial Fibrillation
<b>RAMQ</b>	Quebec Régie de l'Assurance Maladie du Québec (RAMQ) and Med-Echo Administrative Databases

## References:

1. January Craig T, Samuel Wann L., Hugh Calkins, et al. 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society in Collaboration With the Society of Thoracic Surgeons. *Circulation* 2019;140:e125–51. doi:10.1161/CIR.0000000000000665 [PubMed: 30686041]
2. Singer DE, Chang Y, Fang MC, et al. The net clinical benefit of warfarin anticoagulation in atrial fibrillation. *Ann Intern Med* 2009;151:297–305. doi:10.7326/0003-4819-151-5-200909010-00003 [PubMed: 19721017]
3. Lip GYH, Nieuwlaat R, Pisters R, et al. Refining Clinical Risk Stratification for Predicting Stroke and Thromboembolism in Atrial Fibrillation Using a Novel Risk Factor-Based Approach: The Euro Heart Survey on Atrial Fibrillation. *CHEST* 2010;137:263–72. doi:10.1378/chest.09-1584 [PubMed: 19762550]
4. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS) | *European Heart Journal* | Oxford Academic. 10.1093/eurheartj/ehaa612/5899003 (accessed 3 Nov 2020).
5. 1 Recommendations | Atrial fibrillation: management | Guidance | NICE. <https://www.nice.org.uk/guidance/cg180/chapter/1-Recommendations> (accessed 18 Aug 2020).

6. Apostolakis S, Lane DA, Guo Y, et al. Performance of the HEMORR(2)HAGES, ATRIA, and HAS-BLED bleeding risk-prediction scores in patients with atrial fibrillation undergoing anticoagulation: the AMADEUS (evaluating the use of SR34006 compared to warfarin or acenocoumarol in patients with atrial fibrillation) study. *J Am Coll Cardiol* 2012;60:861–7. doi:10.1016/j.jacc.2012.06.019 [PubMed: 22858389]
7. Dan G-A, Iliodromitis K, Scherr D, et al. Translating guidelines into practice for the management of atrial fibrillation: results of an European Heart Rhythm Association Survey. *Eur Eur Pacing Arrhythm Card Electrophysiol J Work Groups Card Pacing Arrhythm Card Cell Electrophysiol Eur Soc Cardiol* 2018;20:1382–7. doi:10.1093/europace/euy094
8. Qiu J, Grine K. Assessing Bleeding Risk in Patients Taking Anticoagulants. *Am Fam Physician* 2017;96:465–6. [PubMed: 29094904]
9. Piccini JP, Singer DE. Putting risk prediction in atrial fibrillation into perspective. *Eur Heart J* 2012;33:1431–3. doi:10.1093/eurheartj/ehs031 [PubMed: 22378796]
10. A Novel User-Friendly Score (HAS-BLED) To Assess 1-Year Risk of Major Bleeding in Patients With Atrial Fibrillation - CHEST [https://journal.chestnet.org/article/S0012-3692\(10\)60585-5/abstract](https://journal.chestnet.org/article/S0012-3692(10)60585-5/abstract) (accessed 19 Aug 2020).
11. Borre ED, Goode A, Raitz G, et al. Predicting Thromboembolic and Bleeding Event Risk in Patients with Non-Valvular Atrial Fibrillation: A Systematic Review. *Thromb Haemost* 2018;118:2171–87. doi:10.1055/s-0038-1675400 [PubMed: 30376678]
12. Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus Warfarin in Patients with Atrial Fibrillation. *N Engl J Med* 2009;361:1139–51. doi:10.1056/NEJMoa0905561 [PubMed: 19717844]
13. Kakkar AK, Mueller I, Bassand J-P, et al. International longitudinal registry of patients with atrial fibrillation at risk of stroke: Global Anticoagulant Registry in the FIELD (GARFIELD). *Am Heart J* 2012;163:13–19.e1. doi:10.1016/j.ahj.2011.09.011 [PubMed: 22172431]
14. Decousus H, Tapson VF, Bergmann J-F, et al. Factors at admission associated with bleeding risk in medical patients: findings from the IMPROVE investigators. *Chest* 2011;139:69–79. doi:10.1378/chest.09-3081 [PubMed: 20453069]
15. Chen TK, Knicely DH, Grams ME. Chronic Kidney Disease Diagnosis and Management. *JAMA* 2019;322:1294–304. doi:10.1001/jama.2019.14745 [PubMed: 31573641]
16. Schulman S, Kearon C. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. *J Thromb Haemost* 2005;3:692–4. doi:10.1111/j.1538-7836.2005.01204.x [PubMed: 15842354]
17. 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;315:1735–49. doi:10.1001/jama.2016.3775 [PubMed: 27022822]
18. Wilson Peter WF, D'Agostino Ralph B, Daniel Levy, et al. Prediction of Coronary Heart Disease Using Risk Factor Categories. *Circulation* 1998;97:1837–47. doi:10.1161/01.CIR.97.18.1837 [PubMed: 9603539]
19. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med* 2004;23:1631–60. doi:10.1002/sim.1742 [PubMed: 15122742]
20. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc* 1999;94:496–509. doi:10.1080/01621459.1999.10474144
21. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. *Stat Med* 2015;34:1659–80. doi:10.1002/sim.6428 [PubMed: 25684707]
22. Steyerberg EW, Harrell FE, Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81. doi:10.1016/s0895-4356(01)00341-9 [PubMed: 11470385]
23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45. [PubMed: 3203132]
24. Carnicelli AP, Hong H, Giugliano RP, et al. Individual Patient Data from the Pivotal Randomized Controlled Trials of Non-Vitamin K Antagonist Oral Anticoagulants in Patients

- with Atrial Fibrillation (COMBINE AF): Design and Rationale: From the COMBINE AF (A Collaboration between Multiple institutions to Better Investigate Non-vitamin K antagonist oral anticoagulant use in Atrial Fibrillation) Investigators. *Am Heart J* 2021;233:48–58. doi:10.1016/j.ahj.2020.12.002 [PubMed: 33296688]
25. Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus Warfarin in Nonvalvular Atrial Fibrillation. *N Engl J Med* 2011;365:883–91. doi:10.1056/NEJMoa1009638 [PubMed: 21830957]
  26. Granger CB, Alexander JH, McMurray JJV, et al. Apixaban versus Warfarin in Patients with Atrial Fibrillation. *N Engl J Med* 2011;365:981–92. doi:10.1056/NEJMoa1107039 [PubMed: 21870978]
  27. Giugliano RP, Ruff CT, Braunwald E, et al. Edoxaban versus Warfarin in Patients with Atrial Fibrillation. *N Engl J Med* 2013;369:2093–104. doi:10.1056/NEJMoa1310907 [PubMed: 24251359]
  28. Connolly SJ, Eikelboom J, Joyner C, et al. Apixaban in Patients with Atrial Fibrillation. *N Engl J Med* 2011;364:806–17. doi:10.1056/NEJMoa1007432 [PubMed: 21309657]
  29. U.S. Food & Drug Administration. SAVAYSA (edoxaban) tablets for oral use. 2015. [https://www.accessdata.fda.gov/drugsatfda\\_docs/label/2015/206316lbl.pdf](https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/206316lbl.pdf)
  30. Qazi JZ, Schnitzer ME, Côté R, et al. Predicting major bleeding among hospitalized patients using oral anticoagulants for atrial fibrillation after discharge. *PloS One* 2021;16:e0246691. doi:10.1371/journal.pone.0246691 [PubMed: 33657116]
  31. Guo C, Ying S, Woosun J. Evaluating Predictive Accuracy of Survival Models with PROC PHREG. *SAS Inst;SAS462–2017*.
  32. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *The TRIPOD Group. Circulation* 2015;131:211–9. doi:10.1161/CIRCULATIONAHA.114.014508 [PubMed: 25561516]
  33. R Core Team. R: A Language and Environment for Statistical Computing. 2023. <https://www.R-project.org/>
  34. Harrell FE. Regression Modeling Strategies (rms) Package.
  35. Therneau TM, Lumley T, Elizabeth A, et al. Survival Analysis: survival.
  36. Ripley B, Benables B, Bates D, et al. MASS.
  37. Vivli Center For Global Clinical Research Data. <https://vivli.org/>
  38. SAS. [www.SAS.com](http://www.SAS.com)
  39. Ocak G, Rookmaaker MB, Algra A, et al. Chronic kidney disease and bleeding risk in patients at high cardiovascular risk: a cohort study. *J Thromb Haemost* 2018;16:65–73. doi:10.1111/jth.13904 [PubMed: 29125709]
  40. Xiaoxi Yao, Shah Nilay D., Sangaralingham Lindsey R., et al. Non-Vitamin K Antagonist Oral Anticoagulant Dosing in Patients With Atrial Fibrillation and Renal Dysfunction. *J Am Coll Cardiol* 2017;69:2779–90. doi:10.1016/j.jacc.2017.03.600 [PubMed: 28595692]
  41. Fang MC, Go AS, Chang Y, et al. A new risk scheme to predict warfarin-associated hemorrhage: The ATRIA (Anticoagulation and Risk Factors in Atrial Fibrillation) Study. *J Am Coll Cardiol* 2011;58:395–401. doi:10.1016/j.jacc.2011.03.031 [PubMed: 21757117]
  42. Cohen M Expanding the Recognition and Assessment of Bleeding Events Associated With Antiplatelet Therapy in Primary Care. *Mayo Clin Proc* 2009;84:149–60. [PubMed: 19181649]
  43. Gage BF, Yan Y, Milligan PE, et al. Clinical classification schemes for predicting hemorrhage: results from the National Registry of Atrial Fibrillation (NRAF). *Am Heart J* 2006;151:713–9. doi:10.1016/j.ahj.2005.04.017 [PubMed: 16504638]
  44. Pavel Osmancik, Dalibor Herman, Petr Neuzil, et al. Left Atrial Appendage Closure Versus Direct Oral Anticoagulants in High-Risk Patients With Atrial Fibrillation. *J Am Coll Cardiol* 2020;75:3122–35. doi:10.1016/j.jacc.2020.04.067 [PubMed: 32586585]
  45. Camm AJ, Cools F, Virdone S, et al. Mortality in Patients With Atrial Fibrillation Receiving Nonrecommended Doses of Direct Oral Anticoagulants. *J Am Coll Cardiol* 2020;76:1425–36. doi:10.1016/j.jacc.2020.07.045 [PubMed: 32943160]



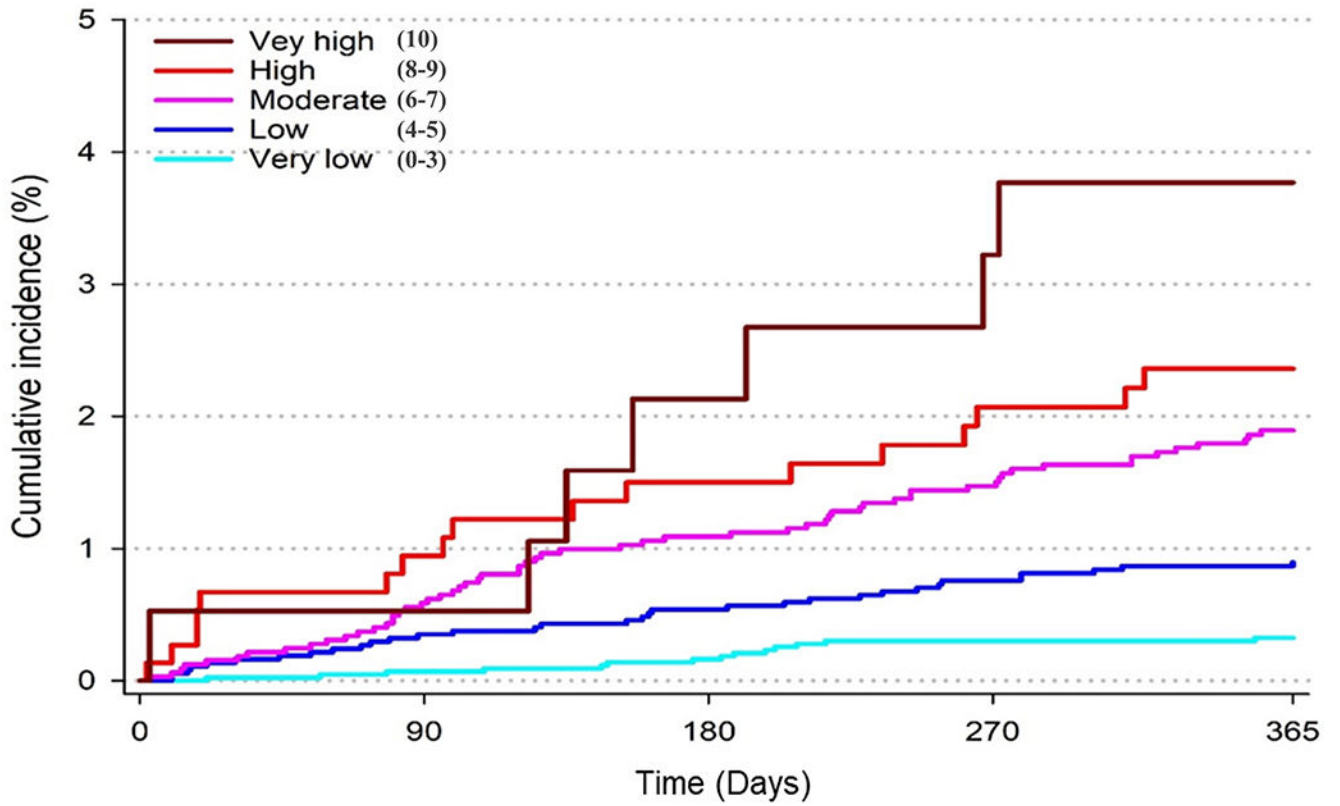
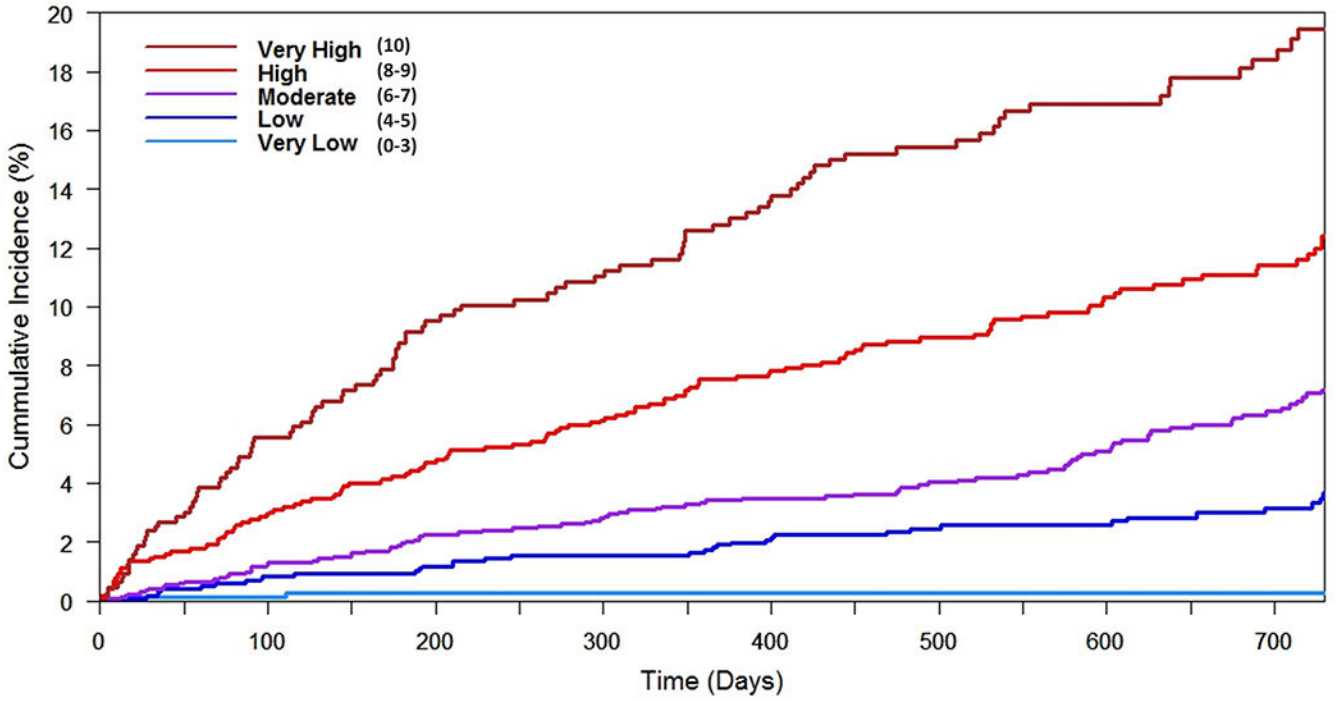
**Clinical Perspective:****What is new?**

- The DOAC Score is a clinical bleeding risk score developed and validated to personalize estimates of bleeding risk for individuals with atrial fibrillation taking direct-acting oral anticoagulants (DOACs).
- The risk score was developed among 5684 patients in the RE-LY trial, further developed in 12,296 individuals in the GARFIELD-AF registry, and then validated among 25,586 individuals in the COMBINE-AF trials as well as 11,945 individuals in the RAMQ administrative database.
- The DOAC Score had moderate discrimination in all cohorts and outperformed the HAS-BLED score.

**What are the clinical implications?**

- The DOAC Score can stratify bleeding risk among patients with atrial fibrillation on DOACs.
- By predicting bleeding risk, clinicians can inform shared decision-making conversations when discussing anticoagulation among patients with atrial fibrillation.

### Major Bleed



**Figure 1. Cumulative Incidence for Bleeding Outcomes by Predicted Risk Category in The Development Cohorts: RE-LY and GARFIELD-AF.**

Presented are the cumulative incidence curves for individuals in the development cohort (RE-LY trial) (1A) and the refinement cohort (GARFIELD-AF registry) (1B). Cumulative incidence curves were based on the risk category assigned to individuals by the DOAC Score. Risk-scores were 0-10, with risk categories assigned as: very low (score 0-3), low (score 4-5), moderate (score 6-7), high (score 8-9) and very high (score 10). Individuals in RE-LY were included if they were in the dabigatran 150 mg twice daily arm of the trial. Individuals in GARFIELD-AF were included if they were on any direct-acting oral anticoagulant (DOAC), irrespective of dose.

A) Cumulative Incidence for Major Bleeding by Risk Category in RE-LY

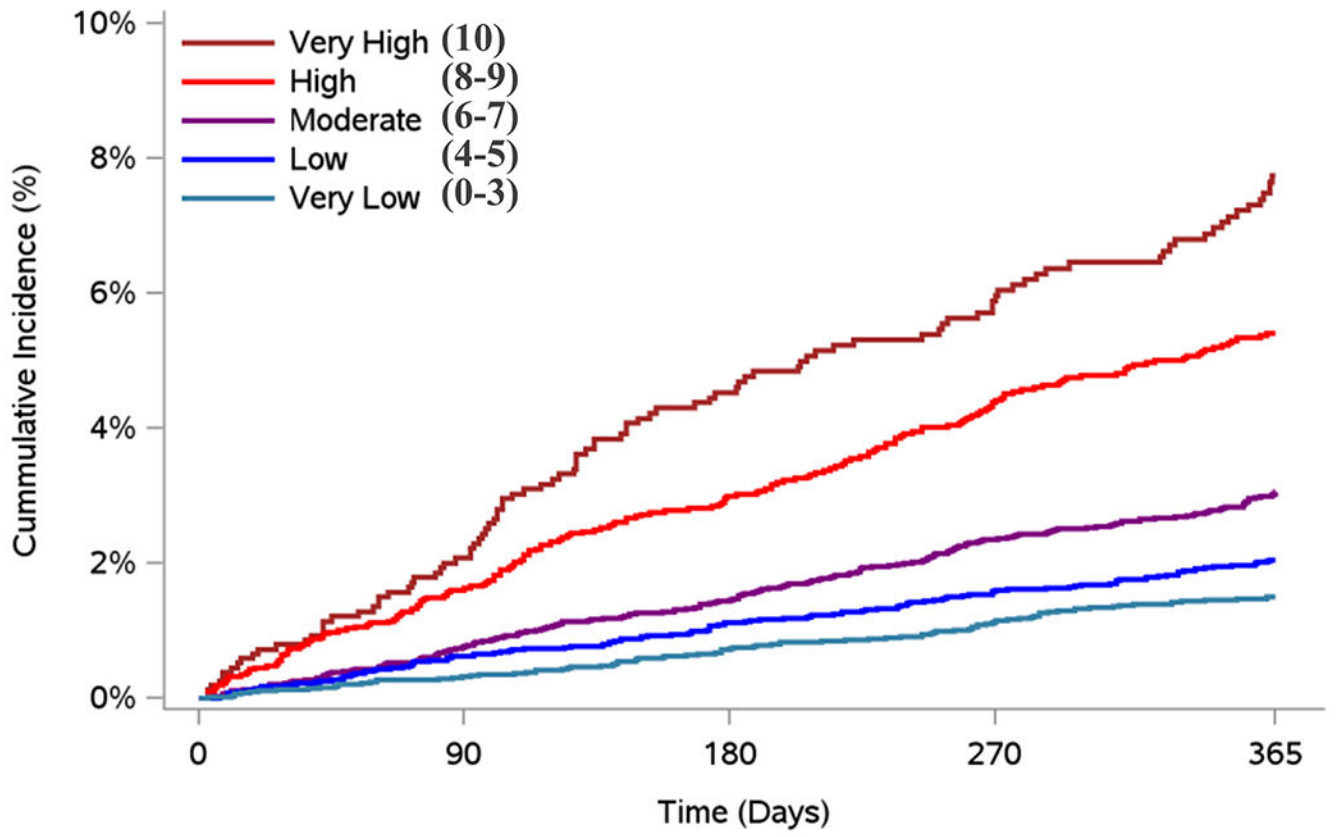
B) Cumulative Incidence for Major Bleeding by Risk Category in GARFIELD-AF

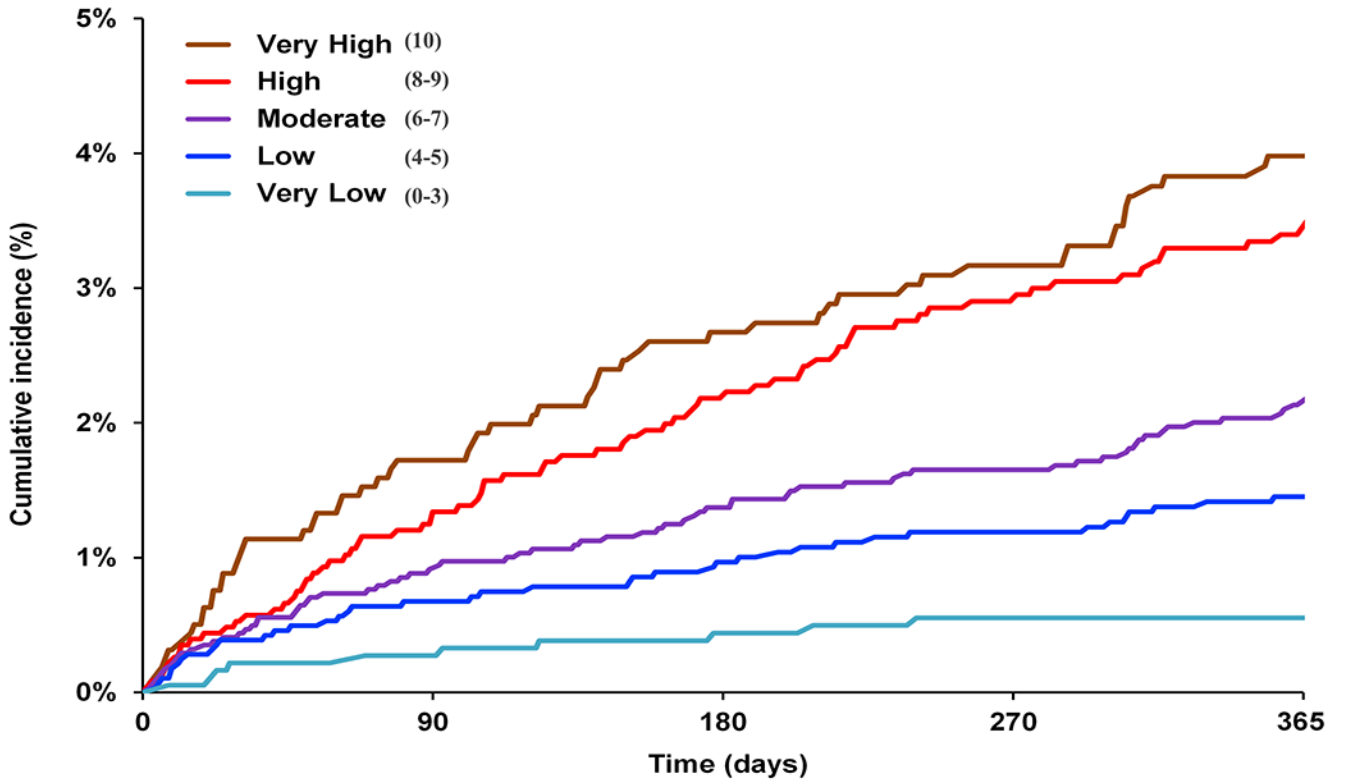
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript





**Figure 2. Cumulative Incidence for Bleeding Outcomes by Predicted Risk Category in the Validation Cohorts: COMBINE-AF and RAMQ.**

Presented are the Cumulative Incidence curves for individuals in the COMBINE-AF clinical trial cohort (2A), and RAMQ administrative database (2B). Cumulative incidence curves were based on the risk category assigned to individuals by the DOAC Score. Risk-scores were 0-10, with risk categories assigned as: very low (score 0-3), low (score 4-5), moderate (score 6-7), high (score 8-9) and very high (score 10). Individuals in COMBINE-AF were included if they were on any direct-acting oral anticoagulant (DOAC). Individuals in RAMQ AF were included if they were on apixaban 5mg twice per day or rivaroxaban 20mg daily.

A) Cumulative Incidence for Major Bleeding by Risk Category in COMBINE-AF  
 B) Cumulative Incidence for Major Bleeding by Risk Category in RAMQ

**Table 1.**  
**Baseline Characteristics of Patients With and Without Bleeding Events in the RE-LY Trial (Development Cohort)**

Presented are baseline characteristics of individuals in the development cohort (RE-LY). Baseline characteristics were stratified by individuals who experienced a major bleeding event and those who did not during the study period. Only individuals in the dabigatran 150 mg twice daily arm of the trial were included. Individuals were followed for the duration of DOAC during the trial.

<b>Major Bleeding Outcome</b>		<b>No (N= 5298)</b>	<b>Yes (N=386)</b>
<b>Mean Age, Years</b>		71.1	75.8
<b>Female, %</b>		36.7%	35.5%
<b>Race/Ethnicity</b>			
	White	69.0%	73.6%
	Black	0.9%	1.6%
	Asian	16.8%	10.9%
	Other	13.3%	14.0%
<b>Mean CHA<sub>2</sub>DS<sub>2</sub>-VASc</b>		3.7	4.3
<b>Mean HAS-BLED</b>		1.7	2.0
<b>Creatinine Clearance, mL/min</b>			
	60	65.3%	48.7%
	30-59	34.2%	49.5%
	<30	0.4%	1.8%
<b>Obesity (Body Mass Index <math>\geq 30</math> kg/m<sup>2</sup>)</b>		34.7%	38.3%
<b>Diabetes</b>		22.5%	32.6%
<b>Hypertension</b>		78.6%	83.9%
<b>Prior Stroke, Transient Ischemic Attack, or Embolism</b>		21.7%	26.7%
<b>Antiplatelet Use</b>			
	Aspirin	44.1%	64.2%
	Dual Antiplatelet	4.4%	9.3%
<b>Nonsteroidal Anti-Inflammatory Drug</b>		64.4%	78.8%
<b>Bleeding History</b>		6.5%	7.8%
<b>Liver Disease</b>		1.2%	0.5%
<b>Smoking History</b>		50.6%	61.1%



**Table 2.****The DOAC Score**

Presented is the final clinical risk prediction scoring system, termed the DOAC Score. Point assignments are based on patient characteristics. The prediction score was developed in the RE-LY trial among participants taking dabigatran 150 mg twice a day and then refined among individuals in the GARFIELD-AF registry among participants on all DOACs. Point assignments were based on the coefficients of these variables in a Cox regression model (presented in the Supplement) for the outcome of major bleeding. The maximum number of allocated points for an individual is 10 points to prevent overestimation of risk in the high-risk groups. Validation was conducted in three large cohorts in individuals on direct-acting oral anticoagulants, including dabigatran, edoxaban, rivaroxaban, and apixaban.

Clinical Risk Prediction Tool	Points
<b>Age, years</b>	
65-69	2
70-74	3
75-79	4
80	5
<b>Creatinine Clearance/eGFR(mL/min)</b>	
30-60	1
<30	2
<b>Underweight (BMI &lt;18.5)</b>	1
<b>Stroke/TIA/Embolism History</b>	1
<b>Diabetes</b>	1
<b>Hypertension</b>	1
<b>Antiplatelet Use</b>	
Aspirin	2
Dual-Antiplatelet	3
<b>Nonsteroidal Anti-inflammatory (NSAID) Use</b>	1
<b>Bleeding History</b>	3
<b>Liver Disease</b>	2
<i>Total Score Range: 0-10 (Maximum 10 points)</i>	

**Table 3.**  
**Predictive Performance of the DOAC Score Compared to HAS-BLED**

Presented is the predictive performance of the DOAC Score compared to HAS-BLED in the development, refinement, and validation cohorts. RE-LY was the development cohort. GARFIELD-AF was used to refine the prediction model. COMBINE-AF and RAMQ were the validation cohorts. C-statistics were compared at one-year. In each cohort, the DOAC Score had stronger predictive performance than the HAS-BLED score.

	N	Number of Events	DOAC Score, C-Statistic	HAS-BLED, C-Statistic	P-value
RE-LY	5,684	386	0.72 (0.71-0.74)	0.60 (0.58-0.62)	<0.001
GARFIELD-AF	12,296	131	0.71 (0.67-0.75)	0.66 (0.62-0.71)	0.025
COMBINE-AF	25,586	692	0.67 (0.64-0.69)	0.63 (0.61-0.65)	<0.001
RAMQ	11,945	258	0.65 (0.61-0.68)	0.58 (0.55-0.62)	<0.001

**Table 4.** **One-Year Bleeding Rates by Risk Category for Individuals in the Development & Validation Cohorts**

Presented are 1-year bleeding rate estimates for individuals in the development and validation cohorts. Event counts are based on one-year censoring. Risk categories were assigned by the DOAC Score based on risk score. Each individual could be assigned a risk score from 0-10, with the following risk categories: very low (score 1-3), low (score 4-5), moderate (score 6-7), high (score 8-9) and very high (score 10).

Risk Category (Risk Score)	RE-LY (N=5684)		GARFIELD-AF (N=12,296)		COMBINE-AF (N=25,586)		RAMQ (N=11,945)	
	(Major Bleeding Events)	1-Year Rate	(Major Bleeding Events)	1-Year Rate	(Major Bleeding Events)	1-Year Rate	(Major Bleeding Events)	1-Year Rate
	N		N		N		N	
<b>Very Low (0-3)</b>	767 (2)	0.8%	4360 (14)	0.3%	6038 (82)	1.5%	1832 (10)	0.6%
<b>Low (4-5)</b>	1249 (21)	1.6%	3735 (33)	0.9%	6630 (123)	2.0%	2834 (40)	1.4%
<b>Moderate (6-7)</b>	1727 (53)	3.4%	3263 (60)	1.9%	7348 (197)	3.1%	3418 (72)	2.1%
<b>High (8-9)</b>	1296 (87)	6.9%	748 (17)	2.4%	4015 (188)	5.4%	2274 (76)	3.3%
<b>Very High (10)</b>	645 (73)	13.9%	190 (7)	3.7%	1555 (102)	7.7%	1587 (60)	3.7%