The HAS-BLED, ATRIA, and ORBIT Bleeding Scores in Atrial Fibrillation Patients Using Non-Vitamin K Antagonist Oral Anticoagulants

Lip, Gregory; Skjøth, Flemming; Nielsen, Peter Brønnum; Kjældgaard, Jette Nordstrøm; Larsen, Torben Bjerregaard

DOI:
10.1016/j.amjmed.2017.11.046

Accepted Manuscript

Title: The HAS-BLED, ATRIA and ORBIT Bleeding Scores in Atrial Fibrillation Patients Using Non-Vitamin K Antagonist Oral Anticoagulants

Author: Gregory Y.H. Lip, Flemming Skjøth, Peter Brønnum Nielsen, Jette Nordstrøm Kjældgaard, Torben Bjerregaard Larsen

PII: S0002-9343(17)31282-2
DOI: https://doi.org/10.1016/j.amjmed.2017.11.046
Reference: AJM 14434

To appear in: *The American Journal of Medicine*

The HAS-BLED, ATRIA and ORBIT bleeding scores in atrial fibrillation patients using non-vitamin K antagonist oral anticoagulants

Gregory Y.H. Lip, MD[1,2*], Flemming Skjøth[1,3*], Peter Brønnum Nielsen[1,4],

Jette Nordstrøm Kjældgaard[3], Torben Bjerregaard Larsen, MD, PhD[1,4]

[1.] Aalborg Thrombosis Research Unit, Department of Clinical Medicine, Faculty of Health, Aalborg University, Aalborg, Denmark.

[2.] Institute of Cardiovascular Sciences, University of Birmingham, United Kingdom

[3.] Unit for Clinical Biostatistics. Aalborg University Hospital, Denmark

[4.] Department of Cardiology, Aalborg University Hospital, Denmark

[*joint first authors]

Correspondence to:

Professor Gregory Y H Lip, MD

Phone +44 121 5075080; g.y.h.lip@bham.ac.uk

Running Heading: Bleeding risk scores in atrial fibrillation

Disclosures

Professor Lip: Consultant for Bayer/Janssen, BMS/Pfizer, Biotronik, Medtronic, Boehringer Ingelheim, Novartis, Verseon and Daiichi-Sankyo. Speaker for Bayer, BMS/Pfizer, Medtronic, Boehringer Ingelheim, and Daiichi-Sankyo. No fees are directly received personally.
Associate Professor Larsen has served as an investigator for Janssen Scientific Affairs, LLC and Boehringer Ingelheim and has served as a speaker for Bayer, BMS/Pfizer, Boehringer Ingelheim, MSD and AstraZeneca. Associate Professor Nielsen has served as a speaker for Boehringer Ingelheim and Bayer. Flemming Skjøth has served as consultant for Bayer.
Dr Kjældgaard – none declared.

Profs Lip and Larsen are guarantors of the paper, taking responsibility for the integrity of the work as a whole.

Highlights

- Limited data for bleeding risk scores are available with non-Vitamin K antagonist OACs(NOACs) in anticoagulated patients with atrial fibrillation(AF).
- In this study contemporary bleeding risk scores(ATRIA, HAS-BLED, ORBIT) showed modest predictive values for major bleeding.
- The HAS-BLED score classified least patients at low risk and achieved the highest clinical usefulness if applying a major bleeding intervention threshold of 2%, whereas benefit from other scores was only evident at higher thresholds.

Abstract

Background: Various bleeding risk scores have been proposed to assess the risk of bleeding in patients with atrial fibrillation (AF) taking oral anticoagulants (OAC). Limited data are available with these scores, in users of non-Vitamin K antagonist OACs (NOACs).

Methods: Using the Danish registries we evaluated and compared the risk classification properties of the HAS-BLED, ATRIA and ORBIT scores for predicting major bleeding in 57,930 atrial fibrillation patients (44.6% female; mean age 73.5 years, SD 11.4; mean CHA2DS2-VASc score 3.2, SD 1.8).

Results. At 1 year follow-up, C-statistics for ATRIA, HAS-BLED and ORBIT were approx. 0.59 with only minor differences between scores. Both ATRIA and ORBIT categorized more patients as 'low risk' (both >83%, when compared to HAS-BLED, only 53%) and qualitatively, the ROC curves revealed higher sensitivity (62.8%) for HAS-BLED compared to ATRIA (29.7%) and ORBIT (37.1%).

The clinical usefulness of scores was evaluated using decision curve analyses at a 1 year perspective. If the intervention threshold is low (<1.7%) the benefit is towards monitoring all patients. If preference is for a major bleeding risk threshold between 1.7-2.0%, most benefit was obtained by using HAS-BLED. ORBIT and ATRIA score provided better benefit for thresholds between 2.0-6.0%.

Conclusion: This analysis of contemporary bleeding risk score stratification in a 'real world' NOAC users population with atrial fibrillation showed modest predictive values using C-statistics. The scores represent different risk thresholds with HAS-BLED classifying least patients at low risk and achieving the highest benefit if applying a major bleeding intervention threshold of approx. 2%,

whereas benefit from using either ATRIA score or ORBIT score was only evident using higher intervention thresholds.

Key words: bleeding, atrial fibrillation, risk stratification

Abbreviations:

AF: Atrial fibrillation.

OAC: Oral anticoagulants.

VKA: Vitamin K antagonists.

NOAC: non-VKA oral anticoagulants

CHA_2DS_2-VASc score: congestive heart failure, hypertension, age ≥75 years [doubled], diabetes, stroke/transient ischemic attack/thromboembolism [doubled], vascular disease [prior myocardial infarction (MI), peripheral arterial disease (PAD), or aortic plaque], age 65-74 years, sex category [female]

HAS-BLED: Hypertension, Age, Stroke, Bleeding tendency/predisposition, Labile INRs, Elderly age/frailty, Drugs such as concomitant aspirin/NSAIDs or alcohol excess

ATRIA: Anticoagulation and Risk Factors in Atrial Fibrillation

ORBIT: Outcomes Registry for Better Informed Treatment of Atrial Fibrillation

Introduction

Stroke prevention is central to the management of patients with atrial fibrillation (AF), and effective stroke prevention requires use of oral anticoagulants (OAC)(1). The latter confers an excess risk of bleeding, and various clinical factors have been associated with bleeding risk(2). These clinical factors have been used to formulate bleeding risk scores, to assess the risk of bleeding in atrial fibrillation patients(2).

Most bleeding risk scores have been derived and/or validated in patients taking Vitamin K antagonists (VKA, eg. warfarin) as the OAC(2). More recently, the non-VKA oral anticoagulants (NOAC) have been increasingly used for stroke prevention, but limited data are available on the comparative predictive and clinical value of various bleeding risk scores, specifically in NOAC users.

Of the various bleeding risk scores, the HAS-BLED(3) score has been used in various guidelines(4), but more recently the ATRIA(5) and ORBIT(6) scores have been proposed as alternative scores that appear applicable to the NOAC era. All scores assign integer valued points to a range of risk factors and use the total points to classify into risk strata (low, intermediate, high risk).

Our objective was to compare the predictive value of the stratification schemes proposed by HAS-BLED, ATRIA, and ORBIT bleeding scores in patients with atrial fibrillation treated with NOACs in a nationwide cohort study, using the Danish registries.

Methods

This study included data from three Danish nationwide databases, which may be linked according to Danish legislation for research purposes. Linkage is enabled by the unique identification number that all Danish citizens hold and is being used throughout all nationwide databases. We used: (i) the Danish National Prescription Registry(7) which holds detailed information on every prescription withdrawal since 1994; (ii) the Danish National Patient Register(8) established in 1977, which includes data for >99% of somatic hospital admissions; and (iii) the Danish Civil Registration System(9) with demographic information. Study population is detailed in the Supplementary Material.

To establish an OAC naïve cohort, we excluded patients with prior experience of any OAC inclusive doses approved for other indications within one year. Eventually, we excluded patients with prior hospital diagnoses indicating valvular atrial fibrillation (mitral stenosis or mechanical heart valves) or venous thromboembolism (pulmonary embolism or deep venous thromboembolism). This population formed the study cohort for the analyses.

Endpoints and baseline variable definitions

Clinical endpoints were extracted from hospital discharge codes in the Danish National Patient Register using the 10th Revision of ICD codes (see **Supplementary Table 1** for specific codes) with follow-up until April 30, 2016. The scores were evaluated on the following bleeding events: intracranial, gastro-intestinal, traumatic intracranial and clinically relevant non-major bleeding reported in total as 'any bleeding' (see **supplementary Table 1** for ICD-10 discharge codes). Primary and secondary inpatient hospital discharge codes were used for endpoint evaluation; to ensure higher validity of the measured outcomes non-emergency ward and outpatient codes were not assessed.

Patient's comorbidities and co-medications at treatment initiation (as listed in **Table 1**) were ascertained from the Danish National Patient Registry and the Danish National Prescription Registry (for code definitions, see **Supplementary Table 1**). Baseline medication was ascertained by the presence of at least one prescription within 365 days prior to study entry, where as comorbidity based on hospital discharge codes included information from hospitalizations and ambulatory visits, but excluding diagnoses coded in emergency wards.

Bleeding scores

Based on patient baseline comorbidity and medication, bleeding risk was ascertained using the risk classifications defined by the HAS-BLED(3) , ATRIA(5), and ORBIT(6) bleeding risk scores (see score definitions in **Supplementary Table 2**). Due to non-availability of data in the national registers, labile INR in the HAS-BLED could not be evaluated, but this criterion was not relevant as all patients were OAC naïve at inclusion and initiated a NOAC treatment.

Renal dysfunction was not uniformly defined in the three scores: in HAS-BLED it is defined as presence of chronic dialysis, renal transplantation, or serum creatinine ≥ 200 m mol/L; in ATRIA: glomerular filtration rate <30 ml/min or dialysis dependent; in ORBIT as insufficient kidney function (eGFR < 60 mg/dL/1.73 m^2). We ascertained renal dysfunction as the presence of prior hospital discharge codes indicating insufficient renal function (see **Supplementary Table 1**; hypertensive kidney disease, acute or chronic glomerulonephritis, hematuria, nephrosis, nephropaty, nephritis, acute or chronic renal insufficiency, polycystic kidney disease). Thus, we applied a generic definition and did not distinguish between bleeding score definitions of renal dysfunction in our study.

The ATRIA and ORBIT scores assigned points for anemia, and the ORBIT score further included information regarding abnormal haemoglobin (<13 mg/dL for males and <12 mg/dL for females) or haematocrit (<40% for males and <36% for females). In the present evaluation, we only included hospital discharge information regarding anemia due to non-availability of laboratory data.

The HAS-BLED score assigns risk scores in the range 0 to 8 and the following risk strata categorization was suggested(3): low risk as scores 0-2 and high risk for scores ≥3. The ATRIA score assigns risk in the range 0 to 10 with the risk strata classification(10): low risk as scores 0-3 and intermediate/high risk for scores ≥4. The ORBIT score assigns risk in the range 0 to 7 and with risk strata classification(6): low risk to scores 0-2 and intermediate/high risk for scores ≥3.

In the present study all scores risk classification were considered as dichotomized to low risk vs intermediate/high risk.

Statistical analysis

Detailed statistical methods are provided in the Supplementary Materials. In brief, risk strata were compared using Cox-proportional hazards regression. Discrimination based on the dichotomized risk classification was evaluated using C-statistics based on time-dependent areas under the ROC-curves, both acknowledging survival data and competing risk for death(11). The net benefit (NB)

for assigning intermediate/high risk was evaluated by decision curve analysis (12)(13). The risk threshold is the anticipated treatment risk at which the utility of treatment equals the cost of avoiding treatment.

Results

A total number of 57,930 OAC naïve non-valvular atrial fibrillation patients were identified and included in this study [**Supplemental figure 1**]. Females accounted for 44.6% and the mean age (SD) was 73.5 (11.4) years. The mean CHA_2DS_2-VASc score for stroke risk was 3.2 (1.8), with hypertension being the most prevalent risk factor (59.0%). Further population baseline characteristics are summarised in **Table 1**.

The distributions of the individual bleeding score levels are shown in **Table 2**, with the mean scores ranging from 1.4 (ORBIT) to 2.4 (HAS-BLED) with ATRIA being intermediate with mean 2.0. The ATRIA and ORBIT scores categorized 12.7% and 16.4% of the study population, respectively, as being at intermediate/high bleeding risk, whereas this proportion using HAS-BLED was 46.8%.

Event rates in relation to bleeding risk scores

The overall event rate (per 100 person-years) of the combined bleeding endpoint was 2.41 at 1 year of follow-up. The 1-year bleeding rates ranged between 0.47 and approximately 11.3 across individual score levels with the lowest rate identified by HAS-BLED score level 0. The score level 0 groups of ATRIA and ORBIT both had bleeding rates of 0.80-0.90 per 100 person-years (**Table 2**). All scores showed positive trends for increased bleeding rates with increasing scores. The crude 1-year bleeding rates per 100 person-years for patients categorized as intermediate/high risk were 3.30 (HAS-BLED), and by combining intermediate and high-risk groups in ORBIT: 5.84, and ATRIA: 6.1 (ATRIA). On score level, the lowest rates in the intermediate/high risk categories were 2.75 (HAS-BLED), 4.72 (ORBIT), and 5.30 (ATRIA). The low risk groups represented rates up to 2.93 (ORBIT) per 100 person-years. See **Supplementary Table 3** for the corresponding rates for 2.5 years follow-up.

The low risk categories had bleeding risk at 1-year follow-up of below 1.9%, whereas the intermediate/high risk category in HAS-BLED conferred an average risk of 3.0%, while for ORBIT intermediate risk was 4.2% and for ATRIA, 4.7%, in high risk groups the risk was about 5.6% based on cumulative incidence (**Figure 1**). The rate differences between the low risk and intermediate risk groups corresponded to hazard rate ratios (HRs) with low risk group as reference: HAS-BLED with HR 1.99 (95% CI 1.77-2.23); ATRIA with HR 2.73 (95% CI 2.29-3.25); and ORBIT with HR 2.61 (95% CI 2.22-3.07), noting that HAS-BLED does not distinguish between intermediate and high risk.

HRs were generally attenuated when evaluated at 2.5 years follow-up (see **Supplementary figure 2**).

Predictive value of bleeding scores

In terms of discrimination, **Table 3** shows C-statistics at 1 year follow-up for ATRIA as 0.59 (95% CI 0.57-0.60), HAS-BLED 0.58 (95% CI 0.57-0.59), and ORBIT 0.61 (95% CI 0.59-0.62), with ORBIT displaying statistically significant difference from both ATRIA and HAS-BLED ($p<0.001$). At the 2.5 years follow-up, comparable statistics were obtained (**Supplementary figure 3**).

Qualitatively, HAS-BLED show higher sensitivity (62.8%) for categorizing intermediate/high risk compared to ATRIA (29.7%) and ORBIT (37.1%) at the expense of reduced specificity for categorization as intermediate/high risk (specificity: HAS-BLED 53.5%; ATRIA 87.6%; ORBIT 84.0%). In this population, the positive predictive values ranged from 3.0% (HAS-BLED) to 5.2% (ATRIA, ORBIT), and with all negative predictive values above 98.2% at 1 year (Table 3). After 2.5 years, the PPVs overall increased by approx. 70% (HAS-BLED: 5.4%, ORBIT: 8.8%, ATRIA: 9.2%) whereas NPV was essentially unchanged (Supplementary Table 4).

Decision curve analysis

The clinical usefulness was evaluated by use of decision curves, as presented in **Figure 3**. The potential benefit of being guided by the score classification is linked to the assumed threshold for intervention, which in this study should be considered as the patient being subjected to more extensive monitoring to avoid bleeding incidences. At a 1 year perspective, and if the threshold is low (<1.7%), the benefit is towards monitoring all patients. If preference is to thresholds ranging 1.7% and 2.0%, most benefit was obtained by using the HAS-BLED score as guidance. The ORBIT or ATRIA scores provided better benefit for thresholds between 2% and 6%.

If the preferred threshold for intervention is above 6%, none of the tested scores will have positive benefit, since the intermediate/high risk categories identified as a whole have an incidence of bleeding at maximum of 6% (**Figure 3**).

The relations between the scores were maintained at 2.5 years of follow-up although the thresholds were shifted upwards (**Supplementary Figure 4**).

Discussion

Our principal finding is that at 1 year follow-up, predictive values (using AUC/C-statistics) for ATRIA, HAS-BLED and ORBIT risk score classifications were broadly similar and performances were modest. Second, both ATRIA and ORBIT categorized more patients as 'low risk' and qualitatively, there was higher sensitivity on the expense of specificity and positive predictive value for HAS-BLED compared to ATRIA and ORBIT. Third, decision curve analyses at a 1 year perspective shows that if preference is for a major bleeding risk threshold between 1.7-2.0%, most benefit was obtained by using HAS-BLED, whereas the ORBIT and ATRIA scores provided better benefit for thresholds between 2.6-6.0%. As far as we are aware this is the largest "real world" analysis of bleeding risk scores in atrial fibrillation patients who are NOAC users, based on an entire nationwide cohort (and not selective insurance provider claims data).

The use of bleeding risk scores has been subject to misinterpretation and misuse(14). Bleeding risk assessment should be part of the holistic management of atrial fibrillation patients being started on antithrombotic therapy. While modifiable bleeding risk factors should be addressed in all anticoagulated patients, a high bleeding risk score per se should be a help to treating physicians by 'flagging up' those patients at risk of bleeding for more regular review and follow-up (which is relevant these days in the era of Electronic Health Records (EHR) (14)).

Since many bleeding risk factors are potentially modifiable, for example, (uncontrolled) hypertension, concomitant use of aspirin or NSAIDs, alcohol excess, etc – a useful bleeding risk score should draw attention to these reversible factors, so they can be addressed(2). In a VKA user, labile INR (as reflected by poor time in therapeutic range, TTR) is a powerful determinant of bleeding (and thromboembolism) risk(15), but this criterion in HAS-BLED is not applicable in NOAC users. Bleeding risk assessment is also a dynamic process, and should be applicable at all stages of the patient management pathway: when first diagnosed on no antithrombotic therapy (or aspirin) and following OAC initiation. A high bleeding risk score is not a reason to withhold OAC, as the net clinical benefit balancing ischaemic stroke reduction against serious bleeding is even greater in such patients(16).

The HAS-BLED score was initially derived from the EuroHeart survey atrial fibrillation population on VKA, and has been subsequently validated in patient cohorts who are not taking any antithrombotic therapy, aspirin and OAC (whether VKA or non-VKA users), as well as in atrial fibrillation and non- atrial fibrillation cohorts in trial and non-trial (ie. 'real world') patients

populations(3) (17–19). The ATRIA bleeding score was derived from the ATRIA community cohort amongst VKA users, and validated in the ROCKET-AF trial population, while the ORBIT score was derived from the ORBIT registry, where most were VKA users(5) (6). In the derivation studies(3) (5) (6), the rates for low risk were in the derivation/validation cohorts at maximum 3.20/1.88 (respectively) for the HAS-BLED score, 2.9 in ORBIT (missing validation cohort), and 0.88/1.27 in ATRIA derivation/validation cohorts. The rates for high-risk groups were at least 19.51/3.74 in HAS-BLED derivation/validation cohorts, 6.8 in ORBIT derivation cohort, and 6.34/4.18 in ATRIA derivation/validation cohorts. Highlighting, that these studies did not apply comparable risk thresholds when assigning risk classification. Withholding these classifications, as in the present study, recent analyses amongst VKA users clearly show that ATRIA and ORBIT would have a significantly poorer predictive value for clinically relevant or major bleeding or ICH compared to HAS-BLED, by not considering the labile INR criterion(20, 21).

In correspondence with the score comparison based on the AMADEUS trials(21) ATRIA and ORBIT categorized >85% as 'low risk' which may lead to non-alerts from EHR, and patients not 'flagged up' for review and follow-up. The lower risk threshold of HAS-BLED thus on the other hand lead to higher sensitivity on the expense of a lower specificity and positive predictive value. As shown in our decision curve analysis, if the preference is for a major bleeding risk intervention threshold between 1.3-1.8%, most benefit was obtained by using the HAS-BLED score, while ORBIT provided better benefit for thresholds between 1.8-3%; and ATRIA at >3%. In comparison a stroke treatment threshold for OAC has been proposed for approximately 1.0-1.7%/year(22), which was based on quality-adjusted life-year analysis. A similar bleeding risk threshold approach to aid decision-making could prove useful for treating physicians. Yet, stroke and bleeding risk stratification may not easily translate into individual patient evaluation, and decisions on life-long antithrombotic treatment should cover the patient as a whole, and not be confined to risk estimates from population-based studies. Nevertheless, our decision curve analysis may hint a preferred score on the expense on another in case an appropriate threshold can be agreed upon.

In accordance, the recent European Society of Cardiology (ESC) guidelines do not recommend a specific bleeding risk score, but tabulates a long list of modifiable, non-modifiable and biomarker-related bleeding risk factors(23). As highlighted above, this precludes use of a simple risk score to aid follow-up decisions or to help flag up patients at risk for more regular review and follow-up. While emphasis on reversible bleeding risk factors are paramount, the suggestion of a biomarker-based bleeding predictor is less useful given that many biomarkers are predictive of bleeding as

well as stroke, myocardial infarction, heart failure, sudden death etc., which could lead to confusion amongst clinicians over which endpoint to focus on. In addition, one may question the generalizability of biomarker predictions of bleeding, since these studies were based on highly selected anticoagulated cohorts in the NOAC randomized trials, where bloods tests are baseline were used for predicting events over many years of follow-up (14)(24–26). Indeed, in a real world anticoagulated atrial fibrillation cohort, the HAS-BLED score performed better than a biomarker based score in predicting major bleeding (27). Hence, clinical management of atrial fibrillation patients would be best served with a simple clinical bleeding risk score that is used appropriately.

Study limitations

The main limitations pertain to the observational nature of our study with a potential of low generalizability due to possible bias from selective prescribing. Specifically, we only included users of NOACs, and atrial fibrillation patients deemed inappropriate for this treatment were not investigated (e.g. end-stage renal kidney disease). The scores differ in some details on the definition of specific risk factors, which could not be ascertained due to lack of laboratory data; these may influence the score distribution. Observational cohorts have a risk of misclassification but the outcomes ascertained have been previously validated with a high positive predictive value (e.g. 97-100% for ischemic stroke)(28). In addition, endpoints were not adjudicated (unlike a trial cohort) and post mortems were not mandated; therefore, some severe bleeding episodes with fatal consequences may not have been fully captured in the coding applied for endpoint analysis. Patient adherence to the NOACs and prescribing practices were not considered. Our data apply to a predominantly white European population, and differential efficacy and safety benefits may be evident between Asians and non-Asians (29, 30).

Conclusions

This analysis of contemporary bleeding risk scores in a 'real world' NOAC user population with atrial fibrillation showed modest predictive performance in terms of C-statistics. The scores represent different risk thresholds with HAS-BLED classifying least patients at low risk and achieving the highest benefit if applying a major bleeding intervention threshold of approx. 2.0%, whereas benefit from using either ATRIA score or ORBIT score was evident using higher intervention thresholds.

Acknowledgements

The Obel Family Foundation partly funded this research by an unrestricted grant. The sponsor had no role the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

All authors had full access to all of the data in this study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors contributed to the design, analysis, interpretation of data, drafting the article, or revising it critically for important intellectual content and approved the final version to be published. This study was conducted fully independent of any pharmaceutical company support or collaboration. The Danish Health Data Authority provided the data material.

Roles: GYHL – study hypothesis, supervised research, drafted paper and undertook critical revisions. FS – statistical analyses and interpretation. All other authors – drafted paper and undertook critical revisions.

REFERENCES

1. Lip GYH, Lane DA. Stroke Prevention in Atrial Fibrillation. JAMA 2015;313:1950.

2. Lip GYH, Andreotti F, Fauchier L, et al. Bleeding risk assessment and management in atrial fibrillation patients. Executive Summary of a Position Document from the European Heart Rhythm Association [EHRA], endorsed by the European Society of Cardiology [ESC] Working Group on Thrombosis. Thromb. Haemost. 2011;106:997–1011.

3. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJGM, Lip GYH. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. Chest 2010;138:1093–100.

4. National-Institute-for-Health-and-Care-Excellence. Atrial fibrillation: the management of atrial fibrillation. (Clinical guideline 180.) 2014. http://guidance.nice.org.uk/CG180. 2014:http://guidance.nice.org.uk/CG180. Available at: http://guidance.nice.org.uk/CG180. [Accessed 17 November 2017]

5. 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.

6. O'Brien EC, Simon DN, Thomas LE, et al. The ORBIT bleeding score: a simple bedside score to assess bleeding risk in atrial fibrillation. Eur. Heart J. 2015; 36:3258-64.

7. Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. Scand. J. Public Health 2011;39:38–41.

8. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand. J. Public Health 2011;39:30–33.

9. Pedersen CB. The Danish Civil Registration System. Scand. J. Public Health 2011;39:22–25.

10. 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.

11. Blanche P, Dartigues J-F, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat. Med. 2013;32:5381–97.

12. Peirce CS. The numerical measure of the success predictions. Science 1884;4:453–454.

13. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med. Decis. Making 2006;26:565–74.

14. Lip GYH, Lane DA. Bleeding risk assessment in atrial fibrillation: observations on the use and misuse of bleeding risk scores. J. Thromb. Haemost. 2016;14:1711–1714.

15. Wan Y, Heneghan C, Perera R, et al. Anticoagulation control and prediction of adverse events in patients with atrial fibrillation: A systematic review. Circ. Cardiovasc. Qual. Outcomes 2008;1:84–91.

16. Olesen JB, Lip GYH, Lindhardsen J, et al. Risks of thromboembolism and bleeding with thromboprophylaxis in patients with atrial fibrillation: A net clinical benefit analysis using a "real world" nationwide cohort study. Thromb. Haemost. 2011;106:739–749.

17. Apostolakis S, Lane DA, Guo Y, Buller H, Lip GYH. 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 atria. J. Am. Coll. Cardiol. 2012;60:861–7.

18. Apostolakis S, Lane DA, Guo Y, Buller H, Lip GYH. Performance of the HEMORR 2 HAGES, ATRIA, and HAS-BLED bleeding risk-prediction scores in nonwarfarin anticoagulated atrial fibrillation patients. J. Am. Coll. Cardiol. 2013;61:386–7.

19. Lip GYH, Banerjee A, Lagrenade I, Lane DA, Taillandier S, Fauchier L. Assessing the risk of bleeding in patients with atrial fibrillation: the Loire Valley Atrial Fibrillation project. Circ. Arrhythm. Electrophysiol. 2012;5:941–8.

20. Proietti M, Senoo K, Lane DA, Lip GYH. Major Bleeding in Patients with Non-Valvular Atrial Fibrillation: Impact of Time in Therapeutic Range on Contemporary Bleeding Risk Scores. Sci. Rep. 2016;6:24376.

21. Senoo K, Proietti M, Lane DA, Lip GYH. Evaluation of the HAS-BLED, ATRIA, and ORBIT Bleeding Risk Scores in Patients with Atrial Fibrillation Taking Warfarin. Am. J. Med. 2016;129:600–607.

22. Eckman MH, Singer DE, Rosand J, Greenberg SM. Moving the tipping point: the decision to anticoagulate patients with atrial fibrillation. Circ Cardiovasc Qual Outcomes 2011;4:14–21.

23. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation. Europace. 2016;18:1609-1678.

24. Hijazi Z, Oldgren J, Lindback J, et al. The novel biomarker-based ABC (age, biomarkers, clinical history)-bleeding risk score for patients with atrial fibrillation: a derivation and validation study. Lancet (London, England) 2016;387:2302–2311.

25. Hijazi Z, Wallentin L, Siegbahn A, et al. N-Terminal Pro–B-Type Natriuretic Peptide for Risk Assessment in Patients With Atrial Fibrillation. J. Am. Coll. Cardiol. 2013;61:2274–2284.

26. Hijazi Z, Oldgren J, Andersson U, et al. Cardiac Biomarkers Are Associated With an Increased Risk of Stroke and Death in Patients With Atrial Fibrillation: A Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) Substudy. Circulation 2012;125:1605–1616.

27. Esteve-Pastor MA, Rivera-Caravaca JM, Roldan V, et al. Long-term bleeding risk prediction in "real world" patients with atrial fibrillation: Comparison of the HAS-BLED and ABC-Bleeding risk scores. Thromb. Haemost. 2017;117;117:1848-1858..

28. Krarup L-H, Boysen G, Janjua H, Prescott E, Truelsen T. Validity of stroke diagnoses in a National Register of Patients. Neuroepidemiology 2007;28:150–4.

29. Wang K-L, Lip GYH, Lin S-J, Chiang C-E. Non–Vitamin K Antagonist Oral Anticoagulants for Stroke Prevention in Asian Patients With Nonvalvular Atrial Fibrillation. Stroke 2015;46:2555–2561.

30. Chiang C-E, Wang K-L, Lip GYH. Stroke prevention in atrial fibrillation: An Asian perspective. Thromb. Haemost. 2014;111:789–797.

Figure 1: Cumulative incidence curves for outcomes at up to 1 years follow-up for each risk classification strata and score and with hazard rate ratio (HR (95%CI)) for intermediate/high risk vs low risk group within risk classification score.

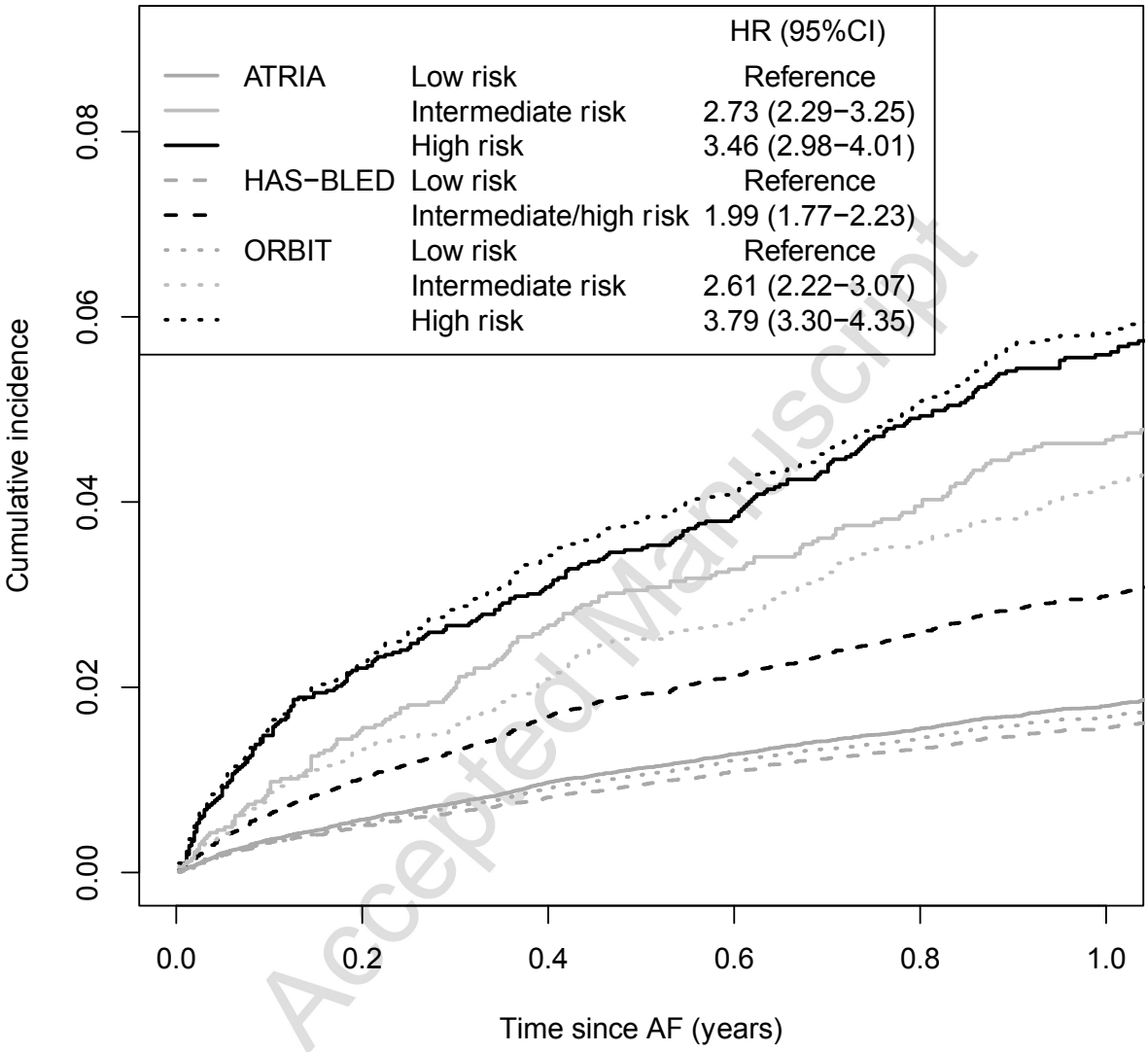


Figure 2: Calibration plots for risk scores and risk strata (low, intermediate high risk) of any bleeding at 1 years follow-up.

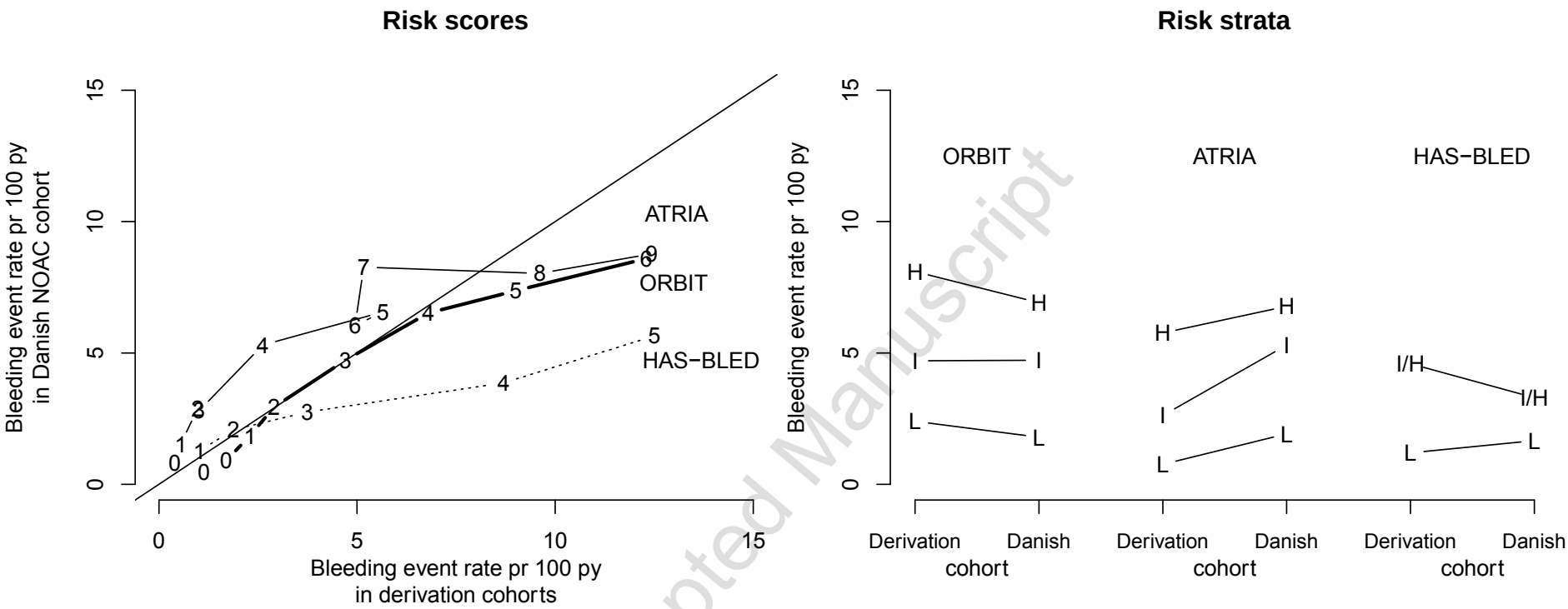


Figure 3 Decision curves curve analysis plots for score risk classification (low, intermediate, high risk) of any bleeding at 1 year follow-up.

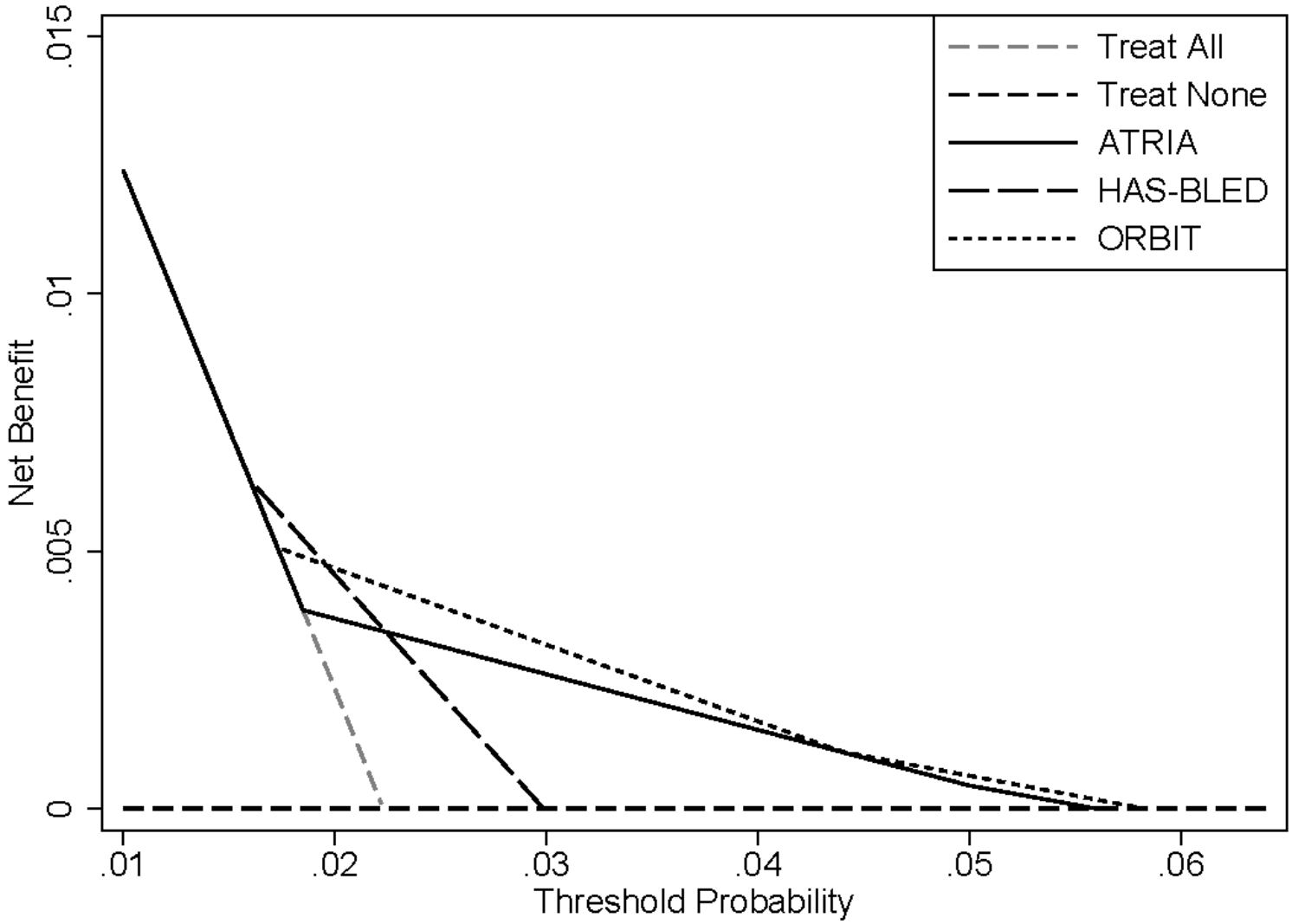


Table 1. Participant characteristics at treatment initiation.

	NOAC N=57,930
Females	44.6 (25,839)
Age, mean (sd)	73.5 (11.4)
Prior AF diagnosis	65.1 (37,706)
CHA_2DS_2-VASc, mean (sd)	3.2 (1.8)
Heart failure or LVD	22.5 (13,053)
Diabetes mellitus	15.2 (8,798)
Vascular diseases	16.2 (9,404)
Hypertension	59.0 (34,153)
CPD	13.3 (7,704)
Prior bleeding	14.2 (8,242)
Kidney diseases	3.4 (1,968)
Aspirin#	39.1 (22,638)
Beta-blocker#	37.8.7 (21,891)
NSAIDs#	22.4 (12,949)
Loop diuretics#	18.0 (10,446)

Abbreviations: sd = Between-subjects standard deviation, CPD= Chronic pulmonary disease, Kidney diseases: Renal dysfunction/kidney transplant/dialysis, NSAIDs = Non-steroidal anti-inflammatory drugs

#At least one prescription within 1 year prior to NOAC initiation.

Table 2. Distribution of individual baseline score levels and risk classification and crude event rates per 100 person-years (number of events) at 1 year follow-up of any bleeding according to bleeding score level and risk classification.

	Baseline score distribution			Crude 1 year event bleeding event rates		
Score level	**ATRIA** % (N)	**HAS-BLED** % (N)	**ORBIT** % (N)	**ATRIA** Rate (N)	**HAS-BLED** Rate (N)	**ORBIT** Rate (N)
0	21.0 (12,158)	4.8 (2,770)	28.0 (16,233)	0.81 (91)	0.47 (12)	0.91 (135)
1	26.5 (15,372)	18.2 (10,567)	35.2 (19,701)	1.53 (215)	1.27 (121)	1.85 (325)
2	15.9 (9,208)	30.2 (17,497)	21.6 (12,517)	2.87 (229)	2.08 (324)	2.93 (316)
3	23.8 (13,805)	29.9 (17,309)	7.8 (4,508)	2.80 (332)	2.75 (418)	4.72 (180)
4	5.6 (3,263)	13.6 (7,877)	6.1 (3,520)	5.30 (146)	3.86 (260)	6.52 (188)
5	2.0 (1,151)	2.9 (1,708)	1.4 (829)	6.56 (60)	5.65 (80)	7.37 (48)
6	3.3 (1,888)	0.3 (202)	1.1 (622)	6.04 (90)	11.33 (18)	8.59 (41)
7	1.3 (769)	-	-	8.27 (50)	-	-
8	0.1 (83)	-	-	8.03 (5)	-	-
9-10	0.4 (233)	-	-	8.77 (15)	-	-
Mean (sd)	2.0 (1.7)	2.4 (1.2)	1.4 (1.3)			
Score risk classification						
Low risk	87.2 (50,543)	53.2 (30.834)	83.6 (48,451)	1.92 (867)	1.66 (457)	1.79 (776)
Intermediate	5.6 (3,263)	46.8 (27,096)	7.8 (4,508)	5.30 (146)	3.30 (776)	4.72 (180)
High risk	7.1 (4,124)	NA	8.6 (4,971)	6.78 (220)	NA	6.91 (277)

Table 3. Discriminative statistics for risk scores using published categorization (low, intermediate, high risk), overall by C-statistics, and by risk thresholds in terms of sensitivity, specificity, negative (NPV) and positive predictive values (PPV) for any bleeding at 1 year follow-up.

Risk score	C-statistics	Threshold	Sensitivity	Specificity	NPV	PPV
ATRIA	0.59 (0.57-0.60)	Intermediate / high risk	29.7	87.6	98.2	5.2
		High risk	17.9	93.1	98.0	5.6
HAS-BLED	0.58 (0.57-0.59)	Intermediate / high risk	62.8	53.5	98.4	3.0
		High risk	-	-	-	-
ORBIT	0.61 (0.59-0.62)	Intermediate / high risk	37.1	84.0	98.3	5.0
		High risk	22.5	91.8	98.1	5.8

SUPPLEMENTARY MATERIAL

Detailed statistical methods

Study population

The study was based on new users of NOAC with no hospital information for treatment for other indications than atrial fibrillation. We identified patients with first-time purchases of each NOAC approved for atrial fibrillation from their respective dates of approval for atrial fibrillation: apixaban (December 10, 2012), dabigatran (August 10, 2011), rivaroxaban (February 1, 2012). Patient inclusion was terminated by February 28, 2016.

Statistical analyses

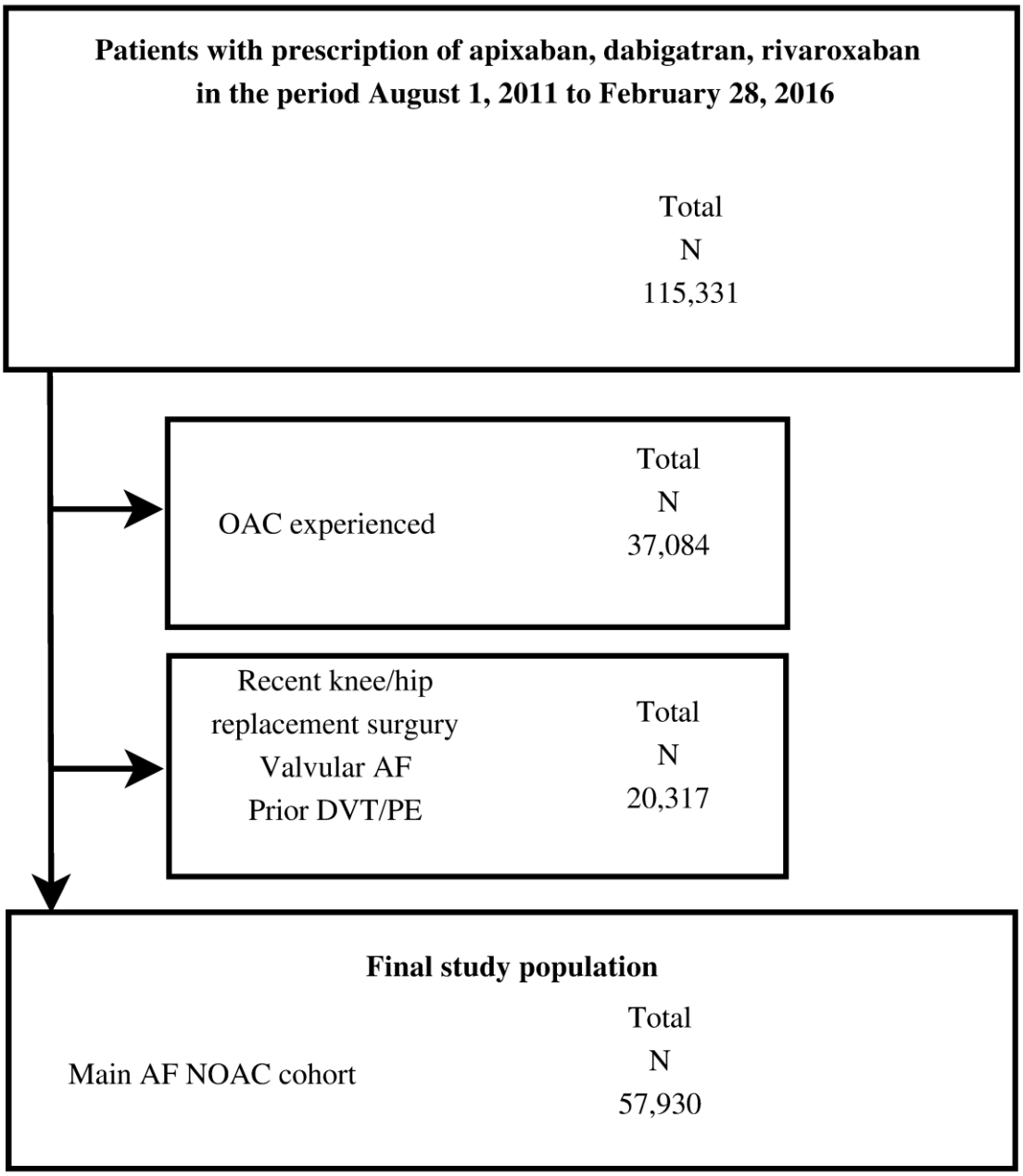
For each score, crude event rates were calculated as number of events divided by person years for each score level and for each risk classification strata. The event risk up to 2.5 years follow-up was depicted in terms of cumulative incidence functions based on the Aalen-Johansen estimator acknowledging competing risk for death. Risk strata were compared using Cox-proportional hazards regression.

To visualize the predictive performance for the three scores from their respective derivation cohort and the data applied in current study, the points from each scoring system were plotted against the event rates of bleeding. Discrimination based on the dichotomized risk classification was evaluated using C-statistics based on time-dependent areas under the ROC-curves, both acknowledging survival data and competing risk for death(11). Also negative and positive predicted values were reported. Confidence intervals for reported measures were based on 500 bootstrap samples.

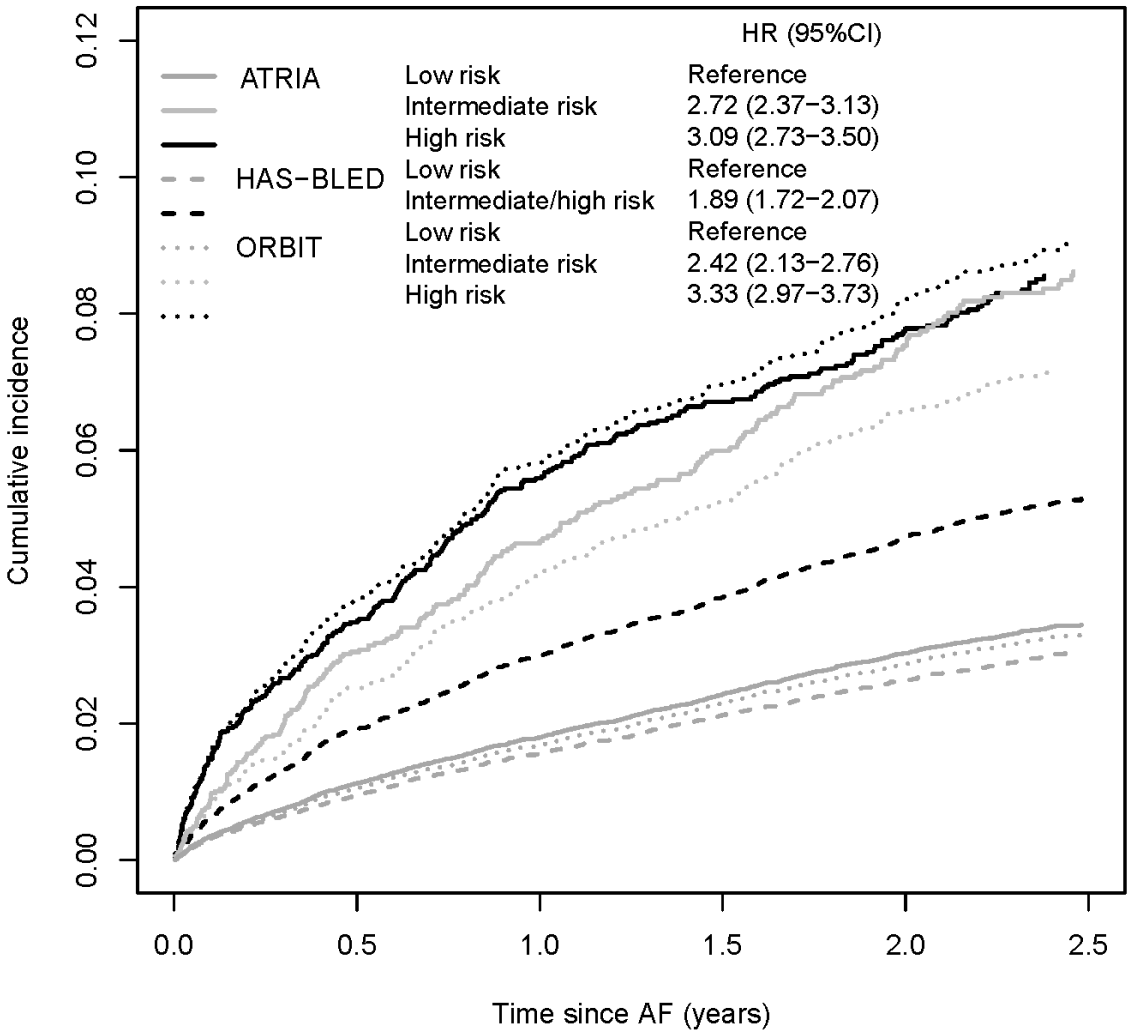
The net benefit (NB) for assigning intermediate/high risk was evaluated by decision curve analysis by weighing the proportion of true positives (*TP*) against the false positives (*FP*) as *(TP - w * FP)/N*, with "*w*" as penalty for false positive(12). In decision curve analysis, w is defined as *p/(1-p)*, where *p* is a selected risk threshold for "intervention"(13). The risk threshold is the anticipated treatment risk at which the utility of treatment equals the cost of avoiding treatment. Under this definition it can be derived that *p/(1-p)* will represent the cost of an unnecessary treatment of the false positive patient, irrespective of the scale of the utility, yet acknowledging the utility difference between treatment and no-treatment of the true positive patient. The two simplest rules are either treat all and treat none. In the treat none case this will result in *NB=0* as no one is considered positive (as *TP=FP=0*). Under the treat all rule NB depend on the anticipated risk threshold. If there is correspondence between the rule and the threshold, i.e. *p=0*, then *NB=TP/N*, which is the prevalence of positives. If the threshold is higher, *p>0,* then NB decrease linearly and reach *zero* at *p=TP/(FP+FP)*, which under the treat all rule again corresponds to the prevalence of positives, as all negatives are classified as false positives. For a dichotomous decision rule (low risk vs high risk) *NB* will coincide with the treat all rule up to the point where risk threshold exceeds the proportion of true positives $p_{low} = TP_{low}/(FP_{low} + FP_{low})$ in the low risk group. A lower risk threshold ($p<p_{low}$) would correspond to classifying all patients as positives. A higher risk threshold ($p>p_{low}$) calls for the use of the decision rule to focus on the high risk population. If the decision rule is sensible, one should expect that the number of true positives will be materially unchanged, whereas the number of false positives should be substantially reduced. In effect, this ought to result in a higher NB than for the treat all strategy. As the threshold *p* increases, *NB* will diminish and reach *NB=0* at $p_{high} = TP_{high}/(FP_{high} + FP_{high})$ the proportion of true positives in the high risk group. In the setting of survival data the proportions are replaced by the corresponding risks, e.g. estimated by Aalen-Johansen cumulative incidence functions.

Stata/MP version 14 and R version 3.1.1 was used for the statistical analysis. A two-sided P value <0.05 was considered statistically significant.

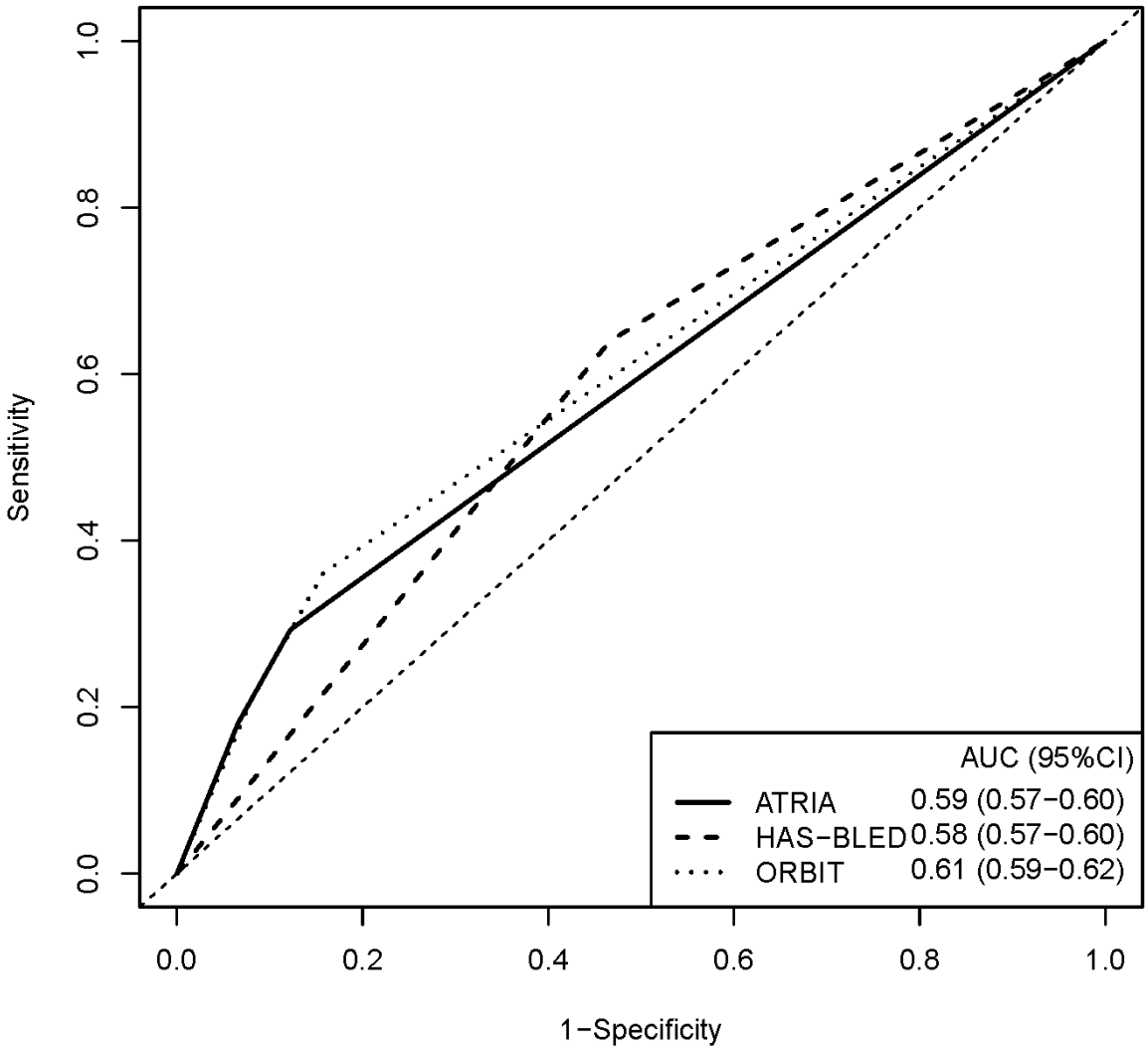
Supplementary figure 1 Flowchart of study population selection



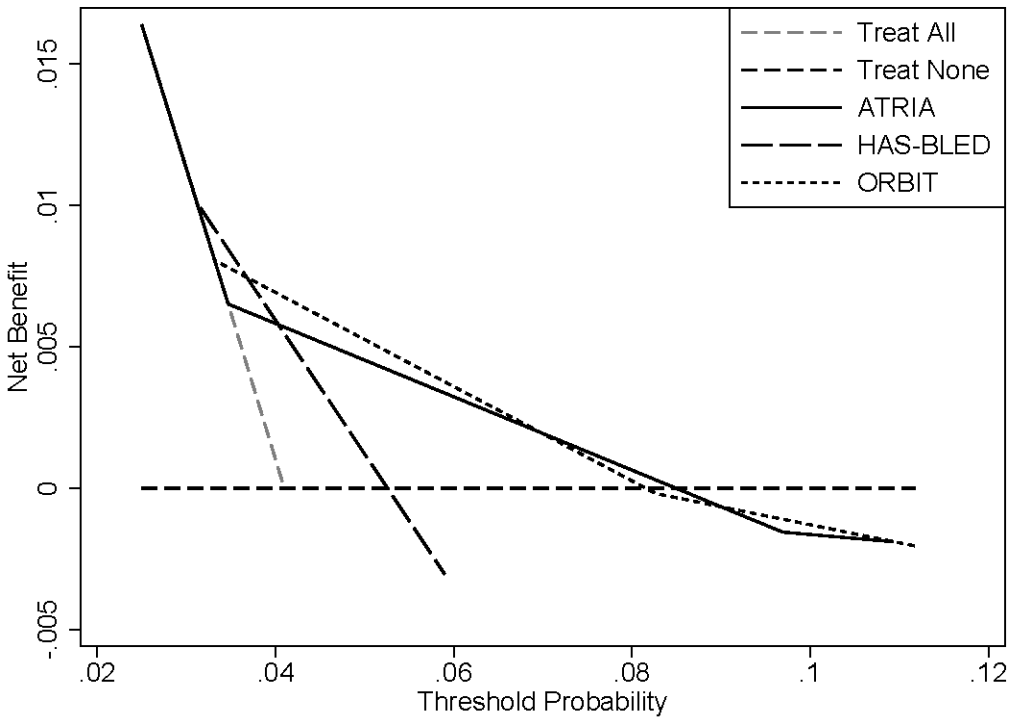
Supplementary figure 2 Cumulative incidence curves for outcomes at up to 2.5 years follow-up for each risk classification strata and score and with hazard rate ratio for intermediate/high risk vs low risk group within risk classification score.



Supplementary figure 3 ROC curves for score risk classification (low, intermediate/high risk) discrimination of any bleeding at 2.5 years follow-up



Supplementary figure 4 Decision curves curve analysis plots for score risk classification (low, intermediate/high risk) of any bleeding at 2.5 years follow-up.



Supplementary Table 1

Definitions on comorbidity and concomitant medication according to ICD-10 codes and ATC-codes. Conditions marked with † was used in the calculation of the CHA_2DS_2-VASc score. Conditions marked with # was used in the calculation of the HAS-BLED score.

	International Classification of Diseases 10th revision (ICD-10) code	Anatomical Therapeutic Chemical (ATC) code
Condition		
†Congestive heart failure	I110 I130 I132 I420 I50	CO3C and C09
†Left ventricular dysfunction	I501 I509	
†#Hypertension		See specified definition*
†Diabetes mellitus	E100 E101 E109 E110 E111 E119	A10
†#Ischemic stroke	I63 I64	
†Systemic embolism	I74	
†#Transient ischemic disease	G45	
†Aortic plaque	I70.0	
†Peripheral arterial disease	I702-I709 I71 I739	
†Myocardial infarction	I21-I23	
#Moderate/severe renal disease	I12 I13 N00 N01 N02 N03 N04 N05 N07 N11 N14 N17 N18 N19 Q61	
#Moderate/Severe liver disease	B150 B160 B162 B190 K704 K72 K766 I85	

Cancer	C	
Chronic pulmonary disease	J40 J41 J42 J43 J44 J45 J46 J47 J60 J61 J62 J63 J64 J65 J67 J684 J701 J703 J841 J920 J921 J982 J983	
Mitral stenosis	I05	
Mechanical heart valve	Z952 Z953 Z954	
#Haemorrhagic stroke – intracranial bleeding	I60 I61 I62	
#Extracranial or unclassified major bleeding	D62 J942 H113 H356 H431 N02 R04 R31 R58	
#Gastrointestinal bleeding	K250 K252 K254 K260 K262 K264 K270 K272 K274 K280 K282 K290 K291	
#Traumatic intercranial bleeding	S063C S064 S065 S066	
#Alcohol	E224 E529A F10 G312 G621 G721 I426 K292 K70 K860 L278A O354 T51 Z714 Z721	
Pulmonary embolism	I26	
Deep venous thromboembolism	I801 I802 I803 I808 I809 I819 I636 I676 I822 I823 I829	
Atrial fibrillation	I48	
Medication		
Apixaban		B01AF02
Dabigatran		B01AE07

Rivaroxaban	B01AF01
Warfarin	B01AA03
Phenprocoumon	B01AA04
#Aspirin	B01AC06
#Clopidogrel	B01AC04
Beta-blockers	C07
Statins	C10
#NSAIDs	M01A
Loop diuretics	C03C
ACE/ARB inhibitors	C09

NSAIDs: Non Steroidal Anti Inflammatory Drugs.

* We identified subjects with hypertension from combination treatment with at least two of the following classes of antihypertensive drugs:

I· Alpha adrenergic blockers (C02A, C02B, C02C)
II· Non-loop diuretics (C02DA, C02L, C03A, C03B, C03D, C03E, C03X, C07C, C07D, C08G, C09BA, C09DA, C09XA52)
III· Vasodilators (C02DB, C02DD, C02DG, C04, C05)
IV· Beta blockers (C07)
V· Calcium channel blockers (C07F, C08, C09BB, C09DB)
VI· Renin-angiotensin system inhibitors (C09)
As well as combination drugs: C07B, C09BB04, C09DA, C09DB, C09DX01, C09DX04

Danish hospital discharge codes have been validated in extensively in the literature, please consult Schmidt et al. Clin Epid 2015;7:449-90 for a comprehensive overview and Sundbøll et al. BMJ Open 2016;6(11).

Supplementary table 2. Bleeding Score definitions

1. HAS-BLED

Component	Prefix	Weight
Hypertension	(Alfa+NonLoop+Vaso+Beta+Calcium+Renin)>1 or combination drug	1
Renal Disease	Renal	1
Liver Disease	Liver	1
Stroke	Istroke or TIA	1
Bleeding	IBleed or MBleed3 or Gbleed2 or TIbleed	1
Age	Age>=65	1
Drugs	Aspirin or clopidogrel or NSAID	1
Alcohol	Alco	1

Risk classification Low: 0-3 Intermediate/High 4-8

2. ATRIA

Component	Prefix	Weight
Anemia	Anemia	3
eGFR<30 or dialysis	Renal	3
Age	Age>=75	2
Bleeding	IBleed or MBleed3 or Gbleed2 or TIbleed	1
Hypertension	(Alfa+NonLoop+Vaso+Beta+Calcium+Renin)>1 or combination drug	1

Risk classification Low: 0-3 Intermediate: 4 High 5-10

3. ORBIT

Component	Prefix	Weight
Age	Age>=75	1
Anemia/red.haem	Anemia	2
Bleeding	IBleed or MBleed3 or Gbleed2 or TIbleed	2
eGFR<60mL/min/1.73m^2	Renal	1
Antiplatelet	Aspirin or thienopyridines	1

Risk classification Low: 0-2 Intermediate: 3 High 4-7

Supplementary Table 3. Crude event rates per 100 person-years (number of events) at 2.5 years follow-up of any bleeding according to bleeding score level and risk classification.

	2.5 year		
Score level	**ATRIA** Rate (N)	**HAS-BLED** Rate (N)	**ORBIT** Rate (N)
0	0.77 (170)	0.42 (21)	0.82 (238)
1	1.23 (347)	1.03 (189)	1.59 (548)
2	2.42 (363)	1.79 (539)	2.49 (508)
3	2.42 (538)	2.18 (642)	3.78 (271)
4	4.47 (230)	3.24 (413)	4.97 (260)
5	4.97 (81)	4.36 (114)	5.77 (66)
6	4.55 (121)	10.22 (28)	6.53 (55)
7	6.58 (71)	-	-
8	6.11 (6)	-	-
9-10	6.94 (19)	-	-
Score risk classification			
Low risk	1.62 (1,418)	1.40 (749)	1.54 (1294)
Intermediate/high risk	4.85 (528)	2.66 (1197)	4.53 (652)

Supplementary table 4. Discriminative statistics for risk scores using published categorization (low, intermediate, high risk), overall by C-statistics, and by risk thresholds in terms of sensitivity, specificity, negative (NPV) and positive predictive values (PPV) for any bleeding at 2.5 year follow-up.

Risk score	C-statistics	Threshold	Sensitivity	Specificity	NPV	PPV
ATRIA	0.57 (0.56-0.58)	Intermediate / high risk	26.8	87.9	96.6	8.7
		High risk	14.9	93.3	96.2	8.7
HAS-BLED	0.57 (0.56-0.58)	Intermediate / high risk	60.6	53.8	97.0	5.3
		High risk	-	-	-	-
ORBIT	0.59 (0.58-0.60)	Intermediate / high risk	32.7	84.4	96.7	8.2
		High risk	18.9	91.9	96.4	9.1