Meta-Analysis

## Comparing the ORBIT and HAS-BLED bleeding risk scores in anticoagulated atrial fibrillation patients: a systematic review and meta-analysis

### Cen Wang<sup>1,\*</sup>, Ye Yu<sup>1,\*</sup>, Wengen Zhu<sup>1</sup>, Jianhua Yu<sup>1</sup>, Gregory Y.H. Lip<sup>3,#</sup> and Kui Hong<sup>1,2,#</sup>

<sup>1</sup>Department of Cardiovascular Medicine, the Second Affiliated Hospital of Nanchang University, Jiangxi, China

<sup>2</sup>The Jiangxi Key Laboratory of Molecular Medicine, Jiangxi, China

<sup>3</sup>University of Birmingham Centre for Cardiovascular Sciences, City Hospital, Birmingham, United Kingdom

<sup>\*</sup>Joint first author

<sup>#</sup>Joint senior author

**Correspondence to:** Kui Hong, **email:** hongkui88@163.com **Keywords:** *atrial fibrillation, ORBIT, HAS-BLED, anticoagulation, major bleeding* 

**Received:** May 04, 2017 **Accepted:** July 25, 2017 **Published:** August 03, 2017

**Copyright:** Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

#### ABSTRACT

Background: The HAS-BLED and ORBIT scores have been proposed to assess bleeding risk in anticoagulated atrial fibrillation patients. We performed a systematic review and meta-analysis to compare the predictive ability by using these two scores.

Materials and Methods: We searched the Cochrane Library, Elsevier and PubMed databases for related studies. Statistical analysis was performed with Revman 5.3 Manager software. We chose the C-statistic to reflect the diagnostic value.

Results: In our seven selected studies, the pooled C- statistic of continuous variables for major bleeding was 0.65 (0.60,0.69) for ORBIT and 0.63 (0.60,0.66) for HAS-BLED. Compared with HAS-BLED, more anticoagulated AF patients (88.45% versus 32.59%) and major bleeding events (75.57% versus 25.57%) were categorized as low risk. The ORBIT score had a 1.21, 1.73 and 1.44-fold elevated risk of major bleeding in the low, intermediate and high risk strata respectively. Calibration analysis demonstrated that the ORBIT score under-predicted major bleeding in the low, intermediate, and high risk stratifications, where a odds ratio of 0.64 (0.37–1.10), 0.63 (0.38–1.05) and 0.64 (0.38–1.06), respectively.

Conclusions: Compared with HAS-BLED, the ORBIT score does not perform better in predicting major bleeding events in anticoagulated atrial fibrillation patients. More anticoagulated AF patients and major bleeding events were categorized as low risk when using ORBIT.

### **INTRODUCTION**

Oral anticoagulants (OAC) are the cornerstones of atrial fibrillation (AF) management for preventing of stroke, thromboembolic events and mortality [1–3]. Bleeding is the major downside of OAC therapy [4, 5].

Many common clinical features are associated with an increase in bleeding risk, and the more common features (many of which are potentially reversible or correctable) have been used to formulate various bleeding risk stratification schema [5]. Nonetheless bleeding risk scores have been subject to misuse, as the objective of performing a bleeding risk assessment is to 'flag up' patients potentially at risk of bleeding for more careful review and follow up, as well as to address the potentially reversible bleeding risk factors [6].

Of the various published bleeding risk scores, the HAS-BLED score (hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile international normalized ratio [INR], elderly [age  $\geq 65$  years], and drugs/alcohol concomitantly) has received widespread attention since it was first

	<b>Risk predictors</b>	Scoring system	<b>Risk stratification</b>
ORBIT	Older age (≥ 74 years) Reduced hemoglobin/anemia Bleeding history Insufficient kidney function Treatment with antiplatelet	1 point for each risk factor	Low risk 0–2 Intermediate risk 3 High risk ≥ 4
HAS-BLED	Hypertension Abnormal renal and/or liver function Stroke Bleeding history Labile INR Elderly (≥ 65 years) Drugs or alcohol concomitant	1 point for each risk factor	Low risk 0–1 Intermediate risk 2 High risk ≥ 3

Table 1: The risk predictors and risk stratifications of bleeding scores

proposed in 2010 [7]. In 2015, O'Brien et al. derived and validated a simple score, ORBIT [8] (older [age  $\geq$ 74 years], reduced hemoglobin/hematocrit/history of anemia, bleeding history, insufficient kidney function, and treatment with antiplatelet), which could be adapted to both VKA and non-VKA users. The ORBIT score was derived from the ORBIT-AF registry [9], with 10,098 AF patients taking warfarin and dabigatran, and validated in the ROCKET-AF trial population (treated with warfarin or rivaroxaban) [8]. According to the ORBIT score, AF patients were categorized as 'low', 'intermediate', and 'high' risk, based on score points '0-2', '3', and ' $\geq$  4', respectively. Table 1 summarises the components of the HAS-BLED and ORBIT scores.

It is unclear whether the ORBIT or HAS-BLED score has a better predictive ability for bleeding in AF patients [8, 10–15]. Our previous meta-analysis compared the HAS-BLED score with several bleeding risk scores (not include the ORBIT score) and showed that the HAS-BLED score had a better ability to predict bleeding risk than the other selected bleeding risk scores [16].

Given the recent interest in the ORBIT score, we performed a systematic review and meta-analysis to compare the predictive ability of bleeding risks using the ORBIT and HAS-BLED scores.

### **MATERIALS AND METHODS**

#### Inclusion and exclusion criteria

We selected studies according to the following inclusion criteria: 1) Types of studies: prospective or retrospective studies reporting the HAS-BLED and ORBIT scores for predicting the bleeding risk; 2) Participants: non-valvular AF patients with VKA and non-VKA anticoagulants; and 3) Outcomes: major bleeding was defined based on the 2005 International Society on Thrombosis and Haemostasis [ISTH] criteria [8]. (i) fatal bleeding and/or (ii) symptomatic bleeding in a critical area or organ (intracranial, intraspinal, intraocular, retroperitoneal, intra-articular or pericardial, or intramuscular with compartment syndrome), and/or (iii) bleeding causing a decrease in the hemoglobin level of 20 g  $L^{-1}$  or more or leading to the transfusion of two or more units of whole blood or red cells.

Studies with insufficient data, not published in English, certain publication types (e.g., conference abstracts, letters, comments, case reports, and reviews) were excluded from this meta-analysis. For those duplicated studies, studies with the longest follow-up or largest sample size were included.

### Literature search

We systematically searched the Cochrane Library, Elsevier and PubMed electronic databases for studies reporting the HAS-BLED and ORBIT scores for predicting the bleeding risk. The included studies were published from January 2010 to June 2017 because the HAS-BLED score was first proposed in 2010 [7] and the ORBIT score was developed in 2015 [8]. Search terms included 'atrial fibrillation', 'ORBIT', 'HAS-BLED', 'anticoagulation' and 'major bleeding'. We did not find other studies in the manual search.

# Data extraction and quality assessment of individual studies

Included studies were selected by two reviewers based on the search terms. The studies were prescreened by C.W., who read the titles and abstracts. The second round of selection involved the complete and careful review of articles by Y.Y., to confirm whether those studies described the bleeding events at each point or each stratification of the HAS-BLED and ORBIT scores. Discrepancies were resolved through discussion or consultation with a third reviewer (K.H.). The Newcastle–Ottawa scale [17] was used to evaluate the quality of all included studies. It contains three parts: cohort selection, cohort comparability, and outcome assessment. Two authors independently assessed both the risk of bias and the quality of the seven studies in our meta-analysis.

#### Statistical analysis

We used Review Manager Version 5.3 (Copenhagen, The Nordic Cochrane Centre, The Cochrane Collaboration, 2014) to analyze the data. Receiver operating characteristic curve (ROC) analysis, expressed by the C-statistic [95% confidence intervals (CIs)] of continuous variables, was applied to evaluate the diagnostic performance in the HAS-BLED and ORBIT scores [18–21]. The Z-statistic

 $(\frac{A_{Z1} - A_{Z2}}{\sqrt{SE_1^2 + SE_2^2}})$  was used to assess the difference

between these two scores.  $I^2$  values  $\leq 25\%$ , 25% to  $\leq 50\%$ , and > 50% represented 'low', 'intermediate' and 'high' heterogeneity, respectively. When I<sup>2</sup> values > 50%, a random-effects model was selected [22]. Both the HAS-BLED and ORBIT scores were divided into three risk stratifications. HAS-BLED = 0–1 and ORBIT = 0–2 were considered as low risk; HAS-BLED = 2 and ORBIT = 3 were defined as intermediate risk; and HAS-BLED  $\geq 3$  and ORBIT  $\geq 4$  were regarded as high risk (Table 1). Odds ratios (ORs) with their 95% CIs were chosen to compare the predictive value between these two scores. In addition,

calibration analysis was used to assess the predictive accuracy of risk scores. The first published study reporting the ORBIT was a predictive model providing the adjusted bleeding rate (events/100 patient-years) across three risk classifications (low, 2.4%; intermediate, 4.7%; and high, 8.1%). Then, we calculated the predicted number of major bleeding events in the subsequent validated studies. The observed number of major bleeding events was collected from all the included studies. OR > 1 represented over-prediction, while OR < 1 represented under-prediction.

#### RESULTS

## Description of the included studies and patients' characteristics

A total of 114 studies were initially identified by the above-mentioned search strategies (42 in Elsevier, 32 in Cochrane Library, and 40 in PubMed). After reading the titles and abstracts, we excluded 69 studies. Next, 24 studies were removed because the studies did not relate to the bleeding events or the selected risk scores (ORBIT or HAS-BLED). Finally, 7 identified studies met our criteria [8, 10–15], and the others were excluded as follows: (i) Certain publication types with no data (n = 3); (ii) Duplicate data without follow-up (n = 4); (iii) Studies not published in English (n = 2); and (iv) Only one risk score (ORBIT or HAS-BLED) used to evaluate bleeding risk (n = 5) (Figure 1).

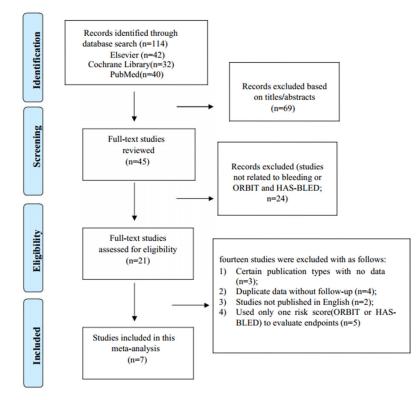


Figure 1: Study search diagram.

Table 2: Patients' characteristics of the 7 included studies

Study	Data source	Follow-up	Patient number	Age (years)	Major Bleeding Definitions	Major Bleeding events	Anticoagulants
O'Brien E C [8]	USA (2010–2012)	Mean, 24 months	7411	75	2005 ISTH criteria	1353	Warfarin DOACs
Proietti M [10]	UK (-)	Median, 19 months	3551	72	NA	127	Warfarin
Senoo K [11]	Netherlands (2003-2005)	NA	2293	71	2005 ISTH criteria	NA	Warfarin
Esteve-Pastor M A [12]	Spain (2013–2014)	Mean, 12 months	1276	74	2005 ISTH criteria	46	Warfarin DOACs
Senoo K [13]	Netherlands (2003-2005)	Median, 10 months	2283	71	2005 ISTH criteria	74	Idraparinux
Abumuaileq R R [14]	Spain (2011–2013)	Mean, 11 months	911	75	2005 ISTH criteria	30	Warfarin
Caro M C [15]	Spain (2013–2014)	Mean, 22 months	969	76	2005 ISTH criteria	101	DOACs

Abbreviations: NA = not available; ISTH = International Society on Trombosis and Haemostasis. DOACs = the new direct oral anticoagulants, apixaban, dabigatran and rivaroxaban.

The participants of our seven included studies come from American, British, Spanish and Netherlands. [8, 10–15] (Table 2). We selected 8079 non-valvular AF patients with anticoagulant therapy (e.g., warfarin, apixaban, dabigatran and rivaroxaban) for further analyses and 348 of them had suffered from major bleeding finally. The 2005 ISTH criteria was used to define the major bleeding in our included studies. Furthermore, the quality of all included studies were evaluated by the Newcastle– Ottawa Scale. Eight or nine stars represented the high quality of these studies (Supplementary Table 1).

# Discrimination analysis between the ORBIT and HAS-BLED scores

The first original article derived and validated the ORBIT score in the ORBIT-AF registry cohort and ROCKET-AF clinical trial populations [8]. The study by Proietti et al. did not provide the C-statistic [10]. Thus, there were a total of seven C-statistics of continuous variables for major bleeding in our 7 selected studies. Four showed that the ORBIT score had a higher C-statistic than the HAS-BLED [8, 12, 14]. The C-statistics ranged from 0.58 to 0.74 (median 0.62) for the ORBIT and from 0.59 to 0.68 (median 0.63) for the HAS-BLED, indicating that both scores had a modest predictive value for bleeding risk.

In our pooled analysis, the C-statistic was 0.65 (0.60, 0.69) for the ORBIT score and 0.63 (0.60, 0.66) for the HAS-BLED score. The Z-statistic between ORBIT and HAS-BLED was 0.72 (p > 0.05), showing that these two scores had a similar discriminative performance for bleeding risk (Supplementary Table 2).

# Major bleeding risks comparing ORBIT and HAS-BLED

### Low risk category

The ORBIT score categorized 88.45% (7146/8079) of anticoagulated AF patients and 75.57% (263/348) of major bleeding events into the low risk category. With HAS-BLED, 32.59% (2633/8079) of total patients and

25.57% (89/348) of major bleeding were categorized as low-risk (Figure 2).

Our pooled analysis showed that the low-risk patients of ORBIT had a 1.21-fold greater risk of major bleeding events compared to those of HAS-BLED (OR = 1.21; 95% CI: 0.93–1.57; p = 0.17;  $I^2 = 0\%$ ; Figure 3).

#### Intermediate risk category

In this strata, 7.15% (578/8079) of anticoagulated AF patients and 12.36% (43/348) of major bleeding events were categorized as intermediate risk when stratified by the ORBIT score. By contrast, the HAS-BLED score categorized 27.59% (2229/8079) of anticoagulated AF patients and 28.45% (99/348) of major bleeding events as intermediate-risk (Figure 2).

In our pooled analysis, the intermediate-risk patients of ORBIT had a 1.73-fold increased risk of major bleeding events when compared with those of HAS-BLED (OR = 1.73; 95% CI: 1.19–2.52; p = 0.004;  $I^2 = 0\%$ ; Figure 3).

#### High risk category

The ORBIT score only categorized 4.39% (355/8079) of anticoagulated AF patients and 12.07% (42/348) of major bleeding events as high-risk. With HAS-BLED, 39.82% (3217/8079) of patients and 45.98% (160/348) of major bleeding events were categorized as high-risk (Figure 2).

Our pooled analysis showed that the high-risk patients of ORBIT had a 1.44-fold elevated risk of major bleeding events compared to HAS-BLED (OR = 1.44; 95% CI: 0.68–3.06; p = 0.34;  $l^2 = 64\%$ ; Figure 3).

#### Calibration analysis of the ORBIT score

Four studies [10, 12, 13, 15] were included in the calibration analysis. All the OR values were less than 1.0 across the three risk strata [low (OR = 0.64; 95% CI: 0.37–1.10; p = 0.11;  $I^2 = 85\%$ ); intermediate (OR = 0.63; 95% CI: 0.38–1.05; p = 0.08;  $I^2 = 3\%$ ); and high (OR = 0.64; 95% CI: 0.38–1.06; p = 0.08;  $I^2 = 0\%$ ] (Figure 4).

The ORBIT score under-predicted risk in low, intermediate and high strata, although there were no differences between the observed and predicted events in the three risk stratifications.

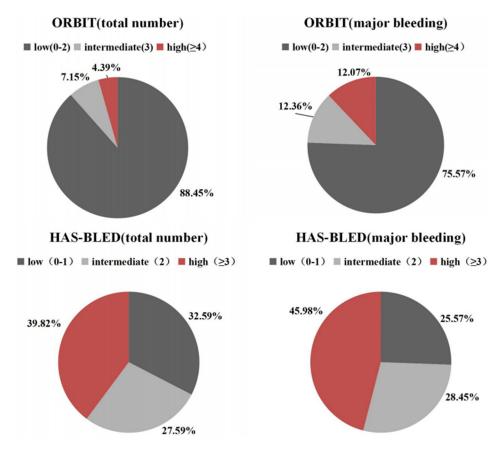


Figure 2: Population distribution in three strata for the ORBIT and HAS-BLED scores.

Senoo, K-2016 Subtotal (95% CI)	70 20 111 62 263	90 117 224 147	Events 37   7 4   41 89   8 (P = 0.66); 35   16 23   25 16	480 401 249 1503 <b>2633</b> I <sup>≠</sup> = 0% 322 493 781	40.3% 9.2% 6.9% 43.6% <b>100.0</b> % 31.5%	M-H, Random, 95% Cl 1.25 [0.82, 1.89] 1.13 [0.48, 2.70] 2.14 [0.78, 5.86] 1.08 [0.72, 1.61] 1.21 [0.93, 1.57] 1.51 [0.77, 2.95] 1.90 [0.76, 4.72] 2.20 [1.11, 4.34]	M-H, Random, 95% Cl
Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 1.38 (f <b>1.1.2 intermediate risk</b> Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	20 111 62 263 <sup>2</sup> = 1.59 P = 0.1 14 7 14	1012 3285 2106 <b>7146</b> 9, df = 3 7) 90 117 224 147	7 4 41 8 (P = 0.66); 35 16 23	401 249 1503 <b>2633</b> I <sup>2</sup> = 0% 322 493 781	9.2% 6.9% 43.6% <b>100.0%</b> 31.5% 17.0%	1.13 (0.48, 2.70) 2.14 (0.78, 5.86) 1.08 (0.72, 1.61) <b>1.21 (0.93, 1.57)</b> 1.51 (0.77, 2.95) 1.90 (0.76, 4.72)	
Esteve-Pastor, M A-2016 Proietti, M-2016 Subtotal (95% CI) Total events Heterogeneity, Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect. Z = 1.39 (f 1.1.2 intermediate risk Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity, Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect. Z = 2.86 (f	20 111 62 263 <sup>2</sup> = 1.59 P = 0.1 14 7 14	1012 3285 2106 <b>7146</b> 9, df = 3 7) 90 117 224 147	7 4 41 8 (P = 0.66); 35 16 23	401 249 1503 <b>2633</b> I <sup>2</sup> = 0% 322 493 781	9.2% 6.9% 43.6% <b>100.0%</b> 31.5% 17.0%	1.13 (0.48, 2.70) 2.14 (0.78, 5.86) 1.08 (0.72, 1.61) <b>1.21 (0.93, 1.57)</b> 1.51 (0.77, 2.95) 1.90 (0.76, 4.72)	
Projetti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 1.39 (f <b>1.1.2 intermediate risk</b> Caro Martínez C-2017 Esteve-Pastor, M A-2016 Projetti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	111 62 263 5 = 1.59 P = 0.1 14 7 14	3285 2106 <b>7146</b> 9, df = 3 7) 90 117 224 147	4 41 8 (P = 0.66); 35 16 23	249 1503 <b>2633</b> I <sup>2</sup> = 0% 322 493 781	6.9% 43.6% <b>100.0%</b> 31.5% 17.0%	2.14 (0.78, 5.86) 1.08 (0.72, 1.61) 1.21 (0.93, 1.57) 1.51 (0.77, 2.96) 1.50 (0.76, 4.72)	
Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 1.39 (f <b>1.1.2 intermediate risk</b> Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	62 263 9 = 0.1 14 7 14	2106 <b>7146</b> 9, df = 3 7) 90 117 224 147	41 89 8 (P = 0.66); 35 16 23	1503 2633 I <sup>≠</sup> = 0% 322 493 781	43.6% <b>100.0%</b> 31.5% 17.0%	1.08 [0.72, 1.61] <b>1.21 [0.93, 1.57]</b> 1.51 [0.77, 2.95] 1.90 [0.76, 4.72]	* * 
Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 1.39 (f <b>1.1.2 intermediate risk</b> Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	263 = 1.59 P = 0.1 14 7 14	<b>7146</b> 9, df = 3 7) 90 117 224 147	89 3 (P = 0.66); 35 16 23	2633   <sup>2</sup> = 0% 322 493 781	<b>100.0%</b> 31.5% 17.0%	<b>1.21 [0.93] 1.57]</b> 1.51 [0.77, 2.95] 1.90 [0.76, 4.72]	
Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 1.39 (f 1.12 intermediate risk Caro Martínez C-2017 Esteve-Pastor, M A-2016 Projetti, M-2016 Senoo, K-2016 Subtotal (95% Cl) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	<sup>2</sup> = 1.59 P = 0.1 14 7 14	9, df = 3 7) 90 117 224 147	3 (P = 0.66); 35 16 23	I <sup>≠</sup> = 0% 322 493 781	31.5% 17.0%	1.51 [0.77, 2.95] 1.90 [0.76, 4.72]	
Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 1.39 (f <b>1.1.2 intermediate risk</b> Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subtotal (95% Cl) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	<sup>2</sup> = 1.59 P = 0.1 14 7 14	90 117 224 147	3 (P = 0.66); 35 16 23	322 493 781	17.0%	1.90 [0.76, 4.72]	- <b>-</b>
Fest for overall effect Z = 1.39 (f 1.1.2 intermediate risk Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity, Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Fest for overall effect: Z = 2.86 (f	P = 0.1 14 7 14	90 117 224 147	35 16 23	322 493 781	17.0%	1.90 [0.76, 4.72]	- <b>-</b>
1.1.2 intermediate risk Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subitotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	14 7 14	90 117 224 147	16 23	493 781	17.0%	1.90 [0.76, 4.72]	
Caro Martínez C-2017 Esteve-Pastor, M A-2016 Proietti, M-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect Z = 2.86 (f	7 14	117 224 147	16 23	493 781	17.0%	1.90 [0.76, 4.72]	
Esteve-Pastor, M A-2016 Proietti, M-2016 Senoo, K-2016 Subitotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	7 14	117 224 147	16 23	493 781	17.0%	1.90 [0.76, 4.72]	+=
Proietti, M-2016 Senoo, K-2016 S <b>ubtotai (95% CI)</b> Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	14	224 147	23	781			
Senoo, K-2016 Subtotal (95% CI) Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f		147			30.4%	2.20 [1.11, 4.34]	
<b>Subtotal (95% CI)</b> Total events Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> Test for overall effect: Z = 2.86 (f	8		25				
Total events Heterogeneity: Tau² = 0.00; Chi² Test for overall effect: Z = 2.86 (F				633	21.1%	1.40 [0.62, 3.17]	
Heterogeneity: Tau² = 0.00; Chi² Test for overall effect: Z = 2.86 (F		578		2229	100.0%	1.73 [1.19, 2.52]	◆
Test for overall effect: Z = 2.86 (F	43		99				
	= 0.90	3, df = 3	3 (P = 0.82);	I <sup>z</sup> = 0%			
113 biab riek	P = 0.0	004)					
Caro Martínez C-2017	17	136	29	167	32.0%	0.68 [0.36, 1.30]	
Esteve-Pastor. M A-2016	19	147	23	382	32.2%	2.32 [1.22, 4.40]	_ <b></b>
Proietti. M-2016	2	42	100	2521	16.6%	1.21 [0.29, 5.08]	
Senoo, K-2016	4	30	8	147	19.1%	2.67 [0.75, 9.53]	
Subtotal (95% CI)		355		3217	100.0%	1.44 [0.68, 3.06]	
Total events	42		160			- / -	
Heterogeneity: Tau <sup>2</sup> = 0.35; Chi <sup>2</sup>	= 8.23	3. df = 3	3 (P = 0.04):	I <sup>≈</sup> = 64%			
Test for overall effect: Z = 0.95 (F							
						F	
						0.	01 0.1 1 10 1 ORBIT score HAS-BLED score

Figure 3: Comparison of major bleeding events in the 'low', 'intermediate' and 'high' risk categories based on the ORBIT and HAS-BLED scores. Abbreviations: CI, confidence interval; df, degrees of freedom; ORBIT, older age [ $\geq$  74 years], reduced hemoglobin/hematocrit/history of anemia, bleeding history, insufficient kidney function, and treatment with antiplatelet agents; HAS-BLED, hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile international normalized ratio [INR], elderly [age  $\geq$  65 years], and drugs/alcohol concomitantly; M-H, Mantel-Haenszel.

#### DISCUSSION

Based on our pooled C-statistic of continuous variables, both the ORBIT and HAS-BLED scores had a modest discriminatory capacity for predicting major bleeding risks. Although our pooled analysis showed that the ORBIT score had an elevated risk of major bleeding in all the three risk strata when compared with HAS-BLED, the P > 0.05 represented no significant differences in low and high stratifications. Compared with the HAS-BLED, more anticoagulated AF patients and major bleeding events were categorized as low risk when using ORBIT. Moreover, the ORBIT score under-predicted risk in all the three strata, but the p value > 0.05 represented that the ORBIT score was an adequate model calibration in predicting bleeding events.

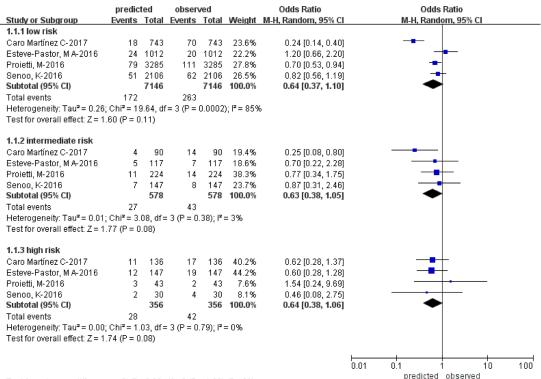
## Adding time in therapeutic range to the ORBIT scores

In recent studies of Vitamin K antagonist (VKA, e.g., warfarin) cohorts, the ORBIT score also underperformed in bleeding prediction compared to HAS-BLED, by not considering the 'labile INR criterion' [10, 11]. This is relevant as VKAs are still widely used as the preferred OAC worldwide. Indeed, the 'labile INR criterion', as measured by time in therapeutic ranges (TTR) [23, 24], is one of the most powerful criteria for bleeding prediction, and is incorporated within the HAS-BLED score.

Abumuaileq et al.found that the ORBIT score with TTR value (C-statistic = 0.74; 95% CI:0.70-0.76) had a higher C-statistic values than the original ORBIT (C-statistic = 0.66; 95% CI, 0.63-0.69) for major bleeding [14]. In the study by Senoo et al., the use of the ORBIT score with TTR values compared with the original ORBIT significantly reclassified +34.8% of the population for major bleeding [11]. According to the study of Proietti et al., when compared with the original ORBIT score with TTR values significantly reclassified +25.08% of the population for major bleeding [10]. Thus, the predictive ability of the ORBIT score with TTR values was improved compared to that of the original ORBIT score.

## Net reclassification improvement (NRI) analysis of the ORBIT score

Nonetheless, a C-statistic cannot correctly and adequately assess the diagnostic ability, and other parameters must be added [25]. The net reclassification improvement (NRI) is a metric that helps us to understand the improvement between event and non-event categories. When a choice of care is presented above or below the given cut-off, the NRI value is better than the C-statistic when evaluating the predictive ability [25, 26]. Three included studies in our meta-analysis used the NRI. For example, Abumuaileq et al. [14] and Proietti et al. [10] found that the ORBIT score reclassified -1.85% and



Test for subaroup differences: Chi² = 0.00, df = 2 (P = 1.00), l² = 0%

#### Figure 4: Forest plot showing calibration analysis of the ORBIT score.

-0.77% of the population for major bleeding (p = 0.9 and 0.392, respectively) when compared with the HAS-BLED. Senoo et al. [13] reported a positive reclassification for any clinically relevant bleeding (NRI: +15.6%, p = 0.007) with the HAS-BLED score when compared with the ORBIT. A significant positive NRI value indicated that the HAS-BLED score performed better than the ORBIT in predicting bleeding events. It was not possible to complete a pooled analysis because of the limited NRI data.

#### **ORBIT versus HAS-BLED score**

Since the HAS-BLED score was first described in 2010, it has been validated in various populations, including AF and non-AF cohorts and various drug treatments (e.g., aspirin, VKA [27] and non-VKA anticoagulants [28-30]) and no antithrombotic therapy cohorts. The HAS-BLED score performs better than other scores in predicting major bleeding and ICH in an Asian/Chinese AF population [31]. For the ORBIT score, O'Brien and colleagues chose the five strongest predictors from a full continuous ORBIT bleeding model, which included numerous independent predictors of major bleeding [8]. In our pooled analysis, we found that this new score categorized more anticoagulated AF patients and major bleeding events as low risk when compared with the HAS-BLED score. And in clinical practice, the physician always pay more attention to the 'high risk' anticoagulated AF patients, they might timely adjust the dose of anticoagulant or even discontinue anticoagulant for the patients. In contrast, the 'low risk' anticoagulated AF patients might get less attention. Unfortunately, the ORBIT score did not meet expectations for evaluating bleeding risks, and several studies have found that the ORBIT only shows a modest or even worse predictive ability when compared with other existing scores [10–15].

#### Limitations

Several evident limitations are in our metaanalysis. First, heterogeneity always exists in three main forms: clinical, methodological and statistical [32], and the inconsistency might have derived from the limited amount of data, variety in study designs and the various types of anticoagulants used (e.g., warfarin, dabigatran, rivaroxaban, and apixaban). Second, the types of included studies were prospective or retrospective studies, and large prospective studies are necessary to validate our findings.

## CONCLUSIONS

In conclusion, our systematic review and metaanalysis showed that the ORBIT score does not superior than the HAS-BLED in predicting major bleeding events of anticoagulated AF patients. Furthermore, when compared with the HAS-BLED score, this new bleeding risk score categorized more anticoagulated AF patients and major bleeding events as low risk.

### **Author contributions**

Hong-K was responsible for the entire project, and Lip G Y revised the draft. Wang-C and Yu-Y performed the systematic literature review and drafted the first version of the manuscript. All authors took part in the interpretation of the results and prepared the final version.

## **CONFLICTS OF INTEREST**

No conflicts of interest exist.

## FUNDING

The authors wish to acknowledge support from the National Natural Science Foundation of China [81530013] and the National Basic Research Program of China [973 Program: 2013CB531100].

## REFERENCES

- January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JJ, Conti JB, Ellinor PT, Ezekowitz MD, Field ME, Murray KT, Sacco RL, Stevenson WG, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society. Circulation. 2014; 130:2071–104. https:// doi.org/10.1161/CIR.000000000000040.
- Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener HC, Heidbuchel H, Hendriks J, Hindricks G, Manolis AS, Oldgren J, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Europace. 2016; 18:1609–78. https://doi.org/10.1093/europace/euw295.
- De Caterina R, Husted S, Wallentin L, Andreotti F, Arnesen H, Bachmann F, Baigent C, Huber K, Jespersen J, Kristensen SD, Lip GY, Morais J, Rasmussen LH, et al. Vitamin K antagonists in heart disease: current status and perspectives (Section III). Position paper of the ESC Working Group on Thrombosis—Task Force on Anticoagulants in Heart Disease. Thromb Haemost. 2013; 110:1087–107. https://doi.org/10.1160/TH13-06-0443.
- Singer DE, Chang Y, Fang MC, Borowsky LH, Pomernacki NK, Udaltsova N, Go AS. The net clinical benefit of warfarin anticoagulation in atrial fibrillation. Ann Intern Med. 2009; 151:297–305. https://doi. org/10.7326/0003-4819-151-5-200909010-00003.
- 5. Lip GY, Andreotti F, Fauchier L, Huber K, Hylek E, Knight E, Lane D, Levi M, Marin F, Palareti G, Kirchhof P. 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. https://doi.org/10.1160/TH11-10-0690.

- Lip GY, Lane DA. Bleeding risk assessment in atrial fibrillation: observations on the use and misuse of bleeding risk scores. J Thromb Haemost. 2016; 14:1711–14. https:// doi.org/10.1111/jth.13386.
- Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. 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. https://doi.org/10.1378/chest.10-0134.
- O'Brien EC, Simon DN, Thomas LE, Hylek EM, Gersh BJ, Ansell JE, Kowey PR, Mahaffey KW, Chang P, Fonarow GC, Pencina MJ, Piccini JP, Peterson ED. The ORBIT bleeding score: a simple bedside score to assess bleeding risk in atrial fibrillation. Eur Heart J. 2015; 36:3258–64. https://doi.org/10.1093/eurheartj/ehv476.
- Cullen MW, Kim S, Piccini JS, Ansell JE, Fonarow GC, Hylek EM, Singer DE, Mahaffey KW, Kowey PR, Thomas L, Go AS, Lopes RD, Chang P, et al. Risks and benefits of anticoagulation in atrial fibrillation: insights from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF) registry. Circ Cardiovasc Qual Outcomes. 2013; 6:461–69. https://doi.org/10.1161/ CIRCOUTCOMES.113.000127.
- Proietti M, Senoo K, Lane DA, Lip GY. 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. https://doi. org/10.1038/srep24376.
- Senoo K, Proietti M, Lane DA, Lip GY. Evaluation of the HAS-BLED, ATRIA, and ORBIT Bleeding Risk Scores in Patients with Atrial Fibrillation Taking Warfarin. Am J Med. 2016; 129:600–07. https://doi.org/10.1016/j. amjmed.2015.10.001.
- Esteve-Pastor MA, Garcia-Fernandez A, Macias M, Sogorb F, Valdes M, Roldan V, Muniz J, Badimon L, Roldan I, Bertomeu-Martinez V, Cequier A, Lip GY, Anguita M, et al. Is the ORBIT Bleeding Risk Score Superior to the HAS-BLED Score in Anticoagulated Atrial Fibrillation Patients? Circ J. 2016; 80:2102–08. https://doi. org/10.1253/circj.CJ-16-0471.
- Senoo K, Lip GY. Predictive abilities of the HAS-BLED and ORBIT bleeding risk scores in non-warfarin anticoagulated atrial fibrillation patients: An ancillary analysis from the AMADEUS trial. Int J Cardiol. 2016; 221:379–82. https:// doi.org/10.1016/j.ijcard.2016.07.100.
- Abumuaileq RR, Abu-Assi E, Raposeiras-Roubin S, Rodriguez-Manero M, Pena-Gil C, Gonzalez-Juanatey JR. Comparison Between 3 Bleeding Scoring Systems in Nonvalvular Atrial Fibrillation Patients. What Can the New ORBIT Score Provide? Rev Esp Cardiol (Engl Ed). 2016; 69: 1112–14. https://doi.org/10.1016/j.rec.2016.05.022.
- 15. Caro MC, Andreu CJ, Flores BP, Valdes M, Bailen LJ, Manzano FS. Comparison of Bleeding Risk Scores in

Patients With Nonvalvular Atrial Fibrillation Starting Direct Oral Anticoagulants. Rev Esp Cardiol (Engl Ed). 2017. https://doi.org/10.1016/j.rec.2017.01.021.

- Zhu W, He W, Guo L, Wang X, Hong K. The HAS-BLED Score for Predicting Major Bleeding Risk in Anticoagulated Patients With Atrial Fibrillation: A Systematic Review and Meta-analysis. Clin Cardiol. 2015; 38:555–61. https://doi. org/10.1002/clc.22435.
- Wells GA, Shea BJ, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of Non-Randomized Studies in Meta-Analysis. Applied Engineering in Agriculture. 2000; 18:727–34. https://doi.org/10.2307/632432.
- Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007; 115:928–35. https:// doi.org/10.1161/CIRCULATIONAHA.106.672402.
- Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics. 2005; 21:3940–41. https://doi.org/10.1093/bioinformatics/ bti623.
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011; 12:77. https://doi.org/10.1186/1471-2105-12-77.
- 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. https://doi.org/10.2307/2531595.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003; 327:557–60. https://doi.org/10.1136/bmj.327.7414.557.
- Wan Y, Heneghan C, Perera R, Roberts N, Hollowell J, Glasziou P, Bankhead C, Xu Y. Anticoagulation control and prediction of adverse events in patients with atrial fibrillation: a systematic review. Circ Cardiovasc Qual Outcomes. 2008; 1:84–91. https://doi.org/10.1038/srep31755.
- Viquez-Jaikel A, Victoria H, Ramos-Esquivel A. Improvement of time in therapeutic range with warfarin by pharmaceutical intervention. Int J Clin Pharm. 2017; 39:41–44. https://doi. org/10.1161/CIRCOUTCOMES.108.796185.
- 25. Pencina MJ, D'Agostino RS, D'Agostino RJ, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008; 27:157–72, 207–12. https://doi. org/10.1007/s11096-016-0403-5.
- 26. Corbanese U. Assessing the performance of the HAS-BLED score: is the C statistic sufficient? Chest. 2011; 139:1247–48. https://doi.org/10.1002/sim.2929.
- Roldan V, Marin F, Fernandez H, Manzano-Fernandez S, Gallego P, Valdes M, Vicente V, Lip GY. Predictive value of the HAS-BLED and ATRIA bleeding scores for the risk of serious bleeding in a "real-world" population with atrial fibrillation receiving anticoagulant therapy. Chest. 2013; 143:179–84. https://doi.org/10.1378/chest.10-2995.

- 28. Apostolakis S, Lane DA, Guo Y, Buller H, Lip GY. 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; 60:861–67. https://doi.org/10.1378/chest.12-0608.
- 29. Apostolakis S, Lane DA, Guo Y, Buller H, Lip GY. 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; 61:386–87. https://doi.org/10.1016/j. jacc.2012.06.019.
- Lopes RD, Al-Khatib SM, Wallentin L, Yang H, Ansell J, Bahit MC, De Caterina R, Dorian P, Easton JD, Erol C, Ezekowitz JA, Gersh BJ, Granger CB, et al. Efficacy and

safety of apixaban compared with warfarin according to patient risk of stroke and of bleeding in atrial fibrillation: a secondary analysis of a randomised controlled trial. Lancet. 2012; 380:1749–58. https://doi.org/10.1016/j. jacc.2012.10.010.

- 31. Guo YT, Zhang Y, Shi XM, Shan ZL, Wang CJ, Wang YT, Chen YD, Lip GY. Assessing bleeding risk in 4824 Asian patients with atrial fibrillation: The Beijing PLA Hospital Atrial Fibrillation Project. Sci Rep. 2016; 6:31755. https:// doi.org/10.1016/S0140-6736(12)60986-6.
- Melsen WG, Bootsma MC, Rovers MM, Bonten MJ. The effects of clinical and statistical heterogeneity on the predictive values of results from meta-analyses. Clin Microbiol Infect. 2014; 20:123–29. https://doi.org/10.1111/1469-0691.12494.