Reliability of the intracerebral hemorrhage score for predicting outcome in patients with intracerebral hemorrhage using oral anticoagulants

Fakiri, M. O.; Uyttenboogaart, M.; Houben, R.; van Oostenbrugge, R. J.; Staals, J.; Luijckx, G. J.

Published in:
European Journal of Neurology

DOI:
10.1111/ene.14336

ORIGINAL ARTICLE

Reliability of the intracerebral hemorrhage score for predicting outcome in patients with intracerebral hemorrhage using oral anticoagulants

M. O. Fakiri[a], M. Uyttenboogaart[a], R. Houben[b], R. J. van Oostenbrugge[b], J. Staals[b] and G.J. Luijckx[a]

[a]*Department of Neurology, University Medical Center Groningen, Groningen; and* [b]*Department of Neurology, Maastricht University Medical Centre, Maastricht, The Netherlands*

Keywords:
intracerebral hemorrhage score, intracerebral hemorrhage, oral anticoagulants, outcome

Received 17 April 2020
revision requested 12 May 2020
Accepted 13 May 2020

European Journal of Neurology 2020, **0:** 1–8

doi:10.1111/ene.14336

Background and purpose: The intracerebral hemorrhage (ICH) score is the most widely used and validated prognostic model for estimating 30-day mortality in ICH. However, the score was developed and validated in an ICH population probably not using oral anticoagulants (OACs). The aim of this study was to determine the performance of the ICH score for predicting the 30-day mortality rate in the full range of ICH scores in patients using OACs.
Methods: Data from admitted patients with ICH were collected retrospectively in two Dutch comprehensive stroke centers. The validity of the ICH score was evaluated by assessing both discrimination and calibration in OAC and OAC-naive patient groups.
Results: A total of 1752 patients were included of which 462 (26%) patients were on OAC. The 30-day mortality was 54% for the OAC cohort and 34% for the OAC-naive cohort. The 30-day mortality was higher in the OAC cohort for ICH score 1 (33% vs. 12.5%; odds ratio, 3.4; 95% confidence intervals, 1.1–10.4) and ICH score 2 (53% vs. 26%; odds ratio, 3.2; 95% confidence intervals, 1.2–8.2) compared with the predicted mortality rate of the original ICH score. Overall, the discriminative ability of the ICH score was equally good in both cohorts (area under the curve 0.83 vs. 0.87, respectively).
Conclusions: The ICH score underestimated the 30-day mortality rate for lower ICH scores in OAC-ICH. When estimating the prognosis of ICH in patients using OAC, this underestimation of mortality must be taken into account.

Introduction

Intracerebral hemorrhage (ICH) is a serious complication of oral anticoagulant (OAC) therapy. Approximately 5–25% of all ICH is related to OAC [1]. Previous studies revealed a 90-day mortality rate of up to 60%, which is much higher than the 30–40% mortality rates observed in OAC-naive ICH [2,3].

Several prognostic models have been developed to aid in decision making in the acute setting [4–8]. The ICH score is the most widely used and validated prognostic model for estimating the 30-day mortality rate [6]. The ICH score (0–6) is the sum of individual points assigned to five different variables: Glasgow Coma Scale (GCS) score 3–4 (2 points) or 5–12 (1 point); age ≥ 80 years (1 point); infratentorial origin (1 point); ICH volume ≥ 30 mL (1 point); and intraventricular hemorrhage (1 point). However, a recent study has shown that the ICH score might underestimate the mortality rates for lower scores (≤3) in patients using OAC [5]. The original score was developed and validated in a more general population of patients with ICH and not for patients using OAC. The population was composed of patients with non-traumatic ICH, including 'primary' and other presumed underlying structural causes such as arteriovenous malformations (AVMs).

Correspondence: M. O. Fakiri, University Medical Center Groningen, Hanzeplein 1, Postbus 30.001, 9700 RB Groningen, the Netherlands (tel.: 031614949989; fax: 031503614227; e-mail: omeed.fakiri@gmail.com).

EUROPEAN JOURNAL OF NEUROLOGY

The use of prognostic models with regard to treatment of acute ICH can lead to self-fulfilling prophesies by withdrawing medical care in patients with a high ICH score [9]. Although an active treatment leads to lower mortality after ICH [10], it is very important to predict prognosis as accurately as possible in these situations, especially with regard to mortality.

Therefore, the aim of this study was to determine the performance of the ICH score for predicting the 30-day mortality rate in the full range of ICH scores in an OAC-ICH cohort. Because of the difference in bleeding rates between direct anticoagulant (DOAC)- and vitamin K antagonist (VKA)-treated patients, a subgroup analysis was performed for patients treated with a DOAC if the data were sufficient [11].

Patient selection

Data were retrospectively collected from an ICH database from the neurology departments of University Medical Center Groningen (UMCG) and Maastricht University Medical Center (MUMC) in the Netherlands. Data of both centers were pooled.

Patient selection at University Medical Center Groningen

The study population included all consecutive patients ≥ 18 years of age, admitted to the emergency department between January 2008 and December 2017.

Patients were included if brain computed tomography confirmed the diagnosis of non-traumatic intraparenchymal ICH. Patients were excluded for secondary causes of ICH as judged by the treating physicians, presentation > 24 h after the ictus, no available GCS score or lack of information on clinical outcome.

Patient selection at Maastricht University Medical Center

All consecutive adult (≥18 years) patients with an imaging-confirmed non-traumatic ICH, seen in the emergency department, inpatient or outpatient clinic in three regional hospitals in the Maastricht area from January 2004 to December 2009 were included. Exclusion criteria were secondary causes of ICH as judged by the treating physicians or non-accessible outcome.

Materials and methods

For both of the cohorts, the demographic data, vascular risk factors (hypertension and diabetes mellitus), use of OAC, international normalized ratio on admission, any neurosurgical intervention and variables necessary to calculate the ICH score were collected. Hypertension and diabetes mellitus were defined as being known from medical history or use of medication. The GCS score was determined on admission to the emergency department. In intubated patients, the first reliable recorded GCS score was used from the mobile medical team in the field. The intracerebral hematoma parameters were measured on the baseline non-contrast computed tomography scan. The hematoma volume was measured, by the local investigators, using the ABC/2 method [12]. The primary endpoint was 30-day mortality.

Survival data were obtained from the hospital registry system. Information on the outcome was requested by letter or telephone from other institutes when patients were transferred within 30 days and did not have a recorded date of death in our hospital registry system.

Statistical analysis

In the statistical analyses, continuous data are represented as means and SD or median with first and third quartiles (Q1–Q3), and categorical data as frequency and percentage. We compared categorical variables using Pearson's chi-squared tests and Fisher's exact tests and continuous data with Mann–Whitney *U*-test.

The 30-day mortality rate was separately determined for both the OAC and OAC-naive cohorts and compared with the model recommended by Hemphill *et al*. [6]. The difference between the various cohorts was analyzed using Pearson's chi-squared tests and Fisher's exact tests.

Reliability of the model for predicting the outcome was investigated with calibration curves and Hosmer–Lemeshow goodness-of-fit test [13–15]. Calibration refers to the agreement between the observed outcome frequencies and probabilities predicted by the model. The plot gives a visual impression of the accuracy of prediction in the whole range of probabilities. The ideal calibration follows a line with an angle of 45°. Discrimination, the ability of the model to accurately classify between dead and alive, was expressed as the area under the receiver operating characteristic curve (AUC). An AUC > 0.8 is considered to reflect good discrimination [13,14,16]. The different AUC values were compared with the method described by Hanley *et al*. [17].

To determine if the variables in the ICH score were independent predictors for 30-day mortality, we performed univariable and multivariable analysis by backward stepwise logistic regression analysis for both the OAC and OAC-naive cohorts. The 30-day

mortality was used as outcome variable and the variables in the ICH score as predictors, odds ratio (OR) and 95% confidence intervals (CI) were calculated. Variables differing significantly in univariable analysis were included in multivariable models of backward stepwise logistic regression analysis. As all of the patients with a GCS score 3–4 in the OAC-ICH group died, we applied a dummy case for this variable to correct for the extreme large OR.

All statistical analyses were performed using (SPSS (version 23.0, Armonk, NY, USA)) and Excel 2010 (Microsoft). A two-sided $P < 0.05$ was considered to be statistically significant.

At the time, the approval of an ethics board was not needed for the UMCG cohort. All of the data that were used were collected during standard medical care. For the MUMC cohort, the ethical committee of MUMC had given its approval in the study performed by Houben *et al*. [5].

Results

A total of 1752 patients were included, 1247 patients from MUMC and 505 patients from UMCG. There were 462 (26%) patients using OAC at admission, 284 from MUMC and 178 from UMCG. Only eight patients used a DOAC, of which three died (one with ICH score 1, one with ICH score 3 and one with ICH score 4). Because of the small numbers, it is not possible to perform a reliable subgroup analysis.

Baseline characteristics are presented in Table 1. Patients using OAC had statistically significantly greater age, higher incidence of diabetes mellitus, hypertension, GCS score 3–4, infratentorial localization, ventricular extension and larger hematoma volumes. The 30-day mortality rate was higher in the OAC cohort in comparison to the OAC-naive cohort (54% vs. 34%; OR, 2.3; 95% CI, 1.8–2.8).

Figure 1 shows the 30-day mortality rate for each ICH score for the OAC and OAC-naive cohorts and the predicted mortality in the original ICH score. The 30-day mortality was higher in the OAC cohort for ICH score 1 (33% vs. 11%; OR, 4.1; 95% CI, 2.4–6.8) and ICH score 2 (53% vs. 36%; OR, 2.0; 95% CI, 1.2–3.2) in comparison to the OAC-naive cohort. There was no difference in 30-day mortality for ICH score ≥ 3 between the cohorts. The 30-day mortality rates for ICH score 1 (33% vs. 12.5%; OR, 3.4; 95% CI, 1.1–10.4) and ICH score 2 (53% vs. 26%; OR, 3.2; 95% CI, 1.2–8.2) in the OAC cohort were higher than the predicted mortality rate of the original ICH score. In the OAC-naive cohort, mortality rates did not differ significantly from the predicted mortality. These results are also shown in the calibration plots (Fig. 2). The Hosmer–Lemeshow test correlation coefficient between observed and predicted probability of death or survival was 54.46 for the OAC-naive cohort with a significance level of <0.001 and 90.88 for the OAC cohort with a significance level of <0.001 (Tables S1 and S2).

The receiver operating characteristic curve is presented in Fig. 3. The AUC was 0.83 (95% CI, 0.79–0.87) in the OAC cohort and 0.87 (95% CI, 0.85–0.89) in the OAC-naive cohort. The AUC for the original ICH score was 0.92 (95% CI, 0.88–0.96). The AUC in the OAC cohort was lower compared with the original ICH score ($P = 0.003$). The AUC from the OAC-ICH was also lower compared with the OAC-naive ICH, but this difference was not significant ($P = 0.07$).

Table 2 shows the association between the different variables in the ICH score and 30-day mortality. All variables showed a significant association with 30-day mortality in the OAC-naive cohort. In the OAC cohort, there was no significant association between infratentorial localization and 30-day mortality in the univariable analysis (OR, 1.02; 95% CI, 0.64–1.6). Furthermore, the weighing of individual predictors differed in multivariable analysis in the OAC cohort in comparison to the OAC-naive cohort.

Discussion

The ICH score underestimates the 30-day mortality rate in the lower range of ICH scores (<3) in OAC-ICH in comparison to OAC-naive ICH. However, the model is a reliable predictor of mortality in OAC-naive patients with ICH.

The main reason for this may be that the original ICH score was developed and validated in a population of patients with ICH who were probably not using OAC. In the original publication of the ICH score (Hemphill *et al*. [6]), use of OAC is not mentioned under presumed causes of ICH (although, for example, use of drugs is mentioned). Also, the original ICH score population was composed of patients with non-traumatic ICH, including other presumed causes such as AVMs, with known lower mortality [18]. However, this seems negligible given that there were no major differences in the 30-day mortality between the OAC-naive cohort and the original ICH score.

The findings of our study are consistent with a previous report [4]. Despite the significant difference in the AUC between the original ICH and the OAC-ICH, the overall discriminative ability of the ICH score is equally good in the OAC and OAC-naive cohorts. This means that the model has a good ability to globally predict mortality. With each point

Table 1 Baseline characteristics

	OAC cohort (1) (n = 462)	OAC-naive cohort (2) (n = 1290)	Total population (n = 1752)	P-values (1 vs. 2)
Sex, male	254 (55)	679 (50)	933 (53)	0.39
Median age (years) (Q1–Q3)	77 (70–83)	73 (61–80)	75 (64–82)	**<0.001**
Age ≥ 80 years	181 (39)	350 (27)	531 (30)	**<0.001**
DM	93 (20)	171 (13)	264 (15)	**<0.001**
Hypertension	336 (74)	670 (52)	1006 (57)	**<0.001**
Median GCS score (Q1–Q3)	13 (8–15)	13 (10–15)	13 (10–15)	**0.004**
GCS score 3–4	62 (13)	110 (9)	172 (10)	**0.002**
GCS score 5–12	160 (35)	436 (34)	596 (34)	0.75
GCS score 13–15	240 (52)	744 (58)	984 (56)	**0.033**
Infratentorial	93 (20)	176 (14)	269 (15)	**0.001**
Median hematoma volume (mL) (Q1–Q3)	19 (6–52)	13 (4–36)	14 (5–40)	**<0.001**
Volume ≥ 30 mL	182 (39)	395 (31)	577 (33)	**0.001**
Ventricular extension	228 (49)	530 (41)	758 (43)	**0.002**
Median INR (IQR)	3.3 (1.5–4.3)	1.0	3.3 (2.5–4.30)	
External ventricular drain	20 (4)	46 (4)	66 (4)	0.56
Hematoma evacuation	12 (3)	42 (3)	54 (3)	0.46
Both	7 (2)	15 (1)	22 (1)	0.62
30-day mortality	248 (54)	433 (34)	681 (39)	**<0.001**

DM, diabetes mellitus; GCS, Glasgow Coma Scale; INR, international normalized ratio; IQR, interquartile range; OAC, oral anticoagulant; Q1–Q3, first and third quartile. Data are given as n (%) unless otherwise stated. Bold values indicates a statistically significant difference with a P-value less than 0.05.

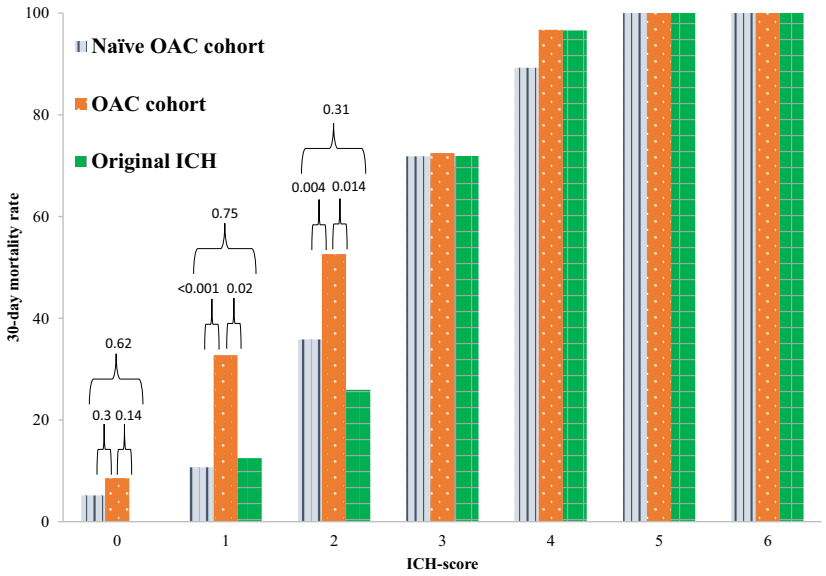

Figure 1 Comparison of the 30-day mortality risk between the oral anticoagulant (OAC) and OAC-naive cohorts and the original intracerebral hemorrhage (ICH) score (showing P-values).

increase, mortality increases linearly and almost all patients with an ICH score of ≥3 die. However, the receiver operating characteristic curve does not take into account the reliability of the prediction for ICH scores separately. Although useful for classification, evaluation of prognostic models should not rely solely on the receiver operating characteristic curve, but should assess both discrimination and calibration [19]. The Hosmer–Lemeshow goodness-of-fit test showed a poor calibration in both cohorts. However, the Hosmer–Lemeshow test has its limitations. One limitation is that it is for overall calibration error, not for any

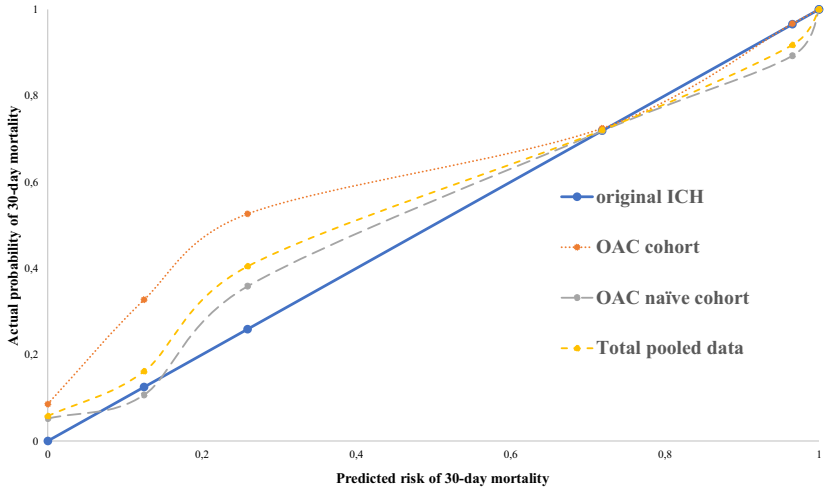

Figure 2 Calibration curves in the intracerebral hemorrhage score predicting 30-day mortality rate. OAC, oral anticoagulant.

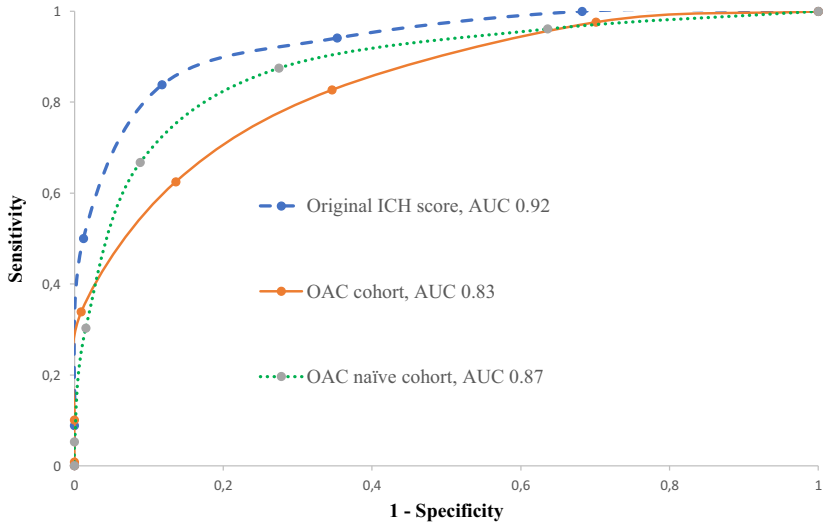

Figure 3 Receiver operating characteristic curve for intracerebral hemorrhage scores. AUC, area under the curve; OAC, oral anticoagulant.

particular lack of fit. As shown in Fig. 1, there are some differences between the observed and predicted outcomes in the OAC-naive cohort, but these are not statistically significant for the individual ICH score. As such, in our opinion, the model remains a useful prognostic test in the OAC-naive population.

In OAC-ICH, the extremely poor calibration in the lower ICH scores can be explained by the fact that infratentorial location was not an independent predictor for mortality in OAC-ICH. Despite a subgroup analysis of the data, we have not found a good explanation for this finding. Furthermore, the weighing of individual predictors differs in the OAC-ICH and OAC-naive cohorts.

The natural history of OAC-ICH differs from that of OAC-naive ICH. In our cohort of almost 1800 patients, the difference in baseline characteristics was significant for almost all of the variables. The overall mortality in OAC-ICH is twice as high as in OAC-naive ICH, similar to the literature [2,3,5]. One of the main reasons for a higher mortality in OAC-ICH could be secondary hematoma growth. We did not systematically assess hematoma growth for all patients. Previous studies have shown that hematoma growth occurs more often in OAC-ICH than in OAC-naive ICH and is associated with a higher mortality [2,20]. Furthermore, greater age and comorbidities such as hypertension and diabetes mellitus are

Table 2 Univariable and multivariable logistic regression analysis for variables predicting 30-day mortality

	30-day mortality					
	Univariable analysis			Multivariable analysis		
	OR	95% CI	*P*-value	OR	95% CI	*P*-value
OAC-naive cohort						
Age > 80 years	1.86	1.45–2.40	<0.001	4.47	2.02–9.92	<0.001
GCS score						
GCS score 13–15	1.0	Reference		1.0		
GCS score 5–12	18.52	9.52–36.01	<0.001	10.6	5.11–22.01	<0.001
GCS score 3–4	136.7	37.27–501.10	<0.001	56.52	14.01–227.90	<0.001
Infratentorial	1.95	1.41–2.67	<0.001	4.01	1.63–9.91	0.003
ICH volume > 30 mL	7.21	5.54–9.37	<0.001	5.78	2.60–12.86	<0.001
Ventricular extension	6.05	4.69–7.79	<0.001	2.11	1.02–4.34	0.043
OAC cohort						
Age > 80 years	1.74	1.20–2.55	0.004	2.01	1.26–3.22	0.003
GCS score						
GCS score 13–15	1.0	Reference		1.0		
GCS score 5–12	4.98	3.23–7.68	<0.001	2.53	1.54–4.16	<0.001
GCS score 3–4	136.4	18.56–1002.00	<0.001	62.98	8.34–475	<0.001
Infratentorial	1.02	0.64–1.60	0.966			
ICH volume > 30 mL	7.11	4.59–11.03	<0.001	3.46	2.07–5.80	<0.001
Ventricular extension	5.86	3.94–8.83	<0.001	3.13	1.96–4.99	<0.001

CI, confidence intervals; GCS, Glasgow Coma Scale; ICH, intracerebral hemorrhage; OAC, oral anticoagulant; OR, odds ratio.

associated with a higher chance of mortality in OAC-ICH [2,4,21]. It is likely that secondary hematoma expansion has more influence on the outcome in low-risk than in high-risk patients (ICH score ≥ 3), as in the latter group the mortality risk is already very high, due to the combination of three or more independent major risk factors (≥80 years of age, hematoma volume ≥ 30 mL, intraventricular hemorrhage, infratentorial origin of hemorrhage and a GCS score between 3 and 12).

Due to the above results, we believe that the lower ICH score is not a reliable prognostic model for predicting the 30-day mortality in OAC-ICH. In a previous study, Houben *et al.* showed that adding 1 point for OAC use to the existing ICH score does not improve the prognostic performance of this score in the overall group [5].

An important strength of this study is the large cohort that consisted of 462 patients with OAC use and 1290 OAC-naive patients. To our knowledge, the only other study that systematically examined the performance of the ICH score in the lower ICH scores in OAC-ICH consisted of 170 subjects, 33 of which used OAC [4]. The original ICH study by Hemphill *et al.* included a total of 152 patients with ICH of diverse underlying causes. Furthermore, the multicenter design makes our study generalizable to daily practice [6].

Nonetheless, our study also has limitations, partly inherent to its retrospective design. Firstly, not all clinical data were completely recorded. Second, there are imbalances in both subpopulation characteristics (Table S3). This could result in heterogeneity in variables of the pooled data affecting the outcome. An explanation for this could be the different time window in which patients were included as well as differences in the inclusion criteria such as inclusion from the outpatient clinic in the MUMC cohort. Furthermore, the ABC/2 scores for both sites were measured by different investigators, which could have influenced the difference in hematoma volumes. The difference regarding primary presentation in the UMCG versus inclusion of patients from regional hospitals in the MUMC cohort probably has little influence, as both hospitals serve a large regional tertiary stroke center. Third, the timing of the GCS evaluation differs between our study and the original article by Hemphill *et al.* [6] in which the GCS score was evaluated at the time of transfer from the emergency department, whereas in this study the GCS score was determined on admission to the emergency department or, if the patient was intubated, the first reliable recorded GCS score was used from the mobile medical team in the field. Different evaluation moments could lead to underestimation of the ICH score in patients who deteriorate during their stay in the emergency department. We chose this moment because, in our center, the decision about invasive treatment is often based on the best ICH and GCS scores, which are often at the time of admission to the emergency department. Fourth, most study subjects used a VKA and

therefore our results cannot be generalized to patients with DOAC-associated ICH. In a recently performed large systematic review and individual patient data meta-analysis of cohort studies comparing DOAC-ICH and VKA-ICH, patients with DOAC-ICH had smaller baseline hematoma volumes and less severe acute stroke syndromes. In this study, the functional outcome was the same at discharge, 1 month or 3 months [20]. Nevertheless, the results of this study should be used with caution in patients with DOAC-associated ICH. Fifth, although DOACs now account for 30–40% of all anticoagulant prescriptions, the majority of the patients are still on VKA [11]. Therefore, the results of this study are still applicable.

In general, the ICH score has good discriminative ability for prediction of mortality in OAC-ICH but underestimates 30-day mortality rate at lower scores. Not all variables of the ICH score are independent predictors or have the same magnitude of association in OAC-ICH in comparison to OAC-naive ICH. Therefore, the use of the ICH score should be applied with caution especially in the lower score range in patients with OAC-ICH.

Disclosure of conflicts of interest

The authors declare no financial or other conflicts of interest.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Performance of the intracerebral hemorrhage (ICH) score: Goodness of fit of general ICH score model using Hosmer–Lemeshow chi-squared statistic test (oral anticoagulant cohort).

Table S2 Performance of the intracerebral hemorrhage (ICH) score: Goodness of fit of general ICH score model using Hosmer–Lemeshow chi-squared statistic test (oral anticoagulant-naive cohort).

Table S3 Baseline characteristics.

References

1. Schols AMR, Schreuder FHBM, van Raak EPM, *et al*. Incidence of oral anticoagulant-associated intracerebral hemorrhage in the Netherlands. *Stroke* 2014; **45**(1)**:** 268–270.
2. Cucchiara B, Messe S, Sansing L, Kasner S, Lyden P. Hematoma growth in oral anticoagulant related intracerebral hemorrhage. *Stroke* 2008; **39**(11)**:** 2993–2996.
3. Cervera Á, Amaro S, Chamorro Á. Oral anticoagulant-associated intracerebral hemorrhage. *J Neurol* 2012; **259**(2)**:** 212–224.
4. Masotti L, Di Napoli M, Godoy DA, *et al*. Intracerebral hemorrhage score in patients with spontaneous intracerebral hemorrhage pretreated and not treated with antithrombotics. *Neurol Clin Neurosci* 2016; **4**(5)**:** 169–175.
5. Houben R, Schreuder FHBM, Bekelaar KJ, Claessens D, van Oostenbrugge RJ, Staals J. Predicting prognosis of intracerebral hemorrhage (ICH): performance of ICH score is not improved by adding oral anticoagulant use. *Front Neurol* 2018; **9:** 100.
6. Hemphill JC, Bonovich DC, Besmertis L, Manley GT, Johnston SC. The ICH score. *Stroke* 2001; **32**(4)**:** 891–897.
7. Ruiz-Sandoval JL, Chiquete E, Romero-Vargas S, Padilla-Martínez JJ, González-Cornejo S. Grading scale for prediction of outcome in primary intracerebral hemorrhages. *Stroke* 2007; **38**(5)**:** 1641–1644.
8. Gregório T, Pipa S, Cavaleiro P, *et al*. Assessment and comparison of the four most extensively validated prognostic scales for intracerebral hemorrhage: systematic review with meta-analysis. *Neurocrit Care* 2019; **30**(2)**:** 449–466.
9. McCracken DJ, Lovasik BP, McCracken CE, *et al*. The intracerebral hemorrhage score: a self-fulfilling prophecy? *Neurosurgery* 2019; **84:** 741–748.
10. Parry-Jones AR, Sammut-Powell C, Paroutoglou K, *et al*. An intracerebral hemorrhage care bundle is associated with lower case fatality. *Ann Neurol* 2019; **86**(4)**:** 495–503.
11. Xian Y, Xu H, O'Brien EC, *et al*. Clinical effectiveness of direct oral anticoagulants vs warfarin in older patients with atrial fibrillation and ischemic stroke. *JAMA Neurol* 2019; **76**(10)**:** 1192–1202.
12. Kothari RU, Brott T, Broderick JP, *et al*. The ABCs of measuring intracerebral hemorrhage volumes. *Stroke* 1996; **27**(8)**:** 1304–1305.
13. Mushkudiani NA, Hukkelhoven CWPM, Hernández AV, *et al*. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. *J Clin Epidemiol* 2008; **61**(4)**:** 331–343.
14. Vergouwe Y, Steyerberg EW. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014; **35**(29)**:** 1925–1931.
15. Rodríguez-Fernández S, Castillo-Lorente E, Guerrero-Lopez F, *et al*. Validation of the ICH score in patients with spontaneous intracerebral haemorrhage admitted to the intensive care unit in Southern Spain. *BMJ Open* 2018; **8**(8)**:** e021719.
16. Uyttenboogaart M, Stewart RE, Vroomen PC, Luijckx G-J, De Keyser J. Utility of the stroke-thrombolytic predictive instrument. *J Neurol Neurosurg Psychiatry* 2008; **79**(9)**:** 1079–1081.
17. Hanley JA, Hajian-Tilaki KO. Sampling variability of nonparametric estimates of the areas under receiver

operating characteristic curves: an update. *Acad Radiol* 1997; **4**(1)**:** 49–58.
18. Murthy SB, Merkler AE, Omran SS, *et al*. Outcomes after intracerebral hemorrhage from arteriovenous malformations. *Neurology* 2017; **88**(20)**:** 1882–1888.
19. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem* 2008; **54**(1)**:** 17–23.
20. Tsivgoulis G, Wilson D, Katsanos AH, *et al*. Neuroimaging and clinical outcomes of oral anticoagulant-associated intracerebral hemorrhage. *Ann Neurol* 2018; **84**(5)**:** 694–704.
21. Kuramatsu JB, Gerner ST, Schellinger PD, *et al*. Anticoagulant reversal, blood pressure levels, and anticoagulant resumption in patients with anticoagulation-related intracerebral hemorrhage. *JAMA* 2015; **313**(8)**:** 824–836.