



External validation of brain arteriovenous malformation haemorrhage scores, AVICH, ICH and R2eD

Basel A. Taweel^{1,2,3} · Conor S. Gillespie^{1,2} · George E. Richardson^{1,2} · Mohammad A. Mustafa^{1,2} · Tamara Ali² · Abdurrahman I. Islim² · Cathal J. Hannan² · Emmanuel Chavredakis^{1,2}

Received: 10 January 2022 / Accepted: 15 March 2022 / Published online: 18 April 2022
© The Author(s) 2022

Abstract

Purpose To externally validate the arteriovenous malformation-related intracerebral haemorrhage (AVICH), intracerebral haemorrhage (ICH), and novel haemorrhage presentation risk score (R2eD) in brain arteriovenous malformations.

Methods Adult patients diagnosed radiologically with an arteriovenous malformation (AVM) at a tertiary neurosurgical centre between 2007 and 2018 were eligible for inclusion. Both the AVICH and ICH scores were calculated for AVM-related symptomatic haemorrhage (SH) and compared against the modified Rankin scale (mRS) at discharge and last follow-up, with unfavourable outcome defined as mRS > 2. R2eD scores were stratified based on presentation with SH. External validity was assessed using Harrel's C-statistic.

Results Two hundred fifty patients were included. Mean age at diagnosis was 46.2 years [SD = 16.5]. Eighty-seven patients (34.8%) had a SH, with 83 included in the analysis. Unfavourable mRS outcome was seen in 18 (21.6%) patients at discharge and 18 (21.6%) patients at last follow-up. The AVICH score C-statistic was 0.67 (95% confidence interval [CI], 0.53–0.80) at discharge and 0.70 (95% CI, 0.56–0.84) at last follow-up. The ICH score C-statistic was 0.78 (95% CI 0.67–0.88), at discharge and 0.80 (95% CI 0.69–0.91) at last follow-up. The R2eD score C-statistic for predicting AVM haemorrhage was 0.60 (95% CI, 0.53–0.67).

Conclusions The AVICH score showed fair-poor performance, while the ICH score showed good-fair performance. The R2eD score demonstrated poor performance, and its clinical utility in predicting AVM haemorrhage remains unclear.

Keywords Arteriovenous malformations · Cerebral haemorrhage · Prognosis · Cerebrovascular disorders

Introduction

Brain arteriovenous malformations (AVMs) are vascular lesions characterised by arteriovenous shunting via a central entanglement of arterial and venous components known as a nidus. These lesions are thought to be congenital and carry

a lifetime risk of haemorrhage, with an estimated annual haemorrhage rate of 2–4% [5, 12].

Haemorrhage is the most common clinical presentation of AVMs (~50%), with AVM haemorrhage representing 2–4% of all haemorrhagic strokes and 1–2% of all strokes [1].

Treatment modalities for AVM include microsurgical resection [25], stereotactic radiosurgery and endovascular therapy, as well as conservative management [18]. In addition to the risk of haemorrhage represented by the natural history of AVMs, each modality of interventional therapy for AVMs carries a risk of neurological morbidity [6, 16], thus careful selection of lesions for interventional therapy is paramount.

Currently, no established paradigm exists to identify lesions that are likely to cause disability later in life, though clinical models have been proposed that attempt this. The R2eD score [10] stratifies AVMs based on both lesion and patient characteristics to ascertain those at risk of

This article is part of the Topical Collection on *Vascular Neurosurgery - Arteriovenous malformation*

✉ Basel A. Taweel
b.taweel@liverpool.ac.uk

¹ Institute of Systems, Integrative and Molecular Biology (ISMIB), University of Liverpool, Liverpool, UK

² The Walton Centre NHS Foundation Trust, Liverpool, UK

³ Clinical Sciences Centre, Aintree University Hospital, Fazakerley, Longmoor Lane, Liverpool L9 7AL, UK

haemorrhage (Table 1). The original intracerebral haemorrhage (ICH) score [13] has been suggested as a predictor of disability following AVM ICH [3]. The AVM-related ICH (AVICH) score [20] is a score based on the original ICH score, which was developed in an attempt to more accurately predict the occurrence of neurological disability following AVM-related ICH (Table 2).

These clinical tools may have the potential to guide management of AVMs, by identifying high-risk lesions that are more likely to bleed and cause neurological disability due to haemorrhage.

This study aims to externally validate these tools using a mixed cohort and to assess their clinical utility.

Methods

Study design and patient population

Data for adult patients (> 18 years) diagnosed with AVMs at a single tertiary neuroscience centre between 2007 and 2018 was retrospectively collected. For inclusion, AVMs must have been confirmed by magnetic resonance imaging (MRI) or digital subtraction angiography (DSA). AVMs that were radiologically identified on computerised tomography scans with angiography (CTA) only, with no further MRI/DSA radiological confirmation were excluded, as a detailed study of certain lesion characteristics required for calculation of both the R2ED and AVICH scores (deep drainage, single feeder, diffuse nidus) could not be reliably assessed. The transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement was used to report the study findings [9]. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Table 1 R2ED score components and component definitions

Component	Definition	R2eD
Race	White	0
	Non-white	2
Location	Other location (not deep)	0
	Deep	1
AVM size	Large ≥ 30 mm	0
	Small < 30 mm	1
Deep venous drainage	No	0
	Yes	1
Monoarterial feeding	No	0
	Yes	1
Score range		0–6

R2eD novel haemorrhage presentation risk score

Table 2 List of the components of the AVICH and ICH scores

Component	Definition	AVICH score	ICH score
Nidus size	< 3	1	/
	3–6	2	/
	> 6	3	/
Deep venous drainage	No	0	/
	Yes	1	/
Eloquence	No	0	/
	Yes	1	/
Age (AVICH thresholds)	< 20 years	1	/
	20–40 years	2	/
	> 40 years	3	/
Age (ICH thresholds)	< 80 years	/	0
	≥ 80 years	/	1
Diffuse nidus	No	0	/
	Yes	1	/
GCS	3–4	0	0
	5–12	1	1
	13–15	2	2
ICH volume	< 30	0	0
	≥ 30	1	1
Intraventricular Haemorrhage	No	0	0
	Yes	1	1
Infratentorial Haemorrhage	No	/	0
	Yes	/	1
Score range		2–13	0–6

AVICH arteriovenous malformation-related intracerebral haemorrhage, ICH Intracerebral haemorrhage, GCS Glasgow coma scale

Baseline clinical and radiological characteristics

Data was collected from electronic patient notes and available imaging. AVM-specific data was collected in accordance with published consensus definitions [23]. The ABC/2 formula [15] was used to determine the volume of the ICH, and the Glasgow Coma Scale on admission was used to determine GCS for the purposes of calculating ICH and AVICH scores. Ethnicity data (required for calculating the R2eD score for AVM patients) was patient reported and was collected from patient records.

Outcomes assessed

The primary outcome for the R2ED score was presentation with an ICH, either at first presentation, or during follow-up. This was conducted with the assumption that the risk of rupture of monitored lesions is identical to that of unknown lesions. Haemorrhages caused by AVM-related flow aneurysms were excluded from analysis. The outcomes for the ICH and AVICH score were modified Rankin scale (mRS) at discharge and last follow-up. Unfavourable outcome for the ICH and AVICH score is defined as mRS > 2. Patient outcomes were identified using electronic patient notes and clinical records.

Statistical analysis

Baseline patient variables found to be significant in scoring models were compared between the original reported patient populations and between our patient cohorts. We assessed whether the individual values of continuous baseline variables lied within the reported ranges of the original study reporting development of the score. Any patient with continuous baseline characteristics lying outside of these ranges was excluded from the analysis.

Binary logistic regression was used to construct a generalised linear model for the purposes of assessing calibration. Regression was performed in accordance with the respective associated outcome (R2eD = occurrence of symptomatic haemorrhage (SH), AVICH and ICH = modified Rankin scale (mRS) > 2). For AVICH and ICH, both mRS after discharge, and at last follow-up were modelled, representing an

assessment of the ability of scores to predict patient disability and dependence both short-term, and long-term. Date of last follow-up was determined to be the date of the most recent physical encounter with any clinician at our tertiary centre. Unique fitted values were ranked to model score outcomes as a continuous variable, for the purposes of constructing calibration plots. The resultant plots were used to visualise concordance of score prediction with observed outcome. Receiver operator curves (ROC) curves were plotted and C-statistics were used to assess discriminatory capability for each score. C-statistics are calculated by measuring the area under receiver operator curves (AUROC). The following grading scale was used to summarise the discriminatory capability of each model based on C-statistic: (1–0.9: excellent, 0.8–0.9: good, 0.7–0.8: fair, 0.6–0.7: poor, 0.5–0.6: very poor) (Table 5). R version 4.0.3 (used packages: pROC v.17.0.1, ggplot 3.3.3) was used for all the statistical analyses.

Table 3 Baseline characteristics of the validation cohort

	All patients (<i>n</i> =250) (%)	Haemorrhage pts (<i>n</i> =83) (%)	Unfavourable outcome patients (<i>n</i> =18) (%)
Age (years)	48.0	45.7	54.0
Male	120 (48.0)	46 (55.4)	8 (44.4)
Female	130 (52.0)	37 (44.6)	10 (55.4)
GCS			
13–15	-	60 (72.2)	9 (44.4)
5–12	-	15 (18.0)	5 (27.8)
3–4	-	8 (0.10)	4 (22.2)
Median ICH volume (cm ³)	-	7.4	10.22
IVH	-	30 (36.1)	10 (55.6)
Infratentorial location	-	30 (36.1)	11 (61.1)
Ethnicity			
White	229 (91.6)	74 (89.2)	16 (88.9)
Non-white	21 (8.4)	9 (10.8)	2 (11.1)
Single feeder		21 (25.3)	2 (11.1)
Small size nidus (< 3 cm)	175 (70.0)	63 (75.9)	14 (77.8)
Large size nidus (≥ 3 cm)	75 (30.0)	20 (24.1)	4 (22.2)
Deep drainage	117 (46.8)	43 (51.8)	11 (61.1)
Deep location	69 (27.6)	30 (36.1)	11 (61.1)
Eloquent location	112 (44.8)	38 (45.8)	11 (61.1)
Diffuse nidus	66 (26.4)	30 (36.1)	7 (38.9)
Spetzler-Martin grade			
I	67 (26.8)	19 (22.9)	4 (22.2)
II	91 (36.4)	36 (43.4)	5 (27.8)
III	56 (22.4)	14 (16.9)	5 (27.8)
IV	29 (11.6)	13 (15.7)	4 (22.2)
V	7 (2.8)	1 (1.2)	0 (0.0)

GCS Glasgow coma scale, ICH: intracerebral haemorrhage, IVH intraventricular haemorrhage

Results

Patient baseline characteristics

Patient baseline characteristics are reported in Table 3. Two hundred fifty patients with AVM (mean age = 46.24 [SD = 16.5], $M:F = 0.92$) were identified. The median duration of follow-up following haemorrhage was 33.74 months. No included patients had any co-morbid cavernous malformations, dural arteriovenous fistulas or hemangioblastomas. Two patients were found to have 2 AVMs each. As distinct nidi could be identified, these were deemed as separate lesions for the purposes of this study. Ranges of continuous variables of the original studies were found to encompass our reported ranges, with no patients excluded due to outlying characteristics.

Patient outcomes

Ninety-three patients (37.2%) presented with SH. After exclusion of 6 flow-aneurysm haemorrhages, 87 patients

remained, with a further 4 patients being excluded due to missing data. Of the included 83 haemorrhages, 74 (89.2%) presented with haemorrhage as the first presentation of their AVM, while 9 (10.8%) AVM haemorrhages were haemorrhages of known, conservatively managed lesions. Eighteen (21.7%) patients were found to have unfavourable mRS (> 2) at discharge, and 18 (21.7%) patients were found to have unfavourable mRS at last follow-up.

External validation results

The resulting ROC curves and calibration plots for each score are visualised in Figs. 1 and 2. The ICH score produced a C-statistic of 0.78 (95% CI 0.67–0.88) when predicting mRS at discharge, and a C-statistic of 0.80 (95% CI 0.69–0.91) when predicting mRS at last follow-up. The AVICH score produced a C-statistic of 0.67 (CI, 0.53–0.80) using discharge mRS and 0.70 (CI 0.56–0.84) using mRS at last follow-up.

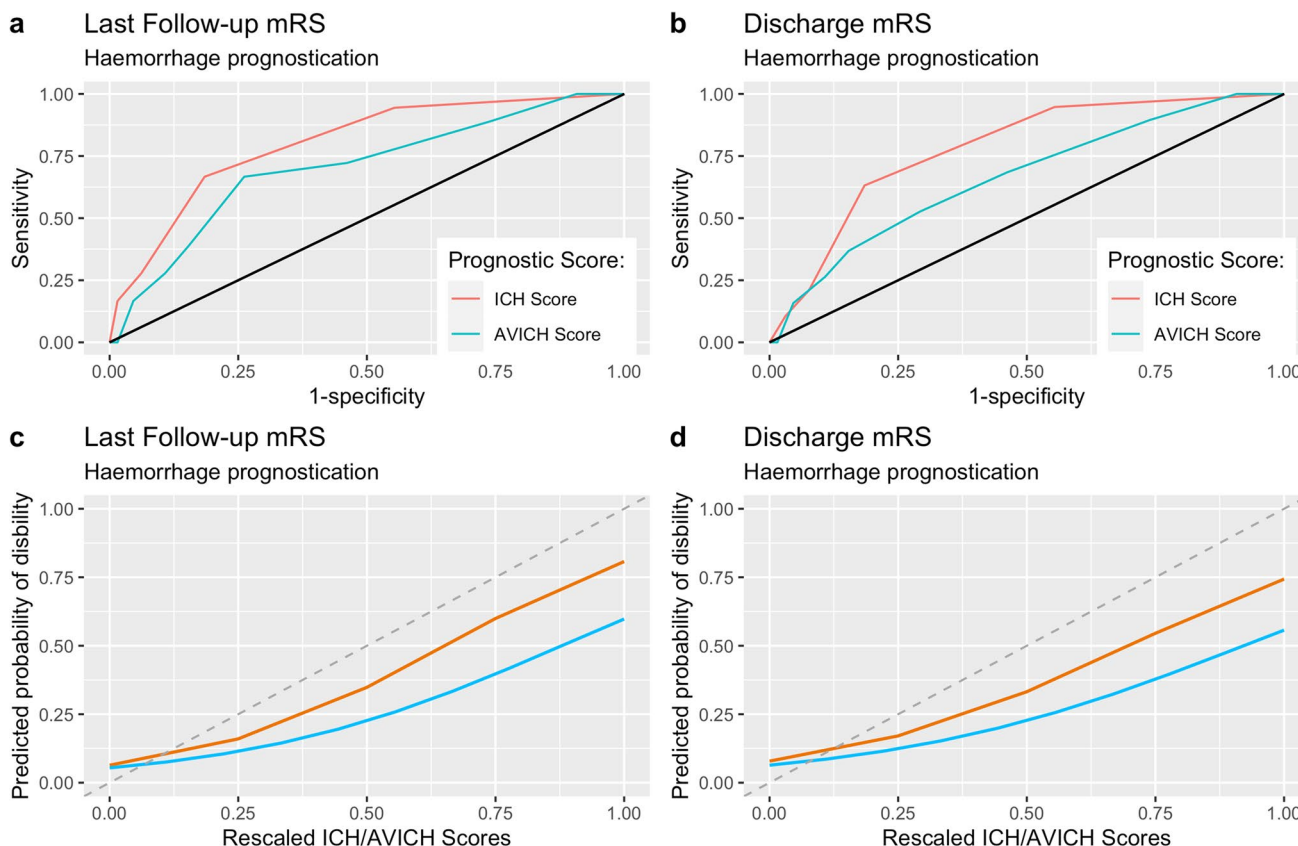


Fig. 1 ROC curves and calibration plots used to assess discriminatory capability and model calibration for the ICH score and the AVICH score. Panels (a) and (c) show the ROC curves and calibration plots,

respectively, for both scores using mRS at last follow-up as the outcome. (b) and (d) utilise mRS at discharge as the outcome of interest

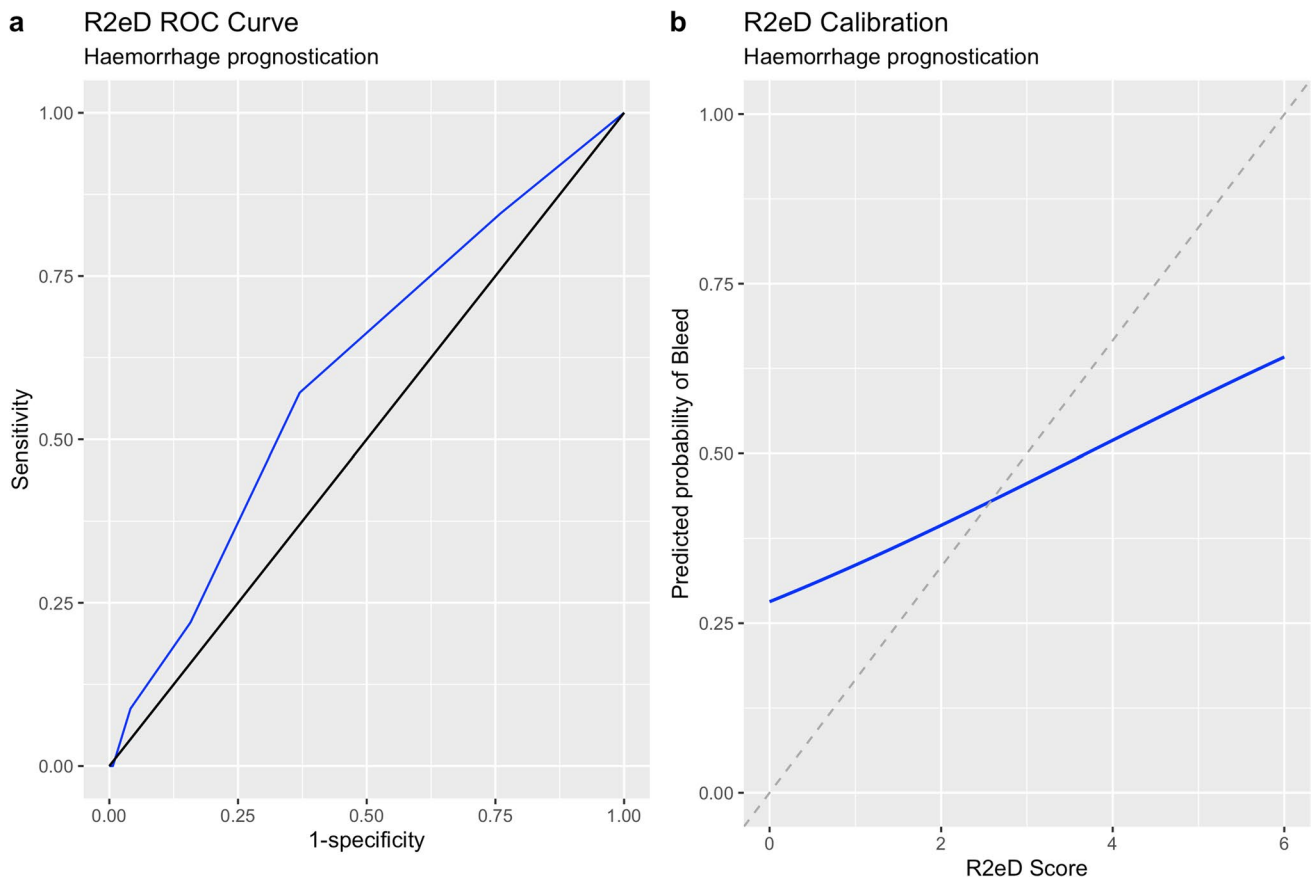


Fig. 2 ROC curves (a) and calibration plots (b) used to assess discriminatory capability and model calibration for the R₂eD score

The R2eD score produced a C-statistic of 0.60 (0.53–0.67) (Table 4). The calibration curve produced for R2eD showed a disparity between predicted and observed results. The ROC curve and calibration curve for the R2eD score are shown in Fig. 2.

Using our aforementioned scale, the ICH score is found to be a good-fair predictor of disability following AVM rupture, while the AVICH score is found to be a fair-poor

Table 4 External validation results, showing the area under curve for each score. For the ICH and AVICH scores, discharge and follow-up mRS are used as short-term and longer-term outcomes, respectively

Score name	Discharge mRS AUC (95% CI)	Follow-up mRS AUC (95% CI)
ICH score	0.78 (0.67–0.88)	0.80 (0.69–0.91)
AVICH score	0.67 (0.53–0.80)	0.70 (0.56–0.84)
Score name	Symptomatic haemorrhage AUC (95% CI)	
R2eD score	0.60 (0.53–0.67)	

mRS modified Rankin scale, AUC area under curve, CI confidence intervals, ICH intracerebral haemorrhage, AVICH arteriovenous malformation-related intracerebral haemorrhage, R2eD novel haemorrhage presentation risk score

predictor. The R2eD score (AUC = 0.60) is found to be a poor and near-non-discriminatory predictor of AVM haemorrhage in our cohort (Table 5).

Discussion

Here we present the largest cohort validation of the R2eD score, in addition to an external validation of the ICH and AVICH scores for AVM rupture. We find that the ICH score is a good-fair predictor of disability following AVM-ICH and that the AVICH score is a fair-poor predictor of

Table 5 The grading scale used to assess Harrel’s C statistic

Grade	AUC score
Excellent	1–0.9
Good	0.9–0.8
Fair	0.8–0.7
Poor	0.7–0.6
Non-discriminatory	0.6–0.5

AUC area under curve

disability. Finally, we find that the R2ED score is a poor predictor of AVM haemorrhage in our cohort (Table 5).

Prior to this study, the R2eD score has been externally validated in a single cohort study of 122 patients [4] showing fair performance as a predictor of haemorrhage (AUC = 0.71); our study therefore represents the largest cohort validation of the R2eD to date.

The AVICH score has been externally validated by the developers of the score in an international multi-centre validation study [19], showing results surpassing that of the ICH score with an AUC of 0.77 for the AVICH score and AUC of 0.71 for the ICH score. Our study is the first independent assessment of the AVICH score in AVM haemorrhage. The ICH score has also been validated externally as a predictor of disability following AVM rupture [3], producing an AUC score of 0.89, with no comparison to the AVICH score.

The differences in the reported discriminatory power of haemorrhage scores between the multi-centre study and others may be explained by smaller event numbers, with only 18 incidences of unfavourable outcome in our cohort, 15 incidences in the initial ICH validation study, and 115 incidences of poor outcome in the multi-centre study. It should also be noted that our validation of the ICH score and the AVICH produced wide intersecting confidence intervals preventing definitive determination of a preferred score based on the discriminative capability reported in our study. For all scores, the AUC values produced were less than that of the original study reporting development of the score, which is expected when validating in an external cohort distinct from the original training cohort [8, 24].

Clinical utility and implications — R2ED score

Grading scales such as the R2ED score are developed with the ultimate purpose of predicting future events, allowing clinicians to use the score to guide intervention in patients with AVM. The R2ED score was developed with retrospective data assuming that the natural history of all AVMs is similar to lesions that undergo conservative management. In this study, (which is also based on retrospective data,) we utilise that same assumption for the purposes of external validation of the R2ED score. We add to this, by including patients with conservatively managed AVMs that were discovered either incidentally or via other non-haemorrhagic symptomatic manifestation. While this technically leads to the inclusion of a cohort of patients that differs from the original training set (known vs. unknown lesions), this inclusion has the advantage of validating the R2ED score in a cohort that is more representative of the ideal clinical use case of the model. As noted by the developers of the R2ED score, the model was constructed based upon the study of characteristics that are observable in lesions that presented with haemorrhage.

This does not, however, mean that these same characteristics are observable in high-risk AVMs that have not yet undergone haemorrhage. These issues arise from the use of retrospective data, and thus further prospective studies are required to determine the true clinical utility of this model and any future model that aims to risk-stratify AVMs. Finally, while the R2ED score aims to be utilised to predict haemorrhage, it does not address lesions that have previously bled, which have been found to have a significantly higher risk of subsequent haemorrhage [11, 22, 27].

Clinical utility and implications— ICH and AVICH scores

For the purposes of establishing a preferred model for predicting patient outcome following AVM rupture, model practicality must be taken into consideration. The ICH score is a model that was proposed in 2001 to assess mortality in all patients presenting with ICH and has been validated extensively for both its original use [7, 14, 21], and for use specifically to predict disability in AVM-ICH presentations [3]. The AVICH score is a bespoke score developed and validated for the purposes of predicting disability following AVM rupture. The use of the AVICH score therefore stipulates that a patient must have a previously known lesion that has bled, or that the lesion must be diagnosed at presentation in order for the model to be applicable. In our cohort, only 9 (10.8%) of our 83 studied bleeds were haemorrhages of known lesions, with the vast majority of AVM haemorrhagic presentations (89.2%) involving previously unknown AVMs. Furthermore, diagnosis of AVMs in the acute setting can prove difficult, as any underlying lesions can be obscured by the presenting haematoma on radiological investigations [2]. Additionally, the obscuring haematoma may impede the study of certain lesion characteristics necessary for use of the score, such as size, deep drainage and diffuseness of the lesion. These factors may obstruct use of the AVICH score in the acute setting to predict patient disability, and thus may preclude the use of AVICH in acute setting decision-making for AVM patients. The ICH score, however, is not subject to this limitation as it is useful both as a predictor of mortality in non-AVM-ICHs and as a predictor of disability in AVM-ICH. Furthermore, the ICH score does not rely on detailed radiological study of AVM characteristics and can be used based on radiological study of the presenting haematoma alone. The ICH score may therefore be applied in all ICH scenarios, irrespective of cause.

Heterogeneity of arteriovenous malformations

Difficulties may arise in constructing models to predict clinical outcomes in AVMs, chief among which is the

wide variability observed in the characteristics of different lesions. The developers of the R2eD score, reported the results of a literature review [10] which aimed to identify all studies of haemorrhage predictors in AVM, including AVM size, deep venous drainage, deep location, the number of feeding artery vessels, venous stenosis, patient ethnicity, and the presence of AVM-associated aneurysms. Wide variability was observed in the results of the identified studies, with some studies reporting the studied characteristic as a significant predictor of haemorrhage, and others reporting the converse. A recent study [26] demonstrated an attempt to utilise machine learning algorithms to construct predictive models in AVMs. The authors reported the produced models as unstable and uncertain. These studies and others [1, 17] indicate that AVM heterogeneity may obstruct the development of a haemorrhage prediction algorithm that encompasses all AVMs. These studies, when considered alongside our findings with relation to the R2eD score, may endorse further sub-setting of AVMs with the aim of finding more consistent patterns amongst similar lesions.

Limitations of study

Our study has a number of limitations. It is a single-centre study based on retrospective patient data. Our cohort had 83 events of AVM haemorrhage for external validation of the R2eD score and 18 events of unfavourable neurological outcome for assessing the ICH and AVICH score, both of which lie under the recommended value of 100 events for validation studies [8]. Additionally, while the ICH score was found to have an AUC value greater than the AUC value for the AVICH score, the low number of events produced wide intersecting confidence intervals meaning that a preferred score cannot be definitively determined.

While the R2eD score showed poor performance in our study, it should be noted that the R2eD score includes ethnicity as a significant predictor of AVM haemorrhage, and our cohort was 91.6% white, which may not effectively illustrate the clinical utility of the R2eD score in more diverse populations.

Conclusion

We present an external validation of three prognostic scores, the ICH, AVICH and R2eD scores that collectively aim to both predict and prognosticate outcome in AVM haemorrhage. We find that the ICH score is a good-fair predictor of disability following AVM haemorrhage, whereas the AVICH score is a fair-poor predictor. We find that the R2eD score is a poor predictor of AVM haemorrhage, and we recommend further prospective studies to help determine the true clinical utility of

these models. Finally, we consider the possibility of an AVM categorisation system as a step to improve the predictive power of future scores that aim to predict AVM haemorrhage.

Author contribution Basel A. Taweel, Abdurrahman I. Islim, Cathal J. Hannan and Emmanuel Chavredakis contributed to the study conception and design. Material preparation, data collection and analysis were performed Basel A. Taweel, Conor S. Gillespie, George E. Richardson, Mohammad A. Mustafa and Tamara Ali. The first draft of the manuscript was written by Basel A. Taweel and all the authors commented on the previous versions of the manuscript. All the authors read and approved the final manuscript.

Funding B. A. T and C. S. G declare a bursary from the Wolfson Foundation and Royal College of Physicians.

There are no conflicts of interest to disclose.

Declarations

Ethics approval Ethical approval was waived by the local Ethics Committee of The Walton Centre NHS Foundation Trust in view of the retrospective nature of the study and all the procedures being performed were part of routine care.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Al-Shahi R, Warlow C (2001) A systematic review of the frequency and prognosis of arteriovenous malformations of the brain in adults. *Brain* 124:1900–1926. <https://doi.org/10.1093/brain/124.10.1900>
2. Amans MR, Dillon WP (2015) Imaging evaluation of arteriovenous malformations. In: Kondziolka DS, Kalani MYS, Higashida RT, Spetzler RF (eds) *Comprehensive management of arteriovenous malformations of the brain and spine*. Cambridge University Press, Cambridge, pp 57–73. <https://doi.org/10.1017/CBO9781139523943.008>
3. Appelboom G, Hwang BY, Bruce SS, Piazza MA, Kellner CP, Meyers PM, Connolly ES (2012) Predicting outcome after arteriovenous malformation-associated intracerebral hemorrhage with the original ICH score. *World Neurosurgery* 78:646–650. <https://doi.org/10.1016/j.wneu.2011.12.001>
4. Bird WA, Hendrix P, Bohan C, Goren O, Schirmer CM, Griesenauer CJ (2021) External validation of the R2eD AVM score to predict the likelihood of rupture presentation of brain arteriovenous score to predict the likelihood of rupture presentation of

- brain arteriovenous malformations. *Neurosurgery*. <https://doi.org/10.1093/neuros/nyab108>
5. Can A, Gross BA, Du R (2017) The natural history of cerebral arteriovenous malformations. In: Elsevier, pp 15–24. <https://doi.org/10.1016/b978-0-444-63640-9.00002-3>
 6. Castel JP, Kantor G (2001) Postoperative morbidity and mortality after microsurgical exclusion of cerebral arteriovenous malformations. Current data and analysis of recent literature. *Neurochirurgie* 47:369–383
 7. Clarke JL, Johnston SC, Farrant M, Bernstein R, Tong D, Hemphill JC (2004) External validation of the ICH score. *Neurocrit Care* 1:53–60. <https://doi.org/10.1385/NCC:1:1:53>
 8. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, Voysey M, Wharton R, Yu LM, Moons KG, Altman DG (2014) External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol* 14:40. <https://doi.org/10.1186/1471-2288-14-40>
 9. Collins GS, Reitsma JB, Altman DG, Moons KGM (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med* 13:1. <https://doi.org/10.1186/s12916-014-0241-z>
 10. Feghali J, Yang W, Xu R, Liew J, McDougall CG, Caplan JM, Tamargo RJ, Huang J (2019) R 2 eD AVM Score. *Stroke* 50:1703–1710. <https://doi.org/10.1161/strokeaha.119.025054>
 11. Gross BA, Du R (2012) Rate of re-bleeding of arteriovenous malformations in the first year after rupture. *J Clin Neurosci* 19:1087–1088. <https://doi.org/10.1016/j.jocn.2011.12.005>
 12. Gross BA, Du R (2013) Natural history of cerebral arteriovenous malformations: a meta-analysis. *J Neurosurg* 118:437–443. <https://doi.org/10.3171/2012.10.Jns121280>
 13. Hemphill JC 3rd, Bonovich DC, Besmertis L, Manley GT, Johnston SC (2001) The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke* 32:891–897. <https://doi.org/10.1161/01.str.32.4.891>
 14. Hemphill JC 3rd, Farrant M, Neill TA Jr (2009) Prospective validation of the ICH Score for 12-month functional outcome. *Neurology* 73:1088–1094. <https://doi.org/10.1212/WNL.0b013e3181b8b332>
 15. Kothari RU, Brott T, Broderick JP, Barsan WG, Sauerbeck LR, Zuccarello M, Khoury J (1996) The ABCs of measuring intracerebral hemorrhage volumes. *Stroke* 27:1304–1305. <https://doi.org/10.1161/01.str.27.8.1304>
 16. Lee B-B, Do YS, Yakes W, Kim DI, Mattassi R, Hyon WS (2004) Management of arteriovenous malformations: a multidisciplinary approach. *J Vasc Surg* 39:590–600. <https://doi.org/10.1016/j.jvs.2003.10.048>
 17. Mansmann U, Meisel J, Brock M, Rodesch G, Alvarez H, Lasjaunias P (2000) Factors associated with intracranial hemorrhage in cases of cerebral arteriovenous malformation. *Neurosurgery* 46:272–280. <https://doi.org/10.1097/00006123-200002000-00004>
 18. Medical, endovascular, and radiosurgical treatment of arteriovenous malformations (2015). In: Kondziolka DS, Kalani MYS, Higashida RT, Spetzler RF (eds) *Comprehensive management of arteriovenous malformations of the brain and spine*. Cambridge University Press, Cambridge, pp 201–326. doi:DOI: <https://doi.org/10.1017/CBO9781139523943.018>
 19. Neidert MC, Lawton MT, Kim LJ, Nerva JD, Kurisu K, Ikawa F, Konczalla J, Dinc N, Seifert V, Habdank-Kolaczowski J, Hatano T, Hayase M, Podlesek D, Schackert G, Wanet T, Gläsker S, Griessenauer CJ, Ogilvy CS, Kneist A, Sure U, Seifert B, Regli L, Bozinov O, Burkhardt J-K (2018) International multicentre validation of the arteriovenous malformation-related intracerebral haemorrhage (AVICH) score. *J Neurol Neurosurg Psychiatry* 89:1163–1166. <https://doi.org/10.1136/jnnp-2017-316259>
 20. Neidert MC, Lawton MT, Mader M, Seifert B, Valavanis A, Regli L, Bozinov O, Burkhardt JK (2016) The AVICH score: a novel grading system to predict clinical outcome in arteriovenous malformation-related intracerebral hemorrhage. *World Neurosurg* 92:292–297. <https://doi.org/10.1016/j.wneu.2016.04.080>
 21. Nisar T, Alchaki A, Hillen M (2018) Validation of ICH score in a large urban population. *Clin Neurol Neurosurg* 174:36–39. <https://doi.org/10.1016/j.clineuro.2018.09.007>
 22. Pollock BE, Flickinger JC, Lunsford LD, Bissonette DJ, Kondziolka D (1996) Factors that predict the bleeding risk of cerebral arteriovenous malformations. *Stroke* 27:1–6. <https://doi.org/10.1161/01.str.27.1.1>
 23. Reporting terminology for brain arteriovenous malformation clinical and radiographic features for use in clinical trials (2001). *Stroke* 32:1430–1442. <https://doi.org/10.1161/01.str.32.6.1430>
 24. Steyerberg EW, Harrell FE Jr (2016) Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* 69:245–247. <https://doi.org/10.1016/j.jclinepi.2015.04.005>
 25. Surgical approaches to treatment of arteriovenous malformations (2015). In: Kondziolka DS, Kalani MYS, Higashida RT, Spetzler RF (eds) *Comprehensive management of arteriovenous malformations of the brain and spine*. Cambridge University Press, Cambridge, pp 113–200. <https://doi.org/10.1017/CBO9781139523943.013>
 26. Tao W, Yan L, Zeng M, Chen F (2021) Factors affecting the performance of brain arteriovenous malformation rupture prediction models. *BMC Med Inform Decis Mak* 21:142. <https://doi.org/10.1186/s12911-021-01511-z>
 27. Yamada S, Takagi Y, Nozaki K, Kikuta K-i, Hashimoto N (2007) Risk factors for subsequent hemorrhage in patients with cerebral arteriovenous malformations. *Journal of Neurosurgery JNS* 107:965–972. <https://doi.org/10.3171/JNS-07/11/0965>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.