AJNR

Prediction of Outcome Using Quantified Blood Volume in Aneurysmal SAH

W.E. van der Steen, H.A. Marquering, L.A. Ramos, R. van den Berg, B.A. Coert, A.M.M. Boers, M.D.I. Vergouwen, G.J.E. Rinkel, B.K. Velthuis, Y.B.W.E.M. Roos, C.B.L.M. Majoie, W.P. Vandertop and D. Verbaan

Prediction of Outcome Using Quantified Blood Volume in Aneurysmal SAH

W.E. van der Steen, H.A. Marquering, L.A. Ramos, R. van den Berg, B.A. Coert, A.M.M. Boers, M.D.I. Vergouwen, G.J.E. Rinkel, B.K. Velthuis, Y.B.W.E.M. Roos, C.B.L.M. Majoie, W.P. Vandertop, and D. Verbaan

ABSTRACT

BACKGROUND AND PURPOSE: In patients with SAH, the amount of blood is strongly associated with clinical outcome. However, it is commonly estimated with a coarse grading scale, potentially limiting its predictive value. Therefore, we aimed to develop and externally validate prediction models for clinical outcome, including quantified blood volumes, as candidate predictors.

MATERIALS AND METHODS: Clinical and radiologic candidate predictors were included in a logistic regression model. Unfavorable outcome was defined as a modified Rankin Scale score of 4–6. An automatic hemorrhage-quantification algorithm calculated the total blood volume. Blood was manually classified as cisternal, intraventricular, or intraparenchymal. The model was selected with bootstrapped backward selection and validated with the R^2, C-statistic, and calibration plots. If total blood volume remained in the final model, its performance was compared with models including location-specific blood volumes or the modified Fisher scale.

RESULTS: The total blood volume, neurologic condition, age, aneurysm size, and history of cardiovascular disease remained in the final models after selection. The externally validated predictive accuracy and discriminative power were high ($R^2 = 56\% \pm 1.8\%$; mean C-statistic $= 0.89 \pm 0.01$). The location-specific volume models showed a similar performance ($R^2 = 56\% \pm 1\%$, $P = .8$; mean C-statistic $= 0.89 \pm 0.00$, $P = .4$). The modified Fisher models were significantly less accurate ($R^2 = 45\% \pm 3\%$, $P < .001$; mean C-statistic $= 0.85 \pm 0.01$, $P = .03$).

CONCLUSIONS: The total blood volume–based prediction model for clinical outcome in patients with SAH showed a high predictive accuracy, higher than a prediction model including the commonly used modified Fisher scale.

ABBREVIATIONS: aSAH = aneurysmal subarachnoid hemorrhage; IPH = intraparenchymal hemorrhage; IQR = interquartile range; IVH = intraventricular hemorrhage; TBV = total blood volume; WFNS = World Federation of Neurosurgical Societies

Aneurysmal subarachnoid hemorrhage (aSAH) is a severe form of stroke caused by the rupture of an intracranial aneurysm.[1] The in-hospital case-fatality rate is approximately 30%, and a large proportion of patients have a poor clinical outcome.[2] Known predictors of outcome are age, neurologic condition at admission, aneurysm size, and hemorrhage volume.[3]

A prediction model including predictors that are quickly and easily available after admission of the patient to the emergency department could assist physicians in making treatment decisions and counseling patients and their families. Such a model including the age, World Federation of Neurosurgical Societies (WFNS) grade on admission, and premorbid history of hypertension, was recently published.[4] Adding the amount of blood, as assessed on the Fisher scale, to their model did not substantially increase the predictive value. This finding is remarkable because the amount of blood has been strongly associated with poor outcome in previous studies.[5,6]

The hemorrhage volume in patients with aSAH is frequently estimated using the Fisher scale, which grades the amount of

blood in both the cisterns and the ventricles.[7] This 4-category radiologic scale is coarse and has only moderate interobserver agreement.[5] An extensive grading scale, such as the Hijdra sum score, has shown a stronger association with outcome. However, its extensiveness makes it less suitable for use in daily practice.[8] Currently, automated quantitative and observer-independent measures are available to determine the hemorrhage volume.[9]

We aimed to develop and validate a prediction model that estimates the risk of poor clinical outcome, including predictors available at admission, using automated quantified total blood volume (TBV) as one of the candidate predictors. Furthermore, we aimed to develop secondary models, including the cisternal, intraventricular (IVH), and intraparenchymal (IPH) blood volumes separately, and a model including the modified Fisher scale and to compare their performances with that of the TBV models.

MATERIALS AND METHODS
Development and Validation Cohort
Patients for the development and validation cohort were collected from the prospective aSAH registries of the Amsterdam University Medical Center and the University Medical Center Utrecht in the Netherlands, respectively. The development cohort consisted of all patients with aSAH admitted to the Amsterdam University Medical Center between December 2011 and December 2016. The validation cohort consisted of patients admitted to the University Medical Center Utrecht between 2013 and 2015. We used the following inclusion criteria: 1) SAH with subarachnoid blood on first admission NCCT or confirmed by xantochromic CSF after lumbar puncture; 2) confirmation of a ruptured aneurysm, diagnosed by either CTA, MRA, or DSA; and 3) 18 years of age or older. We excluded patients for whom the hemorrhage volume could not be segmented due to large movement and/or metal artifacts on admission NCCT. Finally, patients participating in the Ultra-Early Tranexamic Acid After Subarachnoid Hemorrhage (ULTRA) trial (clinicaltrials.gov No. NCT02684812) were excluded.[10] For the retrospective analysis of this registry, the need for informed consent was waived by the local medical ethics committees.

Collected Candidate Predictor Variables
We collected the following clinical candidate predictor variables: age, sex, history of hypertension, history of diabetes, history of cardiovascular disease, and neurologic condition on first admission assessed by the WFNS scale. Collected radiologic candidate predictor variables were the TBV and location-specific blood volumes, including cisternal, IVH, and IPH blood volumes; the modified Fisher grade; aneurysm size (defined as the maximum width or length of the aneurysm suspected of rupture); and aneurysm location (anterior or posterior circulation). The TBV was measured on admission NCCT using a fully automatic hemorrhage-quantification algorithm. The method was based on relative density increases due to the presence of blood. The analysis started by classification of different brain structures by atlas-based segmentation. Thereafter, the density was evaluated to set a tissue-specific threshold for segmentation of blood. Finally, a region-growing algorithm included subtle attenuated parts of the hemorrhagic areas. The TBV was calculated by multiplying the

voxels that were classified as blood by the voxel size and was expressed in centiliters.[9] Most scans were performed on a Somatom AS+ (Ultra-fast Ceramic Detector; Siemens) or Brilliance iCT (solid state detector; Phillips Healthcare). For the analysis, 5- or 3-mm section thickness scans were used. The TBV consisted of all extravasated blood visible on NCCT after ictus, including cisternal, IVH, and IPH blood volumes. The TBV was classified as cisternal, IVH, or IPH by manually outlining the ventricular or intraparenchymal part of the segmented total hemorrhage. In case the patient had recurrent bleeding before treatment, the CT scan after the recurrent bleeding was used instead of the baseline scan to determine the blood volumes. All segmentations were checked and, if necessary, corrected by an experienced radiologist (R.v.d.B.), who was blinded to all clinical data and outcome. The modified Fisher scale was administered by an experienced neurosurgeon.[11]

Outcomes
The clinical outcome was assessed using the mRS in the development cohort and on the Glasgow Outcome Scale in the validation set. The clinical outcome was assessed at 6 months after the SAH in the development cohort and at 3 months in the validation set. A neurovascular research nurse assessed the mRS using a structured interview, either during a visit to the outpatient clinic or by telephone interview. The neurovascular research nurse was blinded to the TBV at admission. Delayed cerebral ischemia was considered as clinical deterioration, defined as the occurrence of new focal neurologic impairment or a decrease of ≥2 points on the Glasgow Coma Scale (with or without new hypodensity on CT) that could not be attributed to other causes, in accordance with the definition proposed by a multidisciplinary research group.[12] The primary outcome was an unfavorable clinical outcome and was defined as an mRS score of 4–6 in the development cohort and as a Glasgow Outcome Scale score of 1–3 in the validation cohort. The secondary outcome was death after the SAH during the observation period.

Statistical Analysis
Baseline variables were compared between the development and validation cohorts using the Fisher exact test for dichotomous and categoric variables, the independent-samples *t* test for normally distributed continuous variables, and the Mann-Whitney *U* test for non-normally distributed continuous variables.

Missing Data. In the development set, 22% of the cases had ≥1 missing variable, with a maximum of 11% missing per variable. In the validation set, 28% of the cases had >1 missing variable, with a maximum of 16% missing per variable. We assumed data to be missing at random. Missing values were imputed with multiple imputation.[13] Using multiple imputation, we created 20 complete imputed datasets of both the development and the validation sets.

Model Development. All candidate predictor variables were included in logistic regression models. Variables with limited predictive value, according to the Akaike information criterion, were stepwise-backward removed until a final model with the lowest

Table 1: Characteristics of the development and validation cohorts[a]

	Development (n = 409)	Validation (n = 317)	P Value
Age (mean) (SD) (yr)	57 (13)	59 (14)	.05
Female sex	277 (68)	222 (70)	.5
History of hypertension	131 (35)	96 (31)	.2
History of cardiovascular disease	75 (20)	61 (20)	1
History of diabetes	25 (7)	18 (6)	.6
WFNS score			.04
I	182 (46)	106 (39)	
II	51 (13)	37 (14)	
III	12 (3)	9 (3)	
IV	68 (17)	40 (15)	
V	81 (21)	74 (27)	
Aneurysm location			.44
Anterior circulation	324 (79)	243 (82)	
Posterior circulation	85 (21)	54 (18)	
Aneurysm size (median) (IQR) (mm)	6 (4–9)	5 (4–8)	.96
Aneurysm treatment			<.001
Coiling	273 (67)	153 (48)	
Clipping	64 (16)	95 (30)	
No treatment	72 (18)	69 (22)	
Total blood volume (median) (IQR) (mL)	29 (12–60)	26 (9–51)	.36
Cisternal blood volume (median) (IQR) (mL)	20 (8–41)	18 (6–34)	.1
Intraventricular blood volume (median) (IQR) (mL)	0.5 (0–2)	0.3 (0–2)	.1
Intraparenchymal blood volume (median) (IQR) (mL)	0 (0–3)	0 (0–3)	.8
Modified Fisher grade			.15
0	20 (5)	12 (4)	
1	26 (6)	9 (3)	
2	7 (2)	8 (3)	
3	92 (23)	83 (27)	
4	262 (64)	201 (64)	
Clinical DCI	109 (27)	61 (29)	.6

Note:—DCI indicates delayed cerebral ischemia.
[a] All values are in No. (%) unless otherwise indicated.

Akaike information criterion was defined. The backward selection procedure was repeated in 100 randomly drawn samples of each of the 20 imputed development datasets using bootstrap resampling, creating 2000 models.[14] Candidate predictors that remained in >50% of these models were included as predictors in our final models.

Model Performance and Internal Validation. The explained variance of the models was evaluated with R^2. The ability to discriminate between patients with or without a poor outcome was assessed with the C-statistic. The agreement between the observed and the predicted outcomes was assessed with a calibration curve. Model performance was internally validated by calculating the optimism-corrected R^2 and C-statistic.[14] These optimism-corrected performance measures were calculated in each of the 20 imputed development sets, and the performance measures were averaged using mean ± SD.

External Validation. The performance of the models in the validation set was evaluated with R^2, the C-statistic, a calibration curve, the calibration slope, and the intercept (calibration-in-the-

large). The calibration slope reflects the average strength of the predictor effects. A calibration slope of 1 indicates perfect agreement between the average predictor effects in the development and validation sets. The intercept reflects the difference between the average of the predicted outcome and the average of the observed outcome. An intercept of zero indicates perfect calibration. The external validation was performed in each of the 20 imputed validation sets, and the performance measures were averaged using the mean (± SD).

Secondary Models. If the TBV remained in the final model, the performance of this model was compared with a location-specific volume model in which the TBV was replaced with the cisternal, IVH, and IPH blood volumes. Furthermore, we created a model in which the TBV was replaced with the modified Fisher scale. The performances of these secondary models were compared with the TBV model by comparing the mean R^2 values in the imputed datasets with the independent-samples t test. The C-statistics of the models were compared using the DeLong test.

Sensitivity Analyses. Because not treating a patient has a strong effect on outcome, the inclusion of patients who were not treated may have biased the model. To determine the predictive value of the model on treated patients, we performed a sensitivity analysis including only patients who received aneurysm treatment. Because the predictive value of the amount of blood may decrease with increasing time between the initial hemorrhage and the admission CT scan, the performance of the TBV model was tested on a subset including only patients who underwent CT within 48 hours after symptom onset. Finally, to determine the effect of the manual corrections of the TBV, a sensitivity analysis including uncorrected TBV instead of corrected TBV was performed.

Analyses were performed using SPSS, Version 24.0.0.1 (IBM); R, Version 3.3.2 (http://www.r-project.org/); and R packages Hmisc, bootstepAIC, rms, pROC, and calibration curves.

RESULTS

The development cohort consisted of 409 patients, of whom 154 (38%) had a poor outcome (including patients who died) and 110 (27%) died within 6 months. The validation cohort consisted of 317 patients, of whom 140 (44%) had a poor outcome (including patients who died) and 80 (28%) died. Eight patients in the development cohort and 2 patients from the validation cohort were excluded due to movement and/or metal artifacts. Characteristics of patients in the development and validation cohorts are shown in Table 1. Patients in the validation cohort more often had a WFNS V grade on admission and more often underwent clipping as aneurysm treatment than patients in the development cohort. No other differences between the development and validation cohorts were found. For the whole group, the median time

Table 2: Model validation

	Internal		External	
	R^2, % (SD)	C-statistic (SD)	R^2, % (SD)	C-statistic (SD)
Unfavorable outcome				
TBV model	52 (0.6)	0.88 (0.01)	56 (1.8)	0.89 (0.01)
mFisher model	43 (0.9)	0.84 (0.01)	45 (3)	0.85 (0.01)
Location-specific model	50 (0.8)	0.88 (0.01)	56 (1)	0.89 (0.00)
Death				
TBV model	48 (0.8)	0.87 (0.01)	50 (1.2)	0.89 (0.00)
mFisher model	42 (1.4)	0.85 (0.01)	42 (2.6)	0.85 (0.01)
Location-specific mode	47 (1.0)	0.87 (0.01)	49 (1.3)	0.88 (0.01)

Note:–mFisher indicates modified Fisher scale.

between the onset of symptoms and CT used for volume measurement was 2.5 hours (interquartile range [IQR] = 1.1–14 hours). For the patients who had recurrent bleeding, the median time between the onset of symptoms and CT was 2.3 hours (IQR = 1.0–12.0 hours).

Model Selection and Performance
The following variables remained in the models after selection: TBV at first admission, WFNS grade at first admission, age, aneurysm size, and history of cardiovascular disease. The mean R^2 of the model was 54% ± 0.5% for poor outcome and 50% ± 1% for death. The mean R^2 values of the included variables were 35% ± 0% for TBV, 29% ± 0.7% for WFNS, 13% ± 0.7% for aneurysm size, 11% ± 0% for age, and 6% ± 1% for a history of cardiovascular disease. The models discriminated well between patients with and without a poor outcome (mean C-statistic = 0.89 ± 0.01) and mortality (mean C-statistic = 0.88 ± 0.01).

Model Validation
Internal validation showed that the optimism of the models was low for both poor outcome and death groups (Table 2). The predictive accuracy and discriminative power of the models in the external validation cohort for both outcome and death were comparable with those of the development cohort (Table 2).

The calibration plot showed a good correlation between predicted and observed outcomes in the validation set, though the models somewhat underestimated the risk of poor outcome (Fig 1A, -B). The mean slope was 1.1 ± 0.5 for poor outcome and 0.97 ± 0.03 for death. The calibration-in-the-large was 0.58 ± 0.03 for poor outcome and 0.12 ± 0.03 for death.

Secondary Models
The location-specific blood volume models showed a comparable explained variance to the TBV model for both poor outcome (mean R^2 = 56% ± 1% versus 56% ± 2%, P = .8) and death (mean R^2 = 49% ± 1% versus 50% ± 1%, P = .1). The mean R^2 values of the individual compartments were 17% ± 0% for cisternal volume, 12% ± 0% for IVH volume, and 14% ± 0% for IPH volume. The discriminative power was also similar for outcome (mean C-statistic = 0.89 ± 0.00 versus 0.89 ± 0.01, P = .4) and death (mean C-statistic = 0.88 ± 0.01 versus 0.89 ± 0.00, P = .66) (Table 2). The calibration

plots of the location-specific models showed a comparable calibration for both outcome and death with the TBV models. (Fig 1C, -D).

The explained variance of the models including the modified Fisher scale in the validation set was significantly lower than the models including the TBV for both poor outcome (R^2 = 45% ± 3% versus 56% ± 2%, P < .001) and death (R^2 = 42% ± 3% versus 50% ± 1%, P < .001). These models showed less accurate discrimination between patients with and without a poor outcome (mean C-statistic = 0.85 ± 0.01 versus 0.89 ± 0.01, P = .03) and death (mean C-statistic = 0.85 ± 0.01 versus 0.89 ± 0.00, P = .01) (Table 2). Calibration plots of the modified Fisher models showed a comparable calibration for outcome and a poorer calibration for death (Fig 1E, -F).

Sensitivity Analyses
After we included only patients who underwent aneurysm treatment, 337 patients remained in the development cohort and 248 patients remained in the validation cohort. Both the explained variance and the discriminative power were lower compared with the whole group (R^2 = 41% ± 1%; mean C-statistic = 0.84 ± 0.01).

After including only patients who had a brain CT within 48 hours after the initial bleeding, 340 and 281 patients remained in the development and validation cohorts, respectively. The explained variance increased (R^2 = 60% ± 1%), while the discriminative power did not change (mean C-statistic = 0.91 ± 0.00).

A sensitivity analysis including uncorrected TBV showed a slightly lower explained variance and discriminative power (R^2 = 49% ± 1%; mean C-statistic = 0.87 ± 0.01) compared with the corrected TBV.

DISCUSSION
The developed prediction models in patients with aSAH including quantified hemorrhage accurately discriminated patients with a favorable outcome from those with an unfavorable clinical outcome. The model including location-specific blood volumes showed similar results compared with the TBV model. Including the TBV in the model resulted in a higher predictive accuracy than including the modified Fisher scale instead.

One previous study also developed an outcome-prediction model including TBV.[15] In that study, various prognostic models for outcome were developed; the model including WFNS grade, age, and total bleeding volume (as continuous variables) had the highest predictive value. The internal performance of this model was similar to that of our models. Aneurysm size and a history of cardiovascular disease were not assessed in that study. An important difference between the previous and current study is the external validation of the developed models. External validation of prediction models is important because developing and validating a model on the same cohort may result in overly optimistic performance estimates.[16] Therefore, our models were validated on a registry of patients with SAH from a different center.
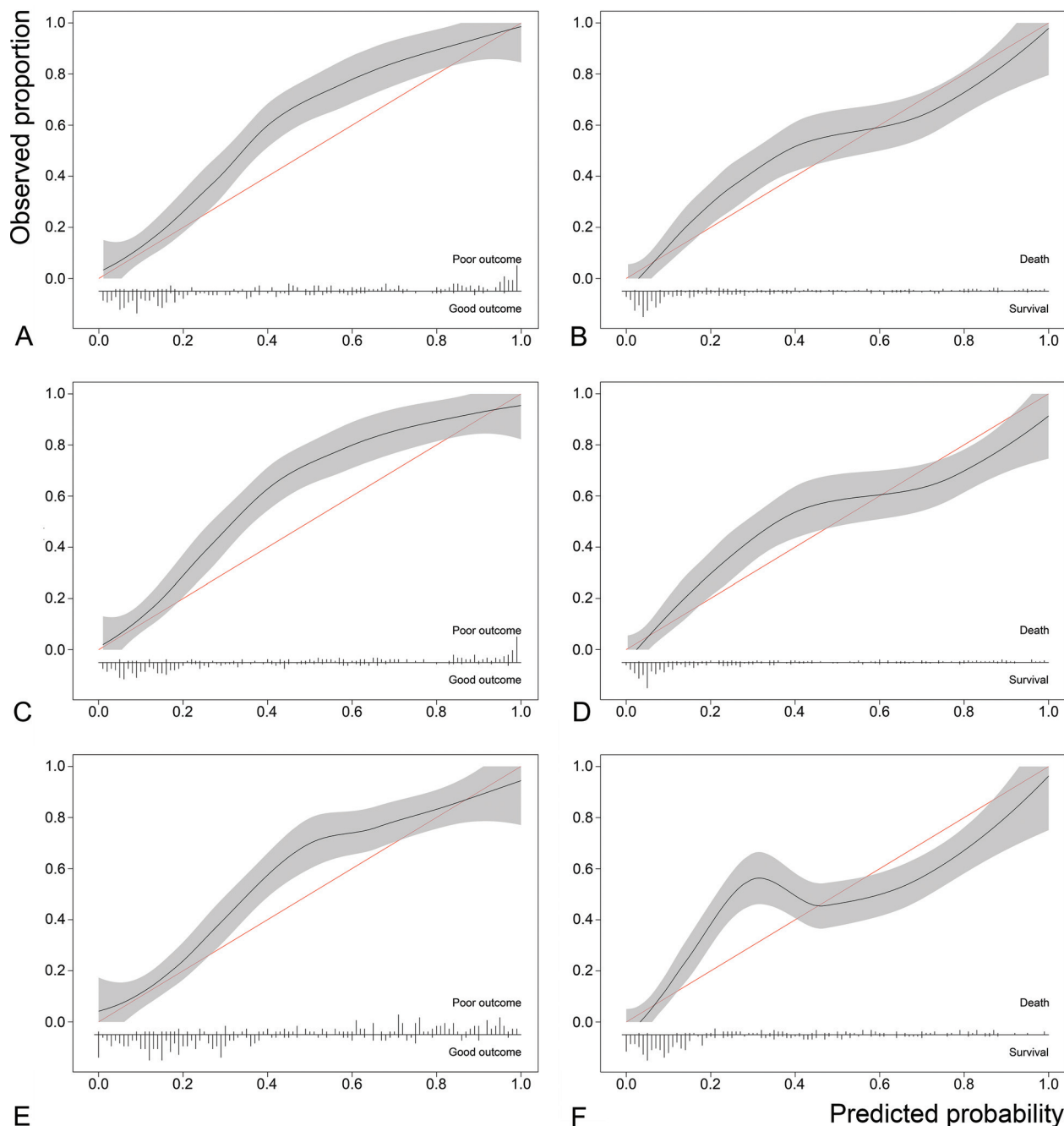
FIG 1. Calibration plots of the TBV model for poor outcome (*A*), the TBV model for death (*B*), the location-specific model for poor outcome (*C*), the location-specific model for death (*D*), the modified Fisher model for poor outcome (*E*), and the modified Fisher model for death (*F*).

The variables that remained in our models after bootstrapped backward selection are comparable with the variables included in previously developed models. A systematic review showed that previously developed models most frequently included age, neurologic condition (assessed with WFNS or Hunt and Hess grade), amount of bleeding (assessed with the Fisher or modified Fisher grade), and aneurysm size.[3] These are all variables that are available shortly after admission to the hospital. These factors not only directly determine the risk of poor outcome but could also increase the risk of the occurrence of late complications like delayed cerebral ischemia and hydrocephalus, which may further determine a patient's risk of poor outcome.[2,17] However, these

late complications are less suitable for early prediction of aSAH outcome.

The prognostic value of imaging variables in patients with aSAH has been debatable. The predictive accuracy of the amount of blood, assessed with the Fisher scale, was low in a large cohort of patients in an SAH trial.[18] Furthermore, in a recently published prediction model derived from that cohort, the addition of the amount of blood to a prediction model did not increase its predictive value.[4] The externally validated explained variance and predictive accuracy of our models were larger than those of previously published models that included a radiologic scale to assess the amount of blood.[3,4,19] One reason for this difference may be that categorization of a

continuous variable leads to loss of information.[20] Furthermore, the Fisher scale and modified Fisher scale have shown only moderate interobserver agreement.[6,21] Automatically quantified TBV is less observer-dependent; this feature may contribute to its higher predictive value.

Considering the cisternal, IVH, and IPH volumes instead of TBV in the model separately did not improve the predictive value of the model. Of these compartments, the amount of blood in the cisterns had the highest explained variance, followed by IPH and IVH volume. From the literature, it is known that the amount of blood in the cisterns is associated with outcome.[5,8,22] Several studies have found an association between IVH and/or IPH volume and poor outcome.[23-25] On the contrary, another study found that a large IPH volume was not associated with poor outcome. All patients included in this study were treated with surgical decompression. In this study, a large number of patients with small IPH volumes also had a poor outcome.[26] As for IVH, it may be that a relatively large blood clot in the ventricles results in less increase in intracranial pressure compared with an equal amount of blood in the subarachnoid space, which may result in a relatively better outcome.

After we included only patients who received treatment, the predictive value was lower compared with the whole group. Currently, the decision not to treat the aneurysm of a patient is based on a variety of factors such as the neurologic status, age, comorbidities, and aneurysm configuration. Patients who are not treated are most likely in poor condition and are estimated to have a very poor prognosis. Therefore, when we excluded those patients from the analysis, only patients with a relatively better prognosis remained. This finding may explain why the predictive value was somewhat lower in treated patients. Nevertheless, the model still showed a high discrimination between patients with and without a poor outcome who were treated. Including non-treated patients in the analyses may introduce some bias because not treating a patient itself will likely lead to a poor outcome. However, information about treatment is frequently not available at admission, and by including both treated and nontreated patients, the model can be applied to all patients who arrive at the emergency department.

An important strength of this study is the use of prospectively collected patient data, which were collected for observational studies in patients with SAH and thoroughly checked by trial nurses and treating clinicians. Furthermore, validation of the prediction model on an external cohort shows that the predictions made by the model are robust and that the model makes accurate predictions for new patients. Only variables that are quickly and easily obtainable after the patient's admission to the hospital were used. This step is necessary because quick decision-making in patients with aSAH is required. Thus, including factors that take considerable time to obtain would limit the clinical applicability of a model. The use of automatic hemorrhage-segmentation techniques to assess the TBV resulted in a more precise measurement of the blood volume compared with coarse grading scales.

On the contrary, the use of automatic hemorrhage-segmentation techniques can also be regarded as a weakness of this study. Hemorrhage-segmentation software needs to be available to apply these models to a hospital population. Furthermore, the currently used automated method required manual correction, which may limit its present usefulness in a clinical setting. However, increasingly more (machine learning) methods to segment structures in CT images are available.[27] The segmentations were corrected by only a single radiologist, which may have introduced some observer bias. Nevertheless, in the sensitivity analysis, we have shown that the influence of the observer on the results was minimal, suggesting that the effect of this single radiologist is modest at most. If a patient had a rebleed, the CT scans after the rebleed were used to determine the TBV. However, no WFNS scores after rebleed were systematically registered, so only admission WFNS scores were used. This use may have led to some limiting of the predictive value of the WFNS grade compared with the TBV. In future studies, use of the modified WFNS scale may further improve the predictive accuracy.[28] A more extensive assessment of the patients' comorbidities, for example using the Charlson Comorbidity Index, could further improve the model.[29] The sample size used for developing and validating these models was relatively small; however, a minimum of 10 cases of poor outcome per variable was still met.[14]

Accurate prediction can assist clinicians in decision-making and improve communication with patients and their families. Furthermore, it may reduce costs by allocating patients to the right intensity of treatment at admission. However, determining whether a model performs well enough to be applied in daily practice is difficult. A C-statistic of 0.89, indicating that the probability that a patient with a poor outcome is given a higher probability of a poor outcome than a patient without a poor outcome, is 0.89, is considered a reliable model.[14,30] A perfectly discriminative model would have a C-statistic of 1.00.

Thus, our model can, with a fairly high amount of certainty, predict which patients with aSAH will have a poor outcome. This study does not show whether clinical decisions based on our models actually improve patient outcome. Furthermore, it does not show whether statistically significant more accurate outcome prediction is also clinically relevant. To assess this question, an impact study that quantifies the effect of using a prognostic model on patient outcome and/or cost-effectiveness in a randomized trial is required.[31] Furthermore, a software package integrating automatic hemorrhage segmentation and clinical values needs to be developed before these models can be more broadly used.

CONCLUSIONS

The TBV-based prediction models for clinical outcome in patients with aSAH have a high predictive accuracy, higher than prediction models including the more commonly used modified Fisher scale. Including location-specific volumes did not improve the quality of the prediction models. The TBV models can accurately predict which patients will have a poor outcome early in the disease process and may aid clinicians in clinical decision-making.

REFERENCES

1. van Gijn J, Kerr RS, Rinkel GJ. **Subarachnoid haemorrhage.** *Lancet* 2007;369:306–18 CrossRef Medline

2. Vergouwen MD, Jong-Tjien-Fa AV, Algra A, et al. **Time trends in causes of death after aneurysmal subarachnoid hemorrhage: a hospital-based study.** *Neurology* 2016;86:59–63 CrossRef Medline

3. Jaja BN, Cusimano MD, Etminan N, et al. **Clinical prediction models for aneurysmal subarachnoid hemorrhage: a systematic review.** *Neurocrit Care* 2013;18:143–53 CrossRef Medline

4. Jaja BN, Saposnik G, Lingsma HF, et al; SAHIT oration. **Development and validation of outcome prediction models for aneurysmal subarachnoid haemorrhage: the SAHIT multinational cohort study.** *BMJ* 2018;360:j5745 CrossRef Medline

5. Kramer AH, Hehir M, Nathan B, et al. **A comparison of 3 radiographic scales for the prediction of delayed ischemia and prognosis following subarachnoid hemorrhage.** *J Neurosurg* 2008;109:199–207 CrossRef Medline

6. Woo PY, Tse TP, Chan RS, et al. **Computed tomography interobserver agreement in the assessment of aneurysmal subarachnoid hemorrhage and predictors for clinical outcome.** *J Neurointerv Surg* 2017;9:1118–24 CrossRef Medline

7. Fisher CM, Kistler JP, Davis JM. **Relation of cerebral vasospasm to subarachnoid hemorrhage visualized by computerized tomographic scanning.** *Neurosurgery* 1980;6:1–9 CrossRef Medline

8. Hijdra A, van Gijn J, Nagelkerke NJ, et al. **Prediction of delayed cerebral ischemia, rebleeding, and outcome after aneurysmal subarachnoid hemorrhage.** *Stroke* 1988;19:1250–56 CrossRef Medline

9. Boers AM, Zijlstra IA, Gathier CS, et al. **Automatic quantification after subarachnoid hemorrhage on non-contrast computed tomography.** *AJNR Am J Neuroradiol* 2014;35:2279–86 CrossRef Medline

10. Germans MR, Post R, Coert BA, et al. **Ultra-early tranexamic acid after subarachnoid hemorrhage (ULTRA): study protocol for a randomized controlled trial.** *Trials* 2013;14:143 CrossRef Medline

11. Frontera JA, Claassen J, Schmidt JM, et al. **Prediction of symptomatic vasospasm after subarachnoid hemorrhage: the modified Fisher scale.** *Neurosurgery* 2006;59:21–27 CrossRef Medline

12. Vergouwen MD, Vermeulen M, van Gijn J, et al. **Definition of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage as an outcome event in clinical trials and observational studies: proposal of a multidisciplinary research group.** *Stroke* 2010;41:2391–95 CrossRef Medline

13. Little R, An H. **Robust likelihood-based analysis of multivariate data with missing values.** *Statistica Sinica* 2004;14:949–68

14. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.* Springer-Verlag; 2009

15. Lagares A, Jiménez-Roldán L, Gomez PA, et al. **Prognostic value of the amount of bleeding after aneurysmal subarachnoid hemorrhage.** *Neurosurgery* 2015;77:898–907 CrossRef Medline

16. Bleeker SE, Moll HA, Steyerberg EW, et al. **External validation is necessary in prediction research: a clinical example.** *J Clin Epidemiol* 2003;56:826–32 CrossRef Medline

17. De Oliveira Manoel AL, Jaja BN, Germans MR, et al. **The VASOGRADE: a simple grading scale for prediction of delayed cerebral ischemia after subarachnoid hemorrhage.** *Stroke* 2015;46:1826–31 CrossRef Medline

18. Jaja BN, Lingsma H, Steyerberg EW, et al; on behalf of SAHIT investigators. **Neuroimaging characteristics of ruptured aneurysm as predictors of outcome after aneurysmal subarachnoid hemorrhage: pooled analyses of the SAHIT cohort.** *J Neurosurg* 2016;124:1703–11 CrossRef Medline

19. van Donkelaar CE, Bakker NA, Birks J, et al. **Prediction of outcome after aneurysmal development.** *Stroke* 2019;50:837–44 CrossRef

20. Altman DG, Royston P. **The cost of dichotomising continuous variables.** *BMJ* 2006;332:1080 CrossRef Medline

21. van der Jagt M, Hasan D, Bijvoet HW, et al. **Interobserver variability of cisternal blood on CT after aneurysmal subarachnoid hemorrhage.** *Neurology* 2000;54:2156–58 CrossRef Medline

22. Dengler NF, Diesing D, Sarrafzadeh A, et al. **The Barrow Neurological Institute Scale revisited: predictive capabilities for cerebral infarction and clinical outcome in patients with aneurysmal subarachnoid hemorrhage.** *Neurosurgery* 2017;81:341–49 CrossRef Medline

23. Jabbarli R, Reinhard M, Roelz R, et al. **Intracerebral hematoma due to aneurysm rupture: are there risk factors beyond aneurysm location?** *Neurosurgery* 2016;78:813–20 CrossRef

24. Kramer AH, Mikolaenko I, Deis N, et al. **Intraventricular hemorrhage volume predicts poor outcomes but not delayed ischemic neurological deficits among patients with ruptured cerebral aneurysms.** *Neurosurgery* 2010;67:1044–52 CrossRef Medline

25. Jabbarli R, Reinhard M, Roelz R, et al. **The predictors and clinical impact of intraventricular hemorrhage in patients with aneurysmal subarachnoid hemorrhage.** *Int J Stroke* 2016;11:68–76 CrossRef Medline

26. Zijlstra IA, van der Steen WE, Verbaan D, et al. **Ruptured middle cerebral artery aneurysms with a concomitant intraparenchymal hematoma: the role of hematoma volume.** *Neuroradiology* 2018;60:335–42 CrossRef Medline

27. Litjens G, Kooi T, Bejnordi BE, et al. **A survey on deep learning in medical image analysis.** *Med Image Anal* 2017;42:60–88 CrossRef Medline

28. Sano H, Satoh A, Murayama Y, et al; members of the 38 registered institutions and WFNS Cerebrovascular Disease & Treatment Committee. **Modified World Federation of Neurosurgical Societies subarachnoid hemorrhage grading system.** *World Neurosurg* 2015;83:801–07 CrossRef Medline

29. Bar B, Hemphill JC. **Charlson comorbidity index adjustment in intracerebral hemorrhage.** *Stroke* 2011;42:2944–46 CrossRef Medline

30. Altman DG, Vergouwe Y, Royston P, et al. **Prognosis and prognostic research: validating a prognostic model.** *BMJ* 2009;338:b605 CrossRef Medline

31. Moons KGM, Altman DG, Vergouwe Y, et al. **Prognosis and prognostic research: application and impact of prognostic models in clinical practice.** *BMJ* 2009;338:b606 CrossRef Medline