A Discriminative Prediction Model of Neurological Outcome for Patients Undergoing Surgery of Brain Arteriovenous Malformations

Julian Spears, MD; Karel G. TerBrugge, MD; Mona Moosavian, BSc; Walter Montanera, MD; Robert A. Willinsky, MD; M. Christopher Wallace, MD, MSc; Michael Tymianski, MD, PhD

Background and Purpose—To develop and validate a discriminative model for predicting neurological morbidity after brain arteriovenous malformation (bAVM) surgery.

Methods—Of 233 consecutive, prospectively enrolled patients undergoing bAVM surgery, the first 175 were used to derive, and the last 58 to validate, the prediction model. Demographic and angiographic factors were related to modified Rankin Scale scores assigned before, within 72 hours, at 7 days and at ≥1 year after surgery to seek predictors of postoperative neurological deficits (modified Rankin Scale score ≥3). These factors included nidus size, eloquence, venous drainage, diffuseness, white matter configuration, arterial perforator supply and associated aneurysms.

Results—Brain eloquence, diffuse nidus and deep venous drainage were significant predictors of early disabling neurological deficits (odds ratios of 4.33, 3.49 and 2.38, respectively). The rounded odds ratios form a weighted 9-point prediction model (maximum scores for eloquence, diffuseness, deep drainage). The score discriminated the probability of experiencing both early (first week) and permanently (at ≥1 year) disabling neurological deficits as follows: 0 to 2: 1.8%, 3 to 5: 17.4%, 6 to 7: 31.6%, >7: 52.9% for early and 0 to 2: 1.8%, 3 to 5: 4.4%, 6 to 7: 18.4%, >7: 32.4% for permanently disabling outcomes. The discrimination of the model was 0.80 with 2.8% optimism. Validation in the second patient cohort revealed good performance at risk stratification.

Conclusions—Relative weights assigned to brain eloquence, diffuse nidus morphology and deep venous drainage of a bAVM provide a simple and discriminative prediction model for neurological outcome after bAVM surgery. (Stroke. 2006;37:1457-1464.)

Key Words: cerebral arteriovenous malformation decision aid decision support models outcome surgery

Brain arteriovenous malformations (bAVMs) impart affected individuals with a life-long risk of hemorrhagic stroke.1 They may also cause seizures, headaches, and progressive neurological deficits.1,2 BAVMs have diverse distributions, morphologies and angioarchitectural associations including brain aneurysms; thus, their clinical and morphological heterogeneity requires that treatments be individualized. When surgery is considered, grading schemes are used to aid in predicting the associated risks.4–6 The most frequently used is the Spetzler-Martin (SM) grading,4 in which ordinal scores are assigned to 3 variables: bAVM size, eloquence of location, and venous drainage pattern. Such grading systems facilitate the stratification of BAVM characteristics between patients and the comparison of clinical outcomes between institutions.4–8 However, simplicity carries the price of a reduced discrimination between grades in the prediction of postoperative neurological outcomes.8 For example, different patients harboring a SM grade III bAVM may be deemed to have different surgical risks owing to features unaccounted for by the grading system, and this has caused some to suggest adjustments to the SM scale.8

Our goal was to develop and prospectively validate a simple but also discriminative prediction model for neurological outcome after bAVM surgery. Thus, in addition to the SM variables above, we also studied the influence of additional angioarchitectural features of BAVMs believed to impact on surgical morbidity.7–12 From this analysis, we generated and internally, as well as prospectively, validated a simple, discriminative model for predicting neurological outcome after bAVM surgery.

Materials and Methods

The University of Toronto Brain AVM Study Group collects an ongoing, prospective database of demographic, clinical, morpholog-
logical and treatment data on all bAVM patients seen at our institution. Between 1990 and 2005, 233 of a total of 1058 bAVM patients had surgery at our hospital. We used the first 175 consecutive patients (derivation cohort) to develop and internally validate the model. Data from the subsequent 58 consecutive patients (validation cohort) were used to validate the model. This 3:1 size ratio between the derivation and validation cohorts is generally accepted as appropriate for prospective validation of a model.11

For model derivation, all morphological bAVM features were analyzed from magnetic resonance images (MRI) and conventional cerebral angiograms by a neuroradiologist (K.G.T.) blinded to all patient data. Three were taken from the SM system: nidus size, surrounding brain eloquence and deep venous drainage. The additional variables were diffuse nidus morphology, a deep white matter configuration, a deep arterial supply and associated aneurysms. All predictors were recorded as binary variables/exposures. Size was dichotomized into <3 cm or ≥3 cm based on MRI measurements (only 2 patients with bAVMs >6 cm underwent surgery). Interobserver variability for key measures was determined between 2 neuroradiologists (K.G.T., R.A.W.) and again between a neuroradiologist (K.G.T.) and a neurosurgeon (M.T.).

For scoring consistency, the physicians assessing the bAVM angioarchitectural variables were required to use solely the criteria contained in the following operational definitions of the variables: Eloquence: as in the SM system4 - a nidus location in sensorimotor, language, and visual cortex, thalamus, hypothalamus, internal capsule, brain stem, cerebellar peduncles and deep cerebellar nuclei. Deep drainage: any nidus drainage into the deep venous system. Deep arterial supply: when arteries penetrated the bAVM through brain parenchyma rather than from the cortical surface. Deep white matter configuration: a nidus located in the deep white matter tracts and not exhibiting the classic wedge-shaped, corticoventricular configuration. Associated aneurysms included prenidal, intranidal or postnidal aneurysms. Aneurysms on arteries remote from the bAVM were excluded from analysis. Diffuse nidus morphology was defined as a nidus in which intervening brain parenchyma was visualized by MRI or was anticipated on cerebral angiography (Figure 1). This contrasts with a compact nidus, which exhibits clearly defined margins and no intervening brain parenchyma (Figure 2).

The primary outcome measures were the occurrence of new neurological deficits (within 72 hours of surgery) and early disabling MRI or was anticipated on cerebral angiography (Figure 1). This as a nidus in which intervening brain parenchyma was visualized by

Statistical Analyses

Data Analysis

Except where otherwise indicated, analyses focused on clinical outcomes scored by mRS. Univariable analysis was used to determine the significance of each variable in predicting a new neurological deficit, early, and permanently disabling neurological outcome. All variables related to early disabling neurological outcome in univariable screens (P<0.10) were entered into a forward, stepwise logistic regression model setting the entry and elimination thresholds at P<0.01. Odds ratios for each predictor variable were rounded to the nearest integer to create relative weightings that reflect the relative importance of each variable.14 From this we derived a model with 4 risk strata for predicting neurological outcome after surgery.

Prediction Model Evaluation

Discrimination of the model measures its ability to classify the patients into those who will incur neurological morbidity after surgery versus those who will not. This was determined by calculating the area under a receiver operator characteristic curve (ROC area) for each outcome (new neurological deficit, early disabling outcome and permanently disabling outcome). An ROC area of 0.5 indicates no discrimination, whereas an ROC area of 1.0 indicates perfect discrimination. The model was calibrated by determining the discrepancies between the predicted and actual outcomes and obtaining the Hosmer-Lemeshow goodness-of-fit statistic.

Prospective validation was performed by applying the model to 58 subsequently treated consecutive patients and observing the trend in the model’s predictive accuracy for each neurological outcome.

Results

Patient and bAVM Characteristics

Demographic and bAVM characteristics of the patient cohort are presented in Table 1. The ages and sex distribution in the derivation and validation cohorts are consistent with previous reports.1 Of the 233 patients (derivation cohort n=175, validation cohort n=58), 38.6% underwent at least 1 embolization as part of their treatment, 5.1% had radiation and 4.3% had both embolization and radiation in addition to surgery. In select cases (1.7%), surgery was performed on patients with progressive neurological deficits (Table 1). The incidence of left and right sided bAVMs, and location in eloquent brain, was comparable in the derivation and validation cohorts.

Interobserver Variability

Interobserver variability (κ statistic) for assessing key angioarchitectural variables, assessment of SM grade and proposed model risk category, was measured both between 2 neuroradiologists (K.G.T., R.A.W.) and between a neuroradiologist (K.G.T.) and a neurosurgeon (M.T.; Table 2). Notably, the determination of SM grade and of our proposed score exhibit similar degrees of agreement, with κ values consistent with previous reports.15,16 Moreover, agreement for determining AVM nidus diffuseness, a significant parameter in our model (below), was as good as or better than the other parameters, including SM variables. In brief, interobserver agreement is considered poor with a κ statistic between 0 and 0.20, fair between 0.21 to 0.40, moderate between 0.41 to 0.60, substantial between 0.61 to 0.80, and near perfect between 0.81 to 1.00.17

Model Derivation

There were no differences between genders, nor an effect of age, nor an impact of preoperative embolization on the occurrence of new neurological deficits, early or permanently disabling neurological outcomes (supplemental Table Ia, available online at http://stroke.ahajournals.org). Therefore, these variables were excluded from further multivariable analysis.

New Neurological Deficits (within 72 hours)

A new neurological deficit was detected in 69 of 175 patients in the derivation cohort (39.4%). All angioarchitectural variables,
Figure 1. Representative diffuse bAVM in the occipital and temporal lobes. A through C, MRI images. D through G, Angiograms. D and F, Right internal carotid angiogram anterior-posterior and lateral views. E and G, Left vertebral artery angiogram anterior-posterior and lateral views.
except for associated aneurysms, were related to new neurological deficits on univariable analyses (supplemental Table Ib), but on multivariable modeling only eloquent brain location and diffuse nidus morphology were predictive (Table 3).

Figure 2. Representative compact bAVM in the left motor area. A through C, MRI images. D and E, Preoperative anterior-posterior and lateral angiogram. F, Preoperative 3 dimensional angiogram. G, Postoperative anterior-posterior and lateral angiogram.

Early Disabling Neurological Outcome (within 7 days)
Early disabling outcomes occurred in 24 (13.7%) or 39 (22.3%) of 175 patients as assessed using GOS and mRS, respectively, suggesting a higher sensitivity of the mRS in
TABLE 1. Demographic and Angioarchitectural Characteristics of Derivation and Validation Cohorts

<table>
<thead>
<tr>
<th></th>
<th>Derivation Cohort n (% of total)</th>
<th>Validation Cohort n (% of total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>175 (100)</td>
<td>58 (100)</td>
</tr>
<tr>
<td>Demographic variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (y) mean, median, range</td>
<td>40, 40, 9–69</td>
<td>39, 36, 14–65</td>
</tr>
<tr>
<td>Gender (M:F, %)</td>
<td>88:87 (50.3:49.7)</td>
<td>26:32 (44.8:55.2)</td>
</tr>
<tr>
<td>Side (RL; %)</td>
<td>74:101 (42.3:57.7)</td>
<td>27:31 (46.5:53.5)</td>
</tr>
<tr>
<td>Presentation (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemorrhage</td>
<td>96 (54.9)</td>
<td>21 (36)</td>
</tr>
<tr>
<td>Seizure</td>
<td>55 (31.4)</td>
<td>29 (50)</td>
</tr>
<tr>
<td>Incidental/Asymptomatic</td>
<td>21 (12.0)</td>
<td>7 (12)</td>
</tr>
<tr>
<td>Progressive neurological deficit</td>
<td>3 (1.7)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Preoperative embolization (%)</td>
<td>61 (34.8)</td>
<td>29 (50)</td>
</tr>
<tr>
<td>Angioarchitectural variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;3 cm</td>
<td>120 (68.6)</td>
<td>48 (82.7)</td>
</tr>
<tr>
<td>3–6 cm</td>
<td>53 (30.3)</td>
<td>10 (17.3)</td>
</tr>
<tr>
<td>&gt;6 cm</td>
<td>2 (1.1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Elocuence (%)</td>
<td>100 (57.1)</td>
<td>25 (43.1)</td>
</tr>
<tr>
<td>Deep venous drainage (%)</td>
<td>70 (40.0)</td>
<td>16 (27.8)</td>
</tr>
<tr>
<td>Diffuse nidus (%)</td>
<td>71 (40.6)</td>
<td>15 (25.9)</td>
</tr>
<tr>
<td>Arterial perforator supply (%)</td>
<td>53 (30.3)</td>
<td>N/R</td>
</tr>
<tr>
<td>Deep white matter (%)</td>
<td>42 (24)</td>
<td>N/R</td>
</tr>
<tr>
<td>configuration (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associated aneurysm* (%)</td>
<td>66 (37.7)</td>
<td>N/R</td>
</tr>
</tbody>
</table>

*Includes prenidal, intranidal and postnidal aneurysms. N/R indicates not recorded.

scoring neurological morbidity. mRS scores were used in all subsequent tests. In univariable analyses, bAVM size, eloquence, deep venous drainage, diffuse nidus and an arterial perforator supply were significant (supplemental Table Ib). On multivariable analysis, the number of independent predictors was reduced to only 3: eloquent nidus location, diffuse nidus morphology, and deep venous drainage (Table 3).

Permanently Disabling Neurological Outcomes (≥1 year)
These occurred in 10 (5.7%) or 21 (12.0%) of 175 patients as assessed using GOS and mRS, respectively. Thirty-day mortality was 1.5% (n=2). One was attributable to a postoperative hemorrhage, the other to a pulmonary embolus. One-year mortality was 2.9% (6 patients). Of the additional 4 deaths, 1 was attributable to a brain hemorrhage and 3 to medical complications of pre-existing neurological disabilities. Additional details are provided (supplemental Table II, available online at http://stroke.ahajournals.org).

Model Construction
This was undertaken using data for early disabling neurological outcomes because too few patients experienced permanently disabling outcomes to provide the requisite statistical power. The model was derived using the 3 significant predictor variables from multivariable analysis: eloquence, diffuse nidus and deep venous drainage. The derived odds ratios (Table 3), when rounded to the nearest integer, form a 9-point stratified risk score in which each predictor’s contribution is related to its relative weight (eloquence=4, diffuse nidus=3, deep venous drainage=2). When applied to the derivation cohort, this 9-point score discriminates the % probability of incurring an early disabling neurological outcome as follows: Low Risk (0 to 2 points)=1.8%, Moderate Risk (3 to 5 points)=17.4%, High Risk (6 to 7 points)=31.6%, Very High Risk (>7 points)=52.9% (Table 4a). In comparison, applying the SM system to the same data produced the following: SM grade I: 2.1%, II: 9.4%, III: 17.3%, IV: 39.1%. There were no Grade V patients.

Prediction of New Deficits and Permanently Disabling Outcomes
When the 9-point model was applied to predicting new neurological deficits (NND) and permanently disabling neurological outcomes (PD), it retained its ability to stratify patients into 4 risk categories (Table 4a) further demonstrating its utility. By contrast, applying the SM system to the same PD outcome suggested a loss of discrimination in predicting disabilities in the lower grades (SM grade I: 2.1%, II: 5.7%, III: 1.9%, IV: 21.7%).

Measures of Discrimination
The discrimination of the model (ROC area; Methods) was determined to test its ability to correctly classify patients into those who will incur neurological morbidity. The model performed well for predicting early disabling neurological outcome (ED) with an ROCED area=0.80 using mRS and 0.82 using GOS, as compared with the ROCED area of 0.76 for the SM system using mRS, and 0.74 using GOS in our derivation cohort. The discrimination of our proposed model for predicting NND and PD was also high, with ROCNND area=0.75 and ROCPD area=0.79 using mRS. By comparison, the ROCED area for the SM system was 0.69, suggesting a loss of discrimination in predicting permanent deficits. An ROC curve area of ≥0.70 is considered clinically useful.18 For example, a well respected model for predicting adverse outcomes after coronary artery bypass grafting had an ROC area of 0.73.19

Internal Model Validation
Optimism is the degree to which a prediction model can overstate its predictive ability. Optimism was estimated by repeatedly calculating the ROC area using 1000 bootstrap resamples of the data set and calculating the difference between the ROC area of the test sample with that of the average of the ROC areas of all bootstrap samples. Optimism was determined as 2.8%. Calibration of the model by predicting early disabling outcomes (Table 4a) showed it to slightly overestimate the risk in the lowest and highest strata and underestimate the risk in the 2 intermediate categories. The Hosmer-Lemeshow statistic was 3.69 (P=0.45) indicating satisfactory goodness-of-fit.

Prospective Model Validation
After model derivation, prospective validation was undertaken in a subsequent 58 consecutive bAVM patients. A NND was detected in 20 of 58 patients (34.5% versus 39.4% in the derivation cohort). Early disabling and permanently disabling...
outcomes occurred in 10 patients (17.2% versus 22.3% in the derivation cohort) and 2 patients (3.4% versus 12.0% in the derivation cohort), respectively. From the prospective application of the model to the validation cohort, it retained its ability to stratify patients into defined risk categories for the incidence of NNDs (0 to 2: 15.6%, 3 to 5: 53.3%, 6 to 7: 60.0%, >7: 100%), ED outcome (0 to 2: 9.3%, 3 to 5: 26.6%, 6 to 7: 30.0%, >7: 0.0%), and PD outcome (0 to 2: 0.0%, 3 to 5: 0.0%, 6 to 7: 20.0%, >7: 0.0%), respectively (Table 4b).

The model slightly underestimated the incidence of new neurological deficits and early poor outcomes in the validation cohort, but clearly shows increasing morbidity associated with increasing risk category.

### TABLE 2. Interobserver Agreement Between Neuroradiologist 1 and 2 and Neuroradiologist 1 and Neurosurgeon for Predictor Variables, SM Grade and Prediction Model Risk Category

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>Neuroradiologist 1 &amp; Neuroradiologist 2</th>
<th>Neuroradiologist 1 &amp; Neurosurgeon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kappa [95% CI]</td>
<td>Kappa [95% CI]</td>
</tr>
<tr>
<td>Eloquence</td>
<td>0.47 [0.24, 0.70]</td>
<td>0.53 [0.30, 0.75]</td>
</tr>
<tr>
<td>Diffuse Nidus</td>
<td>0.62 [0.36, 0.87]</td>
<td>0.67 [0.46, 0.88]</td>
</tr>
<tr>
<td>Deep Venous Drainage</td>
<td>0.72 [0.53, 0.91]</td>
<td>0.70 [0.50, 0.89]</td>
</tr>
<tr>
<td>Size ≥ 3 cm</td>
<td>0.52 [0.21, 0.83]</td>
<td>0.55 [0.29, 0.80]</td>
</tr>
<tr>
<td>Spetzler-Martin Grade*</td>
<td>0.38 [0.19, 0.58]</td>
<td>0.43 [0.25, 0.59]</td>
</tr>
<tr>
<td>Proposed Model Score Category*</td>
<td>0.41 [0.20, 0.62]</td>
<td>0.49 [0.31, 0.68]</td>
</tr>
</tbody>
</table>

*Weighted κ statistic.

### Discussion

A prediction model must balance simplicity and clinical utility. Such utility requires the model to be sufficiently discriminative to separate different types of bAVMs into meaningful risk categories. In the case of the SM system, the authors reported PD neurological outcomes 6 months postoperatively at 0% for grade I-III, and 21% and 17% for grade IV and V bAVM patients, respectively. In another study, bAVM patients graded by the SM system had a disabling neurological outcome at hospital discharge of 8%, 17%, 11%, 19% and 67% for grades I-V, respectively. At 1 year, these figure were 8%, 6%, 4%, 4% and 33% for grades I–V. Thus, the SM system includes parameters that have predictive value.

### TABLE 3. Odds Ratios Generated for Predictors of NNDs and ED Outcomes From Multivariable Analyses

<table>
<thead>
<tr>
<th>Predictors of</th>
<th>NNDs after surgery [95% CI]</th>
<th>Diffuse Nidus Morphology</th>
<th>Deep Venous Drainage</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNDs after surgery [95% CI]</td>
<td>2.77* [1.33, 5.77]</td>
<td>2.77* [1.34, 5.75]</td>
<td>1.66 [0.82, 3.36]</td>
<td>1.76 [0.83, 3.75]</td>
</tr>
<tr>
<td>ED neurological outcome (1 wk) after surgery [95% CI]</td>
<td>4.33* [1.52, 12.38]</td>
<td>3.49* [1.51, 8.08]</td>
<td>2.38* [1.04, 5.44]</td>
<td>N/A</td>
</tr>
</tbody>
</table>

*P<0.01
on outcome, but it stratifies patients into essentially 2 categories: low- and high-risk groups. Our data reflect a similar experience, with patients in SM Grades I–III having low, and SM grade IV having high, risk of incurring a PD outcome. Our work indicates that discrimination may be enhanced by incorporating 3 new aspects: First, by the recognition of nidus diffuseness as a key predictor variable. Second, by assigning relative weights to key predictor variables according to their importance. Third, by grading outcome using the mRS, a system commonly used to measure neurological deficit on multivariable testing, whereas it was important in other classification schemes. As the bAVM size increases, it is more likely to impinge on eloquent brain, have deep venous drainage and have a diffuse component. Thus, these 3 independent predictors likely account for the surgical risks associated with larger bAVMs. However, we cannot exclude the possibility that very large nidus size (>6 cm) could be independently related to surgical risk because only 2 patients in our study had such bAVMs.

Lastly, bAVMs with associated aneurysms have been reported to have a higher hemorrhage rate, and this has motivated some surgeons to operate in order to remove the aneurysm. Because associated aneurysms were unrelated to adverse surgical outcomes, this factor need not influence treatment decisions in bAVM patients unless, of course, the aneurysm is the cause of a brain hemorrhage. The overall clinical results highlight the safety of surgical extirpation of appropriately selected bAVMs. Though the absolute risks associated with each risk stratum in our model may be most applicable to neurosurgical centers with significant experience with bAVMs surgery, we believe that the proposed model is generalizable.
Acknowledgments

J.S. is a recipient of a Heart and Stroke Foundation of Canada and Merck Frosst Canada research fellowship. M.M. is a recipient of an Ontario Heart and Stroke Foundation John D. Schultz Science Student Scholarship award. M.T. is a Clinician-Scientist (Phase II) of the Canadian Institutes of Health Research. The authors wish to acknowledge Dr Fran Cook, Harvard School of Public Health, for his assistance in reviewing the manuscript.

References

A Discriminative Prediction Model of Neurological Outcome for Patients Undergoing Surgery of Brain Arteriovenous Malformations

Julian Spears, Karel G. TerBrugge, Mona Moosavian, Walter Montanera, Robert A. Willinsky, M. Christopher Wallace and Michael Tymianski

Stroke. 2006;37:1457-1464; originally published online May 11, 2006;
doi: 10.1161/01.STR.0000222937.30216.13
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2006 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/37/6/1457

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org//subscriptions/