Published in final edited form as:

J Neurosurg Pediatr. 2013 May ; 11(5): 547–551. doi:10.3171/2013.2.PEDS12465.

Intra and Interrater Reliability of the Pediatric AVM Compactness Score

Fabio A. Frisoli, BS^{1,2}, Shih-Shan Lang, MD^{1,2}, Arastoo Vossough, MD, PhD³, Anne Marie Cahill, MBBCh³, Gregory G. Heuer, MD, PhD^{1,2}, Hisham M. Dahmouch, MBBCh³, Phillip B. Storm, MD^{1,2}, and Lauren A. Beslow, MD, MSCE⁴

¹Department of Neurosurgery, University of Pennsylvania Medical Center

²Division of Neurosurgery, Children's Hospital of Philadelphia

³Department of Radiology, Children's Hospital of Philadelphia

⁴Departments of Pediatrics and Neurology, Yale University School of Medicine

Abstract

Introduction— Cerebral arteriovenous malformations (AVM) have a higher post-resection recurrence rate in children than in adults. Our previous study demonstrated that a diffuse AVM (low compactness score) predicts post-resection recurrence. The aims of this study were to evaluate the intra and interrater reliability of our AVM compactness score.

Methods— Angiograms of 24 subjects assigned a preoperative compactness score (scale of 1 to 3, 1=most diffuse, 3= most compact) in our previous study were re-rated by the same pediatric neuroradiologist 9 months later. A pediatric neurosurgeon, pediatric neuroradiology fellow, and interventional radiologist blinded to each other's ratings, the original ratings, and AVM recurrence also rated each AVM's compactness. Intra and interrater reliability were calculated using the kappa (κ) statistic.

Results— Of the 24 AVMs, scores by the original neuroradiologist were a score of 1 in 6 subjects, 2 in 16 subjects, and 3 in 2 subjects. Intrarater reliability was 1.0. κ among the four raters was 0.69 [95% confidence interval (CI) 0.44–0.89], which indicates substantial reliability. The interrater reliability between the neuroradiologist and neuroradiology fellow was moderate (κ =0.59, 95%CI 0.20–0.89) and was substantial between the neuroradiologist and neurosurgeon (κ =0.74, 95%CI 0.41–1.0). The neuroradiologist and interventional radiologist had perfect agreement (κ =1.0).

Conclusions— Intra and interrater reliability of AVM compactness scoring were excellent and substantial, respectively. These results demonstrate that the AVM compactness score is reproducible. However, the neuroradiologist and interventional radiologist had perfect agreement, which indicates that the compactness score is most accurately applied by those with extensive angiography experience.

Keywords

arteriovenous malformation; AVM; surgery; angiogram; recurrence; pediatric; compactness; reliability

Corresponding Author: Lauren A. Beslow, MD, MSCE, Departments of Pediatrics and Neurology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA, lauren.beslow-kaye@yale.edu.

Disclosures

None of the authors has any disclosures or conflicts of interest.

Introduction

Cerebral arteriovenous malformations (AVMs) in the pediatric population follow a distinct natural history that is different from those in adults. Pediatric AVMs carry an increased yearly risk of hemorrhage with resultant significant morbidity and mortality.^{1,2} Consequently, AVMs are often treated aggressively with microsurgical resection, endovascular embolization, radiosurgery, or a combination of these treatment modalities.^{5,12} While the true post-operative recurrence rate of cerebral AVMs in the pediatric population has not been studied extensively, the literature suggests that it is higher than in adults.^{1,7} Many centers do not follow adult patients with serial vascular imaging after a post-operative image has demonstrated an absence of residual AVM. It is postulated that higher post-operative recurrence in children is due in large part to active angiogenesis of the immature cerebral vasculature.^{6,8,9}

Certain characteristics of the AVM also likely contribute to a higher recurrence risk. The Spetzler-Martin grading scale, which considers AVM nidus size, location in eloquent cortex, and venous drainage, has been widely used to predict the incidence of post-operative neurological complications.¹³ A large cohort study of mostly adult patients with AVM previously demonstrated that diffuseness of the nidus and deep perforating arterial supply are two additional characteristics that predict worse surgical outcome. Patients with non-compact niduses were 38% less likely to have overall improvement in pre-operative symptoms and 5 times more likely to have post-operative clinical deterioration.³ Our recent data on pediatric subjects who underwent surgical resection of cerebral AVMs demonstrated that AVM compactness predicts AVM recurrence after resection.¹¹ Of 24 pediatric subjects who were assigned a preoperative angiographic compactness score by a neuroradiologist (scale of 1–3, 1 designating the most diffuse), all patients with recurrence had a score of 1 ($p = 0.0003$). This represents the first documentation in the literature of a compactness score for pediatric cerebral AVMs. The compactness score was graded by a single neuroradiologist at one time point, thus the reproducibility of the score among raters is unknown.

The aims of the current study are to determine the intra and interrater reliability of angiographic grading of the compactness score for pediatric AVMs. The score may be a helpful tool for radiologists and neurosurgeons in predicting recurrence of cerebral AVMs in children after surgical resection. Such a score could eventually lead to a reduction in the use of serial angiograms in some children at low-risk for recurrent AVM, a reduction that limits exposure to radiation and sedation.

Methods

Study protocol

In our previous study, 24 pediatric patients with angiographic evidence of an AVM were assigned a pre-operative AVM compactness score visually by a board-certified pediatric neuroradiologist.¹¹ This non- computerized compactness score was graded on a scale of 1 to 3 and was different from the computer-aided method of nidus compactness described by Du et al.³ A score of 1 designates a diffuse AVM and a score of 3 designates a compact AVM. For the current study, all clinically obtained pre-operative vascular studies including conventional angiography, CT angiography, and MRI/MRA were used. For assessment of intrarater reliability, the same neuroradiologist re-scored the 24 AVMs 9 months after the original scoring had been performed.

Three new raters, a pediatric neurosurgeon, a pediatric interventional radiologist, and a pediatric neuroradiology fellow, were then trained by the principal neuroradiologist rater to

score AVM compactness. The raters attended a training session with the principal rater to review the definitions of a compactness score of 1 (most diffuse), 2, and 3 (most compact) (Table 1, Figures 1–3). Of note, a diffuse AVM (compactness score of 1) may be large or small, and the rating system is based on the overall compactness of the nidus and perinidal anomalous vessels. Additionally, it is very important to assess the images closely for very small and narrow caliber abnormal vessels that may not be evident immediately on a gross overview of the images. The presence of these very small, narrow caliber abnormal vessels with intervening brain tissue indicates a diffuse AVM that should be scored as 1. The raters scored 8 practice images from patients that were not included in this study and reviewed the scoring as a group to calibrate the ratings. The raters scored all 24 subjects' AVMs in the same order to limit variability for each subject due to rater experience. The raters were blinded to the other raters' scores, to the original scores, and to post-operative recurrence. The institutional review board approved the study.

Statistical analysis

Analyses were conducted using STATA version 11.1 (STATA Corporation, College Station, TX). For the purposes of this study, the original neuroradiologist's scoring was used as the gold standard for reliability calculations. Intra and interrater reliability of the AVM compactness scores were assessed using the parameter kappa (κ). Interrater reliability was assessed across all four raters and also between each new rater and the principal rater. κ was considered moderate agreement if 0.41–0.60, substantial agreement if 0.61–0.80, and almost perfect (excellent) if 0.81–1.00.¹⁰ A two-sided probability value of <0.05 was considered statistically significant.

Results

Patient Characteristics

In our previous study, 24 subjects undergoing an AVM resection were assigned a pretreatment compactness score. Eleven subjects were male (45.8%). The median age at the time of AVM resection was 10.5 years (interquartile range 8–14 years).

Prior Compactness Score

A compactness score of 1 was assigned to six subjects (25%), a score of 2 to sixteen subjects (66.7%), and a score of 3 to two subjects (8.3%). All four subjects with a recurrent AVM had a compactness score of 1, and lower compactness score was associated with recurrence after resection ($p = 0.0003$).¹¹

Intra and Interrater Reliability

The intrarater reliability for the neuroradiologist was 1.0; there was perfect agreement between the initial and repeat scoring. Of the 72 assigned scores from the new study raters (3 raters grading 24 subjects), 64 classifications (88.9%) were the same as the gold standard scores assigned by the attending pediatric neuroradiologist. The κ among the four raters was 0.69 (95% CI 0.44–0.89). The interrater reliability between the neuroradiologist and the neuroradiology fellow was 0.59 (95% CI 0.20–0.89). The fellow misclassified two true 1s as 2s and three true 2s as 1s. The interrater reliability between the neuroradiologist and the pediatric neurosurgeon was 0.74 (95% CI 0.41–1.0). The surgeon misclassified two true 1s as 2s and one true 2 as a 3. The eight discordant scores occurred in seven subjects. There was perfect agreement between the neuroradiologist and the pediatric interventional radiologist ($\kappa = 1.0$).

Discussion

The true incidence of recurrence after surgical resection of AVMs in the pediatric population is not well reported in the literature. A few studies suggest recurrence in children is higher than in adults.^{1,7} AVMs are more prone to hemorrhage in children and thus carry significant morbidity and mortality.^{2,4} Therefore, the ability to predict recurrence after resection has significant implications for monitoring and follow-up imaging. The Spetzler-Martin AVM grading scale has been a long-standing tool used to predict post-operative neurologic complications.¹³ However, this scale was not designed as a tool for predicting recurrence. A study by Du et al in a cohort of mostly adult patients with AVM found that a non-compact AVM nidus was an independent marker for poor clinical outcomes after resection, but recurrence was not specifically examined in relation to diffuseness. In that study they developed a novel computerized method for quantifying the diffuseness of an AVM nidus.³ Our previous study recently demonstrated that a visual pre-surgical angiographic compactness score predicts recurrence. In our cohort, all patients that had a recurrence after complete surgical resection had a diffuse AVM (compactness score = 1) graded by a single pediatric neuroradiologist.¹¹ However, the intra and interrater reliability of the visual compactness score among radiologists and neurosurgeons were not determined.

Accurately identifying diffuse AVMs may prevent the potentially devastating consequences of missing a recurrence by performing adequate follow-up imaging in high-risk patients. However, to be clinically useful, there must be acceptable reliability among practitioners. This study demonstrates that the visual compactness score has moderate-substantial interrater reliability among trained radiologists and neurosurgeons. Both the neuroradiology fellow and neurosurgeon misclassified two true 1s as 2s. This misclassification may lead the clinician to recommend less frequent post-operative surveillance imaging in some patients. Additionally, some subjects were misclassified as having a compactness of 1, which could lead to additional unnecessary imaging. Our raters attended a single training session and reviewed 8 practice cases. Reviewing additional training cases might improve the interrater reliability, particularly for those with less experience in angiographic interpretation. The interrater reliability was excellent (perfect) between the pediatric neuroradiologist and pediatric interventional radiologist, the two raters with the most extensive experience performing and interpreting angiograms. Additionally, the intrarater reliability for the pediatric neuroradiologist was perfect.

Conclusions

The visual pediatric AVM compactness score was demonstrated to predict recurrent AVM after surgical resection in a recent retrospective cohort.¹¹ In the present study, we demonstrated that the visual compactness score had moderate-substantial interrater reliability among radiologists and neurosurgeons with various levels of training and experience. However, to avoid misclassifications, the visual compactness score would be applied better by those with more extensive angiographic experience, demonstrated by the perfect interrater reliability between the pediatric neuroradiologist and interventional neuroradiologist. In the future, the accuracy and usefulness of the pediatric AVM visual compactness score in identifying children with recurrent AVM should be evaluated in a prospective cohort. The pediatric AVM compactness score could lead to a reduction in unnecessary angiograms in children with low-risk lesions for recurrence after complete surgical resection.

Acknowledgments

Funding: Dr. Beslow has been funded by NINDS K12-NS049453, NIH T32-NS007413, and The L. Morton Morley Funds of The Philadelphia Foundation.

References

1. Ali MJ, Bendok BR, Rosenblatt S, Rose JE, Getch CC, Batjer HH. Recurrence of pediatric cerebral arteriovenous malformations after angiographically documented resection. *Pediatr Neurosurg*. 2003; 39:32–38. [PubMed: 12784076]
2. Di Rocco C, Tamburrini G, Rollo M. Cerebral arteriovenous malformations in children. *Acta Neurochir (Wien)*. 2000; 142:145–156. discussion 156–148. [PubMed: 10795888]
3. Du R, Keyoung HM, Dowd CF, Young WL, Lawton MT. The effects of diffuseness and deep perforating artery supply on outcomes after microsurgical resection of brain arteriovenous malformations. *Neurosurgery*. 2007; 60:638–646. discussion 646–638. [PubMed: 17415200]
4. Hladky JP, Lejeune JP, Blond S, Pruvo JP, Dhellemmes P. Cerebral arteriovenous malformations in children: report on 62 cases. *Childs Nerv Syst*. 1994; 10:328–333. [PubMed: 7954502]
5. Hoh BL, Ogilvy CS, Butler WE, Loeffler JS, Putman CM, Chapman PH. Multimodality treatment of nongalenic arteriovenous malformations in pediatric patients. *Neurosurgery*. 2000; 47:346–357. discussion 357–348. [PubMed: 10942007]
6. Kader A, Goodrich JT, Sonstein WJ, Stein BM, Carmel PW, Michelsen WJ. Recurrent cerebral arteriovenous malformations after negative postoperative angiograms. *J Neurosurg*. 1996; 85:14–18. [PubMed: 8683264]
7. Klimo P Jr, Rao G, Brockmeyer D. Pediatric arteriovenous malformations: a 15-year experience with an emphasis on residual and recurrent lesions. *Childs Nerv Syst*. 2007; 23:31–37. [PubMed: 17053936]
8. Kondziolka D, Humphreys RP, Hoffman HJ, Hendrick EB, Drake JM. Arteriovenous malformations of the brain in children: a forty year experience. *Can J Neurol Sci*. 1992; 19:40–45. [PubMed: 1562906]
9. Krayenbuhl HA. Angiographic contribution to the problem of enlargement of cerebral arteriovenous malformations. *Acta Neurochir (Wien)*. 1977; 36:215–242. [PubMed: 848369]
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–74. [PubMed: 843571]
11. Lang SS, Beslow LA, Bailey RL, Vossough A, Ekstrom J, Heuer GG, et al. Follow-up imaging to detect recurrence of surgically treated pediatric arteriovenous malformations. *J Neurosurg Pediatrics*. 2012; 9:497–504.
12. Niazi TN, Klimo P Jr, Anderson RC, Raffel C. Diagnosis and management of arteriovenous malformations in children. *Neurosurg Clin N Am*. 2010; 21:443–456. [PubMed: 20561494]
13. petzler RF, Martin NA. A proposed grading system for arteriovenous malformations. *J Neurosurg*. 1986; 65:476–483. [PubMed: 3760956]

Figure 1. Examples of Compactness Score 1

- A. Lateral cerebral angiogram demonstrating a large diffuse AVM (arrows).
- B. Lateral cerebral angiogram demonstrating a small but diffuse AVM near the vertex (arrows). The AVM does not have a compact nidus. An early draining vein is seen along its superior margin (uppermost arrow).
- C. Image from a cerebral angiogram showing a large diffuse AVM. There is a more compact portion of the AVM with a large early draining vein, but there is also a larger but less compact component of abnormal vascularity, with large amounts of intervening brain tissue as well (arrows), which make this a diffuse AVM with a compactness score of 1.
- D. Lateral cerebral angiogram demonstrating another small but diffuse AVM (arrows). A number of narrow caliber, small abnormal vessels are seen in the region of the AVM, without a compact nidus.

Figure 2.
Cerebral angiogram demonstrating an AVM with compactness score of 2.

Figure 3.
Cerebral angiogram demonstrating a highly compact AVM (compactness score 3).

Table 1

Compactness scoring directions provided to the raters.

Compactness Score		Description
1	Diffuse *	These are the least compact and are rather diffuse AVMs. There is considerable intervening brain tissue among the anomalous vessels. Some may appear to border on cerebral proliferative angiopathy.
2	Intermediate	The nidus is apparent and can be of variable size, but there is some dispersion with a lot of perinidal anomalous vessels.
3	Compact	These are very compact AVMs. They have a prominent nidus with minimal surrounding perinidal anomalous vessels.

* Note: A diffuse AVM (compactness score of 1) does not refer to size. These AVMs can be large or small, and the rating is based on the overall compactness of the nidus and perinidal anomalous vessels.