

Observer Variability of an Angiographic Grading Scale Used for the Assessment of Intracranial Aneurysms Treated with Flow-Diverting Stents

M.D. Joshi, C.J. O'Kelly, T. Krings, D. Fiorella, and T.R. Marotta

ABSTRACT

BACKGROUND AND PURPOSE: Novel angiographic grading scales for the assessment of intracranial aneurysms treated with flow-diverting stents have been recently developed because previous angiographic grading scales cannot be applied to these aneurysms. The purpose of this study was to evaluate the inter- and intraobserver variability of the novel O'Kelly Marotta grading scale, which was developed specifically for the angiographic assessment of aneurysms treated with flow-diverting stents.

MATERIALS AND METHODS: Multiple raters ($n = 31$) from the disciplines of neuroradiology and neurosurgery were presented with pre- and posttreatment angiographic images of 14 aneurysms treated with intraluminal flow diverters. Raters were asked to classify pre- and posttreatment angiograms by using the OKM grading scale. Statistical analyses were subsequently performed with calculation of a generalized multirater κ statistic for assessment of inter- and intraobserver variability and by performing a Wilcoxon signed rank sum test for assessment of group differences.

RESULTS: Variability analysis of the OKM grading scale yielded substantial ($\kappa = 0.74$) and almost perfect ($\kappa = 0.99$) inter- and intraobserver agreement, respectively, with no statistically significant differences between raters with a background of neuroradiology versus neurosurgery or attending physician versus trainee.

CONCLUSIONS: The OKM grading scale for the assessment of intracranial aneurysms treated with flow-diverting stents is a reliable grading scale that can be used equally well by users of varying backgrounds and levels of training. Comparison with interobserver variability of pre-existing angiographic grading scales shows equal or better performance.

ABBREVIATIONS: OKM = O'Kelly Marotta

The development of flow-diverting stents has ushered in an era of new paradigms and possibilities for the endovascular treatment of intracranial aneurysms. These devices are garnering momentum for difficult intracranial aneurysms that are wide-neck, have fusiform configuration, are located in perforator territories, or have complex geometry. Their mechanism of action relies on a reduction of filling of the aneurysmal sac with flow diversion toward the parent vessel, leading to stasis and subsequent thrombosis within the aneurysm. This process usually occurs with time,

and angiographic evidence of aneurysm protection is usually not seen in immediate posttreatment angiograms. Rather, the process of stasis and thrombosis may not be complete for months after initial flow-diverting stent placement. While residual filling is a suboptimal or unacceptable posttreatment angiographic result for an aneurysm treated with endosacular coiling, this can be optimal and acceptable for an aneurysm treated by using flow-diverting devices. Traditional grading scales, such as the 3-point grading scale of Roy and Raymond used for aneurysms treated with coiling or clipping,¹ do not apply to aneurysms treated with flow-diverting stents. For example, after treatment, filling within the aneurysm body would be classified as residual aneurysm and would not be considered a desirable result according to the Roy and Raymond classification, but in the case of flow-diverting stent placement, this would be an expected immediate posttreatment result.

A new grading scale specifically tailored to the angiographic assessment of aneurysms treated with flow-diverting stents was published by O'Kelly et al,² termed the O'Kelly Marotta grading scale. This novel grading scale incorporates 2 dimensions as pa-

Received June 6, 2012; accepted after revision October 5.

From the Department of Medical Imaging (M.D.J., T.K., T.R.M.), University of Toronto, Toronto, Ontario, Canada; Division of Neuroradiology (M.D.J., T.K.), Toronto Western Hospital, Toronto, Ontario, Canada; Division of Neurosurgery (C.J.O.), Department of Surgery, University of Alberta, Edmonton, Alberta, Canada; Department of Neurological Surgery (D.F.), Stony Brook University Medical Centre, Stony Brook, New York; and Division of Diagnostic and Therapeutic Neuroradiology (T.R.M.), St. Michael's Hospital, Toronto, Ontario, Canada.

Please address correspondence to Thomas Marotta, MD, Neurovascular Clinic, Division of Diagnostic and Therapeutic Neuroradiology, St. Michael's Hospital, 30 Bond St, Toronto, ON, M5B 1W8, Canada; e-mail: marottat@smh.ca

<http://dx.doi.org/10.3174/ajnr.A3431>

Figure 1

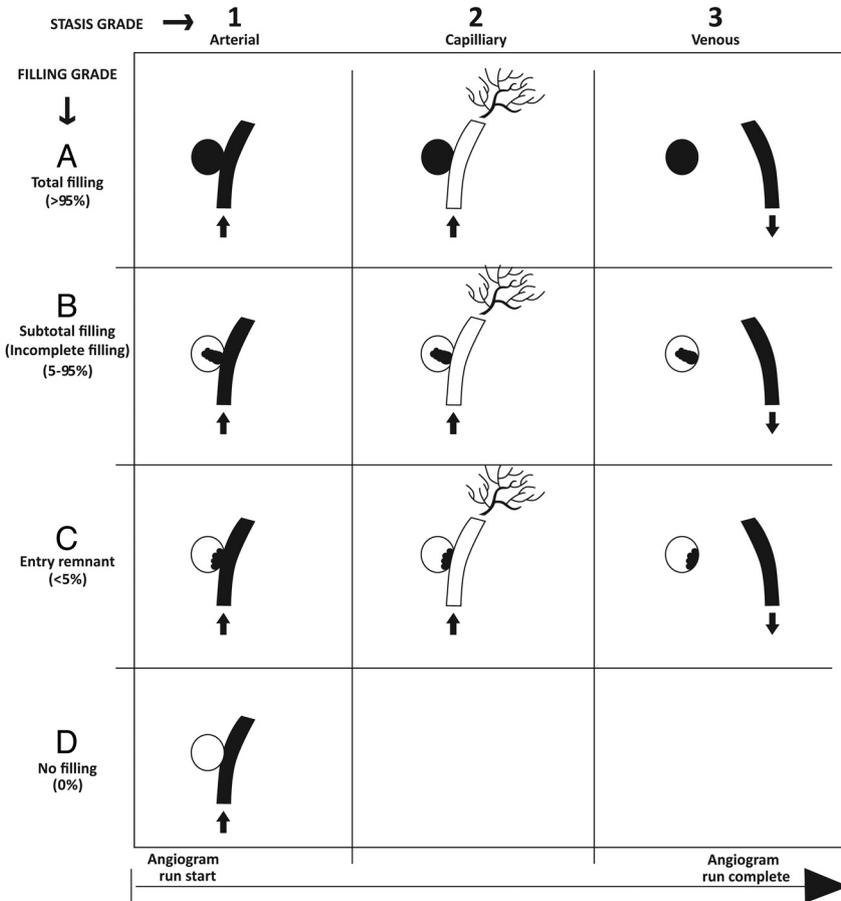


FIG 1. OKM angiographic grading scale for assessment of aneurysms treated with flow-diverting stents. Aneurysms are assigned grades on the basis of the amount of contrast filling of the aneurysm lumen (filling grades, A, B, C, D) and how long contrast persists in the aneurysm lumen with respect to angiographic phase (stasis grades 1, 2, 3). A grade consisting of a letter and a number is assigned to an aneurysm (eg, an aneurysm that fills its lumen completely with contrast persisting within the lumen into the venous phase of the angiogram is assigned grade A3).² Modified from O’Kelly CJ, Krings T, Fiorella D, et al. A novel grading scale for the angiographic assessment of intracranial aneurysms treated using flow diverting stents. *Interv Neuroradiol* 2010;16:133–37 and reproduced with permission from Centauro srl Publishing.

rameters, which reflect the mechanism by which flow-diverting devices accomplish aneurysm protection: reduction in aneurysm sac filling (filling grade), which reflects an anatomic aspect, and promotion of stasis within the aneurysm sac (stasis grade), which reflects a more dynamic or physiologic parameter (Fig 1).

With any novel grading scale, an evaluation of performance in the hands of the evaluating end user is necessary for an assessment of reliability. We present here an analysis of the inter- and intraobserver variability of the OKM grading scale based on evaluation of conventional angiographic images of aneurysms pre- and posttreatment with flow-diverting stents by 31 evaluators with a background in neurointervention from the disciplines of neuroradiology and neurosurgery.

MATERIALS AND METHODS

Fourteen cases of aneurysms, pre- and posttreatment with flow-diverting stents (28 conventional cerebral angiograms), were presented to 31 evaluators from the backgrounds of neuroradiology and neurosurgery. Angiograms were shown in the form of a high-

resolution, scalable, multimedia-enabled, interactive document in Adobe Portable Document Format (Adobe Systems, San Jose, California). Two cases (4 angiograms) were duplicates placed randomly within the set for assessment of intraobserver variability. No patient-identifying, demographic, or flow-diversion device information was provided to any of the evaluators, and none of the evaluators had previous exposure or were involved in the treatment or follow-up of any of the cases presented. Evaluators were presented with a brief written introduction to the OKM grading scale and 2 case examples with answers. There was no personal coaching or verbal introduction given. Evaluators were asked to classify each of the angiograms of the aneurysms according to the OKM grading scale by selecting the appropriate filling grade and stasis grade from a drop-down menu.

Statistical analysis was performed by using statistical software (Statistical Package for the Social Sciences, Version 20; SPSS, Chicago, Illinois). A generalized nonweighted multirater κ statistic was calculated according to the method described by Siegal and Castellan.³ Categorization of κ values was based on Landis and Koch⁴ and reviewed in Fleiss et al.⁵ Differences between groups or raters were analyzed for statistical significance by using a Wilcoxon signed rank sum test.

RESULTS

Thirty-one evaluators, predominantly neurointerventionalists, with backgrounds in neuroradiology ($n = 24$) and neurosurgery ($n = 7$) participated in the study. Evaluators were either fellows or attending physicians. Interobserver variability was assessed by using a generalized multirater κ statistic, which showed substantial agreement across all evaluators with $\kappa = 0.74$, a very low standard error, and a highly favorable P value (Table 1). Intraobserver variability, as determined by random insertion of 4 duplicate angiograms, was also excellent with $\kappa = 0.99$, indicating almost perfect agreement (Table 1).

To determine whether there were any differences in grading based on specialty background or level of training, we performed a Wilcoxon rank sum test. While the number of neuroradiologists participating in the study exceeded the number of neurosurgeons by greater than 3:1, no statistically significant difference existed between the responses provided by either group (Table 1). There was also no statistically significant difference in responses based on level of training (ie, fellow in training versus attending physician), again, with a greater number of attending physicians compared with fellows.

Table 1: Individual and group variability for the OKM grading scale

Variability		Statistical Value
Individual		
Interobserver	Substantial agreement	$\kappa = 0.74$ ($P < .005$) ^a
Intraobserver	Almost perfect agreement	$\kappa = 0.99$ ($P < .005$) ^a
Group		
Neuroradiology vs neurosurgery	No significant difference in interpretation	$P = .329$ ^b
Trainee vs attending	No significant difference in interpretation	$P = .2$ ^b

^a Generalized multirater κ calculated as described in Siegal and Castellani.³

^b Significance determined using a Wilcoxon signed rank sum test.

Table 2: Accuracy of rater response by grading category for the OKM grading scale

Filling Grade	% Correct Response	Stasis Grade	% Correct Response
A	93	1	77
B	84	2	83
C	45	3	90
D	79		

We were also interested in assessing the accuracy of responses with respect to each filling grade and stasis grade category. The percentage of correct responses is tabulated in Table 2. (The correct answer is based on consensus agreement in grading for a given angiogram based on evaluation by publication authors [M.D.J. and T.R.M.]) In general, accuracy was greater than 80% in all but one category, which was filling grade C, which corresponds to the category of entry remnant. According to the OKM grading scale, aneurysms that are pretreatment are graded as filling grade A, total filling, by definition (ie, >95%) even if they are partially thrombosed. This concept appears to have been well-grasped by raters, given the 93% accuracy achieved overall for pre- and posttreatment angiograms and 100% accuracy for grade A when pretreatment subgroup scores were tabulated. Accuracy for filling grades corresponding to 5%–95% aneurysm filling, grade B, and 0% aneurysm filling, grade D, showed 84% and 80% accuracy, respectively.

Evaluators were much less accurate in choosing filling grade C, or entry remnant category, which was reflected by an accuracy rate of 45%, which is substantially less than that in any of the other filling grade categories.

The accuracy of stasis grade scores (ie, grading the degree of contrast stasis within the aneurysm) ranged from 80% to 90%. Much of the variability in determining stasis grade was seen in cases in which filling of an overlapping venous plexus was mistaken for aneurysm filling or vice versa or where subtle stasis and a very small amount of filling persisted into the venous phase but was not detected by the evaluator.

DISCUSSION

The mechanism by which flow-diverting stents divert flow from the aneurysm into the parent vessel leading to progressive stasis, thrombosis, and obliteration of the aneurysm has necessitated the development of new grading scales to assess treatment and predict effect. Widely used and validated scales for grading of coiled aneurysms such as the Roy and Raymond simplified scale⁶ do not apply to those treated with flow-diverting stents for which residual filling on the immediate posttreatment angiograms is acceptable and is the norm. Since the development of the OKM grading

scale, the first published for aneurysms treated by using flow-diverting stents, other grading scales have also been developed. The grading scale by Kamran et al⁷ incorporates 2 main axes: axis I, assessing the degree of aneurysm occlusion, and axis II, assessing parent artery patency. This grading scale has favorable interobserver agreement for both axes (based on assessment by 2 expert raters) for anatomic parameters but does not address gradation of dynamic or physiologic parameters such as contrast stasis in the aneurysm lumen.

This study shows that the OKM grading scale has good inter- and intraobserver variability across many raters ($n = 31$) from varied backgrounds of neuroradiology and neurosurgery with different levels of experience. Our study does have some limitations in that the number of cases that each reader evaluated was limited to the grading of 28 angiograms. However, a high number of raters were recruited to more robustly address interobserver variability. Interobserver κ values may have been higher by selecting raters who were more experienced and coached on use of the grading scale; however, in reality, end-user populations normally comprise those without personal coaching as well as those with variable levels of expertise and the study was designed to capture this aspect. Personal coaching of raters may have also led to a higher degree of accuracy for the entry remnant category, defined as filling of <5%. Part of this is based on the difficulty of visually differentiating such a narrow range of aneurysm filling (ie, >0% to <5%) accurately. From a practical standpoint, the best angiographic view to assess and identify the entry remnant category is one in which the stent and the residual filling segment of the aneurysm are both in maximal profile. This view accomplishes visualization of what would be the aneurysm neck, by conventional nomenclature, and is where subtle residual contrast filling of the aneurysm of <5% would be best visualized.

The entry remnant would appear more like a sessile entity compared with one that is pedunculated or shows layering of contrast appearing like a fluid-fluid level or parfait. The entry remnant also has a propensity to occur in instances in which a small vessel arises from the aneurysm, functioning as a sump, keeping this small portion of the aneurysm persistently filling. However, with increasing and longer cumulative experience with flow-diversion devices, we are seeing that the entry remnant category could also be seen as a separate entity with its own prognostic significance and not only a continuum variable for filling grade. Some of our initial experiences suggest that aneurysms treated with flow-diversion stents that show angiographic appearances consistent with the entry remnant category have a propensity to fail to achieve complete exclusion from circulation. The details of the pathomechanism of this phenomenon are currently being explored (J.P. Cruz, personal communication, June 2011).

Table 3: Comparison of interobserver variability (κ) of the OKM grading scale and other angiographic grading scales

Scale	Parameters	κ	Reference
Aneurysms Flow-diversion stents OKM scale	12 Possibilities Filling grade: A, B, C, D Stasis grades: 1, 2, 3	0.74	This article
Aneurysms Coiled 4-Point Raymond scale	4 Possibilities Complete, dog-ear, residual neck, residual aneurysm	0.50	Cloft et al, 2007 ⁸
Aneurysms Coiled 3-Point scale	3 possibilities Complete, near-complete, incomplete	0.54	Cloft et al, 2007 ⁸
AVMs Spetzler-Martin scale	7 Possibilities Nidus size: 1, 2, 3; eloquence: 0, 1; venous drainage: 0, 1	0.47 0.70	Al-Shahi et al, 2002 ⁹ Iancu-Gontard et al, 2007 ¹⁰

The interobserver variability of the OKM grading scale is comparable with well-established angiographic grading schemes for treated aneurysms (Table 3). Cloft et al⁸ evaluated various well-known aneurysm treatment grading schemes for observer variability with 2, 3, and 4 possibilities for grading based on the angiographic appearance, and reported κ values of 0.63, 0.54, and 0.50, respectively. Better agreement was observed in grading schemes with fewer assignment possibilities. In the OKM scale, there are 12 possibilities for grade assignment from 4 filling grades and 3 stasis grades, with an overall interobserver κ value of 0.74. Moreover, the OKM scale is also comparable with or better than, with respect to interobserver variability, other grading scales that also encompass >1 dimension of observation such as the Spetzler-Martin grading scale for AVM assessment, in which κ values of 0.47⁹ and 0.70¹⁰ have been observed.

CONCLUSIONS

The OKM grading scale for assessment of intracranial aneurysms treated with flow-diverting stents demonstrates good inter- and intraobserver variability across a spectrum of raters of different backgrounds and experience. The scale shows comparable or better interobserver variability as judged by κ values compared with well-established angiographic grading scales. The OKM grading scale has not been formally validated in terms of its ability to predict treatment effect or aid in prognostic stratification; however, this will likely come to fruition as more collective long-term assessments for aneurysms treated with flow-diverting stents become available.

ACKNOWLEDGMENTS

We thank all of the study participants for taking time out of their hectic schedules to participate in our study as well as Dr Hadas Moshonov, JD, M, statistician for discussions and advice regarding statistical analyses.

Disclosures: Cian O’Kelly—UNRELATED: Other: ev3 Endovascular. Comments: I have worked for ev3 as a proctor for Pipeline cases. David Fiorella—RELATED: Consulting

Fee or Honorarium: Covidien/ev3, UNRELATED: *Board Membership:* Vascular Simulations, Computer Vision Systems Laboratories (CVSL), *Consultancy:* Codman/Johnson & Johnson, NFocus, WL Gore, Scientia, *Grants/Grants Pending:* National Institutes of Health (SAMMPRIS),* MicroVention (Low-profile Visualized Intraluminal Support Humanitarian Device Exemption Trial; preclinical device evaluation),* Siemens Healthcare, *Patents (planned, pending, or issued):* ReVasc Inc (acquired by Codman/Johnson and Johnson), *Royalties:* ReVasc Inc (acquired by Codman/Johnson & Johnson), *Stock/Stock Options:* Vascular Simulations, CVSL, *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed:* ev3/Covidien, Codman/Johnson & Johnson. Thomas R. Marotta—RELATED: *Consulting Fee or Honorarium:* for proctoring the Pipeline Embolization Device, UNRELATED: *Patents (planned, pending, or issued):* I have patents for the Endovascular Clip System (no money paid). *Money paid to the institution.

REFERENCES

1. Raymond J, Guilbert F, Weill A, et al. **Long-term angiographic recurrences after selective endovascular treatment of aneurysms with detachable coils.** *Stroke* 2003;34:1398–403
2. O’Kelly CJ, Krings T, Fiorella D, et al. **A novel grading scale for the angiographic assessment of intracranial aneurysms treated using flow diverting stents.** *Interv Neuroradiol* 2010;16:133–37
3. Siegal S, Castellan, NJ. *Nonparametric Statistics for the Behavioral Sciences.* 2nd ed. New York: McGraw-Hill; 1988
4. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74
5. Fleiss JL, Levin BA, Paik MC. *Statistical Methods for Rates and Proportions.* 3rd ed. Hoboken, New Jersey: J Wiley; 2003
6. Roy D, Milot G, Raymond J. **Endovascular treatment of unruptured aneurysms.** *Stroke* 2001;32:1998–2004
7. Kamran M, Yarnold J, Grunwald IQ, et al. **Assessment of angiographic outcomes after flow diversion treatment of intracranial aneurysms: a new grading schema.** *Neuroradiology* 2011;53:501–08
8. Cloft HJ, Kaufmann T, Kallmes DF. **Observer agreement in the assessment of endovascular aneurysm therapy and aneurysm recurrence.** *AJNR Am J Neuroradiol* 2007;28:497–500
9. Al-Shahi R, Pal N, Lewis SC, et al. **Observer agreement in the angiographic assessment of arteriovenous malformations of the brain.** *Stroke* 2002;33:1501–08
10. Iancu-Gontard D, Weill A, Guilbert F, et al. **Inter- and intraobserver variability in the assessment of brain arteriovenous malformation angioarchitecture and endovascular treatment results.** *AJNR Am J Neuroradiol* 2007;28:524–27