AJNR

Interobserver Agreement after Pipeline Embolization Device Implantation

S.H. Suh, H.J. Cloft, G. Lanzino, K. Woodward and D.F. Kallmes

Interobserver Agreement after Pipeline Embolization Device Implantation

S.H. Suh, H.J. Cloft, G. Lanzino, K. Woodward, and D.F. Kallmes

ABSTRACT

BACKGROUND AND PURPOSE: Although flow diversion devices are popular in treatment of aneurysms, angiographic assessment with these devices has rarely been verified by interobserver variability study. The purpose of this study was to determine the interobserver agreement of a 3-point grading system for assessing the angiographic outcome after flow diversion therapy of intracranial, saccular aneurysms and to determine factors affecting such agreement.

MATERIALS AND METHODS: After approval by the institutional review board, 5 independent readers assessed pretreatment and follow-up digital subtraction angiograms from 96 patients treated with the Pipeline embolization device by using a 3-point grading system (complete, near-complete, and incomplete occlusion). "Minor discrepancy" was defined as a difference between any 2 readers of 1 grade, that is, complete vs near-complete or near-complete vs incomplete. "Major discrepancy" was defined as a difference between any 2 readers in which 1 reader noted complete occlusion and the other reader noted incomplete occlusion. We performed statistical analysis for the interobserver agreement by using the intraclass correlation coefficient. Subgroup analyses for discrepancy rate and ICC were performed for previously coiled aneurysms.

RESULTS: The interobserver agreement was excellent (ICC, 0.76; 95% CI, 0.69–0.92). Among 96 cases, there was absolute agreement in 74 (77%), of which 67 had unanimous consensus of "complete" occlusion, 2 "near-complete" occlusion, and 5 "incomplete" occlusion. Discordance between any 2 readers was noted in 22 cases (23%), of which 7 (7.3%) revealed a major discrepancy. Subgroup analysis showed that minor discrepancies were more common among patients previously treated with coils vs those not previously treated with coils (37.5% vs 11.2%; $P < .05$).

CONCLUSIONS: The observer agreement regarding occlusion after PED therapy is excellent. Only a minority of cases demonstrated discrepancy considered as major in this study.

ABBREVIATIONS: CI = confidence interval; ICC = intraclass correlation coefficient; PED = Pipeline embolization device; PUFS = Pipeline for Uncoilable or Failed Aneurysms

Ongoing, angiographic assessment of aneurysms treated by endovascular means represents a standard aspect of patient care, primarily to identify patients considered at risk for future rupture or re-rupture. In an attempt to offer quantitative or semi-quantitative data for angiographic follow-up, various ordinal scales have been proposed. These scales are either descriptive or use estimates of percent volumetric occlusion. Ordinal scales usually use descriptors such as "complete," "near-complete," or "in-complete" occlusion, and some scales use terms such as "neck remnant," "dog ear," "residual neck," and "residual" aneurysm.[1-9] Irrespective of the type of scale used, all are subject to interobserver variability given the relatively subjective nature of the interpretations. Previous studies[10-12] have assessed the degree of interobserver variability for aneurysm occlusion scales, typically demonstrating substantial variability of agreement among readers.

Flow diversion devices, including the Pipeline embolization device (ev3, Irvine, California), represent an important advance in aneurysm therapy, but previous studies[13-16] imprudently adopted a 3-point scale as the follow-up marker, which was the same as for aneurysm embolization by use of coils. Moreover, angiographic outcomes with these devices rarely have been subject to formal interobserver variability studies.

Therefore, the purpose of our current study was to measure interobserver agreement of a 3-point grading system for assessing fol-

low-up angiographic results after PED therapy of intracranial aneurysms.

MATERIALS AND METHODS
Image Acquisition
After approval by our institutional review board, 96 cases were chosen from either a data base of the Pipeline for Uncoilable or Failed Aneurysms registry ($n = 56$) by random selection, or from an institutional data base at the Mayo Clinic in Rochester, Minnesota ($n = 38$), in which all were patients treated at the Mayo Clinic from June 2009 to July 2011. Some of these cases have been published previously in clinical case series,[17,18] but in no previous publication was the issue of interobserver variability specifically assessed. All patients were treated with the PED; 16 patients had a history of aneurysm recurrence after coiling. The locations of the aneurysms were the paraclinoid ICA in 90 cases, the distal ICA in 4, and the basilar artery in 2.

For acquisition of the images, 3D digital subtraction angiography was performed before the procedure, and the working projection was decided to best display the artery and ignore the aneurysm. Subsequently, 2D angiographic images of the primary working projection were taken immediately before and 6 months after PED implantation. In all cases, 2D angiographic images were obtained at the pre-operative and the follow-up visit, but 3D images were available in only 91 cases because of a limited-image data base of the PUFS registry. In 5 cases without 3D images, all of which were included in the PUFS registry, MR angiography was used in 1, CT angiography in 1, and no CT or MRA in 3. Therefore, a total of 546 images from these 96 cases were selected from a data base of the PUFS registry and the Mayo Clinic.

For each case, a digital file was made that consisted of 1 or 2 3D DSA images, both preoperative and follow-up 2D DSA images of the working projection or the conventional angles, and distributed on-line for review.

Readers and Image Interpretation
Each angiogram was evaluated by 5 experienced readers, who worked in 3 different centers in different countries (United States and Korea). There were 4 senior readers ($>$10 years of experience in endovascular aneurysm therapy) and 1 junior reader (S.H.S., with 7 years of experience). Four of 5 reviewers were interventional neuroradiologists, and 1 reviewer was a dual-trained endovascular neurosurgeon. Two readers (D.F.K., G.L.) had participated in the PUFS registry.

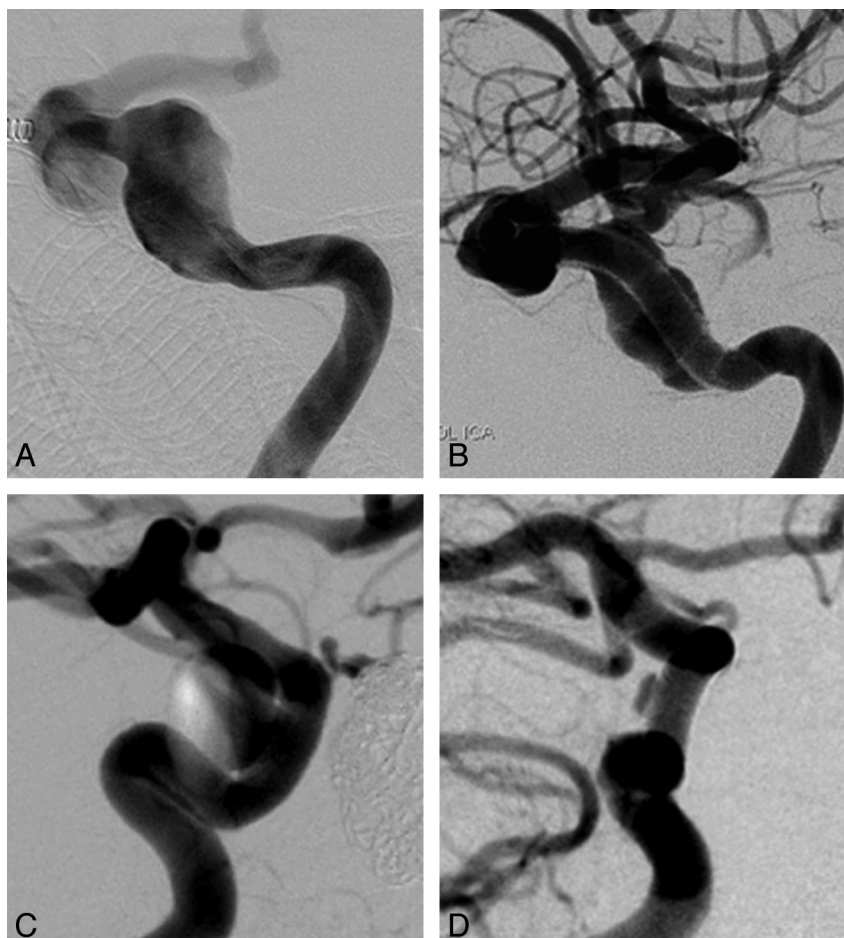Independent of the others' assessment, each reader made his

FIG 1. Examples of the absolute agreement among 5 readers. Fusiform aneurysm of the left ICA before (*A*) and after (*B*) the procedure, which was assessed unanimously as "incomplete" occlusion and paraclinoid aneurysm (*C* and *D*) also assessed unanimously as "near-complete" occlusion.

own assessment by using a 3-point grading scale, which included "complete," "near-complete," and "incomplete" occlusion. For our study, we did not provide any specific training to the readers.

Results were segregated into 3 subgroups, including 1) unanimity, in which all readers noted the same occlusion status; 2) minor discrepancy, in which the greatest discrepancy between any 2 readers was 1 grade, that is, complete vs near-complete or near-complete vs incomplete; and 3) major discrepancy, in which at least 2 readers differed by 2 grades, that is, complete vs incomplete occlusion. For any case with any type of disparity, we defined as "dominant" any reading that was equivalent among 3 or more readers.

Statistical Analysis
To determine interobserver agreement, we calculated the intraclass correlation coefficient with the 2-way random-effects model by using SPSS v.19.0 (SPSS, Chicago, Illinois). The ICC ranges between -1 and $+1$, with scores closer to $+1$ showing better agreement. Interpretation of ICC was as follows: poor, <0.40; fair to good, $0.40-0.75$; and excellent, >0.75.[19,20] Subgroup analysis of ICC was performed for cases with and without prior coiling. The proportions of cases with discrepancy were compared between cases with vs those without prior coiling.

Table 1: Interpretation of the angiographic images for 96 cases with PED implantation

Agreement	No.	Percentage
Absolute	74	77%
Complete	67	
Near-complete	2	
Incomplete	5	
Minor discrepancy		
Dominant[a]	15	16%
Complete	8	
Near-complete	4	
Incomplete	3	
Major discrepancy		
Dominant[a]	7	7%
Complete	2	
Near-complete	2	
Incomplete	1	
Undetermined	2	
Total	96	100%

[a] Dominant indicates case with the same assessment from more than 3 readers.

Table 2: Comparison between "previously coiled" and "not previously coiled" subgroup

	Not Previously Coiled (%)	Previously Coiled (%)	P Value
Absolute agreement	67 (83.8%)	7 (43.8%)	.09
Minor discrepancy	9 (11.2%)	6 (37.5%)	.02
Major discrepancy	4 (5%)	3 (18.7%)	.06
ICC	0.78	0.66	.39
Total	80	16	

RESULTS

The overall ICC was 0.76 (95% CI, 0.69–0.82; $P < .0001$), which represents excellent agreement. Unanimous agreement among all 5 readers was achieved in 74 (77%) of 96 cases, of which 67 cases were read as "complete" occlusion, 2 as "near-complete" occlusion, and 5 as "incomplete" occlusion (Fig 1, Table 1).

Disagreement prevailed between at least 2 readers in 22 cases (23%), of which 15 (16%) were minor disparities and 7 (7.3%) were major disparities. In the minor discrepancy cases, there was dominant "complete" occlusion in 8 cases; dominant "near-complete" occlusion in 4; and dominant "incomplete" occlusion in 3 (Table 1).

The ICC of the "not previously coiled" subgroup was 0.78 (95% CI, 0.72–0.84); for the "coiled" subgroup, it was 0.66 (95% CI, 0.45–0.84). Both groups represented fair to good agreement. For the "not previously coiled" vs the "coiled" subgroup, there was absolute agreement in 67 vs 7 cases (83.8% vs 43.8%; $P > .05$); minor discrepancy in 9 vs 6 cases (11.2% vs 37.5%; $P < .05$); and major discrepancy in 4 vs 3 cases (5% vs 18.7%; $P > .05$) (Table 2).

DISCUSSION

In our study, we demonstrated excellent agreement among 5 reviewers by using a 3-point graded response to evaluate the angiographic outcome after flow diversion therapy of intracranial aneurysms. Clinically relevant discrepancies, which we defined as "major" discrepancy, were uncommon in our study. Although the subgroup of previously coiled aneurysms in our series was small, we noted significantly higher rates of discrepancy when coils were present vs when they were absent. Thus, we suggest that observer variability in the angiographic assessment of flow diversion devices will be high, except in cases of prior coil embolization. When clinical trials related to flow-diverting devices are still ongoing, a reliable and valid scoring system is mandatory in designing these trials, and this finding may affect the design and implementation of such trials.

Cloft et al[10] reported observer variability of several aneurysm assessment scales by 2 experienced observers and in follow-up angiogram assessments by using a 3-point response scale. In that study, the concordant rate and the κ value for interobserver agreement were 72% and 0.67, respectively. Tollard et al[11] described validation of a 3-point grading scale in scoring angiographic results of coiled aneurysms by 10 readers and showed a 21%–60% concordant rate and a generalized κ value of approximately 0.4. Daugherty et al[12] proposed that interobserver variability of a 5-point scale regarding the need and type of treatment of recurrent coiled aneurysms presented 0.27 of the median κ value as an agreement of 5 observers over 27 cases. Compared with these prior studies, our study showed a higher concordant rate and better, consistent agreement among readers. It is clear that the lack of coils to obscure the aneurysm-parent artery interface will yield improved concordance among readers. Furthermore, to a greater extent, our patient population was limited to large ICA aneurysms, which may have improved concordance.

Our study had some limitations. As noted, this study was mostly restricted to aneurysms of the paraclinoid ICA, which may have limited the generalizability of our findings. There was clustering of outcomes, with most of the cases demonstrating complete occlusion, which may have lowered the statistical value.[21,22] The definition of "near-complete" or "incomplete" occlusion should be considered specifically in PED embolization because the healing mechanism of the PED is different than that of the coil, and this terminology might create confusion with clinicians regarding decisions for subsequent treatment vs observation. Moreover, in cases of multiple aneurysms, it is necessary to clarify

FIG 2. Examples of the major discrepancy among 5 readers. Saccular aneurysm of the left cavernous ICA before (*A*) and after (*B*) the procedure. After exclusion of the large aneurysm, another small aneurysm was detected on the anterior genu of the ICA.

which is the target aneurysm intended by the physician performing the procedure. Misinterpretation of the targeted aneurysm could produce some discrepancies in such a study (Fig. 2). Owing to a limited image data base in the PUFS registry, there may have been selection bias of the images in our study. Finally, although we have defined minor and major discrepancies, the clinical ramifications of such discrepancies remain unclear.

CONCLUSIONS

The observer agreement regarding occlusion after PED therapy is excellent. Only a minority of cases demonstrated discrepancy considered as major in our study, and there was a significantly higher discrepancy in a subgroup of the aneurysms treated with coils and the PED despite its small proportion.

REFERENCES

1. Byrne JV, Sohn MJ, Molyneux AJ, et al. **Five-year experience in using coil embolization for ruptured intracranial aneurysms: outcomes and incidence of late rebleeding.** *J Neurosurg* 1999;90:656–63
2. Cognard C, Weill A, Spelle L, et al. **Long-term angiographic follow-up of 169 intracranial berry aneurysms occluded with detachable coils.** *Radiology* 1999;212:348–56
3. Kuether TA, Nesbit GM, Barnwell SL. **Clinical and angiographic outcomes, with treatment data, for patients with cerebral aneurysms treated with Guglielmi detachable coils: a single-center experience.** *Neurosurgery* 1998;43:1016–25
4. Meyers PM, Schumacher HC, Higashida RT, et al. **Reporting standards for endovascular repair of saccular intracranial cerebral aneurysms.** *J Vasc Interv Radiol* 2009;20:S435–50
5. Murayama Y, Nien YL, Duckwiler G, et al. **Guglielmi detachable coil embolization of cerebral aneurysms: 11 years' experience.** *J Neurosurg* 2003;98:959–66
6. Murayama Y, Viñuela F, Duckwiler GR, et al. **Embolization of incidental cerebral aneurysms by using the Guglielmi detachable coil system.** *J Neurosurg* 1999;90:207–14
7. Raymond J, Roy D, Bojanowski M, et al. **Endovascular treatment of acutely ruptured and unruptured aneurysms of the basilar bifurcation.** *J Neurosurg* 1997;86:211–19
8. Roy D, Raymond J, Bouthillier A, et al. **Endovascular treatment of ophthalmic segment aneurysms with Guglielmi detachable coils.** *AJNR Am J Neuroradiol* 1997;18:1207–15
9. Thornton J, Debrun GM, Aletich VA, et al. **Follow-up angiography of intracranial aneurysms treated with endovascular placement of Guglielmi detachable coils.** *Neurosurgery* 2002;50:239–49; discussion 249–50
10. 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
11. Tollard E, Darsaut TE, Bing F, et al. **Outcomes of endovascular treatments of aneurysms: observer variability and implications for interpreting case series and planning randomized trials.** *AJNR Am J Neuroradiol* 2012;33:626–31
12. Daugherty WP, Rad AE, White JB, et al. **Observer agreement regarding the necessity of retreatment of previously coiled recurrent cerebral aneurysms.** *AJNR Am J Neuroradiol* 2011;32:566–69
13. Berge J, Biondi A, Machi P, et al. **Flow-diverter Silk stent for the treatment of intracranial aneurysms: 1-year follow-up in a multicenter study.** *AJNR Am J Neuroradiol* 2012;33:1150–55
14. Byrne JV, Beltechi R, Yarnold JA, et al. **Early experience in the treatment of intra-cranial aneurysms by endovascular flow diversion: a multicentre prospective study.** *PloS One* 2010;5
15. Lubicz B, Collignon L, Raphaeli G, et al. **Flow-diverter stent for the endovascular treatment of intracranial aneurysms: a prospective study in 29 patients with 34 aneurysms.** *Stroke* 2010;41:2247–53
16. Nelson PK, Lylyk P, Szikora I, et al. **The Pipeline embolization device for the intracranial treatment of aneurysms trial.** *AJNR Am J Neuroradiol* 2011;32:34–40
17. Brinjikji W, Cloft HJ, Fiorella D, et al. **Estimating the proportion of intracranial aneurysms likely to be amenable to treatment with the Pipeline embolization device.** *J Neurointerv Surg* 2013;5:45–98
18. Crobeddu E, Lanzino G, Kallmes DF, et al. **Marked decrease in coil and stent utilization following introduction of flow diversion technology.** *J Neurointerv Surg* 2012 Apr 27 [Epub ahead of print]
19. Rosner B, ed. *Fundamentals of Biostatistics.* Belmont, California: Duxbury Press; 2005
20. Sampat MP, Whitman GJ, Stephens TW, et al. **The reliability of measuring physical characteristics of spiculated masses on mammography.** *Br J Radiol* 2006;79 Spec No 2:S134–40
21. Feinstein AR, Cicchetti DV. **High agreement but low kappa: I. The problems of two paradoxes.** *J Clin Epidemiol* 1990;43:543–49
22. Feinstein AR. **Misguided efforts and future challenges for research on "diagnostic tests".** *J Epidemiol Community Health* 2002;56:330–32