

Interobserver Variability in Retreatment Decisions of Recurrent and Residual Aneurysms

J.S. McDonald, R.E. Carter, K.F. Layton, J. Mocco, J.B. Madigan, R.G. Tawk, R.A. Hanel, S.S. Roy, H.J. Cloft, A.M. Klunder, S.H. Suh, and D.F. Kallmes

ABSTRACT

BACKGROUND AND PURPOSE: The degree of variation in retreatment decisions for residual or recurrent aneurysms among endovascular therapists remains poorly defined. We performed a multireader study to determine what reader and patient variables contribute to this variation.

MATERIALS AND METHODS: Seven endovascular therapists (4 neuroradiologists, 3 neurosurgeons) independently reviewed 66 cases of patients treated with endovascular coil embolization for ruptured or unruptured aneurysm. Cases were rated on a 5-point scale recommending for whether to retreat and a recommended retreatment type. Reader agreement was assessed by intraclass correlation coefficient and by identifying cases with a “clinically meaningful difference” (a difference in score that would result in a difference in treatment). Variables that affect reader agreement and retreatment decisions were examined by using the Wilcoxon signed-rank test, Pearson χ^2 test, and linear regression.

RESULTS: Overall interobserver variability for decision to retreat was moderate (ICC = 0.50; 95% CI, 0.40–0.61). Clinically meaningful differences between at least 2 readers were present in 61% of cases and were significantly more common among neuroradiologists than neurosurgeons ($P = .0007$). Neurosurgeons were more likely to recommend “definitely retreat” than neuroradiologists ($P < .0001$). Previously ruptured aneurysms, larger remnant size, and younger patients were associated with more retreat recommendations. Interobserver variability regarding retreatment type was fair overall 0.25 (95% CI, 0.14–0.41) but poor for experienced readers 0.14 (95% CI, 0–0.34).

CONCLUSIONS: There is a large amount of interobserver variability regarding the decision to retreat an aneurysm and the type of retreatment. This variability must be reduced to increase consistency in these subjective outcome measurements.

ABBREVIATIONS: CI = confidence interval; ICC = intraclass correlation coefficient; IQR = interquartile range; mRS = modified Rankin Scale

The propensity for some types of aneurysms treated with coil embolization to undergo recanalization with time has prompted routine adoption of surveillance imaging. Such imaging is performed to identify aneurysms considered prone to rupture or rerupture; these identified aneurysms then are retreated,

either with endovascular or open surgical therapy. Because the rate of rupture or rerupture of any coiled aneurysm, irrespective of angiographic finding, is extremely low, there are no robust data to guide decisions regarding the need for retreatment in a given case. Confident estimates of the rerupture rate for specific types of aneurysm remnant would require enormous clinical trials that likely will never be carried out.

In current practice the decision to treat or not treat a residual or recurred aneurysm remains subjective. Indeed, even within recent, randomized controlled trials, marked variation in retreatment rates was seen among sites.^{1–4} The degree of variation in retreatment recommendation among operators is relatively poorly defined. In addition, patient characteristics that influence retreatment decisions have never previously been studied systematically. To advance our understanding of the degree of interobserver variability regarding retreatment decisions, and to examine physician and patient features that contribute to such variability,

Received May 22, 2012; accepted August 2.

From the Departments of Radiology (J.S.M., H.J.C., A.M.K., S.H.S., D.F.K.), Health Sciences Research (R.E.C.), and Neurosurgery (H.J.C., D.F.K.), Mayo Clinic, Rochester, Minnesota; Department of Radiology (K.F.L.), Baylor University Medical Center, Dallas, Texas; Departments of Neurological Surgery and Radiology and Radiological Sciences (J.M.), Vanderbilt University Medical Center, Nashville, Tennessee; Department of Neuroradiology (J.B.M.), St. Georges Hospital, London, UK; Department of Neurosurgery (R.G.T., R.A.H.), Mayo Clinic, Jacksonville, Florida; and Department of Neurosurgery (S.S.R.), University of Illinois at Chicago, Chicago, Illinois.

Please address correspondence to Jennifer S. McDonald, PhD, Department of Radiology, Mayo Clinic, 200 1st St SW, Rochester, MN 55905; e-mail: mcdonald.jennifer@mayo.edu

<http://dx.doi.org/10.3174/ajnr.A3326>

we carried out a multireader study of recurred and residual aneurysms.

METHODS

Study Population

Institutional review board approval was obtained for this retrospective study. Records for patients treated with endovascular coil embolization for either ruptured or unruptured aneurysms at our institution between 2006 and 2010 were evaluated. A subset of these data were previously published without analysis of the impact of patient data on recommendations.⁵ Inclusion criteria were the presence of any aneurysm remnant and available angiographic follow-up at 6 months or greater, along with relevant clinical data. Clinical data that were retrieved included patient age, sex, status of aneurysm at time of initial treatment (ruptured versus unruptured), size of aneurysm remnant (maximum dimension, dome:neck ratio, height:neck ratio), duration of follow-up, smoking history, family history of aneurysm, and clinical status at follow-up by using the modified Rankin scale.⁶

Review of Cases

Seven endovascular therapists, 4 neuroradiologists and 3 neurosurgeons representing 6 medical centers, independently reviewed angiographic and clinical data for each case. Experience level for each clinician was subcategorized into <5 years, 5–10 years, and 10+ years. Readers were asked to offer recommendations regarding the need for retreatment by using a 5-point graded scale along with what the recommended treatment would be (Table 1).

Statistical Analysis

The overarching goal of the study was to quantify agreement among readers and examine variables associated with differences in recommendations with each case. Statistical analyses were performed by using JMP (v.9, SAS Institute, Cary, NC) and R version 2.15 (<http://www.r-project.org/>).⁷ Continuous variables were presented as median and interquartile range and categorical variables were presented as percentage. Analyses described below were performed for all readers, and were also stratified by reader specialty (neuroradiology or neurosurgery) and experience level (<5 years, 5–10 years, and 10+ years). Reader and case variables were compared between groups by using the Wilcoxon signed-rank test, Pearson χ^2 test, and linear regression. Statistical significance was defined as $P < .05$.

The intraclass correlation coefficient was used to assess interreader agreement in decision to retreat and type of retreatment, where an ICC result ≥ 0.8 indicated very good agreement, 0.60–0.79 good agreement, 0.40–0.59 moderate agreement, 0.20–0.39 fair agreement, and < 0.20 poor agreement. Reader agreement was also assessed by identifying cases with a “clinically meaningful difference” in ratings, defined as a case with at least one grade 1 rating and at least one grade 4 or grade 5 rating, or a case with at least one grade 5 rating and at least one grade 1 or grade 2 rating. That is, a clinically meaningful difference is one in which at least one reviewer recommended definitely or probably do not retreat while at least one other reviewer recommended definitely retreat, and visa versa.

Recommendations on the decision to retreat were analyzed by

Table 1: Treatment options

| Options |
|------------------------------|
| Retreat rating scale |
| 1) Definitely do not retreat |
| 2) Probably do not retreat |
| 3) Unsure |
| 4) Probably retreat |
| 5) Definitely retreat |
| Retreatment recommendation |
| Surgical clipping |
| Simple coiling |
| Flow diversion |
| Balloon-assisted coiling |
| Stent-assisted coiling |

Table 2: Case demographics

| Demographics | n = 66 |
|-----------------------------------|----------------|
| Sex (female) | 46 (70%) |
| Age (yr) | 59 (52–65) |
| Aneurysm type | |
| Ruptured | 45 (68%) |
| Unruptured | 21 (32%) |
| Aneurysm remnant | |
| Maximum dimension (mm) (range) | 3.3 (2.7–6.0) |
| Dome/neck ratio (range) | 0.9 (0.8–1.0) |
| Height/neck ratio (range) | 0.9 (0.6–1.2) |
| Months since embolization (range) | 6.6 (5.8–10.2) |
| mRS score | |
| 0 | 21 (32%) |
| 1 | 28 (42%) |
| 2 | 9 (14%) |
| 3 | 8 (12%) |
| Smoking history | |
| Current | 22 (33%) |
| Previous | 25 (38%) |
| Never | 19 (29%) |
| Family history | |
| None | 57 (86%) |
| First-order relative | 7 (11%) |
| Second-order relative | 2 (3%) |

calculating the mean rating score for all cases and for all readers and within reader subgroups. Recommendations on the retreatment type were analyzed for each case and categorized by the number of times a reader selected 1 of the 5 treatment options.

RESULTS

Case Demographics

A total of 66 patients were included in this study (Table 2). The median age was 59 years (IQR = 52–65) and 70% (46/66) of patients were female. Most of the cases (68%, 45/66) were initially ruptured aneurysms. A third of the sample were current smokers (22/66) and most (86%, 57/66) had no family history of aneurysms.

The Effect of Reader and Patient Variables on Interobserver Agreement

Correlations among readers for decision to retreat and retreatment type are summarized in Table 3. Interobserver agreement regarding decision to retreat was moderate, with an ICC of 0.50 (95% CI, 0.40–0.61) for all 7 readers. Agreement among neuro-radiologists tended to be lower than that among neurosurgeons,

Table 3: Interclass correlation coefficients

| Decision to Retreat | ICC (95% CI) |
|------------------------------------|------------------|
| All readers | 0.50 (0.40–0.61) |
| Neuroradiologists | 0.48 (0.35–0.60) |
| Neurosurgeons | 0.59 (0.46–0.71) |
| Readers with <5 years' experience | 0.45 (0.30–0.60) |
| Readers with ≥10 years' experience | 0.52 (0.37–0.65) |
| Treatment type | |
| All readers | 0.25 (0.14–0.41) |
| Neuroradiologists | 0.20 (0.07–0.37) |
| Neurosurgeons | 0.25 (0.09–0.46) |
| Readers with <5 years' experience | 0.46 (0.28–0.63) |
| Readers with ≥10 years' experience | 0.14 (0–0.34) |

but this difference did not reach statistical significance due to overlapping confidence intervals. Interobserver agreement regarding treatment type was fair (ICC = 0.25 [95% CI, 0.14–0.41]). Again, neuroradiologists had lower agreement than neurosurgeons, though this difference was not significant. Agreement in less experienced readers was substantially higher (ICC = 0.46) than the poor agreement observed with more experienced readers (ICC = 0.14).

A total of 40 (61%) cases had a clinically meaningful difference in rating when examining all readers (Table 4). When examining only readers within their specialty, neuroradiologists had a significantly higher number of cases with a clinically meaningful difference in rating (47% [31/66]) compared with neurosurgeons (18%, [12/66]; $P = .0007$). Experience level did not affect the number of cases with clinically meaningful differences, with less experienced readers showing a similar number of discrepant cases as compared with more experienced readers (26% [17/66] versus 30% [20/66]; $P = .70$). No patient variables were significantly different between cases that had a clinically meaningful difference and those that did not.

The Effect of Reader Variables on the Decision to Retreat

The distribution of 462 individual recommendations from 7 readers on decision to retreat is shown in Table 4. There was significant heterogeneity in recommendations by reader. When the individual ratings were utilized, neurosurgeons were more likely to recommend retreatment (grade 5, 48% [95/198] compared with neuroradiologists, 34% [89/264]; $P < .0001$). In contrast, neuroradiologists were more likely to recommend no further treatment (grade 1, 23% [61/264] compared with neurosurgeons, 11% [21/198]; $P < .0001$). Readers with less experience (<5 years) had similar rates of recommendations to definitely retreat (43% [85/198]) and definitely not retreat (14% [27/198]) compared with readers with more experience (10+ years, 40% [80/198]; $P = .92$ and 18%, [35/198]; $P = .33$, respectively). Mean treatment scores show similar trends in specialty and experience subgroups.

The Effect of Patient Variables on Decision to Retreat

Cases that were recommended for retreatment by all readers (grade 4 or 5) were more likely to be patients with ruptured aneurysms ($P = .0251$) compared with cases ranked 3 or lower. This was observed in both the neuroradiologist ($P = .0131$) and neurosurgeon ($P = .0149$) subgroups. Cases recommended for retreatment also had significantly larger aneurysm remnants

(median maximum dimension of 4.7 mm versus 3.1 mm, $P = .0160$), larger remnant dome:neck ratios (median ratio 1.1 versus 0.9, $P = .0095$), and larger remnant height:neck ratios (median ratio 1.2 versus 0.8, $P = .0006$) compared with cases where retreatment was not recommended. These trends in remnant dome:neck ratio and height:neck ratio were observed in both neuroradiologist ($P = .0119$ and $P = .0004$, respectively) and neurosurgeon ($P = .0165$ and $P < .0001$, respectively) subgroups. Conversely, cases that were recommended for no retreatment by all readers (grade 1 or 2) were more likely to be older patients ($P = .0290$) and patients with unruptured aneurysms ($P = .0324$), and have significantly lower remnant height:neck ratios (0.5 versus 0.9, $P = .0113$) compared with cases ranked 3 and higher. These trends were observed in both neuroradiologist ($P = .0138$, $P = .0418$, and $P < .0001$, respectively) and neurosurgeon ($P = .0290$, $P = .0324$, and $P = .0113$, respectively) subgroups. Other patient variables had no significant effect on the decision to retreat.

The Effect of Reader Variables on Retreatment Type

Reader recommendations for retreatment type are summarized in Table 5. Overall, readers most frequently selected stent-assisted coiling (33%) and simple coiling (21%). Neuroradiologists selected simple coiling and balloon-assisted coiling more frequently than neurosurgeons (25% versus 16%; $P = .0244$ and 19% versus 9%; $P = .0025$, respectively). Conversely, neurosurgeons selected stent-assisted coiling and flow diversion more frequently than neuroradiologists (41% versus 28%; $P = .0185$ and 21% versus 13%; $P = .0486$). Readers with less experience selected stent-assisted coiling more frequently than readers with more experience (50% versus 17%; $P < .0001$). Conversely, readers with more experience selected surgical clipping and flow diversion more frequently than less experienced readers (21% versus 6%; $P < .0001$ and 26% versus 11%; $P = .0004$, respectively).

DISCUSSION

In the current study we have demonstrated that, between at least 2 of 7 endovascular therapists, clinically meaningful differences in retreatment of recurrent or residual aneurysms would have been made in more than half of all cases. That is, at least one therapist would have recommended strongly against retreatment while at least one other therapist would have recommended for retreatment in 61% of the 66 cases in our cohort. Further, the mode of retreatment varied widely among readers. Notably, agreement regarding type of retreatment was poor for experienced practitioners. These findings suggest strongly that, far from representing an objective outcome, the decision to retreat a given aneurysm remains subjective and suffers from profound variability among practitioners.

We evaluated numerous patient and practitioner variables to uncover factors that would influence retreatment recommendations. Recommendations for retreatment were made more frequently for younger patients and those whose treated aneurysms were initially ruptured. Neurosurgeons were more likely to recommend retreatment, but type of retreatment was highly variable across all practitioner variables. Experience level of the provider had no significant effect on retreatment recommendations.

Neuroradiologists and neurosurgeons recommended use of a

Table 4: Retreatment ratings by reader

| Individual Readers | Decision to Retreat | | | | | Mean Rating (SD) | Clinically Meaningful Difference Cases | P Value (CMD Cases) | P Value (Grade 5) | P Value (Grade 1) | Total |
|--|------------------------------|----------|----------|----------|-----------------------|------------------|--|---------------------|--------------------|--------------------|-------|
| | 1) Definitely Do Not Retreat | 2 | 3 | 4 | 5) Definitely Retreat | | | | | | |
| Reader 1: neuroradiologist, <5 years | 11 | 15 | 12 | 13 | 15 | 3.1 (1.4) | | | | | 66 |
| Reader 2: neuroradiologist, 5–10 years | 20 | 17 | 1 | 12 | 16 | 2.8 (1.6) | | | | | 66 |
| Reader 3: neuroradiologist, ≥10 years | 21 | 9 | 2 | 13 | 21 | 3.1 (1.7) | | | | | 66 |
| Reader 4: neuroradiologist, ≥10 years | 9 | 9 | 0 | 11 | 37 | 3.9 (1.5) | | | | | 66 |
| Reader 5: neurosurgeon, <5 years | 6 | 4 | 6 | 10 | 40 | 4.1 (1.3) | | | | | 66 |
| Reader 6: neurosurgeon, <5 years | 10 | 9 | 2 | 15 | 30 | 3.7 (1.5) | | | | | 66 |
| Reader 7: neurosurgeon, ≥10 years | 5 | 14 | 1 | 21 | 25 | 3.7 (1.4) | | | | | 66 |
| Only neuroradiologists | 61 (23%) | 50 (19%) | 15 (6%) | 49 (19%) | 89 (34%) | 3.2 (1.6) | 31 (47%) | .0007 | <.0001 | <.0001 | 264 |
| Only neurosurgeons | 21 (11%) | 27 (14%) | 9 (5%) | 46 (23%) | 95 (48%) | 3.8 (1.4) | 12 (18%) | | | | 198 |
| Only <5 years' experience | 27 (14%) | 28 (14%) | 20 (10%) | 38 (19%) | 85 (43%) | 3.6 (1.5) | 17 (26%) | .71 ^a | .0198 ^b | .0090 ^b | 198 |
| Only 5–10 years' experience | 20 (30%) | 17 (26%) | 1 (2%) | 12 (18%) | 16 (24%) | 2.8 (1.6) | – | | .92 ^a | .33 ^a | 66 |
| Only ≥10 years' experience | 35 (18%) | 32 (16%) | 3 (2%) | 45 (23%) | 80 (40%) | 3.5 (1.6) | 20 (30%) | | | | 198 |
| All readers total | 82 (18%) | 77 (17%) | 24 (5%) | 95 (21%) | 184 (40%) | 3.5 (1.6) | 40 (61%) | | | | 462 |

Note:—CMD indicates a difference in score that would result in a difference in treatment.

^a Comparing level groups with <5 years' and ≥10 years' experience only.

^b Comparing level groups with <5 years', 5–10 years', and ≥10 years' experience.

Table 5: Retreatment type by reader

| Specialty and Experience | Retreatment Type | | | | | Total |
|--|-------------------|----------------|----------------|--------------------------|------------------------|-------|
| | Surgical Clipping | Simple Coiling | Flow Diversion | Balloon-Assisted Coiling | Stent-Assisted Coiling | |
| Reader 1: neuroradiologist, <5 years | 1 | 16 | 13 | 5 | 31 | 66 |
| Reader 2: neuroradiologist, 5–10 years | 9 | 6 | 5 | 26 | 20 | 66 |
| Reader 3: neuroradiologist, ≥10 years | 6 | 22 | 13 | 4 | 1 | 46 |
| Reader 4: neuroradiologist, ≥10 years | 20 | 17 | 1 | 12 | 16 | 66 |
| Reader 5: neurosurgeon, <5 years | 3 | 23 | 0 | 5 | 35 | 66 |
| Reader 6: neurosurgeon, <5 years | 7 | 2 | 6 | 7 | 23 | 45 |
| Reader 7: neurosurgeon, ≥10 years | 12 | 4 | 32 | 4 | 14 | 66 |
| Only neuroradiologists | 36 (15%) | 61 (25%) | 32 (13%) | 47 (19%) | 68 (28%) | 244 |
| Only neurosurgeons | 22 (12%) | 29 (16%) | 38 (21%) | 16 (9%) | 72 (41%) | 177 |
| Only <5 years' experience | 11 (6%) | 41 (23%) | 19 (11%) | 17 (10%) | 89 (50%) | 177 |
| Only 5–10 years' experience | 9 (14%) | 6 (9%) | 5 (8%) | 26 (39%) | 20 (30%) | 66 |
| Only ≥10 years' experience | 38 (21%) | 43 (24%) | 46 (26%) | 20 (11%) | 31 (17%) | 178 |
| All readers total | 58 (14%) | 90 (21%) | 70 (17%) | 63 (15%) | 140 (33%) | 421 |

flow diverter 70 times out of 421 total recommendations (17%). This finding was surprising considering that such usage would be outside of the licensed indication for this device. However, because this study was international in scope, FDA approval was not considered as an essential part of the study design. Further, off-label use in the United States of approved neurointerventional devices is extremely common, so these recommendations in our study likely reflect practice in some medical centers.

Previous studies have also reported substantial interobserver variability between readers in regard to the evaluation and retreatment of cerebral aneurysms.^{5,8,9} Our study expands upon these findings with a larger and more diverse group of readers and a larger group of aneurysm cases. Further, this study is the first to incorporate and analyze the effect of patient variables, including sex, aneurysm type, smoking status, and age, on reader decisions

and variability. Daugherty et al⁵ reported that reader specialty had no effect on the decision to retreat or on the type of treatment, whereas our study found that neurosurgeons were significantly more likely to retreat and recommended different types of treatment compared with neuroradiologists. This difference in results between studies may be attributable to different readers or a more diverse group of readers used in the studies or the effect of introducing patient variables into the reader decision-making process.

This study has several limitations. First, we did not provide aneurysm-specific information, including location or size, to the readers and therefore did not examine whether aneurysm information affects retreatment decisions or interobserver variability. Aneurysm variables were outside of the focus on this study and would have created an excessive number of variables for this relatively small case size. However, our study did incorporate rem-

nant size in our analysis. Second, the number of readers used in this study is relatively small, especially for analyzing differences in recommendations between experience levels. Finally, the readers in this study made their decisions about whether to retreat and retreatment type based upon one posttreatment angiographic study and limited clinical demographics. In a real-world setting, clinicians typically incorporate several follow-up angiographic studies and additional patient characteristics, as well as patient anxiety and preferences, into the decision-making process. Additional studies using larger groups of readers, larger numbers of cases, and incorporating more patient and follow-up imaging variables should therefore be performed.

CONCLUSIONS

Our study shows that there is a large amount of interobserver variability regarding both the decision to retreat an aneurysm and the type of treatment to be used. The extent of and contributors to this variability must be examined to reduce reader variability and increase consistency in these subjective outcome measurements.

Disclosures: Jennifer S. McDonald—UNRELATED: Grants/Grants Pending: GE Healthcare.* J. Mocco—UNRELATED: Consultancy: Lazarus Effect, NFocus, Edge Therapeutics, Stock/Stock Options: Blockade Therapeutics. Harry Cloft—UNRELATED: Grants/Grants Pending: Cordis Endovascular.* David F. Kallmes—UNRELATED: Grants/Grants Pending: ev3,* Benvenue,* MicroVention,* Sequent,* Micrus;* Royalties: University of Virginia Patent Foundation, Payment for Development of Educational Presentations: ev3,* CareFusion,* Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed: MicroVention.* *Money paid to the institution.

REFERENCES

1. Johnson SC, Gholkar A. **The MAPS (Matrix and Platinum Science) Trial: primary results.** In: *Proceedings of the Eighth Annual Meeting of the Society of NeuroInterventional Surgery*, Colorado Springs, Colorado. July 25–28, 2011
2. Molyneux AJ, Fox A, Sneade M, et al. **Cerecyte Coil Trial: angiographic and clinical outcomes of endovascular coiling in patients with ruptured and unruptured intracranial aneurysms treated with Cerecyte coils compared with bare platinum coils—final results of a prospective randomized trial.** In: *Proceedings of the 50th Annual Meeting of the American Society of Neuroradiology and the ASNR Symposium*, New York. April 21–26, 2012
3. Molyneux AJ, Clarke A, Sneade M, et al. **Cerecyte coil trial: angiographic outcomes of a prospective randomized trial comparing endovascular coiling of cerebral aneurysms with either Cerecyte or bare platinum coils.** *Stroke* 2012;43:2544–50
4. White PM, Lewis SC, Gholkar A, et al. **Hydrogel-coated coils versus bare platinum coils for the endovascular treatment of intracranial aneurysms (HELPS): a randomised controlled trial.** *Lancet* 2011; 377:1655–62
5. 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
6. Farrell B, Godwin J, Richards S, et al. **The United Kingdom transient ischaemic attack (UK-TIA) aspirin trial: final results.** *J Neurol Neurosurg Psychiatry* 1991;54:1044–54
7. R Development Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2012
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. 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