AJNR

Effect of Core Laboratory and Multiple-Reader Interpretation of Angiographic Images on Follow-Up Outcomes of Coiled Cerebral Aneurysms: A Systematic Review and Meta-Analysis

I. Rezek, G. Mousan, Z. Wang, M.H. Murad and D.F. Kallmes

Effect of Core Laboratory and Multiple-Reader Interpretation of Angiographic Images on Follow-Up Outcomes of Coiled Cerebral Aneurysms: A Systematic Review and Meta-Analysis

I. Rezek, G. Mousan, Z. Wang, M.H. Murad, and D.F. Kallmes

ABSTRACT

BACKGROUND AND PURPOSE: Reported rates of recanalization following coil embolization vary widely across studies. Some confounders are known to affect outcomes but others remain questionable. In the current study, we assess differences in reported angiographic outcomes for cerebral aneurysms treated with coil embolization as a function of single vs multiple readers and site investigator vs core laboratory settings.

MATERIALS AND METHODS: Our systematic review covered 1999–2011 by using Ovid MEDLINE and EMBASE. Search terms were subarachnoid hemorrhage, intracranial aneurysms, endovascular treatment, and coiling. Inclusion criteria were >50 aneurysms and available imaging follow-up. Study characteristics of interest were readers at the treating site(s) or at an independent core imaging facility, single vs multiple readers, number of aneurysms treated, mean aneurysm size, mean follow-up time, coil type, initial rupture status, and angiographic follow-up. We defined "unfavorable angiographic outcome" as either "recanalization," <90% occlusion, or "incomplete occlusion."

RESULTS: There were 104 (2.6%) of 4022 studies that fulfilled our inclusion criteria, comprising a total of 22,134 treated aneurysms, of which 15,969 (72.1%) had reported angiographic follow-up. The overall unfavorable outcome rate was 17.8% (2955/15,969 aneurysms). Eight (7.7%) of 104 studies reported core laboratory readings in which the pooled rate of unfavorable outcomes was 0.23 (95% CI, 0.19–0.28) compared with 0.16 (95% CI, 0.14–0.18) in readings from the treating sites ($P < .001$). The multivariate meta-regression suggested that core laboratory interpretation was significant for unfavorable outcomes (OR, 5.60; 95% CI, 2.01–15.60; $P = .001$), after adjustment for initial rupture status, aneurysm size, follow-up duration, and coil type. No significant association was found with use of multiple readers.

CONCLUSIONS: Core laboratory studies tend to report higher rates of unfavorable outcomes compared with self-reported studies.

ABBREVIATIONS: CI = confidence interval; ISAT = International Subarachnoid Aneurysm Trial; MeSH = Medical Subject Headings; OR, odds ratio

Endovascular coiling has become the first-line treatment of cerebral aneurysms after the ISAT. Current recommendations[1] suggest ongoing angiographic or MR angiographic follow-up to ensure radiographic stability and direct subsequent management, including the possibility of re-treating the patient. Reported rates of "unfavorable" angiographic outcome vary markedly among studies. Some of this reported variability may reflect real differences in outcome because rupture status, time since treatment, aneurysm size, and neck width have been shown to substantially affect rates of recanalization.[2-6] Other important factors deter-

mining rates of "unfavorable" outcome relate not to the actual angiographic result but rather from differences in reporting nomenclature, with numerous and different scales and terminology used across studies. Furthermore, angiographic interpretation even of the same scale, as with any other diagnostic tests, may be subject to intraobserver and interobserver variability.

In addition to the numerous factors listed above, other study design features might influence reported outcomes. For example, it has been shown in the cardiology literature[7] that site readings (ie, angiographic readings done by the operators themselves) may be significantly different than readings performed in an independent core laboratory facility. Furthermore, use of multiple readers has been hypothesized to affect reported outcomes.[8]

The current literature focusing on angiographic recanalization after coil embolization includes numerous and different strategies for image interpretation, including those done at the treating facility as well as those in a core laboratory. In addition, some reported studies rely on single observers,[9-11] whereas others report

multiple-reader outcomes.[12-14] However, the impact of setting (site readings vs core facility interpretation) as well as single-reader vs multiple-reader studies remains poorly studied. In our current study, we assessed differences in reported angiographic outcomes for cerebral aneurysms treated with coil embolization as a function of single vs multiple readers and site investigator vs core laboratory settings.

MATERIALS AND METHODS

Search of the Literature

We used identical search criteria as in a recent systematic review[15] to cover the period from January 1999 to December 2011. The search covered Ovid MEDLINE and EMBASE databases and was performed by a librarian at our institution. The following key words as MeSH terms and text words were used in relevant combinations: subarachnoid hemorrhage, intracranial aneurysm, endovascular treatment, and coiling in both *AND* and *OR* combinations. We surveyed abstracts from major scientific meetings in 2011 and 2012 to identify any additional study that used independent core laboratory readings.

Inclusion criteria were greater than 50 aneurysms reported and imaging follow-up with DSA or MR angiography. DSA outcomes were preferably included when available. If not available, MR angiography was then considered. Exclusion criteria were traumatic, dissecting, mycotic, or flow-related aneurysms; stent-treated aneurysms without coiling; and noncoiled embolic agents used to perform either aneurysm or parent vessel coil occlusions. Studies with subgroups that used different imaging modalities were considered only for their DSA-followed group. When the same patient population was the subject of several publications, only the study with the largest cohort was included.

The primary outcome in our study was defined as an "unfavorable" angiographic outcome. We considered the longest duration of reported angiographic results for each study when more than 1 phase of follow-up was reported. Unfavorable angiographic outcome was defined as any degree of recanalization noted on the follow-up images with comparison to the immediate posttreatment results. Terms such as aneurysm recurrence, new filling of aneurysm lumen, and regrowth were considered synonymous with recanalization. If recanalization was not reported in a study, then the unfavorable angiographic outcome was defined as either < 90% degree of occlusion or class 3 on the Raymond scale, which is defined as any opacification of the aneurysm sac.[16]

Two reviewers [I.R., G.M.] independently evaluated the articles in the librarian's primary list and selected studies that fulfilled the design criteria. From each study, we extracted the study center and the reporting settings (whether self-reported or through an independent core laboratory and whether assessed by single or multiple readers when provided), the number of coiled aneurysms, number of aneurysms that had available follow-up, mean aneurysm size, follow-up results, rupture status, mean duration of follow-up, and coil types. Studies with subsets, in which different types of coils were used, were considered separate cohorts for statistical analysis.

"Core laboratory" was defined as an explicit statement in the methodology that an independent image interpretation facility and staff interpreted images. Within the Core Laboratory cohort of studies, where possible we identified whether the studies were assessed by a single reader or by multiple readers. Site readings (noncore laboratory) were defined as 1) an explicit statement that images were interpreted by the treating physician, 2) images were interpreted by more than 1 independent reader at the same institution, or 3) not mentioned.

Assessment by a single reader or by multiple readers was determined from the studies' methods. Studies were included in the multiple-reader group if an explicit statement in the methods was mentioned. Studies were included in the single-reader group if either an explicit statement was mentioned or not mentioned.

Statistical Analysis

To compare baseline characteristics, we used the Student t test to assess the difference on mean aneurysm size, percentage of initially ruptured aneurysms, and mean follow-up duration. The χ^2 test was conducted to assess the number of studies that used bare platinum coils. We calculated the rates of unfavorable angiographic outcomes from each study. The CIs of the rates were estimated by the Jeffreys method.[17] We then pooled the overall rate of unfavorable angiographic outcomes by using the DerSimonian and Laird random-effects methods after log-transforming the rates.[18,19] Analysis of variance was used to test the difference between the natural logarithm of the pooled rate of unfavorable angiographic outcomes on 1) whether a core laboratory ran the assessment (vs site assessment) and 2) whether images were assessed by a single reader (vs more than 1 reader). We constructed multivariate nested random-effects meta-regression models to further explore the heterogeneity of core laboratory and multiple readers after adjusting for baseline rupture, aneurysm size, follow-up time, and coiling device.[19]

We used the I^2 statistic to measure the overall heterogeneity across the studies, where $I^2 > 50\%$ suggests high heterogeneity.[20] Publication bias was assessed by the Egger regression asymmetry test.[21] We conducted all statistical analyses by using STATA version 12 (StataCorp, College Station, Texas).

RESULTS

The literature review encompassed 4019 articles published between January 1999 and December 2011, as well as 3 core laboratory studies identified from 2011–2012 annual meetings[22-24] (Fig 1). Of 4022 studies, 3918 (97.4%) were excluded for the following reasons: 1) <50 treated aneurysms (2887 articles [71.8%]); 2) no available angiographic follow-up (746 articles [18.6%]); 3) no patient series such as editorial letters and nonhuman experimental models for cerebral aneurysms in rabbits, dogs, or pigs (256 articles [6.4%]); 4) duplicate publishing about the same patient population (21 articles [0.5%]); and 5) use of treatment options other than coiling such as stents alone or parent artery embolization (8 articles [0.2%]). As such, 104 studies (3%) met all inclusion criteria and were included in our analysis (see On-line Appendix for a full list of included studies).

The total number of treated aneurysms among these 104 included studies was 22,134, of which 15,969 (72.1%) had available angiographic follow-up outcomes. The mean aneurysm size was reported in 72 (69.2%) of 104 studies. The overall mean aneurysm size was 7.8 mm. The mean duration of follow-up was reported in
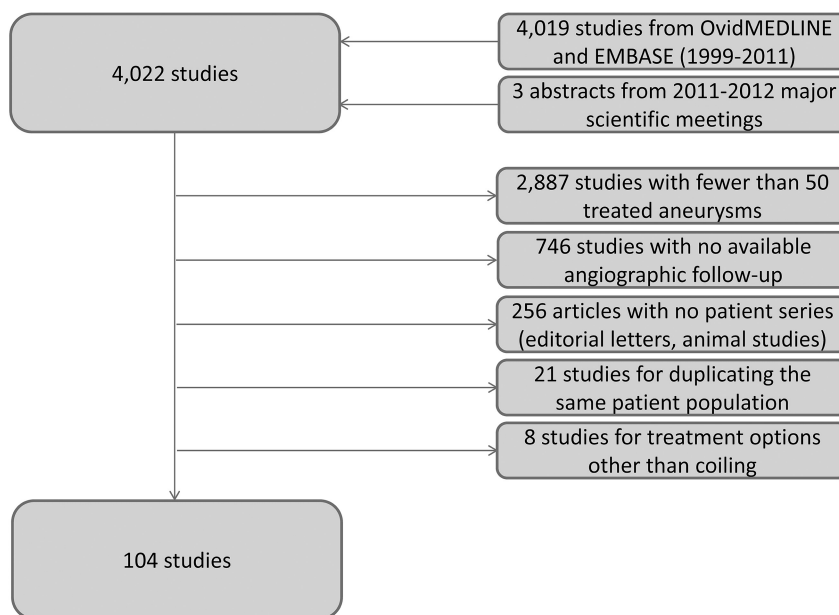
FIG 1. This figure depicts the search strategy in a flow chart along with reasons for exclusion.

Table 1: Baseline characteristics of study groups

Type of Interpretation	No. of Studies (%)	Mean Aneurysm Size (IRQ) (mm)	P Value	Initially Ruptured Aneurysms (%)	P Value	Mean Follow-Up Duration (IRQ) (mo.)	P Value	Bare Platinum-Coiled Aneurysms (%)	P Value
Noncore laboratory	96 (92.3)	7.8 (9.5)	.91	11,844/17,631 (67.2)	.07	19.7 (67.4)	.18	13,100/19,385 (67.6)	.16
Core laboratory	8 (7.7)	7.6 (0)		853/2232 (38.2)		11.2 (2.5)		1072/2749 (39)	
Single-reader	73 (70.2)	7.8 (7.9)	.72	9018/14,578 (61.9)	.08	15.9 (45.4)	.02	10,044/16,127 (62.3)	.39
Multiple-reader	31 (29.8)	7.6 (6.0)		3679/5285 (69.6)		25.5 (66.0)		4128/6007 (68.7)	

IRQ indicates interquartile range.

Table 2: Unadjusted, pooled rate of unfavorable outcomes

	Rate	95% CI	I^2 (%)	P Value[a]
Overall	0.170	0.153–0.190	90.3	—
Noncore laboratory interpretation	0.163	0.144–0.184	90.6	<.001
Core laboratory interpretation	0.234	0.193–0.283	85.2	
Single-reader	0.168	0.148–0.191	90.8	.06
Multiple-reader	0.174	0.142–0.213	89.1	

[a] Analysis of variance was used to test the difference between categories.

Table 3: Multivariate random-effects meta-regression

Variable	OR	SE	95% CI	P Value
Core laboratory vs noncore laboratory	5.60	2.87	2.01–15.60	.001
Single-reader vs multiple-reader studies	0.85	0.53	0.24–2.95	.794

SE indicates standard error.

101 studies (97.1%). The overall mean duration of follow-up was 18.8 months. Rupture status was described in 93 studies (89.4%). The average percentage of ruptured cases in the enrolled studies was 61%.

Eight (7.7%) of 104 studies used independent core laboratories. Thirty-one of the 104 studies (29.8%) used multiple readers. Two of these studies, CLARITY[24] and French Matrix Registry,[25] were core laboratory studies.

Table 1 presents baseline factors that have previously been associated with unfavorable outcomes, including aneurysm size, follow-up duration, and initial rupture status. Mean aneurysm size was similar among groups. Compared with core laboratory studies, site-read studies had nonsignificantly longer mean follow-up ($P = .18$), nonsignificantly larger proportions of initially ruptured aneurysms ($P = .07$), and nonsignificantly larger number of studies using bare platinum coils ($P = .16$). Compared with single-reader studies, multiple-reader studies had significantly longer mean follow-up ($P = .02$), nonsignificantly larger proportions of initially ruptured aneurysms ($P = .08$), and nonsignificantly larger number of studies using bare platinum coils ($P = .39$).

The pooled rates of unfavorable outcomes are shown in Table 2. Overall, the rate of unfavorable outcomes from all studies was 0.17 (95% CI, 0.15–0.19). Among core laboratory studies, the rate was 0.23 (95% CI, 0.19–0.28), which was significantly higher than the rate among noncore laboratories (0.16; 95% CI, 0.14–0.18; $P < .001$). However, no significant difference was found if the assessment was run by a single reader or by multiple readers ($P = .06$).

We used a random-effects meta-regression model to further explore the effects of core laboratory (yes vs no) and multiple image readers (1 reader vs multiple readers) on the rate of unfavorable outcomes after adjusting for baseline rupture, aneurysm size, follow-up time, and coiling device. Table 3 summarizes the results of the meta-regression model. After these adjustments,

there remained a significant increase in unfavorable outcomes from core laboratories compared with noncore laboratories (OR, 5.60; 95% CI, 2.01–15.60; $P = .001$). Again, no significant association was found between the unfavorable outcomes and the assessment of a single reader vs multiple readers (OR, 0.85; 95% CI, 0.24–2.95; $P = .794$).

Across the studies, substantial heterogeneity was observed in all of the pooled outcome estimates ($I^2 > 50\%$). The Egger regression asymmetry test suggested potential publication bias in this study ($P < .001$). A total of 97.1% of the included studies were observational studies. Only 3 (2.9%) of 104 studies were randomized controlled trials.[2,22,23] However, these studies were not designed to address our questions. We treated these randomized controlled trials as observational studies along with the rest. With use of the GRADE framework, the overall quality of this evidence (ie, confidence in the estimates) is low.[26,27]

DISCUSSION

In the current study, we demonstrated that reported unfavorable outcomes are more frequent in studies that use an independent core laboratory compared with studies that do not. Indeed, the rate of unfavorable outcomes was 43.6% greater in core laboratory compared with noncore laboratory studies. These differences are especially remarkable because the aneurysms from noncore laboratory studies had been followed longer and were more likely to have been initially ruptured than those from core laboratory cohorts. Both of these features would have been expected to increase, rather than decrease, rates of unfavorable outcomes. As such, our study strongly suggests that the use of the core laboratory itself may predict worse reported outcomes compared with noncore laboratory studies. The explanation for this observation may be that independent readers bring a more objective approach to interpretation compared with site readings. Furthermore, core laboratories usually use a limited number of experienced observers compared with treating sites, which may use observers with different levels of experience. As such, core laboratory utilization would be recommended for planning and interpretation of future trials.

Previous echocardiography studies[28-30] have demonstrated the superiority of core laboratory interpretations compared with site readings regarding variability and precision of study results. Recent standards from the American Society of Echocardiography recommend use of echocardiographic core laboratories for future trials.[7] Our current study focuses on the role of core laboratories in neuroradiology and also notes substantial differences between core laboratory and single-center studies.

Few prior studies focusing on aneurysm therapy have provided data on the effect of site vs core laboratory readings. Pierot et al[31] reported that rates of incomplete occlusion immediately after embolization were doubled by the use of a core laboratory compared with site readings. Our study is consistent with such findings regarding significantly different, and worse, outcomes noted by core laboratory than by site readings.

Our study had many limitations. First, we could not identify the readers' training and level of experience in most studies. Readings of neurosurgeons, neuroradiologists, or neurologists would arguably be subject to variations. Second, we were forced to combine patients who were monitored by either DSA or MR angiography despite possible technique-related variations. Most studies describe series of patients that reflect the center's practice. MR angiography was used at later stages without clear categorization of the patients and outcomes. Also, incomplete reporting of outcomes caused us to exclude most of the studies. Relatively few core laboratory studies were available, yet the 8 studies represent all of the available core laboratory studies to date. On the basis of variability in reported angiographic outcomes, we were forced to group multiple types of unfavorable results, including incomplete occlusion and Raymond class 3, with recanalization. Follow-up duration was variable as well, which was adjusted for in the analysis. Other variables that may affect outcomes, such as wide-neck aneurysms and anatomic location, were not clearly necessarily described; thus, we did not consider them. Finally, in this systematic review, we did not find any randomized controlled trials specifically designed to answer our questions. Observational studies are subject to high risk for bias because of baseline imbalance and potential outcome confounding factors. Also, ecologic bias resulting from pooling observational studies may have also affected our results. Thus, because of lack of high-quality evidence, we cannot rule out the possibility of a different conclusion. Further studies are clearly needed to provide higher-quality evidence.

CONCLUSIONS

Core laboratories tend to report higher rates of unfavorable outcomes compared with self-reporting centers. The current findings suggest that the higher unfavorable rates reflect reality rather than artificial inflation resulting from other confounders. Multiple-reader assessment, however, does not result in similar differences. Being more objective and having more experience than self-reporting centers, core laboratories should strongly be considered to evaluate angiographic outcomes for future trials and clinical practice.

REFERENCES

1. Connolly ES Jr, Rabinstein AA, Carhuapoma JR, et al. **Guidelines for the management of aneurysmal subarachnoid hemorrhage: a guideline for healthcare professionals from the American Heart Association/American Stroke Association.** Stroke 2012;43:1711–37
2. White PM, Lewis SC, Gholkar A, et al. **Hydrogel-coated coils versus bare platinum coils for the endovascular treatment of intracranial aneurysms (HELPS): a randomized controlled trial.** Lancet 2011;377:1655–62
3. Kole MK, Pelz DM, Kalapos P, et al. **Endovascular coil embolization of intracranial aneurysms: important factors related to rates and outcomes of incomplete occlusion.** J Neurosurg 2005;102:607–15
4. Fernandez Zubillaga A, Guglielmi G, Viñuela F, et al. **Endovascular occlusion of intracranial aneurysms with electrically detachable coils: correlation of aneurysm neck size and treatment results.** AJNR Am J Neuroradiol 1994;15:815–20
5. Sluzewski M, van Rooij WJ, Slob MJ, et al. **Relation between aneurysm volume, packing, and compaction in 145 cerebral aneurysms treated with coils.** Radiology 2004;231:653–58

6. Tamatani S, Ito Y, Abe H, et al. **Evaluation of the stability of aneurysms after embolization using detachable coils: correlation between stability of aneurysms and embolized volume of aneurysms.** *AJNR Am J Neuroradiol* 2002;23:762–67

7. Douglas PS, DeCara JM, Devereux RB, et al. **Echocardiographic imaging in clinical trials: American Society of Echocardiography standards for echocardiography core laboratories: endorsed by the American College of Cardiology Foundation.** *J Am Soc Echocardiogr* 2009;22:755–65

8. Dahiya A, Bolen M, Grimm RA, et al. **Development of a consensus document to improve multireader concordance and accuracy of aortic regurgitation severity grading by echocardiography versus cardiac magnetic resonance imaging.** *Am J Cardiol* 2012;110:709–14

9. Wakhloo AK, Gounis MJ, Sandhu JS, et al. **Complex-shaped platinum coils for brain aneurysms: higher packing density, improved biomechanical stability, and midterm angiographic outcome.** *AJNR Am J Neuroradiol* 2007;28:1395–400

10. Standhardt H, Boecher-Schwarz H, Gruber A, et al. **Endovascular treatment of unruptured intracranial aneurysms with Guglielmi detachable coils: short- and long-term results of a single-centre series.** *Stroke* 2008;39:899–904

11. Bracard S, Abdel-Kerim A, Thuillier L, et al. **Endovascular coil occlusion of 152 middle cerebral artery aneurysms: initial and midterm angiographic and clinical results.** *J Neurosurg* 2010;112:703–08

12. Berenstein A, Song JK, Niimi Y, et al. **Treatment of cerebral aneurysms with hydrogel-coated platinum coils (HydroCoil): early single-center experience.** *AJNR Am J Neuroradiol* 2006;27:1834–40

13. Grunwald IQ, Papanagiotou P, Struffert T, et al. **Recanalization after endovascular treatment of intracerebral aneurysms.** *Neuroradiology* 2007;49:41–47

14. Shankar JJ, Lum C, Parikh N, et al. **Long-term prospective follow-up of intracranial aneurysms treated with endovascular coiling using contrast-enhanced MR angiography.** *AJNR Am J Neuroradiol* 2010;31:1211–15

15. Ferns SP, Sprengers ME, van Rooij WJ, et al. **Coiling of intracranial aneurysms: a systematic review on initial occlusion and reopening and retreatment rates.** *Stroke* 2009;40:e523–29

16. Roy D, Milot G, Raymond J. **Endovascular treatment of unruptured aneurysms.** *Stroke* 2001;32:1998–2004

17. Clopper C, Pearson ES. **The use of confidence or fiducial limits illustrated in the case of the binomial.** *Biometrika* 1934;26:404–13

18. DerSimonian R, Laird N. **Meta-analysis in clinical trials.** *Control Clin Trials* 1986;7:177–88

19. Freemantle N, Cleland J, Young P, et al. **Beta blockade after myocardial infarction: systematic review and meta regression analysis.** *BMJ* 1999;318:1730–37

20. Higgins JP, Thompson SG, Deeks JJ, et al. **Measuring inconsistency in meta-analyses.** *BMJ* 2003;327:557–60

21. Egger M, Davey Smith G, Schneider M, et al. **Bias in meta-analysis detected by a simple, graphical test.** *BMJ* 1997;315:629–34

22. Molyneaux AJ, Fox A, Smeade 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—finals results of a prospective randomized trial.** In: *Proceedings of the American Society of Neuroradiology 50th Annual Meeting and the Foundation of the ASNR Symposium.* New York, NY: American Society of Neuroradiology; April 21–26, 2012.

23. Claiborne Johnson S, McDougal C, 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

24. Pierot L, Cognard C, Ricolfi F, et al. **Mid-term anatomic results after endovascular treatment of ruptured intracranial aneurysms with Guglielmi detachable coils and Matrix coils: analysis of the CLARITY series.** *AJNR Am J Neuroradiol* 2012;33:469–73

25. Pierot L, Leclerc X, Bonafé A, et al. **Endovascular treatment of intracranial aneurysms with Matrix detachable coils: midterm anatomic follow-up from a prospective multicenter registry.** *AJNR Am J Neuroradiol* 2008;29:57–61

26. Balshem H, Helfand M, Schunemann HJ, et al. **GRADE guidelines: 3. Rating the quality of evidence.** *J Clin Epidemiol* 2011;64:401–06

27. Guyatt GH, Oxman AD, Vist G, et al. **GRADE guidelines: 4. Rating the quality of evidence–study limitations (risk of bias).** *J Clin Epidemiol* 2011;64:407–15

28. Hole T, Otterstad JE, St. John Sutton M, et al. **Differences between echocardiographic measurements of left ventricular dimensions and function by local investigators and a core laboratory in a 2-year follow-up study of patients with an acute myocardial infarction.** *Eur J Echocardiogr* 2002;3:263–70

29. Baur LH, Schipperheyn JJ, van der Velde EA, et al. **Reproducibility of left ventricular size, shape and mass with echocardiography, magnetic resonance imaging and radionuclide angiography in patients with anterior wall infarction. A plea for core laboratories.** *Int J Card Imaging* 1996;12:233–40

30. Oh JK. **Is core laboratory essential for using echocardiography in clinical trials? Controlled vs random error.** *Eur J Echocardiogr* 2002;3:245–47

31. Pierot L, Cognard C, Ricolfi F, et al. **Immediate anatomic results after the endovascular treatment of ruptured intracranial aneurysms: analysis in the CLARITY series.** *AJNR Am J Neuroradiol* 2010;31:907–11