AJNR

Reliability and Agreement of 2D and 3D Measurements on MRAs for Growth Assessment of Unruptured Intracranial Aneurysms

K.M. Timmins, H.J. Kuijf, M.D.I. Vergouwen, M.J. Otten, Y.M. Ruigrok, B.K. Velthuis and I.C. van der Schaaf

Reliability and Agreement of 2D and 3D Measurements on MRAs for Growth Assessment of Unruptured Intracranial Aneurysms

K.M. Timmins, H.J. Kuijf, M.D.I. Vergouwen, M.J. Otten, Y.M. Ruigrok, B.K. Velthuis, and I.C. van der Schaaf

ABSTRACT

BACKGROUND AND PURPOSE: Reliable and reproducible measurement of unruptured intracranial aneurysm growth is important for unruptured intracranial aneurysm rupture risk assessment. This study aimed to compare the reliability and reproducibility of 2D and 3D growth measurements of unruptured intracranial aneurysms.

MATERIALS AND METHODS: 2D height, width, and neck and 3D volume measurements of unruptured intracranial aneurysms on baseline and follow-up TOF-MRAs were performed by two observers. The reliability of individual 2D and 3D measurements and of change (growth) between paired scans was assessed (intraclass correlation coefficient) and stratified for aneurysm location. The smallest detectable change on 2D and 3D was determined. Proportions of growing aneurysms were compared, and Bland-Altman plots were created.

RESULTS: Seventy-two patients with 84 unruptured intracranial aneurysms were included. The interobserver reliability was good-to-excellent for individual measurements (intraclass correlation coefficient > 0.70), poor for 2D change (intraclass correlation coefficient < 0.5), and good for 3D change (intraclass correlation coefficient $= 0.76$). For both 2D and 3D, the reliability was location-dependent and worse for irregularly shaped aneurysms. The smallest detectable changes for 2D height, width, and neck and 3D volume measurements were 1.5, 2.0, and 1.9 mm and 0.06 mL, respectively. The proportion of growing unruptured intracranial aneurysms decreased from 10% to 2%, depending on the definition of growth (1 mm or the smallest detectable changes for 2D and 3D).

CONCLUSIONS: The interobserver reliability of the size measurements of individual 2D and 3D unruptured intracranial aneurysms was good-to-excellent but lower for 2D and 3D growth measurements. For growth assessment, 3D measurements are more reliable than 2D measurements. The smallest detectable change for 2D measurements was larger than 1mm, the current clinical definition of unruptured intracranial aneurysm growth.

ABBREVIATIONS: ICC = intraclass correlation coefficient; SDC = smallest detectable change; UIA = unruptured intracranial aneurysm

In the adult population, the prevalence of unruptured intracranial aneurysms (UIAs) is around 3%.[1] Intracranial aneurysm rupture leads to SAH with a high case fatality rate. The PHASES (Population, Hypertension, Age, Size, Earlier subarachnoid hemorrhage and Site) study found the 5-year rupture risk of UIAs to be, on average, 3.4% (0.5%–17.8%), depending on patient and aneurysm characteristics.[2] When one makes a treatment decision, the risk of aneurysm rupture is weighed against the complication risk of treatment. Aneurysm size is a key determinant in the prediction models of rupture risk.[2,3] If a multidisciplinary team decides against preventive aneurysm treatment, the UIA is followed up with repeat TOF-MRA or CTAs to detect potential aneurysm growth. Growth is an additional rupture risk factor,[4] and if detected, preventive treatment should be considered. TOF-MRA has been shown to systematically underestimate the size and volume of the aneurysm compared with the criterion standard DSA.[5] However, noninvasive TOF-MRA is the first-choice imaging method for follow-up imaging in clinical practice because neither contrast agent administration nor radiation exposure is required.[6,7]

Assessment of UIAs is performed by taking 2D size measurements of aneurysms on MRA/CTA using electronic calipers. The 3D nature of UIAs makes 2D measurements difficult and

dependent on optimal orientation in multiplanar imaging. The 2D measurements by human observers are reported to have comparable mediocre reproducibility on both CTAs and MRAs.[8,9] These UIA measurements are relevant when comparing aneurysm size in a follow-up scan to assess aneurysm growth. Aneurysm growth is defined as an increase in either 2D height or width of at least 1 mm.[10] A reliable measurement method with good agreement is important for risk assessment. In this context, the reliability depends on the variability of the aneurysm sizes among patients. The agreement describes the interobserver measurement error and is characteristic of the measurement method itself. Without knowledge of reliability and agreement, it is unclear whether a measured change in aneurysm size between baseline and follow-up scans represents real growth or is attributable to observer or scan variations.

In this study, we investigated the reliability and reproducibility of 2D size and 3D volume measurements of UIAs and change in aneurysm size and volume between baseline and follow-up MRAs. For an agreement measure, we calculated the smallest detectable change (SDC) and assessed agreement using Bland-Altman plots.

MATERIALS AND METHODS

Study Population

We included 72 patients from a series of patients with UIAs from the University Medical Center Utrecht who met the following inclusion criteria: 1) A TOF-MRA was available at both the baseline admission scan and follow-up, 2) the follow-up scan was performed at least 6 months after the baseline scan, and 3) the patient had at least 1 untreated UIA present on both baseline and follow-up MRA. Any treated aneurysm in these subjects was excluded from this study. The most recent follow-up scan in which the UIA remained untreated and unruptured was used. The scans had an in-plane resolution range of 0.175–1.04 mm and a section thickness range of 0.399–1.2 mm. All scans were obtained from 2004 to 2019. Due to the nature of the scans, protocols varied, but all scans were obtained on 1T, 1.5T, or 3T scanners with a median TR of 23 ms and a median TE of 6.4 ms across all scans. This retrospective study required no formal consent from participants. The data that support the findings of this study are available from the corresponding author on reasonable request.

Measurements

2D Measurements. Manual 2D measurements of the UIAs were performed on the IntelliSpace Portal (Phillips Healthcare). Measurements were obtained using electronic calipers on the TOF-MRAs, which could be rotated in the software. The aneurysm height, width, and neck were measured on the TOF-MRAs on a 0.1-mm scale[11,12] as shown in parts A and C in the Figure. Aneurysm height was defined as the maximum distance from the aneurysm neck to the dome. Aneurysm width was measured perpendicular to the height along the maximum width of the UIA. The neck was measured as the maximum width of the UIA where it attached to the parent vessel. Observers determined whether the UIA shape was regular or irregular.
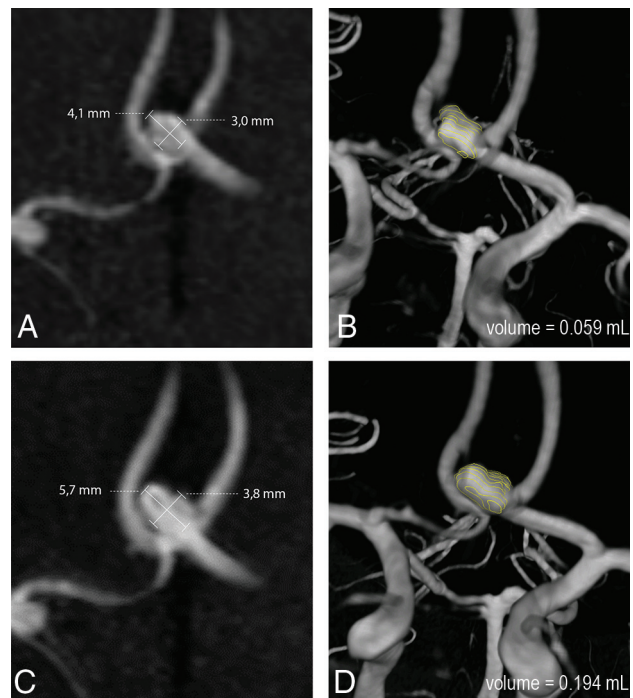
FIGURE. Baseline (*A* and *B*) and follow-up (*C* and *D*) TOF-MRA with an anterior communicating artery aneurysm that shows growth when measured in 2D (*A* and *C*) and in 3D (*B* and *D*).

All 2D measurements were performed independently by 2 observers. The observers were a neuroradiologist (I.C.v.d.S., with 15 years of experience) and a general radiologist (M.J.O., with 10 years of experience, including cerebral MRA evaluation). Individual measurements were first obtained on the baseline scan, then on the follow-up scan of the same patient. The observers were not blinded to the time order of the scans and had the baseline for comparison, as is standard in clinical practice.

3D Measurements. For 3D measurement, the UIAs were segmented from the TOF-MRAs using in-house-developed software implemented in MeVisLab (MeVis Medical Solutions). A contour was drawn around the outline of the aneurysm on axial slices, and the parent vessels were not included (Figure). The UIA volume (in milliliters) was determined on the basis of the voxels contained within the contours and the MRA voxel size. Annotations were performed independently by two observers, first on the baseline scan, followed by the follow-up scan of the same patient. The observers were the neuroradiologist (I.C.v.d.S.) and a trained medical student (D.S.).

Statistical Analysis

First, the interobserver reliability of the individual 2D measurements (height, width, neck) and 3D measurements (volume) of the aneurysms was determined. Second, on the basis of the 2D and 3D size measurements, changes in size and volume (growth) between paired baseline and follow-up scans for 2D (difference in height, width, and neck in mm) and 3D (volume difference in milliliters) were calculated. Third, the interobserver reliability of these changes in size (2D) and volume (3D) measurements was assessed by computing the intraclass correlation coefficient

Table 1: Interobserver size and volume measurements[a]

Parameters	Height (mm)	Width (mm)	Neck (mm)	Volume (mL)
Observer A measurement	3.4 (2.4–4.4)	3.4 (2.2–4.5)	2.6 (2.0–3.5)	0.0278 (0.0117–0.0578)
Observer B measurement	3.4 (2.5–4.6)	3.2 (2.2–4.3)	2 (1.6–2.8)	–
Observer C measurement	–	–	–	0.0227 (0.0090–0.0470)
Absolute Diff$_{obs}$	0.4 (0.2–0.6)	0.45 (0.2–0.8)	0.55 (0.3–1.0)	0.0091 (0.0036–0.0206)
ICC$_{agreement}$ (95% CI)	0.93 (0.90–0.95)	0.85 (0.80–0.89)	0.74 (0.31–0.87)	0.98 (0.97–0.98)

Note:—— indicates no measurement; ICC$_{agreement}$, Intraclass correlation coefficient on absolute agreement between observers' measurements; Absolute Diff$_{obs}$, absolute difference between observers' measurements.

[a] 2D and 3D measurements of the aneurysms by observers A and B for 2D and observers A and C for 3D. Total: 168 aneurysms, including both baseline and follow-up scans. Each measurement is provided as a median (quartiles 1–3). Reliability is in the bottom row as an ICC on absolute agreement (95% confidence interval).

Table 2: Interobserver change measurements[a]

Parameters	Height Change (mm)	Width Change (mm)	Neck Change (mm)	Volume Change (mL)
Observer A measurement	0.2 (−0.1–0.7)	0.1 (−0.2–0.4)	0 (−0.2–0.4)	0.00001 (−0.0043–0.0011)
Observer B measurement	0.1 (−0.1–0.5)	0.1 (−0.2–0.6)	0 (−0.1–0.3)	–
Observer C measurement	–	–	–	0.0015 (−0.0054–0.0106)
Absolute Diff$_{obs}$	0.4 (0.2–0.7)	0.4 (0.2–0.6)	0.3 (0.1–0.7)	0.0057 (0.0024–0.0135)
ICC$_{agreement}$ (95% CI)	0.46 (0.27−0.61)	0.45 (0.26−0.60)	0.26 (0.06−0.46)	0.76 (0.65−0.76)

Note:—— indicates no measurement; ICC$_{agreement}$, Intraclass correlation coefficient on absolute agreement between observers' measurements; Absolute Diff$_{obs}$, absolute difference between observers' measurements.

[a] Change between baseline and follow-up measurements of the 2D height, width, neck and 3D volume of the aneurysm by observers A and B for 2D and observers A and C for 3D. Total: 84 baseline–follow-up pairs measured by 2 observers. Substantial positive differences between baseline and follow-up may indicate growth of the aneurysm. Each measurement is provided as a median (quartiles 1–3). Reliability of the differences is provided in the bottom row as the ICC on absolute agreement (95% confidence interval).

(ICC). The ICC was calculated using a single-measurement, absolute-agreement, 2-way random-effects model.[13] An ICC above 0.9 represents excellent reliability; between 0.75 and 0.90, good reliability; between 0.5 and 0.75, moderate reliability; and lower than 0.5, poor reliability.[13,14] The interobserver reliability for detecting change in 2D and 3D measurements was compared in regular and irregular aneurysms.

The SDC was computed on the basis of the 2D and 3D measurements to assess the interobserver agreement. The SDC represents the minimal change that an aneurysm measurement must show to ensure that the observed change is real and not just due to measurement error. For both 2D size and 3D volume measurements, we calculated the standard error of measurement (SEM) using the ICC previously determined. The SDC was calculated from the standard error, $SEM_{agreement}$, $SDC = 1.96 \sqrt{2} \ SEM_{agreement}$, where $SEM_{agreement} = SD \sqrt{1 - ICC_{agreement}}$, and SD is the standard deviation of all measurements.[15]

Bland-Altman plots for the interobserver difference between the change in 2D and 3D measurements between baseline and follow-up scans were created to assess agreement. The difference between each observer and the overall mean of both observers was calculated and plotted. The limits of agreement from the mean (±1.96 SD) were determined. Measurements outside the limits of agreement were considered outliers.

The number of UIAs with change in 2D height and/or width measurements larger than 1 mm, the current clinical definition of aneurysm growth,[10] was determined. Next, the number of UIAs with a change in 2D height and/or width and volume larger than the determined 2D and 3D SDCs was determined. The proportion of UIAs showing growth based on the 1-mm clinical definition versus the proportion of UIAs with growth based on the SDCs was compared.

Finally, a subanalysis was performed stratifying the reliability of change measurements for aneurysm location: anterior cerebral or communicating artery, internal carotid artery, posterior communicating artery, MCA, and posterior circulation.

All data analyses were conducted using Pandas, SciPy, and Pengouin[16] toolboxes with Python 3.7 (https://www.python.org/downloads/release/python-370/).

RESULTS

We included 72 patients with 84 UIAs. The mean age was 53 years (range, 27–73 years), and 71% were women. Most patients had 1 UIA ($n = 63$). The median time between baseline and follow-up scans was 4.7 years (range, 0.9–13.1 years). The median aneurysm height was 3.4 mm (range, 0.8–15 mm). Twenty-two percent of aneurysms were located at the anterior cerebral artery/anterior communicating artery, 27% at the ICA or posterior communicating artery, 38% at the MCA, and 13% in the posterior circulation. The Figure shows an example of a growing aneurysm measured in 2D and 3D.

The interobserver reliability of the 2D size and 3D volume measurements is summarized in Table 1. The ICC of the individual 2D size measurements was excellent for height (0.93), good for width (0.85), and moderate for the neck (0.74). The ICC for the individual 3D volume measurement was excellent (0.98).

The ICCs for the change in measurements (growth) between the paired baseline–follow-up scans for the 2 observers are shown in Table 2. The ICC for the change in 2D measurements was poor for height (0.46), width (0.45), and neck (0.26). The ICC for the change in 3D volume measurements was good (0.76). Irregularly shaped aneurysms had a lower reliability for 2D change in height and width (ICCs = 0.23, 0.38) and 3D change in volume (ICC = 0.60) than for regular aneurysms (ICCs = 0.57, 0.47, 0.83, respectively).

On the basis of the standard error of measurement for agreement, between the 2 observers, the SDC for 2D measurements was 1.5 mm for height, 2.0 mm for width, and 1.9 mm for neck. The SDC for 3D volume measurement was 0.062 mL.

The Online Supplemental Data show Bland-Altman plots for the interobserver difference between the change in 2D and 3D

measurements between baseline and follow-up scans. The Bland-Altman plots show that there are 4–6 outliers that fall outside the limits of agreement for all change measurements between size and volume. About half of these outliers (55%) were the same for 2D and 3D and were classified as irregularly shaped by the observers. There was no relation between aneurysm size and the outliers.

The number of UIAs with a change in size measurements larger than 1 mm and a change in size and volume larger than the SDCs is shown in the Online Supplemental Data. The proportion of UIAs with growth based on the definition of 1 mm was 10%, compared with 2% when using a 1.5-mm change in height as a cutoff value (SDC for 2D height) or a 2.0-mm change in width as a cutoff value (SDC for 2D width) and a 0.062-mL change (SDC for 3D) as cut-off value.

The Online Supplemental Data indicate the reliability of the change in measurements in different locations. The reliability was found to be location-dependent for both 2D and 3D; however, 3D measurements were more reliable than 2D measurements across all locations (ICC > 0.5).

DISCUSSION
In this study, interobserver reliability was better for 3D than 2D measurements of UIAs, both for individual size and detection of change in size (growth). Overall, the interobserver reliability of both 2D and 3D measurements was lower for the detection of change (growth) compared with measurements on individual scans. The SDC between the baseline and follow-up scan for 2D measurements was substantially larger than the current clinical definition (1 mm), and proportions of UIAs showing growth decreased more than three-quarters depending on the growth definition.

Many studies have investigated MRAs for UIA diagnosis.[17] However, few studies have investigated the interobserver reliability of 2D measurements from individual MRAs of patients, and no studies have fully investigated the reliability and agreement of growth measurements between baseline and follow-up MRAs of the same patient. The results of studies for individual 2D height and width measurements are similar to our findings, with the lowest reliability for measuring the neck. Kim et al[8] studied intra- and interobserver individual 2D measurement variability of 33 aneurysms with a mean size of 5.1 mm, finding an ICC of 0.83–0.99 on MRAs with the lowest reliability for the neck measurement (ICC = 0.83-0.86). Mine et al[18] compared the diagnosis and measurements of UIAs between DSAs and MRAs. Three readers assessing 56 aneurysms in MRAs determined an interobserver agreement between individual 2D maximal diameter as moderate-to-substantial ($\kappa = 0.53$–0.66) and the neck measurement as fair-to-moderate ($\kappa = 0.20$–0.41). The lower ICC for the neck is likely due to difficulty in defining an aneurysm neck, particularly if there are branching vessels emerging from the neck. This lower measurement reliability for neck measurements may have implications for treatment-planning and complication risk assessment.[19] For aneurysm growth assessment, the neck measurement is less important because height and width measurements are commonly used.[10,11]

With ever-improving image analysis techniques, 3D measurements of UIAs[20-22] are more commonly investigated, but little is known of their reliability or reproducibility for individual size and growth measurement of UIAs in TOF-MRAs. D'Argento et al[23] found no significant difference in intra- and interobserver variability of automatic and manual 2D size measurements of UIAs on 3D DSAs and CTAs.

We determined the ICC of absolute agreement to include the systematic error of both observers and random residual errors. A substantially lower ICC for change measurements (growth) between paired baseline–follow-up scans was determined, relative to measurements from individual scans. The ratio of the systematic measurement error compared with the individual aneurysm size is smaller than the ratio of the measurement error compared with the change in aneurysm size. Thus, a small measurement error in individual measurements can have a larger influence on the subsequent change measurements in paired scans.

The interobserver agreement in 2D and 3D measurements was assessed by determining the SDC. The SDC for both 2D and 3D measurements was relatively large, compared with the median aneurysm size (3.4 mm) and median aneurysm volume (0.025 mL). For example, for 2D height, the SDC of 1.5 mm was about half of median aneurysm height. This study has a large proportion of small aneurysms, and the ratio of the SDC to aneurysm size would be better (lower) in larger aneurysms. However, because most patients who undergo follow-up MRAs have small UIAs, our population represents the clinical situation. The SDC for the 2D measurements is larger than the 1 mm used in the current definition of aneurysm growth.[10] The number of UIAs showing growth according to threshold values of the SDC of 2D and 3D measurements decreased by more than three-quarters compared with this 1-mm threshold. This finding shows the influence of the thresholds for growth definition and has potential important clinical consequences for treatment decisions based on aneurysm growth.

The Bland-Altman plots (Online Supplemental Data) show that 3D interobserver differences were more similar than 2D measurements because the measurements were closer together. Most outliers for both the 2D and 3D measurements were irregularly shaped. We also found that irregular aneurysms had a lower interobserver reliability for detecting change in both 2D size and 3D volume measurements. Irregular aneurysm shape is a risk factor for rupture.[12,24] 2D measurements and shape assessment of aneurysms are influenced by the selected viewing angle. 3D volume measurements allow a more complete shape of the UIA to be assessed with a single, rotationally-invariant measure. Furthermore, 3D segmentation may allow quantitative shape assessment of UIAs, which would be potentially beneficial in risk assessment.[12]

We found that aneurysm location affects the reliability of 2D and 3D measurements. We found that the reliability of 3D volume measurements was higher and more consistent for all locations than 2D size measurements.

There were some limitations in our study. One limitation was that the 3D measurements were determined from segmentations based on 2D annotations on axial slices, which is time-consuming and the aneurysm neck definition could be difficult, particularly when the parent vessel did not lie in-plane. Furthermore, the difference in experience of the second observers for 2D (radiologist) and 3D (student) measurements may have introduced bias. If this had influenced our results, it would be toward less agreement for

the 3D measurement between the student and the neuroradiologist. However, we found higher agreement in 3D than in 2D.

Second, most scans had small aneurysms with a median diameter of 3.4 mm (range, 0.8–15 mm). The population of patients with small UIAs is, however, representative of patients who undergo follow-up imaging. Because rupture risk increases with aneurysm size, the larger UIAs are more often treated. The protocol and quality of the MRAs between baseline and follow-up differed in some cases, possibly resulting in measurement differences, but they are realistic for clinical practice.

This study investigates TOF-MRAs only because this is the preferred imaging method for follow-up of UIAs.[7]

Our findings of a large SDC for 2D size measurements may have implications for the definition of clinical aneurysm growth[10] and growth/rupture models. This subject requires further study because it would have important consequences for rupture and treatment assessment of UIAs. 2D and 3D measurements cannot be directly compared, but instead a standard growth definition should be used for both. The higher reliability of 3D measurements compared with 2D measurements implies that 3D measurements may be important for accurate assessment of aneurysm growth on TOF-MRA. Automatic or semi-automatic 3D UIA segmentation would allow faster and less operator-dependent aneurysm volume measurement for standard 3D growth assessment, alongside quantitative 3D morphologic characterization of UIAs.

CONCLUSIONS

This study found that 3D change measurements are more reliable than 2D with regard to assessing the change in size and volume measurements of UIAs. The SDC for 2D measurements was found to be larger than the current definition for clinical growth, suggesting that more studies into the reliability of 2D measurement on MRA should be performed. This study opens the door for development and incorporation into of automatic and semi-automatic segmentations and volumetric growth assessments of UIAs into clinical practice.

REFERENCES

1. Vlak MH, Algra A, Brandenburg R, et al. **Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis.** *Lancet Neurol* 2011;10:626–36 CrossRef Medline
2. Greving JP, Wermer MJ, Brown RD, et al. **Development of the PHASES score for prediction of risk of rupture of intracranial aneurysms: a pooled analysis of six prospective cohort studies.** *Lancet Neurol* 2014;13:59–66 CrossRef Medline
3. Tominari S, Morita A, Ishibashi T, et al. Unruptured Cerebral Aneurysm Study Japan Investigators. **Prediction model for 3-year rupture risk of unruptured cerebral aneurysms in Japanese patients.** *Ann Neurol* 2015;77:1050–59 CrossRef Medline
4. Villablanca JP, Duckwiler GR, Jahan R, et al. **Natural history of asymptomatic unruptured cerebral aneurysms evaluated at CT angiography: growth and rupture incidence and correlation with epidemiologic risk factors.** *Radiology* 2013;269:258–65 CrossRef Medline
5. Takao H, Murayama Y, Ishibashi T, et al. **Comparing accuracy of cerebral aneurysm size measurements from three routine investigations: computed tomography, magnetic resonance imaging, and digital subtraction angiography.** *Neurol Med Chir (Tokyo)* 2010;50:893–99 CrossRef Medline
6. Thompson BG, Brown RD, Amin-Hanjani S, et al. American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, and Council on Epidemiology and Prevention; American Heart Association; American Stroke Association. **Guidelines for the Management of Patients with Unruptured Intracranial Aneurysms: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association.** *Stroke* 2015;46:2368–2400 CrossRef Medline
7. Malhotra A, Wu X, Gandhi D, et al. **Management of small, unruptured intracranial aneurysms.** *World Neurosurg* 2020;135:379–80 CrossRef Medline
8. Kim HJ, Yoon DY, Kim ES, et al. **Intraobserver and interobserver variability in CT angiography and MR angiography measurements of the size of cerebral aneurysms.** *Neuroradiology* 2017;59:491–97 CrossRef Medline
9. Forbes G, Fox AJ, Huston J, et al. **Interobserver variability in angiographic measurement and morphologic characterization of intracranial aneurysms: a report from the International Study of Unruptured Intracranial Aneurysms.** *AJNR Am J Neuroradiol* 1996;17:1407–15 Medline
10. Hackenberg KA, Algra A, Salman RA; the Unruptured Aneurysms and SAH CDE Project Investigators, et al. **Definition and prioritization of data elements for cohort studies and clinical trials on patients with unruptured intracranial aneurysms: proposal of a multidisciplinary research group.** *Neurocrit Care* 2019;30:87–101 CrossRef Medline
11. Backes D, Vergouwen MDI, Tiel Groenestege AT, et al. **PHASES score for prediction of intracranial aneurysm growth.** *Stroke* 2015;46:1221–26 CrossRef Medline
12. Backes D, Rinkel GJ, Greving JP, et al. **ELAPSS score for prediction of risk of growth of unruptured intracranial aneurysms.** *Neurology* 2017;88:1600–06 CrossRef Medline
13. Koo TK, Li MY. **A guideline of selecting and reporting intraclass correlation coefficients for reliability research.** *J Chiropr Med* 2016;15:155–63 CrossRef Medline
14. Kottner J, Audigé L, Brorson S, et al. **Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed.** *J Clin Epidemiol* 2011;64:96–106 CrossRef Medline
15. de Vet HC, Terwee CB, Knol DL, et al. **When to use agreement versus reliability measures.** *J Clin Epidemiol* 2006;59:1033–39 CrossRef Medline
16. Vallat R. **Pingouin: statistics in Python.** *JOSS* 2018;3:1026 CrossRef
17. White PM, Teasdale EM, Wardlaw JM, et al. **Intracranial aneurysms: CT angiography and MR angiography for detection: prospective blinded comparison in a large patient cohort.** *Radiology* 2001;219:739–49 CrossRef Medline
18. Mine B, Pezzullo M, Roque G, et al. **Detection and characterization of unruptured intracranial aneurysms: comparison of 3T MRA and DSA.** *J Neuroradiol* 2015;42:162–68 CrossRef Medline
19. Algra AM, Lindgren A, Vergouwen MD, et al. **Procedural clinical complications, case-fatality risks, and risk factors in endovascular and neurosurgical treatment of unruptured intracranial aneurysms: a systematic review and meta-analysis.** *JAMA Neurol* 2019;76:282–93 CrossRef

20. Leemans EL, Cornelissen BM, Slump CH, et al. **Comparing morphology and hemodynamics of stable-versus-growing and grown intracranial aneurysms.** *AJNR Am J Neuroradiol* 2019;40:2102–10 CrossRef Medline

21. Chan SH, Wong KS, Woo YM, et al. **Volume measurement of the intracranial aneurysm: a discussion and comparison of the alternatives to manual segmentation.** *J Cerebrovasc Endovasc Neurosurg* 2014;16:358 CrossRef Medline

22. Piotin M, Gailloud P, Bidaut L, et al. **CT angiography, MR angiography and rotational digital subtraction angiography for volumetric assessment of intracranial aneurysms: an experimental study.** *Neuroradiology* 2003;45:404–09 CrossRef Medline

23. D'Argento F, Pedicelli A, Ciardi C, et al. **Intra- and inter-observer variability in intracranial aneurysm segmentation: comparison between CT angiography (semi-automated segmentation software stroke VCAR) and digital subtraction angiography (3D rotational angiography).** *Radiol Med* 2021;126:484–93 CrossRef Medline

24. Lindgren AE, Koivisto T, Björkman J, et al. **Irregular shape of intracranial aneurysm indicates rupture risk irrespective of size in a population-based cohort.** *Stroke* 2016;47:1219–26 CrossRef Medline