Ultrasound-determined diameter measurements are more accurate than axial computed tomography after endovascular aortic aneurysm repair

Sukgu M. Han, MD,[a] **Kaushel Patel, MD,**[b] **Vincent L. Rowe, MD,**[a] **Susana Perese, BS, RVT,**[a] **Aaron Bond, PhD,**[a] and **Fred A. Weaver, MD, MMM,**[a] *Los Angeles, Calif*

Objective: **This study evaluated the correlation of ultrasound (US)-derived aortic aneurysm diameter measurements with centerline, three-dimensional (3-D) reconstruction computed tomography (CT) measurements after endovascular aortic aneurysm repair (EVAR).**

Methods: **Concurrent CT and US examinations from 82 patients undergoing post-EVAR surveillance were reviewed. The aortic aneurysm diameter was defined as the major axis on the centerline images of 3-D CT reconstruction. This was compared with US-derived minor and major axis measurements, as well as with the minor axis measurement on the conventional axial CT images. Correlation was evaluated with linear regression analyses. Agreement between different imaging modalities and measurements was assessed with Bland-Altman plots.**

Results: **The correlation coefficients from linear regression analyses were 0.92 between CT centerline major and US minor measurements, 0.94 between CT centerline major and US major measurements, and 0.93 between CT minor and centerline major measurements. Bland-Altman plots showed a mean difference of 0.11 mm between US major and CT centerline measurements compared with 5.38 mm between US minor and CT centerline measurements, and 4.25 mm between axial CT minor and centerline measurements. This suggested that, compared with axial CT and US minor axis measurements, US major axis measurements were in better agreement with CT centerline measurements. Variability between major and minor US and CT centerline diameter measurements was high (standard deviation of difference, 4.27-4.84 mm). However, high variability was also observed between axial CT measurements and centerline CT measurements (standard deviation of difference, 4.36 mm).**

Conclusions: **The major axis aneurysm diameter measurement obtained by US imaging for surveillance after EVAR correlates well and is in better agreement with centerline 3-D CT reconstruction diameters than axial CT. (J Vasc Surg 2010;51:1381-9.)**

Endovascular aortic aneurysm repair (EVAR) has been shown to have lower short-term morbidity and mortality than open abdominal aortic aneurysm (AAA) repair, but its long-term outcomes have not been fully elucidated, making life-long post-EVAR surveillance essential.[1] Criteria for successful EVAR include absence of endoleak and stent-related problems along with a decrease in maximum aneurysmal diameter. Particularly, continual expansion of the aneurysm sac diameter is suggestive of endoleak or endotension and is thought to be predictive of future risk of rupture.[2,3]

Although three-dimensional (3-D) reconstruction is routinely used for the preoperative evaluation of AAA, the current standard for postoperative surveillance is still axial CT scan without 3-D reconstruction.[4] However, in the AAA patient population where advanced age and coexisting morbidities such as renal insufficiency are common, repeated radiation and iodinated contrast exposure may lead to adverse outcomes.[1]

Furthermore, previous studies have documented significant interobserver and intraobserver variability in maximal aortic diameter measurements from cross-sectional CT images.[5-8] Specifically, axial images on the CT scans may overestimate the true aortic diameter caused by the oblique cut of the images relative to the angle of the aorta.[9,10] Various proposed methods to minimize such error include using calipers while taking measurements, having fewer radiologists and surgeons reading CT scans, and taking the minor axis diameter measurement on a CT slice.[8-10]

Duplex ultrasound (DUS) imaging is a noninvasive modality that may be a useful alternative to CT scans for post-EVAR surveillance. It has been proposed that ultrasound technologists can correct for the angulation error seen in CT scan measurements by placing the US probe perpendicular to the course of the aorta.[8] In addition, US imaging has a high degree of correlation with CT scans and a similar degree of variability in AAA diameter measurements.[8] However, specific methods of US diameter measurements have been inconsistent at best.[11]

From the Division of Vascular Surgery and Endovascular Therapy, CardioVascular Thoracic Institute, University of Southern California[a]; and Department of Surgery, Kaiser Sunset Medical Center.[b]

Competition of interest: none.

Presented at the Western Vascular Society Annual Meeting, Tucson, Ariz, Sept 19-20, 2009.

Reprint requests: Fred A. Weaver, MD, MMM, University of Southern California, Division of Vascular Surgery and Endovascular Therapy, 1520 San Pablo, Ste 4300, Los Angeles, CA 90033 (e-mail: fweaver@surgery.usc.edu).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a competition of interest.

0741-5214/$36.00

doi:10.1016/j.jvs.2010.01.033

This lack of a standardized US imaging protocol for AAA diameter measurements is illustrated by the heterogeneous methods used in some of the landmark aneurysm studies. For example, anteroposterior (AP) diameter measurements on screening US images were used in the United Kingdom Small Aneurysm Trial (UKSAT),[12] whereas the Multicentre Aneurysm Screening Study (MASS)[13] trials used transverse diameters. Although past studies have used the axial CT scan as the standard against which US imaging was evaluated, recent advances in vascular image processing, particularly 3-D reconstruction and availability of orthogonal, centerline aorta images, may provide a new gold standard in diameter measurements.[14]

In turn, maximal aortic diameter measurements determined by US imaging and axial CT scans can now be evaluated against the centerline measurements. Therefore, our study assessed the accuracy of the aortic diameter measurements at the region of maximal dilatation obtained from US imaging and axial CT with reference to centerline CT reconstructions in patients after EVAR repair. Specifically, we evaluated different methods of diameter measurements on US imaging by including major and minor axes measurements in our comparison.

METHODS

The protocol for this study was approved by University of Southern (USC) California Institutional Review Board (IRB). Patient consent was exempted because of the retrospective nature of our review.

Patients. From 2003 to 2007, 204 patients underwent EVAR for infrarenal aortic aneurysm. EVAR was performed with Excluder (W. L. Gore and Assoc, Flagstaff, Ariz), AneuRx (Medtronic, Minneapolis, Minn), or Zenith (Cook, Bloomington, Ind) endografts. All cases were performed with the patient under general anesthesia in a dedicated operating suite with fixed angiographic equipment (Phillips Allua 2001, Bothell, Wash).

The concurrent studies were routinely ordered according to our institutional protocol during the study period. DUS was the sole imaging used in 48 patients with elevated serum creatinine (≥1.5 mg/dL). Concurrent studies were not performed in 69 patients because they wanted their follow-up imaging done at a local imaging center. Eighty-seven patients underwent a CT scan and US imaging ≤2 weeks of one another during their follow-up and were selected for review in our study. The mean post-EVAR follow-up was 4.3 months (range, 1-14.5 months). The mean and median preoperative aneurysm sizes were 5.68 and 5.50 cm, respectively, by centerline measurements on 3-D CT reconstruction.

DUS imaging. All DUS scans were performed by one of the four certified vascular technologists in the USC University Hospital vascular laboratory. Philips ATL HDI 5000 US machines (Oceanside, Calif) with 4-MHz transducers were used for scans. According to the Intersocietal Commission for the Accreditation of Vascular Laboratories (ICAVL)-approved institutional protocol, patients were

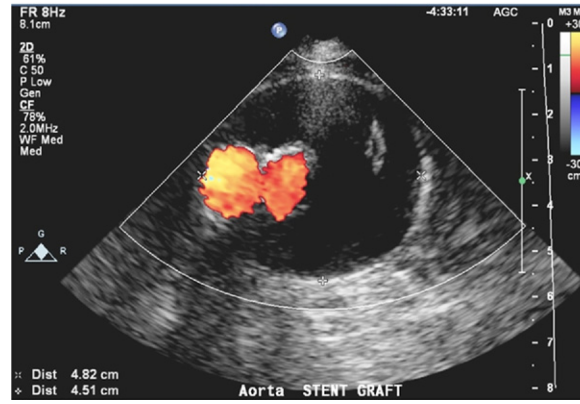

Fig 1. The ultrasound-determined anteroposterior and transverse maximal aortic diameter is shown. The larger of the two measurements was recorded as the major axis measurement, and the smaller as the minor axis measurement.

instructed to fast before their scan. The probe was routinely maintained perpendicular to the aortic blood flow by the technologists to yield the most accurate orthogonal diameters. To ensure the perpendicular orientation of the probe, the entire length of the aorta was first surveyed with B mode to acquire a general idea of the regions of angulation.

Second, color flow was visualized while the probe was oriented in a transverse fashion. Because of the pulsatile nature of aortic blood flow, when the probe is oriented perpendicular to it, a mix of alternating red and blue is visualized in the color window. When the probe is oriented obliquely, flow is mainly seen in one predominant color.[15]

Cross-sectional image was captured during maximal dilation of the aortic sac during the cardiac cycle. Then, the certified technologists used on-screen calipers to obtained maximal anteroposterior (AP) and transverse diameter measurements of the aorta from adventitia to adventitia (Fig 1).

All of the images and measurements were reviewed by a single vascular surgeon (K.P.). The study excluded five technically difficult studies, due to inability to view the entire aorta secondary to bowel gas or patient body habitus. Diameter measurements were entered into an Excel spreadsheet (Microsoft Corp, Redmond, Wash) and the images were archived for reference. The technologists performing the studies were blinded to the results of the CT scan.

CT scan and 3-D reconstruction. Abdominal and pelvic CT scans were performed with approximately 100 mL of nonionic contrast ≤2 weeks of the US study in each patient. All scans were performed with Lightspeed Pro-16, a multidetector spiral CT scanner (General Electric, Piscataway, NJ) with 16 detectors. The detector width was 1.5 mm with a collimation of 2.5 mm, and the image thickness was 1 to 2 mm. The CT images were reviewed on the Synapse picture archiving and communication system (Fujifilm USA, Valhalla, NY) on a Dell Desktop Optiplex GX620 computer (Dell Computers, Round Rock, Tex). Axial images were reviewed to obtain the maximal aortic

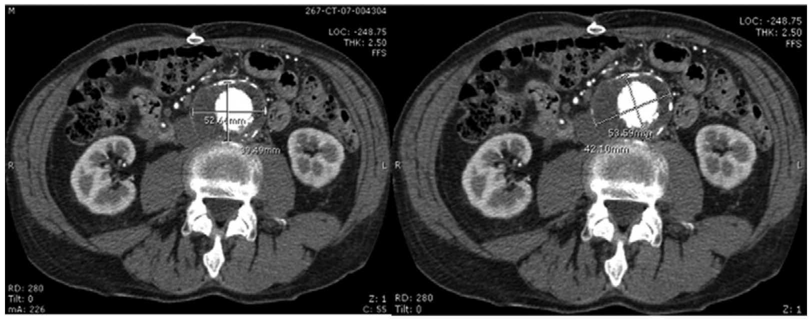

Fig 2. The computed tomography-determined **(left)** anteroposterior and transverse maximal aortic diameter and **(right)** the major and minor axis aortic diameter are shown.

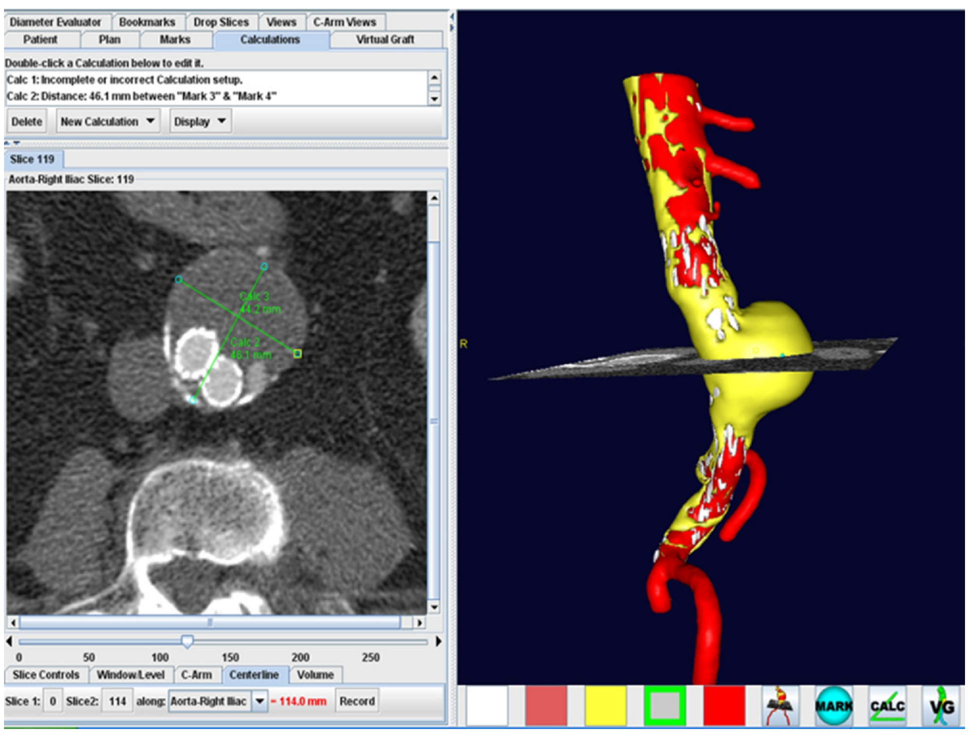

Fig 3. Major and minor aortic diameters are shown as determined by centerline M2S reconstruction.

diameter. All images were reviewed by the same vascular surgeon (K.P.), who used on-screen calipers to measure AP, transverse, and major and minor axis aortic diameters from adventitia to adventitia (Fig 2). Diameter measurements were entered into an Excel spreadsheet.

The spiral CT data were processed by Medical Images & Data Management Services (Evanston, Ill) to generate a 3-D model of the aorta. The 3-D images were reviewed by the same vascular surgeon, who was blinded to the patient name and measurements from US and axial CT scan. The M2S Preview software (West Lebanon, NH) provided a platform for orthogonal diameter measurements on a 3-D aorta model (Fig 3). The entire length of the infrarenal aorta was assessed for the segment with the maximal aneurysm size. At this segment, the AP, transverse, and major and minor axes diameters were measured from adventitia to adventitia and were entered into an Excel spreadsheet.

Statistical analysis. On the US measurements**,** the major axis was defined as the larger of the AP and transverse diameter measurements, and the minor axis was defined the smaller of the two. To assess correlation between different measurement modalities, linear regression analysis was performed between the centerline major axis measurements in the 3-D CT reconstructions and the US major and minor axes measurements. Similarly, linear regression was performed between the CT minor axis and the centerline major axis measurements in the 3-D reconstructions.

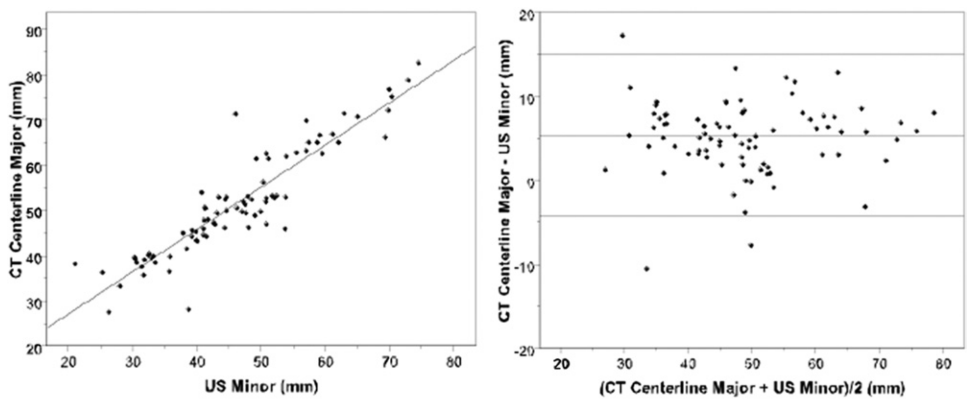

Fig 4. Left, A linear correlation is shown between computed tomography *(CT)* centerline images and the minor axis measured on ultrasound *(US)* imaging. **Right,** A Bland-Altman plot shows US minor and CT centerline diameter measurements.

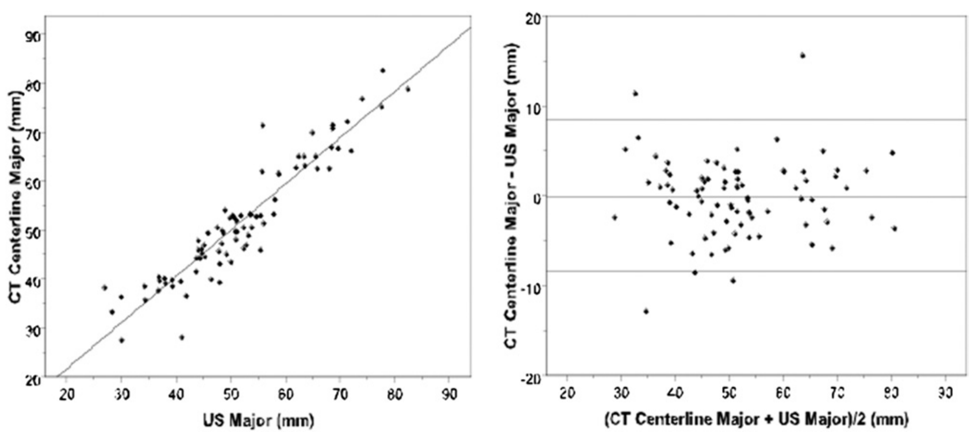

Fig 5. Left, A linear correlation is shown between the computed tomography *(CT)* centerline diameter and the ultrasound *(US)* major axis diameter. **Right,** A Bland-Altman plot shows US major and CT centerline diameter measurements.

Bland-Altman plots were performed between each method to assess the agreement between measurement modalities. In Bland-Altman analysis, the difference between two measurement methods is plotted against their average for each patient.[16] Calculating the mean difference and the standard deviation (SD) of difference allows one to quantitatively assess how close the measurements from two different methods are to each other and how scattered they are collectively. Mean values are presented with SD. Two-tailed 95% confidence intervals (CIs) were calculated for each mean SD difference value. Statistical significance was noted at $P = .05$.

To eliminate interobserver variability, a single vascular surgeon (K.P.) performed all CT and CT centerline measurements and reviewed all US measurements. To avoid observer bias, the US technologists were blinded to the CT measurements, and the vascular surgeon was blinded to the patient name and took measurements in random order. After excluding technically difficult US studies, 82 patients were analyzed.

RESULTS

CT centerline major vs US minor. CT centerline major (51.0 mm [SD, 11.97mm]) was larger than US minor (46.3 mm [SD, 11.7 mm]) measurement. The correlation between CT centerline major and US minor was analyzed by linear regression (Fig 4), and there was a direct relationship ($R^2 = 0.92$). In the Bland-Altman plot, the mean difference between the two measurements was 5.38 mm (SD, 4.84 mm; Fig 4). Therefore, US minor measurements deviated from CT centerline major measurements by average of 5.38 mm, which indicated a substantial disagreement. There was a significant variability in measurements suggested by the SD difference of 4.84 mm (95% CI of difference, 0.33 to 6.43 mm; $P = .05$).

CT centerline major vs US major. The CT centerline major (51.0 mm [SD, 11.97 mm]) was slightly smaller than the US major (51.7 mm [SD, 11.8 mm]). The correlation between CT centerline major and US major was

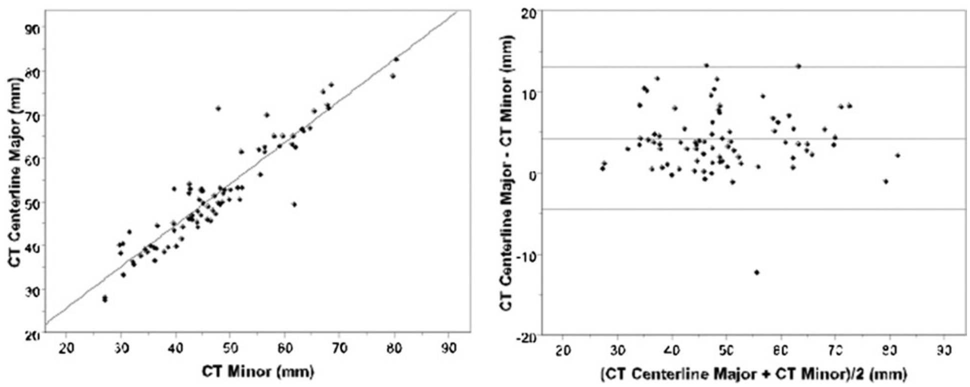

Fig 6. Left, A linear correlation is shown between computed tomography *(CT)* centerline diameter and minor axis diameter measurements. **Right,** A Bland-Altman plot is shown of CT centerline diameter and minor axis diameter measurements.

analyzed by linear regression (Fig 5), and there was a direct relationship (R^2 = 0.94). In the Bland-Altman plot, the mean difference between the two measurements was 0.11 mm (SD of difference, 4.26 mm; Fig 5). Therefore, US major measurements deviated from the CT centerline major measurements by only 0.11 mm on average. This suggested a close agreement, but there was a significant variability in measurements indicated by the SD of difference of 4.26 mm (95% CI of difference, –1.03 to 0.81 mm; P = .05).

CT centerline major vs axial CT minor. CT centerline major (51.0 mm [SD, 11.97 mm]) was larger than axial CT minor (47.4 mm [SD, 11.7 mm]). The correlation between CT centerline major and axial minor was analyzed by linear regression (Fig 6), and there was a direct relationship (R^2 = 0.93). In the Bland-Altman plot, the mean difference between the two measurements was 4.25 mm (SD of difference, 4.37 mm; Fig 6). Therefore, axial CT minor measurements deviated from CT centerline major measurements by 4.25 mm on average. This suggested a substantial disagreement, but there was a significant variability in measurements indicated by the SD of difference of 4.25 mm (95% CI of difference, 3.30-5.19 mm; P = .05).

DISCUSSION

Accurate measurement of the AAA diameter is of paramount importance for determination of surgical indication and for postoperative surveillance. Historically, however, the AAA diameter assessment has been inconsistent. Even the landmark studies such as the UK Small Aneurysm Trial and the MASS trials used US imaging differently to measure the diameter. This heterogenous nature in AAA diameter measurements should be considered in interpreting the results of the landmark trials and applying them to clinical practice.[12,13]

Numerous studies have compared US and axial CT-determined AAA diameter measurements, pointing out the systematic differences between them. Only a few of those studies incorporated 3-D reconstruction of the CT images, which is now considered the most accurate method to measure aortic diameters.[9,11] Furthermore, most of the previous studies focused on evaluating the accuracy of US imaging as a screening tool for determining indications for surgery. Although CT scan with 3-D reconstruction is unlikely to be replaced by US imaging in the preoperative evaluation of a patient, one can argue that US imaging has a greater potential in its clinical effect in postoperative surveillance. Therefore, we directly evaluated the measurement modalities by focusing on patients who were undergoing post-EVAR surveillance scans.

Table. Correlation coefficients (R^2) and differences among major and minor computed tomography *(CT)* and ultrasound *(US)* measurements with respect to 3-D CT centerline measurements

Variable	*US major*	*US minor*	*CT minor*
R^2	0.94	0.92	0.93
Mean difference, mm	−0.11	5.38	4.25
SD of difference, mm	4.26	4.84	4.37

SD, Standard deviation.

Both major and minor axes US measurements showed a good correlation to the CT centerline measurements, with correlation coefficients ≥0.92 (Table). This range of correlation coefficient (0.89 to 0.93) was similar to ranges in previous studies evaluating US imaging and CT scan in maximal aortic diameter measurements.[1,11] In addition, the mean differences and SDs of difference between measurements observed in our study were similar to the 3 mm (range, 4.29-5.3 mm) obtained in a study by d'Audiffret et al.[7] Although d'Audiffret used transverse diameter measurements for both CT and US imaging, our study assessed specific methods of US measurements by evaluating major and minor axes diameters. CT scans have been reported to overestimate the true aortic diameter because of oblique slicing of the CT scan in an angulated aorta.[1,7,8,13] Therefore, the minor axis on CT scan is considered the more accurate measurement of the true aortic diameter.[6,11,14] Furthermore, advent of 3-D reconstruction of CT scans allows digital correction of aortic angulation, thus providing the diameter measurement perpendicular to the aortic centerline.[10,11,14] Our study incorporated these considerations by comparing minor axis measurements in conventional CT scans and centerline major axis measurements in 3-D reconstruction CT images.

Previous studies comparing US- and CT-determined aortic diameter measurements found that CT consistently resulted in larger measurements than US imaging.[1,7-9,11] This difference may be partially due to the overestimation of aortic diameter by CT scans due to angulation, and US imaging is thought to correct for aortic angulation by the technologists maintaining the probe perpendicular to blood flow.[9,11] This is consistent with the previous observation that the difference between CT- and US-derived diameter measurements increases with larger aortic angulation.[9] Assuming complete correction of aortic angulation by US imaging, one would expect the major axis measurement in US images to be more accurate than the minor axis measurement. To our knowledge, however, no previous study has evaluated the accuracy of different US methods in aortic diameter measurements. Although d'Audiffret et al[1] used transverse measurements on US imaging, Manning et al[11] evaluated AP measurements, indicating the lack of a standardized US method in diameter assessment. Therefore, we investigated this issue by evaluating both minor and major axis US measurements in our study.

The correlation coefficient was the highest, at 0.94, between US major and CT centerline major measurements, followed by 0.92 between US minor and CT centerline major measurements. The high correlation coefficient between the US and CT centerline major measurement was similar to the 0.93 between CT minor and CT centerline major measurements. Therefore, compared with the axial CT minor measurements, which is the current standard surveillance at most vascular practices, US imaging appeared to have an equivalent correlation with the 3-D centerline measurements.

Although linear regression analyses demonstrate similar correlation of US and CT scans with respect to the CT centerline in post-EVAR aortic diameter measurements, it does not prove clinical interchangeability of US and CT imaging. To assess equivalence, we performed an agreement analysis as described by Bland and Altman.[16] The mean difference was the lowest at 0.11 mm between US major and CT centerline, compared with 5.38 mm between US minor and CT centerline, suggesting that US major axis measurements were in better agreement to CT centerline than US minor axis measurements. Similarly, the mean difference of 0.11 mm between US major and CT centerline measurements was much lower than the 4.25 mm between axial CT minor and CT centerline measurements. Therefore, in the setting of equivalent correlation coefficients, Bland-Altman analyses suggest that the US major axis measurement is the best approximation of the aortic diameter as determined by CT centerline major measurements.

There was high variability, however, measured by the SD of difference ranging from 3.98 to 4.84 mm between US and CT centerline. High variability was also observed between conventional axial CT measurements and centerline CT measurements (SD of difference, 4.36 mm). Despite our effort to eliminate interobserver variability by using a single vascular surgeon for all diameter measurements, there appears to be some variability in interpretation of CT images as well as centerline CT images, as suggested by the relatively high SDs. However, the level of variability seen in our study was lower than the SDs of differences of 4.9 to 8.0 mm reported by Manning et al.[11] Furthermore, the variability seen in our study should be interpreted in the context of the optimal accuracy that one can expect from axial CT scan, the current gold standard.

Although no study to our knowledge has directly evaluated the variability of aortic diameter measurements in post-EVAR patients, studies have reported a significant interobserver and intraobserver variability in the interpretation of CT scan measurements in preoperative patients with AAA. In a multicenter Veterans Administration study, Lederle et al[8] reported that an intraobserver and interobserver difference of ≥5 mm in CT measurements was common.[8] Cayne et al[6] reported the SD of difference of 4.4 to 6.3 mm, equivalent to our largest SD of difference, 4.48 mm. Our findings are also consistent with the previous reports by Jaakkola et al,[7] which showed similar interobserver variability in AP and transverse diameter measurements taken from US images and CT scans.

Therefore, the level of variability that we observed in US aortic diameter measurements appears to be equivalent to that of axial CT scan, the current gold standard. Specifically, interoperator variability among the four US technologists is partially responsible for the SDs of differences seen in our US measurements. Nevertheless, the equivalent level of variability between US images and CT scans in our study indicates that, by instituting a standardized US protocol, one can minimize interoperator variability. Furthermore, the 95% CIs of differences (P = .05) indicate that the difference between US major and CT centerline measurements lies well within the clinically acceptable limits of ±2 mm, as defined by Jaakkola et al[7] and Lederle et al[8] (Fig 7). Therefore, the observed level of variability as reflected by the SD of difference did not negate the statistical significance of the agreement between US major and CT centerline measurements.

Our study has some limitations. First, this was a single-center, cross-sectional study with patients at different stages of follow-up after EVAR. Therefore, no direct information on the evolution of aneurysm size and the ability of US imaging and CT scans to detect the changes in diameter over time was available. Second, a single observer reviewed all images and determined every measurement. Third, endoleak detection, an important component of post-EVAR surveillance, was not evaluated.

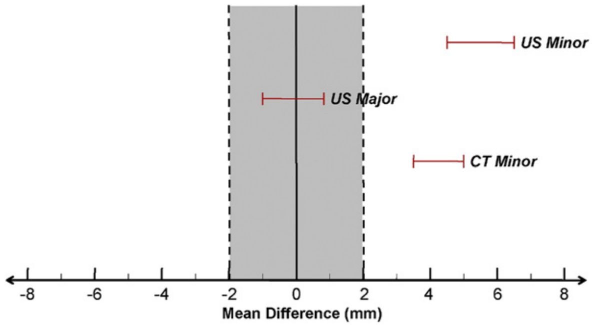

Fig 7. The 95% confidence intervals (P = .05) are shown for mean differences obtained from Bland-Altman analyses.

CONCLUSIONS

Despite these limitations, our findings illustrate the accuracy of aortic aneurysm diameter measurements in post-EVAR US surveillance, when performed by certified vascular technologists in an accredited vascular laboratory and according to a carefully devised and standardized US protocol, as described in Methods. Although the high variability suggests that one cannot use US imaging and conventional CT measurements interchangeably, consistent use of an US protocol will likely provide aortic diameter measurements that are of equal or superior reliability to the current CT measurements in post-EVAR patients. Specifically, the major axis measurement by US imaging appears to be the more accurate estimation of the aortic diameter, as reflected by the 3-D CT measurements. Further investigations on standardization of US surveillance protocol may demonstrate improved accuracy of US imaging.

AUTHOR CONTRIBUTIONS

Conception and design: FW
Analysis and interpretation: SH, KP, VR, FW
Data collection: SH, KP, SP
Writing the article: SH, FW
Critical revision of the article: SH, KP, VR, FW
Final approval of the article: FW
Statistical analysis: SH, KP, AB
Obtained funding: FW
Overall responsibility: FW

REFERENCES

1. d'Audiffret A, Desgranges P, Kobeiter H, Becquemin JP. Follow-up evaluation of endoluminally treated abdominal aortic aneurysms with duplex ultrasonography: Validation with computed tomography. J Vasc Surg 2001;33:42-50.
2. Meier GH, Parker FM, Godziachvili V, Demasi RJ, Parent FN, Gayle RG. Endotension after endovascular aneurysm repair: the Ancure experience. J Vasc Surg 2001;34:241-7.
3. Torsello GB, Klenk E, Kasprzak B, Umscheid T. Rupture of abdominal aortic aneurysm previously treated by endovascular stentgraft. J Vasc Surg 1998;28:184-7.
4. Balm R, Jacob MJ. Use of spiral computed tomographic angiography in monitoring abdominal arotic aneurysms after transfemoral endovascular repair. Tex Heart Inst J 1997;24:200-3.
5. Aarts NJ, Schurink GW, Schultze Kool LJ, Bode PJ, van Baalen JM, Hermans J, et al. Abdoimnal aortic aneurysm measurements for endovascular repair: intra- and interobserver variability of CT measurements. Eur J Vasc Endovasc Surg 1999;18:475-80.
6. Cayne NS, Veith FJ, Lipsitz EC, Ohki T, Mehta M, Gargiulo N, et al. Variability of maximal aortic aneurysm diameter measurements on CT scan: significance and methods to minimize. J Vasc Surg 2004;39:811-5.
7. Jaakkola P, Hippelainen M, Farin P, Rytkonen H, Kainulainen S, Partanen K. Interobserver variability in measuring the dimensions of the abdominal aorta: comparison of ultrasound and computed tomography. Eur J Vasc Endovasc Surg 1996;12:230-7.
8. Lederle FA, Wilson SE, Johnson GR Reinke DB, Littooy FN, Acher CW, et al. Variability in measurements of abdominal aortic aneurysms. J Vasc Surg 1995;21:945-52.
9. Sprouse LR, Meier GH, Parent FN, DeMasi RJ, Glickman MH, Barber GA. Is ultrasound more accurate than axial computed tomography for determination of maximal abdominal aortic aneurysm diameter? Eur J Vasc Endovasc Surg 2004;28:28-35.
10. Sprouse LR, Meier GH, Parent FN, DeMasi RJ, Stokes GK, LeSar CJ, et al. Is three-dimensional computed tomography reconstruction justified before endovascular aortic aneurysm repair? J Vasc Surg 2004;40:443-7.
11. Manning BJ, Kristmundsson T, Sonesson B, Resch T. Abdominal aortic aneurysm diameter: a comparison of ultrasound measurements with those from standard and three-dimensional computed tomography reconstruction. J Vasc Surg 2009;50:263-8.
12. Ashton HA, Buxton MJ, Day NE, Kim LG, Marteau TM, Scott RA, et al. The Multicentre Aneurysm Screening Study (MASS) into the effect of abdominal aortic aneurysm screening on mortality in men: a randomized controlled trial. Lancet 2002;360:1531-9.
13. The U.K. Small Aneurysm Trial: design, methods and progress. The UK Small Aneurysm Trial participants. Eur J Vasc Endovasc Surg 1995;9:42-8.
14. Chaikof EL, Blankensteijn JD, Harris PL, White GH, Zarins CK, Bernhard VM, et al. Reporting standards for endovascular aortic aneurysm repair. J Vasc Surg 2002;35:1048-60.
15. McPharlin M, Rumwell C. Vascular technology: an illustrated review. Pasadena, CA: Davies Publishing; 2009.
16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;2:307-10.

Submitted Sep 22, 2009; accepted Jan 14, 2010.

DISCUSSION

Dr Niren Angle *(San Diego, Calif)*. Members and invited guests of the Western Vascular Society. It is my privilege to be invited to discuss a paper at this meeting and, moreover, to discuss a very interesting paper such as the one that was just presented. The USC [University of Southern California] group has put forth the proposition that ultrasound measurement of aortic aneurysm diameter is best correlated with centerline measurement of diameter on 3-D [three-dimensional] CT [computed tomography] reconstruction. In doing so, they have offered up a thesis, which at least intuitively, creates cognitive dissonance in my troubled mind but is certainly supported by the data that they present.

The authors stipulate, and I think correctly, that measurement of aneurysm diameter on axial CT is subject to much variability and error depending on the fastidiousness of the examiner and the plane in which the aorta is measured (ie, elliptical vs straight line). In this regard, centerline measurements are probably more accurate, as has been shown, but even that current belief is likely to be tempered in the future. That notwithstanding, the authors then compared ultrasound, standard axial CT, and M2S centerline measurements of aortic diameter to determine their correlation and also the variability of these measurements.

To examine this, two methods were used—linear regression and correlation, and also Bland-Altman plots. It is very useful that both these measures were used because there are limits to each method.

Correlation coefficient is a measure of the degree of association between two quantities; it does not measure how closely they agree, its use in comparing two methods that purport to measure the same thing is inappropriate, and quoting *P* values in such circumstances is meaningless. The Bland-Altman plot, which is so commonly used now in medical literature for comparative statistics, with over 10,000 citations in papers, provides a quantitative measure of how close the measurements are to each other.

The conclusion here appears to be that: (1) Using centerline major as the standard to measure by, measurement of the aortic diameter by ultrasound in the major axis is the best correlated with the least mean difference. (2) Measurement of aortic diameter in the minor axis on ultrasound had good correlation, but the mean difference in measurements was 5.38 mm in the Bland-Altman plot. (3) Axial CT measurements in the minor axis also had a mean difference of 4.25 mm.

Ultrasound is a very operator-dependent technique, and as such, it is difficult to understand that one can hope to obtain this degree of precision vs a CT scan, where although there may be variability in the measurement of the image cuts, ultrasound may have variability in the images obtained as well as the measurement. Also, the quality of the image and resolution is also dependent on many factors.

To what degree are these data only reliable at USC because of vascular technologists that have presumably honed the art of aortic ultrasound? I suspect that the interobserver variability, if the universe of ultrasound examiners was expanded beyond the ones in this study, would be much much larger than with CT scan assessment.

In reference to the Bland-Altman plots, the mean difference between CT centerline vs CT minor was 4.25 mm, whereas CT centerline vs US minor was 5.38 mm. The difference between CT centerline and ultrasound major was 0.11 mm. Does this mean that the axial CT is more accurate than is ultrasound measurement in the minor axis?

The minor axis measurement on CT is more accurate than major axis measurement and this is accepted. Why does the same principle not hold true for ultrasound major versus minor axis?

How much faith is one to have that the ultrasound technologist is reliably able to image the aorta perpendicular to blood flow, as this appears to be the requisite condition for this assessment to be valid? How does one confirm this?

In the comparison of centerline flow to ultrasound in the major axis, the mean difference between 2 measurements was 0.11 mm but the standard deviation collectively was 4.26 mm. Can you comment on this? Does the agreement between 2 measurements become less precise with different sizes of the aorta? Because if the mean difference between measurements is 0.11 mm, and this was maintained over a range of samples, why is the standard deviation so much larger, in contrast to the other comparisons?

I want to congratulate Dr Han and his colleagues and Dr Weaver for a very thought-provoking paper and for a very well written manuscript. I drank a few martinis learning the subtleties of the Bland-Altman analysis, and I feel like I know them personally. All I can say is that I am glad I am a vascular surgeon. I must say, however, unless the USC group gets a big RV to drive their ultrasound technologists all over Southern California to do these beautiful studies, it will be hard for me to be pulled away from the CT scanner as my test of choice for pre- and post-EVAR. Thank you again for the privilege and my congratulations to the authors.

Dr Sukgu M. Han. To what degree are these data only reliable at USC because of vascular technologists that have presumably honed the art of aortic ultrasound? I suspect that the interobserver variability, if the universe of ultrasound examiners was expanded beyond the ones in this study, would be much much larger than with CT scan assessment.

I think that is always a concern when it comes to any measurements that are done on US. I think if one were to pool all the US technologists in the world together, the variability would indeed be much larger, but that variability would probably come from the heterogenous nature of individual institutional protocols.

Our study and previous institutional studies by others, including Drs Sprouse, Jaakkola, and Lederle papers published in *JVS*, have shown comparable degree of variability within US measurements to axial CT scans. Therefore, this demonstrates that establishing a strict US protocol and properly training US technologists ***can*** minimize the variability all the way down to that of CT scan.

In reference to the Bland-Altman plots, the mean difference between CT centerline vs CT minor was 4.25 mm whereas CT centerline vs US minor was 5.38 mm. The difference between CT centerline and ultrasound major was 0.11 mm. Does this mean that the axial CT is more accurate than is ultrasound measurement in the minor axis?

We cannot safely say that because of the degree of variability and closeness of the mean differences. Another way to look at it is that the 95% CI [confidence interval] for axial CT and US minor measurements overlapped, so we cannot say that one is better than the other.

The minor axis measurement on CT is more accurate than major axis measurement and this is accepted. Why does the same principle not hold true for ultrasound major versus minor axis?

Well, that was the question we asked ourselves. No studies have really compared major vs minor axis measurements on US. Intuitively, we thought that if our US protocol can correct for angulation error, then we should be trusting the major axis measurements the same way that we would for the CT centerline, and our data strongly support that.

How much faith is one to have that the ultrasound technologist is reliably able to image the aorta perpendicular to blood flow, as this appears to be the requisite condition for this assessment to be valid? How does one confirm this?

After our data, we have a lot of faith in our US technologists at USC. There are two ways to ensure that this happens. First, when the aorta is being surveyed with B mode, the technologists examine the longitudinal images in order to get an idea of regions of dilatation and angulation. Second, we use color flow while the probe is being oriented. Because of pulsatile nature of aortic flow, when the probe is oriented perpendicular to it, we see a mix of alternating red and blue, compared to when it is oriented obliquely, we see mainly one predominant color. This, particularly second maneuver, is not always done, for example in radiology department US.

In the comparison of centerline flow to ultrasound in the major axis, the mean difference between 2 measurements was 0.11 mm, but the standard deviation collectively was 4.26 mm. Can you comment on this? Does the agreement between 2 measurements become less precise with different sizes of the aorta? Because if the mean difference between measurements is 0.11 mm, and this was maintained over a range of samples, why is the standard deviation so much larger, in contrast to the other comparisons?

That was our concern as well when we looked at this relatively large variability, and that is why we performed 95% CI of mean differences from our gold standard. CI demonstrates that in comparison, the agreement seen between US major and CT centerline, and the lack thereof for US minor and axial CT minor, was a statistically significant observation. To answer your second question regarding the effect of increasing aortic size on agreement, we did not perform graded Bland-Altman plot, so I don't have a precise answer for you, but looking at the plot, it appears that the difference dots are scattered evenly with increasing size of the aorta.