Variation in Maximum Diameter Measurements of Descending Thoracic Aortic Aneurysms Using Unformatted Planes versus Images Corrected to Aortic Centerline

N. Rudarakanchana a,b,*, C.D. Bicknella,b, N.J. Cheshire a,b, N. Burfitt c, A. Chapman c, M. Hamady a,c, J.T. Powell d

a Department of Surgery and Cancer, Imperial College, London, UK
b Imperial Vascular Unit, Imperial College Healthcare NHS Trust, London, UK
c Department of Interventional Radiology, Imperial College Healthcare NHS Trust, London, UK

WHAT THIS PAPER ADDS
Current guidelines for intervention for descending thoracic aortic aneurysm (dTAA) do not state specific methods or anatomical planes for measurement of threshold diameters. Standardized measurement protocols are required in order to synthesize the accumulating evidence for the use of thoracic endovascular repair and best patient management. This study shows that for dTAA corrected diameter measurements (perpendicular to aortic centerline after image-processing) on CT scans are on average smaller, have higher repeatability, and are subject to less interobserver variability than axial diameter measurements. Therefore, these corrected external diameters should be adopted as the reporting standard for dTAA measurement in the future.

Objective: Evaluation of variation in descending thoracic aortic aneurysm (dTAA) diameters measured on CT scans in different planes and by different observers and the potential impact on treatment decisions.

Methods: CT angiography of dTAA (N = 20) were assessed by three specialists, with measurements repeated after 1 month. Calliper measurements of maximum external diameters were made on unformatted images and perpendicular to the aneurysm centerline after image processing (corrected). Repeatability was assessed using Bland–Altman plots.

Results: Maximum corrected diameter measurements were smaller than axial measurements (66.3 ± 7.9 mm vs. 74.9 ± 20.9 mm, p < .001). Both intraobserver and interobserver variation were less for corrected than for axial measurements (mean intraobserver differences 5.0 ± 3.8 mm vs. 11.8 ± 9.3 mm, p < .001; mean interobserver differences 2.8 ± 2.5 mm versus 10.4 ± 14.0 mm, p < .001) and interobserver variation increased with aneurysm diameter for maximum axial but not corrected measurements. Using corrected rather than axial measurements could have changed treatment decisions in two patients (10%) using a treatment threshold diameter of 55 mm and 10 patients (50%) using a threshold of 65 mm.

Conclusion: Corrected diameters were smaller than axial diameters, could be measured with higher repeatability, and were subject to less interobserver variability. Using corrected versus axial measurements would have changed management decisions in up to half of the cases in this study.

© 2013 European Society for Vascular Surgery. Published by Elsevier Ltd. All rights reserved.

Article history: Received 26 July 2013, Accepted 22 September 2013, Available online 3 October 2013

Keywords: Thoracic aortic aneurysm, CT scan, Dimensional measurement accuracy, Endovascular procedure

INTRODUCTION
The incidence of descending thoracic aortic aneurysm (dTAA) is increasing: epidemiology data from Sweden report a rise of 52% in men and 28% in women over the 15 years to 2002, to 9–16 per 100,000 population per year.1 Whilst open surgical repair of dTAA continues to be associated with high mortality and morbidity rates, the development of thoracic endovascular stent grafts and associated improvements in operative survival has resulted in an exponential increase in thoracic endovascular aneurysm repair (TEVAR), with a particularly marked increase in those 75 years and older.2 However, despite this increase in numbers of patients undergoing operations, natural history data concerning the risk of aneurysm rupture and the evidence base for threshold diameters at which dTAA repair become beneficial are weak and limited to single-centre series.3 The definition and measurement of maximum diameter are not reported in detail and have not been standardized. Perhaps unsurprisingly therefore, guidelines for treatment of dTAA vary in recommended threshold diameters, from 5.5 to 7.0 cm.3-8
Although treatment decisions for aneurysms are based on a number of factors, including aneurysm morphology and rate of growth, patient age, gender, anaesthetic risk, functional status, and comorbidities, maximum aortic diameter is the key parameter used to predict rupture risk and therefore is central in directing clinicians whether to offer surveillance or surgical repair and in the formulation of guidelines. Despite this, the exact definition of maximum diameter and measurement methods are not clearly described or uniformly used, even for relatively more common abdominal aortic aneurysm. Computed tomography (CT) is the gold standard imaging modality for the thoracic aorta and current image-processing software enables sophisticated three-dimensional and multiplanar reconstructions (MPRs). Diameters may be measured on unformatted images in a single plane (axial, sagittal, or coronal) or on processed images following manual adjustment of planes perpendicular to the centreline of the aorta. Precise measurement methods for dTAA should specify: (a) plane of acquisition, (b) axis of measurement, (c) position of callipers, internal or external, (d) selected diameter. Given that the thoracic aorta may be highly tortuous, using different planes for measurement of dTAA may be expected to produce variable results.

Previous studies have investigated the variability in measurement of dTAA diameters in patients with connective tissue disease and by technologists or non-experts using semi-automated analysis techniques. A recent study has reported the variability in measurement of proximal and distal landing zone diameters in patients with thoracic aortic disease (dTAA or penetrating aortic ulcer). However, no study to date has reported the variability in measurement of dTAA diameters by expert vascular radiologists. In this study, we assess (a) the intraobserver and interobserver repeatability of three expert vascular radiologists in measuring, according to a defined protocol, (b) dTAA diameters in different planes, and (c) how such variation might impact treatment decisions according to current guidelines.

MATERIALS AND METHODS

Subjects

The National Research Ethics Service (UK) confirmed that ethical approval was not required for this study. The most recently managed 20 patients with fusiform degenerative dTAA and CT scans, with reported diameter ranges from 50 to 130 mm, were identified from a prospectively collected database of patients with dTAA at St Mary’s Hospital, Imperial College Healthcare National Health Service Trust, London, UK.

Image acquisition

Imaging was performed using the Brilliance iCT 256-slice scanner (Philips Healthcare, Best, The Netherlands) in accordance with standard clinical protocol with the following parameters: 200 mA, 120 kEV, 1 mm slice thickness, matrix 1024 x 1024. The field of view was adjusted according to patient body habitus and image acquisition was performed using non-electrocardiographic gated algorithms. Intravenous iodinated contrast was administered in conjunction with image acquisition and image series timed for opacification of the aorta.

Descending TAA measurements

The anonymized CT scans were analysed by experienced specialists in vascular interventional radiology (M.H., N.B., A.C.), all holders of the Fellowship of the Royal College of Radiologists. Measurements were performed on a single dedicated CT workstation (Extended Brilliance Workspace, V3.5.0.2254, Philips Medical Systems). Experts were blinded to all clinical data and previously reported measurements. Each expert independently assessed the CT scans in random order following the protocol established for this study. For each scan, experts recorded four measurements using workstation software, with callipers from outer wall to outer wall of the aneurysm: maximal diameters in axial, sagittal, and coronal cross-sections (without attempting to correct for aortic tortuosity), and maximal diameter in any plane following double-oblique MPRs with manual correction to aortic centerline (Fig. 1). All measurements were rounded to the nearest millimetre.

Intraobserver and interobserver variability

To assess repeatability, each expert repeated all measurements on all CT scans, including creation of new MPR images, after an interval of at least 1 month, again with scans in random order and observers blinded to previous measurements. For each method of measuring dTAA diameter, intraobserver variation was determined by calculating the difference between the repeated diameter measurements taken by each expert for each scan (measurement A — measurement B) and deriving the mean and standard deviation of these differences. Bland—Altman plots were used to assess repeatability and repeatability coefficients (the value within which the absolute test—retest difference would be expected to lie 95% of the time) were calculated for each expert’s measurements. For each method of measuring dTAA diameter, the difference in diameter measurements between each expert and every other expert (expert 1 — expert 2, expert 1 — expert 3, expert 2 — expert 3) for each scan was calculated. The mean and standard deviation of these differences were used to determine the interobserver variability.

Timed measurements

To investigate time taken to obtain different measurement diameters, a different specialist in interventional radiology (A.Al-S.) performed timed maximum external diameter measurements on 10 randomly selected CT scans from the same cohort, in axial cross-section and following double-oblique MPRs with manual correction to centerline. Measurement time in seconds was assessed using a stopwatch,
with timings beginning once the CT scan was uploaded (therefore including preparation time for MPR images).

**Impact on treatment decisions**

To examine the potential impact of different measurement methods on clinical decisions for treatment of dTAA, current guidelines for threshold diameters were applied using each measurement method and the number of patients meeting criteria for surgical repair by each method compared.

**Statistical analysis**

Data were analysed with the Statistical Package for Social Sciences version 20 (IBM SPSS, Chicago, IL, USA). Data were explored for normality using histograms and Kolmogorov–Smirnov tests. Repeatability was assessed using Bland–Altman plots, with statistical significance calculated using repeated measures ANOVA, Mann–Whitney U, Wilcoxon signed rank, and paired Student t tests. A p value < .05 was considered to indicate statistical significance.

**RESULTS**

*Diameter measurement in different aortic planes*

The reporting of coronal, sagittal, axial, and corrected maximum external diameters by a single observer is shown in Fig. 2. Measurements made by this observer in the three unformatted planes (axial, coronal, and sagittal) were not significantly different (axial mean diameter of 81.6 ± 19.8 mm, a sagittal mean 86.3 ± 16.8 mm, and a coronal mean 90.0 ± 17.6 mm: repeated measures ANOVA p = .370). However, on average, the maximum external diameters taken on corrected images were smaller than those measured in the unformatted planes: expert 1 recorded a corrected mean diameter of 68.1 ± 8.3 mm (p = .001). The differences between maximum external axial diameter and corrected external diameter for all three experts are shown in Fig. 3. This shows that maximum axial diameters are larger than maximum corrected diameters (axial mean diameter 74.9 ± 20.9 mm vs. corrected mean diameter 66.3 ± 7.9, p < .001) and that this difference increases with increasing aneurysm diameter. For example, for aneurysm diameters in the range 50–74 mm, the mean difference between maximum axial and corrected measurements was 3.2 ± 5.2 mm compared with a difference of 27.8 ± 14.6 mm for aneurysms in the range 75–130 mm.

Further assessment of interobserver and intraobserver variability focused on axial (as the most commonly viewed CT format) and corrected diameters only.

*Intraobserver and interobserver variation*

Intraobserver variation was significantly greater for maximum axial diameters than corrected diameters (mean intraobserver differences 11.8 ± 9.3 mm vs. 5.0 ± 3.8 mm,
Fig. 4 shows a Bland–Altman plot for (a) maximum axial diameter measurements and (b) maximum corrected diameter measurements by the three assessors. There was better repeatability for maximum corrected diameter measurements than with maximum axial measurements (repeatability coefficients for corrected diameters were 11.9, 10.9, and 13.2 and for maximum axial diameters 25.1, 33.9, 26.0, for the three assessors).

Interobserver variation was significantly greater for maximum axial diameters than corrected diameters (mean interobserver differences 10.4 \( \pm \) 14.0 mm vs. 2.8 \( \pm \) 2.5 mm, \( p < .001 \)). Fig. 5 shows the mean difference in measured diameters between observers plotted against the mean measured dTAA diameter, for (a) maximal axial diameter measurements (Fig. 5A) and (b) maximum corrected diameter measurements (Fig. 5B). This shows that interobserver variation increased with increasing aneurysm diameter for maximum axial diameter measurements but not for corrected diameter measurements. For aneurysm diameters 50–74 mm the mean difference between experts’ maximum axial measurements was 6.1 \( \pm \) 5.5 mm and for aneurysms measuring for 75–130 mm it was 13.8 \( \pm \) 7.1 mm (\( p = .022 \)). The mean difference in experts’ maximum corrected measurements was 2.8 \( \pm \) 1.8 mm for aneurysm diameters 50–74 mm and for aneurysms measuring 75–130 mm it was 2.5 \( \pm \) 2.3 mm (\( p = .682 \)).

Timed measurements

Time required to obtain maximum diameter measurement was significantly faster for axial diameter measurements than corrected diameter measurements (15 \( \pm \) 2.4 s versus 31 \( \pm \) 6.0 s, \( p < .001 \)).

Potential impact of measurement variation on clinical decisions

Current guidelines for treatment of dTAA are summarized in Table 1. Using maximum corrected diameters rather than maximum axial measurements could have changed treatment decisions in two out of 20 patients (10%) using a treatment threshold diameter of 55 mm and 10 patients (50%) using a threshold diameter of 65 mm. Using the lower treatment threshold diameter (55 mm), in both patients using maximum corrected measurements rather than axial diameters would have changed the recommended treatment from TEVAR to clinical surveillance. Using the higher treatment threshold diameter (65 mm), in seven cases using maximum corrected measurements rather than axial diameters would have changed the recommended treatment from TEVAR to clinical surveillance, whilst in three cases management according to guidelines would have changed from clinical surveillance to TEVAR.
Aneurysm diameter is one of the most important factors in determining intervention for degenerative dTAA and is usually prognostic for overall patient outcome. However, current guidelines stating threshold diameter for intervention do not specify the exact method for diameter measurement in detail and whether the method used is reproducible. For example, internal diameters will be smaller than external diameters and in tortuous aortas axial diameters may be much greater than corrected diameters. This has particular importance now for three reasons. First, multiplanar CT reconstructions, with correction to centerline, are becoming routine. Second, the recent evolution in endovascular technology has extended treatment options for patients with dTAA. Third, the overall repair rate is rapidly increasing, driven by increasing numbers of TEVARs. Here we show that the repeatability, both intraobserver and interobserver, of dTAA diameter is far better for corrected diameters than for any of the unformatted planes. Moreover, corrected diameters are significantly smaller than the axial or any other of the unformatted diameters, even for the smaller aneurysms. Therefore, we propose that external corrected aortic diameters become the new reporting standard for dTAA.

DISCUSSION

Aneurysm diameter is one of the most important factors in determining intervention for degenerative dTAA and is usually prognostic for overall patient outcome. However, current guidelines stating threshold diameter for intervention do not specify the exact method for diameter measurement in detail and whether the method used is reproducible. For example, internal diameters will be smaller than external diameters and in tortuous aortas axial diameters may be much greater than corrected diameters. Therefore, we propose that external corrected aortic diameters become the new reporting standard for dTAA.

This has particular importance now for three reasons. First, multiplanar CT reconstructions, with correction to centerline, are becoming routine. Second, the recent evolution in endovascular technology has extended treatment options for patients with dTAA. Third, the overall repair rate is rapidly increasing, driven by increasing numbers of TEVARs. Here we show that the repeatability, both intraobserver and interobserver, of dTAA diameter is far better for corrected diameters than for any of the unformatted planes. Moreover, corrected diameters are significantly smaller than the axial or any other of the unformatted diameters, even for the smaller aneurysms. Therefore, we propose that external corrected aortic diameters become the new reporting standard for dTAA.

Previous studies have assessed the ability of technicians to measure dTAA diameters accurately and the ability of experts to measure aortic diameters in patients with Marfan or other genetic syndromes. Each of these studies reported significant measurement variability. A recent study has reported on measurement variability for selection of...
As may be expected, the interobserver variation between experts reported here for corrected diameter measurements is lower than that reported between experts in this study advised that external diameters were easier to define than internal diameters, our protocol focused on external diameter measurements only. Since the three experts in this study advised that external diameters were easier to define than internal diameters, our protocol focused on external diameter measurements only. We have not sought to define the relationship between tortuosity of the dTAA and measurement variation, nor the variation in diameters according to the use of different workstations, or electrocardiograph-gated versus non-gated CT scans, although these would be interesting questions for future study. We have not investigated the variability using automated correction of images to aortic centerline, a method available on current workstations. However, the accuracy of this method is significantly affected by the tortuosity of the descending aorta and manual correction is usually still required. Since the three experts in this study advised that external diameters were easier to define than internal diameters, our protocol focused on external diameter measurements only. We have not investigated the variability of internal diameters. Method protocols should specify whether internal or external diameters are measured: given that the likely thickness of the aneurysm wall is 2–3 mm, this may

Table 1. Guidelines for treatment threshold diameters for descending thoracic aneurysm.

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Threshold diameter (cm)</th>
<th>Measurement method</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coady et al.</td>
<td>1999</td>
<td>6.5</td>
<td>Not specified</td>
<td>Single-centre study</td>
</tr>
<tr>
<td>Svensson et al.</td>
<td>2008</td>
<td>5.5</td>
<td>Not specified</td>
<td>Expert consensus from Society of Thoracic Surgeons</td>
</tr>
<tr>
<td>Hiratzka et al.</td>
<td>2010</td>
<td>5.5</td>
<td>External maximal diameter perpendicular to axis of blood flow</td>
<td>American Heart Association guidelines</td>
</tr>
<tr>
<td>Rimbau et al.</td>
<td>2013</td>
<td>5.5</td>
<td>Not specified</td>
<td>European Society of Vascular Surgery guidelines</td>
</tr>
</tbody>
</table>

TEVAR devices but did not include measurements of maximum diameter. No study has reported on the repeatability of measurements made by experts for patients with degenerative data, where tortuosity is an important problem which is not usually seen in patients with connective tissue disorders.

This study confirms that the variation in dTAA diameter measurements by vascular radiologists on unformatted planes (axial, sagittal, or coronal) is wide, both between observers and for repeated measurements by the same person, on average varying 10–12 mm. This is despite these experts using a defined protocol for measurement (e.g. maximum external aneurysm diameter in specified planes). Such unformatted planes were probably used in the earlier single-centre series used to define intervention thresholds and the risk of aneurysm rupture. Measurements taken on formatted images using double-oblique MPRs were, on average, smaller than those taken on unformatted images and, importantly, were more reliable, with less variability between observers and for repeated measurements by the same person, on average 3 and 5 mm respectively. In addition, variation in corrected diameter measurements did not appear to increase with increasing aneurysm diameter. Perhaps unsurprisingly, it was faster to obtain axial diameter measurements than corrected diameter measurements, although this included the time required to prepare MPR images.

Parallel findings to those reported here have been published on variability of measurements of genetically triggered thoracic aortic aneurysms, with the authors also recommending manual double-oblique corrected diameters for aortic measurement. As may be expected, the interobserver variation between experts reported here for corrected diameter measurements is lower than that reported between non-experts for both semi-automated MPR and manual centerline measurement.

Interestingly, repeatability of semi-automated MPR measurements of non-aneurysmal thoracic aorta does not differ significantly from manual double-oblique corrected diameter measurements.

With increasing endovascular experience, the care of many patients with dTAA is likely to transfer from the cardiothoracic surgeon to the endovascular surgeon. As the vascular research community embarks on gathering evidence about the safe and appropriate use of TEVAR for dTAA, it becomes vitally important that measurement protocols for determining aneurysm diameter are standardized so that data can be properly and consistently compared both within trials and between trials and other studies.

The majority of the studies that currently guide treatment decisions do not state a clear protocol for measurement of aneurysm diameters. Therefore, it is possible that the variation in diameter measurements may in fact be even greater than that seen in this study, with consequent inconsistencies in patient management. In this study, up to a half of treatment decisions (recommending TEVAR or clinical surveillance only) would have varied depending on the specific method used for diameter measurement, either corrected MPRs or unformatted planar images.

Most of our knowledge about the natural history of dTAA and rupture risk is derived from reports from a landmark single-centre study, published in a series of papers by the Yale group beginning nearly a quarter of a century ago. This study reported a hinge point of 7 cm: below this rupture risk was calculated at 2–6%, above this rupture risk was said to be 28%. Again, no explicit protocol for aneurysm diameter measurement is given in these publications but presumably these were made on unformatted images and no time dimension is given to the rupture risk. Imaging capabilities have improved enormously since such studies commenced. Cardiovascular risk prevention strategies also have improved, with recent indications that statins improve patient prognosis. For all these reasons, the conclusions derived earlier may be unreliable when applied to 21st-century patients.

Limitations of this study include the small number of observers studied, the number of patient scans measured and use of a single workstation. However, this mimics practice in most centres and hence provides a reasonable estimate of measurement variability for dTAA diameters. We have not sought to define the relationship between tortuosity of the dTAA and measurement variation, nor the variation in diameters according to the use of different workstation software, or electrocardiograph-gated versus non-gated CT scans, although these would be interesting questions for future study. We have not investigated the variability using automated correction of images to aortic centerline, a method available on current workstations. However, the accuracy of this method is significantly affected by the tortuosity of the descending aorta and manual correction is usually still required. Since the three experts in this study advised that external diameters were easier to define than internal diameters, our protocol focused on external diameter measurements only. We have not investigated the variability of internal diameters. Method protocols should specify whether internal or external diameters are measured: given that the likely thickness of the aneurysm wall is 2–3 mm, this may
significantly affect the reported size. The protocol for this study required observers to measure diameters in unformatted axial, sagittal, and coronal images without attempting to correct to aortic centerline. Whilst such measurements may not be used in practice in many centres, the reported impact of different measurement methods on treatment decisions here does not seek to estimate reality, but rather to illustrate the degree of error possible in the absence of standardised measurement protocols for dTAA.

The importance of standardising aneurysm diameter measurement is vital for evidence synthesis. Although mortality following TEVAR for dTAA is now very low for good risk patients (30-day mortality 2–3%), the risk of stroke and paraplegia remains significant, with 1–5% of patients sustaining a perioperative neurological event and 2–3% suffering permanent neurological deficit whilst other issues such as renal replacement therapy are reported rarely. Whilst conservative management with clinical and imaging surveillance and best medical therapy may be advocated in light of this, this must be viewed in balance against very high mortality rates following ruptured dTAA and poor outcomes following emergency repair (reported perioperative mortality for TEVAR 19–28%). Precise calculation of rupture risk is therefore crucial to ensuring that patients with small dTAA do not undergo repair where the risks of rupture are lower than the risks involved in TEVAR. It is also important in making certain that patients are appropriately offered treatment when the risk of rupture is higher. Rupture risk prediction is currently based upon aneurysm diameter, probably from unformatted planes, and improved information is pivotal to providing evidence-based treatment and improving overall patient outcomes, whether managed with surveillance, TEVAR or even open surgical repair. Standardized protocols using corrected aneurysm diameters should be applied to predicting the risk of rupture.

Results of this study suggest that manual maximum external diameter measurements using double-oblique MPR images with correction to aortic centerline are more consistent than those using unformatted planar images. This measurement method should be used as the reporting standard for all future studies of dTAA and guidelines.

ACKNOWLEDGMENTS

We would like to thank Dr Ali Al-Safi for performing the timed measurements in this study and Dr Emanuela Falaschetti for advising on statistical analyses.

CONFLICT OF INTEREST

None.

FUNDING

This research was funded by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare National Health Service (NHS) Trust and Imperial College London. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, or the Department of Health.

REFERENCES


