Reproducibility of ECG-gated Ultrasound Diameter Assessment of Small Abdominal Aortic Aneurysms

K. Bredahl a,*, N. Eldrup b, C. Meyer c, J.E. Eiberg a, H. Sillesen a

a Department of Interventional Radiology, University Hospital of Copenhagen, Copenhagen, Denmark
b Department of Cardio-Thoracic and Vascular Surgery, Aarhus University Hospital, Denmark
c Cardoz, Copenhagen, Denmark

WHAT THIS PAPER ADDS
Ultrasound evaluation of abdominal aortic aneurysm is influenced by unclear vessel wall definitions and variations during the cardiac circle. We present an ECG-gated method with off-line reading to estimate the maximum aneurysm diameter, controlling for both vessel wall delineation and cardiac cycle variation. This ultrasound method was applied during a clinical trial testing a pharmaceutical compound’s potential to reduce aneurysm growth rate. This validation study shows that ECG-gated ultrasound recordings transferred and subsequently analysed at a core facility have very low inter- and intra-operator variability, potentially enabling the detection of even small growth rates.

Objective: No standardised ultrasound procedure to obtain reliable growth estimates for abdominal aortic aneurysms (AAA) is currently available. We investigated the feasibility and reproducibility of a novel approach controlling for a combination of vessel wall delineation and cardiac cycle variation.

Design: Prospective comparative study.

Methods: Consecutive patients (N = 27) with an AAA, attending their 6-month control as part of a medical treatment trial, were scanned twice by two ultrasound operators. Then, all ultrasound recordings were transferred to a core facility and analysed by a third person. The AAA diameter was determined in four different ways: from the leading edge of adventitia on the anterior wall to either the leading edge of the adventitia (method A) or leading edge of the intima (method B) on the posterior wall, with both measurements performed in systole and diastole.

Result: Inter-operator reproducibility was ±3 mm for all methods applied. There was no difference in outcome between methods A and B; likewise, end-diastolic measurement did not improve reproducibility in preference to peak-systolic measurement.

Conclusion: The use of a standardised ultrasound protocol including ECG-gating and subsequent off-line reading with minute calliper placement reduces variability. This may be of use in developing protocols to better detect even small AAA growth rates during clinical trials.

* Corresponding author. K. Bredahl, Department of Vascular Surgery, University Hospital of Copenhagen, Rigshospital-3111, Blegdamsvej 9, 2100 Copenhagen, Denmark. Tel.: +45 29 47 22 69; fax: +45 35 45 31 11.
E-mail address: kimbobank@yahoo.dk (K. Bredahl).
1078-5884/$ – see front matter © 2012 European Society for Vascular Surgery. Published by Elsevier Ltd. All rights reserved.
Article history: Received 15 September 2012, Accepted 10 December 2012, Available online 17 January 2013
Keywords: Aortic aneurysm, Abdominal, Ultrasonography, Imaging, Observer variation, Electrocardiogram

INTRODUCTION
Abdominal aortic aneurysms (AAAs) are common in older men (4% of men >65 years) and the natural history includes expansion and rupture. As few patients survive an acute incident of rupture (overall mortality 80–94%), prophylactic intervention is preferable in patients suitable for repair. Only patients with large AAAs (max. diameter >5.0–5.5 cm) have a natural risk of rupture exceeding the procedural risk, and for this reason patients with small asymptomatic aneurysms are handled conservatively with life-style changes, medical prevention and regular ultrasound surveillance. The key component in the current surveillance strategy for small AAAs is evaluation of the maximum aneurysm diameter, which today is considered the most important risk factor for rupture. A reliable method for repeated assessment of maximum AAA diameter is required, especially for decision-making in cases near the threshold for surgical referral (5–5.5 cm), for determining rapid growth accurately and for evaluating the effects of new treatments on aneurysm growth. With reported mean annual growth rates of 2–3 mm, a high reproducibility will be required to allow correct detection of small changes in AAA growth rates in individuals and reduce the risk of type II errors in clinical trials.
Computer tomography (CT) is generally perceived to be accurate and reproducible, but inter-operator variability is more than 5 mm in 17% of cases and the procedure includes exposure to radiation.\(^6\) The most commonly used imaging modality in screening for and follow-up for AAA is ultrasound, reflecting its non-invasive and harmless nature, cost-effectiveness, wide availability, and the historical fact that ultrasound was the imaging modality used when the threshold (5.5 cm) for intervention was determined.\(^7\) The exact methodology, however, for measuring aneurysm diameter as reported in the literature is diverse and often not fully specified, if mentioned at all.\(^8\) The reproducibility of ultrasound diameter assessment (anterior-posterior) is furthermore hampered by a wide range of variability, from ±2 mm to ±10 mm.\(^9\) Thus, ultrasound imaging has been the first-line image modality for years even though no commonly accepted standardised image acquisition and reading protocol exists. In two separate reports, the variability of AAA diameter measurement due to pulse wave propagation and the exact position of calliper placement on the vessel wall has been outlined.\(^10,11\) We present a standardised ultrasound method controlling for both the variation during the cardiac cycle and calliper placement according to the vessel wall layers and structures delineated on ultrasound imaging. Our aim was to evaluate the feasibility and reproducibility of a novel approach using ECG (electrocardiogram)-gated ultrasound and off-line reading of maximum anterior-posterior AAA diameter.

**METHODS**

The reproducibility study was carried out as a sub-study of the “AORTA trial”, a phase II randomised, controlled, multicentre clinical trial (ClinicalTrials.gov Identifier NCT01354184) evaluating the ability of a new pharmaceutical compound to attenuate aneurysm expansion. In the “AORTA trial”, 484 patients were screened and 326 patients were randomised. One hundred and fifty eight patients were screening failures, either because they did not fulfil the inclusion criterion of the trial (maximum AAA diameter within 39–49 mm) or due to inadequate image quality. Out of the 326 randomised patients we randomly selected and consecutively enrolled a group of patients \((N = 27)\) who were scheduled for a 6-month follow-up visit during February 2012.

The measurement procedure was divided into two steps: the ultrasound scan procedure and the subsequent off-line reading process by which all measurements for further analysis were made.

In a mutual blinded setup, all patients were scanned by two experienced (±10 years) ultrasound operators. For intra-operator variability assessment the patient was scanned by the same operator both morning and afternoon. Hence, as each patient \((N = 27)\) was examined twice by the same operator, there were 54 concurrent video recordings available for inter-operator assessment.

The patients were scanned in the supine position, ECG electrodes were attached and all the scans were performed with a CX-50 US system (Philips\(^\text{®}\), Bothell, Washington, USA) using a 5 MHz curved array transducer (C5-1). In the case of obesity, patients were investigated slightly turned to the side showing the best image of the aorta. The B-mode gain level was adjusted so that echoes were just seen in the aortic lumen, and two focal zones were adjusted to the anterior and posterior wall, resulting in an average frame rate of 15–25 per second.

The aorta was interrogated from the superior mesenteric artery to the level of the iliac bifurcation. For proper orientation and to ensure that the scanning angle was perpendicular to the aneurysm centre line, both cross-sectional and longitudinal imaging was used. Once the section of the aorta with the greatest diameter was identified, the continuous ECG recording was activated and a 10-sec video sequence was recorded. All stored videos were copied onto CDs and shipped to the core lab at the University Hospital of Copenhagen. All readings were performed by a third person, a ‘core reader’ (KB), using software for ultrasound images and recordings (QLAB v. 9.1, Philips, Bothell, Washington, USA). This allows ECG gating and minute calliper placement, and subsequently conversion of all measurements into an Excel file.

The delineation between actual vessel wall and the edge of the lumbar vertebrae or ligaments may rely on the link and communication with the lateral vessel wall, which are best evaluated on cross-sectional scan. On longitudinal section scans all echo lines extend in the same direction. Furthermore, an incorrect scanning angle may as well occur in longitudinal section as in a cross-sectional scan view. For these reasons, all measurements were only performed on the cross-sectional video recordings. By reviewing the recorded videos in QLAB, maximum dilatation (peak-systole) was visually identified. Subsequently, end-diastole was identified, typically by scrolling 3–4 frames backwards. In total, 20 measurements were made on each 10-sec video, using two different measurement methods (A and B) evaluated in peak-systole and end-diastole, and thus producing 4 main results, each derived from the mean of 5 measurements. In the case of arrhythmia, measurements were only performed after a ‘regular’ heart beat.

Leading edge of adventitia at the anterior wall was defined as the first echo line assessed to have continuity to the lateral vessel wall. Leading edge of intima on posterior wall was defined as the first echo line representing the lumen—intima interface, whereas leading edge of adventitia was defined as the second echo line representing the media—adventitia interface (Fig. 1).

According to the above, we defined two different methods measuring maximum diameter, both performed in cross-sectional mode using the anterior-posterior axis. Method A was from the leading edge adventitia (anterior wall) to the leading edge adventitia (posterior wall). Method B was from the leading edge adventitia (anterior wall) to leading edge intima (posterior wall) (Fig. 1).

In the core lab, a time delay of 14 days was used between evaluating recordings performed by the first and the second operator, likewise evaluating first and second recordings for the same operator.
The reproducibility of the reading process in the core lab was assessed by comparing the first set of video recordings performed by the first operator twice. In order to make the readings as independent as possible, a 2-week interval was intercalated between readings.

Statistics

Evaluation of the methods A and B was performed using the Bland Altman method, where the differences between paired measurements of the same subject are plotted against the mean outcome, showing the mean difference and the upper and lower limits of agreement given by the mean ± 1.96 × SD (standard deviation). Inter- and intra-operator reproducibility coefficients were expressed as 1.96 × SD. To compare means we used the Student’s t-test. Observed variability differences between methods were tested for homogeneity of variances using Levene’s test. All calculations were performed using SPSS v. 18.01 (SPSS Inc, Chicago, IL, USA).

RESULTS

Patients

Twenty seven patients with a small asymptomatic AAA of mean diameter 45.1 mm (ranging from 39.9 to 57.6 mm), measured using method A peak-systole, were included.

Inter-operator variability

Inter-operator variability measures were assessed, and all methods showed a variability of around ± 3 mm with reproducibility coefficients (1.96 SD) ranging from 3.0 mm to 3.4 mm (Fig. 2). Method A systole numerically had the lowest observed mean difference of 0.02 mm, although this was not statistically significantly different (Student’s t-test) from method B end-diastole with a mean difference of 0.4 mm ($\rho = 0.17$).

There were no statistical differences ($\rho = 0.49$) in the variances between methods A and B (Levene’s test for variance of homogeneity comparing method A end-diastole with the greatest SD vs method B peak-systole with the lowest SD), although measurements obtained by method B peak-systole numerically were found to have the smallest SD = 1.5 mm.

Intra-operator variability

Intra-operator variability for operators 1 and 2 is shown in Table 1. For operator 2 showing the greatest mean difference between measurements obtained by method B peak-systole and method A end-diastole, we performed Levene’s test, but could not prove any significant difference ($\rho = 0.26$).

Variability of the core reading process

There was no statistical difference in the reproducibility of the different methods even though method A peak-systole numerically had less variability than method B end-diastole ($\rho = 0.314$) (Table 2).

Cardiac cycle and method variability

The mean difference between maximum diameters obtained in peak-systole (45.3 mm) and end-diastole (44.7 mm) for method A was 0.6 mm, ranging from 0.1 to 1.3 mm. The mean difference between maximum diameters obtained in peak-systole (43.7 mm) and end-diastole

Figure 1. Cross-sectional illustration of abdominal aortic aneurysm seen on ultrasound imaging. The posterior wall is defined by two echogenic lines: the lumen-interface representing the leading edge of intima and the media–adventitia interface representing the leading edge of adventitia. The intima-media complex is seen in between these defined lines.
The difference between methods A and B peak-systole was 1.6 mm, range 0.4–6.4 mm, and the difference between methods A and B end-diastole was 1.7 mm, range 0.6–6.5 mm.

**DISCUSSION**

We investigated an ECG-gated ultrasound method measuring maximum aortic aneurysm diameter, and demonstrated reproducibility within ±3 mm for all methods. Since the reported average annual growth rate of AAAs with diameter 30–55 mm ranges between 2 and 6 mm, we believe the difference between the two operators was small enough to be considered acceptable.

**Table 1. Intra-operator variability for the two operators.**

<table>
<thead>
<tr>
<th>Operator Method</th>
<th>Operator 1</th>
<th>Operator 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac cycle</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Systole</td>
<td>Diastole</td>
<td>Systole</td>
</tr>
<tr>
<td>Mean difference (mm)</td>
<td>0.04</td>
<td>−0.3</td>
</tr>
<tr>
<td>Lower LoA (mm)</td>
<td>−2.9</td>
<td>−3.2</td>
</tr>
<tr>
<td>Upper LoA (mm)</td>
<td>2.8</td>
<td>2.6</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.5</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Method A: leading edge adventitia anterior wall — leading edge adventitia posterior wall. Method B: leading edge adventitia anterior wall — leading edge intima posterior wall. LoA: limit of agreement = mean ± 1.96 × SD (standard deviation).
K. Bredahl et al.

Table 2. Assessment of the single core reader.

<table>
<thead>
<tr>
<th>Cardiac cycle</th>
<th>Method A</th>
<th>Method B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Systole</td>
<td>Diastole</td>
</tr>
<tr>
<td>Mean diff. (mm)</td>
<td>−0.3</td>
<td>−0.2</td>
</tr>
<tr>
<td>Lower LoA (mm)</td>
<td>−1.3</td>
<td>−1.3</td>
</tr>
<tr>
<td>Upper LoA (mm)</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>SD (mm)</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Method A: leading edge adventitia anterior wall — leading edge adventitia posterior wall. Method B: leading edge adventitia anterior wall — leading edge intima posterior wall. LoA: limit of agreement = mean ± 1.96 × SD (standard deviation).

3 mm, we find that our standardised procedure can be used for the detection of even small changes in AAA diameter in individuals, to reduce the risk of type II error in clinical trials. The reported reproducibility of ultrasonic diameter measurement (anterior-posterior) varies in the available literature between ±2 mm and ±10 mm. Among the published reports, an inter-operator standard deviation of 1.1 mm has been determined in two studies. These results appear to be slightly better than ours, but as normal aortas and aortas smaller than our inclusion criteria were included in this study, the results are not comparable. The uncertainty with regard to acquiring the cross-sectional image at the exact anatomical location of the maximum diameter increases with increasing aneurism size and, consequently, the standard deviation is expected to increase when scanning and measuring larger aneurisms. Therefore, our results for inter- and intra-operator variability are comparable with the best reported results for aneurisms within the diameter range of 39–49 cm (inclusion criteria in the ‘AORTA trial’).

Previous reports of reproducibility measures were primarily motivated by the purpose of validating ultrasound imaging in the specific context of screening programs for AAAs. This study was motivated by the need to detect even small differences in the incremental growth rate in a randomised controlled trial testing a new pharmaceutical compound for halting growth of AAA and, in contrast to screening programs, demanding higher standards for producing reliable estimates of growth. As more experienced operators show better reproducibility coefficients compared with less experienced sonographers, variability may simply depend on level of experience.

Previous studies have, however, not taken into account that measurement of the maximum AAA diameter is influenced by the pulse wave propagation, found on average to be 1.94 mm between diastole and systole and with a wide range (0–4.7 mm). In addition, systematic placement of the calliper according to ultrasonically well-defined structures is important, as the variability due to this factor has been reported to be up to 6 mm between inner-to-inner versus outer-to-outer vessel wall. The UK AAA screening group has recently demonstrated that measuring inner-to-inner vessel wall diameter improves reproducibility in preference to measuring outer-to-outer vessel wall diameter. This approach, however, appears not to be sufficiently sensitive to account for all potential sources of variability. In the clinical trial setting, and maybe also for clinical follow-up of small asymptomatic AAAs, a different standard may be required in order to be able to detect even small changes in growth rates. Controlling the variability related to unclear vessel wall definitions, plaque or thrombus, and variations during the cardiac cycle is the scope of this novel approach, and to our knowledge these issues have never been addressed or thoroughly investigated before.

The most distinctive ultrasound reflection is obtained when ultrasound passes from a layer of low acoustic impedance to a layer of high acoustic impedance (poor to strong reflection). This suggests that the correct calliper placement on the anterior wall would be the interface between the surrounding tissue and adventitia (the leading edge of the adventitial layer of the anterior wall). For the same reasons, measuring intima-media thickness is measured on the posterior wall rather than the anterior wall. AAA expansion can also be accompanied by the development of atherosclerotic disease, which changes the thickness of the vessel wall and the echo details seen on the ultrasound image. On the posterior wall, optimally two echo lines separated by the intima-media complex are visualized, representing the lumen—intima interface and the anterior margin of adventitia, respectively. Atherosclerotic disease leading to plaque formation takes place in the intima-media complex. Using the first echo line on the posterior wall (method B), concurrent growth of plaque may potentially mask actual expansion of the aneurysm. In contrast, tunica adventitia is not affected by atherosclerotic disease as opposed to the intima-media complex, and consists of collagen performing high acoustic impedance. Moreover, considering an image resolution of 1–2 mm, a thin intima-media complex may not be clearly visualized in detail and, as previously suggested, thrombus and intima may be difficult to differentiate. Our results did not demonstrate any advantages in using method A in preference to method B; however, as our present reproducibility study did not include study over time, this potential limitation of method B cannot be evaluated from our data at present and should be addressed in future research.

An inflammatory response would not affect this method on the posterior wall, but we recognize that expansion of the adventitial layer would affect the measurement on the anterior wall. Although not part of the present study, measuring AAA wall thickness over time may be an interesting issue.

We demonstrated a difference between measurements in peak-systole and end-diastole of 0.6–0.7 cm, which is smaller but in line with previous investigations (mean difference of 1.94 mm), since they included larger aneurysms and our study involved a smaller sample. Some ultrasound operators measure a maximum dilatation simply by visual appearance, but considering that pulse wave propagation from the left ventricle in systole only lasts a fraction (i.e. 1/10th) of the cardiac cycle combined with a frame-rate presentation of approximate 15–25, potential
variability in measurement at the exact moment of maximum dilatation exists. During end-diastole the changes in aortic diameter from frame to frame will be the least and the reproducibility would theoretically be highest, but this has never been confirmed and would challenge the concept of using peak-systolic measures. In contrast to one study, our data do not indicate a significantly reduced inter-operator variability in favour of end-diastole. It has been suggested that the more distensible aneurysms tend to rupture more often than stiff aneurysms; we therefore find it more suitable to base our clinical decisions on the measurement performed in systole. For use in daily clinical practice, we think measuring at maximum dilatation by simple visual appearance may be as good as performing peak-systole measurement assisted by ECG-gating, but this needs to be confirmed in a reproducibility study.

The strength of this study is that we used experienced operators, and 10-sec video recordings for subsequent standardised analysis by a core reader. On the other hand, we are aware that we only used two operators, and the group of patients was selected from a randomised clinical trial with pre-selection criteria applied relating to aneurism size and image quality. Moreover, with the use of only one core reader to control the variability of the reading process, which was only ±1 mm, the majority of the variability can be attributed to image acquisition.

Other specifics and conditions speculated to have influence on diameter measurement were not evaluated in this study; for example, plane of acquisition (transverse or longitudinal), axis (anterior-posterior, any direction, and transverse), choice of transducer frequency (defining axial and spatial resolution) and choice of either anatomical or aeurismal reference.

In conclusion, we present a new ECG-gated method for ultrasonic measurement of AAA showing reproducibility within ±3 mm. In clinical practice we recommend measurements from leading edge adventitia anterior wall to leading edge adventitia posterior performed during peak-systole.

ACKNOWLEDGEMENTS

Project Nurse Bente Culmsee (operator 1) and Sidsel Nissen (operator 2) are thanked for standardized data sampling.

FUNDING

None.

CONFLICT OF INTEREST

Co-authors Eldrup N and Sillesen H are both member of the Steering Committee in the “AORTA Trial” and received honoraria from Cardoz AB.

Co-author Sillesen H has received a research grant and honorarium from Philips Ultrasound.

Co-author Meyer C is Chief Medical Officer in Cardoz AB sponsoring the “AORTA Trial”.

REFERENCES