


Intra- and Interobserver Variability of 4D Ultrasound Examination of the Infrarenal Aorta

Wojciech Derwich, MD , Antonia Wiedemann, MD, Andreas Wittek, MSc in Eng, Natalie Filmann, M. Sc., Christopher Blase, M. Sc., Thomas Schmitz-Rixen, MD

Received October 4, 2020, from the Department of Vascular and Endovascular Surgery, University Hospital Frankfurt Goethe University, Frankfurt am Main, Germany (W.D., A.W., T.S.-R.); Personalised Biomedical Engineering Lab, Frankfurt University of Applied Sciences, Frankfurt am Main, Germany (A.W., C.B.); Department of Mechanical Engineering, University of Siegen, Siegen, Germany (A.W.); Institute for Biostatistics and Mathematical Modeling, Goethe University, Frankfurt am Main, Germany (N.F.); and Department of Biological Sciences, Goethe University, Frankfurt am Main, Germany (C.B.). Manuscript accepted for publication December 24, 2020.

All of the authors of this article have reported no disclosures.

Address correspondence to Wojciech Derwich, MD, Department of Vascular and Endovascular Surgery, University Hospital Frankfurt Goethe University, Theodor-Stern-Kai 7, Frankfurt am Main 60590, Germany.

Email: wojciech.derwich@kgu.de

Abbreviations

2D-US, Two-dimensional ultrasound; 3D-US, three-dimensional ultrasound; 4D-US, four-dimensional ultrasound; AUROC, area under receiver operating curve; CI, confidence interval; CT, computer tomography; ICC, intraclass correlation coefficient; IID, independent and identically distributed; ROC, receiver operating characteristic; SD, standard deviation

doi:10.1002/jum.15622

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Objectives—The four-dimensional ultrasound (4D-US) enables imaging of the aortic segment and simultaneous determination of the wall expansion. The method shows a high spatial and temporal resolution, but its in vivo reliability is so far unknown for low-measure values. The present study determines the intraobserver repeatability and interobserver reproducibility of 4D-US in the atherosclerotic and non-atherosclerotic infrarenal aorta.

Methods—In all, 22 patients with non-aneurysmal aorta were examined by an experienced examiner and a medical student. After registration of 4D images, both the examiners marked the aortic wall manually before the commercially implemented speckle tracking algorithm was applied. The cyclic changes of the aortic diameter and circumferential strain were determined with the help of custom-made software. The reliability of 4D-US was tested by the intraclass correlation coefficient (ICC).

Results—The 4D-US measurements showed very good reliability for the maximum aortic diameter and the circumferential strain for all patients and for the non-atherosclerotic aortae (ICC >0.7), but low reliability for circumferential strain in calcified aortae (ICC = 0.29). The observer- and masking-related variances for both maximum diameter and circumferential strain were close to zero.

Conclusions—Despite the low-measured values, the high spatial and temporal resolution of the 4D-US enables a reliable evaluation of cyclic diameter changes and circumferential strain in non-aneurysmal aortae independent from the observer experience but with some limitations for calcified aortae. The 4D-US opens up a new perspective with regard to noninvasive, in vivo assessment of kinematic properties of the vessel wall in the abdominal aorta.

Key Words—4D-US; intra- and interobserver variability; reliability; strain

Introduction

Functional analysis of the infrarenal aorta assumes more importance in view of the development of ultrasound technology. Two-dimensional ultrasound (2D-US) enabled an assessment of the aortic morphology with the determination of the maximum diameter and the evaluation of the pulsatility of the aortic wall during the heart cycle. Subsequently, the introduction of three-dimensional ultrasound (3D-US) expanded the possibilities of sonographic diagnostics with complex imaging of the three-dimensional aortic morphology comparable to the precision

of computer tomography (CT) scans and enabled the geometry-based stress calculation of the infrarenal aorta.¹ Thanks to further improvements to the temporal resolution in the 3D-US imaging, 4D-US (3 spatial planes + time axis) evolved and made way for an era of complex kinematic analyses of pulsating structures such as the heart and vessels.^{2,3}

The clinical application of the strain analysis in previous studies has focused on the effects of increased aortic stiffness on left ventricular function.⁴ However, the introduction of real-time 3D speckle-tracking ultrasound (4D-US) enables the imaging of the entire aortic segment and thus the simultaneous determination of the longitudinal and circumferential expansion.⁵ In comparative studies, the age-related and pathology-specific changes in the strain pattern in the infrarenal aorta were demonstrated⁶ and characterized by corresponding changes in distensibility.⁷ The use of 4D-US provides information about the geometry of the aorta and local deformation of the aortic wall during the heart cycle. Such specific data about diastolic reference geometry and deformation enable biomechanical analyses with the characterization of the deformation pattern, determination of the patient-individual material properties,⁸ and description of the functional and morphological heterogeneity of the aortic wall in the healthy and aneurysmal aorta.^{6,9}

Ultrasound is a cost-effective, noninvasive, radiation-free, and readily available diagnostic tool. The main disadvantage of aortic sonography is a relevant intra- and interobserver variability. The 2D imaging method is associated with low variability for anterior–posterior aortic diameter measurements described by high intraobserver repeatability coefficients of 1.6–4.4 mm but poor interobserver reproducibility coefficients up to –10.5 to 10.4 mm.¹⁰ When compared to the CT-based anterior–posterior diameter measurement as the clinical gold standard, the 2D-US-based diameter measurements show a mean deviation of 6 mm and a variability of 9.4 mm. In contrast, the 3D-US imaging-based measurements show a significantly reduced mean deviation of –1.3 mm ($p < .001$) and range of variability of 6.6 mm ($p = .009$).¹¹ However, the accuracy of 3D strain imaging has not yet been sufficiently validated. In a direct comparison of myocardial strain imaging by MRI, 2D, and 4D ultrasound, Onokata et al found

the lowest observer variability and the highest intraclass correlation for 4D ultrasound. However, the heart muscle shows a pronounced physiological deformation with the global longitudinal strain that ranges from 15.9 to 22.1% during the cardiac cycle.¹² The deformation of the infrarenal aorta occurs mainly in the circumferential direction, is significantly lower than in the heart, and is described for circumferential strain by $10 \pm 6.9\%$ in young subjects and $2.5 \pm 1.5\%$ in older atherosclerotic patients.⁶ Bihari et al highlighted ex vivo high-measurement accuracy of the 4D US on a silicone model with maximum differences of 0.24, 0.54, and 0.76 mm in the x , y , and z directions, respectively, when compared to the measurements using video photogrammetry.¹³ However, the reliability of the 4D-US, especially when considered with regard to the small size of the measured values, is so far unknown for the healthy and diseased aortic wall. In this study, we determine the intra- and interobserver variability of 4D-US in the non-aneurysmal atherosclerotic and non-atherosclerotic infrarenal aorta. First, we compare the accuracy of the maximum diameter measurement by the clinically used 2D-US and 4D-US. As a second step, we evaluate the intraobserver repeatability and interobserver reproducibility of 4D-US in young and old atherosclerotic patients. Finally, we examine the impact of single analysis steps on the strain analysis.

Materials and Methods

Considering 2 investigators with 2 repetitions per patient, double masking process per examination, and known dispersion of circumferential strain in the population by $10 \pm 6.9\%$ in young subjects and $2.5 \pm 1.5\%$ in older atherosclerotic patients, a case number of 10 ± 1 subjects per group was calculated by the Institute for Biostatistics and Mathematical Modeling. After obtaining a written declaration of consent, we examined, in our prospective study, 22 patients with a normal caliber, non-aneurysmal infrarenal aorta at the Department for Vascular and Endovascular Surgery of the University Hospital Frankfurt Goethe University, Frankfurt am Main, Germany. Three patients were excluded due to incomplete recordings. The study design was approved by the local ethics committee (ethics

Table 1. Epidemiological Data for Patients With Non-atherosclerotic ($n = 9$) and Atherosclerotic Aorta ($n = 10$)

	Non-atherosclerotic Aorta	Atherosclerotic Aorta
Sex (male/female)	8/1	9/1
Smoking	2	3
Diabetes mellitus	0	4
Arterial hypertension	1	5
PAD	0	6
Stroke	0	2
Coronary heart disease	0	4

committee business number: 275/17). The patients were divided into 2 groups of non-atherosclerotic aorta ($n = 9$) and atherosclerotic aorta ($n = 10$) (Tables 1 and 2).

Data Collection

The patients were examined by 2 observers after a rest period of 5 minutes in a supine position. All recordings were made during a breathing break at the end of exhalation. The first examiner was a vascular surgeon experienced in ultrasound application; the second examiner was a student of the medical faculty. All examinations were performed successively in the same session without significant breaks. Because of the clear anatomical identifiability, the distal aortic segment close to the aortic bifurcation was examined for further analysis. The systolic and diastolic blood pressure was determined before and after the examination using sphygmomanometry on the brachial artery. The maximum anterior–posterior end-systolic aortic diameter was measured by 2D-US in the axial plane in accordance with the leading-edge-technique with the 6.0 MHz concave probe 6C1. The 4D-US

examination was performed with a 4D echocardiography system (Artida, Toshiba Medical Systems Corporation, Otawara Japan) using a 4D transthoracic probe (PST-25SX, 1–4 MHz phased array matrix transducer, Toshiba). A complete real-time 3D image was created by summing the movement recordings of 6 subvolumes of $90^\circ \times 15^\circ$, each, over 6 successive cardiac cycles. Data collection was ECG-triggered starting with end-diastole (Figure 1A).

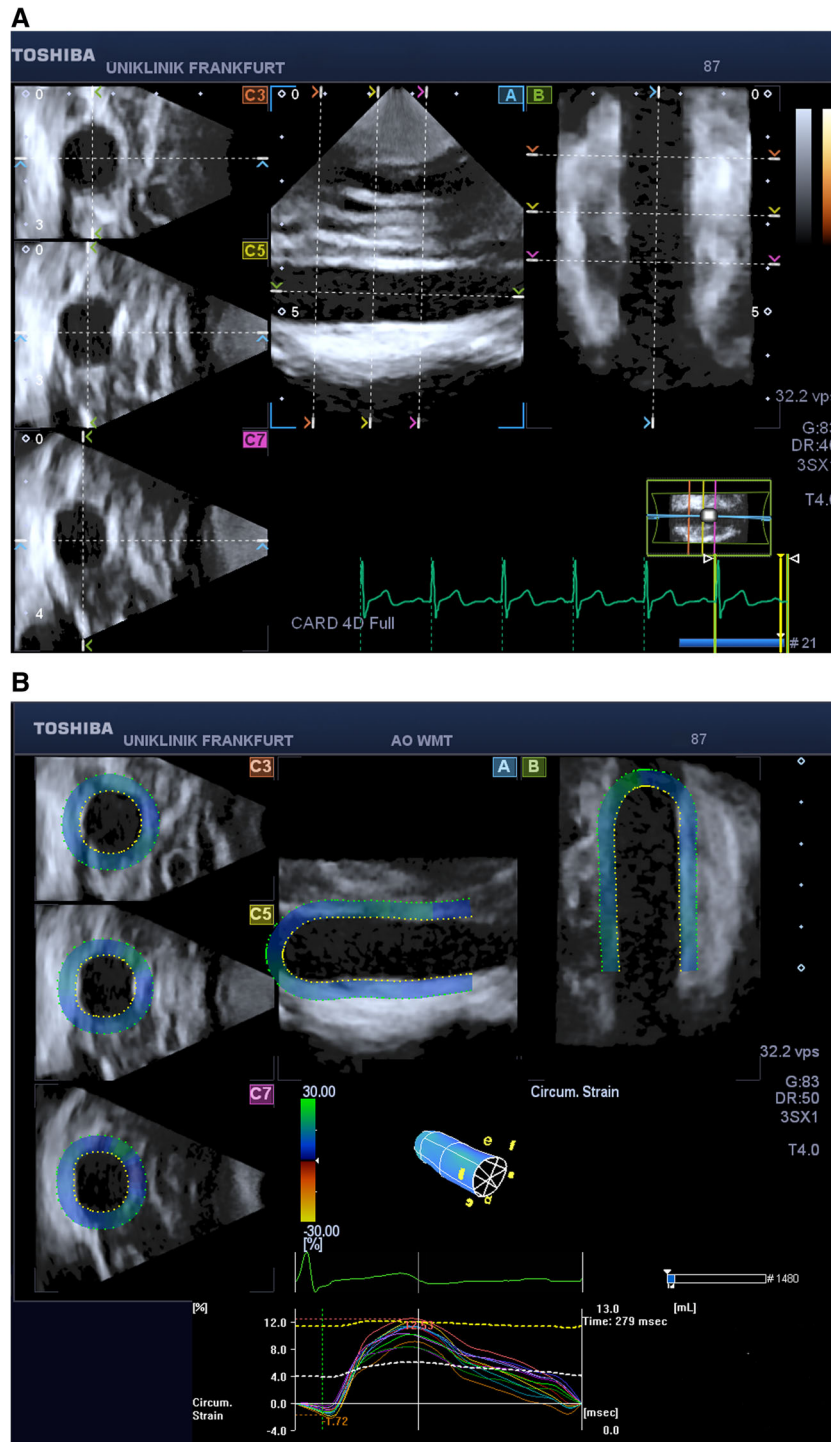
Application of the Speckle Tracking Algorithm

The native recordings were processed in the external UltraExtend Workstation with the Advanced Cardiac Package (ACP, Toshiba Medical Systems). Each examination was masked by the 2 examiners. Post-processing involved manual marking of the inner contour of the aortic wall in the end-diastole and application of the speckle tracking algorithm. The native image was a three-dimensional data set and enabled a multi-planar reconstruction. An optimal presentation of the target structure was achieved by adapting the centerline. The orientation in the three-dimensional space was based on the global Cartesian coordinate system with the following assignment: x -axis from cranial to caudal, y -axis from anterior to posterior, and z -axis from the right to the left side of the body. The manual masking of the aortic wall was carried out successively in the sagittal (A-plane) and frontal (B-plane) planes in a counterclockwise direction. After confirming the process, the system automatically completed the spatial segmentation and marked the corresponding contours in 3 transverse planes. This automated step could be corrected manually. After confirmation of the completed masking process, the speckle tracking algorithm was activated in which it was the system that defined the grayscale

Table 2. Physiological Data for Patients With Non-atherosclerotic ($n = 9$) and Atherosclerotic Aorta ($n = 10$)

	Non-atherosclerotic Aorta		Atherosclerotic Aorta	
	Median	Range	Median	Range
Age	27	(23; 37)	69	(47; 86)
Body height (m)	1.8	(1.65; 1.86)	1.73	(1.53; 1.96)
Body weight (kg)	75	(65; 88)	72	(57; 108)
Body mass index	24	(19; 29)	26	(19; 28)
Systolic blood pressure (mmHg)	120	(115; 145)	140	(115; 160)
Diastolic blood pressure (mmHg)	75	(50; 100)	70	(60; 90)
Anterior–posterior aortic diameter in 2D-US	15	(13; 18)	17	(16; 29)

Figure 1. Image processing with a real-time 3D speckle tracking algorithm. **(A)** The image shows sagittal **(A)** and frontal **(B)**, longitudinal sections, and 3 cross-sections (C3, C5, and C7) of the volumetric data. **(B)** After the optimal setting of the centerline, the inner contour of the aortic wall is marked manually.



pattern of a three-dimensional cube and tracked it during the cardiac cycle. The cardiac software that was used for determining the wall strain in the heart required the marking of an apex that had no corresponding anatomical structures in the displayed tubular subject like that of an aorta. The resulting fictive pseudo-apex had to be removed from the data sets before further analysis (Figure 1B).

3D Wall Motion Tracking

The Advanced Cardiac Package enables the tracking of wall volumes in a three-dimensional space throughout the heart cycle. The examined structure is divided into small template volumes ($10 \times 10 \times 10$ mm cubes). The center of a cube is defined by a position point. Each tracked position point is moved in the consecutive recordings to estimated positions in the three-dimensional space. The motion vector for these position changes is calculated using the template matching technique. In the next image, the template volumes are formed again and compared with the original cube. The best match leads to the determination of the new estimation position. Finally, an interpolation of the motion vectors is carried out with a 3D interpolation algorithm so that the motion can be interpolated for any x points on the aortic wall.³

In collaboration with Toshiba Medical Systems, a customized interface was developed that enabled the export of wall movement data in ASCII format as spatial coordinates. The tracking points were collected as x , y , and z coordinates of 36×36 nodes (36 layers with 36 nodes each along the circumference). After exclusion of artificial nodes in the region of the pseudo-apex, up to 30×36 nodes described the 3D shape of the luminal side of the aortic wall. Each layer consisting of 36 nodes described an aortic cross-section along the axis of the imaged segment. These data were obtained at different times throughout an entire cardiac cycle such that ultimately the cyclic movement of the wall was mapped.

Determination of Wall Deformation

Further processing was carried out using the specially designed software, written with Visual Basic® for applications (Microsoft Corp., Redmond, WA). All calculations were carried out based on the position coordinates. First, the artificial nodes that were not located in the aortic wall and formed the pseudo-apex

had to be deleted (Figure 2iA,B). These nodes were detected in a 3D model, which was based on the primary 36×36 nodes. The methods for calculating aortic diameter and circumferential strain of the imaged aortic segment for each temporal frame throughout the cardiac cycle have been described in detail by Wittek and Karatolios.² For each aortic cross-section that was defined by 1 layer of 36 nodal points, a center point was determined. A straight centerline of the aortic segment was obtained as a connection of the center points of the 2 end cross-sections. The local radius of a nodal point on the aortic wall was obtained as the shortest distance between this nodal point and the centerline (Figure 3).

The radius (and, consequently, the diameter) of the aortic segment for each temporal frame was obtained as the mean of the local radii that were determined for this frame. The diastolic and systolic diameters were identified as the minimum and maximum diameters that were observed throughout the whole cardiac cycle, respectively. The systolic diameter was used to describe the static configuration of the wall. In contrast, the dynamic configuration of the aortic wall was described by the circumferential strain, which was defined as a deformation of the aorta in the surrounding direction in relation to its diastolic reference shape, expressed mathematically as $\varepsilon = (d_{\text{sys}} - d_{\text{dia}})/d_{\text{dia}}$, where d_{sys} and d_{dia} are the end-systolic and end-diastolic diameters, respectively. The cyclic registration of strain amplitude of all speckles with the temporal resolution of 22–26 frames/s enabled the precise definition of diastolic and systolic aortic diameters and patient-specific strain. The circumferential strain was chosen as the most significant parameter for the investigation of 4D-US reliability.^{8,7}

Determination of Variability

The 4D-US examination of the infrarenal aorta is a two-stage imaging method. In the first step, a time-resolved scan of the aorta is carried out. In the second step, the contours of the aortic wall are manually marked on all 3 planes in the commercially available workstation. The overall result is, thus, influenced by the recording of the examination and by the manual masking of the initial images as part of the intraobserver variability. The final result can be affected by the variability of the parameter in the test objects themselves and by the differences in the

Figure 2. Three-dimensional model of the infrarenal aorta with artificial nodes with (A) and without (B) pseudo-apex, and (C) systolic displacement vector field (colored pegs) mapped on subdivided diastolic reference geometry of aortic segments.

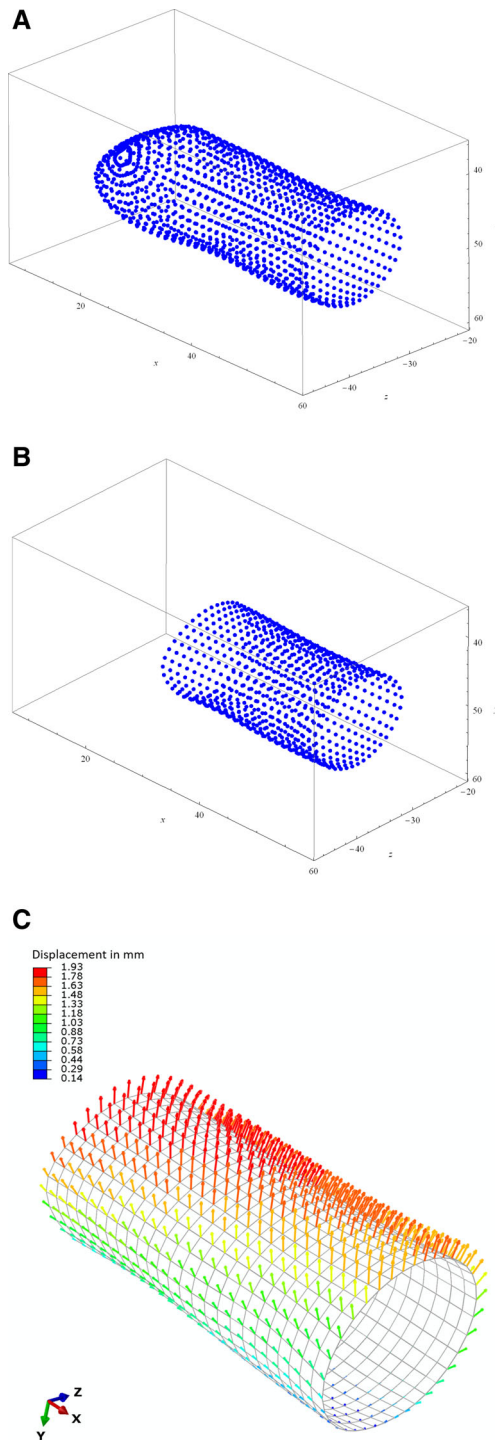
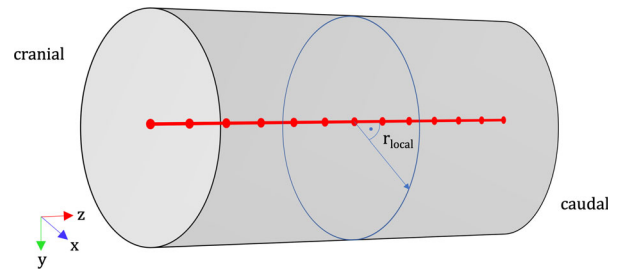


Figure 3. The luminal shape of an imaged aortic segment with centroids of aortic cross-sections along the axis of the segment and straight centerline of the segment. Exemplarily, the local radius r_{local} of a nodal point on the wall is shown.



examination quality between the several examiners, that is, the interobserver variability. The temporal resolution of 4D-US can contribute to the overall error.

Statistical Analysis

Statistical analysis was done using R (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria. Packages: lme4 and lattice) and BiAS. (Epsilon-Verlag, Hochheim Darmstadt, Germany, version 10.14–11/2014). For the descriptive analysis of the kinematic parameters, mean (with 95% confidence interval [CI]), standard deviation (SD), median, and minimum and maximum were calculated. For group comparisons, the nonparametric Wilcoxon–Mann–Whitney U test was used. All tests were two-sided and a significance level of 0.05 was used. The measurement of maximum diameter from 2D- and 4D-US examination was compared by using the Bland–Altman analysis method. The mean value of the paired differences, the coefficient of repeatability equal to $1.96 \times \text{SD}$ of the differences, and the limits of agreement (LoA) equal to $\text{mean} \pm 1.96 \text{ SD}$ were determined. Clinical thresholds of 5 mm for the coefficient of repeatability and $[-5; +5]$ mm for the LoA were applied to diameter analysis.

To measure the reliability among multiple measurements of the kinetic parameters, we calculated the different types of ICCs (Figure 4). Inpatient correlation coefficient (ICC[inpatient]) values were estimated via this random-effects model.

$$Y_{ijk} = \mu + s_i + r_j + q_k + e_{ijk}, \quad i = 1, 2, \dots, 19, \quad j = 1, 2, \quad k = 1, 2 \quad (1)$$

where Y_{ij} is the measurement of the j th rater with the k th masking in the i th patient, μ is the fixed effect, s_i the random effect of the i th patient, independent and identically distributed (IID) with $N(0, \sigma_s^2)$, r_j is the random effect of the j th observer, iid $\sim N(0, \sigma_r^2)$, q_k is the random effect of the k th masking, iid $\sim N(0, \sigma_k^2)$, and e_{ijk} is the measurement error, iid $\sim N(0, \sigma_e^2)$. In the following, the empirical estimates of variance components will also be denoted by σ_s^2 , σ_r^2 , σ_k^2 , and σ_e^2 .

The measurements of an experiment can be said to be reliable if the differences between observers' observations, repetitions, and masking are small compared to the differences between the individuals. The ICC as ICC[intrapatient] measures the relation of the variance that is attributed to a random factor as follows:

$$\text{ICC}[\text{intrapatient}] = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_r^2 + \sigma_k^2 + \sigma_e^2}. \tag{2}$$

For interpretation of the ICC values, we followed the suggestion of Landis and Koch wherein an ICC of 0–0.20 denotes a slight, 0.21–0.40 denotes a fair, 0.41–0.60 denotes a moderate, 0.61–0.80 denotes a substantial, and 0.81–1.00 denotes (an almost) perfect reliability.¹⁴

Analogously, it is also possible to estimate the influence of the observer σ_r^2 and the masking σ_k^2 in relation to the total variance:

$$\text{ICC}[\text{observer}] = \frac{\sigma_r^2}{\sigma_s^2 + \sigma_r^2 + \sigma_k^2 + \sigma_e^2} \text{ and} \tag{3}$$

$$\text{ICC}[\text{masking}] = \frac{\sigma_k^2}{\sigma_s^2 + \sigma_r^2 + \sigma_k^2 + \sigma_e^2}. \tag{4}$$

It must be noted that high ICC[observer] and high ICC[masking] across patients are not desirable as this would suggest that measurements mainly depend on the observer and masking characteristics, respectively, and not on the patient characteristics.¹⁵

Furthermore, a receiver operating characteristic (ROC) analysis was conducted including the estimations of the area under receiver operating curve (AUROC) of all kinetic parameters (with 95% CI), and the DeLong test was used to evaluate the diagnostic accuracy of the kinematic parameters. Optimal cutoff values were obtained by optimizing the Youden's index, whereby only 1 measurement per patient was taken into account, that is, the first examination of observer 1 and his own masking.

Results

Descriptive Statistics for Non-Atherosclerotic and Atherosclerotic Aorta

The atherosclerotic patients were an older group and, as a result, characterized by a bigger aortic diameter than the young non-atherosclerotic patients (mean 18.4 ± 3.9 mm versus 16.0 ± 1.4 mm, $p < .001$) with the coefficient of variation for intrapatient value variability being 7.2 and 3.9%, respectively. The circumferential strain was significantly lower in the older patients compared to non-atherosclerotic probands (mean 2.4 ± 1.0 versus $9.7 \pm 3.6\%$, $p < .001$) with the

Figure 4. The influence of the observer experience by ultrasound examination on the final result is estimated by a variance analysis for values 1, 2, 5, and 6 versus values 3, 4, 7, and 8. The influence of observer experience by masking on the final result is examined by a variance analysis of values 1, 3, 5, and 7 versus values 2, 4, 6, and 8.

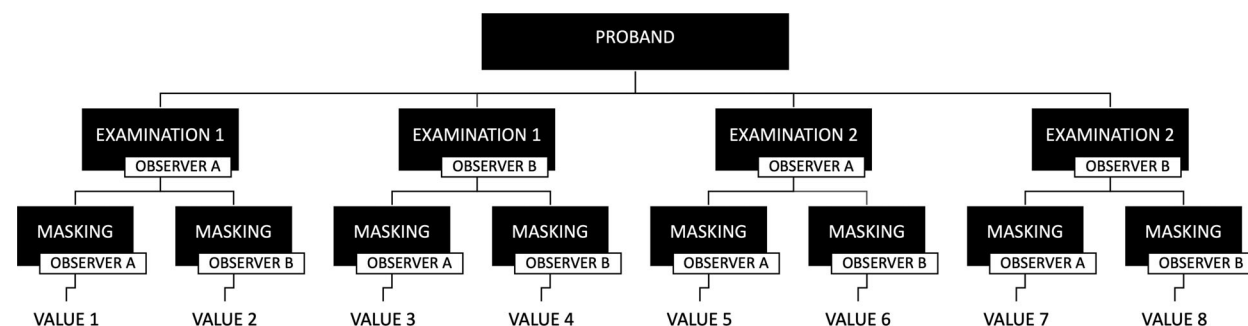


Table 3. Kinematic Differences Between the Patients with Non-atherosclerotic ($n = 9$) and Atherosclerotic ($n = 10$) Aortae as Measured by 4D-US

Parameter	Group	Mean \pm SD	95% CI	Median	Min; Max	p^*
Maximum diameter (mm)	Non-atherosclerotic aorta	16.0 \pm 1.4	15.7; 16.3	15.7	13.8; 18.6	<.001
	Atherosclerotic aorta	18.4 \pm 3.9	17.5; 19.2	18.0	12.7; 28.4	
Circumferential strain (%)	Non-atherosclerotic aorta	9.7 \pm 3.6	8.8; 10.6	9.3	3.8; 19.3	<.001
	Atherosclerotic aorta	2.4 \pm 1.0	2.1; 2.6	2.4	0.5; 4.7	

*Wilcoxon-Mann-Whitney test

coefficient of variation for inpatient value variability being 33 and 13.7%, respectively (Table 3).

Comparison of Maximum Diameter Measurement with 2D and 4D Ultrasound

In patient-related comparisons (Table 4), we observed a strong correlation between the measurement of maximum aortic diameter using 2D- and 4D-US for all patients ($r = .92$, $p < .001$) with a mean difference of -0.17 mm, a coefficient of repeatability of 3.19 mm, and 95% LoA equal to -3.36 to 3.02 mm as per the Bland-Altman-analysis (Figure 5).

Reliability of 4D Ultrasound Examination on the Infrarenal Aorta

The patients with non-atherosclerotic aorta showed a narrow range of maximum aortic diameter between 13.8 and 18.6 mm in comparison to older patients (12.7–28.4 mm). However, the young patients were characterized by a wide range of circumferential strain of 3.8–19.3% in contrast to a narrow range of 0.5–4.7% in the case of the atherosclerotic aorta (Figure 6). The 4D-US measurements showed very good reliability with regard to the determination of the maximum aortic diameter and the circumferential strain for all the patients together and for the non-atherosclerotic aortae (ICC >0.7). However, the reliability of the method was low for the circumferential strain in atherosclerotic aortae with low strain amplitude, but there was a high variance between the observers (ICC = 0.29) (Table 5).

Impact of the Observers' Experiences on the Final Result in Performing the 4D-US and Masking

Observer- and masking-related variances for maximum diameter (Table 6) and circumferential strain (Table 7) were close to zero, and thus the experience of the observer by performing examination and masking had a low impact on the quality of the 4D-US and its final result.

ROC Analysis

The patients with non-atherosclerotic and atherosclerotic aortae cannot be differentiated by the maximum diameter. However, circumferential strain with the optimal cutoff by 4.22% separates the 2 groups with a total misclassification rate of 0% (Table 8).

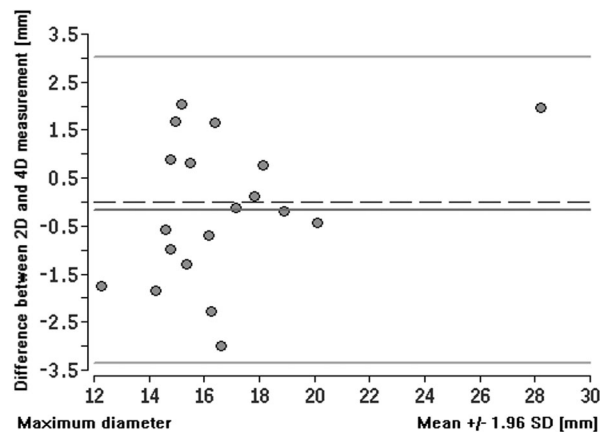
Discussion

Ultrasound examination has become an effective screening method in the diagnosis of the infrarenal aorta.¹⁶ Nevertheless, Beales et al describe a high variability of the diameter measurements in the infrarenal aorta¹⁰ that depends on the selection of the image plane, the aortic segment with the maximum diameter, and the parallax error. The introduction of 3D-US technology promised an increase in the accuracy of measurement by optimizing the centerline.¹⁷ However, the reproducibility of the diameter measurement in 3D-US is described only in a few studies. Ghulam et al were able to establish an excellent interoperator reproducibility for measuring the maximum diameter

Table 4. Maximum Diameter Measurement of the Non-atherosclerotic and Atherosclerotic ($n = 19$) Aorta in 2D- and 4D-US

		Mean \pm SD	95% CI	Median	Min; Max
Non-atherosclerotic and atherosclerotic aorta ($n = 19$)	2D	16.6 \pm 3.7	15.0; 18.3	15.8	11.4; 29.2
	4D	16.8 \pm 3.1	15.4; 18.2	16.0	13.2; 27.2

Figure 5. The Bland–Altman plot for the maximum diameter of the non-atherosclerotic and atherosclerotic aorta ($n = 19$) in 2D- and 4D-US.



in 3D-US with an ICC of 0.97.¹⁸ In a study comparing the 3D-US and 3D-CT centerline diameter, Bredahl et al reported the interobserver reproducibility coefficient to be 3.7 and 3.2 mm, respectively ($p >.05$). However, the mean ultrasound dual-plane diameter and the 3D-US centerline diameter were 2.6 mm and 1.8 mm smaller than the mean 3D-CT centerline diameter, respectively ($p <.01$).¹⁹ Since the

pulsatility of the aorta influences the systolic–diastolic diameter difference, 4D-US enables an adequate determination of the maximum aortic diameter with the selection of the optimal time window in the end-systole. We were able to demonstrate high reproducibility of the diameter determination in 2D- and 4D-US with a coefficient of repeatability of 3.19 mm and 95% LoA from 3.36 to 3.02 mm. The reliability of the results obtained by the 2 observers was very high for both non-atherosclerotic and atherosclerotic aortae.

Thus, the development of 4D-US opens up new avenues for describing the kinematic properties of the aorta.¹³ The deformation of the aortic wall can be differentiated by the circumferential strain in different clinical situations.⁶ The young aortae are characterized by a significantly higher mean circumferential strain compared to the atherosclerotic aortae as presented in this study and as described by the optimal cutoff value being at 4.22% between both the groups. With regard to the low strain amplitude and calcification-related imaging artifacts, the verification of the reliability of 4D-US appears to be justified. The intraclass correlation for both groups indicates a very high agreement between the 2 observers. The subgroup analysis, on the other hand, confirms the

Figure 6. Distribution of maximum diameter (A) and circumferential strain (B) in the non-atherosclerotic and atherosclerotic infrarenal aorta measured and masked by 2 observers in 4D-US, where E, examination; O, observer, and M, masking.

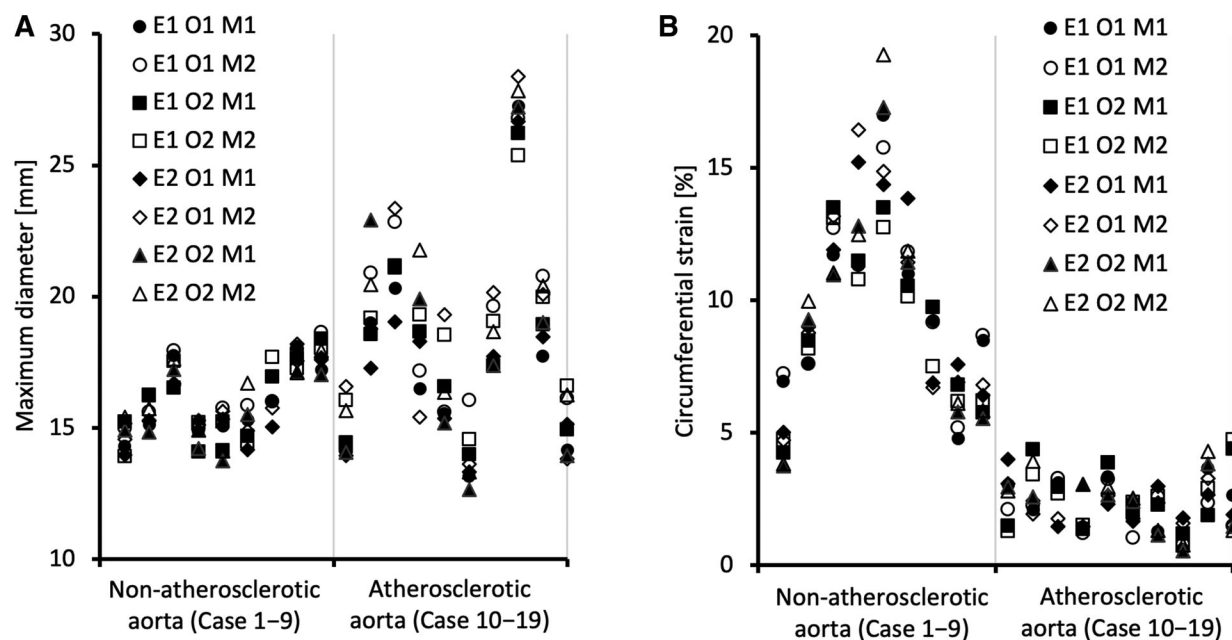


Table 5. Patient-related Intraclass Correlation for the Non-atherosclerotic Aortae ($n = 9$), Atherosclerotic Aortae ($n = 10$), and Both the Groups

Parameter	Patient-related ICC		
	All Patients	Non-atherosclerotic Aorta	Atherosclerotic Aorta
Maximum diameter	0.88	0.70	0.88
Circumferential strain	0.95	0.89	0.29

perfect reliability of the strain assessment in 4D-US for non-atherosclerotic aorta but shows a fair agreement in the atherosclerotic aortae only. The reduced reliability in atherosclerotic aortae can, possibly, be explained by calcification-related artifacts. Because of the difference in the acoustic impedance of abdominal soft tissue and calcifications, the ultrasound waves are strongly reflected at the surfaces of calcifications and aortic areas that are situated in the “shadows” of calcifications. Since the ultrasound waves fall at a sharp angle on the lateral segment of the aortic wall, the calcification causes a strong beam scattering resulting in an inadequate imaging of the lateral aortic circumference. The ventral and dorsal portions of the aortic wall, on the other hand, reflect the echoes attenuated on the calcifications and enable a correct measurement of the anterior–posterior diameter with high reliability. The 4D-US image is created by summing the 3 spatial planes with the changes on the time axis as a fourth component. The inadequate mapping of the target structure at one level causes artifacts in the 4D model, which in turn affects the reliability of the circumferential strain assessment.

The ICC assumes values between -1 and $+1$. Negative values suggest that the variation within the repetition measurements of the same subject is greater than the variation between the different measured subjects. The deviations in the repeated measurement of the same subject are so large that the actual differences between the subjects can no longer be verily recognized due to the uncertainty of the measurement. Positive values, on the other hand,

indicate that the variation between the subjects is greater than the variation in the repeated measurements of the same subject, that is, that the differences between the subjects can be determined by the measurements. The ICC does not depend solely on the quality of the reproducibility of the measurement as it weighs the variability between the subjects against the variability of the repeated measurements both low variability of the repeated measurements of the same subjects (ie, a high reproducibility) and large differences between the subjects (ie, the selection of the study’s collective result in good ICC values). The circumferential strain shows great differences between the atherosclerotic and the non-atherosclerotic subject collective. The non-atherosclerotic subjects are characterized not only by a higher strain but also by a high interpatient value variability, whereas the atherosclerotic subjects all have very similar strain values. Thus, the repeated measurements still allow a distinction between the different subjects with small differences. From this point of view, we are, at present, able to carry out population-based analysis, but the method does not enable us to follow the patient-individual processes describing the deformation of the aortic wall in the same patient and consecutive examination. The intrapatient variability may be namely higher than the temporal development of the kinematic changes in the aortic wall. The need for further development of the proposed strain analysis should concentrate on the maximum reduction of the observer-related variability by multiple masking of the same recording and the formation of the averaged

Table 6. Impact of Examination Experience on the Reliability of 4D-US Presented as an Intraclass Correlation Between the Measurements for all Strain Components for Non-atherosclerotic, Atherosclerotic aorta, and Both the Groups

Parameter	ICC[observer]		
	All Patients	Non-atherosclerotic Aorta	Atherosclerotic Aorta
Maximum diameter	0.00	0.00	0.00
Circumferential strain	0.00	0.01	0.03

Table 7. Impact of Masking Experience on the Reliability of 4D-US Examination Presented as an Intraclass Correlation Between the Measurements for All Strain Components for Non-atherosclerotic, Atherosclerotic aorta, and Both the Groups

Parameter	ICC[masking]		
	All Patients	Non-atherosclerotic Aorta	Atherosclerotic Aorta
Maximum diameter	0.03	0.03	0.04
Circumferential strain	0.00	0.00	0.00

model. The next step would also be the formation of an averaged model from the 2 studies of the corresponding aortic segment.

The maximum aortic diameter can be measured with acceptable deviations even by examiners with varying experience in 2D ultrasound.¹⁰ Ghulam et al observed a good agreement among ultrasound technicians (mean difference -0.08 mm, LoA, -3.17 ; 3.00 mm) by the AAA diameter sizing with the 2D- and 3D-US techniques. With regard to interrater reproducibility, they concluded 3D-US to be superior to 2D-US.¹⁸ For the first time, we were able to assess the reproducibility of 4D-US between the experienced observer and medical student without ultrasound experience. The variance of the measured parameters was extremely low between the examiner and the student for both the maximum diameter determination and the circumferential strain. In our study, the experience of the observer in performing the examination itself and in the post-processing of the recordings shows no significant influence on the overall reliability of the 4D-US, and this holds good for a wide range of clinical applications.

Conclusion

The 4D-US enables an adequate determination of the maximum aortic diameter regardless of calcification grade with a high agreement with 2D-US. The assessment of circumferential strain can be performed with 4D-US with very high reliability and low intra- and interobserver variability. However, calcifications can reduce the accuracy of the strain evaluation in

atherosclerotic aortae. In any case, the experiences of the observer in performing the examination and post-processing the images do not significantly affect the final result. Thus, the 4D-US opens up new avenues for the use of kinematic parameters in the functional diagnostic of the infrarenal aorta.

Limitations

The 4D-US shows high reliability in determining the maximum diameter and circumferential strain in vessels with few tons of calcifications. However, the calcifications increase the variability of the circumferential strain due to insufficient imaging at the lateral segments of the aortic circumference. The adjustment of the registration angle does not lead to any significant improvement in image quality. Obesity has no negative effect on image quality because the phased array probe with a transmission frequency of 4–4.5 MHz allows sufficient penetration depth. In this case, the fat tissue serves as a standoff distance with uniform propagation of the ultrasound signal and allows homogeneous imaging of the aorta. On the other hand, given the short distance between a probe and aorta in slim patients, we could observe echo-rich artifacts directly under the abdominal wall with difficult registration of the anterior and anterolateral segments of the aorta. To avoid movement of the artifacts, the examination was carried out with apnea of approx. 10 s after expiration. The pause of breath especially bothered the patients with a chronic obstructive disease and necessitated repeated examination.

Table 8. ROC Analysis for Maximum Diameter and Circumferential Strain in the Non-atherosclerotic and Atherosclerotic Aorta for Both the Groups ($n = 19$)

	Optimal Cutoff	Total Misclassification Rate	AUC	95% CI for AUC	p Value for DeLong Test
Maximum diameter (mm)	18.54	0.35	0.58	(0.29; 0.86)	>.05
Circumferential strain (%)	4.22	0.00	1.0	(1.0; 1.0)	<.001

Meteorism was not a disruptive factor in 4D-US even without any special preparation of the patient being carried out.

Acknowledgment

This work was funded by the Adolf Messer Foundation and the Ministry of Higher Education, Research and the Arts (Hessisches Ministerium für Wissenschaft und Kunst, Hesse, Germany), and their support is gratefully acknowledged. The authors appreciate the contribution of Prof. Eva Herrmann from the Institute of Biostatistics and Mathematical Modeling, Goethe University for the constructive suggestions and critical comments regarding the statistical evaluation of the 4D-US reliability. Open Access funding enabled and organized by ProjektDEAL.

References

1. van Disseldorp EMJ, van Dronkelaar JJ, Pluim JPW, van de Vosse FN, van Sambeek MRHM, Lopata RGP. Ultrasound based wall stress analysis of abdominal aortic aneurysms using multi-perspective imaging. *Eur J Vasc Endovasc Surg* 2020; 59:81–91.
2. Wittek A, Karatolios K, Fritzen CP, et al. Cyclic three-dimensional wall motion of the human ascending and abdominal aorta characterized by time-resolved three-dimensional ultrasound speckle tracking. *Biomech Model Mechanobiol* 2016; 15:1375–1388.
3. Seo Y, Ishizu T, Atsumi A, Kawamura R, Aonuma K. Three-dimensional speckle tracking echocardiography. *Circ J* 2014; 78:1290–1301.
4. Rong LQ, Kim J, Gregory AJ. Speckle tracking echocardiography. *Curr Opin Cardiol* 2019; 35:1.
5. Karatolios K, Wittek A, Nwe TH, et al. Method for aortic wall strain measurement with three-dimensional ultrasound speckle tracking and fitted finite element analysis. *Ann Thorac Surg* 2013; 96:1664–1671.
6. Derwich W, Wittek A, Pfister K, et al. High resolution strain analysis comparing aorta and abdominal aortic aneurysm with real time three dimensional speckle tracking ultrasound. *Eur J Vasc Endovasc Surg* 2016; 51:187–193.
7. Wittek A, Derwich W, Fritzen C-P, Schmitz-Rixen T, Blase C. Towards non-invasive in vivo characterization of the pathophysiological state and mechanical wall strength of the individual human AAA wall based on 4D ultrasound measurements. *ZAMM Zeitschrift Fur Angew Math Und Mech* 2018; 98:2275.
8. Wittek A, Blase C, Derwich W, Schmitz-Rixen T, Fritzen C-P. Characterization of the mechanical behavior and pathophysiological state of abdominal aortic aneurysms based on 4D ultrasound strain imaging. *Proc SPIE Int Soc Opt Eng* 2017; 10333:1033303.
9. Derwich W, Wittek A, Hegner A, Fritzen CP, Blase C, Schmitz-Rixen T. Comparison of abdominal aortic aneurysm sac and neck wall motion with 4D ultrasound imaging. *Eur J Vasc Endovasc Surg* 2020; 60:539–547.
10. Beales L, Wolstenhulme S, Evans JA, West R, Scott DJA. Reproducibility of ultrasound measurement of the abdominal aorta. *Br J Surg* 2011; 98:1517–1525.
11. Bredahl K, Taudorf M, Long A, et al. Three-dimensional ultrasound improves the accuracy of diameter measurement of the residual sac in EVAR patients. *Eur J Vasc Endovasc Surg* 2013; 46:525–523.
12. Potter E, Marwick TH. Assessment of left ventricular function by echocardiography. *The case for routinely adding global longitudinal strain to ejection fraction. JACC Cardiovasc Imaging* 2018; 11:260–274.
13. Bihari P, Shelke A, Nwe TH, et al. Strain measurement of abdominal aortic aneurysm with real-time 3D ultrasound speckle tracking. *Eur J Vasc Endovasc Surg* 2013; 45:315–323.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–174.
15. Chen G, Taylor PA, Haller SP, et al. Intraclass correlation: improved modeling approaches and applications for neuroimaging. *Hum Brain Mapp* 2018; 39:1187–1206.
16. Wanhainen A, Verzini F, Van Herzelee I, et al, eds. Editor's choice – European Society for Vascular Surgery (ESVS) 2019 clinical practice guidelines on the management of abdominal aorto-iliac artery aneurysms. *Eur J Vasc Endovasc Surg* 2019; 57:8–93.
17. Long A, Rouet L, Debreuve A, et al. Abdominal aortic aneurysm imaging with 3-D ultrasound: 3-D-based maximum diameter measurement and volume quantification. *Ultrasound Med Biol* 2013; 39:1325–1336.
18. Ghulam QM, Kilaru S, Ou SS, Sillesen H. Clinical validation of three-dimensional ultrasound for abdominal aortic aneurysm. *J Vasc Surg* 2020; 71:180–188.
19. Bredahl K, Sandholt B, Lönn L, et al. Three-dimensional ultrasound evaluation of small asymptomatic abdominal aortic aneurysms. *Eur J Vasc Endovasc Surg* 2015; 49:289–296.