*Original Research*

ULTRASOUND

Ultrasound
2019, Vol. 27(3) 148–155

DOI: 10.1177/1742271X19830742
journals.sagepub.com/home/ult
SAGE

# A precise, reproducible method for measuring ultrasound probe slice thickness using a Gammex 403 phantom

**Steven J Jackson and Stephen Russell**

## Abstract

A precise, reproducible method for measuring ultrasound probe slice thickness has been developed for linear array ultrasound probes. The method used a custom built jig to draw the probe along the surface of a Gammex 403 phantom, with the image plane parallel to the filaments within the phantom. Still images at 0.5 mm intervals are saved for post-processing using in-house software. Slice thickness measurements with a precision of 0.1 mm are obtained. The method was shown to give reproducible estimates of probe slice thickness at several depths to within 0.4 mm during repeat tests. The method was able to provide information about the slice thickness of different sections of the probe face. It is expected that the method can quantify changes in probe performance due to lens wear or replacement over time that may elude both in-plane and in-air reverberation-based tests. A total of 18 linear probes were tested across eight centres, including six specialist vascular ultrasound centres.

## Keywords

Ultrasound beam, slice thickness, Gammex Phantom, quality assurance, physics

Date received: 11 May 2018; accepted: 15 January 2019

## Introduction

The B-mode imaging performance of an ultrasound scanner has a large influence on its diagnostic capability. Many aspects of probe and scanner design have an effect on the imaging performance. Quality assurance (QA) of ultrasound systems is recommended to ensure reliability of results and to check for deterioration in performance.[1] QA tests routinely make use of one or more tissue mimicking test objects or phantoms.

The spatial resolution of an ultrasound system is an important measure of imaging performance. The ultrasound beam has a characteristic resolution in three directions: axial, lateral (both in-plane) and elevation or slice thickness (out of plane). The ultrasound beam is focused in-plane by manipulating the timing of ultrasound pulses transmitted and received at individual elements of the transducer array. The beam is focused in the elevation direction by a cylindrical acoustic lens bonded to the probe face. Slice thickness and its variation with depth are determined by the elevation aperture of the probe face and the intrinsic focal depth of the acoustic lens.

Axial and lateral resolution measurements are well established as part of routine QA using a simple, calliper-based method such as the one described in IPEM Report 102.[2] Changes in these parameters over time could indicate problems with the beam forming or reconstruction software.

Several techniques exist to measure the slice thickness resolution; these will be reviewed shortly. Changes in slice thickness are expected, if the shape or thickness of the acoustic lens is altered from its original specification. This is known to occur primarily through probe

The Christie NHS Foundation Trust, Manchester, UK

**Corresponding author:**
Steven J Jackson, The Christie NHS Foundation Trust, Manchester M20 4BX, UK.
Email: Steven.Jackson@christie.nhs.uk

cleaning with wipes that have some abrasive action. A change in the slice thickness of the beam has the potential to affect diagnosis because smaller lesions may be obscured by the partial volume effect in the elevation direction if slice thickness is larger than expected.

In many QA programs, the only parameters measured that depend on the shape of the lens (and therefore affect the slice thickness) are the low echo penetration depth (the depth to which speckle structures can be seen) and the appearance of the in-air reverberation pattern. A reduction in low echo penetration may be caused by beam widening in the elevation direction at depth, whilst non-uniformity of the in-air reverberation pattern is characteristic of localised wear or damage to the lens. Neither measurement gives direct information about the slice thickness itself. A recent survey of ultrasound probe conditions found 39 probes out of 219 with detectable lens damage or in-air reverberation non-uniformity that can be expected to impact the slice thickness of the probe.[3]

## Established and recent techniques for measuring ultrasound beam slice thickness

A specialist phantom with a diffuse reflective plane positioned at 45 degrees to the vertical is often used to measure slice thickness at arbitrary positions along the beam.[4,5] The axial length of the visible plane on the resulting B-mode image gives an approximate measure of slice thickness at the depth of the plane section imaged. With careful positioning, the system is capable of reproducible results, but the requirement for a specialist phantom for one QA measurement prohibits its inclusion in a regional service.

The Skolnik method uses a probe orientated at 45 degrees to the phantom walls to image known filaments within a phantom.[6] The lateral extent of each filament imaged with this probe orientation is equal to the slice thickness at that depth. An acceptable initial tolerance for this, and the in-plane calliper-based resolution tests mentioned above, is ±40%, suggesting there is room for improvement in the reproducibility of the measurement.[2] Further, the method is not independent of lateral resolution.

The resolution integral method associated with the Edinburgh pipe phantom produces a single metric of ultrasound grey-scale imaging performance based on the visibility or otherwise of several anechoic pipes with depth within a specially created phantom.[7] The single metric produced by the method is dependent on the slice thickness of the ultrasound beam with depth, but the slice thickness obtained is one of a discrete set of values, corresponding to set width of pipes within the phantom.

Another specialist phantom allows a measurement of both lateral and elevational slice thickness as a continuous (pixel by pixel) function of depth.[8] The phantom comprises two different materials arranged in a $2 \times 2$ chequerboard pattern. The experimental method relies on the incremental motion of the probe across the boundary between two phantom materials, with up to 0.1 mm resolution, and a subsequent measurement of the step change in signal between the regions.

A method reported recently involves the capture of multiple images, spaced 1 mm apart, of a phantom with multiple crossed filaments. The method allows an interpolated beam profile to be constructed out of the captured images in both the lateral and elevation directions simultaneously.[9] This paper further demonstrated the importance of the acoustic lens by comparing two images of a random void phantom, one acquired with an intact lens and one with a defective lens. (The random void phantom comprises open pore foam, with random pore sizes filled with degassed saline solution.) The amount of fine detail in the random void phantom that is obscured by the defective lens is striking.

Recent methods have similarities to our work in that they adopt an incremental image capture and post-processing approach. However, both require a specialist phantom that precludes inclusion in a regional QA service at present.

## Motivation and aim

A recent trend is to replace acoustic lenses rather than entire probes when poor performance or significant wear is discovered. This may produce clinically relevant changes in probe performance that are not well quantified. A precise, routinely available measure of the beam slice thickness would allow the tracking of changes in lens properties due to wear, or quantification of any difference before and after lens replacement.

The aim of this study was to demonstrate a precise, reproducible method that can be used routinely to measure slice thickness across the imaged field at a range of depths using a standard 2D imaging phantom. The technique aims to allow quantification of changes in measured slice thickness at different stages of a transducer's working life.

## Method

A bespoke jig was built that allowed a linear array ultrasound probe to be moved along the surface of a Gammex-RMI Model 403 phantom via a screw thread mechanism. The Model 403 phantom is widely used in our region and contains thin filaments in a vertical ladder arrangement, embedded within tissue mimicking

material at depths of 1 cm, 2 cm and in 2 cm intervals thereafter down to 16 cm. The probe was positioned with its scan plane parallel to the plane of the ladder of filaments. The filaments appeared in the ultrasound image and disappeared as the scan plane approached and then passed through the plane of the filaments. Figure 1 shows the phantom, jig and probe, with arrows indicating the motion of the probe and a ring around the filaments.

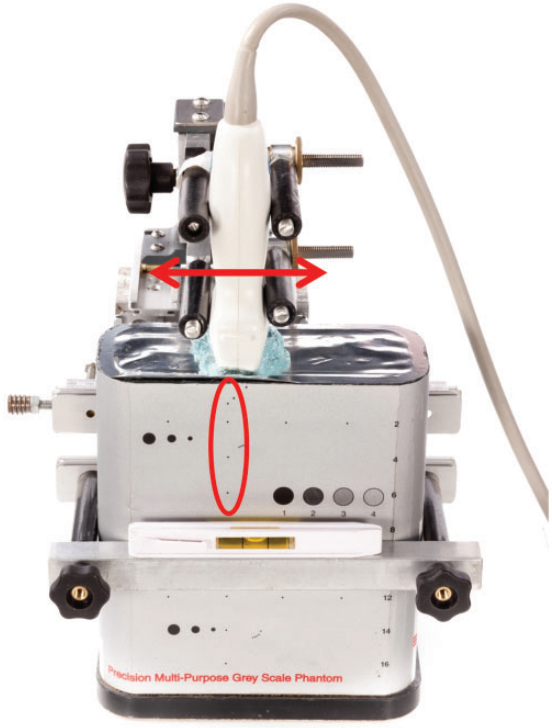


**Figure 1.** Phantom, jig and probe setup. The double headed arrow shows the direction of probe movement. The ring shows the wires in the phantom used in the test.

The probe was initially offset from the filaments such that none were visible in the image. It was then moved along the phantom in 0.5 mm increments determined by a millimetre precision ruler and fiducial needle system incorporated into the design of the jig. After every 0.5 mm, an image was captured and stored for post processing. Around 15 images were required to fully visualise all filaments across the phantom. Three such passes per probe were collected during each data acquisition session, for example the first left-to-right, the second right-to-left and the final one left-to-right again. Around 15 minutes was required to set up the jig, save the required images and export them for offline post processing.

An in-house MATLAB script was used for all offline post-processing (MATLAB, R2013b, The MathWorks, Inc., Natick, MA, United States). For each probe position, a composite of the three saved images was formed by assigning a value to each image pixel equal to the average of the corresponding pixel values in the saved images. A composite image is desirable as an alternative to averaging the results of three individual tests as it simplifies extended analysis, as discussed below.

For each composite image, the presence or otherwise of the filament at its expected position was automatically recorded by the script. To do this, a region of interest was placed across the image and an average of the information across each row of pixels was taken to produce a one-dimensional (1D) profile of image brightness versus depth, as shown in Figure 2.

A peak occurred in the 1D profile when the ultrasound beam was wide enough to insonate the respective filament from its current longitudinal position. Peaks were identified as places where the maximum pixel

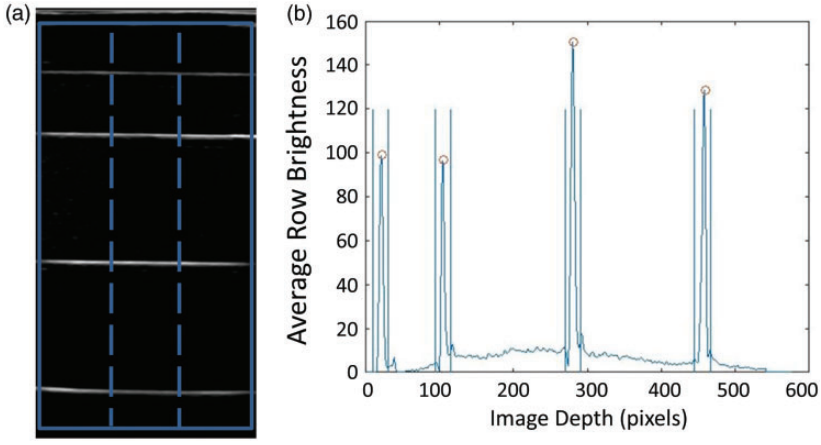


**Figure 2.** (a) Example composite image obtained at one probe position. The dotted line overlay showing the separation of the method into three regions of the probe face for an extension to the test. (b) Example 1D profile obtained by averaging the pixel values in each row of (a). Each detected peak corresponds to a visible wire in the phantom for this probe position.

value within a 20 pixel region covering each known filament depth in the 1D profile was greater than 15 units above background signal for each filament. Each image pixel has a value between 0 and 255, with zero indicating no echoes received from the test object at that location and 255 representing a strongly reflective feature, depending on the choice of gain. Background was defined as the mean of 10 pixels surrounding each 20 pixel region known to contain a filament, five directly above and five directly below. The mean background signal was subtracted from the peak signal as a form of background correction.

For each known filament depth, the amount of background corrected peak signal with longitudinal position was collated and plotted. A polynomial fit to the data was calculated and overlaid, as shown in Figure 3. The slice thickness at that depth was defined as the full-width half maximum of the polynomial fit. Slice thickness measurements were calculated to 0.1 mm precision. The post-processing script outputted results in under 10 minutes, including time required to appropriately sort the initial images.

An extended method was devised that applied the above method to the left hand, central and right hand regions of the probe face, as shown in Figure 2(a). It was anticipated that this extension would allow an investigation into different levels of wear across the probe face.

A total of 18 linear probes were tested across eight centres, including six specialist vascular ultrasound centres. The scanner settings chosen were minimally adapted clinical pre-sets for vascular imaging. The depth was altered during different acquisition sessions to enable visualisation of four (or five) available filament depths of 1, 2, 4, 6 (and 8) cm within the phantom for linear probes with minimum frequency of 3 MHz, and three filaments at depths of 1, 2 and 4 cm for probes with minimum frequency of 4 MHz or greater. The in-plane focus position was retained from the clinical pre-sets, and if automatically adjusted when the depth was extended for the tests, the new focus position was retained. See Table 1 for further details of the probes used and supplementary material for individual measurements from the three most common probe designs tested.

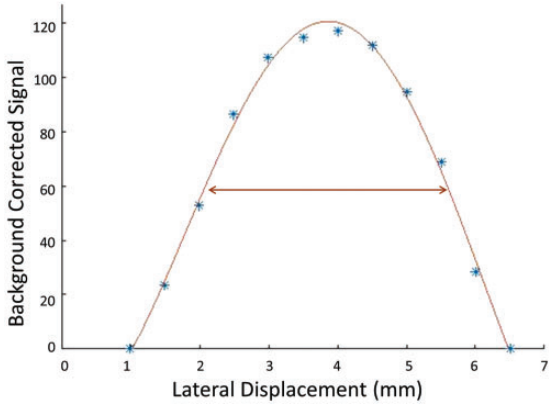


**Figure 3.** Example probe slice profile obtained using the described method. Asterisks show the background corrected signal obtained for one wire in the phantom across all composite images for the tested probe. The full width half maximum (FWHM) of the curve fitted through the data points was defined as the slice thickness at this wire depth for the probe.

## Tests completed

### *Test 1: Reproducibility*

Three probes were tested on two occasions with scan parameters set to reproducible values. Probe A (Siemens 14L5) remained in the jig between tests carried out one after another. Probe B (Philips L9-3) was removed from the jig, then repositioned and tested 20 minutes later. Probe C (Siemens VFX13-5) was removed from the jig, and the jig removed from the phantom, with everything reassembled and re-tested the next day. Measurements from 10 filaments were used for this test.

### *Test 2: Is calculating slice thickness from a composite image of three frames for each probe position equivalent to taking the average from three individual passes along the phantom?*

The script was extended to produce a slice thickness measurement for each of the three longitudinal passes

**Table 1.** Details of the manufacturer names for the linear (L) probes included in the study.

| Probe | Manufacturer | Number of probes imaged |
|---|---|---|
| L12-3 | Philips | 6 |
| L9-3 | Philips | 4 |
| L9-3U | Mindray | 4 |
| LM16-4U | Mindray | 1 |
| L12-4 s | Mindray | 1 |
| 14L5[a] | Siemens | 1 |
| VFX13-5[a] | Siemens | 1 |
| | Total | 18 |

[a]A probe that was used for the reproducibility test only.

across the phantom individually. The mean of these three measurements was compared with the slice thickness measurement calculated from the composite image. It was expected that there would be no difference between the mean slice thickness of the three separate passes across the phantom and the slice thickness calculated from the composite image. Measurements from 64 filaments across 16 probes were used for this test.

### *Test 3: Can differences in slice thickness across the surface of the probe be quantified?*

For a probe with a uniform acoustic lens, there should be no difference between slice thickness measurements in the left, right and central regions of the image and that measured using the full width. In the case where a filament is visible in only the left hand part of an image, that filament may not produce a detectable peak in the 1D average profile across the whole probe face, whilst producing a peak in a 1D profile obtained from only the left hand third of the image.

The extended method described above and shown via the dotted lines in Figure 2(a) was applied to 64 filaments across 16 probes. The mean difference between the slice thicknesses measured using the full field of view and those obtained in the left hand, central and right hand regions of the probe face was calculated.

## Results

Bar charts were prepared for each probe tested to show the measured slice thickness at each filament depth, as shown in Figure 4. Such charts can be used as baselines for future tests of the probes. The error bars are standard deviation of the slice thickness measurements in three regions across the probe face, as described in test 3 above.

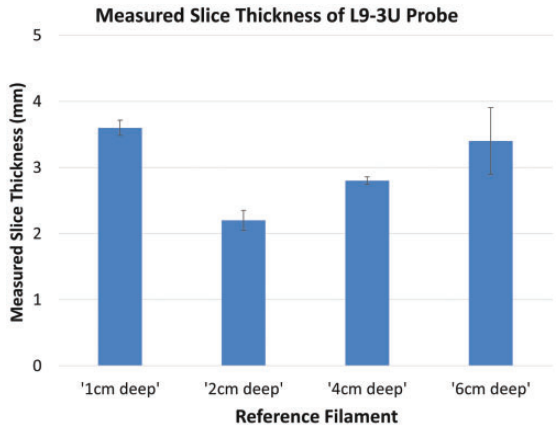


**Figure 4.** Results of the test for one probe. The bars showing the measured slice thickness at the given depths, the error bars are the standard deviation of the measurements obtained by performing the test across three regions of the probe face.

### *Test 1: Reproducibility*

A Bland-Altman plot of pairwise difference in measured slice thickness between test and re-test versus the average of the two measurements is shown in Figure 5, with the usual limits of agreement included. The coefficient of repeatability was calculated.

The coefficient of repeatability is defined as the value below which the absolute difference between two values measured with different methods may be expected to lie, with 95% confidence.

A two-tailed paired t-test for difference in the mean slice thickness yielded $p = 0.51$, and the null hypothesis of no difference between the test and the re-test was accepted. The coefficient of repeatability was 0.4 mm, suggesting that 95% of data acquired with a test and re-test will differ by less than 0.4 mm using this method.

### *Test 2: Is taking a composite image justified?*

A Bland-Altman plot of pairwise difference in measured slice thickness between the composite image and average of three passes across the phantom versus the average of the two measurements is shown in Figure 6, with the usual limits of agreement included.

A two-tailed paired t-test for difference in the mean slice thickness obtained from a composite image and from an average of three measurements yielded $p = 0.94$, and the null hypothesis of no difference between methods was accepted. The coefficient of repeatability was 0.10 mm, suggesting that 95% of

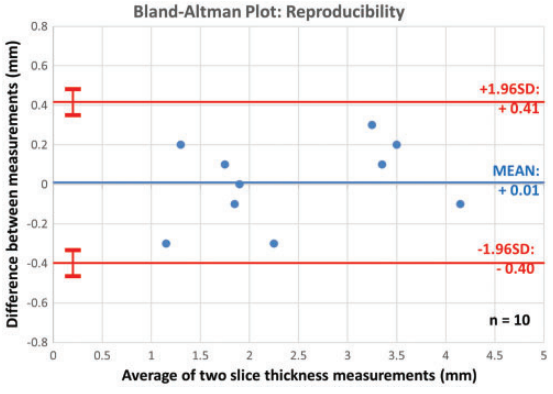


**Figure 5.** Bland-Altman plot showing the difference in measured slice thickness between a test and re-test for 10 wires across three probes.

slice thickness measurements obtained with the composite image method and with an average of three tests on individual passes across the phantom will differ by less than 0.1 mm.

### *Test 3: Can differences in slice thickness across the probe face be quantified?*

Bland-Altman plots of pairwise differences between the slice thickness measurement obtained from the full field of view and each of three sub-regions of the composite image versus the average of the two measurements are shown in Figure 7, with the usual limits of agreement included.

The maximum absolute difference between slice thickness measured in each sub-region and the full field of view (across all 64 filaments) is shown in Table 2, along with the results of two-tailed paired t-tests for the difference in mean slice thickness obtained from each sub-region versus the full field of view and the calculated coefficients of repeatability for this investigation.


**Figure 6.** Bland-Altman plot showing the difference between the slice thickness calculated using the composite image technique and an average of three passes across the phantom, for 64 wires across 16 probes.

## Discussion

### *Reproducibility*

The results suggested there is no statistically significant difference between a test and re-test of three probes, and the method is therefore reproducible. The intrinsic accuracy of the method was not tested but was not thought to be necessary for the purposes of the study, because the method is simply intended to track changes in performance. The coefficient of repeatability was 0.4 mm, which is of the order of the incremental distance between individual frame acquisitions and represents a small percentage of the largest slice thicknesses measured with the method. The coefficient of repeatability metric is inversely proportional to square root of sample size and the sample size was small.

### *Test 2*

The results suggested performing analysis only on the composite image is statistically no different from performing three individual tests and averaging the results. This justified the creation of a composite image and allowed the three images obtained at each probe position to be analysed in one image.

The composite image contains two frames from passes in one direction and only one frame from a pass in the opposite direction. There remains the possibility that the method is dependent on the direction the probe first passes across the phantom, due to the known asymmetric gripping of the probe by the jig.

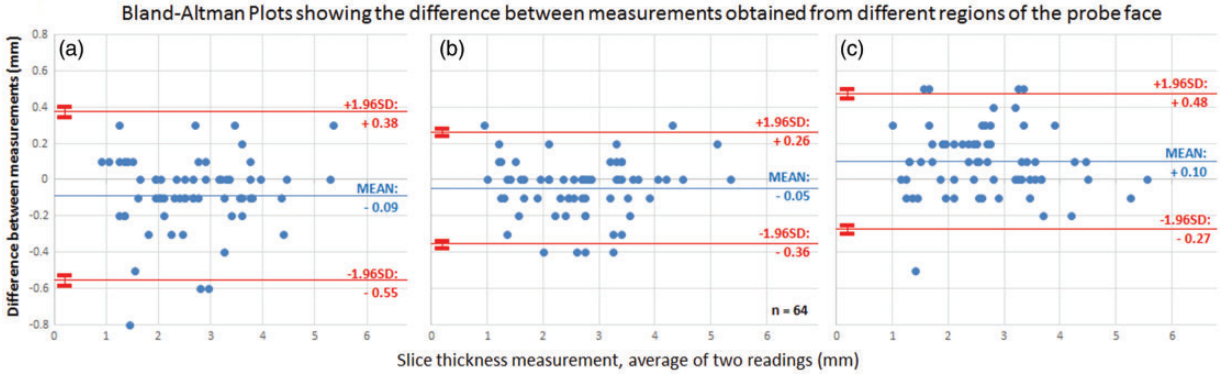


**Figure 7.** Bland-Altman plots showing the difference between slice thickness measured via three regions of interest (ROI) across the probe face versus the slice thickness measured across the whole probe face: (a) the left hand third of the image, (b) the central third of the image, (c) the right hand third of the image (ROI's as shown in Figure 2(a)).

**Table 2.** Statistical analysis of the differences between the slice thicknesses calculated using the full field of view composite image and using left hand, central and right hand pillars across the image, respectively.

| (n = 64) | Left hand region | Central region | Right hand region |
|---|---|---|---|
| Maximum absolute difference (mm) | 0.7 | 0.4 | 0.6 |
| Two-tailed t-test for difference in mean | p = 0.15 | p = 0.02 | p = 0.03 |
| Coefficient of repeatability (mm) | 0.4 | 0.3 | 0.4 |

Preliminary investigation could not detect any difference, but the sample size was too small for inclusion.

### *Test 3*

The Bland-Altman plots in Figure 7 show noticeable differences between the slice thicknesses measured using different sub-regions of the probe face versus the probe face as a whole. Table 2 shows that there were statistically significant differences between slice thicknesses obtained for the central and right hand regions of the probe face, when compared against the composite image-derived slice thickness across the whole probe face. This suggested the method was able to measure different slice thicknesses across the probe face and therefore provide information relating to lens changes. The values obtained for the maximum absolute difference between the regions and the full probe face across the three regions further highlighted this potential.

The limits of agreement of the Bland-Altman plots (the ± 1.96 times standard deviation lines) give evidence that the edges of ultrasound probes are more affected by wear than the central parts of the probe face, which could correlate to the wiping of coupling gel off the edges of the probe.

The coefficients of repeatability for these tests were similar to those obtained for the reproducibility test but, crucially the sample size was over six times greater, lending confidence that measured differences between sub-regions and the full field of view is a real effect that can be measured.

The possibility that this measured effect was due to non-orthogonal positioning of the ultrasound probe on the phantom was considered, although every effort was made to ensure this was not the case. It was reasoned that non-orthogonal positioning of a probe on the phantom surface with a uniform lens should give the same result as an orthogonally positioned one (up to a very small trigonometrical correction factor); the only difference in the acquired data would be the frame number in which each filament first appears.

When one end of the probe was lifted from the phantom surface in relation to the other end, this was apparent because the filaments in the resulting image no longer appeared horizontal. This was immediately visible during set up and was eliminated before image acquisition.

Finally, if the probe was positioned such that the handle of the probe and the probe face were not vertically aligned, then the frame in which each filament first appeared in the image would differ to the vertical probe case, but the measured width of the filaments would be the same, up to a very small trigonometrical correction factor.

Further investigation into the physical condition of the acoustic lens would be beneficial in the cases where the greatest differences between one region of the probe face and the full field of view were observed. Measuring slice thickness for three regions of the probe face was attempted for simplicity. An extension to a pixel by pixel slice thickness profile would be achievable and could be investigated.

## Limitations and further work

There were many limitations of the study due to the equipment available. The phantom must have a removable dam in order to accommodate linear probes moving across its surface. Newer models of this phantom do not have this feature. Further, the sole phantom used for the tests appeared to sustain damage due to the design of the jig pushing the probe into the surface of the phantom, such that it developed a split and leaked coupling fluid, although this may have been unrelated. It was not anticipated that the damage would have an effect on the results obtained from the testing, as the only thing potentially compromised was coupling between the probe and the test object and the images were displayed with the expected level of signal.

The Gammex 403 phantom only contained filaments at 1 cm, 2 cm, 4 cm etc., meaning a limited amount of data points were acquired throughout. The Gammex 404 phantom is designed for linear probes and contains filaments spaced at 0.5 cm intervals. This would have been ideal for the study but was unavailable in the region. Further work could

include a continuation of the tests with this phantom, although this would require another phantom to be transported along with existing test equipment in the case of off-site measurements, which was intended to be avoided.

A small number of linear probes were too wide to fit across the phantom. A multi-stage test could be attempted for these probes. Similarly, it was not possible to assess the whole surface of a curvilinear probe using the method. A re-designed jig could accommodate different probe angulations that would allow a curvilinear probe to be tested in segments across its surface, although significantly more time would be required to perform the test.

Further work could include a check of the intrinsic accuracy of the method against needle hydrophone measurements and a comparison of the results obtained using this method with any of the tests described in the Introduction section.

Collecting annual data and observing any trends is the logical next step for this method. The bar chart representation is straightforward to prepare. The inclusion of the standard deviation of the slice thicknesses measured using three regions of the probe face as an error bar gives some information about the level of suspected wear encountered to the probe face, which could also be tracked for trends.

It would be instructive to collect data on the amount of clinical use each probe undergoes in between annual tests in order to estimate the wear and attempt to correlate it to the observed probe slice thickness changes. Some record of the cleaning and composition of the lens would also be informative, as it is known that softer lenses on higher frequency probes can suffer more damage due to cleaning.

## Conclusion

A precise and reproducible method for measuring ultrasound beam slice thickness has been developed for linear ultrasound probes. It is anticipated that the method could be used to quantify differences in wear across the surface of the probe.

### Acknowledgements

We are grateful to Independent Vascular Services for access to their devices and to Patricia Amata for assistance during the project.

### Contributors

SR conceived the study. SJJ designed the study, acquired the data, conceived and implemented the post processing method, analysed the results. SR made suggestions to improve the validation of the method. SJJ wrote up the final version of the manuscript. All authors reviewed and approved the final version of the manuscript.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Ethics Approval

This is a phantom study, no ethical approval or required.

### Funding

The author(s) received no financial support for the research, authorship and publication of this article.

### Guarantor

SR.

### ORCID iD

Steven J Jackson http://orcid.org/0000-0002-9195-0049

### References

1. Dudley N, Russell S, Ward B, et al. BMUS guidelines for the regular quality assurance testing of ultrasound scanners by sonographers. *Ultrasound* 2014; 22: 8–14.
2. Institute of Physics and Engineering in Medicine. *Report 102: quality assurance of ultrasound imaging systems*. York: IPEM, 2010.
3. Dudley N and Woolley DJ. A multicentre survey of the condition of ultrasound probes. *Ultrasound* 2016; 24: 190–197.
4. Goldstein A. Slice thickness measurements. *J Ultrasound Med* 1988; 7: 487–498.
5. Goodsitt M, Carson PL, Witt S, et al. Real-time B-mode ultrasound quality control test procedures. Report of AAPM Ultrasound Task Group No. 1. *Med Physics* 1998; 25: 1385–1406.
6. Skolnick ML. Estimation of ultrasound beam width in the elevation (slice thickness) plane. *Radiology* 1991; 180: 286–288.
7. Moran CM, Inglis S and Pye SD. The resolution integral – a tool for characterising the performance of diagnostic ultrasound scanners. *Ultrasound* 2014; 22: 37–43.
8. Rowland DE, Newey VR, Turner DP, et al. The automated assessment of ultrasound scanner lateral and slice thickness resolution: use of the step response. *Ultrasound Med Biol* 2009; 35: 1524–1534.
9. Doblhoff G, Satrapa J and Coulthard P. Recognising small image quality differences for ultrasound probes and the potential of misdiagnosis due to undetected side lobes. *Ultrasound* 2017; 25: 35–44.