Automated performance assessment of ultrasound systems using a dynamic phantom

J Joy¹, F Riedel², AA Valente³, S Cochran¹ and GA Corner²

¹Institute of Medical Science & Technology, University of Dundee, UK

²Department of Medical Physics, NHS Tayside, UK

³Department of Physics, University of St. Andrews, UK

Corresponding author: Joyce Joy. Email: jjoy@dundee.ac.uk

Abstract

Quality assurance of medical ultrasound imaging systems is limited by repeatability, difficulty in quantifying results, and the time involved. A particularly interesting approach is demonstrated in the Edinburgh pipe phantom which, with an accompanying mathematical transformation, produces a single figure of merit for image quality from individual measurements of resolution over a range of depths. However, the Edinburgh pipe phantom still requires time-consuming manual scanning, mitigating against its routine use. This paper presents a means to overcome this limitation with a new device, termed the Dundee dynamic phantom, allowing rapid set-up and automated operation. The Dundee dynamic phantom is based on imaging two filamentary targets, positioned by computer control at different depths in a tank of 9.4% ethanol–water solution. The images are analysed in real time to assess if the targets are resolved, with individual measurements at different depths again used to calculate a single figure of merit, in this case for lateral resolution only. Test results are presented for a total of 18 scanners in clinical use for different applications. As a qualitative indication of viability, the figure of merit produced by the Dundee dynamic phantom is shown to differentiate between scanners operating at different frequencies and between a relatively new, higher quality system and an older, lower quality system.

Keywords: Ultrasound, resolution integral, quality assurance, dynamic phantom, test object

Ultrasound 2014; **22**: 199–204. DOI: [10.1177/1742271X14549591](https://doi.org/10.1177/1742271X14549591)

Introduction

Ultrasound imaging, now used for approximately 25% of diagnostic scans,¹ is a safe and inexpensive modality whose image quality is still improving through developments in research and industry. Ultrasound imaging is also now used increasingly to guide interventions in real time, where accuracy is important to avoid unnecessary complications.^{2,3} Since a suboptimal scanner could lead to misdiagnosis or inaccurate guidance in the hands of a non-specialist user, action should be taken to ensure that scanner performance remains satisfactory in use.

Different approaches exist for testing parameters such as spatial resolution and dynamic range, using anatomical and geometrical tissue mimicking materials.^{4–6} However, the limitations of these quality assurance (QA) methods are widely recognized.^{7–9} They include difficulty in quantifying results, operator dependence, and the time and thus the cost involved in taking the measurements. The importance of these limitations is increasing as the use of ultrasound imaging by non-specialist users grows, with a corresponding

increase in the number of scanners designed for non-specialist use.

Novel phantoms have been introduced to address some of these issues. One such is the Edinburgh pipe phantom (EPP),¹⁰ which allowed the innovation of the production of a single figure of merit (FoM) for grey scale imaging to improve the comparability of scanners and probes and to allow monitoring of performance with time. The developers of the EPP describe its use to characterize pre-clinical and clinical ultrasound scanners^{10–14} based on analysis involving plotting depth as a function of inverse resolution. They term their FoM a resolution integral (RI),¹⁰ as its mathematical description involves integration of a measurement of resolution over a range of depths

$$RI = \int_0^{\infty} L(\alpha) d\alpha \quad (1)$$

In equation (1), $\alpha = 1/a$, where a = beam width/2, this taking into account lateral and axial resolution and slice thickness; and $L(\alpha)$ = depth range for which the beam width is less than $2/\alpha$. Data calculated in this way have

Figure 1 (a) Schematic diagram of Dundee dynamic phantom (DDP). (b) Prototype realization of DDP for proof of principle

been shown to be consistent and reliable and to correlate well with clinical perception. However, the results still have a subjective basis, as they require manual scanning and judgment of target images, and the procedure remains time consuming.

In the work reported here, we show that it is possible to combine the very useful presentation of a measure of image quality as a single FoM with a less-subjective, automated phantom design, termed the Dundee dynamic phantom (DDP). Presently, the DDP measures only lateral resolution over a range of depths and we have therefore adapted equation (1) to obtain what we term the FoM_{DDP} , a single measure of quality based on lateral resolution resulting from integration over a range of depths with the expression

$$\text{FoM}_{\text{DDP}} = \int_0^{\infty} L(\alpha') d\alpha' \quad (2)$$

where $\alpha' = 1/a'$, $a' = \text{lateral beam width}/2$ and $L(\alpha') = \text{depth range for which the beam width is less than } 2/\alpha'$.

Lateral resolution refers to the minimum separation, perpendicular to the ultrasound beam, between targets in tissue that a scanner can present distinctly as two reflectors.¹⁵ Consider starting with a brightness profile of two identical targets which can be easily resolved. If the targets are moved closer together, their brightness profiles merge and, at a certain point, the distance between the targets is equal to the beam width. Continued movement of the targets makes the intensity profiles overlap, until a separation is achieved beyond which there is no discernible reduction in the superimposed brightness of the two targets and they can no longer be resolved. It is this process that is automated by the DDP, over a range of depths, and which we have taken in this paper as a proxy for overall scanner performance.

Methods

The DDP is based around two 0.3 mm diameter monofilamentary nylon targets held taut and moved by

corresponding pairs of rare earth magnets inside and outside a tank containing a liquid medium (Figure 1(a)). The magnets are used to allow the integrity of the tank to be maintained and to avoid arms protruding outside its spatial envelope. Their position is controlled by two stepper motors, one to move the targets axially, i.e. up and down in the tank, and the other to move them laterally, i.e. together and apart (Figure 1(b)). To allow only two stepper motors to be used, the midpoint between the targets is always in the centre of the tank, with a split-thread screw connected to one stepper motor controlling lateral target separation. Both targets must also always be at the same axial distance, with the other stepper motor moving them simultaneously.

The movement of the stepper motors is controlled by a processor chip, programmed in the C computer language, with communication based on the virtual instrument software architecture (VISA, National Instruments, Berkshire, UK) through a universal serial bus interface. The DDP has four photodetectors to set upper, lower, inner, and outer motion limits, ensuring that the motors stop when the targets reach their upper/lower and inner/outer limits, respectively. The precise positions of the targets are not subject to real-time sensing so their positions are calculated relative to the limits. The axial and lateral motion step sizes are 0.8 and 1 mm, respectively, and the detail engineering of the attachment of the target filaments to the magnets allows a minimum separation below 1 mm. In use, the ultrasound transducer is placed with its face immersed in the liquid medium, the lowest point for a curvilinear transducer, at the targets' upper limit. In this way, depth and separation of the targets are always known accurately, with respect to the limits set in hardware.

To mimic use of an ultrasound system in the clinic fully, the test tank should be filled with liquid tissue mimicking material to match tissue in terms of both speed of sound, c , and attenuation. However, since the main objective of this paper and of the DDP is to demonstrate the feasibility of an automated approach to QA, a simple 9.4% ethanol/water mixture was used, with $c = 1540 \text{ m/s}$ at 21°C .¹⁶ This

Figure 2 Ultrasound images of (a) resolved and (b) unresolved targets

ultrasound medium, lacking attenuation, was deliberately selected to enable easier visualization of targets during the development stages of the DDP and easier understanding of the principle of operation. In use, it was found to have the advantage of providing very high contrast, which aids target identification.

Software control of the stepper motors within a feedback loop is achieved through computer code (LabVIEW, Version 8.5, National Instruments, Berkshire, UK). A frame grabber (IMAQ PCI-1405, National Instruments, Berkshire, UK) is used to obtain images in digital form for processing. The frame grabber used for the results presented here was a low cost item with only a composite video input. This is suitable for most older ultrasound scanners, capturing analogue images from the printer output. For more modern imaging systems, one approach is to tap the signal to the display monitor, using a signal converter to convert from HDMI or VGA as necessary. If this is impossible, an alternative is to use a low cost, high resolution video camera to record data from the screen in a way exactly resembling user operation.

At each position of the targets, an image from the scanner under test is captured, as in the examples shown in Figure 2. In these images, side lobes are visible because of the low attenuation of the ultrasound medium. The captured image is then analysed by the control software to calculate lateral resolution for a given depth. No manual intervention is required for a given scanner setting once the scanner and phantom are set up, the transducer is mounted, and the measurement process launched, thus achieving the key aim of automatic operation.

System setup

Measurements taken with the DDP are divided into three depth ranges, low, medium, and high, according to scanner system settings. Prior to the start of the measurement process, the scanner is optimized manually for each depth setting so that only the two targets can be seen against an almost completely black background. In this way, there is still some subjectivity in the process, but this can be minimized, e.g. as we have done by using maximum target contrast, as shown in Figure 2. Moreover, if scanner settings are maintained from one set of measurements to another, later

set, then direct comparison is possible, to gauge potential performance deterioration.

An alternative approach worth considering would be to record the grey-scale image of a single target at different positions then to measure the full width half maximum values for each position. This has the advantage that it would provide measures of both lateral and potentially, axial resolution, allowing calculation of RI rather than just FoM_{DDP} as in the demonstrator system here. However, it is possible that it would be much more subject to machine settings since a mid-range grey scale, rather than essentially binary image would be needed, and it has more potential for error owing to the particular characteristics of a given filamentary target.

The present DDP operating software takes input from the user to define the range of depths of scanning, taking into account a specific scanner set up. The targets are then moved laterally apart and axially downwards, away from the transducer, by a number of steps corresponding to the user's specification of 'Start Depth' on the front panel and an image is captured and analysed. If the targets are resolved, then they are stepped inwards and the image capture and analysis cycle is repeated until the targets are not resolved or they reach the very small, sub-mm inner limit set by the present DDP hardware. The targets are then moved apart again and downwards one step and the image capture and analysis cycle is repeated, continuing until the user-specified 'Final Depth' is reached. At this point, the set of depths over which the lateral resolution is measured and the corresponding resolutions are saved for further analysis, including integration to transform the multiple individual measurements into a FoM_{DDP} and the test is complete. A flowchart of this process is shown in Figure 3, with user intervention required only at the beginning and to view the results at the end.

Results

Eighteen scanner-transducer combinations were tested in Ninewells Hospital (NHS Tayside, Dundee, UK). The names of the scanners and their host departments have been anonymized in the results but the scanners were categorized to indicate their expected performance

Figure 3 System operation flow chart

range based on three parameters: age, price range, and frequency range. Each test was repeated multiple times to determine if the values for FoM_{DDP} were consistent.

Two scans of $L(\alpha')$ are superimposed on one another in Figure 4 to demonstrate repeatability; this format corresponds to the parameters needed to calculate FoM_{DDP} . Figure 5 shows the results for the 18 scanner-transducer combinations that were tested, classified according to frequency in MHz and transducer geometry, linear (L) or curvilinear (CL). Most importantly, the values for the FoM_{DDP} show an approximately linear relationship with frequency, as expected, and also allow some differentiation of L and CL probes, the latter generally designed for higher frequencies and having a higher FoM.

To further demonstrate the basis for the FoM being somewhat independent of frequency and the potential usefulness of the DDP, two specific ultrasound scanners, A and B, were selected. System A was a relatively old scanner, from the early 1990s, with a 7.5 MHz linear array transducer and System B was from the early 2000s, with a 5 MHz linear array transducer. Figure 6

Figure 4 Plot showing two separate measurements of $L(\alpha')$ versus $1/a$ for a single 9 MHz probe, illustrating repeatability: first scan (blue), second scan (red)

Figure 5 Graph showing the lateral resolution values of the transducer-system combinations under medium depth settings: curvilinear probe (blue); linear probe (red)

Figure 6 Plot of lateral resolution for two transducers of different age: System A – linear array, $f = 7.5$ MHz, from 1990s (red squares); System B – linear array, $f = 5$ MHz, from 2000s (orange circles)

shows $L(\alpha')$ for these two scanner-transducer combinations. It is clear that the newer scanner, System B, maintains resolution better with depth than the older machine, System A, and it is significant that this is achieved with a lower frequency transducer.

Discussion and conclusions

An automated, closed-loop system for QA of ultrasound imaging systems has been realized in the form of the DDP. Although realized as a preliminary proof of principle, this can nevertheless run unattended and requires no manual intervention to test a scanner with particular settings after a relatively short setup procedure, saving time and cost and increasing objectivity. The results, which are highly repeatable, suggest preliminary agreement with expectations, based on a single FoM_{DDP} defined in equation (2) based on principles established from work done elsewhere on the RI.¹⁰

Even though the DDP produced repeatable lateral resolution values for all of the scanner-transducer combinations tested, some drawbacks have been identified which need further consideration. The ultrasound image of the targets in Figure 4 indicates that the motion of one of the targets lags behind the other. This is attributed to ineffective coupling of pairs of magnets inside and outside the liquid medium container and could be resolved by modifying the target mountings, either with a liquid container with a low friction coating or by replacing the magnets with arms holding the targets, albeit that the latter increases the overall physical envelope of the system.

The DDP is intended to offer a way to assess and compare ultrasound imaging systems quantitatively, in support of purchasing decisions, delivery acceptance, and ongoing QA. It has been designed to allow low cost manufacture and to require minimal manual intervention in use. Calculation of a parameter we have termed the FoM_{DDP} to transform multiple individual measurements into a single FoM has been shown to differentiate scanners on the basis of operating frequency and age as expected.

In further work, it is planned to develop the phantom by replacing the magnets with arms holding the targets and allowing the orientation of the targets to be changed to 45° with respect to the transducer to make the results more directly comparable with those of other phantoms used to determine the RI, but with greater automation cutting costs and reducing subjectivity. With these improvements, the method could be developed to test imaging systems rigorously and to set benchmarks.

DECLARATIONS

Competing interests: The authors have no conflicts of interest to declare.

Funding: Joyce Joy is supported by a Scottish Universities Physics Alliance (SUPA) INSPIRE studentship funded through the Scottish Funding Council SPIRIT scheme.

Ethical approval: Not applicable

Guarantor: JJ

Contributorship: AAV, FR, and JJ researched literature and conceived the study. JJ did the data analysis and wrote the manuscript. SC and GAC supervised the work and reviewed the manuscript before submission. All authors reviewed and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

We are grateful for the support for this work from the Scottish Opto-electronics Association and Diagnostic Sonar, Ltd, Livingston, UK.

REFERENCES

1. NHS Imaging and Radiodiagnostic activity in England. National Statistics, <http://www.england.nhs.uk/statistics/wpcontent/uploads/sites/2/2013/04/KH12-release-2012-13.pdf> (2013, accessed 10 September 2013)
2. Corner G, Grant C. *Physics of Ultrasound In: Principles and Practice of Regional Anesthesia*, 4th edn. Oxford, UK: Oxford University Press, 2013
3. Madsen EL, Ernest L. Quality assurance for grey-scale imaging. *Ultrasound Med Biol* 2000;**26**:48–50

4. Dudley NJ, Griffith K, Houldsworth G, et al. A review of two alternative ultrasound quality assurance programmes. *Eur J Ultrasound* 2001;**12**:233–45
5. Sipilä O, Mannila V, Vartanen E. Quality assurance in diagnostic ultrasound. *Eur J Radiol* 2011;**8**:519–25
6. Mannila V, Sipilä O. Phantom based quality assurance measurements in B-mode ultrasound. *Acta Radiol Short Rep* 2013;**2**:PMC C3863967
7. Boyden J. Is quality assurance of ultrasound equipment necessary? – The benefits and objections. *Ultrasound* 2003;**11**:30–32
8. Donofrio NM, Hanson JA, Hirsch JH. Investigating the efficacy of current quality assurance performance tests in diagnostic ultrasound. *J Clin Ultrasound* 1984;**12**:251–60
9. Keeble C, Wolstenhulme S, Davies AG, et al. Is there agreement on what makes a good ultrasound image? *Ultrasound* 2013;**21**:118–23
10. Moran CM, Ellis W, Janeczko A, et al. The Edinburgh pipe phantom: Characterising ultrasound scanners beyond 50 MHz. *J Phys* 2011;**279**:1–7
11. Pye SD, Ellis W, Macgillivray T. Medical ultrasound: A new metric of performance for grey-scale imaging. *J Phys Conf Ser* 1 2004;**187**–92
12. Moran CM, Pye SD, Ellis W, et al. A comparison of the imaging performance of high-resolution ultrasound scanners for pre-clinical imaging. *Ultrasound Med Biol* 2011;**37**:493–501
13. Moran CM, Ellis W, Pye SD, Smart S. Characterising the performance of a high resolution ultrasound scanner for pre-clinical ultrasound imaging. *Proc IEEE Ultrasonics Symp* 2008;**1724**–7
14. MacGillivray TJ, Ellis W, Pye SD. The resolution integral: visual and computational approaches to characterizing ultrasound images. *Phys Med Biol* 2010;**17**:5067–88
15. Hoskins P, Thrush A, Martin K, et al. *Diagnostic Ultrasound: Physics and Equipment*. London: Greenwich Medical Media, 2003:65
16. Martin K, Spinks D. Measurement of speed of sound in ethanol/water mixtures. *Ultrasound Med Biol* 2001;**27**:289–91