IPEM-Translation 2 (2022) 100007

Contents lists available at ScienceDirect

IPEM-Translation

journal homepage: www.elsevier.com/locate/ipemt

Full Length Article

Assessing the performance of 360° ultrasound probes designed for endoanal ultrasound

James Harkin*, Steve Perring

University Hospitals Dorset NHS Foundation Trust, Longfleet Rd, Poole BH15 2JB, United Kingdom

ARTICLE INFO

Keywords:
Endoanal ultrasound
Quality assurance
Quality control
Phantom
Medical physics
Ultrasound
QA
QC
Python
Programming

ABSTRACT

A bespoke phantom has been designed, with clinically relevant features for endoanal ultrasound (EAUS), capable of rigorously assessing the performance of 360° ultrasound probes.

The performance of three, commercially available, anorectal probes, capable of producing both 2D and 3D images, was assessed. One of the probes was also assessed in two states: before failure and after a repair to correct a failure. For each probe the signal to noise ratio (SNR), contrast to noise ratio (CNR), penetration depth, resolution, focus depth, distance accuracy and low contrast object detectability (LCOD) were assessed at varying dynamic ranges, receive gains and operating frequencies. A Python program (SAUQA) was developed to semi-automate the analysis.

In general the measured parameters varied as expected. However, at intermediate receive gains, adjusting the receive gain resulted in the SNR, CNR, penetration depth and LCOD varying in an unexpected manner. The reason for this is not known, but because it was exhibited by all probes it is believed to be related to the ultrasound machine itself and/or an inherent characteristic of the probe design.

The quantitative results suggest that all probes tested offer an effective method of assessing the integrity of the Internal Anal Sphincter (IAS) and the repair of the probe appears to have been successful. However, differences between the probes were observed both quantitatively and qualitatively, with the original probe providing the best results for EAUS.

In light of the results, a recommendation was made, to the EAUS service at University Hospitals Dorset, to adjust the default machine start-up settings for EAUS.

1. Introduction

Endoanal ultrasound (EAUS) is an established technique to review the anatomical condition of the anal muscle complex and is considered the gold standard for evaluating anal sphincter pathology in the investigation of anal incontinence [1].

The introduction of 3D EAUS allowed further opportunity to characterise the severity and extent of anal injury. It has been shown to be effective in identifying small obstetric defects after childbirth [2], and in identifying and characterising fistulae [3]. The internal anal sphincter (IAS) and external anal sphincter (EAS) are close to the probe face. Typically the IAS is the most prominent feature, is <3 mm thick and appears within 5 mm of the probe-tissue interface [4–6]. In order to achieve sufficient resolution, a high frequency ultrasound (US) beam is used (typically ≥10 MHz), which allows for the visualisation of the fine structure of the IAS (Fig. 1).

The EAUS service at University Hospitals Dorset (UHD) is typical of many hospitals in the UK: using a BK Medical® Anorectal 3D 20R3 transducer, in combination with BK 3000 US machine (BK Medical UK, Droitwich, Worcestershire, United Kingdom) [7,8]. The 20R3 probe is a single element, multi frequency transducer, which is cylindrical, with an external diameter of 17 mm. Inside the transducer head, two crystals

---

Abbreviations: CNR, Contrast to noise ratio; CoV, Coefficient of variation; DR, Dynamic range; DS, Dataset; EAS, External anal sphincter; EAUS, Endoanal ultrasound; FOV, Field of view; GSCT, Grey scale contrast target; HF, High frequency; IAS, Internal anal sphincter; LCOD, Low contrast object detectability; LF, Low frequency; QA, Quality assurance; SNR, Signal to noise ratio; TGC, Time gain compensation; TMM, Tissue mimicking material; UHD, University Hospitals Dorset; US, Ultrasound.

* Corresponding author.
E-mail addresses: jameswharkin@gmail.com (J. Harkin), steven.perring@uhd.nhs.com (S. Perring).

https://doi.org/10.1016/j.ipemt.2022.100007
Received 26 May 2022; Received in revised form 29 August 2022; Accepted 31 August 2022

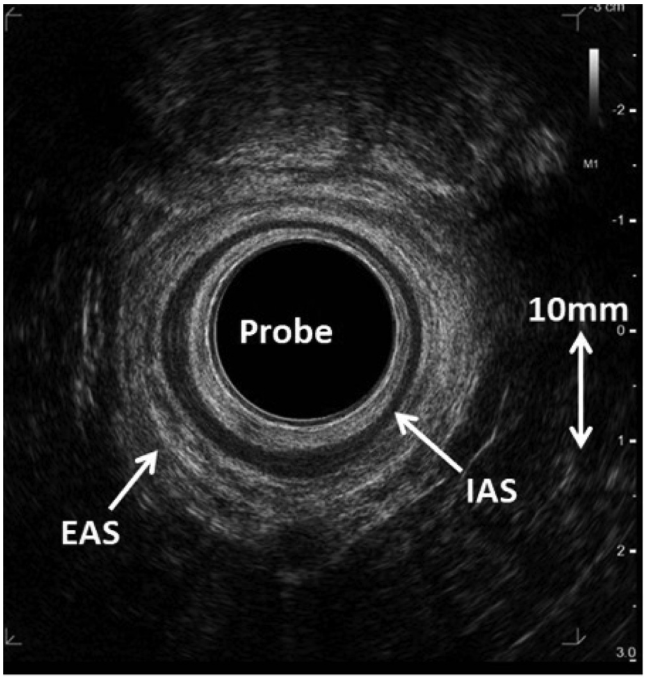


Fig. 1. Typical EAUS image of the anal complex, depicting the IAS and EAS. The IAS shows as a dark band, typically <3 mm thick and within 5 mm of the probe face.

are situated back-to-back. The user has the option of selecting the high frequency crystal (HF mode) or the low frequency crystal (LF mode). HF mode allows for the system to operate at 10 MHz, 13 MHz or 16 MHz and LF mode operates at 6 MHz, 9 MHz or 12 MHz (manufacturer quoted). The transducer head images perpendicular to the probe face and rotates (around the probe's long axis) at a known frequency. This allows for the reconstruction of the images over a 360° field of view (FOV). The crystal assembly can also be stepped along the probe's long axis to produce 3D images. Potential image artefacts are avoided, by ensuring all moving parts are fully encapsulated and do not touch the patient.

At UHD, the performance of US machines used for diagnosis are assessed monthly by the users, using a simple set of quick to perform tests, as recommended by BMUS [9]. Where possible the performance is also quantitatively assessed annually using a test object; following the guidelines in IPEM report 102 [10]. Quantitative quality assurance (QA) phantoms are useful tools for this testing and can allow for the assessment of resolution, penetration depth, measurement accuracy and contrast-detail.

A search revealed the lack of any commercial phantom that would be suitable for assessment of an EAUS probe. Of the commercially available phantoms suitable for inter-cavity US probes, the most appropriate for assessment of EAUS was found to be the CIRS ATS 540 phantom (CIRS, Inc., Norfolk, VA, USA) [11]. However, the central probe well (bore) diameter and features included in the phantom, were unsuitable for application to EAUS. In particular:

- There are no features within 10 mm of the TMM surface, to test at the depth anal features would appear.
- The minimum low contrast feature size was 2mm, which is unlikely to be a challenge for an EAUS machine to image.
- There was no variation in features along the central axis of the phantom, meaning there was no opportunity to test the 3D function of the probe.
- The tissue mimicking material (TMM) used, urethane rubber, is a particularly poor mimicking material at high US frequencies [12].

Due to these limitations, the CIRS ATS 540 was deemed unsuitable for rigorously assessing a 3D 360° EAUS probe and an alternative solution was sought.

A search of the literature did not highlight other groups who had manufactured an in-house solution optimised for EAUS probes. Rodgers et al. [13] have previously assessed the performance of a 3D 360° transvaginal US system. Using an agar phantom mimicking the geometry of the pelvis with brachytherapy needles inserted, they assessed the system's ability to produce geometrically accurate images, with sufficient contrast between the needles and agar TMM. However, their phantom was very specific to brachytherapy and did not contain low contrast targets (mimicking the IAS), suitable for assessing low contrast object detectability, or grey scale contrast targets (GSCTs), suitable for assessing the signal to noise ratio (SNR) and contrast to noise ratio (CNR). Phantoms have also been designed for assessing the performance of intravascular US systems [14–17], however these phantoms are typically designed to assess the systems ability to accurately visualise a vessel and do not contain a combination of high and low contrast structures within the TMM. Another option was the Edinburgh pipe phantom [18], which allows for the assessment of both intravascular [19] and endoscopic ultrasound systems [20]; operating between 2 and 15 MHz. However, while the resolution, depth of field and LCOD can be assessed using this phantom, the phantom does not feature grey scale targets. Also the phantom does not enclose the probe, meaning the full 360° FOV can not be assessed at the same time.

2. Methods

2.1. Phantom design

A bespoke phantom was designed and CIRS (CIRS Inc., Norfolk, VA, USA) were approached to construct the phantom. The bespoke test tool (Fig. 2) was manufactured to the following specification (target depths are defined from the centre of the phantom):

- Donut design, with a bore diameter of 20 mm. The bore is capable of accommodating other similar probes including Pentax J10 gastroscopes (Pentax Medical, Hamburg, Germany) [21], BK Endocavity probes (BK Medical UK, Droitwich, Worcestershire, United Kingdom) [22–24], Olympus Radial Rectal probes (Olympus, Hamburg, Germany) [25] and THD EAUS probes (THD, Correggio, Italy) [26]. Coupling between the probe and tissue mimicking material (TMM) is achieved using Baxter (8010 Zurich, Switzerland) sterile water for irrigation.
- Test features (radially positioned unless otherwise stated):
  - Anechoic targets, of diameter 1 mm, 2 mm and 4 mm, suitable for assessing low contrast object detectability.
  - High Contrast Nylon Monofilament Targets (manufactured by CIRS), of diameter (~76 μm) 0.003" +0.0004", −0.0002", suitable for assessing resolution and distance accuracy in three dimensions.
  - GSCTs of varying relative backscatter with respect to the background, suitable for assessing the CNR and SNR. Relative backscatters of +9.7 dB, +7.0 dB, -5.7 dB and -9.3 dB (±1.5 dB) were chosen to be comparable to the UltraiQ phantom set (Cablon Medical, Leusden, The Netherlands) [27] in use at UHD for routine B-Mode US QA.
  - Uniform background region.
- Zerdine used for background TMM, anechoic targets and GSCTs

Fig. 3 shows an US image of the bespoke phantom.

2.2. Image analysis

The test tool allows for the calculation of multiple, quantitative image quality metrics. With the aim of reducing the time required for image analysis and reducing inter/intra user variability in the results from

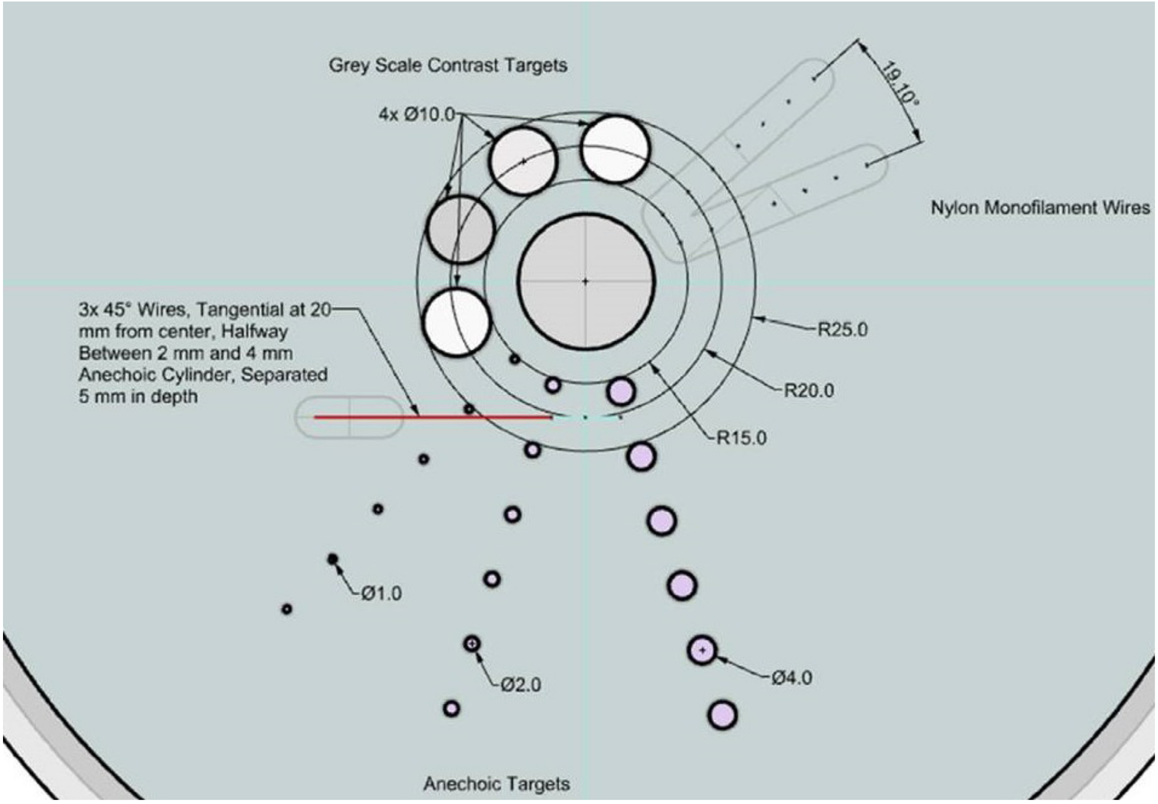
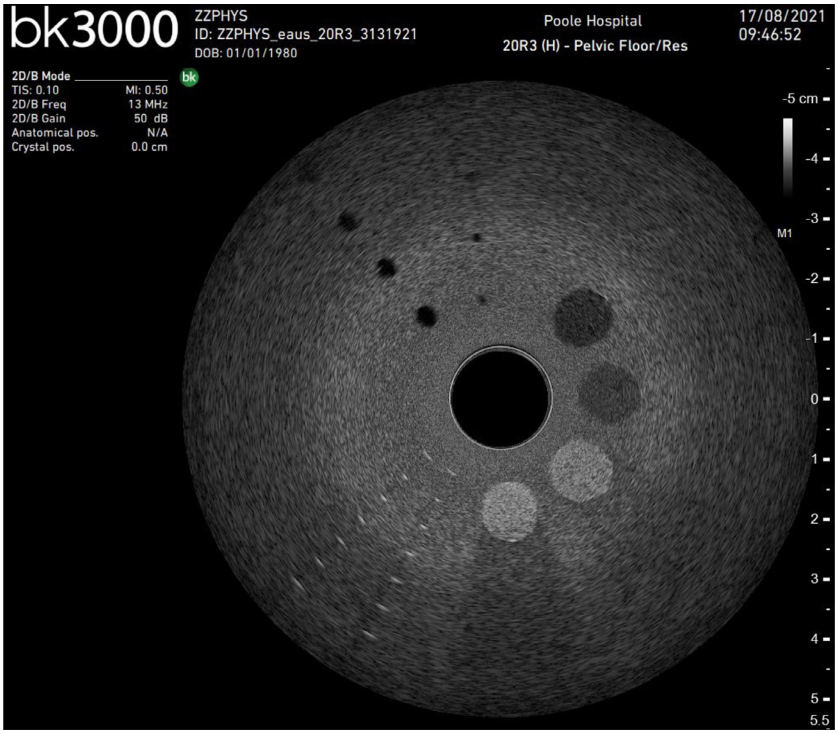


Fig. 2. Axial drawing of EAUS phantom.


bk3000
ZZPHYS
ID: ZZPHYS_eaus_20R3_3131921
DOB: 01/01/1980
Poole Hospital
20R3 (H) - Pelvic Floor/Res
17/08/2021
09:46:52

2D/B Mode
TIS: 0.10 MI: 0.50
2D/B Freq 13 MHz
2D/B Gain 50 dB
Anatomical pos. N/A
Crystal pos. 0.0 cm

Fig. 3. US image of bespoke phantom, acquired using the repaired probe.

the analysis, a semi automated program was developed. This program is referred to as SAUQA [28] and was written in Python 3.8 [29]. SAUQA is only designed to analyse images in one 360° imaging plane and does not produce any image quality metrics for the third dimension.

2.3. Grey scale contrast targets (GSCTs)

The regions used for calculations using the GSCTs are circles of radius 4 mm (80% of the GSCT radius).

Noise in an US image originates from multiple sources. Noise can be modelled as the combination of an additive and multiplicative source. Additive sources include electronic and thermal noise, which exhibit a Gaussian distribution and the multiplicative source (often referred to as "speckle") exhibits a Rayleigh distribution. Speckle is an interference pattern produced when US pulses encounter objects of comparable size to the wavelength of the US pulse [30]. In order to approximate the levels of noise in an image, the standard deviation (σ) in a uniform region is used.

Signal to noise ratio The SNR of a target is calculated using Eq. 1 [31], where $\langle S_i \rangle$ is the average signal of the target of interest and σ_0 is the standard deviation of the background region.

$$SNR_i = \frac{\langle S_i \rangle}{\sigma_0} \tag{1}$$

Contrast to noise ratio The contrast can be defined as the difference in signal between a background region and the target region, where $\langle S_0 \rangle$ is the average signal of a background region and $\langle S_i \rangle$ is the average signal of the target region:

$$Contrast = \langle S_i \rangle - \langle S_0 \rangle \tag{2}$$

The Contrast to Noise Ratio (CNR) of each target is calculated using Eq. 3 [31], where σ_i is the standard deviation of the target of interest.

$$CNR = \frac{\langle S_i \rangle - \langle S_0 \rangle}{\sqrt{\sigma_i^2 + \sigma_0^2}} \tag{3}$$

Penetration depth

For the purpose of this investigation, the penetration depth was defined by the depth at which the mean background intensity had reduced by 6 dB [32]. To calculate the penetration depth (Fig. 4), the average intensity of the GSCTs is assumed to be directly proportional to the attenuation of the target. Using the theoretical attenuations and $\langle S_i \rangle$ for each GSCT, a linear regression is performed, and the Python package SciPy [33] used, to calculate the theoretical average intensity of a target −6 dB below background.

Next, the average intensity of a background region at 20 mm depth is calculated. The background region is then stepped away from the centre of the phantom radially. The distance from the centre of the phantom at which the average intensity in the background region is less than the theoretical average intensity of a target −6 dB below background is specified as the penetration depth.

2.3.1. Nylon monofilament targets

SAUQA asks the user to select the approximate location of the radial nylon monofilament targets (Fig. 5(a)). The centre of the targets is automatically calculated using the region of maximum intensity of a Gaussian blurred image (Fig. 5(b)). The OpenCV-Python [34] function gaussianblur is used to blur the image with a kernel size of (1,1) and a kernel standard deviation of 1. This kernel size was found to be optimal for automatically calculating the specific location of the targets.

Distance accuracy The distance between the centres of the targets in each radial set is calculated. The distance between the targets at the same radial depth is also calculated.

Resolution Both radial and tangential resolutions are calculated for each of the selected nylon monofilament targets (Fig. 5). The rotation angle between the radial axis of the targets and the x axis of the image is calculated using the location of the centre of the target and centre of the image. A 5 mm by 5 mm region of interest around the target is rotated by the calculated angle, which puts the radial dimension in the y axis of the rotated image and the tangential axis in the x axis of the rotated image. The rotated image has the same pixel size and spacing as the original image. Resolution is calculated using 1D slices of the rotated image through the centre. The SciPy [33] function signal.peakwidths is used to calculate the peak width of the 1D slices with the relative height set to 0.5 (equivalent to the −6 dB width) [32].

2.3.2. Anechoic cylinders

The code calculates the CNR for each Anechoic cylinder centre selected by the user. The CNR calculation differs from the GSCT CNR in the ROIs used, so this will be referred to as 'Low Contrast Object Detectability' (LCOD). *Low contrast object detectability* ROIs can be seen in Fig. 6. For an anechoic cylinder of radius R, a circle of radius 80% of the anechoic cylinder radius is used as the signal ROI. The background ROI is a disk centred on the same location as the signal ROI and with internal radius 120% of the anechoic cylinder diameter. The external radius of the background ROI is chosen so the areas of the signal and background ROI are equal.

2.3.3. Data export

The results from the analysis are exported to Excel using the Python module openpyxl.

2.4. Image acquisition

2.4.1. Code performance and reproducibility of results

With the aim of assessing the repeatability and reliability of the calculated image quality metrics, images were acquired with consistent acquisition parameters: time gain compensation (TGC) controls centred, FOV = 110 mm, frequency = 13 MHz, receive gain = 50 dB, DR = 70 dB. 10 images were acquired with the probe held by a clamp stand to ensure that images were acquired in the same location and 10 images were acquired with the probe held by a user. In the latter, the probe was removed from the phantom and reinserted between each acquisition.

The first image of the dataset for the clamp stand was analysed 10 times by SAUQA and also 10 times by a user in ImageJ. SAUQA was also used to analyse the 10 images with the probe held by the clamp stand and the 10 images with the probe held by a user. In each case, the coefficient of variation (CoV) in the image quality metrics was calculated. For the penetration depth calculation using ImageJ, rather than stepping a background ROI radially, the profile of a linear segment was plotted. Also the linear fit of the GSCTs average intensities was performed using the Python package SciPy [33].

2.4.2. Probe comparison

With the aim of assessing if the phantom could rigorously assess the performance of US probes, four data sets were acquired across three 20R3 probes. Each dataset was acquired at the maximum FOV in HF mode (110 mm), and covers the full useable frequency range, dynamic range (DR) and receive gain range of the machine. Unless otherwise stated, the images were acquired with the following acquisition parameters: receive gain 50 dB (median receive gain in HF mode), DR 70 dB (median DR in HF mode) and frequency 13 MHz (median frequency in HF mode). The probes were held by a clamp stand to ensure that, for each data set, images were acquired in the same location.

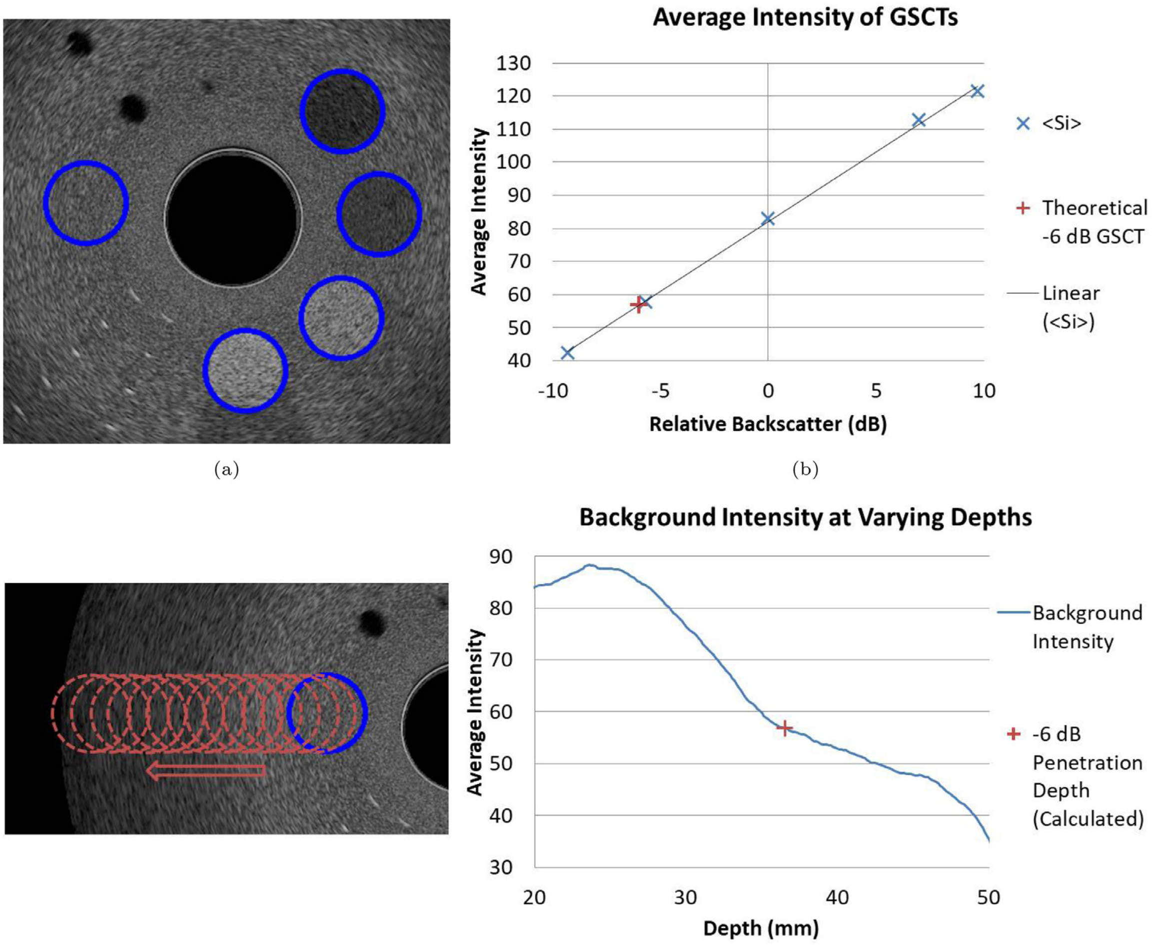
Datasets:

- DS1a: Original probe before repair (acquired 15/04/2021)
- DS2: Loan Probe (acquired 30/06/2021)
- DS1b: Original probe after repair (acquired 19/08/2021)
- DS3: Replacement probe (acquired 14/02/2022)

3. Results

3.1. Code performance

Table 1 shows the average results and CoV of quantitative parameters calculated using different methods (ImageJ/ SAUQA) and of different data sets acquired with the same acquisition parameters (the same image 10 times, 10 images acquired with the probe held by a stand and 10 images acquired with the probe being removed and reinserted between acquisitions).



(a) (b) (c) (d)

Fig. 4. Penetration depth calculation: (a) Average intensity of GSCTs calculated. (b) Linear regression performed and theoretical intensity of a −6 dB GSCT at 20 mm depth calculated. (c) Background ROI stepped 1px radially and at each point the average intensity is calculated. (d) Penetration depth calculated as the depth at which the average background intensity is equal to the theoretical intensity of a −6 dB GSCT at 20 mm depth.

3.2. SNR

With a fixed DR, the SNR of the GSCTs was measured with varying receive gain (Fig. 7). With a fixed receive gain, the SNR of the GSCTs was measured at different DRs (Fig. 8).

3.3. CNR

With a fixed DR, the CNR of the GSCTs was measured with varying receive gain (Fig. 9). With a fixed receive gain, the CNR of the GSCTs was measured with varying DR (Fig. 10).

3.4. Penetration depth

With a fixed DR of 70 dB, the −6 dB penetration depth was calculated with varying frequency and receive gain in LF and HF mode. In LF mode the penetration depth in each case was greater than the FOV so only the HF mode results are displayed in Fig. 11. Penetration depths are specified from the centre of the phantom. In HF mode, pixels at a depth of > 50 mm are set to 0, so any calculated penetration depth of > 50 mm will not reflect the true penetration depth. In the displayed graphs (Fig. 11), calculated penetration depths of ≥50 m have all been set to 50 mm (this is relevant to DS2 and DS3).

3.5. Resolution

For the purpose of measuring resolution, it is assumed that the arc length is comparable to the tangential length of the arc. This is valid for small arc lengths in comparison to the radius of the arc.

The resolution of the high contrast targets, in the tangential (Fig. 12) and radial (Fig. 13) direction, was measured with varying frequency. For each target depth, the resolutions reported are the average of both rows.

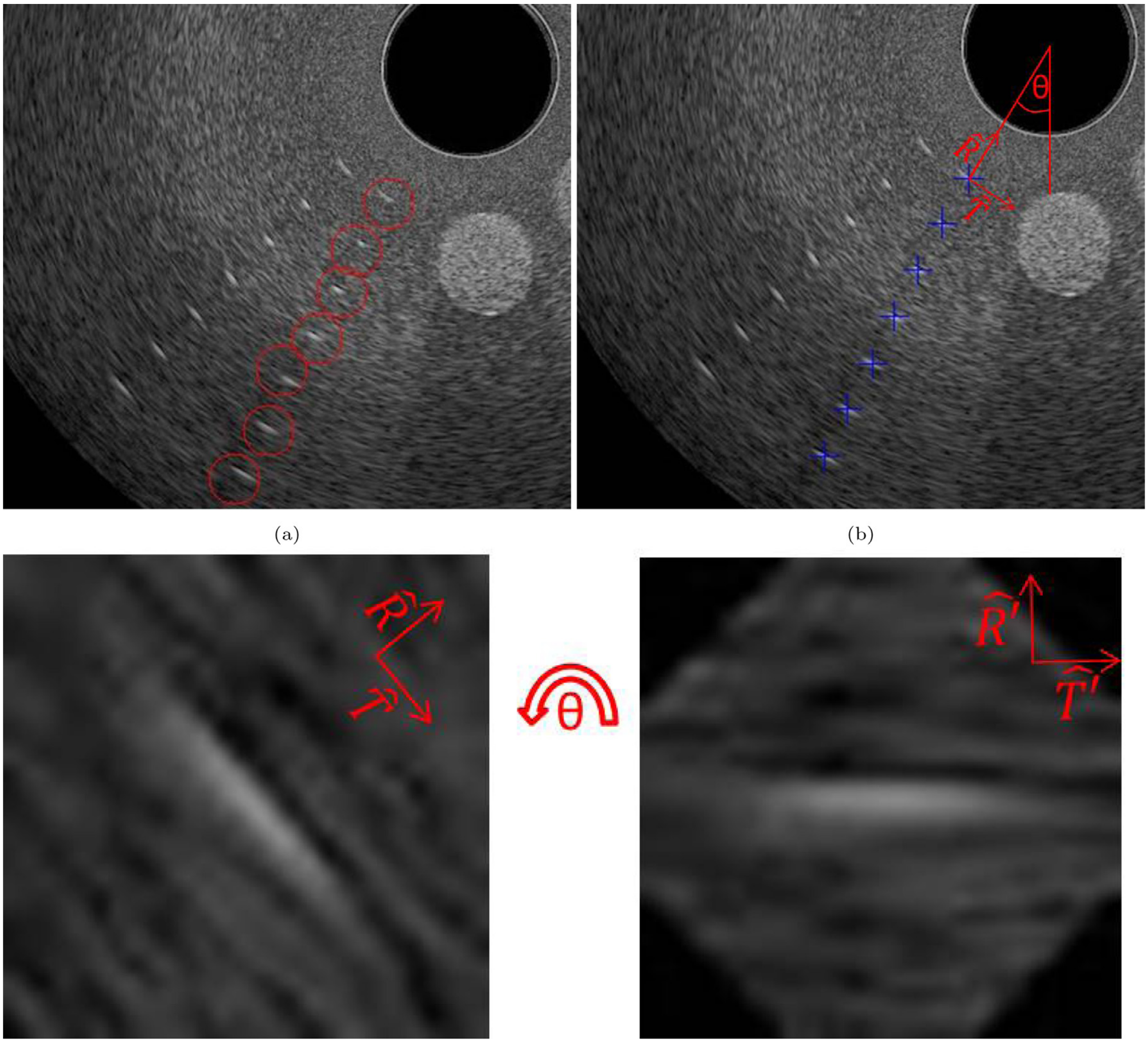
3.6. Distance accuracy

Theoretical radial distances are all 5 mm and theoretical chord distances were calculated assuming an angle of 19.1° between rows. The percentage difference between the measurements and the theoretical distances is reported (Fig. 14). For radial distances, the results reported are the average of both rows.

3.7. Low contrast object detectability

The LCOD (see Fig. 6) and contrast (see Eq. (2)) of the first 4mm anechoic target at 20 mm depth was measured at different receive gains for each dataset (Fig. 15).

The LCOD and contrast of the anechoic targets was also measured at different DRs and receive gains (Fig. 16).



(a) (b)

(c)

Fig. 5. Nylon monofilament targets: (a) Approximate locations selected by user. (b) Centre of the targets automatically calculated. (c) Radial ($\hat{R}$) and tangential ($\hat{T}$) resolution of each target is calculated as the $\hat{y}$ and $\hat{x}$ resolution of the target in the rotated ROI.

Table 1
Results and indicative variability of quantitative parameters calculated using different methods and of different data sets acquired with the same acquisition parameters.

Image quality metric		Mean result (CoV)			
		ImageJ analysis same image	SAUQA: Same image	SAUQA: Held by stand	SAUQA: Held by user
SNR	−9.3 dB Target	2.26 (0.021)	2.28 (0.014)	2.27 (0.015)	2.14 (0.087)
	−5.7 dB Target	3.37 (0.023)	3.39 (0.015)	3.37 (0.015)	3.19 (0.076)
	Background	5.09 (0.014)	5.12 (0.013)	5.07 (0.015)	5.15 (0.045)
	7 dB Target	6.75 (0.025)	6.79 (0.013)	6.71 (0.016)	6.49 (0.033)
	9.7 dB Target	7.44 (0.025)	7.43 (0.013)	7.36 (0.015)	7.21 (0.013)
CNR	−9.3 dB Target	2.22 (0.0075)	2.22 (0.0086)	2.21 (0.0098)	2.41 (0.17)
	−5.7 dB Target	1.36 (0.018)	1.36 (0.0050)	1.34 (0.0083)	1.59 (0.26)
	7 dB Target	1.04 (0.050)	1.05 (0.012)	1.04 (0.016)	0.870 (0.29)
	9.7 dB Target	1.46 (0.041)	1.43 (0.012)	1.43 (0.0073)	1.32 (0.089)
Penetration Depth (mm)		35.2 (0.022)	35.3 (0.015)	34.6 (0.023)	35.7 (0.042)
LCOD	4 mm Target 1	3.34 (0.016)	3.56 (0.0079)	3.61 (0.011)	3.70 (0.069)
	2 mm Target 1	1.54 (0.10)	1.58 (0.099)	1.77 (0.099)	1.83 (0.11)
Radial Distance Targets 1–2 (mm)		4.99 (0.040)	5.00 (0)	5.00 (0.0024)	4.95 (0.010)
Arc Distance Targets 1 (mm)		4.85 (0.072)	5.00 (0)	4.94 (0.013)	5.15 (0.030)
Radial Resolution Target 1 (mm)		0.395 (0.058)	0.440 (0.0074)	0.450 (0.028)	0.460 (0.058)
Tangential Resolution Target 1 (mm)		1.54 (0.077)	1.82 (0.013)	1.88 (0.046)	1.64 (0.065)

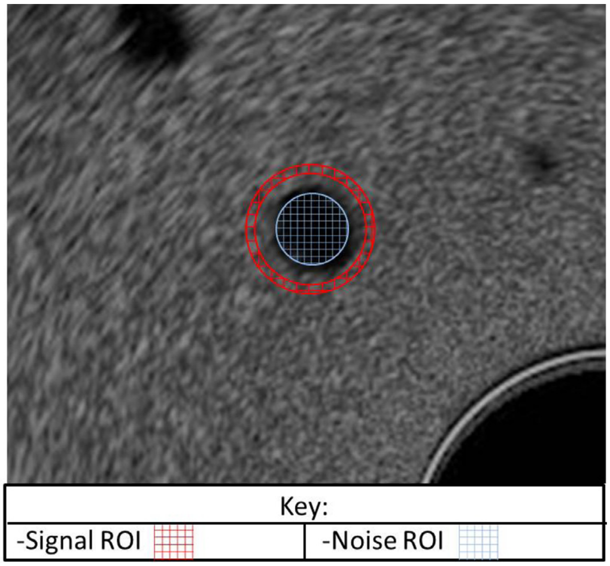


Fig. 6. ROIs used for LCOD assessment. Signal ROI in blue and noise ROI in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

The aim of this study was to provide an introduction to a new method of assessing the performance of 360° EAUS systems. The study is limited in its scope and it is hoped these limitations will be addressed in the future. In particular:

- Only one probe model was assessed - Multiple factors such as probe diameter, focus depth and operating frequency could affect how well the phantom is able to rigorously assess the performance of probes.
- The accuracy of the construction of the phantom has not been independently verified - The geometric accuracy could be assessed through further measurement with different US systems from other manufacturers, or CT imaging.
- Errors in the results has not been calculated - The CoV in Table 1 gives some indication in the repeatability of the calculated results, but only with one set of acquisition parameters; this may change for different acquisition parameters. In order to assess the reproducibility and uncertainty of the results, more data would need to be acquired. It should also be noted that the accuracy of the results is not known, given an independent verification of the phantom construction hasn't been performed. However, for QA this isn't necessarily an issue: Once baselines have been established, changes in performance can still be detected provided results are reproducible.

4.1. *Code performance*

Across the parameters in Table 1, there was good agreement between results when analysed using ImageJ and in SAUQA. Using a two tailed, Welch's *t*-test the only parameters producing significantly different results (with $p < 0.05$), were the LCOD of the 4 mm target, $(13) = 12$, $p < 0.01$; radial resolution, $t(9) = 6.1, p < 0.01$ and tangential resolution measurements $t(9) = 7.3, p < 0.01$.

For each parameters in Table 1 the average percentage deviation from the mean was also calculated. Using a single tailed, Welch's *t*-test a larger percentage deviation was observed across all parameters, $t(287) = 6.0, p < 0.01$, when analysing using ImageJ ($M = 2.9$, SD = 3.0), than SAUQA ($M = 1.2$, SD = 2.1). Taking each parameter individually, the only parameters where SAUQA didn't provide a significantly lower average percentage deviation than ImageJ (with $p < 0.05$), were −9.3 dB CNR $t(15) = -0.058$, $p = 0.48$; −5.7 dB CNR $t(17) = 1.5$, $p = 0.079$; 0 dB CNR $t(16) = 0.71$, $p = 0.24$; penetration depth $t(17) = 1.5$, $p = 0.071$; LCOD 4 mm $t(9) = 1.2$, $p = 0.12$ and LCOD 2 mm $t(14) = -0.12$, $p = 0.46$. There were no parameters where ImageJ provided a significantly smaller deviation than SAUQA. These results suggest that SAUQA will offer a more repeatable method of calculation than ImageJ, while also being quicker to use.

Across all parameters in Table 1 (and using a single tailed Welch's *t*-test), an increase was observed in the average percentage deviation from the mean, when comparing the SAUQA analysis on the same image ($M = 1.2$, SD = 2.1), to the 10 different images with the probe held by a stand ($M = 1.2$, SD = 2.1); $t(317) = 2.1$, $p = 0.016$. An additional increase in variability was observed, $t(175) = 6.5$, $p = 0.016$, when comparing the results with the probe held by the stand ($M = 1.7$, SD = 2.0) and the results when the probe was removed and reinserted between acquisitions ($M = 6.5$, SD = 9.0). Given this increase in variability of results observed, it is believed that the different positioning of the probe between datasets will have an impact on the results. However, by holding the probe in a consistent location for each dataset, using a stand, this should reduce the variability in calculated results within each dataset.

4.2. *Receive gain change*

As the receive gain is increased, the SNR increased smoothly for all GSCTs, apart from between ~45 dB and ~55 dB, where

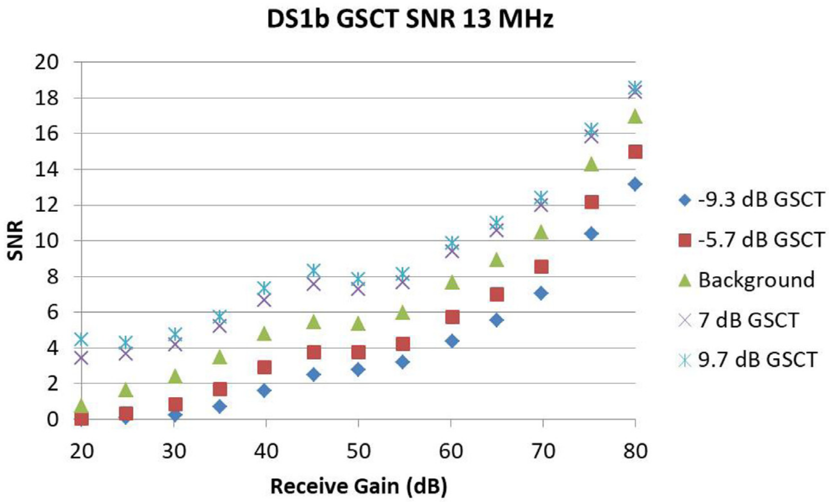


Fig. 7. GSCT SNR at different receive gains for DS1b.

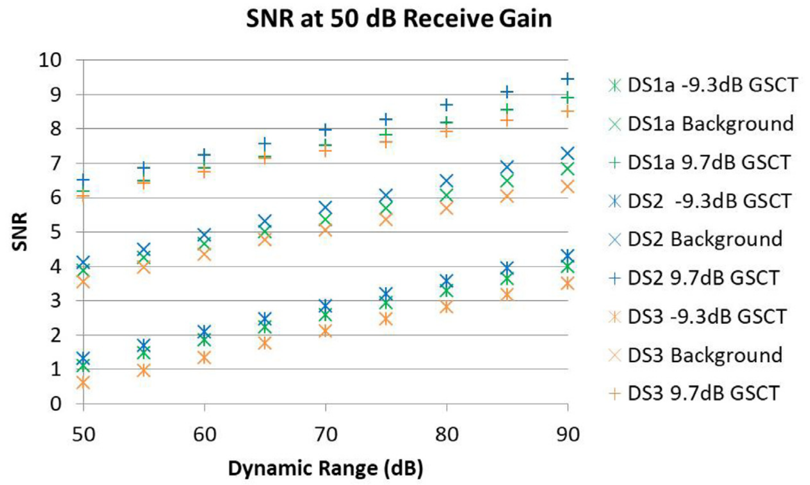


Fig. 8. GSCT SNR at different DRs for each probe.

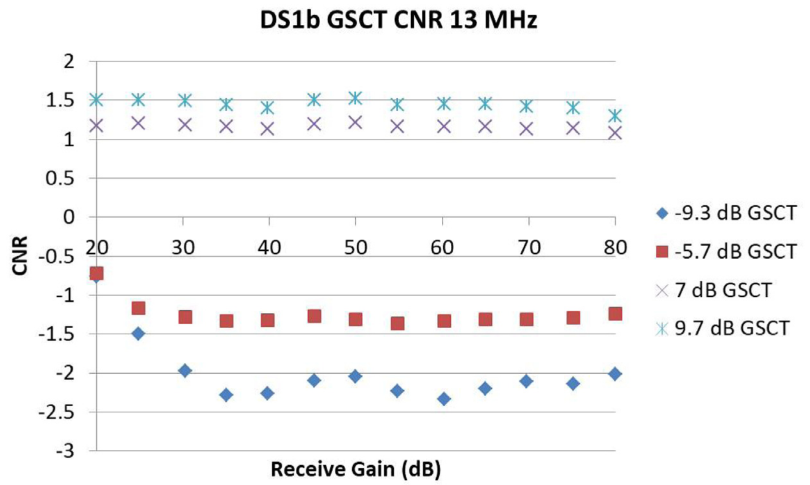


Fig. 9. GSCT CNR at different receive gains for DS1b.

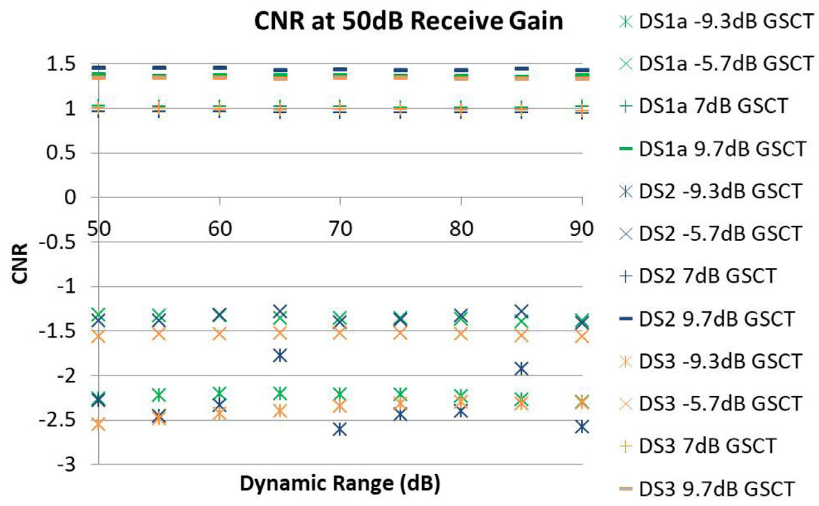
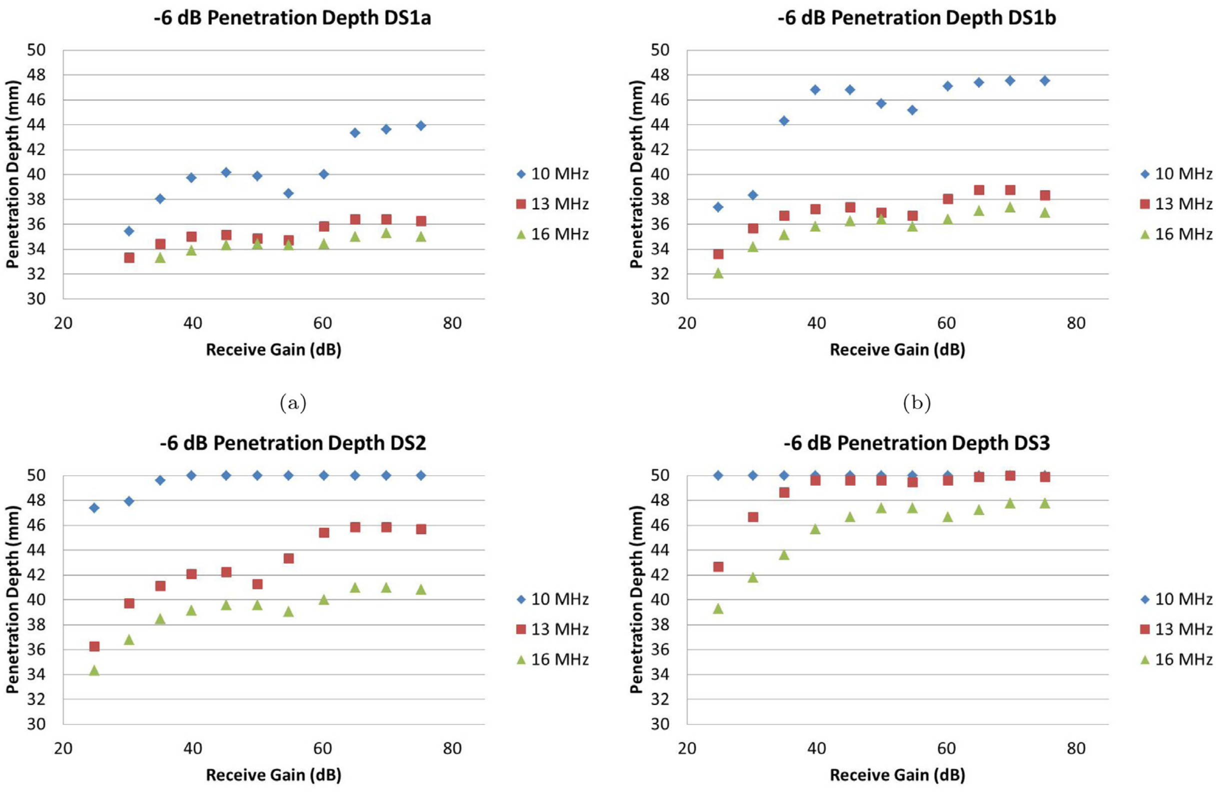


Fig. 10. GSCT CNR at different DRs for each probe.

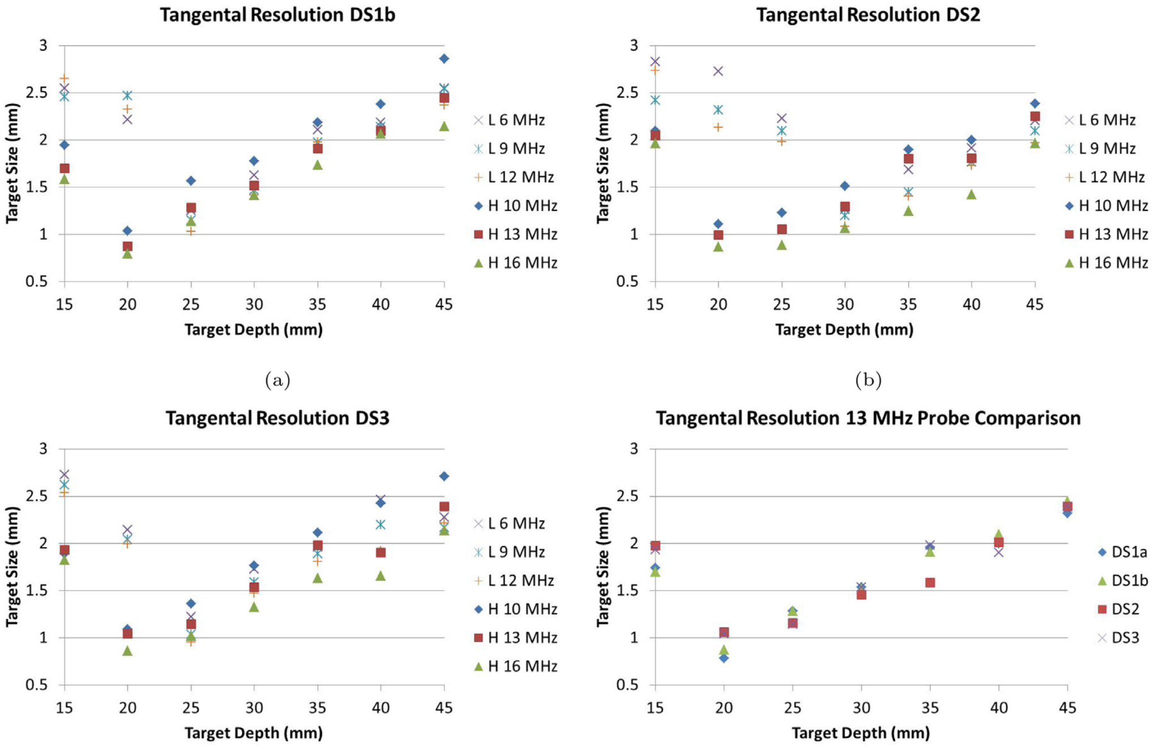
there was a small reduction in SNR (Fig. 7). The dip in SNR at intermediate receive gains was exhibited in each dataset. At very low receive gains, the lower backscatter targets offered virtually no signal.

For the higher backscatter targets, the CNR was broadly consistent across the receive gain range (Fig. 9). However, for the lower backscatter targets, the CNR improved as the receive gain was increased, until ~30 dB receive gain for the −5.7 dB GSCT and ~35 dB receive gain for the −9.5 dB GSCT. This is believed to be due to the low signal offered by the lower backscatter targets at low receive gains. At high receive gains, the CNR remains broadly consistent for the −5.7 dB target. However, for the −9.5 dB target there is a dip in CNR between ~40 dB and



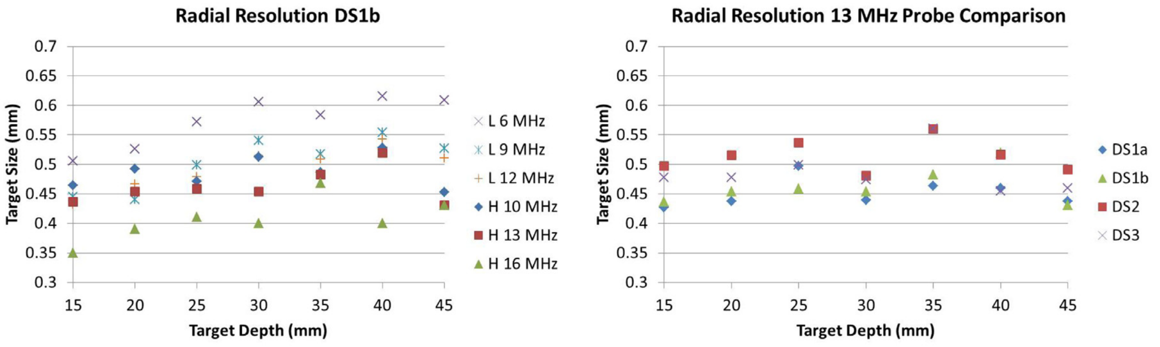
(a) (b) (c) (d)

Fig. 11. −6 dB penetration depth at different receive gains and frequencies of DS1a (a), DS1b (b), DS2 (c), and DS3 (d). Calculated penetration depths of ≥50 mm have all been set to 50 mm in the graphs.



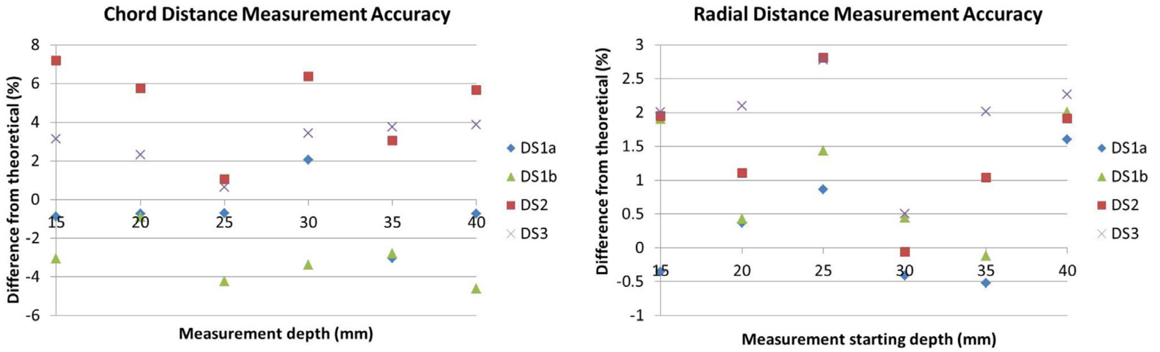
(a) (b) (c) (d)

Fig. 12. Resolution of high contrast targets in the tangential direction. (a,b and c) the results for the different probes, (d) a comparison between the probes at 13 MHz.



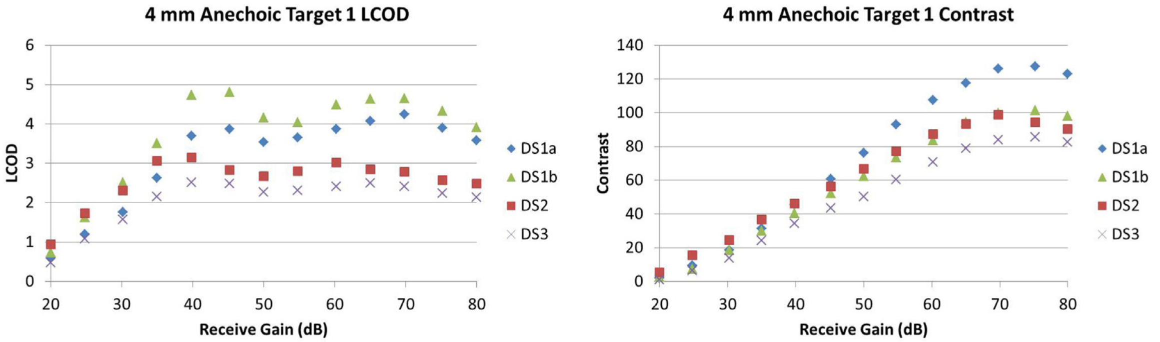
(a) (b)

Fig. 13. Resolution of high contrast targets in the radial direction. (a) the results for DS1b at different frequencies, (b) a comparison between the probes at 13 MHz.



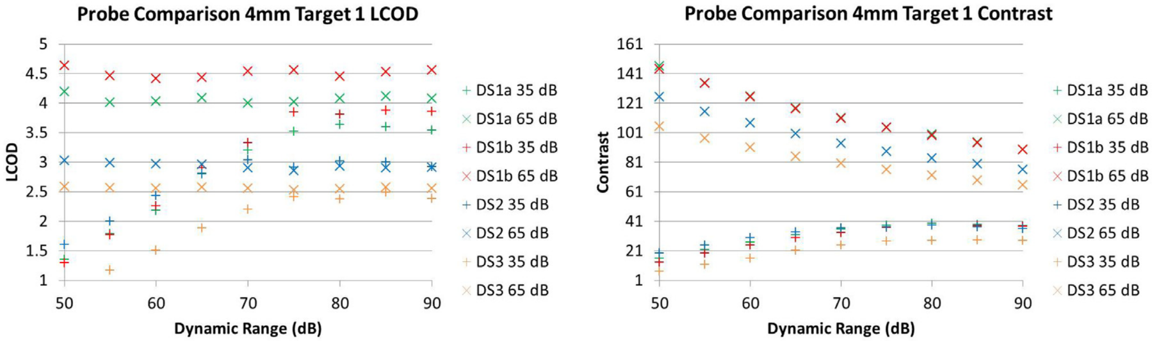
(a) (b)

Fig. 14. Distance accuracy between high contrast targets, in chord (a) and radial (b) direction.



(a) (b)

Fig. 15. 4mm anechoic target LCOD (a) and contrast (b) at different receive gains.



(a) (b)

Fig. 16. 4 mm anechoic target LCOD (a) and contrast (b) at different DRs and receive gains. Legend titles indicate the receive gain for each series.

~60 dB receive gain. The dip in CNR at intermediate receive gains was exhibited in each dataset.

For all datasets the −6 dB penetration depth was found to increase with receive gain up to ~40 dB, at which point there was a small reduction in penetration depth up to ~55 dB (Fig. 11). At higher receive gains, there is a small increase in penetration depth, before the penetration depth levels off. At low receive gains, there is a large drop off in the signal provided by the background TMM resulting in the poor penetration depth.

The LCOD also increases as receive gain is increased, up to ~40 dB, after which the LCOD becomes broadly consistent (Fig. 15). As the receive gain is increased, the signal provided by the background TMM increases and therefore the contrast of the anechoic objects increases. The noise will also increase, but at lower attenuations the rate at which it increases is less than the increase in background signal. A slight reduction in LCOD was exhibited at high receive gains as a result of a drop in contrast. Similar to the SNR, CNR and penetration depth, each dataset exhibited a drop in LCOD at an intermediate receive gain between ~40 dB and ~55 dB (Fig. 15(a)).

The cause of the discrepancies at intermediate receive gains is not known. However, given the contrast is approximately linear in this region and the LCOD is not (Fig. 15), the noise measured appears to be the main factor resulting in this discrepancy. The fact each dataset exhibited similar behaviour, suggests that the discrepancy is caused by the US machine and/or an inherent characteristic of the probe design.

4.3. DR change

For an intermediate (50 dB) receive gain, the SNR increases as DR increases; in an approximately linear manner for all targets (Fig. 8). The GSCT CNR is also broadly consistent with changing DR (Fig. 10). As the DR is increased, the contrast and noise both reduce, resulting in a consistent CNR.

Similar behaviour is seen in the LCOD (Fig. 16). For receive gains above 50 dB, both the contrast and noise decrease with increasing DR. This results in a broadly consistent LCOD across the full DR. However, for low receive gains, the signal offered by the TMM is very small resulting in a low contrast between the background TMM and the anechoic cylinders. For low receive gains and DRs the contrast approaches zero. This results in contrast and LCOD increasing with increasing DR, before stabilising at high DRs.

4.4. Frequency change

The focus depth in the arc direction is assumed to be comparable to the depth of the smallest resolution target in the tangential direction. Fig. 12 suggests that adjusting the frequency within each mode (LF/HF) has little impact upon the arc focus depth. However, changing between LF and HF mode has a substantial effect on the arc direction focus depth; with the focus depth in LF mode being ~5 mm deeper in LF mode than HF mode.

Theoretically, higher frequencies should result in improved resolution at the expense of poorer penetration [35]. Fig. 11 shows that lower frequencies resulted in a greater penetration depth and Fig. 12 shows that, for each mode (LF/HF), the tangential resolution of the high contrast targets decreases (improves) with increasing frequency.

The behaviour of the radial resolution is less clear (Fig. 13(b)). This is believed to be due to the radial resolution being comparable to the radial speckle size. However, at 16 MHz the radial resolution is consistently smaller than at lower frequencies and the radial resolution at 6 MHz appears to be larger than higher frequencies. It should be noted that there is little difference in the measured radial resolution and there is overlap between the measurements at different frequencies.

4.5. Probe comparison

Across all parameters, the results from DS1a and DS1b suggests the repair of the probe was successful. No difference could be observed qualitatively and the differences observed quantitatively were generally smaller than the differences between probes.

The SNR of the GSCTs was lower in DS3 than the other datasets (Fig. 8). However, each dataset produced broadly similar results for GSCT CNR and SNR at different receive gains and DRs and the discrepancy in the smoothly changing penetration depth/CNR/SNR with receive gain happens at similar receive gains for each dataset.

All probes exhibited similar radial and tangential resolutions (Figs. 12 and 13). However, the arc direction focus depth appears to vary between the probes. In LF mode, the arc direction focus depth appears to be the same for DS1b and DS3 (~ 25 mm), however for DS2 the focus depth appears to be deeper (~ 30 mm). In HF mode, DS2 and DS3 have similar resolution high contrast targets at 20 mm and 25 mm depth, suggesting the arc direction focus depth is somewhere between. For DS1b, the arc direction focus depth is likely closer to 20 mm depth.

The distance accuracy exhibited by all datasets was excellent and very similar. In a radial direction, all measurements taken were within 4% of the expected value of 5 mm and in the chord direction, all measurements were within 8% of that expected. A larger error was expected in the chord direction than in the radial direction because the tangential resolution is larger than the radial resolution. DS2 and DS3 did exhibit a slightly greater measurement error from the expected than the DS1a and DS1b. However, because the actual distance between the targets is not know, it is difficult to be certain which is more accurate. On average, the chord measurements were smaller than expected for DS1a and DS1b and larger than expected for DS2 and DS3. It should be noted that the speed of sound of the US beam will be different inside the TMM to that expected in vivo. It is known that speed of US within Zerdine varies with temperature [12]; with lower temperatures resulting in lower acoustic velocities. If an US beam travels slower than expected through the phantom, this will result in the reflections from the US taking longer to be received by the probe and hence structures may appear more distant.

Comparing the results from the anechoic targets, there were some differences between the probes in terms of LCOD and contrast. For the first set of anechoic targets, DS1a and DS1b offered greater LCOD and contrast than DS2 or DS3 (Figs. 15 and 16). This suggests that anechoic targets would be easier to distinguish from the background TMM with the old probe rather than the loan/replacement probe.

The -6dB penetration depth did vary between the datasets (Fig. 11), with DS3 providing the deepest penetration depth, followed by DS2, DS1b and finally DS1a (in HF mode).

4.6. Relating results to in vivo

The TMM used in the phantom construction is Zerdine. While this is a better tissue mimic than ATS rubber, the acoustic properties of Zerdine will differ from in vivo tissue. In particular, the attenuation coefficient of the TMM, differs from that expected in in vivo tissue at higher frequencies ≥~ 6 MHz [12].

For IAS assessment, the ideal focus depth for most patients would be at ~5 mm from the probe surface (~13.5 mm depth). The arc direction focus depth varies between HF and LF mode, with the focus depth being ~5 mm shallower in HF mode than LF mode. In HF mode the arc focus depth was measured at ~20 mm depth (~10 mm from the TMM surface). Therefore, while the focus depth may be different in vivo to in the phantom, it is expected that HF mode will offer a more optimal focus depth than LF mode for IAS assessment. DS1a/b provided a focus depth closer to the typical IAS depth, than DS2 or DS3.

With a typical IAS thickness of ~3 mm, there is also the requirement that the radial resolution is less than 3 mm and ideally significantly less. In HF mode, all datasets offered a radial resolution of ~0.5 mm, however DS1a/b performed better than DS2 or DS3 at most target depths and in particular at the typical depth of the IAS (Fig. 13(b)). The radial resolution also tended to be better at higher frequencies and in HF mode rather than LF mode.

In terms of visualising the IAS, the other important factor to consider is contrast between the IAS and tissue surrounding it; with the IAS offering little signal in comparison to the tissue surrounding it. For this reason the IAS is expected to behave in a similar manner to the anechoic cylinders, particularly the cylinders at the shallowest depth. The LCOD and contrast of the anechoic cylinders tended to be best at higher receive gains and intermediate DRs. At very high and low receive gains and at very high and low DRs the anechoic cylinders were less easy to visualise quantitatively and in some cases could not be observed by the user qualitatively. DS1a/b also quantitatively performed better than DS2 and DS3. It should be noted that the optimal values for DR and receive gain may be different for the IAS in vivo. However, the results in the phantom do suggest that adjusting the receive gain and DR can have a large impact on how easily the IAS can be observed.

While DS2 and DS3 provided a greater penetration depth than DS1a and DS1b, this is not expected to have a significant impact in vivo. Even at the highest frequency, all probes we able to provide excellent contrast and signal at the typical depth of the IAS.

4.7. *In vivo preset optimisation*

On the BK 3000, adjusting the FOV and receive gain is performed using two physical dials. However, adjusting the DR/frequency requires the mouse, which takes longer and is more difficult to do. For this reason, users rarely adjust the frequency and DR when acquiring patient images. It is therefore particularly important that the default frequency/DR at start-up is appropriate. Currently the default start-up settings are: HF mode, frequency = 13 MHz, receive gain = 50 dB and DR = 70 dB.

While HF mode is appropriate, ideally the system would start at 16 MHz, because this should offer improved resolution at the typical depth of the IAS.

At an intermediate receive gain, adjusting the DR did not quantitatively affect many of the metrics significantly. 65 dB DR is an appropriate default because the LCOD of the first 4 mm anechoic target at 50 dB receive gain peaked at a DR of 65 dB.

Between receive gains of ~40 dB and ~60 dB, the SNR, CNR, penetration depth and LCOD all varied in unexpected manners. For example, the LCOD of the first 4 mm anechoic was lower at 50 dB receive gain than both 40 dB receive gain and 70 dB receive gain (Fig. 15). While the phantom results may not directly translate to in vivo, it may be advisable to adjust the default to a higher receive gain of 60 dB. This is still in the middle of the range of available receive gains and multiple metrics peaked at ~60 dB receive gain.

5. Conclusion

- The bespoke phantom, in combination with the analysis package, allows for the rigorous quantitative assessment of 360° US probes designed for EAUS.
- At intermediate receive gains adjusting the receive gain resulted in the SNR, CNR, penetration depth and LCOD varying in an unexpected manner. The reason for this is not known, but because it was exhibited by all probes it is believed to be related to the BK 3000 US machine itself and/or an inherent characteristic of the probe design.
- The repair of the probe appears to have been successful, with DS1a and DS1b providing comparable results.
- The parameters measured suggest that all probes tested offer an effective method of assessing the integrity of the IAS. The system offers appropriate SNR, CNR, LCOD, penetration depth, distance accuracy and the resolution for EAUS.
- The old (DS1a) and repaired (DS1b) probe, appeared to performed better than the loan (DS2) and replacement (DS3) probes: with a shallower focus depth, better visualisation of low contrast targets, better resolution, and better distance accuracy.
- Adjusting the receive gain, DR, frequency and frequency mode had significant effects on the qualitative and quantitative appearance of the image. For the machine tested, appropriate start-up settings would be: HF mode, frequency = 16 MHz, DR = 65 dB and receive gain = 60 dB. Users should also ensure they adjust both the DR and receive gain to achieve the optimal image quality.

Funding

None.

Ethical approval

Not required.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank CIRS for manufacturing the bespoke phantom.

References

[1] Z. Abdool, A. H. Sultan, R. Thakar, Ultrasound imaging of the anal sphincter complex: a review, 2012, doi:10.1259/bjr/27314678.
[2] P. Tejedor, J. Plaza, I. Bodega-Quiroga, M. Ortega-López, D. García-Olmo, C. Pastor, The role of three-dimensional endoanal ultrasound on diagnosis and classification of sphincter defects after childbirth, J. Surg. Res. (2019), doi:10.1016/j.jss.2019.06.080.
[3] I.S. Almeida, U. Jayarajah, D.P. Wickramasinghe, D.N. Samarasekera, Value of three-dimensional endoanal ultrasound scan (3D-EAUS) in preoperative assessment of fistula-in-ano, BMC Res. Notes (2019), doi:10.1186/s13104-019-4098-2.
[4] M.B. Nielsen, C. Hauge, O.Ø. Rasmussen, M. Sørensen, J.F. Pedersen, J. Christiansen, Anal sphincter size measured by endosonography in healthy volunteers, Acta Radiol. (1992), doi:10.3109/02841859209172033.
[5] D.E. Fenner, J.S. Kriegshauser, H.L. Henry, R.W. Beart, A. Weaver, J.L. Cornelia, Anatomic and physiologic measurements of the internal and external anal sphincters in normal females, Obstet. Gynecol. (1998), doi:10.1016/S0029-7844(97)00678-9.
[6] A. Sboarina, A. Minicozzi, C. Segattini, F. Leopardi, F. Lombardo, V. Passeri, G. Scudo, M. Bencivenga, A. Fenzi, C. Cordiano, Shape and volume of internal anal sphincter showed by three-dimensional anorectal ultrasonography, Eur. J. Radiol. (2012), doi:10.1016/j.ejrad.2011.04.024.
[7] User Guide 20R3 Transducer, 2019, 16-01278-EN-07.
[8] bk3000 & bk5000 Ultrasound Systems 16-01487-EN-03, 2020, Systems 16-01487-EN-03.
[9] N. Dudley, S. Russell, B. Ward, P. Hoskins, BMUS guidelines for the regular quality assurance testing of ultrasound scanners by sonographers, Ultrasound (2014), doi:10.1177/1742271X13511805.
[10] N. Dudley, T. Evans, P. Hoskins, A. Watson, H. Starritt, Quality Assurance of Ultrasound Imaging Systems Report No. 102, Technical Report, IPEM, 2010.
[11] Rectal Scan Phantom Model ATS 540 DS 081518, 2018.
[12] J.E. Browne, K.V. Ramnarine, A.J. Watson, P.R. Hoskins, Assessment of the acoustic properties of common tissue-mimicking test phantoms, Ultrasound Med. Biol. (2003), doi:10.1016/S0301-5629(03)00053-X.
[13] J.R. Rodgers, K. Surry, D. D'Souza, E. Leung, A. Fenster, 360-degree 3D transvaginal ultrasound system for high-dose-rate interstitial gynaecological brachytherapy needle guidance, Medical Imaging 2017: Ultrasonic Imaging and Tomography, 2017, doi:10.1117/12.2254082.
[14] J.H. Sung, J.S. Jeong, Development of high-frequency (>60 MHz) intravascular ultrasound (IVUS) transducer by using asymmetric electrodes for improved beam profile, Sensors (Switzerland) (2018), doi:10.3390/s18124414.
[15] S.K. Nadkarni, H. Austin, G. Mills, D. Boughner, A. Fenster, A pulsating coronary vessel phantom for two- and three-dimensional intravascular ultrasound studies, Ultrasound Med. Biol. (2003), doi:10.1016/S0301-5629(02)00730-5.
[16] J.L. Evans, K.H. Ng, S.G. Wiet, M.J. Vonesh, W.B. Burns, M.G. Radvany, B.J. Kane, C.J. Davidson, S.I. Roth, B.L. Kramer, S.N. Meyers, D.D. McPherson, Accurate three-dimensional reconstruction of intravascular ultrasound data: spatially correct three-dimensional reconstructions, Circulation (1996), doi:10.1161/01.CIR.93.3.567.

[17] H. Nisar, J. Moore, R. Piazza, E. Maneas, E.C. Chen, T.M. Peters, A simple, realistic walled phantom for intravascular and intracardiac applications, Int. J. Comput. Assist. Radiol. Surg. (2020), doi:10.1007/s11548-020-02201-3.

[18] C. Moran, W. Ellis, A. Janeczko, D. Bell, S. Pye, The edinburgh pipe phantom: characterising ultrasound scanners beyond 50 MHz, in: Journal of Physics: Conference Series, vol. 279, IOP Publishing, 2011, p. 012008.

[19] C. McLeod, C.M. Moran, K.A. McBride, S.D. Pye, Evaluation of intravascular ultrasound catheter-based transducers using the resolution integral, Ultrasound Med. Biol. 44 (12) (2018) 2802–2812.

[20] S. Inglis, A. Janeczko, W. Ellis, J.N. Plevris, S.D. Pye, Assessing the imaging capabilities of radial mechanical and electronic echo-endoscopes using the resolution integral, Ultrasound Med. Biol. 40 (8) (2014) 1896–1907.

[21] Ultrasound Video Gastroscope EUS-J10 Fact Sheet Rev1, 2021.

[22] Endocavity Biplane Transducer Type 8848 User Guide, 2016.

[23] X14L4 (9038) Transducer Product Data Sheet 16-01665-01, 2018a.

[24] E14CL4b (9048) Transducer Product Data Sheet 16-01282-07, 2018b.

[25] Olympus Ultrasonic Probe RU-75M-R1/RU-12M-R1 Flyer 1848.

[26] THD EAUS Probe 12 MHz Technical Datasheet Rev01, 2018.

[27] UltraiQ General Purpose Phantom Specifications. 2020.

[28] J. Harkin, SAUQA v1.0 24/02/2022, 2022. https://github.com/harkin7765/SAUQA.

[29] F.L. Drake, Python 3.8.13 documentation, 2021.

[30] H. Li, J. Wu, A. Miao, P. Yu, J. Chen, Y. Zhang, Rayleigh-maximum-likelihood bilateral filter for ultrasound image enhancement, Biomed. Eng. Online (2017), doi:10.1186/s12938-017-0336-9.

[31] K.M. Kempski, M.T. Graham, M.R. Gubbi, T. Palmer, M.A. Lediju Bell, Application of the generalized contrast-to-noise ratio to assess photoacoustic image quality, Biomed. Opt. Express (2020), doi:10.1364/boe.391026.

[32] J.M. Thijssen, G. Weijers, C.L. de Korte, Objective performance testing and quality assurance of medical ultrasound equipment, Ultrasound Med. Biol. (2007), doi:10.1016/j.ultrasmedbio.2006.09.006.

[33] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, A. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A.P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C.N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D.A. Nicholson, D.R. Hagen, D.V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G.A. Price, G.L. Ingold, G.E. Allen, G.R. Lee, H. Audren, I. Probst, J.P. Dietrich, J. Silterra, J.T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J.L. Schönberger, J.V. de Miranda Cardoso, J. Reimer, J. Harrington, J.L.C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N.J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P.A. Brodtkorb, P. Lee, R.T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T.J. Pingel, T.P. Robitaille, T. Spura, T.R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y.O. Halchenko, Y. Vázquez-Baeza, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods (2020). arXiv: 1907.10121.

[34] K. Rahman, Abid, A. Mordvintsev, OpenCV-Python Open Source Computer Vision 4.6.0 Documentation, 2022.

[35] A. Ng, J. Swanevelder, Resolution in ultrasound imaging, Contin. Educ. Anaesthesia, Crit. Care Pain (2011), doi:10.1093/bjaceaccp/mkr030.