Phantom-based quality assurance measurements in B-mode ultrasound

Mannila, Vilma

Sage
2013

Mannila, V & Sipilä, O 2013, 'Phantom-based quality assurance measurements in B-mode ultrasound', Acta Radiologica Short Reports, vol. 2, no. 8, 2047981613511967. https://doi.org/10.1177/20479816135

http://hdl.handle.net/10138/178755
10.1177/2047981613511967

publishedVersion

Acta Radiologica Short Reports
2(8) 1–4
DOI: 10.1177/2047981613511967
arr.sagepub.com
SAGE

Phantom-based quality assurance measurements in B-mode ultrasound

Vilma Mannila[1,2] and Outi Sipilä[1,2]

Abstract

Background: Recommended phantom-based quality assurance measurements in B-mode ultrasound (US) may be tedious. For the purpose of cost-effective US quality assurance it is important to evaluate measurements that effectively reflect the quality of US scanner.

Purpose: To find out which recommended phantom-based quality assurance measurements are effective in detecting dead or weak transducer elements or channels in US scanners when visual image analysis and manual measurements are used.

Material and Methods: Altogether 66 transducers from 33 US scanners were measured using a general purpose phantom and a transducer tester. The measurements were divided into two groups. Group I consisted of phantom-based uniformity measurement, imaging the air with a clean transducer (air image) and measuring the transducer with the transducer tester, and group II of phantom-based measurements of depth of penetration, beam profile, near field, axial and lateral resolution, and vertical and horizontal distance accuracy. The group II measurements were compared to group I measurements.

Results: With group I measurements, the results with 20% of the transducers were found defective. With 35% of the transducers the results were considered defective in group II measurements. Concurrent flaws in both groups were found with 11% of the transducers.

Conclusion: Phantom-based measurements of depth of penetration, beam profile, near field, axial and lateral resolution, and vertical and horizontal distance accuracy did not consistently detect dead or weak transducer elements or channels in US scanners.

Keywords
Ultrasound, quality assurance, quality control, acceptance testing

Date received: 17 April 2013; accepted: 15 October 2013

Introduction

Quality assurance (QA) methods for B-mode ultrasound (US) are often based on phantom studies (1–8), conventionally performed with manual measurements and visual image analysis (1–4,6,7). Computer programs for automatic image analysis have also been utilized to increase the objectivity, e.g. in reference (8). In addition, transducer testers for evaluating the functionality of the individual elements are available (9,10).

International standards and recommendations include a range of routine tests with tissue mimicking phantoms (1–5), which are quite tedious to perform. For cost-effective US quality assurance, the tests should reliably reflect the quality of US scanner. According to the recent standards and recommendations by the American College of Radiology (ACR) (2), American Institute of Ultrasound in Medicine (AIUM) (3), and Institute of Physics and Engineering in Medicine (IPEM) (4), either visual or automatic methods can be used for analysis, while in the guidelines by the European Federation of Societies for

---

[1] HUS Helsinki Medical Imaging Center, Helsinki University Central Hospital, Helsinki, Finland
[2] Department of Physics, University of Helsinki, Helsinki, Finland

Corresponding author:
Vilma Mannila, HUS Helsinki Medical Imaging Center, Helsinki University Central Hospital, P.O. Box 340, 00029 HUS, Finland.
Email: vilma.mannila@hus.fi

Ultrasound in Medicine and Biology (EFSUMB) (5) only automatic image analysis is recommended.

In reference (11), phantom-based uniformity measurements, imaging the air with a clean transducer (air image), and measurements with a transducer tester were compared in order to evaluate which measurements were useful detecting faults related with dead or weak transducer elements or channels in US scanners. The results indicated the usefulness of the methods. On the other hand, none of the methods was adequate alone. Hangiandreou et al. (12) have also found image uniformity assessment a useful measurement in routine QA. The purpose of this study was to evaluate the effectiveness of other recommended phantom-based QA measurements to detect dead or weak transducer elements or channels in US scanners. The most straightforward approach, visual image analysis and manual measurements, was used. The phantom-based measurements included in this study were depth of penetration, beam profile, near field, axial and lateral resolution, and vertical and horizontal distance accuracy (1–5). Depth of penetration and distance accuracy measurements are mandatory in all of the abovementioned standards.

To detect changes in most of the phantom measurements, usually baseline values are needed. Since the baseline specifications in this study were not available from the manufacturers nor from the acceptance measurements, the limits for abnormal QA parameter values for the depth of penetration and for the spatial resolution measurements were estimated utilizing standard deviations in groups of similar transducers operated in similar US scanners.

Material and Methods

Altogether 33 US scanners with 66 transducers from three different manufacturers were included in this study (Table 1). The transducers were divided into 14 groups consisting of similar transducer-scanner combinations. Each group consisted of three to 12 transducers.

The measurements were divided into two groups. Group I consisted of a phantom-based uniformity measurement, scanning air with a clean transducer (air image), and measuring the transducers with a transducer tester. In the uniformity measurement and in the air image, clear vertical streaks or asymmetry starting from the very top of the image were interpreted as defects (11). For the transducer tests, a Sonora FirstCall aPerioTM system (Sonora Medical Systems, Inc., Longmont, CO, USA) was utilized (13). The results were categorized as in Mårtensson et al. (9), meaning that at least two to four dead or weak elements were required for a transducer to be classified as non-functional. The results from the measurement group I were considered defective, if the uniformity measurement or the air image or the transducer test produced a defective result.

Measurement group II consisted of standard phantom-based measurements for B-mode imaging: depth of penetration, beam profile, near field, axial and lateral resolution, and vertical and horizontal distance accuracy (1–5). A CIRS General Purpose Phantom Model 040 (CIRS Tissue Simulation and Phantom Technology, Norfolk, VA, USA) was utilized (14). The uniformity measurement in measurement group I was performed with the same phantom. Attenuation

Table 1. Measured US scanners and transducers.

Manufacturer	Scanner model	Models (*n*)	Purchase years	Median purchase year	Transducer model	Transducers (*n*)
Aloka	SSD-5500 / SSD-5500SV / SSD5000	10	1998–2004	2001	UST-5545	5
					UST-5712	3
					UST-9126	4
					UST-9119	4
GE	LOGIQ 9	14	2001–2007	2005	M12L	12
					10L	4
					4C	11
					3.5C	4
					8C	4
	LOGIQ S6	3	2006	2006	M12L	3
					10L	3
					4C	3
Toshiba	Aplio XG	3	2008	2008	PVT-375BT	3
	Aplio/Aplio 80	3	2002–2004	2003	PVT-375BT	3

coefficient of 0.5 dB/cm/MHz was used in this study. All the measurements were carried out by the same physicist. Identical scanner–transducer combinations were measured using the same scanning protocols. The main principle in the protocols was to include minimum processing of the signal, e.g. to switch off advanced imaging functions. All phantom-based parameters were measured using two frequencies; the highest and the lowest. The analysis was performed visually with manual measurements.

The averages and standard deviations were calculated for the depth of penetration, and axial and lateral resolution for every transducer–scanner group. The depth of penetration was the depth where the noise started to dominate over the speckle. The values less than the average minus two standard deviations indicated a defective value. The axial and lateral resolutions were checked visually from filaments with different separations (axial: 0.5, 1, 2, 3, 4, and 5 mm; lateral: 1, 2, 3, 4, and 5 mm) in the phantom. Values higher than the average plus two standard deviations were considered as defects.

The result of a near field resolution measurement was defective if the wire closest to the surface (1 mm) did not appear as a dot in the image. If the narrowest point in the diverging beam did not coincide with the position of the lateral focus, the beam profile was considered defective. For the vertical and horizontal distance accuracy the suggested defect levels by Goodsitt et al. (1), 2% or 2 mm and 3% or 3 mm, respectively, were utilized. The measured distance in the vertical direction was 20–140 mm depending on the depth of penetration of the transducer. The measured horizontal distance depended on the width of the field of view being 20–40 mm in the depth of 30 mm and 40–120 mm in the depth of 90 mm.

The results from measurement group II were considered defective if any of the abovementioned phantom measurement was defective in either frequency measured, except for the vertical and horizontal distance accuracy. With these parameters, a defective result with both frequencies was required to minimize possible effects due to measurement accuracy.

Group II measurements were compared to group I measurements to determine their efficacy in detecting dead or weak transducer elements or channels in US scanners.

Results

In measurement group I, results with 13 of the 66 transducers (20%) were considered defective. With 23 transducers (35%) a defective result was detected with group II measurements. Concurrent flaws, i.e. a defective result in the both groups, were found with seven transducers (11%). With these seven transducers, the defective result in group II measurements was detected in the vertical distance accuracy (five transducers), horizontal distance accuracy (one transducer), or in the near field resolution (one transducer).

With six transducers (9%), group II measurements had no defects although a defect was found with group I measurements. On the other hand, with 16 transducers (24%) defective image quality was only detected with group II measurements. With 15 of these 16 transducers, defects were found in the vertical or horizontal distance accuracy, and with one transducer in the near field resolution measurement. Defective results in measuring the depth of penetration, beam profile, axial resolution, or lateral resolution were not found in any of the 66 transducers.

Although the sizes of the groups, utilized to estimate the limits for the defective values in the depth of penetration and axial and lateral resolution measurements, varied from three to 12, there was no statistical correlation between the sizes of the groups and the standard deviations of the measured parameters.

Discussion

This study focused on finding effective phantom-based QA measurements for detecting signal defects when visual image analysis and manual measurements were used. Size and resolution of the monitor in some US scanners may have affected the measurement accuracy. Positioning of the transducer could have contributed to the results, especially with the micro-convex transducers. Also, manual measurements can produce quite large intra- and inter-observer variability (15). The criterion for the defective result in the near field resolution measurement was based on the previous experience and may not have been fair to older transducers.

Unlike in our study, fault detection from phantom measurements is commonly based on baseline values from the acceptance checks and on a long-term follow-up. However, a long-term follow-up does not necessarily help in detecting deteriorating image quality either (16), not even when using the most recommended measurements of depth of penetration or distance accuracy (12). The type of the transducer may also have an impact on the results, especially in the measurements with small target, because of the different number of elements utilized in forming a small part of the image by different transducers.

In conclusion, phantom-based measurements of depth of penetration, beam profile, near field, axial and lateral resolution, and vertical and horizontal distance accuracy did not consistently detect dead or weak transducer elements or channels in US scanners.

References

1. Goodsitt MM, Carson PL, Witt S, et al. Real-time B-mode ultrasound quality control test procedures. Report of AAPM Ultrasound Task Group No. 1. Med Phys 1998;25:1385–1406.
2. American College of Radiology (ACR). ACR technical standard for diagnostic medical physics performance monitoring of real time ultrasound equipment. Reston, VA: ACR, 2011.
3. American Institute of Ultrasound in Medicine (AIUM). Routine Quality Assurance for Diagnostic Ultrasound Equipment. Laurel, MD: AIUM, 2008.
4. Institute of Physics and Engineering in Medicine (IPEM). Quality Assurance of Ultrasound Imaging Systems, Report No 102. York: IPEM, 2010.
5. Kollmann C, Dekorte C, Dudoley NJ, et al. Guideline for Technical Quality Assurance (TQA) of Ultrasoud devices (B-Mode) – Version 1.0 (July 2012). EFSUMB Technical Quality Assurance Group. Ultraschall Med 2012;33:544–549.
6. Tradup DJ, Hangiandreou NJ, Taubel JP. Comparison of ultrasound quality assurance phantom measurements from matched and mixed scanner–transducer combinations. J Appl Med Phys 2003;4:239–247.
7. Balbis S, Meloni T, Tofani S, et al. Criteria and scheduling of quality control of B-mode and Doppler ultrasonography equipment. J Clin Ultrasound 2012;40:167–173.
8. Thijssen JM, Weijers G, de Korte CL. Objective performance testing and quality assurance of medical ultrasound equipment. Ultrasound Med Biol 2007;33:460–471.
9. Mårtensson M, Olsson M, Segall B, et al. High incidence of defective ultrasound transducers in use in routine clinical practice. Eur J Echocardiogr 2009;10:389–394.
10. Mårtensson M, Olsson M, Brodin L-Å. Ultrasound transducer function: annual testing is not sufficient. Eur J Echocardiogr 2010;11:801–805.
11. Sipilä O, Mannila V, Vartiainen E. Quality assurance in diagnostic ultrasound. Eur J Radiol 2011;80:519–525.
12. Hangiandreou NJ, Stekel SF, Tradup DJ, et al. Four-year experience with a clinical ultrasound quality control program. Ultrasound Med Biol 2011;37:1350–1357.
13. FirstCall aPerio^{TM} Operator's manual. 11552 Revision A. Sonora Medical Systems, Inc.
14. CIRS. General purpose multi-tissue ultrasound phantom, Model 040: User guide and technical information.
15. Sipilä O, Blomqvist P, Jauhiainen M, et al. Reproducibility of phantom-based quality assurance parameters in real-time ultrasound imaging. Acta Radiol 2011;52:665–669.
16. Dudley NJ, Griffith K, Houldsworth G, et al. A review of two alternative ultrasound quality assurance programmes. Eur J Ultrasound 2001;12:233–245.