

First Time Implementation of a Physics Ultrasound (US) Testing Program: Analysis and New Lessons

K. Stiles M.S., R. Makkia Ph.D, D. Vergara M.S., DABR, A. Dohatcu Ph.D, DABR

Yale New Haven Hospital Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, CT



Background

Ultrasound (US) Imaging Annual Equipment Performance Evaluation is required for American College of Radiology® (ACR)-accredited facilities since 2015¹. Failure rates for US equipment have previously been evaluated for multi-year longitudinal trends², optimal frequency for testing in order to detect failures^{3,4}, and for typical passing rates⁵.

Purpose

During the implementation of a new annual US Quality Control (QC) testing program by our healthcare system-wide in-house diagnostic medical physics team, an engineer made a comment that our failure rates were higher than other clients'. Our aim is to investigate our failure rates after the first year of physics testing, compare our results to published data, and determine any trends in tester performance.

Methods

The annual QC tests performed were: **Physical and Mechanical Inspection**, **Image Uniformity and Artifact Survey**, and scanners' **Electronic Image Display Performance**. Per ACR, **Image Uniformity** assessment of transducers consists of three checks, verifying that:

- (A) "The average brightness at the edge of the scan is the same as that in the middle."
- (B) "There are no vertically or radially oriented shadows from array element dropout."
- (C) "There are no brightness transitions between focal zones."

These checks are graded on a scale from 1 to 3, meaning pass, conditional pass with requirements, and fail, respectively. The other tests were based on pass/fail criteria. A summary of these tests is shown in Table 1.

Table 1: Summary of tests and their corresponding criteria and grading methods as per the ACR¹

Test	Criteria	Grading
Physical and Mechanical Inspection	Transducers	Check cable, housing, and transmitting surface for cracks, separations, discoloration.
	Transducers parts	Check for connectivity. All transducer ports should be tested with at least one transducer.
	Power cord	Check for cracks, discoloration or damage to cable & plug.
	Control	Check operation of switches & knobs, note any burnt out bulbs.
	Monitor, VCR, Camera	Check for cleanliness and operation of controls, and that they are fastened securely to the ultrasound equipment.
Image Uniformity and Artifact Survey	Wheels	Check the wheel locks in working condition.
	Dust filters	Check for cleanliness; clean or replace if necessary.
Ultrasound Electronic Image Display Performance	SMPT Pattern Calibration	Verify maximum luminance > 90 cd/m ² , minimum luminance < 1.2 cd/m ² , all greyscale steps are properly calibrated.
	SMPT Pattern Observation	0-5% contrast visible 95-100% contrast visible Each grey level step visible Borders and lines are straight No distortion or misalignment Alphanumeric characters look sharp High-contrast line pair visible No streaking in uniform rectangles
UN-80 Pattern	Luminance % difference non-uniformity is ≤ 15%	1- Pass 2- Conditional Pass 3- Fail

Our first annual QC was performed on 106 different units of 4 different manufacturers, 11 different models, and a total of 451 transducers. The US equipment tested is summarized in Table 2.

The phantom used for testing image uniformity was the CIRS® (Norfolk, VA, USA) General Purpose Ultrasound Phantom Model 054GS. The RaySafe™ (Billdal, Sweden) X2 Light Sensor was used to test acquisition display image performance.

Methods (cont.)

Testing data was tracked not only in regards to failure rate but also by tester (Figure 1). Some of the physicists who tested were experienced in US scanner evaluation, while others were not and still in training. Years of experience were defined by number of years testing at least 15 US units every 12 months.

Table 2: Our number of ultrasound units and transducers tested, broken down by vendor and model.

Vendors and Models	Number of Units	Number of Transducers
BK Medical	1	1
Flex Focus 400	1	1
Philips	87	361
Affinity	1	4
CX50	1	1
EPIQ SG	17	75
EPIQ 7G	6	32
EPIQ Elite	8	32
IE33	8	29
iU22	45	194
Siemens	15	66
Acuson S2000	14	64
Juniper	1	2
Zonare	4	17
ZS3	4	17
Grand Total	106	451

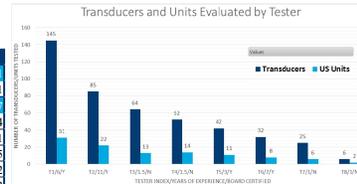


Figure 1: Number of transducers and US units tested by each tester. Testers are given as their index/years of experience/board certified.

Results

The failure rate for the **Physical & Mechanical Inspection** test was 11%, with over 90% of the failures being attributed to mechanical faults of the unit. The failure rates for the **Image Uniformity and Artifact Survey** checks for A, B, and C were 0.7%, 13.1%, and 0.4%, respectively.

Out of the total, 24.3% of units failed at least part of the **Ultrasound Electronic Image Display Performance** test, with the vast majority of those failures attributed to the SMPT Greyscale Calibration test. The breakdown of testing results is shown in Table 3.

Table 3: Results from physics tests broken down by testers T1-T8. Corresponding columns represent the number of tests per physicist that pass and fail, respectively. (Percentage of tester total). Totals for our physics group and literature are shown in the last two rows.

Tester	Physical and Mechanical Inspection		SMPT Greyscale Calibration		SMPT Pattern Observation		UN-80 Pattern	
	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail
T1	27 (87.1%)	4 (12.9%)	23 (74.2%)	8 (25.8%)	31 (100%)	0 (0%)	30 (96.8%)	1 (3.2%)
T2	17 (77.3)	5 (22.7%)	19 (86.4%)	3 (13.6%)	22 (100%)	0 (0%)	21 (95.5%)	1 (4.5%)
T3	12 (92.3%)	1 (7.7%)	10 (76.9%)	3 (23.1%)	13 (100%)	0 (0%)	13 (100%)	0 (0%)
T4	13 (92.9%)	1 (7.1%)	10 (74.4%)	4 (28.6%)	14 (100%)	0 (0%)	14 (100%)	0 (0%)
T5	11 (100%)	0 (0%)	10 (90.1%)	1 (9.1%)	11 (100%)	0 (0%)	10 (90.9%)	1 (9.1%)
T6	7 (100%)	0 (0%)	7 (100%)	0 (0%)	7 (100%)	0 (0%)	7 (100%)	0 (0%)
T7	5 (83.3%)	1 (16.7%)	3 (50%)	3 (50%)	6 (100%)	0 (0%)	6 (100%)	0 (0%)
T8	2 (100%)	0 (0%)	1 (50%)	1 (50%)	2 (100%)	0 (0%)	1 (50%)	1 (50%)
Group	94 (88.7%)	12 (11.3%)	83 (78.3%)	23 (21.7%)	106 (100%)	0 (0%)	102 (96.2%)	4 (3.8%)
Literature	25.1% ²		-		-		-	

Tester	Transducer Check A			Transducer Check B			Transducer Check C		
	Pass	Conditional Pass	Fail	Pass	Conditional Pass	Fail	Pass	Conditional Pass	Fail
T1	142 (97.9%)	2 (1.4%)	1 (0.7%)	95 (65.5%)	35 (24.1%)	15 (10.3%)	135 (93.1%)	9 (6.2%)	1 (0.7%)
T2	85 (100%)	0 (0%)	0 (0%)	40 (47.1%)	25 (29.4%)	20 (23.5%)	85 (100%)	0 (0%)	0 (0%)
T3	55 (85.9%)	7 (10.9%)	2 (3.1%)	50 (78.1%)	8 (12.5%)	6 (9.4%)	53 (82.8%)	11 (17.2%)	0 (0%)
T4	52 (100%)	0 (0%)	0 (0%)	34 (65.4%)	11 (21.2%)	7 (13.5%)	52 (100%)	0 (0%)	0 (0%)
T5	33 (78.6%)	9 (21.4%)	0 (0%)	31 (73.8%)	7 (16.7%)	4 (9.5%)	40 (95.2%)	2 (4.8%)	0 (0%)
T6	28 (87.5%)	4 (12.5%)	0 (0%)	26 (81.3%)	5 (15.6%)	1 (3.1%)	26 (81.3%)	5 (15.6%)	1 (3.1%)
T7	25 (100%)	0 (0%)	0 (0%)	16 (64%)	5 (20%)	4 (16%)	23 (92%)	2 (8%)	0 (0%)
T8	6 (100%)	0 (0%)	0 (0%)	3 (50%)	1 (16.7%)	2 (33.3%)	6 (100%)	0 (0%)	0 (0%)
Group	426 (94.5%)	22 (4.9%)	3 (0.7%)	295 (65.4%)	97 (21.5%)	59 (13.1%)	420 (93.1%)	29 (6.4%)	2 (0.4%)
Literature	13.90% ²			8			27.1% ⁴		

Results (cont.)

Compared with literature, table 3 shows that the failure rates that we observed for our US equipment were much lower for **Physical and Mechanical Inspection** (11.3% vs. 25.1%)² and similar or lower for the **Image Uniformity and Artifact Survey** tests (13.1% vs 13.9%/27.1%)⁴. No literature was found to compare with our results for the scanners' **Electronic Image Display Performance**. More research is needed to determine whether the correlation is due to tester performance or equipment bias.

Figure 2B demonstrates that testers with fewer years of experience or those that were not yet board certified are more likely to fail **Image Uniformity and Artifact Survey** check B than to conditionally pass it. There were insufficient grades of conditionally pass and failure for checks A and C to draw a conclusion.

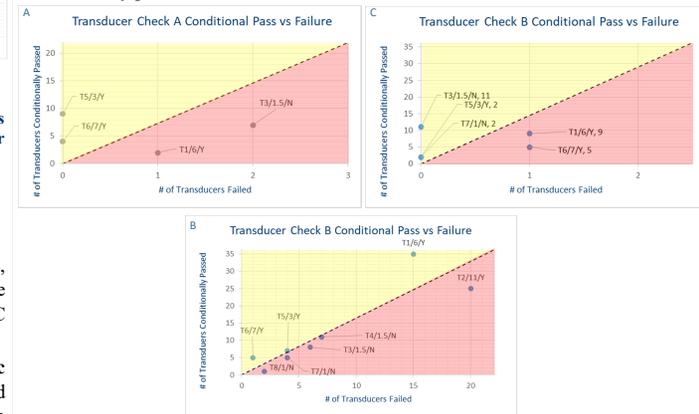


Figure 2: Ratio of Conditional Pass/Failure for Uniformity and Artifact Survey test A, B, and C for each Tester (Tester/Years of Experience/Board Certified). Being in the red section corresponds to a higher rate of failures as opposed to conditional pass than the group. Being in the yellow section corresponds to a higher rate of conditional passes as opposed to failures than the group. The midline (dotted) represents the rate at which our physics group conditionally passes vs. fails.

Figure 2 is not intended to evaluate individual testers' quality of testing, but rather to illustrate their conditional passing rate against the group rate.

Conclusions

The review of a new US QC annual physics testing program at our institution was performed. The engineer's comment was investigated, and our failure rates were found to be on the lower end compared to literature. Further research is needed into tester bias and other confounding variables.

References

- ¹American College of Radiology®, Reston, VA, USA.
- ²N. Hangiandreou, S. Stelck, D. Tradup, K. Gorny, D. King, *Ultrasound in Medicine & Biology*, Volume 37, Issue 8, 2011 / <http://amos3.aapm.org/abstracts/pdf/67-17535-89300-119.pdf> (last accessed on 28 February 2022)
- ³M. Mårtensson, M. Olsson, B. Segall, A. G. Fraser, R. Winter, L. Brodin, *European Journal of Echocardiography*, Volume 10, Issue 3, May 2009.
- ⁴M. Mårtensson, M. Olsson, L. Brodin, *European Journal of Echocardiography*, Volume 11, Issue 9, October 2010.
- ⁵O. Sipilä, V. Mannila, E. Vartiainen, *European Journal of Radiology*, Volume 80, Issue 2, 2011.