MAYO CLINIC

Quality Assurance in Ultrasound

Nicholas J Hangiandreou PhD
hangiandreou@mayo.edu

Donald J Tradup RDMS, Scott F Stekel BS, Zaiyang Long PhD, Jacinta E Browne PhD, Daniel Gomez-Cardona PhD

Mayo Clinic – Rochester MN

AAPM Spring Clinical Meeting
Zagzebski-Carson Distinguished Lectureship On Medical Ultrasound
April 5, 2020

Learning objectives

1. Identify the basic steps in a team-based approach to assessing ultrasound imaging systems prior to purchase
2. Understand current techniques for routine quality control
3. Describe emerging techniques in ultrasound quality assurance

NJH has no conflicts of interest to disclose.

Overview of the elements of an ultrasound quality assurance (QA) program

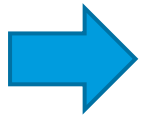
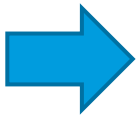
1. **Pre-purchase scanner evaluation**
2. Acceptance testing
3. Initial set-up of measurement package and DICOM SR, and
4. Cross-calibration of quantitative measurement tools, betwe
5. Initial preset/ image quality optimization
6. Configuration management of scanner fleet
7. **Quality control and accreditation maintenance**
8. On-going image quality optimization and troubleshooting
9. Evaluation and translation of new imaging techniques into clinical practice
10. On-going participation in practice efficiency and quality improvement initiatives
11. Ultrasound physics and technology education for staff and trainees

Pre-purchase ultrasound scanner evaluation: A team-based approach

- Last year our Vascular and General ultrasound practices initiated a fleet-replacement effort with the goal of purchasing 45-50 new, premium US scanners over a ~2 –year period
- Our ultrasound physics team proposed a comprehensive evaluation process for assessing many potential candidate scanners and identifying the one(s) best suited for our clinical practice
  - Traditional “bring it in and try it out” approach but with more preparation and data gathering
- Employing a team…
  - Shares work so no group is overwhelmed
  - Builds ownership in the purchase decision across the practice
  - Assess aspects of system performance that physics can not effectively do, e.g. evaluating usability and ergonomics

- **Upon completion, leaders of the radiologist, sonographer, and administrative groups reviewed with their groups a summary of the process, results, and decision (or asked physics to do so) → *Success***

Evaluation tasks

• Vendor communication and logistics of on-site assessment	• Administration
• Technical questionnaire	• Physics team
• Safety testing, scanner set-up for patient scanning, networking to PACS	• Equipment service engineer
• Scanning patients side-by-side with current clinical scanner, with image comparison in PACS and data collection	• Physics team (preparation)
• Usability and ergonomics	• Sonographers
 • **Subjective assessment of image performance for clinical tasks**	• Radiologists and sonographers
• Lab testing of specialized functionality	• Sonographers and physics team
• Scanning volunteers side-by-side with current clinical scanner, with image comparison in PACS and data collection	• Physics team and IT (preparation), sonographers, radiologists
 • **Objective image performance assessment using phantoms**	• Physics team

Subjective assessment of image performance for clinical tasks

- List specific image views from clinical exam protocols, for side-by-side back-scanning with candidate scanners
  - Emphasize clinical utility, not aesthetic preference
- Rating form that benchmarks performance vs current scanner

GENERAL IMAGING
1. Abdomen
Long Distal Aorta
Long Liver / IVC
Trans Lt Liver (showing IVC & LHV)
Trans Rt Liver (high showing dome), both subcostal and intercostal
MPV Gray Scale
Long GB
CHD / CBD
Long Liver / Rt Kidney
Long Rt Kidney
Trans Lt Liver Linear Transducer
Trans Rt Liver Linear Transducer
2. Liver Transplant
(Include all of the above listed abdomen images with these color and spectral images)
MHA
MPV
HV's

VASCULAR
1. Iliac Veins & Arteries (Prerenal Tx, Stents, Grafts etc.)
(If bilateral choose side with more disease)
CIA Bif Grayscale
CIA Bif Color
EIV Upper Spectral with Color
If stent is present: Grayscale at Endpoints
2. EVAR (general scan & contrast)
Sac Area with Color (looking for leak)
Prox Endpoint Grayscale and Color
Dist Endpoint Grayscale and Color
IMA Area
3. Hemodialysis Access
AVF Spectral with Color (check scale range)
Brachial Art Dist to AVF Spectral with Color
Vein Confluence Spectral
SCV Spectral with Color
Flow Volume

Radiologist US Vendor Trial Feedback Form: Vendor Name

(To be filled out for each patient)

Pt. Clinic #: _____ Exam Type: _____

Date: _____ Initials: _____

1. How does overall image quality compare with current machine name?

☐ Unacceptable ☐ Worse, but acceptable ☐ Same ☐ Superior

2. Was there particular anatomy/pathology seen better or worse with the trial scanner?

Trial Scanner Image (Identify by anatomy, image number or annotation)	Compared to the equivalent current machine name image is the image				Comment on difference
	Unacceptable	Worse, but acceptable	Equivalent	Better	

Additional Comments (please use back of sheet if needed):

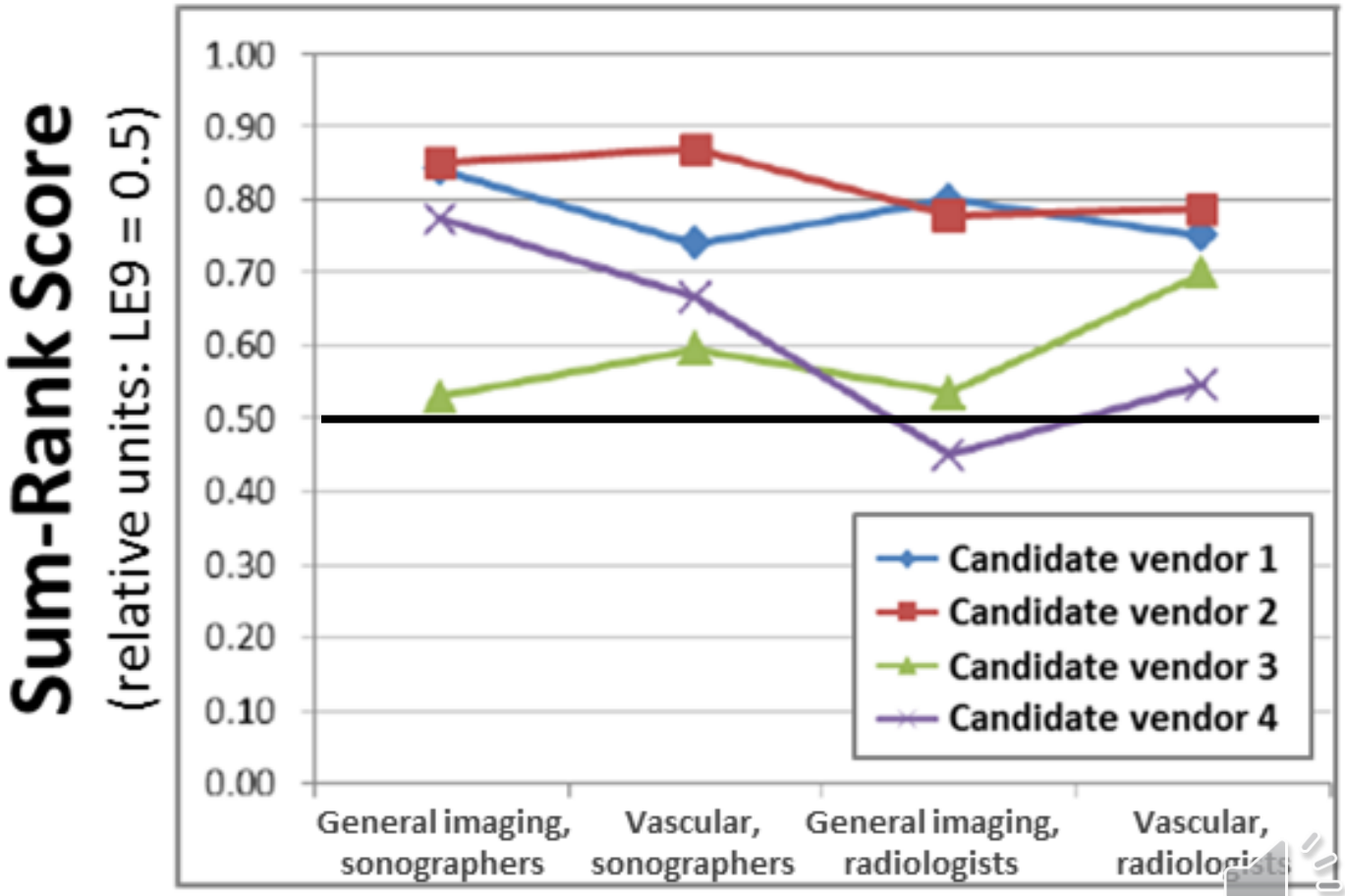
Statistical analysis of subjective image quality rating data

- One week evaluations in each imaging area for each candidate scanner yielded n~20 sets of sonographer and radiologist feedback forms
  - Statistical hypothesis testing can be performed, and significant differences are seen (highlighted values below)
  - All performance measures are benchmarked against that of the current clinical scanner

Sum-Rank Score

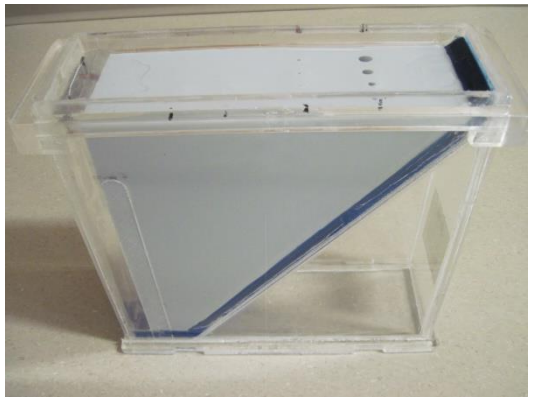
	Gen-S	Vasc-S	Gen-R	Vasc-R
CV 1	0.84	0.74	0.80	0.75
CV 2	0.85	0.87	0.78	0.79
CV 3	0.53	0.60	0.54	0.70
CV 4	0.77	0.67	0.45	0.55

(dummy data)

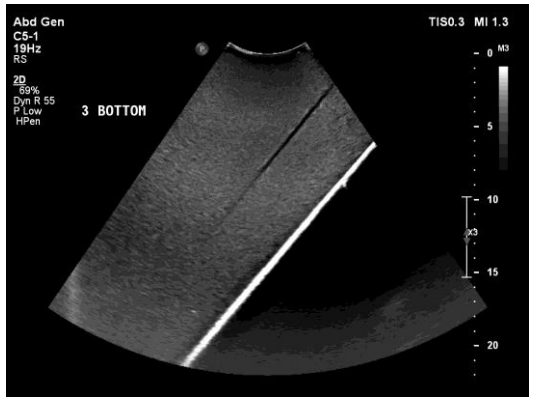


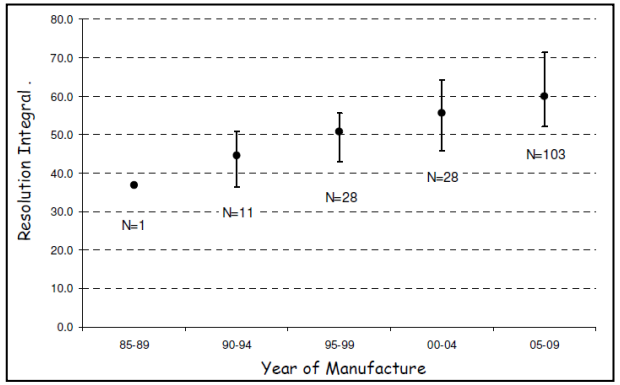
Objective image performance assessment using phantoms

- Primary emphasis on task-based performance, e.g. based on imaging of echogenic or anechoic spherical targets or cylinders
  - Ideally a single performance metric could be computed, integrating together multiple aspects of image quality
- Our group is working with the **Resolution Integral** measured using the Edinburgh pipe phantom, as the basic measure of scanner performance
  - Well-described in the literature, e.g. Moran, Inglis, and Pye; "The Resolution Integral – a tool for characterizing the performance of diagnostic ultrasound scanners," Ultrasound 2014; 22:37-43)



Edinburgh Pipe Phantom (EPP)

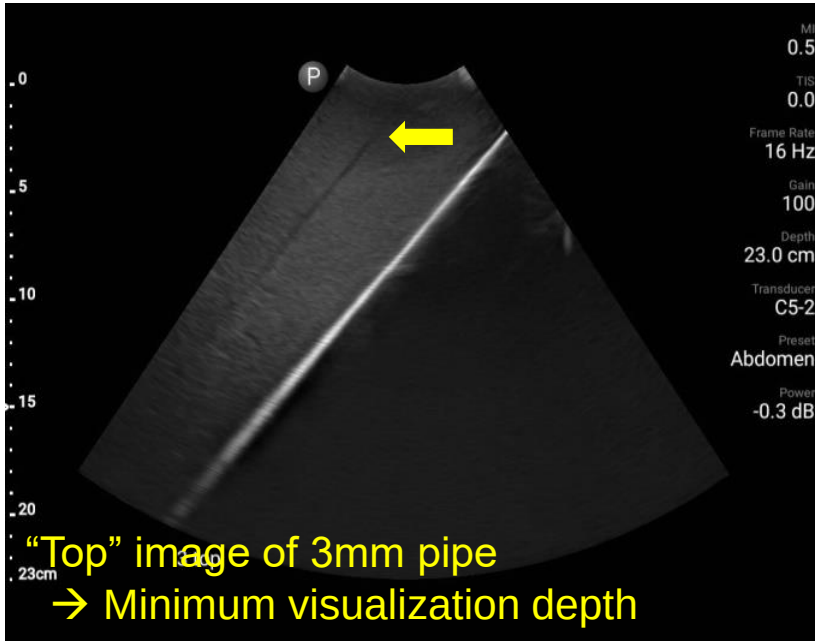


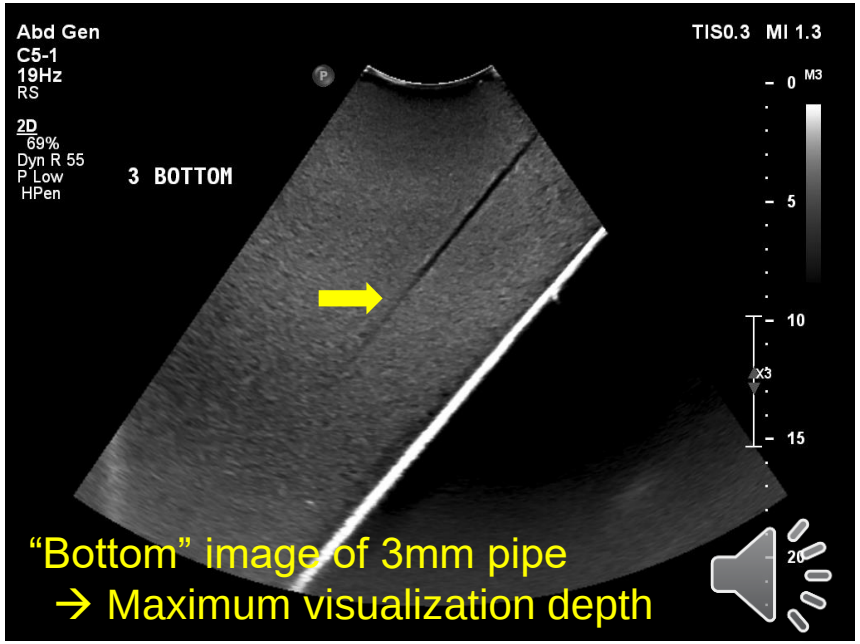


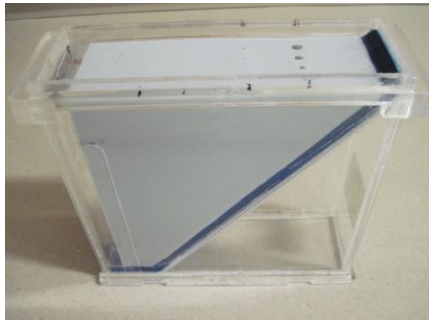
Pye SD and Ellis W, *Journal of Physics,* 2011

Resolution integral measurement process

- The original resolution integral approach involves visually determining the depth ranges over which cylindrical anechoic targets (“pipes”) of different diameters can be visualized in the Edinburgh Pipe Phantom.
- The depth limits of visibility are evaluated by visual inspection of pipe images separately adjusted to optimize visualization at the minimum and maximum depths
  - This is done for all pipe diameters present in the phantom (8mm, 6mm, 4mm, 3mm, 2mm, 1.5mm, 1mm, 0.5mm, and 0.4mm)
  - The depth range of visualization is then calculated for each pipe diameter as the difference of the maximum and minimum visualization depths



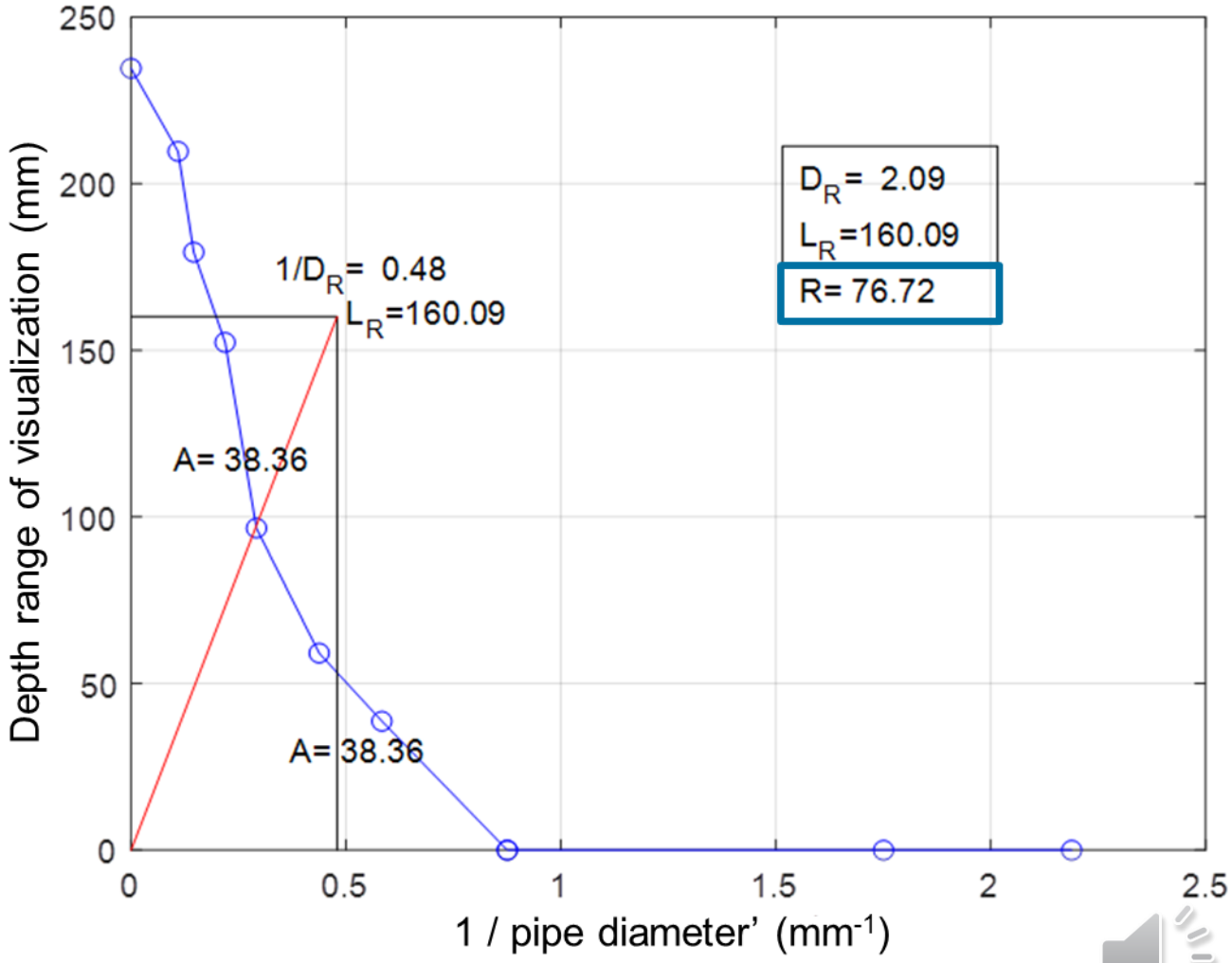




MAYO CLINIC

Resolution integral measurement process (continued)

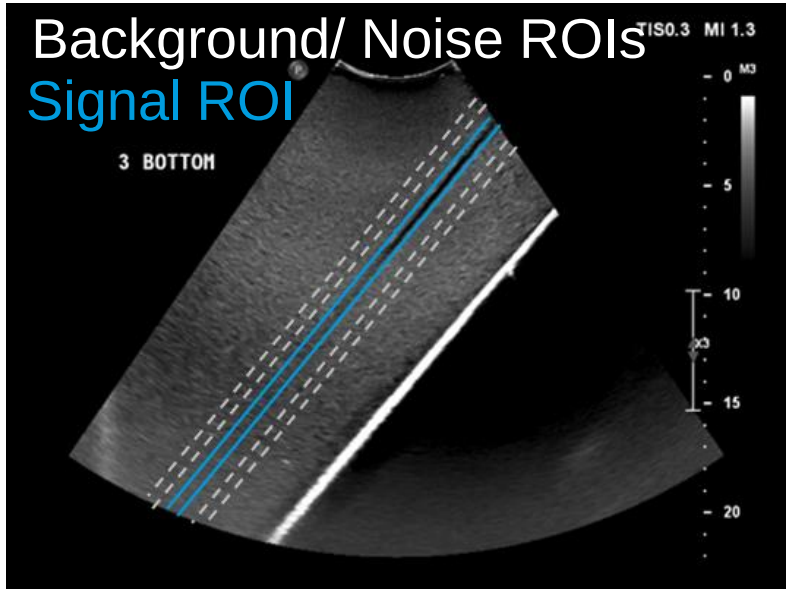
- **Overall system performance is described by the Resolution Integral, R, which aggregates visualization capability over all pipes:**
  - Depth range of visualization for each pipe diameter is plotted against *the inverse* of the pipe diameter
  - These data points form a curve bounded on both x- and y-axes
  - **The unit-less resolution integral value, R, is equal to the area under this curve**
    - The bisector of this area can be used to determine characteristic spatial resolution (D_R) and depth of field (L_R), which can distinguish transducers used for different applications, e.g. abdominal or small parts

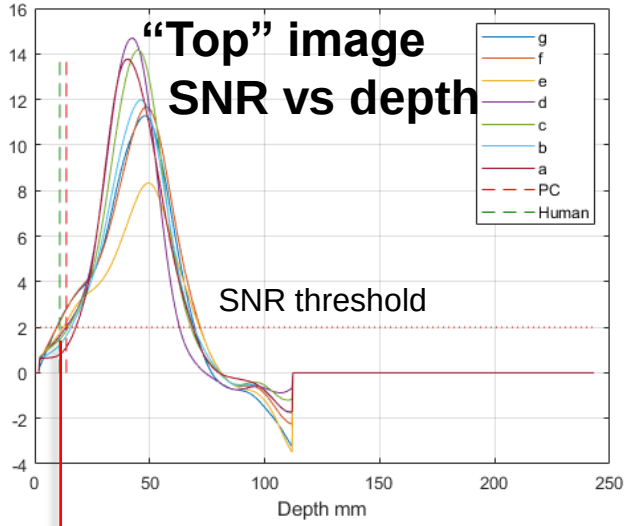


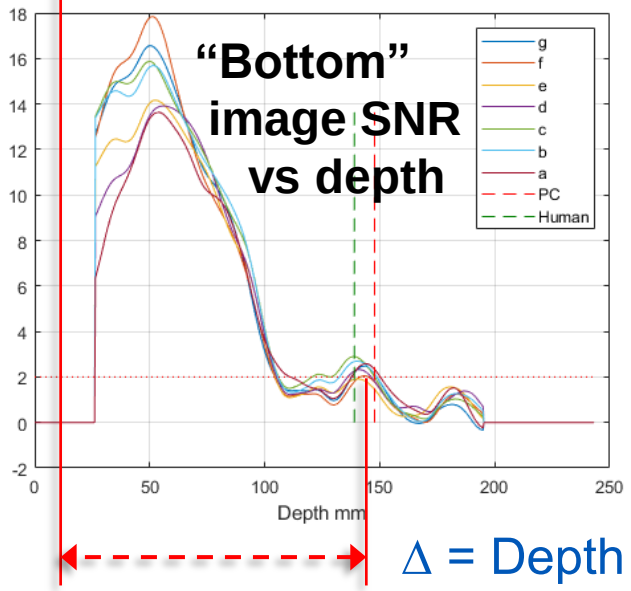
MAYO CLINIC

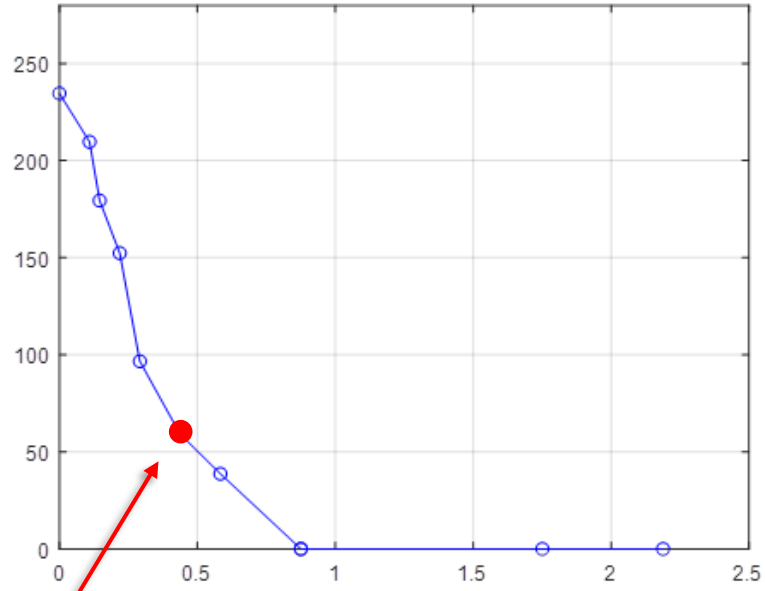
Objective SNR-based determination of depth range of “visualization”

For each pipe diameter, we acquire multiple images of the pipe and background gel





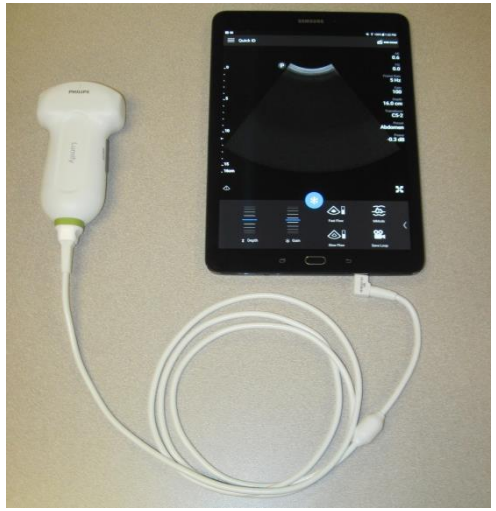


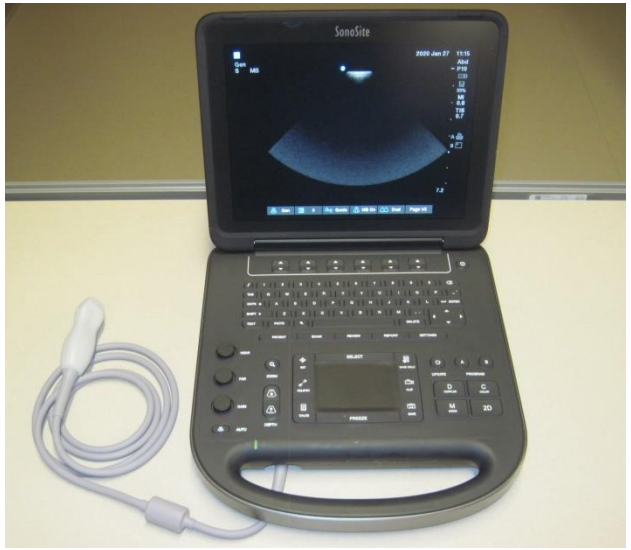


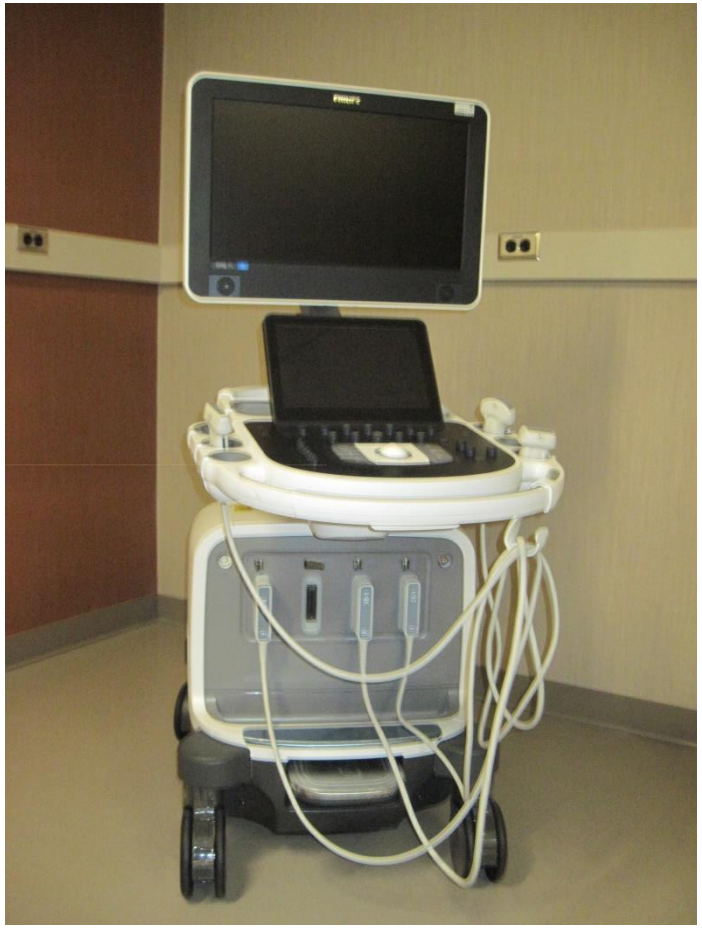
Δ = Depth range of “visualization”

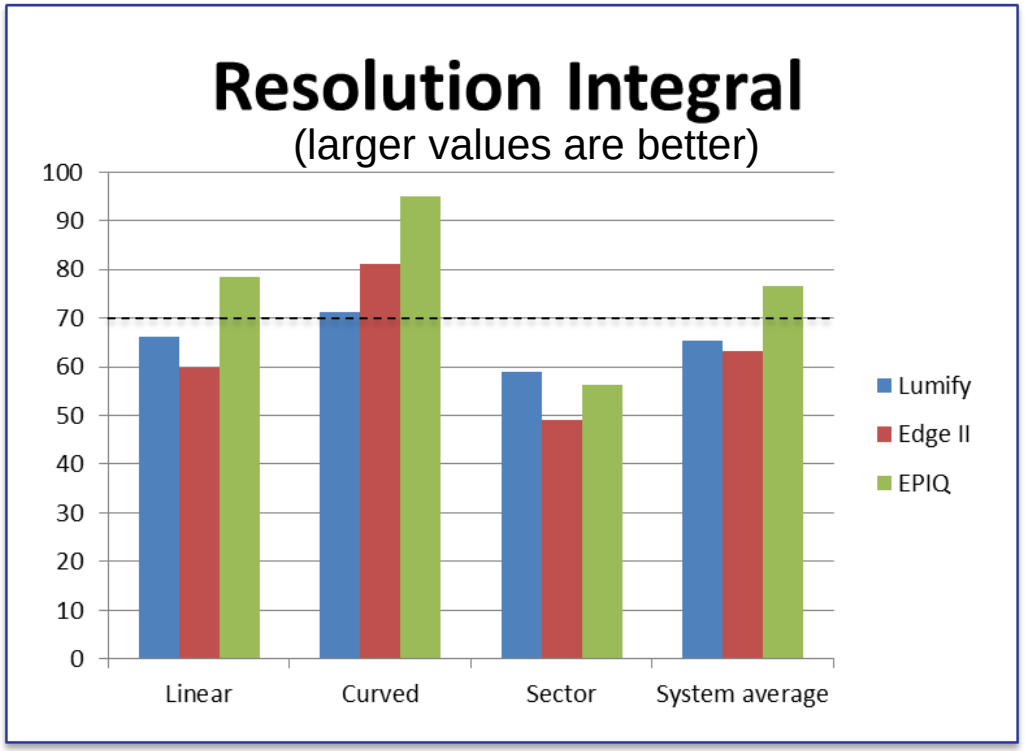
MAYO CLINIC

Sample resolution integral results (visual image assessment)







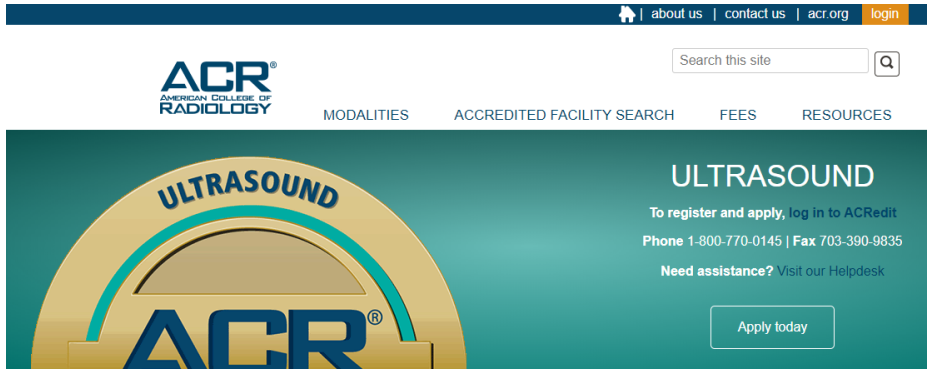


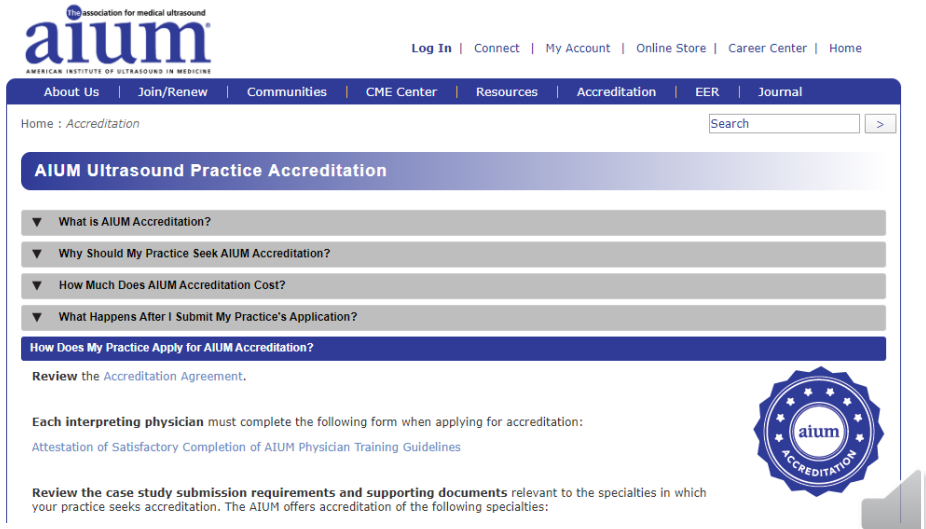
Resolution integral measurements for tablet-based Philips Lumify, laptop-form factor Sonosite Edge II, and premium Philips EPIQ ultrasound scanners. A dashed line is shown for R=70, which is an estimated general reference performance level for systems tested between 2015 and 2019 (extrapolated from Pye and Ellis, Journal of Physics, 2011)

Quality control and accreditation maintenance: Approaches for providing services remotely

- What annual physics services are required by ACR and/or AIUM Ultrasound Accreditation programs?
  - Uniformity assessment/ artifact survey
  - Monitor brightness and calibration, overall display quality
    - Scanner display
    - Primary interpretation workstation
  - Mechanical inspection of transducers and scanner
  - System sensitivity/ maximum depth of visualization
  - Distance measurement accuracy
  - Contrast resolution (optional)
  - Spatial resolution (optional)

- Is annual testing really quality control?
- Could (some) tests be performed remotely?





Quality control and accreditation maintenance: Approaches for providing services remotely

- Assessment of image uniformity and presence of artifacts is the most productive US QC test we do
- These artifacts tend *not* to be reported by clinical users

ELSEVIER

Ultrasound in Med. & Biol., Vol. 37, No. 8, pp. 1350–1357, 2011

0301-5629/$ - see front matter

doi:10.1016/j.ultrasmedbio.2011.05.007

● *Technical Note*

FOUR-YEAR EXPERIENCE WITH A CLINICAL ULTRASOUND QUALITY CONTROL PROGRAM

NICHOLAS J. HANGIANDREOU, SCOTT F. STEKEL, DONALD J. TRADUP, KRZYSZTOF R. GORNY, and DEIRDRE M. KING

Department of Radiology, Mayo Clinic, Rochester, MN, USA

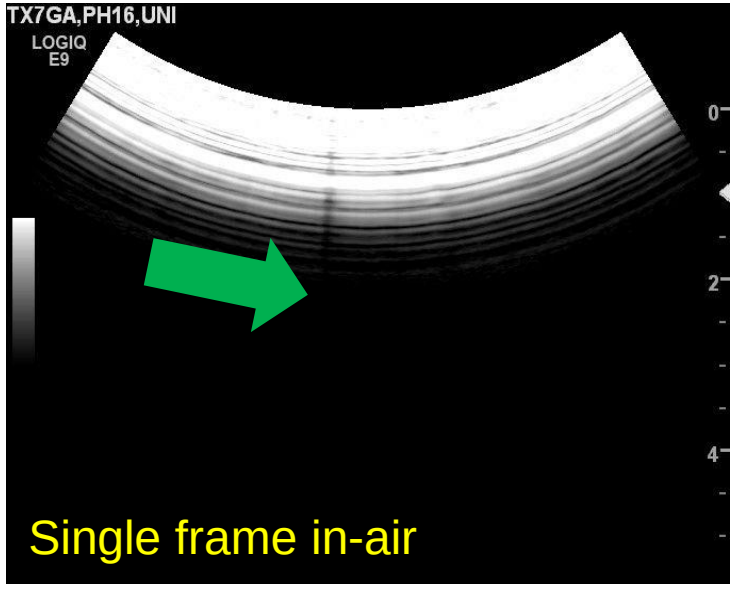
(*Received* 11 *March* 2011; *revised* 6 *May* 2011; *in final form* 9 *May* 2011)

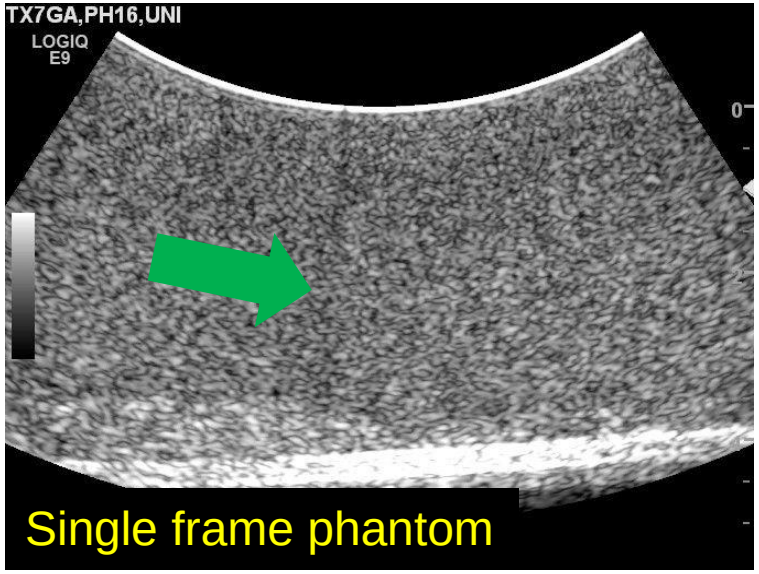
Table 4. The number and percentage of equipment failures discovered by each detection method over the 4-year analysis period

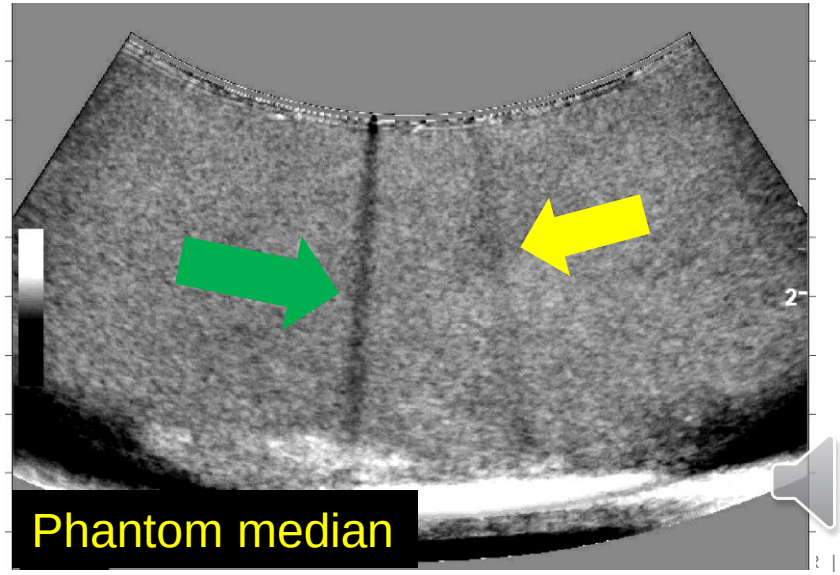
Evaluation method	Number of detected failures	Percent of detected failures
Mechanical integrity	47	25.1
Image uniformity	124	66.3
Distance measurement accuracy	0	0.0
Maximum depth of penetration	3	1.6
Clinical use	13	7.0
Total	187	100.0

Assessing uniformity with phantoms

- Use soft, uniform phantoms that can couple to entire face of curved probes
- Inspect phantom images while scanning live and moving the probe to acquire images of changing speckle field, to smooth out speckle, increasing sensitivity
  - Optimize scan parameters to maximize sensitivity to artifacts
  - Also inspect in-air images
- Can also store clips of phantom images and process to generate single frame image showing the median value across the frames at each pixel location – *smoothing of speckle increases sensitivity*

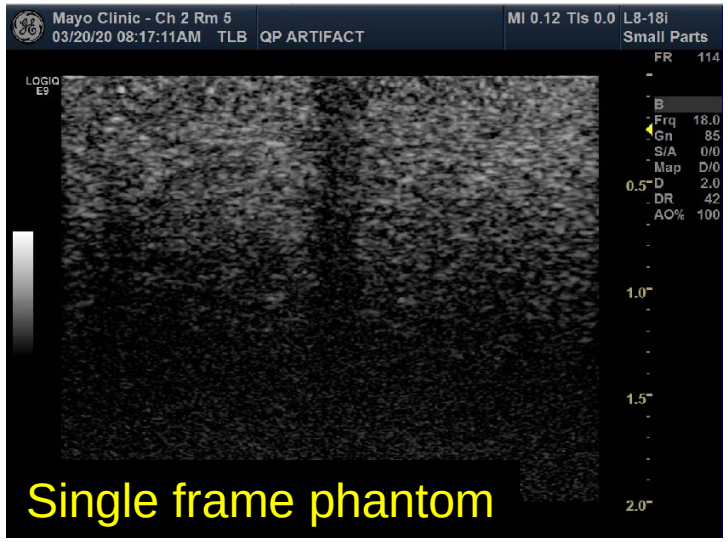


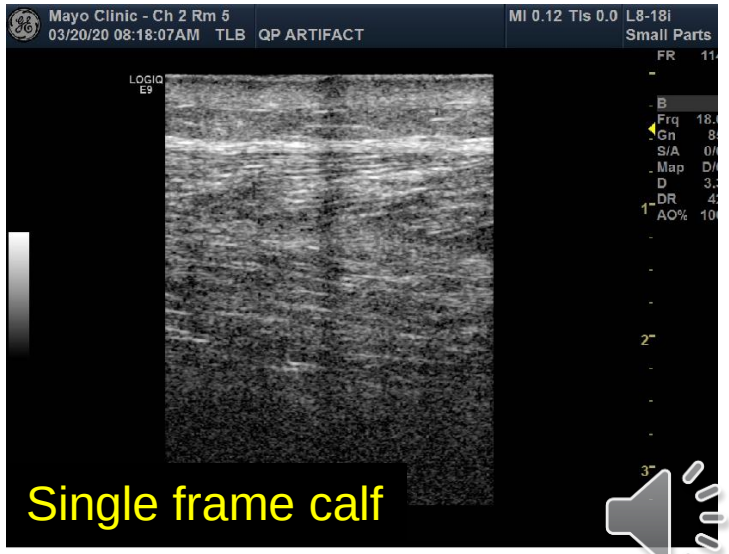




Potential pitfalls in uniformity testing

- Is the artifact due to an actual equipment defect?
  - Inspect the probe face for debris
  - Assure that the probes is properly coupled to the phantom, and no bubbles are present
  - Remove and re-seat the probe in its connection port to assure no dust or debris is present
- Is the defect in the probe or the scanner, i.e. the port or channel?
  - Check the probe in other ports (and other scanners if available)
  - Check other probes in the same port
- Is the defect severe enough to warrant failing and replacing the probe?
  - Check artifact while flexing or otherwise manipulating the cable
  - Check artifact conspicuity in image of anatomy
- We have not commonly noted a gradual degradation in artifact severity: These appear abruptly, and get worse abruptly → Damage through use
  - **Frequent uniformity testing would be helpful**
  - Users will not reliably report even severe artifacts





 |

Can uniformity artifacts be detected using clinical images? *Yes!*

Received: 29 November 2016 | Revised: 1 November 2017 | Accepted: 17 November 2017

DOI: 10.1002/acm2.12248

MEDICAL IMAGING

WILEY

Method for automatic detection of defective ultrasound linear array transducers based on uniformity assessment of clinical images — A case study

Robert Lorentsson[1,2] | Nasser Hosseini[1] | Jan-Olof Johansson[1] | Wiebke Rosenberg[1] | Benny Stenborg[1] | Lars Gunnar Månsson[1,2] | Magnus Båth[1,2]

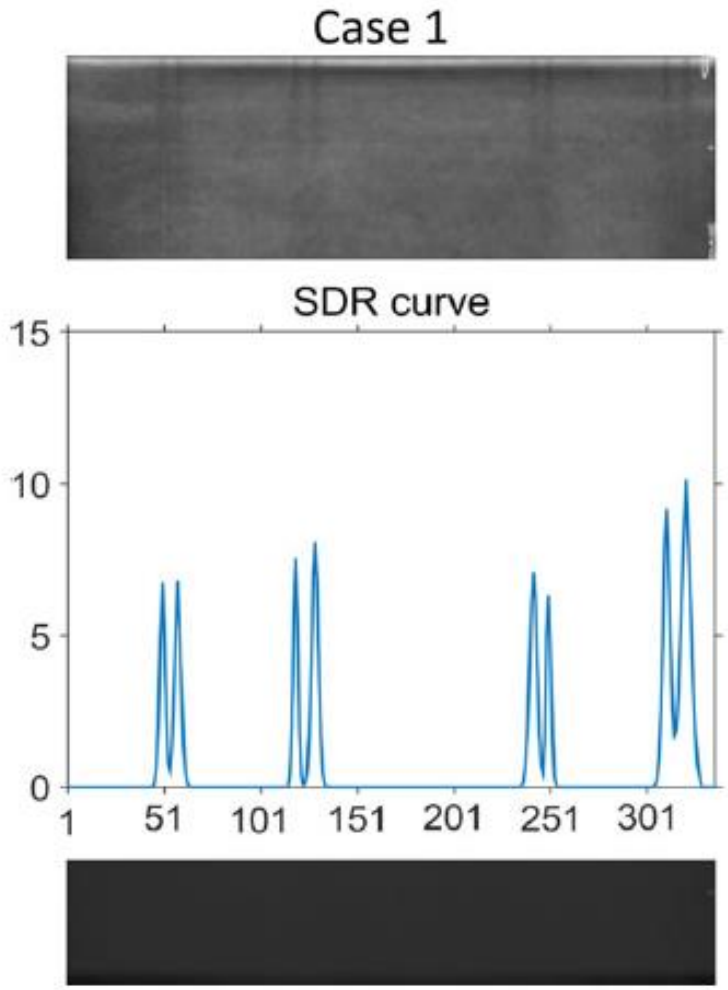
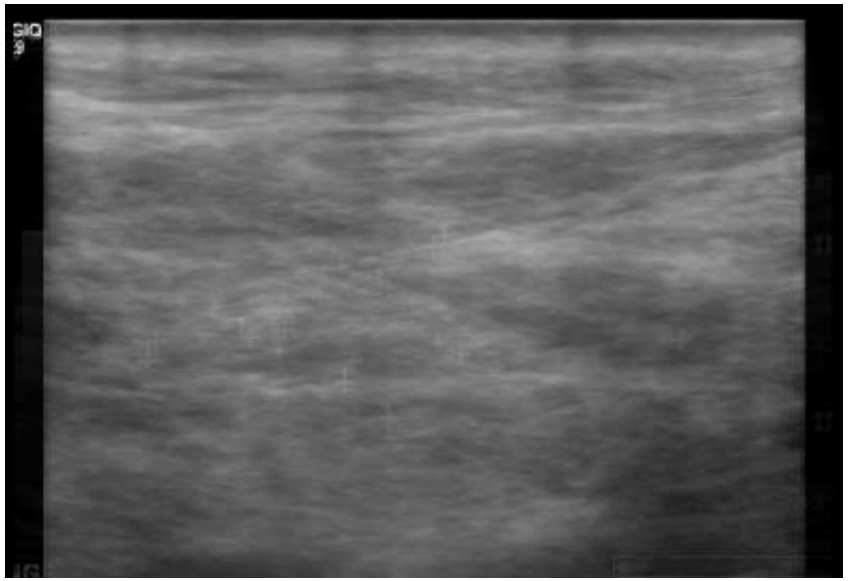


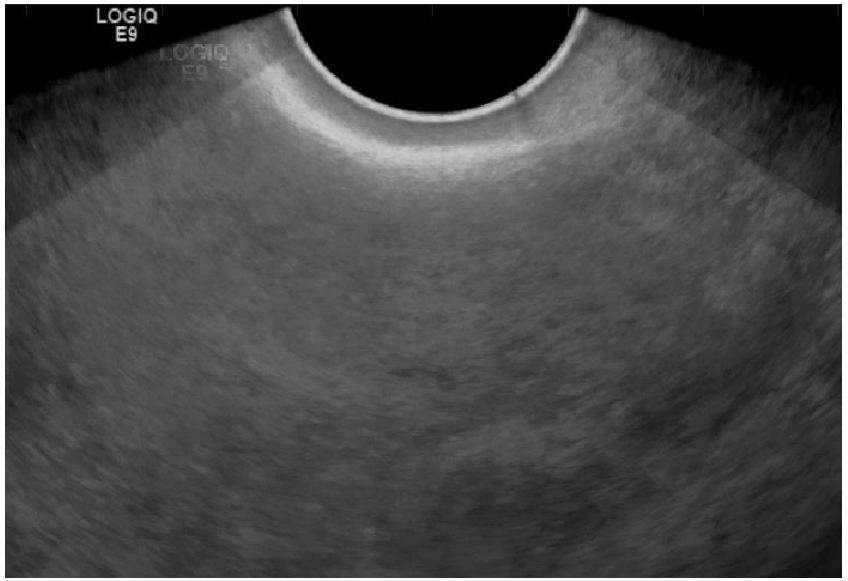
FIG. 4. An example of an SDR curve for Case 1, representing the darker streaks in the median image in the background.

Some key general steps in the automated process…

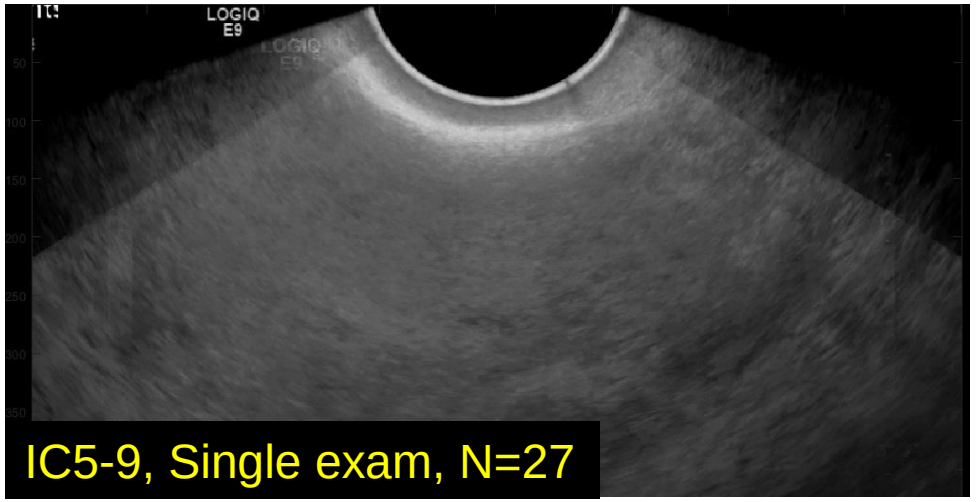
- Obtain a feed of all clinical ultrasound images in DICOM format (LAN or WAN)
- Sort grayscale images from each unique transducer
- Group images for combination into single uniformity image
  - Re-grid to consistent pixel size, and co- register
  - Normalize (increased) contrast level and brightness vs depth
  - Compute median of all pixels at each image location

Visually inspect median images for artifacts

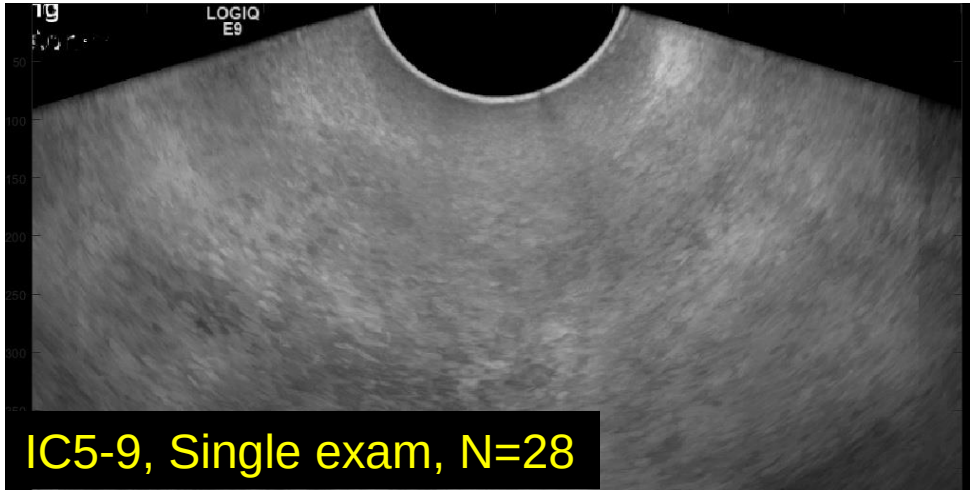




Examples



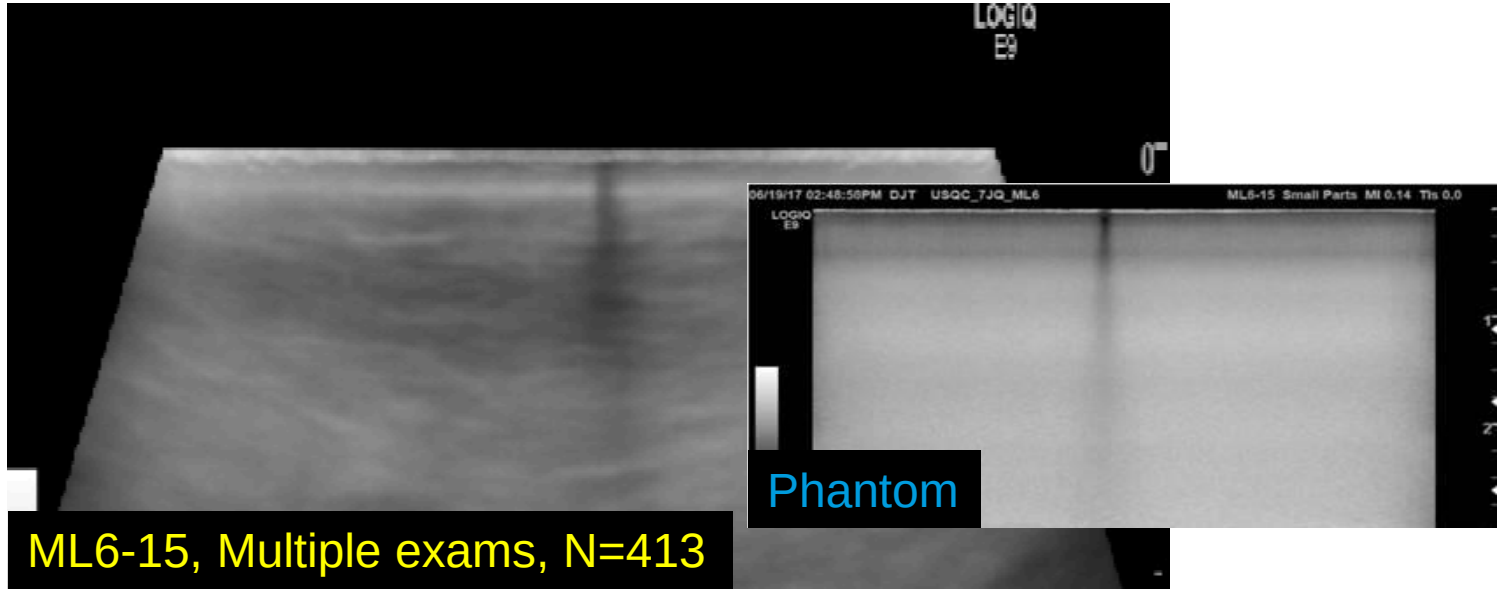
IC5-9, Single exam, N=27



IC5-9, Single exam, N=28

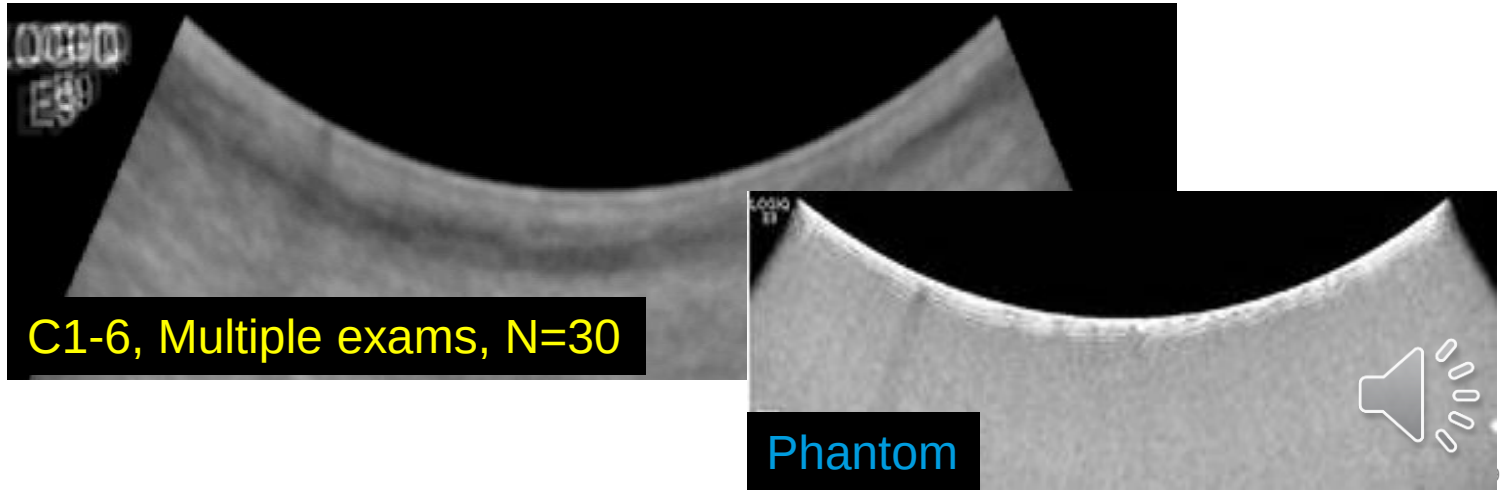
Artifact detected in a probe “pool” shared by multiple scanners

Ferrero et al: Assessing ultrasound probe uniformity from clinical images: proof of feasibility for a variety of probe models. AAPM 2019.



ML6-15, Multiple exams, N=413

Phantom



C1-6, Multiple exams, N=30

Phantom

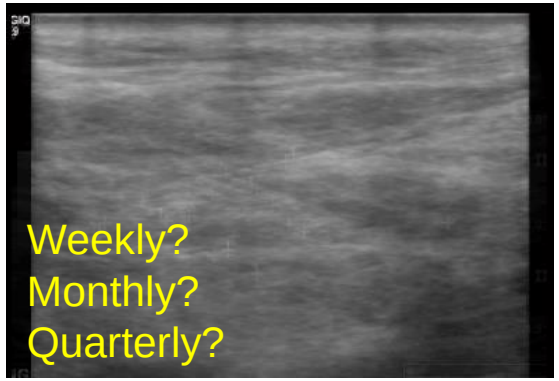
Hurdles to implementation

- Identification of each unique probe (Serial numbers in DICOM header?)
- Identification of US image region for scaling and registration (Pixel mask in DICOM header?)
- How many images to combine?
  - More images → Greater sensitivity and fewer images to review
    Easier automated detection? Fewer false alarms? Less sensitive to flex artifacts
  - Fewer images → Greater specificity for actionable defects

Form both of these images?

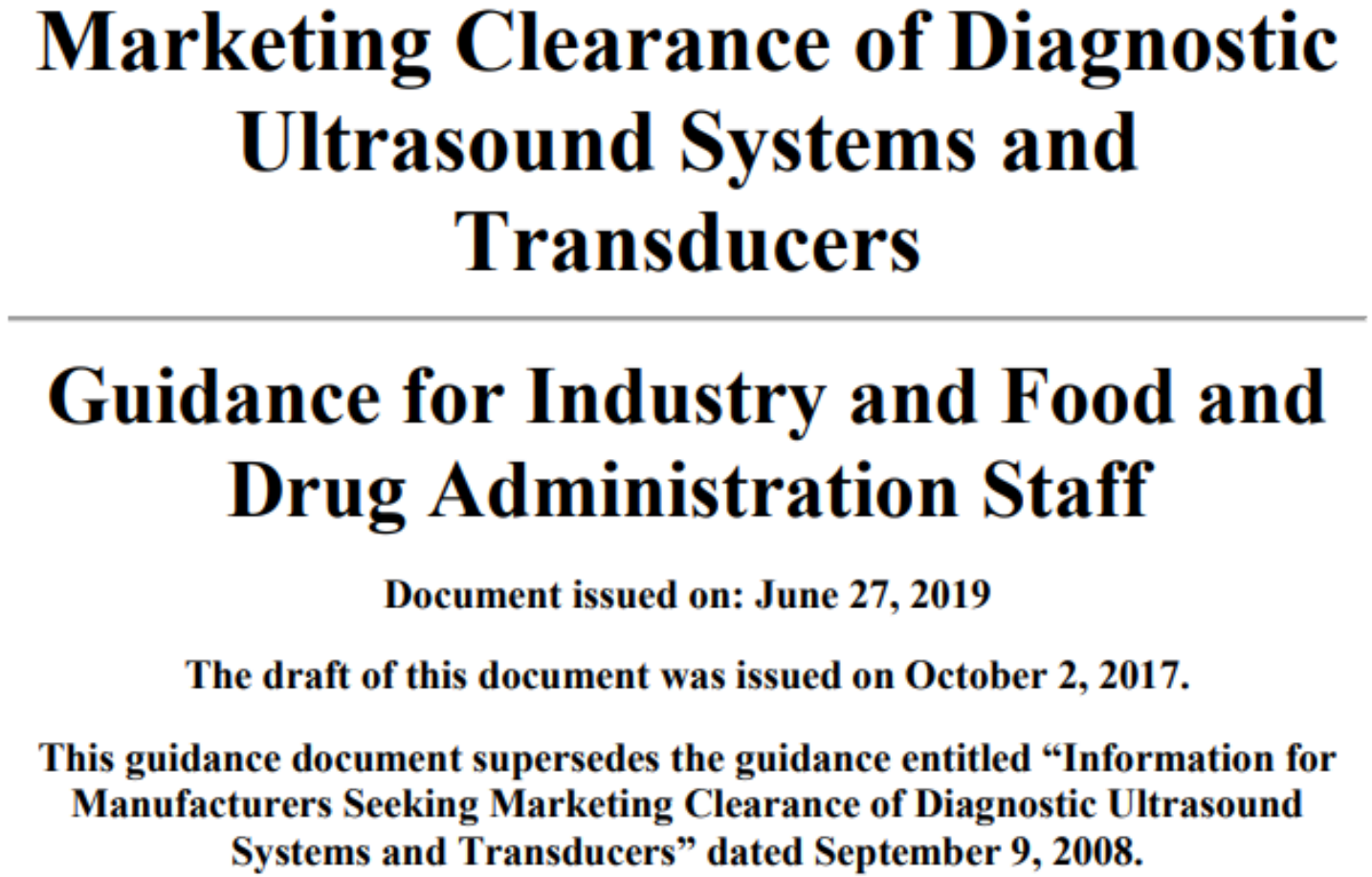
- Development and validation of reliable, automated artifact detection

- **Verification that detected artifact is due to actual equipment defect is still needed**
  - This approach an **adjunct** to annual testing using a phantom, not a replacement



June 2019: A gift from the government!

- An FDA ultrasound guidance document released in June 2019 included recommendation of a "transducer element check"
  - All scanners are already (likely) capable of automated self-checks of probe function, but this information is not shared with user
  - FirstCall systems provided this capability, but this seemed to be reverse-engineered, a probe set-up for testing was not easy
- Document contained many "should"s (and one "hope" in a webinar transcript), but so far no "shall"s or "must"s



Marketing Clearance of Diagnostic Ultrasound Systems and Transducers

Guidance for Industry and Food and Drug Administration Staff

Document issued on: June 27, 2019

The draft of this document was issued on October 2, 2017.

This guidance document supersedes the guidance entitled "Information for Manufacturers Seeking Marketing Clearance of Diagnostic Ultrasound Systems and Transducers" dated September 9, 2008.

https://www.fda.gov/media/71100/download

What is specified?

- Array element tests should be performed each time any transducer is connected or activated
- Test results should be made available to the system users
- Test results should specify array locations where poor performance is detected

Transducer Element Check

- Integrated tests of transducer performance each time a transducer is connected to the main system or activated
- The transducer performance test should be accessible by competent technical personnel, such as operators or service personnel
- While the FDA appreciates that different performance specifications may be necessary for transducers based on the application and system configuration, each device should include some level of testing. For example, an impedance check of each transducer element may provide a preliminary evaluation of the element integrity and function.
- Device manufacturers implement methods to communicate the results of the transducer performance tests to the operators, and identify regions of the image that have been compromised by transducer malfunction
- This integrated test feature would also generate a report on the performance of the probe under test for documentation, generally including a list of elements or smallest available patches of elements that have been compromised
- This integrated test should also be available to the operators to initiate any time when a particular probe is suspected of failure

What is missing?

- Details
  - Remote access to test results (DICOM SR?)
  - Alert if a potential problem is detected?
  - Specification that the report should include actual performance data, not just a simple Pass/ Fail msg
    - *Each clinical practice must be able to determine their own acceptable performance levels*
    - Uniformity images from clinical exams should be useful for characterizing clinical impact of defects … *These two methods seem quite complementary*

Transducer Element Check

FDA

- Integrated tests of transducer performance each time a transducer is connected to the main system or activated
- The transducer performance test should be accessible by competent technical personnel, such as operators or service personnel
- While the FDA appreciates that different performance specifications may be necessary for transducers based on the application and system configuration, each device should include some level of testing. For example, an impedance check of each transducer element may provide a preliminary evaluation of the element integrity and function.
- Device manufacturers implement methods to communicate the results of the transducer performance tests to the operators, and identify regions of the image that have been compromised by transducer malfunction
- This integrated test feature would also generate a report on the performance of the probe under test for documentation, generally including a list of elements or smallest available patches of elements that have been compromised
- This integrated test should also be available to the operators to initiate any time when a particular probe is suspected of failure

Conclusions

- **There is tremendous opportunity for medical physicists to contribute *in valuable ways* to an ultrasound practice quality assurance program**
  - A team-based scanner selection process can best set up a clinical practice for future success with a new scanner, whether a practice is buying 2 scanners or 20 scanners
    - Involving a diverse team (radiologists, sonographers, administrators, equipment service, medical physicists, medical physicist assistants) shares the workload
    - Participation by many staff will allow many in the practice to have some ownership in the final decision
    - The evaluations and decision are evidence-based and well-documented, which increases confidence in the final decision ,and facilitates the funding approval process
  - Developing methods to remotely provide required services can improve quality and lower cost, thereby increasing value or physics service for all practices, remote or nearby
    - Uniformity assessment from clinical images and scanner transducer element check data will both be extremely beneficial, and should be very complementary
    - *Influencing scanner vendors to facilitate remote system management and access to diagnostic test data will be critical to these efforts*

QUESTIONS & ANSWERS