AAPM Working Group on Quantitative B-mode Ultrasound Quality Control: Software for assessment of transducer artifacts

Sandra Larson, PhD
University of Michigan Medical Center, Ann Arbor, MI,

Mitch Goodsitt, PhD
Paul Carson, PhD
Zheng Feng Lu, PhD
Nicholas Hangiandreou, PhD
Sara Salari, MS

Ernest Madsen, PhD
James Zagzebski, PhD
Thaddeus Wilson, PhD
Wenzheng Feng, MS
Evan Boote, PhD

What is the most common problem found in ultrasound QC testing?

20% 1. Distance measurement errors.
20% 2. Problems with image uniformity.
20% 3. Inadequate spatial resolution.
20% 4. Reduced depth of penetration/visibility.
20% 5. Inadequate contrast resolution.

10

What is the most common problem found in ultrasound QC testing?

Correct Answer:

- **Problems with image uniformity.**

These problems are commonly due to defective transducer elements.

References:

- Stekel, S.F., Hangiandreou, N.J., and Tradup, D.J. (2009) “Four-Year Experience with an Ultrasound Quality Control Program”, *J. Ultrasound Med* 28: S148-149 (abstract).
- Lu, Z., Nickoloff, E., Dutta, A., and So, J. (2006) “Quality Control Testing of Diagnostic Ultrasound Systems: Experience in Testing 72 Systems in 11 years”, *Med Phys* 33: 2210 (abstract).

Uniformity is Subjective

- **The hardest question to answer: When is an artifact clinically relevant?**
- **Many of us just identify artifacts and let the user/physician in charge decide whether to take any action.**
- **Some respond by reviewing clinical images to determine if the problem is clinically relevant.**
- **Others will respond by calling service for any defect regardless of relevancy.**

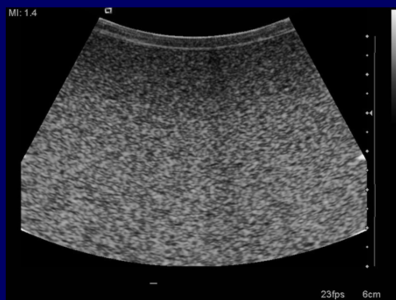
Transducer Example #1

Example 1: How would you evaluate this transducer?

20% 1. Uniform – no artifacts.
20% 2. Minor damage – no action needed.
20% 3. Significant damage – transducer usable, but repair may be needed.
20% 4. Major damage – transducer should not be used until repaired.
20% 5. More information is needed to evaluate appropriately.

10

Transducer Example #2



Example 2: How would you evaluate this transducer?

20%	1. Uniform – no artifacts.
20%	2. Minor damage – no action needed.
20%	3. Significant damage – transducer usable, but repair may be needed.
20%	4. Major damage – transducer should not be used until repaired.
20%	5. More information is needed to evaluate appropriately.

10

How would you evaluate this transducer?

Correct Answer (if there is one)**:**

- More information is needed to evaluate appropriately**.**

The final decision as to whether an artifact is relevant is a clinical one (most physicists are not qualified to determine clinical relevancy).

However, we may be called upon to give an opinion. It would be helpful to have an objective means of evaluating the transducer.

Reference: Thijssen, J.M., Weijers, G., De Korte, C.L., "Objective Performance Testing and Quality Assurance of Medical Ultrasound Equipment", *Ultrasound in Med & Biol* 33: 460-471.

Working Group on Quantitative B-Mode Ultrasound QC

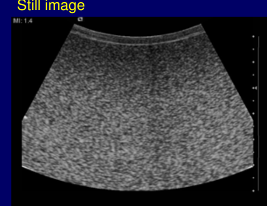
- **The goal of our current project is to produce an objective means to evaluate transducer artifacts.**
- **We have produced software intended to analyze transducer artifacts from stored cine loops of uniform images.**
- **Our software is written as a Plug-in for the open source software Image J (available at http://rsbweb.nih.gov/ij/).**

Viewing Defective Element Artifacts

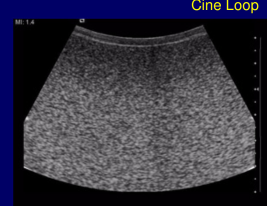
- **The best way to view these artifacts is with a uniform phantom.**
- **Decrease the image depth to get a good view of phantom-transducer interface.**
- **You may need to adjust the focal zone number or location to reduce user- and software-induced artifacts.**
- **Move the transducer to make sure you are not viewing an air bubble in the coupling gel.**

Viewing Defective Element Artifacts

Still image



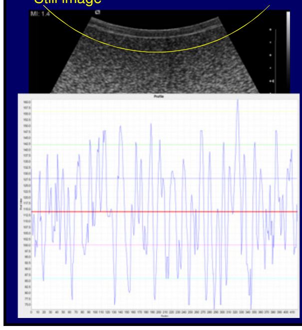
Cine Loop



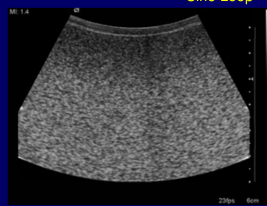
A profile across this image will not show this defect.
It will be lost in the noise.

Viewing Defective Element Artifacts

Still image



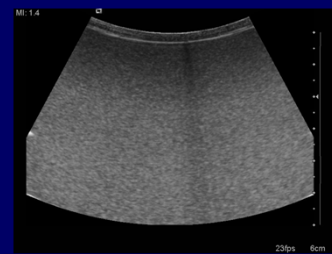
Cine Loop



A profile across this image will not show this defect. It will be lost in the noise.

Viewing Defective Element Artifacts

We can decrease the noise and better visualize artifacts by averaging this cine loop.



Quantitative Analysis

- Quantitative analysis from very noisy images is problematic.
- Therefore, we chose to use the median values across the stack of images to produce a low-noise median image.
- Our project seeks to use this median image to evaluate defective element artifacts.

Note: We chose to use the median (not an average) because ultrasound images are typically log-compressed, and with the median, we can perform all of our calculations without linearizing the image.

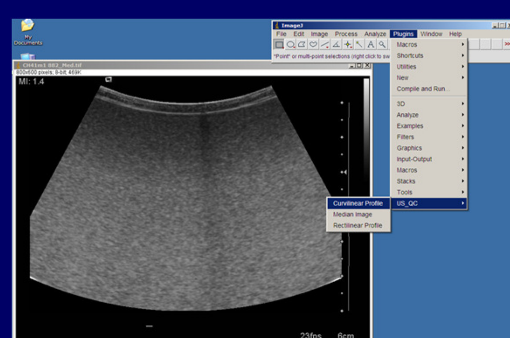
The Software

- Image J can read most image formats, but not compressed DICOM. Compressed DICOM images have to be decompressed before using them with our software.
- The program is designed to work with DICOM and tiff images (but may also work with other formats like jpeg).

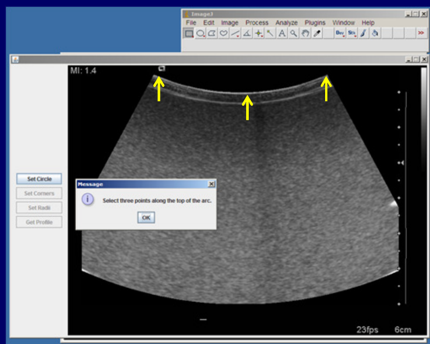
The Software

- One Plug-in converts the images to grayscale (if needed) and creates the median image.
- A second Plug-in allows the user to place a region-of-interest on the image.
- Within that region of interest, a median is computed for each column of data. Those median values are then plotted versus lateral position.

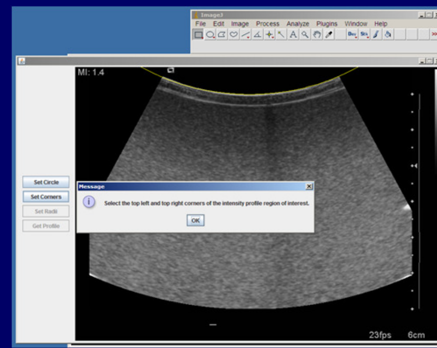
Create a profile from the median image



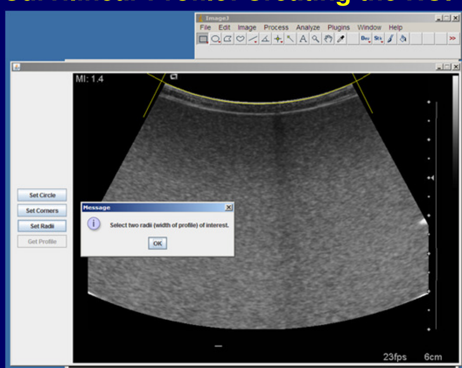
Curvilinear Profile: Creating the ROI



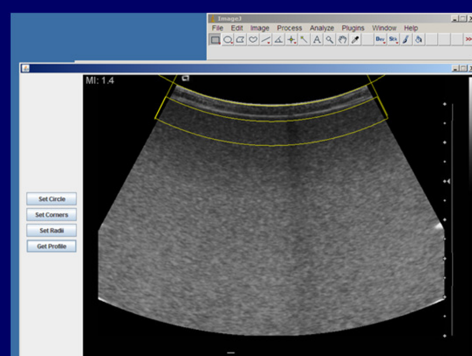
Curvilinear Profile: Creating the ROI



Curvilinear Profile: Creating the ROI



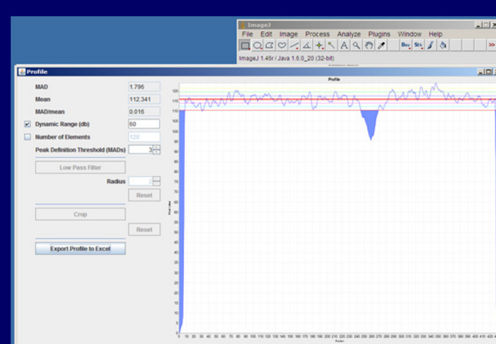
Curvilinear Profile: Creating the ROI



The ROI

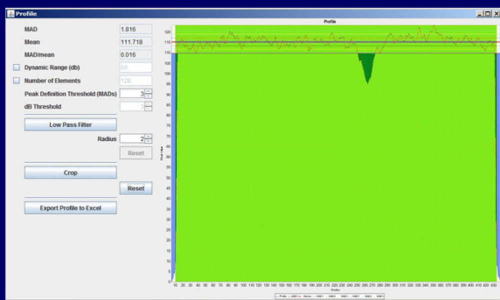
- Within the selected ROI, the program takes the median of all pixels along a ray normal to the transducer surface. A ray is analyzed for each pixel at the upper bound of the ROI.
- The program plots those median values against the lateral pixel location.

Curvilinear Profile Result

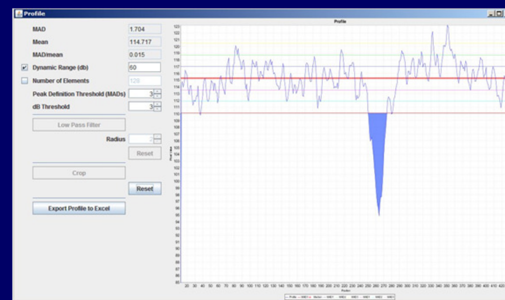


Cropping the ROI

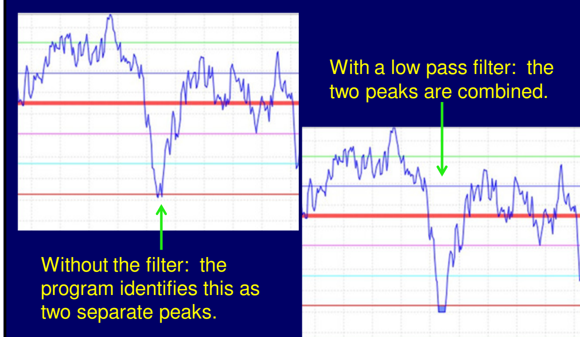
(don't want to include the dark edges in the profile)



Profile Results



When a Low Pass Filter is useful:



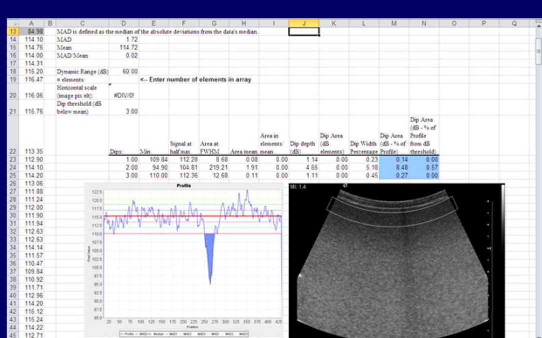
Profile Results

- **MAD is the Median Absolute Deviation.**

$$\mathrm{MAD} = median_i\left(\left|X_i - median_j\left(X_j\right)\right|\right)$$

- MAD is a measure of statistical dispersion which is more robust than the standard deviation.
- The profile is plotted with a default threshold of 3 MADs. Any data below this threshold is included as a potential artifact.

Profile Results Excel File



Profile Results Excel File

Dips:	Min	Signal at half max	Area at FWHM	Area/mean	Area in elements/mean	Dip depth (dB)	Dip Area (dB elements)	Dip Width Percentage	Dip Area (dB - % of Profile)	Dip Area (dB - % of Profile from dB threshold)
1.00	59.53	67.07	81.30	1.09	0.00	3.53	0.00	4.10	3.89	0.00
2.00	71.16	72.88	63.44	0.85	0.00	0.81	0.00	4.88	1.96	0.00

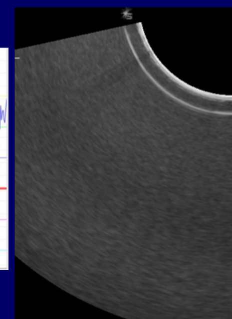
- Dips are defined as areas that extend below the median by 3 or more MADs.
- Mean: mean of all signal (intensity) values in the profile.
- Min: the minimum signal value of the dip.
- Signal at half "max": Mean – (Mean – Min)/2
- Area at FWHM: the area in pixel values integrated over the scan lines in the valley below the Mean, and between the bounds of the "half-max" of the dip.

Profile Results Excel File

Dips:	Min	Signal at half max	Area at FWHM	Area/mean	Area in elements/mean	Dip depth (dB)	Dip Area (dB elements)	Dip Width Percentage	Dip Area (dB - % of Profile)	Dip Area (dB - % of Profile from dB threshold)
1.00	59.53	67.07	81.30	1.09	0.00	3.53	0.00	4.10	3.89	0.00
2.00	71.16	72.88	63.44	0.85	0.00	0.81	0.00	4.88	1.96	0.00

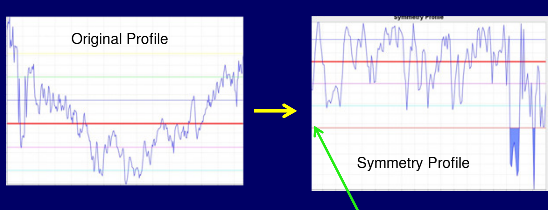
- Area in elements/mean:
  (Area/Mean)*(# transducer elements/256)
- Dip depth (dB): a measure of signal intensity loss
  dynamic range x (Mean – Min)/256
- Dip Width Percentage: with of the dip at the threshold divided by the width of the profile.
- Dip Area (dB-% of profile): the Area at FWHM expressed in dB and as a percent of total profile width.

What if the profile is curved?



Symmetry Profile

We can split the data in half and subtract the left side from the right side (folded like a book). This should remove the curvature and help to visualize any peaks.



The working group has compiled uniformity images from many transducers. We have performed a reader study to establish a baseline evaluation of the images from reviewers who are experienced in ultrasound quality control.

We will use this data to evaluate the profile data. Our goal is to find an objective way to mimic our reader study results.

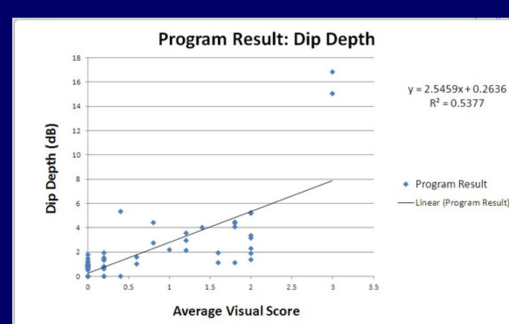
Our First Reader Study

5 readers (all experienced in ultrasound quality control testing) reviewed 61 cine image clips taken from 33 transducers.

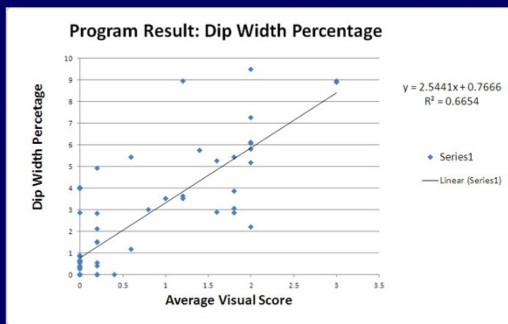
They rated the severity of each perceived non-uniformity on a 3-point scale (Minor, Significant, Major).

The next slides show the correlation between the average reviewer scores and the profile data from our program.

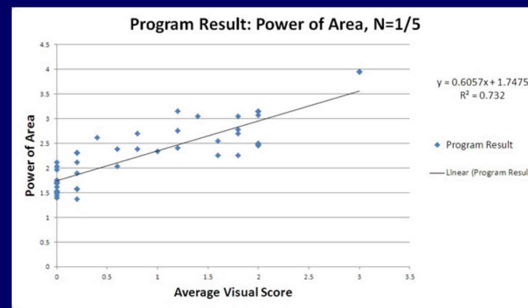
Study Result: Signal Loss in the Dip



Study Result: Dip Width



Study Result: Area of the Dip



Lessons Learned

Readers felt that our 3-point grading scale was insufficient to capture both the level of signal loss and the size of the non-uniformity.

We also learned that the readers needed better training (more detailed instructions regarding what qualifies as an artifact and experience in scoring images before the actual reader study).

Lessons Learned

We may be able to get better agreement with our program measurements if we do a second reader study with a refined visual grading scale and better training of the readers.

In order to do a second reader study, we need more images from previously untested transducers.

Seeking Volunteers

We would like to solicit your help to acquire sufficient images to complete a second reader study.

We would provide you a uniformity phantom (on loan) and the software. If you can provide image loops from at least 20 transducers (at least 10 of which should have some level of non-uniformity), please contact Sandra Larson at sclarson@umich.edu to volunteer.

Acknowledgement

We would like to thank our programmer, Christopher Hanes, for sticking with us through many changes and additions to the software.