AIUM Webinar: System-Dependent Factors Influencing Shear Wave Speed Measurements for Liver Fibrosis Characterization
Mark L. Palmeri, MD, PhD
Aired on 7/19/2016

Post-lecture Questions and Answers

1. Given the dependence of the SWS estimate on the frequency content, should manufacturers report some metric of the wave's frequency content with every SWS measurement?
That's a great question, and it's one that actually the committee has gone back and forth on multiple times. The consideration that was one of the more dominant things of how the end user being the clinician would interpret what is presented by a given system. So one of the challenges is that to fully appreciate why someone would be reported with the frequency content would require them to understand how to then interpret that in the context of a system that would report a different frequency. That was then the motivation for trying to tie all of the systems to a common frequency that they should all have - which is 140 Hz - that an MRE system could measure. And MRE was chosen as a good comparative modality because it's such a narrow band discrete frequency that it hones in and is not affected by that slope or gradient of speeds of adjacent frequency bands, and luckily the MRE community has been very collaborative in this regard, plus it will also help down the road clinically that an MRE system should hopefully agree with an ultrasound elasticity system, though right now as I mentioned the MRE systems measure closer to 6o Hz whereas the ultrasound systems are up at a higher 140 Hz.

2. Why a median of 10? Why not a mean of 10?
That's another great question. A median was chosen at first historically in the literature – there was a lot of traction in the literature from systems such as the fibroscan which actually made a recommendation of median. The main reason I think the community as a whole has stayed with that recommendation is that it is not uncommon to have some spurious measurements that clearly do not fall within kind of a normal distribution of variance, and by using the median metric, we're able to de-emphasize the influence of some of those outliers and really converge more on what we think is an accurate representation of the distribution of variability and not askew with that. That being said, the number of measurements being 10 and the use of median is not set in stone, and again, QIBA profiles are meant to evolve with time, and as the systems get more robust and tighter in agreement, there might be opportunity to actually drop the number of recommended measurements.

3. What's the reason to perform the samplings in the same area when we're studying a diffuse, but not uniform disease?
Another great question. So if someone were to add a confounder right now of biological variability of the heterogeneity of the disease on top of the current variability that we're trying to overcome, we know that we would expect some degree of change just with spatial distribution, and right now the emphasis has been put on trying to get the most representative metric at a given location rather than trying to get an overall reconstruction for the entire liver. This is a bit different than the MRE data as the pictures I showed alluded to. They actually get a wave field that encompasses most of the organ, and they probably get a more overall global measurement of stiffness than the ultrasonic systems which have a much tighter region of interest. I think that if you are looking for characterizing more of the heterogeneity of stiffness across the liver, probably the current recommendation would lean more towards taking 10 measurements at a given location and then doing that at the disparate locations spatially throughout the liver to build up that map so that you have higher confidence in reported speed at a given location.

4. How is the performance of these systems when imaging inclusions/heterogeneity?
Another great question, and exactly not what most of them are designed to do right off the bat. At least the point measurement systems, so the 2D shear wave speed imaging systems are actually specifically trying to look at delineating structures with contrast. It's more commonly been applied to in the breast than in the liver, but could be done in the liver. I'm going to defer that to say right now that's an active research question. We have not as a group through QIBA tried to hone in on what the accuracy and variability is as a function of pixel size or region of interest. There is a decent amount of literature that has started to come out that has specifically looked at how the confounders of shear wave reflections - its structural boundaries – influence things using different filtering schemes to specifically remove the spurious artifact components, but right now I'd say that's an active source of research more than something that's been standardized for clinical use.

5. What was the frequency they are talking about for us systems, ie 60 vs 140 hz? Was that the shear wave repetition frequency or us interrogation pulse?
What the different spectral values I quoted are is effectively the frequency content of the resulting shear wave. So in MRE data they actually externally vibrate with a continuous wave source that vibrates at 60 Hz and so the waves that are generated in the liver are temporally locked in at that 60 Hz rate. The ultrasound systems generate a more impulsive excitation to generate the shear waves and so they're considered to be more broad band, so they have more of a range of frequency contents and the 140 Hz would be the more energetic band in those resulting sheer waves. So it's really reflecting the morphology of the results in shear waves rather than any sort of repetition frequency or something like that.

6. **Is QIBA also considering releasing recommendations on ways to avoid biological confounders (fasting, hydration, ...)?**
That is an active part of the document. There is a lot of very recent literature where clinicians have been looking at that. As I alluded to in the clinical part of the profile, there are recommendations in fasting for at least 4 hours, depth of respiratory suppression for doing measurements, patient orientation. I think that there's a lot of opportunity to really delineate the other clinical confounders more, and active areas of research, and I think that the 1st version of the QIBA profile will have some of these included and future versions will be able to build on that a lot more.

7. In clinical settings, will patient body habitus eg. high BMI vs. low BMI affect the measurement of SWS in liver? Do the overlying soft tissues affect the measurement?
Another great question. Right now BMI is one of the factors that typically trends with the disease state so it's a little tricky sometimes to delineate this. We do know from the literature that shear wave speeds can be reconstructed in people with BMIs as high as in the 60s have been reported. No studies that I have seen have actually correlated BMI directly with that. We do know some systems do not work as well in ascites, so for overlying tissues, things that require the surface punch, like the fibroscan, have a bit more challenge going through ascites because shear waves don't couple through the fluids. Radiation force systems do not have as much of that challenge, but all of the systems will be challenged by potentially lots of different heterogeneous layers of soft tissue, specifically subcutaneous fat and the sort. Those things will degrade image quality, and that makes the tracking of the shear waves more complicated, and that's why one of the recommendations for optimizing success with shear wave measurements is actually to get the best B-mode image possible. That's the best qualitative sign of likelihood of getting good shear wave data you can reconstruct.

8. Could you comment on the QIBA choice of m/s - many clinicians use kPa, following Echosens choice.
As I've alluded to and as the person asking the question has, Echosens had a lead in the literature on this and has established some of the protocols that have been adopted. Meters per second was chosen I think by the QIBA group because that is what a lot of the underlying radiation force systems were actually directly reconstructing, and then Young's modulus was typically being reconstructed from that step. That doesn't necessarily mean that one or the other is "better" except for the fact that there is not a linear relationship between kPa and m/s so the quadratic dependence of the speed can actually expand the dynamic range of the measurements related to fibrosis relative to the stiffness, and that was seen by some as an advantageous metric.

9. Could physical rheological techniques be used to know the exact stiffness of phantoms instead of using MRE at a proposed Hz? Would that be more accurate?
Another great question. Something QIBA considered early-on. The challenge with a lot of the rheological systems that were investigated in early generation studies was actually limitations due to the fact that they worked in very low frequencies relative to the higher frequencies of radiation force systems and some discrepancies that were very amplified in those low frequency modes. So ideally – yes. We would be able to have kind of an on-the-bench standard. That's something that we're working towards.

10. How will be managed differences between measurements obtained with ultrasound imaging systems and FibroScan, especially when patients will be seen by both gastroenterologists-hepatologists (using FibroScan) and radiologists (using US imaging systems)?
If the ARF-based US systems do adopt conformance to a standardized phantom measurement at 140 Hz, then there is an opportunity to provide a calibration factor that may allow the end user to scale a FibroScan measurement made at closer to 60 Hz and reported as a Young's modulus to an equivalent shear wave speed at 140 Hz (similar to how MRE data was converted to equivalent shear wave speeds shown during the webinar). This conversion, however, has not been directly discussed by the QIBA committee and will be an active topic of future discussion.

11. What is the standard method for experimentally calibrate the elastic and viscoelastic phantom materials referred to ?
The elastic and viscoelastic phantoms used in the Phase I and II studies, respectively, were calibrated using a custom ARF-based shear wave sequence on a Verasonics system at Duke University. Moving forward, the MRE system operating at 140 Hz is being proposed as a new calibration standard for both types of phantoms.

12. Is it not questionable to consider biopsy the gold standard as the control tool to parametrise the efficiency of non invasive tools?
It definitely is questionable, and is a topic that was directly addressed in the SRU consensus document on this topic, and specifically commented on by pathologists: http://dx.doi.org/10.1148/radiol.2015150619

13. What about the confounding factors?

This is a very open-ended question. In addition to the system-dependent factors that were discussed explicitly during the webinar, along with the clinical procedure recommendations that were outlined (all recommended to help reduce confounding factors), there are many other potential confounding factors that are active areas of research. These include liver fibrosis disease etiology (e.g., infectious, alcoholic, auto-immune), comorbidities, hydration status, etc. The primary goal of the QIBA effort is to reduce system-based factors as much as possible, and then allow these "agreeing" systems to be used across the board for larger-scale clinical studies where these biological confounding factors can be studied more effectively.

14. Was there an obvious difference based on the method of shear wave generation ie comb vs single push location vs supersonic?

The commercial systems that were tested did not delineate the type of ARF excitation that was being used in the version of software being tested, so the Phase II study will not attempt to draw a conclusion about the ARF configuration and potential bias/variance in resultant measurements. As was presented, site variability was almost an equal contributor to system variability, and we are working to determine if we have enough data to delineate true system-based trends in our existing Phase II data.