Radiological Society of North America/Quantitative Imaging Biomarker Alliance Shear Wave Speed Bias Quantification in Elastic and Viscoelastic Phantoms

Mark L. Palmeri, MD, PhD, Andy Milkowski, MS, Richard Barr, MD, PhD, Paul Carson, PhD, Mathieu Couade, PhD, Jun Chen, PhD, Shigao Chen, PhD, Manish Dhyani, MD, Richard Ehman, PhD, Brian Garra, MD, Albert Gee, PhD, Gilles Guenette, Zaegyoo Hah, PhD, Ted Lynch, PhD, Michael Macdonald, PhD, Ravi Managuli, PhD, Veronique Miette, PhD, Kathryn R. Nightingale, PhD, Nancy Obuchowski, PhD, Ned C. Rouze, PhD, D. Cody Morris, Shana Fielding, Yufeng Deng, PhD, Derek Chan, Kingshuk Choudhury, PhD, Siyun Yang, PhD, Anthony E. Samir, MD, Vijay Shamdasani, PhD, Matthew Urban, PhD, Keith Wear, PhD, Hua Xie, PhD, Arinc Ozturk, MD, Bo Qiang, MS, Pengfei Song, PhD, Stephen McAleavey, PhD, Stephen Rosenzweig, PhD, Michael Wang, PhD, Yoko Okamura, MS, Glen McLaughlin, PhD, Yuling Chen, PhD, David Napolitano, PhD, Lindsey Carlson, PhD, Todd Erpelding, PhD, Timothy J. Hall, PhD

Received June 24, 2020, from the Duke University, Durham, North Carolina, USA (M.L.P., K.R.N., N.C.R., D.C.M., S.F., Y.D., D.C., K.C., S.Y.); Siemens Medical Solutions, Issaquah, Washington, USA (A.M., S.R.); The Surgical Hospital at Southwoods, Boardman, Ohio, USA (R.B.); University of Michigan, Ann Arbor, Michigan, USA (P.C.); SuperSonic Imagine, Aix-en-Provence, France (M.C.); Mayo Clinic, Rochester, Minnesota, USA (J.C., S.C., R.E., M.U., B.Q., P.S.); Massachusetts General Hospital, Boston, Massachusetts, USA (M.D., A.E.S., A.O.); US Food and Drug Administration, Silver Spring, Maryland, USA (B.G., K.W.); Zonare Medical Systems, Mountain View, California, USA (A.G., G.M., Y.C., D.N.); Toshiba Medical Research Institute, Redmond, Washington, USA (G.G.); Samsung Medison, Seoul, South Korea (Z.H.); CIRS, Inc, Norfolk, Virginia, USA (T.L.); GE Healthcare, Milwaukee, Wisconsin, USA (M.M., M.W.); Hitachi Healthcare, Twinsburg, Ohio, USA (R.M.); Echosens, Paris, France (V.M.); The Cleveland Clinic, Cleveland, Ohio, USA (N.O.); Philips Research, Cambridge, Massachusetts, USA (V.S., H.X.); University of Rochester, Rochester, New York, USA (S.M.); Canon Medical Systems Corp, Otawara, Japan (Y.O., T.E.); and University of Wisconsin, Madison, Wisconsin, USA (L.C., T.J.H.). Manuscript accepted for publication November 29, 2020.

We gratefully acknowledge the support of CIRS, Inc, for providing the phantoms used in this study. We are also grateful to the Radiological Society of North America (RSNA) for covering the costs of shipping the phantoms to the individual sites. Additional thanks to the RSNA Quantitative Imaging Biomarker Alliance (QIBA) staff for their help in coordinating group teleconferences, phantom shipping, and data exchange, and Julian Lee for assistance in data collection. Special thanks to Daniel Sullivan, MD, for fostering the QIBA vision and providing an infrastructure for these studies and collaborations between academics and industry. Funding for these studies was provided by the RSNA, National Institutes of Health (NIH) National Institute of Biomedical Imaging and Bioengineering (NIBIB) contracts HHSN268201500021C and HHSN268201000050C, NIH National Institute of Diabetes and Digestive and Kidney Diseases grants R01DK092255 and R01DK106957, and NIH NIBIB grants R01EB002132 and R01EB001981. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the US Food and Drug Administration.

Address correspondence to Mark L. Palmeri, MD, PhD, CIEMAS 1427, 101 Science Drive, Duke University, Durham, NC, 27705 USA.

E-mail: mark.palmeri@duke.edu

doi:10.1002/jum.15609

Objectives—To quantify the bias of shear wave speed (SWS) measurements between different commercial ultrasonic shear elasticity systems and a magnetic resonance elastography (MRE) system in elastic and viscoelastic phantoms.

Methods—Two elastic phantoms, representing healthy through fibrotic liver, were measured with 5 different ultrasound platforms, and 3 viscoelastic phantoms, representing healthy through fibrotic liver tissue, were measured with 12 different ultrasound platforms. Measurements were performed with different systems at different sites, at 3 focal depths, and with different appraisers. The SWS bias across the systems was quantified as a function of the system, site, focal depth, and appraiser. A single MRE research system was also used to characterize these phantoms using discrete frequencies from 60 to 500 Hz.

Results—The SWS from different systems had mean difference 95% confidence intervals of ±0.145 m/s (±9.6%) across both elastic phantoms and ± 0.340 m/s (±15.3%) across the viscoelastic phantoms. The focal depth and appraiser were less significant sources of SWS variability than the system and site. Magnetic resonance elastography best matched the ultrasonic SWS in the viscoelastic phantoms using a 140 Hz source but had a − 0.27 ± 0.027-m/s (−12.2% ± 1.2%) bias when using the clinically implemented 60-Hz vibration source.

Conclusions—Shear wave speed reconstruction across different manufacturer systems is more consistent in elastic than viscoelastic phantoms, with a mean difference bias of < ±10% in all cases. Magnetic resonance elastographic measurements in the elastic and viscoelastic phantoms best match the ultrasound systems with a 140-Hz excitation but have a significant negative bias operating at 60 Hz. This study establishes a foundation for meaningful comparison of SWS measurements made with different platforms.

Key Words—acoustic radiation force; elasticity; phantom; Quantitative Imaging Biomarker Alliance; shear wave; viscoelasticity

The Radiological Society of North America (RSNA) created the Quantitative Imaging Biomarker Alliance (QIBA) with imaging system manufacturers, academics, clinicians, and representatives from the United States federal government (eg, Food and Drug Administration, National Institutes of Health, and National Institute of Standards and Technology) to advance the concept of converting "imaging systems" to "measurement systems." Quantitative Imaging Biomarker Alliance profiles are developed for each measurement system that provide specific claims of what biomarker performance is possible when following the QIBA protocol, with the ultimate intent being to validate the profile across imaging systems with phantoms. The ultrasonic shear wave speed (SWS) biomarker committee was formed in 2012, with the purpose of developing a protocol and data analysis methods to allow direct comparison of SWS measurements made with different commercial systems, with the current clinical application being to estimate liver fibrosis. Several systems that measure SWS in the liver are commercially available, and many articles report that these measurements can differentiate fibrosis stages.[1,2] Shear wave elasticity imaging[3] methods implemented by several manufacturers, including both point SWS measurements and 2-dimensional shear wave elastography,[4] have been cleared by the Food and Drug Administration, and the technology has already reduced the number of liver biopsy procedures performed in Asia and Europe, as reflected in the National Institute for Health and Care Excellence guidelines for the management of viral hepatitis and the role of shear wave elasticity imaging in diagnosing and following disease progression in these patient populations.[5]

Literature suggests that SWS measurements depend on the measurement system.[1,2,6,7,8] These system differences cause clinical uncertainty and slow the adoption of this technology by the clinical community. Given the need for serial assessments of liver fibrosis and the impracticality of serial liver biopsy, providing a consistent SWS measurement that is system agnostic would improve the impact of this technology to noninvasively stage liver fibrosis.

A crucial step toward understanding sources of bias in SWS estimates is performing parametric studies in calibrated phantoms across all of the different manufacturer systems to study potential confounding factors, including the focal depth, material stiffness and viscosity, and appraiser. Phantoms may be elastic, which are relatively easy to fabricate, or viscoelastic, which are more difficult to fabricate but more closely mimic human liver. The SWS is independent of the shear wave frequency content in elastic media, but it depends on the frequency in viscoelastic media. Viscosity causes dispersion in the propagating shear waves, which means that the resultant SWS is dependent on the frequency content of the shear wave, with higher-frequency components of the shear wave propagating faster than lower-frequency components.[9] The frequency content of the generated shear wave can be affected by the spatial and temporal acoustic radiation force focal configurations used to generate the shear waves and the stiffness of the tissue and is also dependent on the how the shear wave displacements are estimated by using echoes from tracking beams.[10] Some commercial systems use tissue displacement data acquired from a single reference in the tissue before the acoustic radiation force is applied, whereas other systems estimate tissue velocity data using a progressive referencing sequence after the acoustic radiation force generates the shear wave.[4] Velocity data represent the first time derivative of the displacement data and therefore inherently have higher appreciable frequency content than the displacement data, making them a potential source of SWS differences between systems in viscoelastic media.[9] In these studies, we have calculated both the group SWS, which refers to the speed of a broadband pulse containing many frequencies, and the phase SWS, which refers to the speed of monochromatic waves as a function of frequency.

We first conducted an elastic phantom study (phase I) to evaluate first-order intersystem measurement differences in the absence of material viscosity.[11] We then conducted a viscoelastic phantom study (phase II) to evaluate how systems performed in materials with viscosity, which more realistically match the material properties of human liver tissue. For both phase I and II studies, comparative measurements were made with a research magnetic resonance elastography (MRE) system as a nonultrasonic modality that can also independently characterize stiffness and dispersion and is used clinically to characterize liver fibrosis.[12] Additionally, MRE allows for

multiple discrete excitation frequencies to be used to generate shear waves in the phantoms, which is not possible with the clinical ultrasound systems and allows for more direct characterization of the dispersive properties of these phantoms.

The phase I and II studies allowed us to quantify the bias of SWS measurements between different commercial ultrasonic shear wave elasticity imaging systems and an MRE system in elastic and viscoelastic phantoms. These analyses serve as a foundation for the claims and protocols in the first QIBA ultrasonic SWS profile.[13]

Materials and Methods

Phantom Calibration

Elastic Phantoms (Phase I)

Phase I studies were conducted from January 2012 to December 2013. Eleven pairs of elastic phantoms (E178*; Table 1) with nominal SWSs of 1.0 and 2.0 m/s, herein referred to as the "soft" and "stiff" elastic phantoms, respectively, were fabricated by CIRS, Inc (Norfolk, VA). These nominal speeds were chosen on the basis of the speeds associated with normal and fibrotic livers in the literature, where accurate resolution of the speed is important for clinical diagnosis.[2] The phantoms were homogeneous cylinders that were 100 mm in diameter and height, except for a pair of phantoms designed for MRE measurements (E1788) that were 200 mm in diameter and 120 mm in height to reduce standing wave reflections off the phantom walls.

The SWSs in all of the phantoms were measured at Duke University using a Vantage research scanner (Verasonics, Inc, Kirkland, WA) sequence (Table 2), following the procedure outlined in Appendix 1.[14] A grand mean was calculated across all of the phantom measurements and used as a normalization factor to compensate for the SWS bias due to fabrication variability among the phantoms. Ten replicate measurements were made across 10 different speckle realizations (transducer positions) with 3 different focal depths (40, 60, and 80 mm) in each phantom, where each speckle realization was obtained by rotating the phantom about a common location using a rotation platform. Group and phase SWS measurements were made by the methods described previously[9] and are available for download (https://github.com/RSNA-QIBA-US-SWS/VerasonicsPhantomSequences).

Viscoelastic Phantoms (Phase II)

Phase II studies were conducted from January 2014 to March 2016. Three viscoelastic (phase II) phantoms (E2297; Table 1) were characterized at Duke University using a Verasonics research scanner, following the procedure outlined in Appendix 1.[14] In contrast to the phase I studies: (1) 16 replicate measurements, instead of 10, were performed in each phantom; and (2) 3 different stiffness phantoms, instead of 2, were measured with a given system at each imaging site.

Viscoelastic phantoms (phase II) can be susceptible to more fabrication replication variability, so for the phase II study, a single set of phantoms were shipped to each of the different measurement sites. The stiffnesses and viscosities chosen for the phase II phantoms represent different degrees of normal through fibrotic livers (supporting data presented in "Results").

To characterize how the phantom dispersion represents that of the human liver, we compared the group speeds derived from displacement and velocity data in these phantoms to group speeds derived from displacement and velocity data in healthy and diseased human livers. All human data were acquired in an Institutional Review Board–approved study that

Table 1. Phantoms Fabricated by CIRS, Inc, and Measured as Part of These Phase I and II Studies, Including Their Designated Use in These Studies

Phantom Label	Elastic/Viscoelastic	Phase	Purpose
E1786-[1–10]	Elastic	I	Intersystem comparison
E1787-[1–10]	Elastic	I	Intersystem comparison
E1788-[1,2]	Elastic	I	Ultrasound-MRE comparison
E2297-[A1, B3, C1]	Viscoelastic	II	Intersystem and ultrasound-MRE comparison

has already been published.[15,16] Although the data acquisitions in the human data were done with a system and sequence not used to image the phase II phantoms, we were not interested in the absolute agreement of speeds between the different systems. We instead evaluated the ratio of the group speeds estimated with each type of data, where a nonunity ratio indicates dispersion, which should be relatively independent of bias between the different systems.

Site Measurement Protocol

The phantoms were distributed among 12 sites for measurements on commercial clinical SWS-capable systems, including FibroScan (Echosens, Paris, France), EPIQ 5 (Philips Healthcare, Amsterdam, the Netherlands), Acuson S2000/S3000 (Siemens Healthineers, Munich, Germany), Aixplorer (SuperSonic Imagine, Aix-en-Provence, France), HiVision Ascendus (Hitachi Healthcare, Tokyo, Japan), LOGIQ E9 (GE Healthcare, Chicago, IL), RS80 (Samsung Healthcare, Seoul, Korea), Aplio 500 (Canon Medical Systems Corp [formerly Toshiba], Otawara, Japan), and ZS3 (Mindray [formerly Zonare], Shenzhen, China) as well as Verasonics research systems at Duke University and the Mayo Clinic. It should be noted that in the phase I study (2012–2013), only 5 of the systems were available at the time for phantom measurements, whereas all of the systems were available for the phase II phantom measurements (2014–2016). The systems and sites in our analysis were assigned arbitrary letter designations (A–K) to maintain their anonymity throughout the study, and there was no correlation between the letter designations between the phase I and II studies (ie, system A in phase I is not necessarily system A in phase II).

For the phase I study, each site had at least 3 appraisers scan each phantom 10 times at each focal depth [30, 45, and 70 mm, which differed from the phase I calibration measurements described in "Elastic Phantoms (Phase I)"] with a handheld transducer, with each combination of appraiser and focal depth repeated for 3 trials in random order relative to the other appraisers. A single appraiser at each site was used in the phase II study, and 16 replicate measures were made in these phantoms.

The order of data acquisition was randomized for phantoms, appraisers, depths, and imaging systems (if >1 was used) to allow for an accurate statistical investigation of results. Participants were all blinded to the intermediate results of others measurement sites. All of these data were then analyzed to estimate the bias in SWS estimates across different systems, measurement sites, focal depths, and appraisers. If a system did not report the SWS (C_T) directly but instead reported the Young modulus (E) or shear modulus (μ), those moduli were converted to the SWS assuming an isotropic, incompressible, elastic material assumption:

$$c_T = \sqrt{\frac{\mu}{\rho}} = \sqrt{\frac{E}{3\rho}}, \tag{1}$$

where ρ represents the density of the phantom material (as quoted on the phantom label or assumed to be 1000 kg/m^3). Since curvilinear arrays were used to image phantoms with flat surfaces, a coupling solution was used to match the sound speed of the phantom material to minimize index of a refraction mismatch that could bias SWS estimates.[17,18]

A statistical analysis of variance was performed to evaluate which variables in our study (phantom, system, site, appraiser, and focal depth) led to significant differences ($P < .01$) between reported SWSs. A Tukey mean difference analysis was also performed to evaluate trends in the bias among systems and sites. Linear regression was used to evaluate for the bias as a function of focal depth for each system. All

Table 2. Acoustic Radiation Force Excitation and Displacement Tracking Parameters Used on a Verasonics Research Scanner With a Philips C5–2 Curvilinear Array to Measure all of the Phase I Elastic Phantoms Before Distribution to Individual Measurement Sites

Excitation Parameters		Tracking Parameters	
Frequency, MHz	**2.4**	**Frequency, MHz**	**3.2**
Focal depths, mm	40, 60, 80	Transmit F-number	Plane wave
F-number	F/2.0	Receive F-number	F/2.0
Duration, μs	400 (960 cycles)	Pulse repetition frequency, kHz	5

statistical analysis was performed with the Statsmodels and SciPy packages in Python version 3.8.[19,20]

Magnetic Resonance Elastography

Magnetic resonance elastography on the phase I and II phantoms was performed at a single research site. To generate shear wave propagation in the phantoms, a square MRE electromechanical shear driver (64 × 64 × 3.0 mm) was placed on top of the phantom with light compression to maintain mechanical coupling. The driver frequency ranged from 60 to 500 Hz, with MRE performed at each discrete frequency. To better match the bandwidth of ultrasonic SWS systems, the shear wave frequencies used in the MRE measurements were expanded to included higher values than those used in clinical MRE of the liver (typically 60 Hz).[12]

Shear wave propagation images were acquired by using a 3-dimensional MRE wave imaging sequence on a single-channel coil and a 1.5-T Signa scanner (GE Healthcare, Waukesha, WI). The following major parameters were used in the study: field of view, 216 mm; matrix, 128 × 128; repetition time, 1600–2314 milliseconds; echo time, 62.7–119 milliseconds; slice thickness/spacing, 3.5/0 mm; 16 slices; motion sensitivity, 4.5–25.2 μm/π radians; motion sensitivity direction, *x*/*y*/*z*; and axial imaging plane.

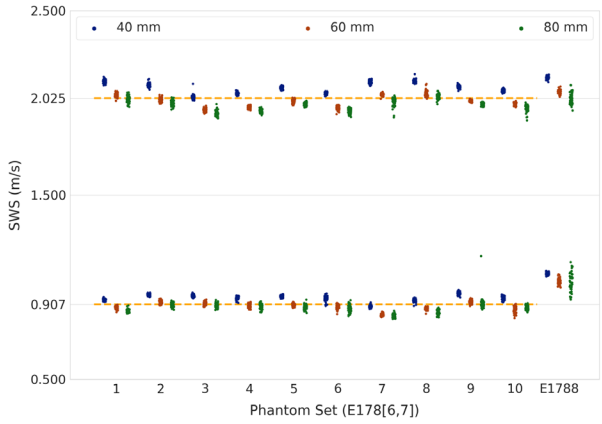
A 3-dimensional MRE direct inversion algorithm was used to process wave images and compute elastograms.[21] The model-free direct inversion algorithm provides calculated images depicting the magnitude, real part, and imaginary part of the complex shear modulus. Region-of-interest measurements were obtained from each of the images over a large area of each phantom.

The complex shear modulus $[G^*(\omega)]$ was calculated as $G^*(\omega) = G_r(\omega) + i\, G_i(\omega)$, where G_r and G_i are the real and imaginary parts of the complex shear modulus as a function of the angular frequency (ω). Using this complex shear modulus, the phase velocity (v_s) can be expressed as

$$v_s(\omega) = \sqrt{\frac{2}{\rho}\frac{|G^*(\omega)|^2}{G_r(\omega) + |G^*(\omega)|}}. \quad (2)$$

Before the MRE examinations, the phantoms were allowed to equilibrate to 20°C for at least 8 hours before measurements were made.

Figure 1. Calibration measurements on all of the softer (E1786) and stiffer (E1787) elastic ultrasound phantoms and the phantom set designated for comparison to MRE measurements (E1788) using a research scanner sequence at 3 different focal depths (40 [blue], 60 [red], and 80 [green] mm). The dashed orange line in each plot represents the grand mean of all measurements made in the ultrasound phantoms for each plot: 0.907 ± 0.033 (3.7%) m/s and 2.025 ± 0.051 (2.5%) m/s for the soft and stiff phantoms, respectively. A given phantom set's mean difference from these grand means was used as a corrective factor to normalize for this fabrication variability between different phantom pairs.



Results

Elastic Phantoms (Phase I)

Figure 1 shows the calibration measurements made on all of the elastic phantoms. Figures 2 and 3 show the aggregated SWS measurements grouped by unique site and system, respectively. There were statistically significant differences in SWS measured between soft and stiff phantoms ($P < .01$), between different systems ($P < .01$), at different sites ($P < .01$), and as a function of the focal depth ($P < .01$).

Figure 4 shows a Tukey mean difference plot for aggregate systems and sites, using the normalization data (Figure 1) as the reference measurement for each phantom. These data had a mean difference 95% confidence interval (CI) of ±0.145 m/s (±9.6%) between the soft and stiff phantoms. Table 3 shows the focal depth bias for each system in each elastic phantom.

At each site, there was not a significant difference in SWS acquisitions between different replicate acquisition procedures ($P > .05$). Differences between appraisers were significant ($P < .01$), but the variance associated with appraiser differences (0.00176 m/s) was less than 7% compared to the variance associated with system, site, and focal depth (0.0266 m/s).

Figure 5 shows the group and phase SWS estimates made in a pair of the phase I elastic phantoms, along with MRE estimates of the phase SWS in a customized set of elastic MRE phantoms. As summarized in Table 4, the soft phantom (E1786-1) had a statistically significant ($P < .001$; 0.85/0.77) 10.4% increase in the group SWS using velocity instead of displacement data and a 0.30-(m/s)/kHz linear increase ($R^2 = 0.87$) in the phase velocity. The stiff phantom did not have a statistically significant difference in the group SWS ($P > .05$), with only a 0.03-(m/s)/kHz linear increase ($R^2 = 0.43$) in the phase velocity.

Figure 2. Aggregate SWS data in the soft (blue) and stiff (green) elastic phantoms measured at different sites, where some sites had multiple systems available for measurement. Each system at each site was used by 3 appraisers who made 10 replicate measurements at each of the focal depths (30, 45, and 70 mm) in each phantom. In some cases (sites D, E, F, J, K, L, and M), coarser quantization (rounding to the nearest 0.1 m/s) of the reported SWS by some or all of the site systems is apparent.

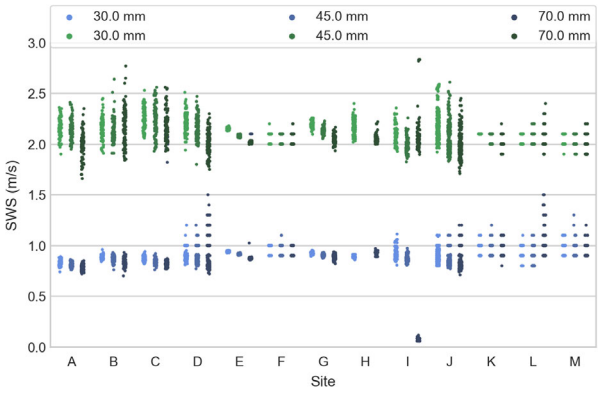


Figure 3. All of the elastic phantom data grouped by unique system. Some systems were used at only a single measurement site, whereas other systems were used at multiple measurement sites. Note that a single system (B) appears to report SWS with coarser quantization (0.1 m/s) compared to the other systems.

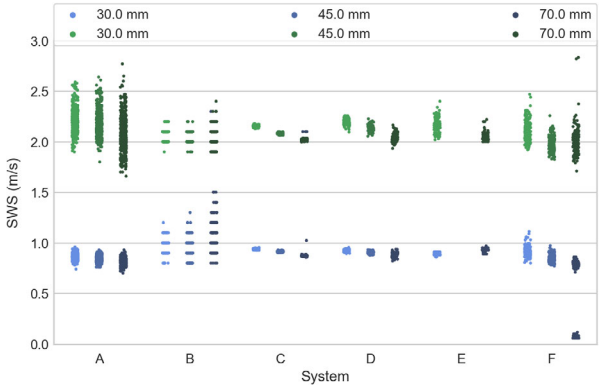


Figure 4. Tukey mean difference plots for the aggregated phase I systems (top) and sites (bottom), using the normalization data (Figure 1) as the reference measurement for each phantom. The colors in each plot represent the same system/site, respectively. Note that system/site biases are not necessarily consistent across the soft and stiff phantoms (eg, a system with a negative bias in the soft phantom may have a positive bias in the stiff phantom).

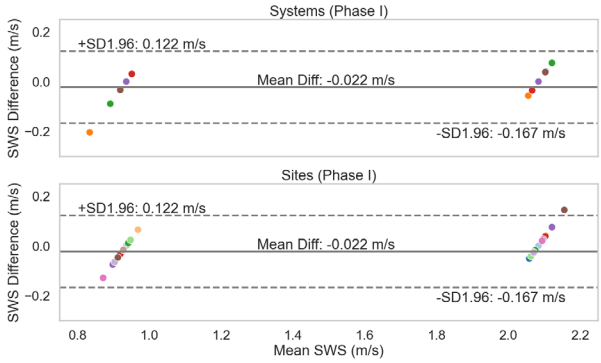


Table 3. Focal Depth Bias as a Function of Different Systems in the Phase I Phantoms, Calculated by Simple Linear Regression

System	Phantom	Focal Depth Slope, (m/s)/mm	R^2
A	Soft	−0.00063	0.01
	Stiff	−0.0035	0.14
B	Soft	0.000032	3.0e−06
	Stiff	−0.00091	0.06
C	Soft	0.00046	0.002
	Stiff	−0.0033[a]	0.95[a]
D	Soft	−0.00091	0.46
	Stiff	−0.0037[a]	0.75[a]
E	Soft	0.0011[a]	0.76[a]
	Stiff	−0.0028	0.40
F	Soft	−0.012	0.47
	Stiff	−0.0021	0.08

[a]Non-negligible bias with moderate-to-good linear regression coefficients.

Viscoelastic Phantoms (Phase II)

Figure 6 shows the comparison of different systems measuring the group SWS in the phase II phantoms. There was a statistically significant difference in the SWS measured between each of the 3 phase II phantoms ($P < .01$), along with a statistically significant difference between SWS measurements as a function of the system ($P < .01$), site ($P < .01$), and focal depth ($P < .01$).

Figure 7 shows a Tukey mean difference plot for different systems using the normalization data from the Verasonics calibrations as the reference measurements for each phantom. This figure shows a mean difference 95% CI of ±0.340 m/s (±15.3%) across all 3 phantoms.

Table 4. Comparison of the Dispersion Estimated in the Phase I and II Phantoms Using the Verasonics Ultrasound System and MRE

Phantom	Ultrasound, (m/s)]/kHz	MRE, (m/s)/kHz
E1786-1	0.30 (R^2 = 0.87)	NA
E1787-1	0.03 (R^2 = 0.43)	NA
E1788-1	NA	0.60 (R^2 = 0.970)
E1788-2	NA	0.20 (R^2 = 0.970)
E2297-A1	0.61 (R^2 = 0.92)	3.0 (R^2 = 0.99)
E2297-B3	0.78 (R^2 = 0.96)	3.2 (R^2 = 0.98)
E2297-C1	0.78 (R^2 = 0.92)	3.8 (R^2 = 0.99)

Linear regression of the phase velocity data was performed by using the frequency ranges shown for each phantom in Figures 5 and 8. NA indicates not applicable.

Figure 8 shows the displacement- and velocity-based group SWS reconstructions in the 3 viscoelastic phantoms, along with their corresponding phase velocity curves. As summarized in Table 4, the E2291-A1 phantom had a (2.17/1.61) 35% increase in the velocity group SWS compared to the displacement group SWS, with a 0.61-(m/s)/kHz linear increase in the phase velocity (R^2 = 0.92) with the frequency; the E2297-B3 phantom had a (2.77/2.12) 31% increase in the velocity/displacement group SWS with a 0.78-(m/s)/kHz linear increase in the phase velocity (R^2 = 0.96) with the frequency; the E2297-C1 phantom had a (3.33/2.55) 31% increase in the velocity/displacement group SWS with a 0.78-(m/s)/kHz linear increase in the phase velocity (R^2 = 0.92) with the frequency. Figure 9 shows the group SWS calculated displacement and velocity shear wave data in both the phase I (elastic) and phase II (viscoelastic) phantoms compared to previously reported in vivo human data.[15,16] The human data had an increase in the velocity/displacement group SWS of (2.26/1.78) 27% ± 5.6% across all fibrosis stages. Table 5 shows the focal depth bias for each system in each viscoelastic phantom.

Matched MRE measurements made at discrete excitation frequencies ranging from 60 to 200 Hz are shown in Figure 10. The MRE data represent the mean measurements (with negligible error bars) superimposed on the aggregate ultrasonic SWS data for all systems and sites at a focal depth of 45 mm. The corresponding dispersion slopes are summarized in Table 4. The-140 Hz MRE excitation frequency

Figure 5. Group and phase SWS measurements in one pair of the phase I elastic phantoms made by using the Verasonics research scanner sequences and processing code at a focal depth of 45 mm, derived from shear wave displacement (Disp) and from shear wave velocity (Vel). The circles in each plot represent the mean of 10 independent acquisitions, and the error bars represent the 95% CI for each measurement. Magnetic resonance elastographic measurements were made at discrete frequencies of 140, 180, 200, 300, 400, and 500 Hz. The slopes of linear fits to these phase velocities, which are indicative of undesired dispersion (frequency-dependent phase velocity) in these elastic phantoms, are summarized in Table 4.

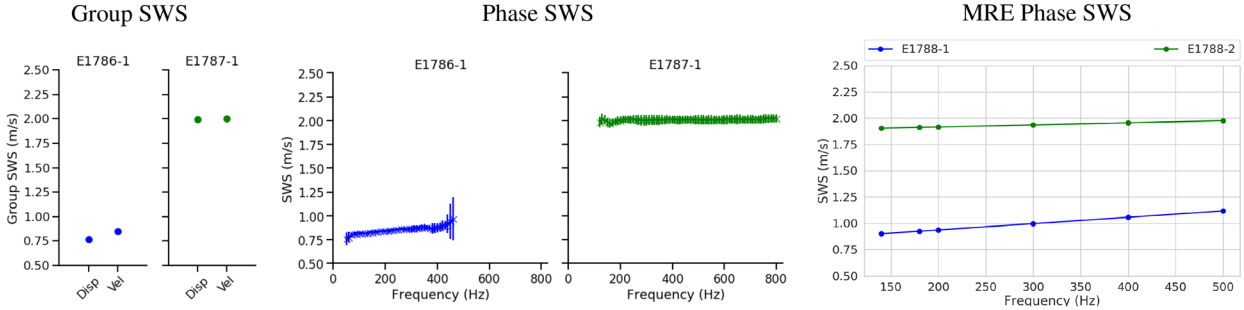
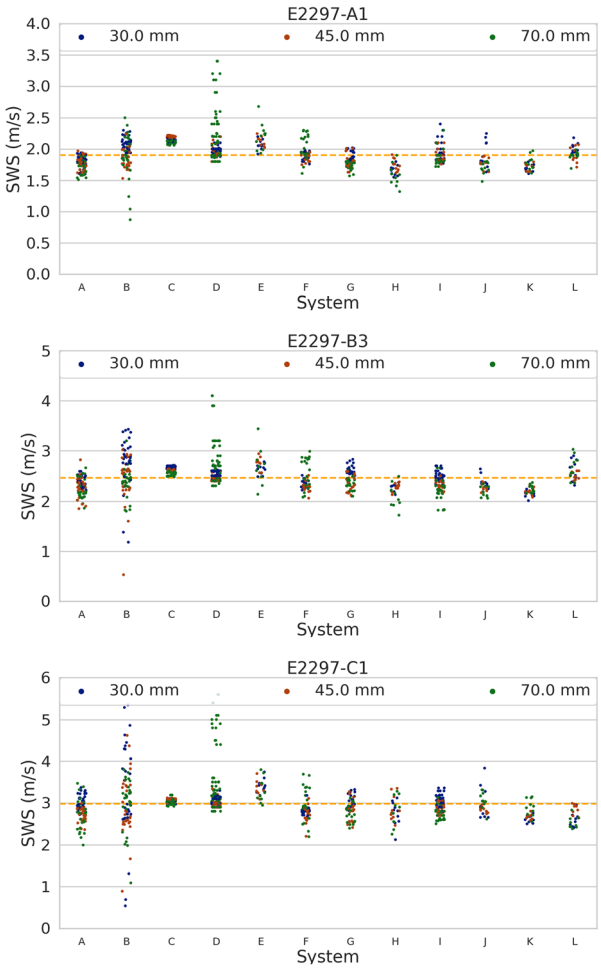


Figure 6. Phase II phantoms measured with different systems with 3 different focal depth configurations (30, 45, and 70 mm). The orange line on each plot represents the grand mean value across all systems for each phantom.

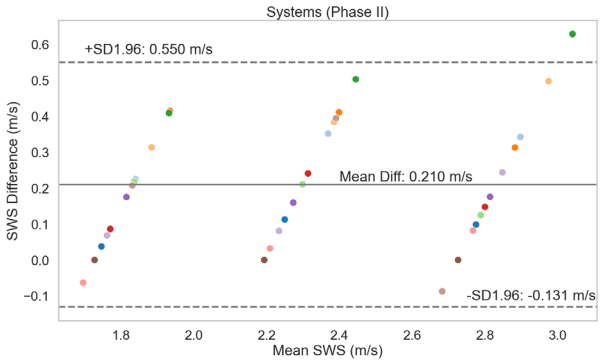


matched the mean of the ultrasonic SWS measurements, but the clinically implemented 60-Hz excitation had a -0.27 ± 0.027-m/s $(-12.2\% \pm 1.2\%)$ bias. Figure 11 shows the SWS measured in the phase II phantoms at 4 time points ranging from August 2014 to September 2015 to evaluate their temporal stability.

Discussion

The phase I elastic study revealed several interesting findings. We found that the mean difference between

Figure 7. Tukey mean difference plots for the aggregated phase II systems using the data from the calibration Verasonics system as the reference measurements for each phantom. The colors in each plot represent the same system. Note that system biases are not necessarily consistent across the different phantoms (eg, a system with a negative bias in one phantom may have a positive bias another phantom).



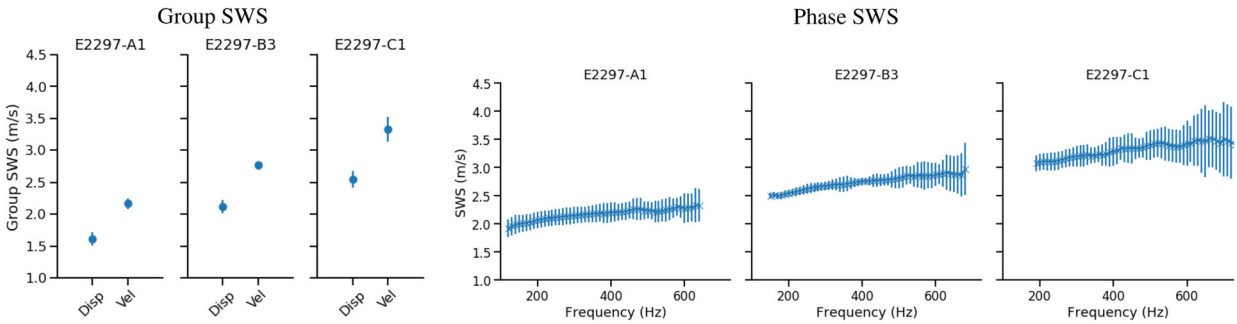
systems had a 95% CI of ±9.6%. It can be noted that one system (B) reported values with coarser quantization (0.1 m/s) compared to the other systems. In addition to system variability, site variability was also appreciable, even when the same system was being used at difference sites. Although there were biases associated with each system and site, those biases were not necessarily consistent in both the soft and stiff phantoms (eg, a system that had a negative bias in the soft phantom may have had a positive bias in the stiff phantom).

The focal depth bias was, in general, a less significant confounding factor than system and site variability. There was one outlier case (system C in the stiff elastic phantom) that did have an appreciable −0.132-m/s bias across the 30- to 70-mm focal depth range ($R^2 = 0.95$), although that system did not have such a bias in the soft elastic phantom or the viscoelastic phantoms.

Although the phase I studies did show a statistically significant difference in SWSs measured between different appraisers at a given measurement site, the appraiser was much less of a confounding factor compared to the system and site differences. That being said, these studies were conducted in phantoms and did not capture the challenges of imaging livers in patients, in whom differences in appraisers could be significant.

System and site differences were also present in the phase II viscoelastic phantoms, and the mean

Figure 8. Group and phase velocities calculated in the 3 phase II viscoelastic phantoms that were distributed to all of the measurement sites. The error bars represent the 95% CI over 16 independent measurements. As expected, these viscoelastic phantoms have higher group SWS estimated when by using velocity (Vel) data instead of displacement (Disp) data (left plots). This same trend is seen in the positive slope of the corresponding phase velocity curves (right plots). In the phase velocity plots, note that the frequency range of the reconstructed phase velocities increases as a function of increasing stiffness, and the variance of the estimated phase velocity increases at higher frequencies because of the lower signal-to-noise ratio at these higher frequencies. The slopes of the linear-fit phase velocity lines are summarized in Table 4.



difference 95% CI (±7.8%) was greater than that in the elastic phantoms. These viscoelastic phantoms matched the distribution of group SWS displacement/velocity ratios that we observe in human data, indicating similar amounts of dispersive characteristics in these phantoms. Both of the elastic phantoms showed minimal dispersion with the use of these group SWS ratios. It should be noted, however, that MRE yielded significantly greater linear dispersion slopes than the ultrasound system phase velocity analysis for these phase II phantoms. The source of this discrepancy has not been resolved and will be a focus of future studies.

The best agreement between MRE and the aggregated ultrasonic SWS measurements in the viscoelastic phantoms occurred at an excitation frequency of 140 Hz, but the lower 60-Hz excitation used in clinical MRE could lead to lower MRE values for liver stiffness in the literature than ultrasound systems.[12,22]

The ultrasound system SWS distribution for the softest viscoelastic phantom (E2297-A1) in Figure 10 shows a bimodal distribution. Such a distribution may be indicative that some systems are reconstructing group SWSs using displacement data (leading to the lower distribution), whereas others may be using velocity data (leading to the higher distribution). Such separation of these populations could be lost in the stiffer phantoms as the variability of the reconstruction using either displacement or velocity data increases. It should be noted that this bimodal distribution explanation is simply a hypothesis, as each manufacturer did not reveal how it calculated its group SWS metrics. If the data type (displacement/velocity) is a source of this variability, then a manufacturer consensus on what data to use in calculating the group SWS or implementation of a bias reduction factor could help provide better consistency of reported SWSs between systems.

Because proprietary processing algorithms and scanner sequencing could not be disclosed by manufacturers in this study, we cannot conclude what the sources of the intersystem bias were in these studies. To allow researchers in academics, industry, and clinical practice to have a common platform to perform ultrasonic SWS measurements, we created standardized shear wave acquisition sequences on a Verasonics research scanner[14] that can be used to test tissue-mimicking phantoms, along with postprocessing code to estimate the group SWS using the displacement and velocity as the raw input data into the reconstruction algorithms,[9] as presented in this study. In addition to estimating the group speeds, the reconstruction code also estimates the phase velocity over the more energetic bandwidth of the shear wave signal. These sequences and postprocessing software are openly available on GitHub (https://github.com/RSNA-QIBA-US-SWS/VerasonicsPhantomSequences) and will be incorporated in the first generation of the QIBA profile for

the ultrasonic SWS (http://qibawiki.rsna.org/index.php/Profiles). In addition to these phantom studies and associated experimental sequences and post-processing code, elastic and viscoelastic digital phantoms based on finite-element methods have been developed and released to the community to use for algorithm development and validation.[23]

The work presented in this article represents the culmination of several years of effort with evolving methods between the two phases of the study and, in turn, had some limitations. The use of a grand mean normalization across all of the phantom pairs fabricated for the phase I study allowed all of the phantoms to be compared to a nominal reference value but complicated studies that involved relative measurements made on any singleton pair of phantoms. Circulating the same sets of phantoms, as was done in phase II, placed the comparative burden on the longitudinal stability of these phantoms, which did appear relatively consistent across the duration of the study.

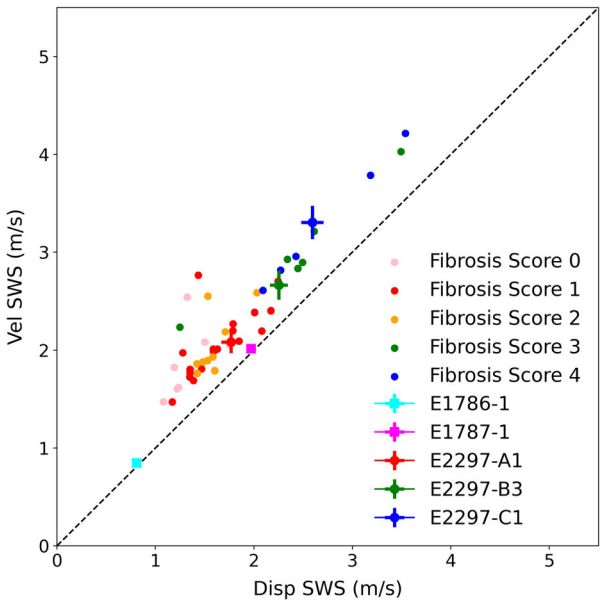
The collection of data over several years may have led to different software versions being installed on systems that were deemed the same in our analysis. All of the system software and models used in this study may have been older than the latest-generation system SWS elasticity tools and algorithms. The recording of the specific scanner software version is considered to be just as important as the recording of

Table 5. Focal Depth Bias as a Function of Different Systems in the Phase II Phantoms, Calculated by Simple Linear Regression

System	Phantom	Focal Depth Slope, (m/s)/mm	R^2
A	E2297-A1	−0.0019	0.08
	E2297-B3	−0.0042	0.14
	E2297-C1	−0.0070	0.16
B	E2297-A1	−0.0047	0.095
	E2297-B3	−0.0096	0.11
	E2297-C1	−0.0073	0.019
C	E2297-A1	−0.00095	0.13
	E2297-B3	−0.0031[a]	0.76[a]
	E2297-C1	−0.00054	0.029
D	E2297-A1	0.0064	0.14
	E2297-B3	0.0055	0.10
	E2297-C1	0.016	0.21
E	E2297-A1	0.0034	0.14
	E2297-B3	0.00084	0.0039
	E2297-C1	−0.0025	0.038
F	E2297-A1	0.0046	0.25
	E2297-B3	0.0065	0.24
	E2297-C1	0.0051	0.085
G	E2297-A1	−0.0039	0.34
	E2297-B3	−0.0063	0.32
	E2297-C1	−0.0073	0.24
H	E2297-A1	−0.0027	0.11
	E2297-B3	−0.0047	0.21
	E2297-C1	0.00066	0.0012
I	E2297-A1	−0.0017	0.047
	E2297-B3	−0.0070	0.40
	E2297-C1	−0.0075	0.36
J	E2297-A1	−0.0052	0.26
	E2297-B3	−0.0046	0.37
	E2297-C1	0.0026	0.025
K	E2297-A1	0.0024	0.20
	E2297-B3	0.0024	0.25
	E2297-C1	0.0074	0.47
L	E2297-A1	−0.0026	0.17
	E2297-B3	0.0019	0.027
	E2297-C1	−0.0024	0.038

[a]Non-negligible bias with a moderate-to-good linear regression coefficient.

Figure 9. Comparison of group SWS calculated with displacement and velocity data in the phase I and II phantoms compared to equivalent processing of in vivo human data at varying fibrosis stages. The dashed line represents a unity ratio between the velocity- and displacement-based group SWSs that would be indicative of an elastic material; data points above this line would indicate a dispersive material. In the phase II phantoms, the group SWS calculated by using velocity data was 32% ± 1.9% greater than by using displacement data, whereas in the human data, the velocity-based group SWS was 27% ± 5.6% greater than the displacement-based group SWS across all fibrosis stages.



system and transducer models in the proposed QIBA profile.

These studies did not evaluate the differences that exist between different ultrasound systems in the presence of in vivo confounding factors, such as physiologic motion and challenging imaging artifacts, such as clutter and finding good acoustic windows. Additionally, whereas the range of stiffnesses and viscosities in the phase I and II phantoms represent realistic values that have been measured in healthy and fibrotic livers, they do not represent the full range of material parameters that may be encountered when estimating the SWS in the liver.

The results of these elastic and viscoelastic phantom studies have been incorporated into the measurement protocols described in the QIBA ultrasonic SWS profile to minimize interinstitutional and intersystem variability, and the inclusion of future standardized phantom and clinical SWS measurements will allow the profile to be refined in future revisions.[13]

In conclusion, elastic phantom measurements made across different manufacturer systems and different measurement sites had a mean difference 95% CI of ±0.145 m/s (±9.6%) across both phantoms, whereas viscoelastic phantoms had a mean difference 95% CI of ±0.340 m/s (±15.3%). The focal depth and appraiser were not appreciable sources of variability compared to the system and site. The best agreement between ultrasound systems and MRE in the elastic and viscoelastic phantoms was with an MRE excitation frequency of 140 Hz; the clinically implemented excitation frequency of 60 Hz had a − 12.2% bias, which could be a source of discrepancy in the literature between MRE and ultrasound systems characterizing liver fibrosis with the SWS. This study establishes a foundation for meaningful comparison of diagnostic SWS measurements made with different platforms.

Figure 10. Violin distributions of aggregate ultrasound SWS data across all systems and sites at a focal depth of 45 mm for each phase II phantom compared to discrete MRE measurements made at frequencies ranging from 60 to 200 Hz. The black box within each violin plot represents the interquartile range of the data, with the white circle representing the median value. Vertical lines extend away from each violin distribution to represent 1.5× the standard deviation of the data. The surrounding shape represents the probability density of the data.

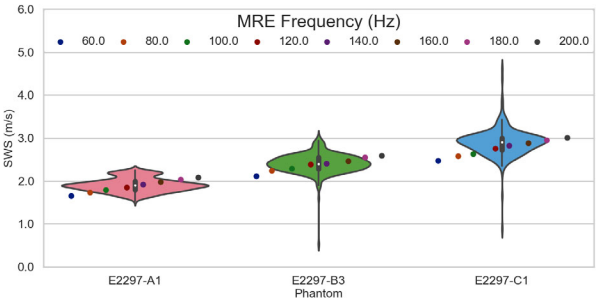
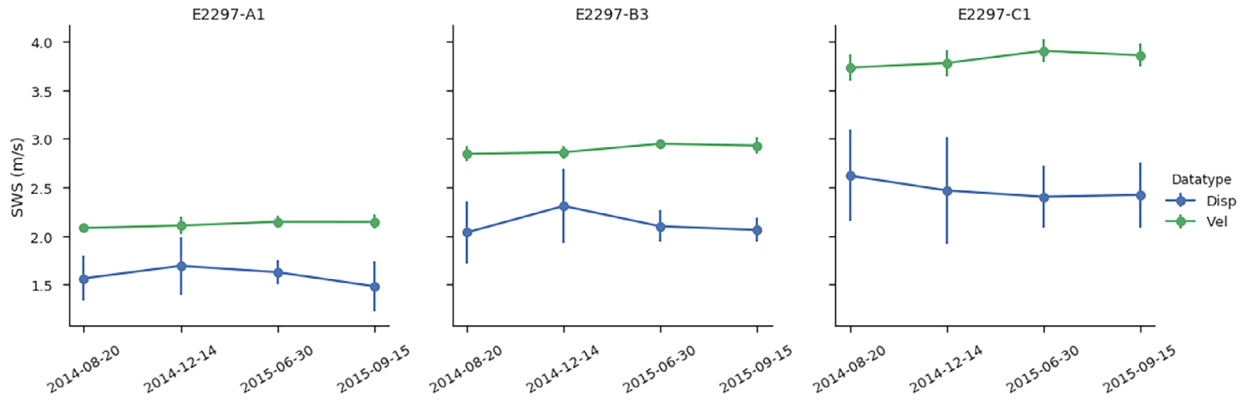


Figure 11. Measurements showing the longitudinal stability of the phase II phantoms using the group SWS calculated by using displacement and velocity data as representative metrics. The error bars represent the standard deviation over 16 independent measurements.



Appendix 1

Verasonics Data Acquisition Procedure

The following steps outline the procedure used to acquire phantom data:

1. Place the phantom on an optical isolation table to reduce room vibration artifacts. To aid in acquiring multiple independent speckle realizations at the same location in the phantom, the phantom can be placed on a rotating platform to avoid having to lift the transducer between acquisitions.
2. Remove the cover of the phantom, and pour enough saline solution to ensure adequate acoustic coupling with the transducer at a matched sound speed.
3. Secure the C5–2 transducer in a ring stand, and lower it onto the phantom.
4. Connect the transducer to the Verasonics scanner.
5. Initialize the Verasonics Vantage software (switch into the Verasonics directory in MATLAB [The MathWorks, Natick, MA]), and type "activate" in the command window. Run the C5–2 shear wave MTL setup script (https://github.com/RSNA-QIBA-US-SWS/VerasonicsPhantomSequences/blob/master/SetUpC5_2Shear_wave_MTL.m).
6. The setup script will save the acquisition structures to a MATLAB output file and display a VSX command in the command window.
7. Run this command in MATLAB to launch the Verasonics imaging graphical user interface. This interface will display live B-mode images.
8. Change the push voltage to 60 V (or adjust as necessary depending on the stiffness of the phantom to the voltage required for shear wave data with adequate displacement).
9. Click on the live B-mode image to acquire in-phase/quadrature (IQ) shear wave data. This will save 2 IQ data files (real and imaginary components of the data), as well as a parameters file in the indicated directory.
10. In the directory containing the IQ data, run the displacement processing using genDispMTL.m, which will generate an output file ([timestamp]_fromIQ_arfidata.mat).
11. Rotate the phantom to obtain a different speckle realization. Ensure that the transducer is appropriately coupled to the phantom, and repeat the acquisition until there are an adequate number of replicate displacement data.

Verasonics sequences and postprocessing code for the generated data are available for download (https://github.com/RSNA-QIBA-US-SWS/VerasonicsPhantomSequences).[14]

References

1. Ferraioli G, Filice C, Castera L, et al. WFUMB guidelines and recommendations for clinical use of ultrasound elastography, part 3: liver. *Ultrasound Med Biol* 2015; 41:1161–1179.
2. Barr RG, Ferraioli G, Palmeri ML, et al. Elastography assessment of liver fibrosis: Society of Radiologists in Ultrasound consensus conference statement. *Radiology* 2015; 276:845–861.
3. Arvazyan A, Rmen PS, Udenko O, et al. Shear wave elasticity imaging: a new ultrasonic technology of medical diagnostics. *Ultrasound Med Biol* 1998; 24:1419–1435.
4. Doherty JR, Trahey GE, Nightingale KR, Palmeri ML. Acoustic radiation force elasticity imaging in diagnostic ultrasound. *IEEE Trans Ultrason Ferroelectr Freq Control* 2013; 60:685–701.
5. National Institute for Health and Care Excellence. Hepatitis B (chronic): diagnosis and management. National Institute for Health and Care Excellence website. https://www.nice.org.uk/guidance/cg165. Accessed February 2018.
6. Ferraioli G, De Silvestri A, Lissandrin R, et al. Evaluation of intersystem variability in liver stiffness measurements. *Ultraschall Med* 2019; 40:64–75.
7. Nadebaum DP, Sood S, Gibson RN. Variability of liver shear wave measurements using a new ultrasound elastographic technique. *J Ultrasound Med* 2017; 37:647–656.
8. Long Z, Tradup DJ, Song P, et al. Clinical acceptance testing and scanner comparison of ultrasound shear wave elastography. *J Appl Clin Med Phys* 2018; 19:336–342.
9. Rouze NC, Deng Y, Trutna CA, Palmeri ML, Nightingale KR. Characterization of viscoelastic materials using group shear wave speeds. *IEEE Trans Ultrason Ferroelectr Freq Control* 2018; 65:780–794.
10. Palmeri ML, Deng Y, Rouze NC, Nightingale KR. Dependence of shear wave spectral content on acoustic radiation force excitation duration and spatial beamwidth. *Proceedings of the 2014 IEEE International Ultrasonics Symposium*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2014:1105-1108.
11. Hall TJ, Milkowski A, Garra B, et al. RSNA/QIBA: Shear wave speed as a biomarker for liver fibrosis staging. *Proceedings of the*

2013 IEEE International Ultrasonics Symposium. Prague, Czechia: Institute of Electrical and Electronics Engineers; 2013:397–400.

12. Venkatesh SK, Yin M, Ehman RL. Magnetic resonance elastography of liver: technique, analysis, and clinical applications. *J Magn Reson Imaging* 2013; 37:544–555.
13. Radiological Society of North America. QIBA profile: ultrasound measurement of shear wave speed for estimation of liver fibrosis. Quantitative Imaging Biomarker Alliance, profile stage: public comment. Radiological Society of North America website. http://qibawiki.rsna.org/index.php/Profiles. Accessed October 2019.
14. Deng Y, Rouze NC, Palmeri ML, Nightingale KR. Ultrasonic shear wave elasticity imaging sequencing and data processing using a Verasonics research scanner. *IEEE Trans Ultrason Ferroelectr Freq Control* 2017; 64:164–176.
15. Palmeri ML, Wang MH, Rouze NC, et al. Noninvasive evaluation of hepatic fibrosis using acoustic radiation force–based shear stiffness in patients with nonalcoholic fatty liver disease. *J Hepatol* 2011; 55:666–672.
16. Morris DC, Rouze NC, Palmeri ML, Nightingale KR. Group shear wave based viscoelastic parameter estimation in SWEI: analysis of sources of bias. *Proceedings of the 2017 IEEE International Ultrasonics Symposium*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2017:1-4.
17. Zhao H, Song P, Urban MW, et al. Bias observed in time-of-flight shear wave speed measurements using radiation force of a focused ultrasound beam. *Ultrasound Med Biol* 2011; 37:1884–1892.
18. Deng Y, Rouze NC, Palmeri ML, Nightingale KR. On system-dependent sources of uncertainty and bias in ultrasonic quantitative shear-wave imaging. *IEEE Trans Ultrason Ferroelectr Freq Control* 2016; 63:381–393.
19. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. *SciPy* 2010 website; 2010; 1:92–96.
20. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020; 17:261–272.
21. Oliphant TE, Manduca A, Ehman RL, Greenleaf JF. Complex-valued stiffness reconstruction for magnetic resonance elastography by algebraic inversion of the differential equation. *Magn Reson Med* 2001; 45:299–310.
22. Manduca A, Oliphant TE, Dresner MA, et al. Magnetic resonance elastography: non-invasive mapping of tissue elasticity. *Med Image Anal* 2001; 5:237–254.
23. Palmeri ML, Qiang B, Chen S, Urban MW. Guidelines for finite-element modeling of acoustic radiation force–induced shear wave propagation in tissue-mimicking media. *IEEE Trans Ultrason Ferroelectr Freq Control* 2017; 64:78–92.