RSNA/QIBA Shear Wave Speed as a Biomarker for Liver Fibrosis Staging: Elastic (Phase I) and Viscoelastic (Phase II) Phantom Studies

Mark L. Palmeri, Andy Milkowski, Timothy J. Hall, Brian Garra, Nancy Obuchowski, Kathy R. Nightingale, Ned C. Rouze, Ted Lynch, Matt Urban, Shigao Chen, Jun Chen, Richard Ehman, Paul Carson

## Keywords

QIBA, acoustic radiation force, shear wave, phantom, elasticity

# Abstract

Commercial shear elasticity systems have been available to clinicians for several years for characterizing liver stiffness. Shear wave speed (SWS) is a common metric that these systems can report, and we have conducted an interlaboratory study of shear wave speed estimation using multiple commercial systems in homogeneous, isotropic elastic tissue-mimicking phantoms. The commercial systems captured in this study included the Echosens™ FibroScan, Philips EPIQ 5. Supersonic Imagine Aixplorer®, Siemens ACUSON S2000™, and two research configurations implemented on Verasonics Vantage research scanners. Measurements were also made with a Magnetic Resonance Elastography (MRE) system. Eleven pairs of elastic phantoms of two different stiffnesses were manufactured by CIRS, Inc. Overall, SWS reported across different systems was consistently within 5% of the grand mean of measurements made by all systems. Statistically-significant differences were observed between different manufacturer systems and the same system at different measurement sites. Different appraisers and replicate measurements at a given were not sources of significant variability in SWS measurements. A negative bias in SWS was seen in all systems as a function of increasing focal depth, though that trend can be minimized in these phantoms by using a transducer coupling fluid with a sound speed that matches that of the phantoms. A repository of research scanner shear wave sequences and processing code has been created to allow manufacturers to have a calibration standard to compare their systems against for compliance with the QIBA US SWS profile and to achieve more consistent results for more rapid clinical adoption of this technology to evaluate liver fibrosis.

# Introduction

In 2008 the Radiological Society of North America (RSNA) created the Quantitative Imaging Biomarker Alliance (QIBA) Alliance (QIBA) with pharmaceutical companies, imaging system manufacturers, academics, clinicians and representatives from the USA federal government (e.g., FDA, NIH, NIST) to advance the concept of converting “imaging systems” to “measurement systems.” The Alliance is organized by Modality Committees and within these committees are Technical Committees whose efforts involve specific classes of biomarkers. Each technical committee is supported by one or more Subcommittees organized for specific tasks. The goal is to create QIBA/UPICT (Uniform Protocol for Imaging in Clinical Trials) protocols that specify methods for data acquisition, analysis and interpretation as well as QIBA Profiles that will provide specific claims of what can be accomplished by following the QIBA Protocol. The intent is to validate the profile across imaging systems with phantoms and volunteers working with other organizations such as drug and instrument companies and clinical trials organizations.

The Ultrasound Modality Committee was formed in 2012. Our only Technical Committee (TC) is the Ultrasound Shear Wave Speed TC. We are developing a protocol and data analysis methods to allow direct comparison of shear wave speed (SWS) measurements in liver for staging fibrosis. Several systems that measure SWS in the liver are commercially available, and many articles are reporting the ability to differentiate among fibrosis stages. Existing literature suggests that different estimates of shear wave speed in the same liver are obtained with different measurement systems (see, for example, \cite{Sporea2012}). These differences can cause uncertainty and a lack of technology adoption in the clinical community. Given the need for serial assessment of liver fibrosis, and the impracticality of serial liver biopsy, providing a common SWS estimate among systems would make these studies more clinically viable and speed adoption of the technology.

The first step toward understanding sources of bias and variance in SWS estimates is to perform a study in relatively simple phantoms. Our intent was to measure SWS in a set of equivalent phantoms containing homogeneous, isotropic and nearly completely elastic material. The lack of a loss component in the complex modulus was expected to ease direct comparison among the various SWS estimation systems (for example, the frequency content of the shear wave wouldn't matter in such media). The methods, preliminary results and conclusions of our initial investigation are described below.

# Methods

## Phantom Calibration

Eleven pairs of elastic phantoms with nominal shear wave speeds of 1.0 and 2.0 m/s were fabricated by CIRS, Inc. (Norfolk, VA). The phantoms were homogeneous cylinders that were 10 cm in diameter and height, except for a pair of phantoms designed for MRE measurements that were 20 cm in diameter and 12 cm in height to reduce standing wave reflections off the phantom walls. The shear wave speeds in all the phantoms were measured at Duke University using a Verasonics research scanner sequence (Table 1) to correct for fabrication variability between the phantoms. Ten replicate measurements were made across 10 different speckle realizations at 3 different focal depths (40, 60 and 80 mm) in each phantom.\

Table : Acoustic radiation force excitation and displacement tracking parameters used on a Verasonics research scanner with a Philips C5-2 curvilinear array to measure all the phantoms before distribution to individual measurement sites.

|  |  |
| --- | --- |
| **Excitation Parameters** |  |
| Frequency | 2.4 MHz |
| Focal Depths | 40, 60, 80 mm |
| F-number | F/2.0 |
| Duration | 405 µs (960 cycles) |
| **Tracking Parameters** |  |
| Frequency | 3.2 MHz |
| Transmit F-number | Plane-wave |
| Receive F-number | F/2.0 |
| Pulse Repetition Frequency | 5 kHz |

Group SWS was estimated using displacement data and a RANSAC estimation algorithm [CITE WANG].

## Magnetic Resonance Elastography (MRE)

There were three major steps to measure stiffness on QIBA phase II phantoms (CIRS) using MR Elastography (MRE): 1) Shear wave generation using shear driver: To generate shear wave propagation in the phantoms, a square MRE electromechanical PVC shear driver (6.4 cm X 6.4 cm X 0.3 cm) was placed on top of the phantom with a light compression for keeping the shear mechanical coupling. The driver frequency ranged from 60 Hz to 200 Hz (20Hz interval), with MRE performed at each mono frequency respectively. 2) Shear wave Imaging sequence: To acquire shear wave propagation images in the phantoms using a 3D MRE wave imaging sequence, the following major parameters were used in the study: FOV = 21.6 cm, matrix = 128X128, TR = 1600 - 2314 ms, TE = 62.7 - 119 ms, slice thickness/spacing = 3.5/0 mm, 16 slices, motion sensitivity (MENC) = 4.5 - 25.2 micron/pi radians, motion sensitivity direction = x/y/z, axial imaging plane. One channel head coil and a 1.5T GE Signa scanner (Waukesha, WI, USA) were used. 3) Complex shear modulus calculation: To process wave images and compute elastograms, 3D MRE direct inversion (DI) algorithm was used [1]. ROIs were then drawn in the middle 8 slices of stiffness map. Mean and standard deviation of magnitude complex shear stiffness (kPa), as well as its real and imaginary parts (kPa) were reported. Wave speed, phase velocity (m/s) was calculated by converting complex shear modulus according to = + ,   , whereas and are the real and imaginary parts of the complex shear modulus (kPa), and is the phase velocity (m/s)

Prior to the MRE exams, phantoms were kept in different rooms at different temperatures (20°C and 23.9°C) overnight for at least 8 hours to achieve temperature equilibrium.

MRE measurements were made at the Mayo Clinic using a research system at frequencies of 140, 180, 200, 300, 400, 500 Hz. Note that these frequencies are greater than those of clinical MRE systems (60 Hz), but closer to the expected bandwidth of the ultrasound SWS systems. All MRE measurements were made in a room at a constant 20 °C temperature. Complex shear moduli were measured in the phantoms, converted to SWS, these SWS were fit using linear regression over the measured bandwidth.

## Site Measurement Protocol

FILL ME IN WITH MORE DETAILS FROM THE OFFICIAL PROTOCOL

The phantoms were distributed among 15 sites (12 included in the current analysis) for measurements on commercial clinical SWS-capable systems (Fibroscan, Philips, Siemens S2000, SSI Aixplorer) as well as experimental systems under development in some labs.

Before data acquisition at the various sites commenced, the SWS TC members agreed to have measurements performed at three depths relevant for liver evaluation (3.0, 4.5 and 7.0 cm). Each site had at least three appraisers scan each phantom at least three times to estimate the bias and components of variance in SWS estimates among systems. The order of data acquisition was randomized for phantoms, appraisers, depths, and imaging systems (if more than one was used) to allow for detailed statistical investigation of results. Participants were all blinded to the intermediate results of others.

# Results

Figure 1 shows the measurements made on all the phantoms using the same research sequence at Duke using a Verasonics research scanner with the configuration detailed in Table 1. These data were used to normalize all the phantoms to minimize phantom fabrication variability as a confounder when comparing the measurements made at the different sites with different systems.

Figure 2 and Figure 3 show the aggregate SWS measurements from all of the sites and systems grouped by unique site and system, respectively, for the soft (blue) and stiff (green) phantoms at all 3 focal depths (3.0, 4.5 and 7.0 cm). All the measurements were made with 3 different appraisers, with 10 replicate measurements per unique appraiser / focal configuration / speckle realization. No temporal dependence was found in the data acquisition process when randomized for acquisition order among appraisers, phantoms, and acquisition depths.

Table 2 summarizes the system bias and variability across all the measurement sites. Table 3 summarizes the bias and variability of SWS measurements made at the different unique sites, highlighting the exclusion of one site from the analysis since it was an extreme outlier that otherwise biased all the statistical conclusions.

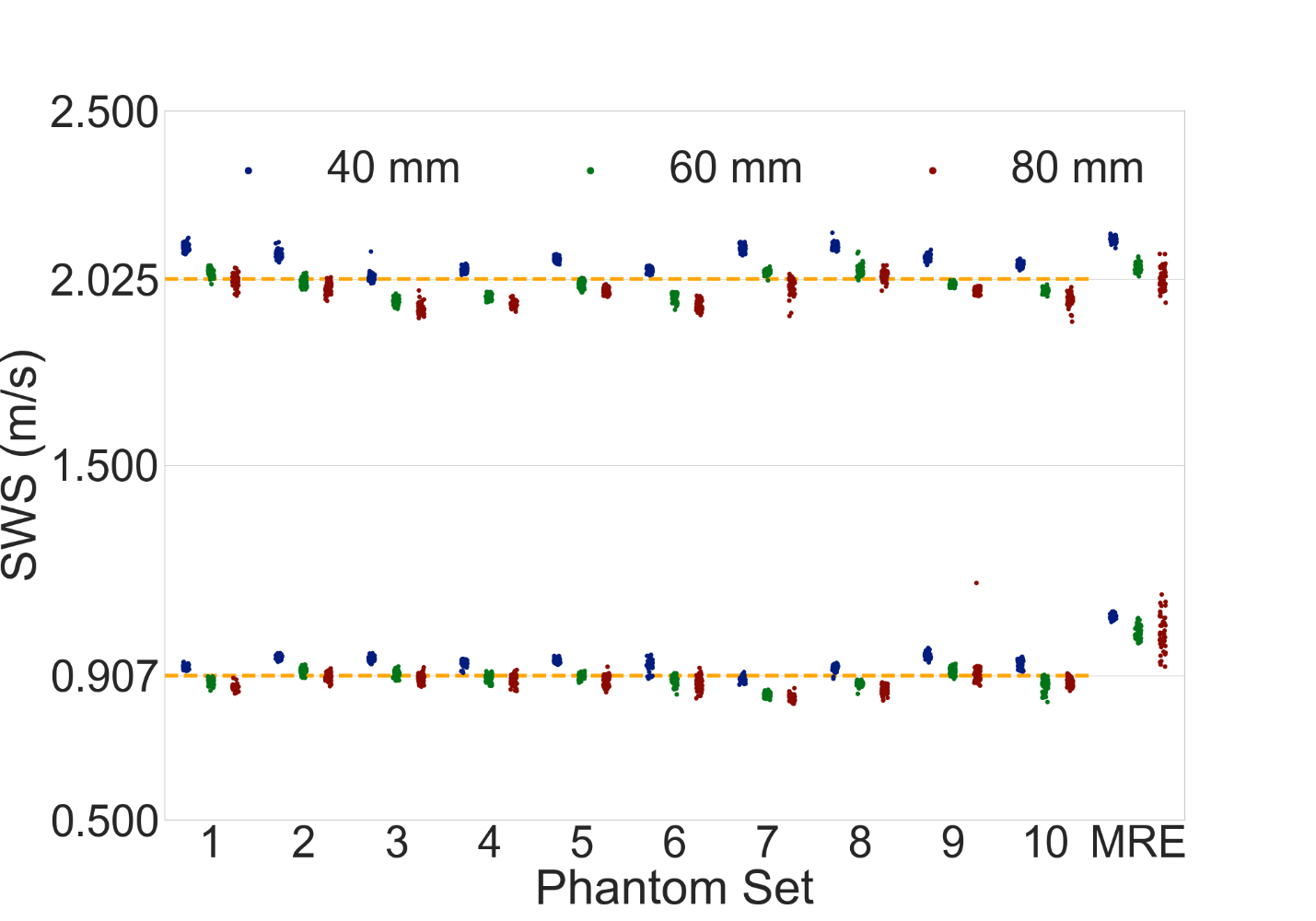
[](figs/duke_all_phantoms.png)

Figure : Calibration measurements on all the softer (E1786) and stiffer (E1787) elastic ultrasound phantoms and MRE phantoms (E1788) using a Duke sequence at 3 different focal depths (40, 60 80 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±0.33 m/s and 2.025±0.51 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.

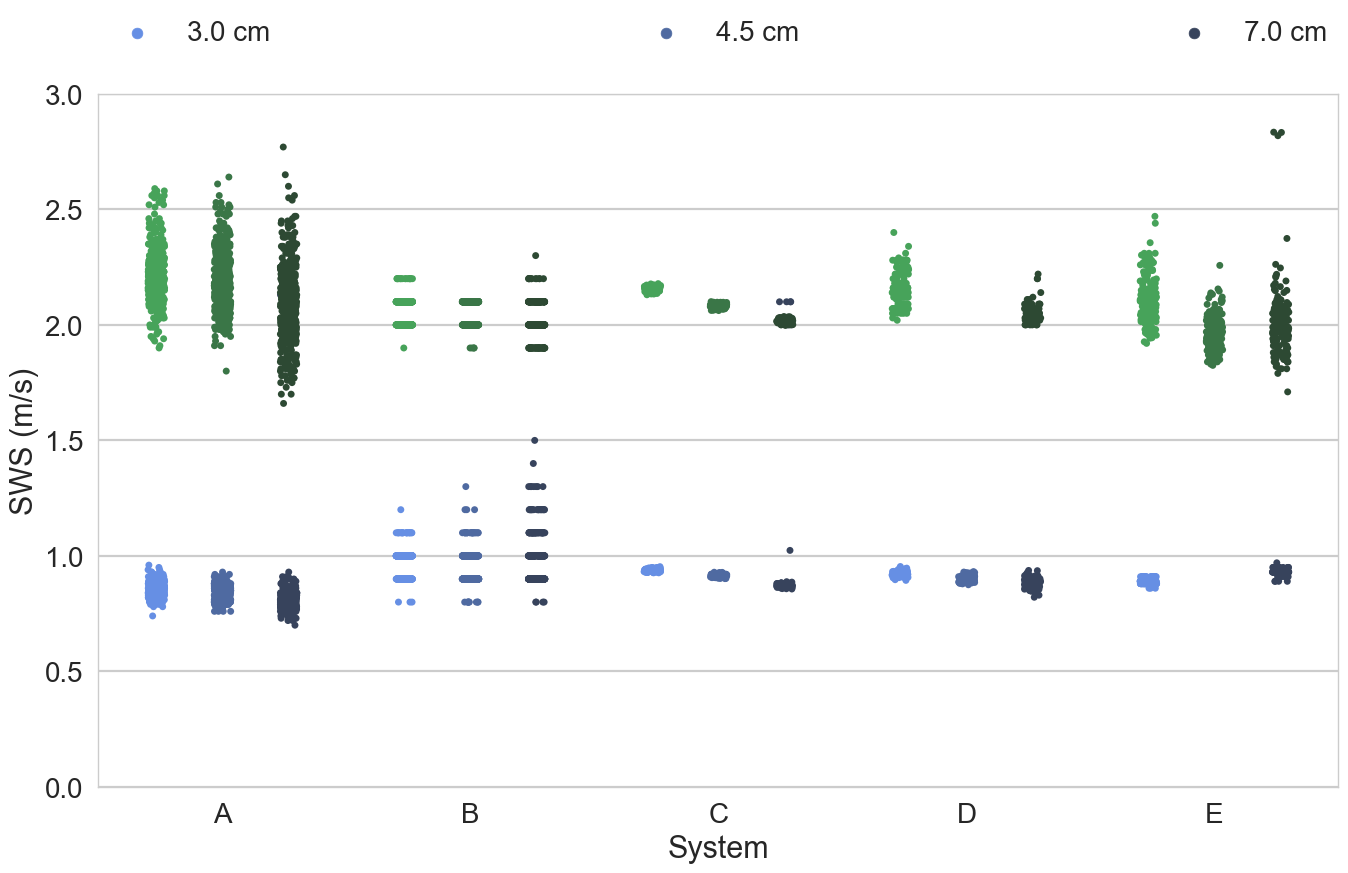
[](figs/sws_system_alldepths.png)

Figure 4 and Figure 5 show the aggregate SWS measurements for 2 different systems that were each used at 5 different measurement sites. These plots highlight the differences in both the distribution of SWS measurements for each system, and the relative precision of the reported SWS on each system.

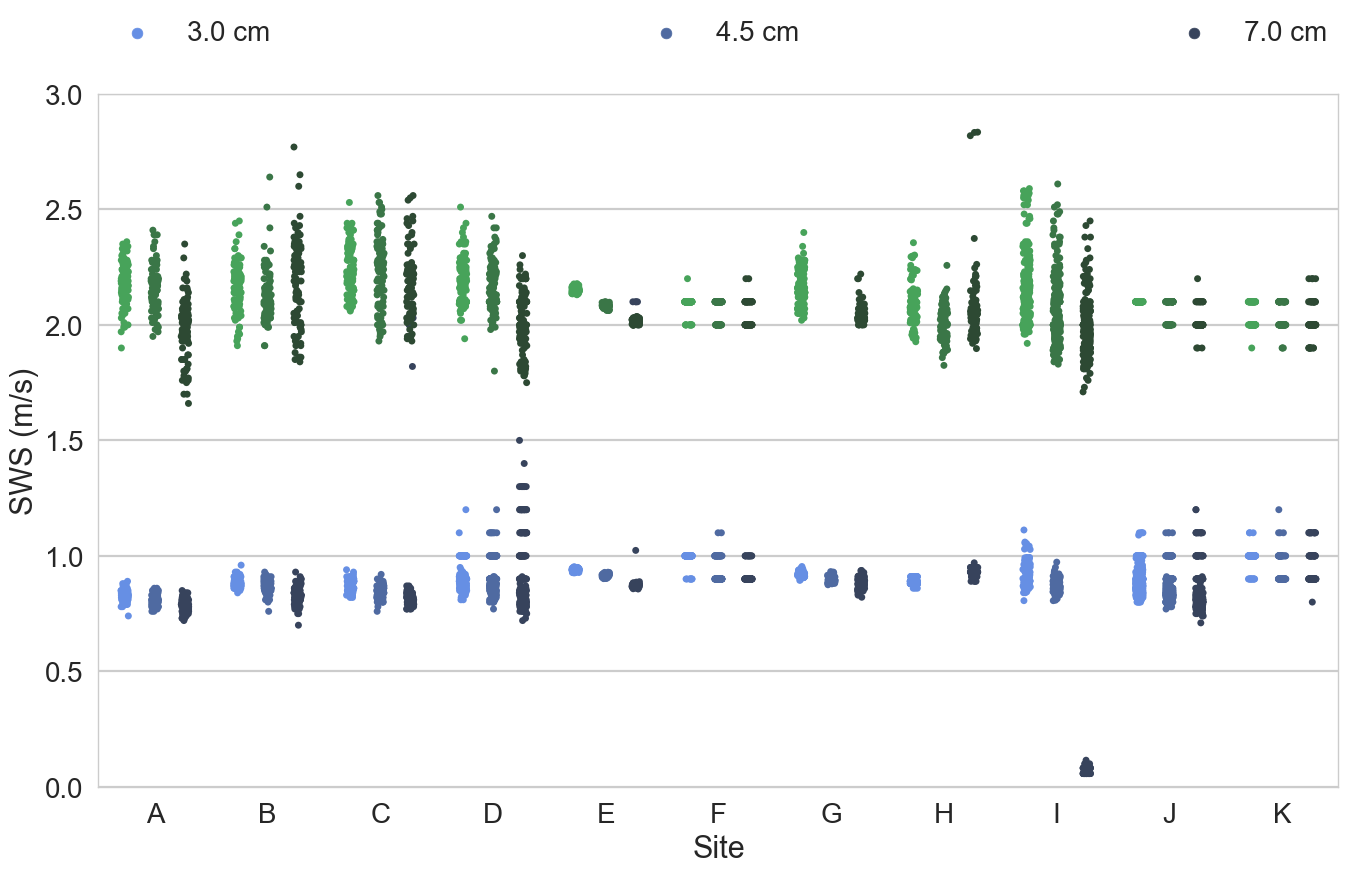
[](figs/sws_site_alldepths.png)

Figure : 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 (3.0, 4.5 and 7.0 cm) in each phantom.

Figure : All the elastic phantom data presented in Figure 2 grouped by unique system instead of measurement site. Some systems were used at only a single measurement site, while other systems were used at multiple measurement sites.

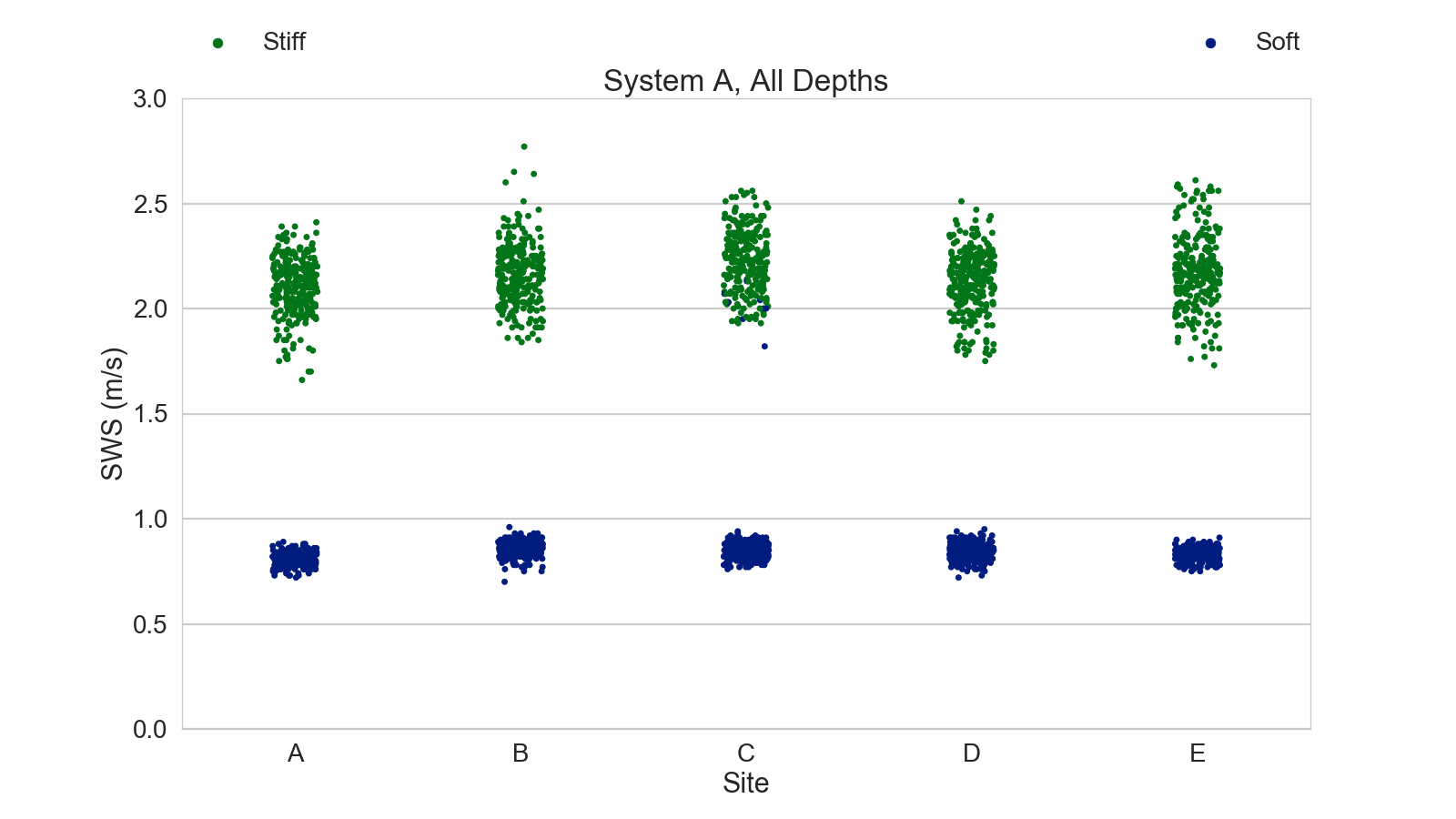
[](figs/sws_systemA_alldepths.png)

Figure : Aggregate SWS measurements in the soft (blue) and stiff (green) elastic phantoms from all appraisers and focal depths for System A that was used at 5 different measurement sites. Note that at Site C there were several soft phantom data points that were erroneously categorized as being from the stiff phantoms; such transcription errors were not retrospectively corrected in the statistical analysis.

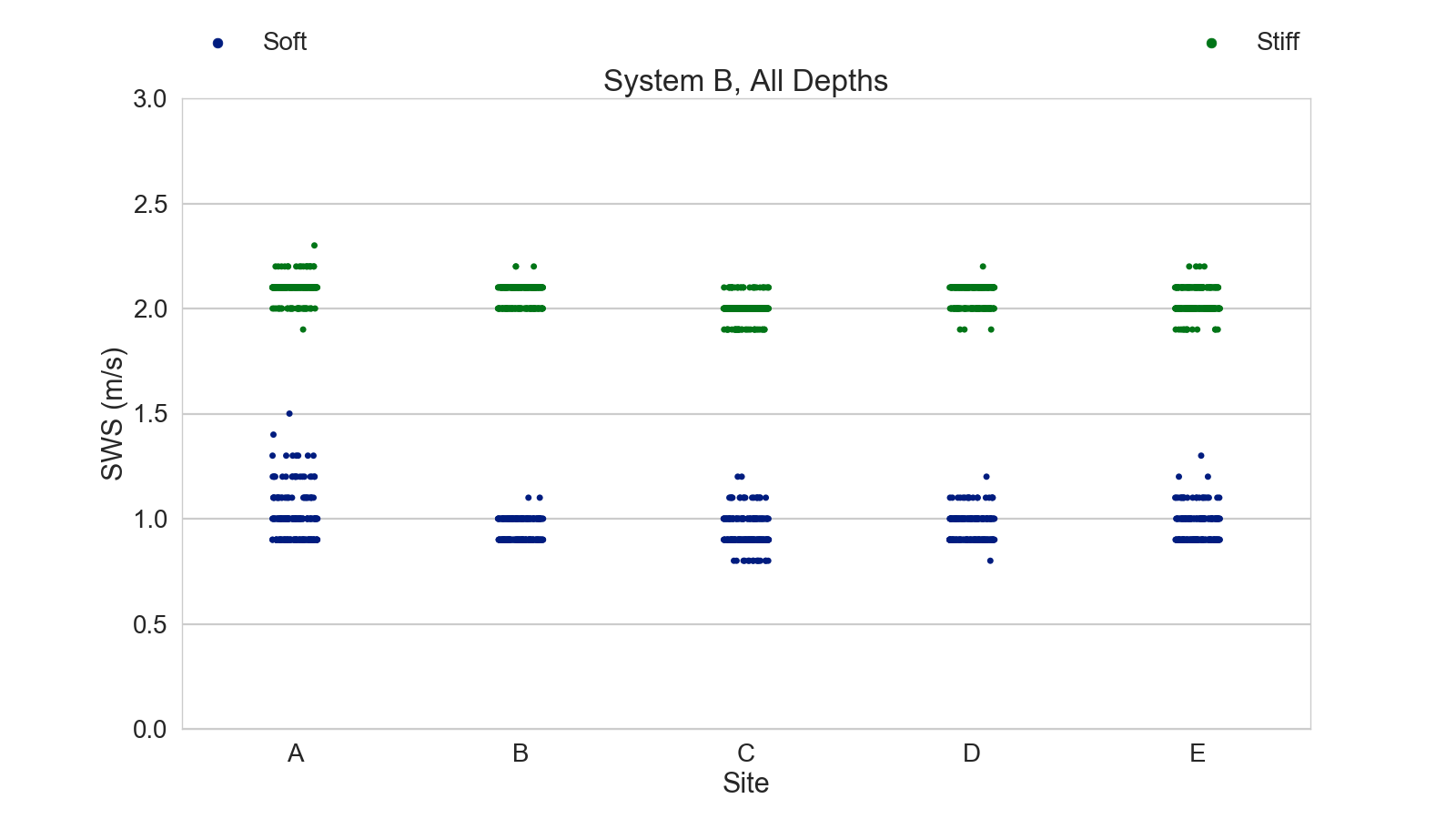


Figure : Aggregate SWS measurements in the soft (blue) and stiff (green) elastic phantoms from all appraisers and focal depths for System B that was used at 5 different measurement sites. Notice the difference in the precision of the reported SWS for this system compared to the precision of the system shown in Figure 4.

|  |  |  |
| --- | --- | --- |
| **System** | **Mean Bias (%)** | **Mean Std Dev (%)** |
| **A** | 0. | 2. |
| **B** | -9. | 5. |
| **C** | +2. | 7. |
| **D** | +2. | 5. |
| **E** | +2. | 2. |

Table : Mean system bias and standard deviations across all measurement sites and focal configurations.

|  |  |  |
| --- | --- | --- |
| **Site** | **Mean Bias (%)** | **Mean Std Dev (%)** |
| **A** | 0. | 5. |
| **B** | 0. | 1. |
| **C** | 0. | 2. |
| **D** | +3. | 10. |
| **E** | +1. | 3. |
| **F** | +2. | 2. |
| **~~G~~** | ~~+622.~~ | ~~53.~~ |
| **H** | -13. | 6 |
| **I** | +4 | 8 |
| **J** | -1. | 5 |
| **K** | +2. | 4 |
| **L** | -2. | 6 |
| **M** | +9. | 7 |
| **N** | +1. | 6 |

Table : Mean SWS bias and standard deviation across all Phase I phantom measurements made with all systems as a function of unique site. Site G data was excluded from all of the analysis in this manuscript given that it was an extreme outlier compared to all of the other site data.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Df** | **Sum Sq** | **Mean Sq** | **F** | **p-value** |
| **Phantom** | 10 | 4.00 | 0.40 | 35.65 | < 0.001 |
| **System** | 5 | 1.18 | 0.24 | 20.94 | < 0.001 |
| **Site** | 9 | 0.58 | 0.06 | 5.71 | < 0.001 |
| **Focal Depth** | 1 | 0.03 | 0.03 | 3.10 | 0.079 |
| **Residuals** | 838 | 9.41 | 0.01 |  |  |

Table : ANCOVA results for the Phase I phantoms, where appraiser and replicate measurements were shown to not be sources of significant differences and were aggregated in this analysis.

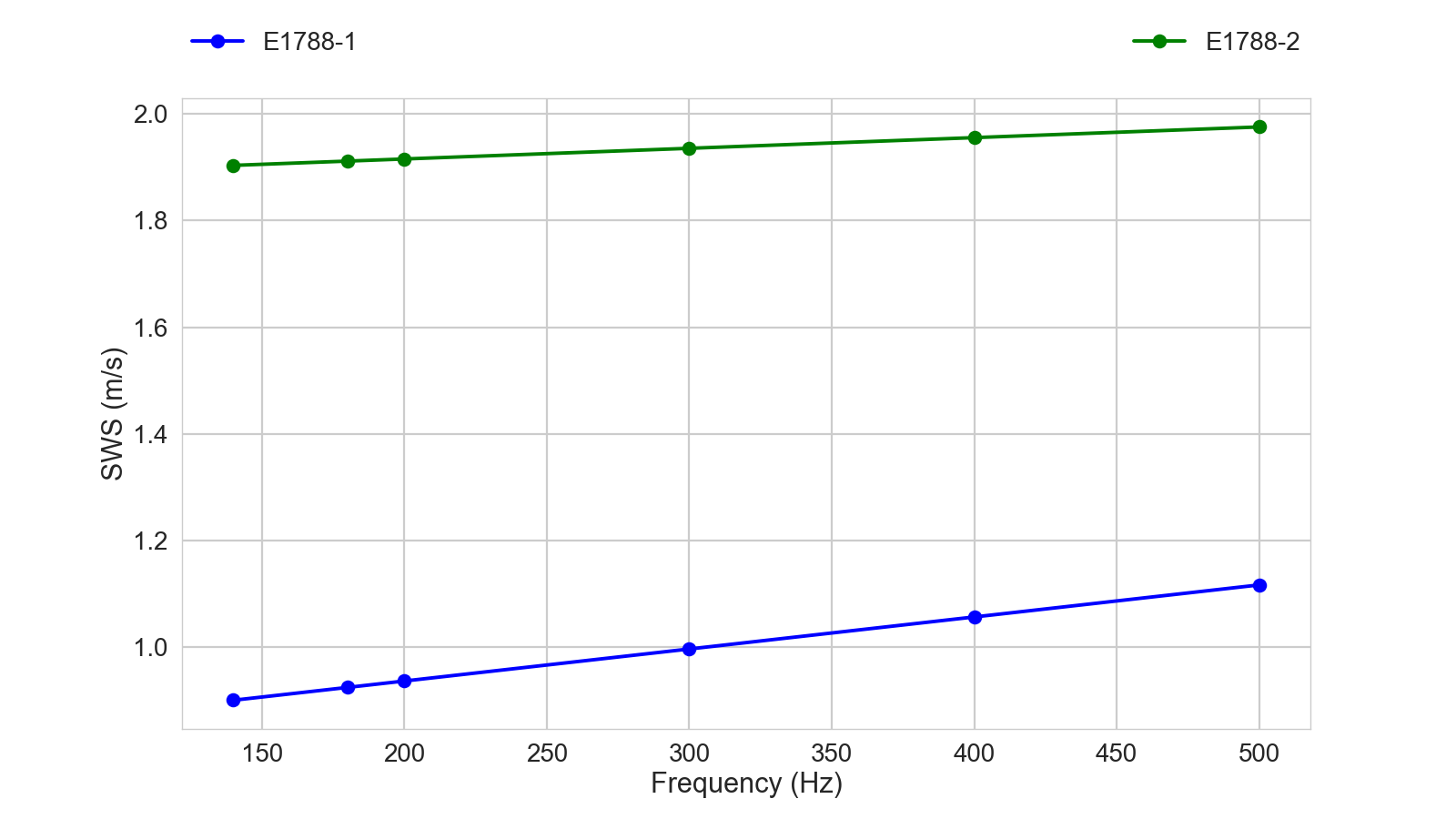


Figure : MRE measurements of the soft (blue) and stiff (green) phantoms. The data markers represent the mean SWS measurements at discrete frequencies of 140, 180, 200, 300, 400 and 500 Hz. The linear fits for each phantom had R2 = 0.9698 and R2 = 0.9689 for the soft and stiff phantoms, respectively.

Figure 7 shows the comparison of different systems measuring the SWS in the 3 Phase II phantoms at 3 different focal depths (3.0, 4.5 and 7.0 cm). The orange line for each phantom plot represents the grand median value across all the measurement systems for each phantom.

Figure 8 shows how the same manufacturer and model system used at different measurement sites can yield biases and differences in variance for the 3 different phantoms.

Table 5 shows the ANCOVA analysis for the Phase II phantom study.

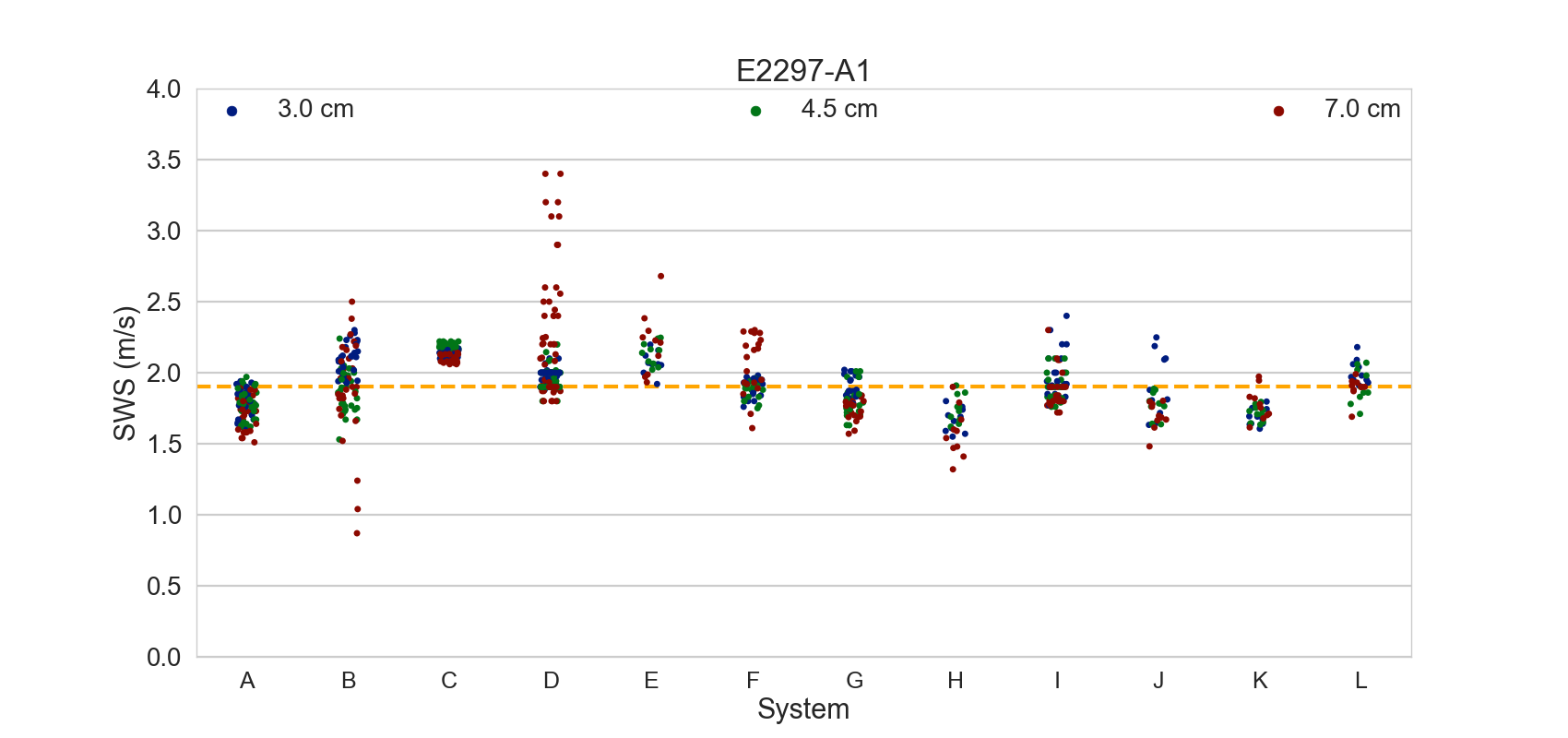
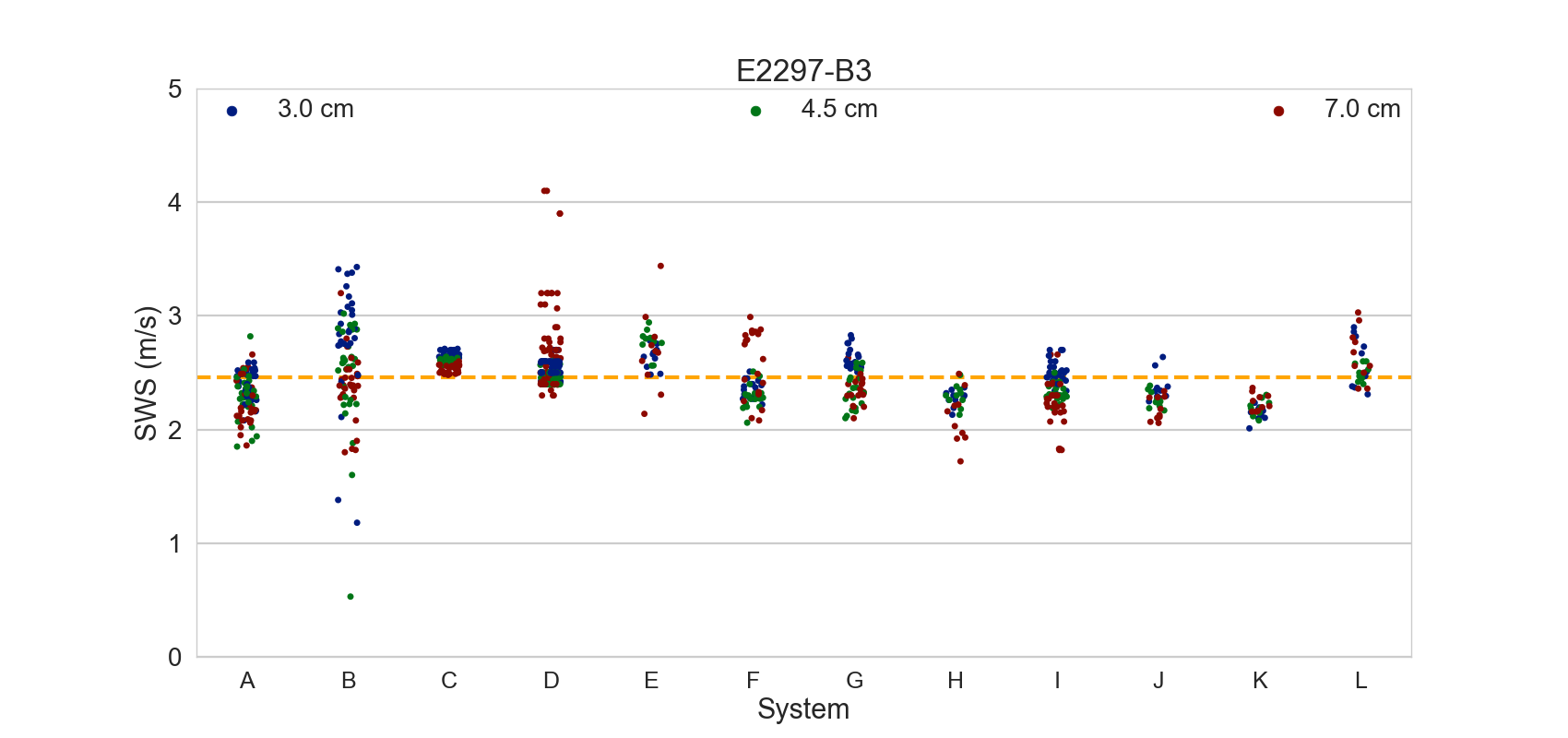
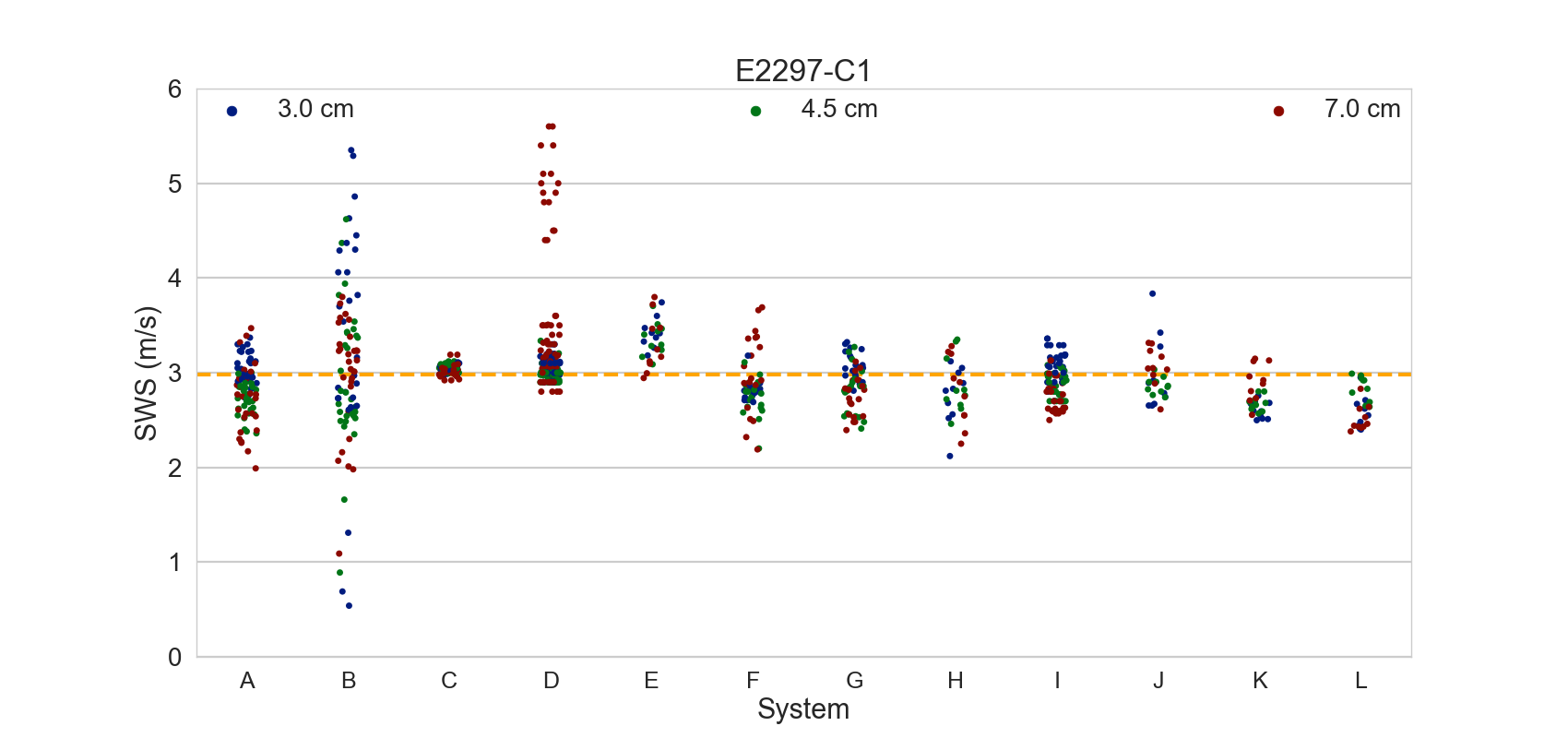
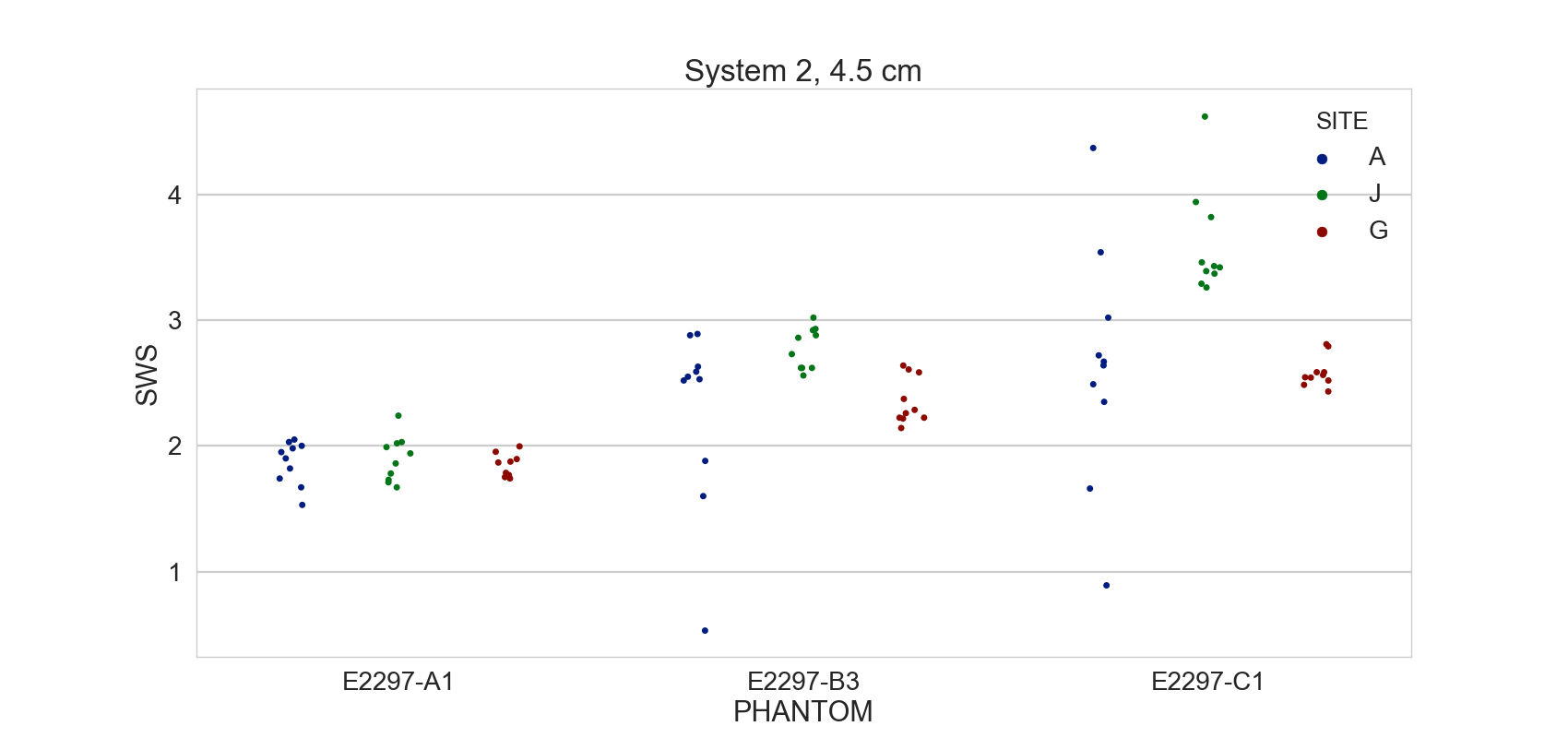
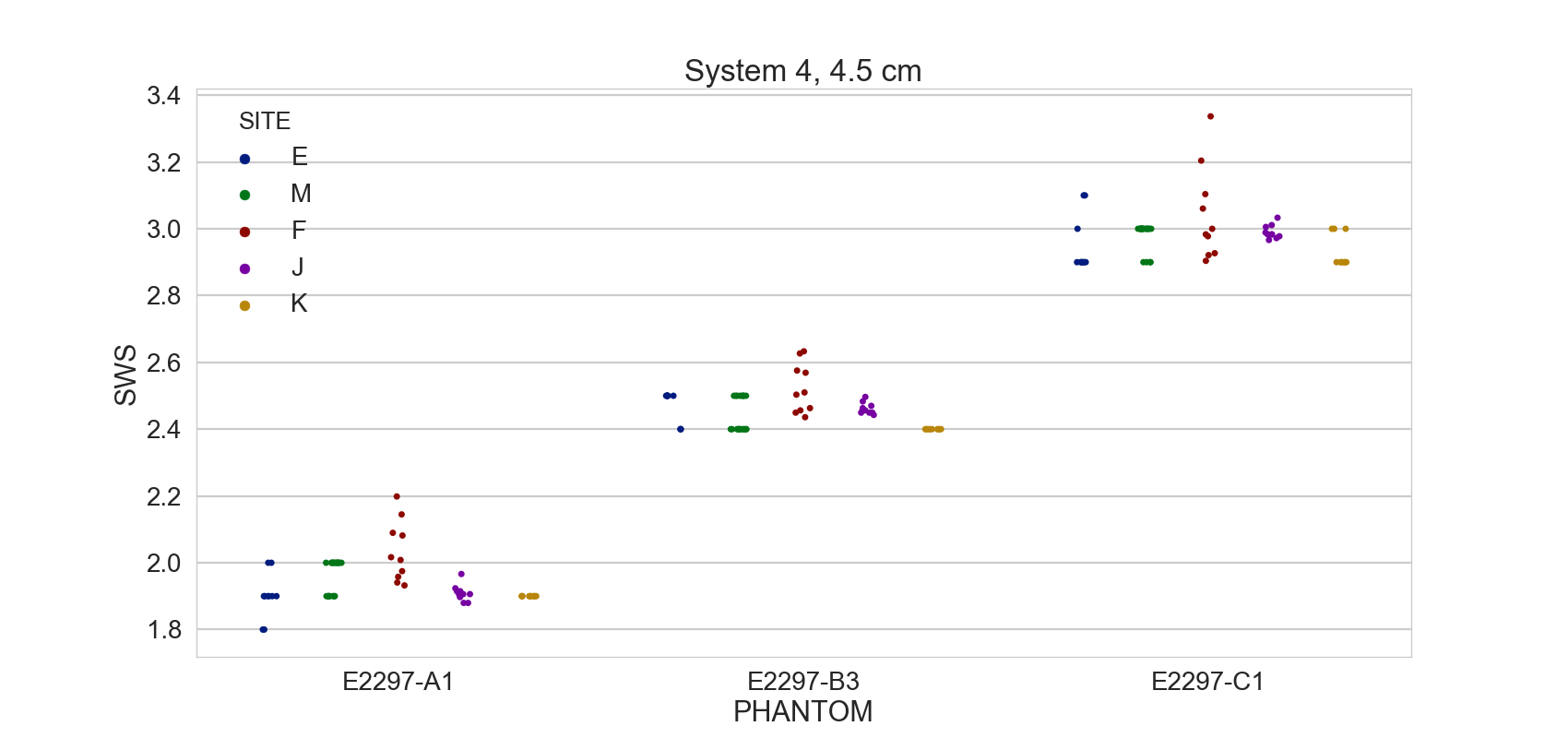


Figure : Phase II phantoms measured with different systems with 3 different focal depth configurations (3.0, 4.5 and 7.0 cm). The orange line on each plot represents the grand median value across all systems for each phantom.

Figure : Examples of the differences in measured SWS in the Phase II phantoms when using the same manufacturer and model system at different measurement sites for the 4.5 cm focal depth. Notice the bias and differences in variance as a function of site. These two systems were example systems that we used at least 3 different sites for this comparison.



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Df** | **Sum Sq** | **Mean Sq** | **F** | **p-value** |
| **Phantom** | 2 | 451.30 | 225.65 | 2617.63 | < 0.001 |
| **System** | 11 | 47.43 | 4.31 | 50.01 | < 0.001 |
| **Site** | 13 | 34.32 | 2.64 | 30.63 | < 0.001 |
| **Focal Depth** | 2 | 4.39 | 2.20 | 25.48 | < 0.001 |
| **Residuals** | 2405 | 207.32 | 0.09 |  |  |

Table : ANCOVA results in the Phase II phantoms.

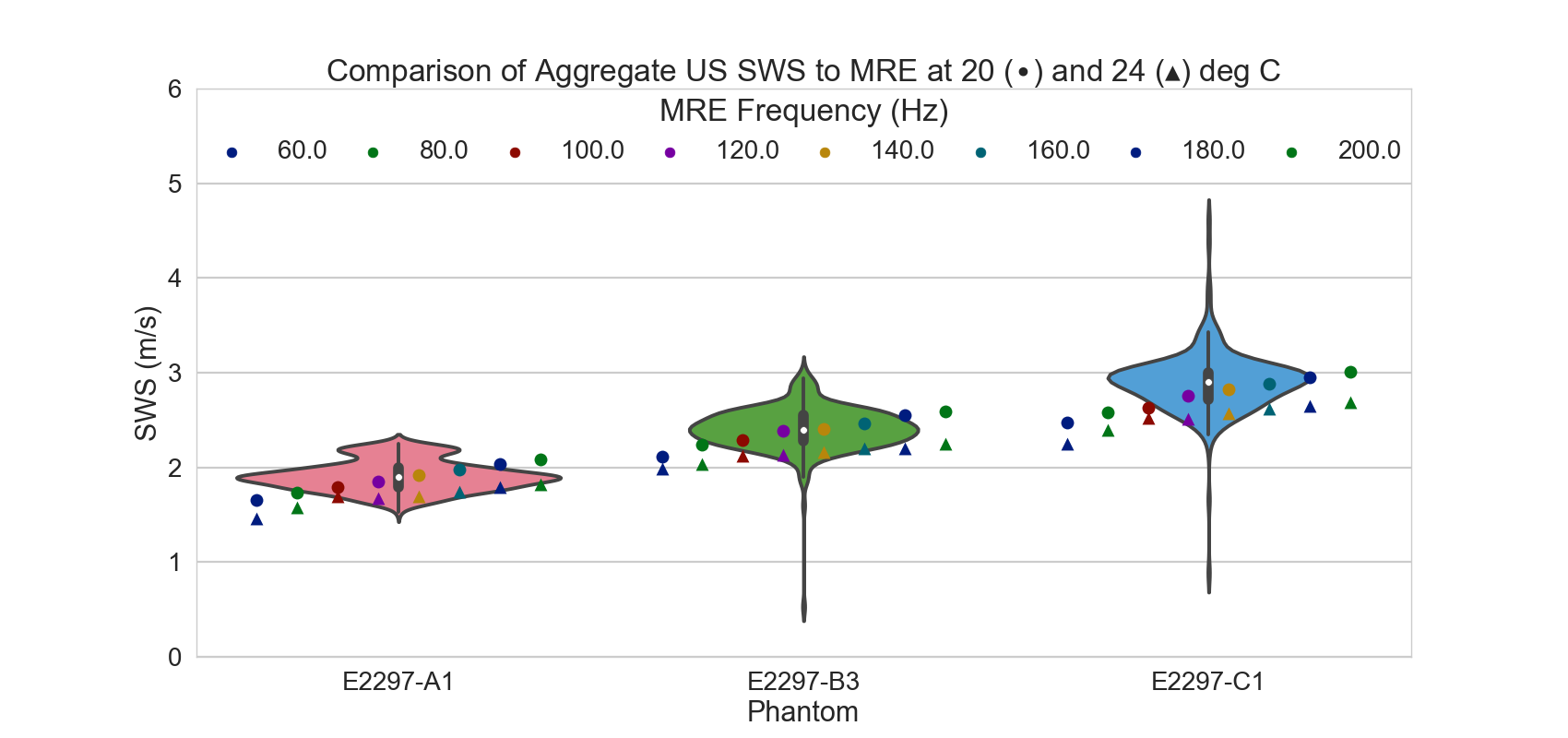


Figure : Comparison of aggregate ultrasound SWS data—shown as violin distributions—across all systems and sites at a focal depth of 4.5 cm for each phantom, compared with discrete MRE measurements made at frequencies ranging from 60—200 Hz at temperatures of 20 (circles) and 24 (triangles) degrees Celsius. 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 1.5x the standard deviation of the data. The surrounding shape represent the probability density of the data. Note that—as expected for a viscoelastic material—each phantom demonstrates an increase in reconstructed MRE shear wave speed with increasing frequency. The warmer environment leads to an overall decrease in the shear stiffness of the phantoms. Overall, MRE measurements made at frequencies ranging from 100-200 Hz match the overall ultrasound SWS measurements.

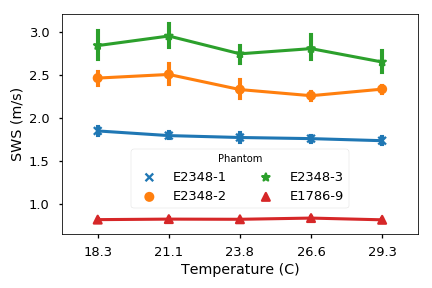


Figure : Temperature dependence of a Phase I (E1786-9) and equivalent Phase II (E2348-[1-3]) phantoms as a function of temperature ranging from 18—30 °C, as controlled using a water bath. SWS measurements were made using a commercial Siemens S2000 implementation.

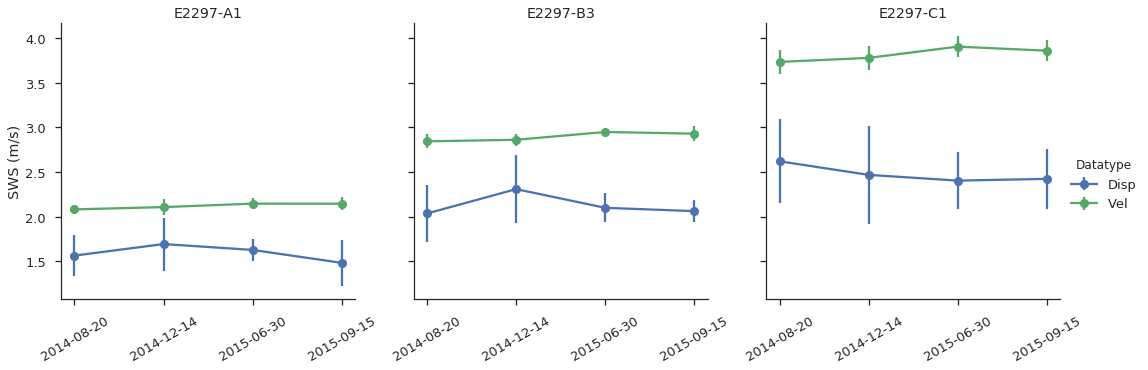
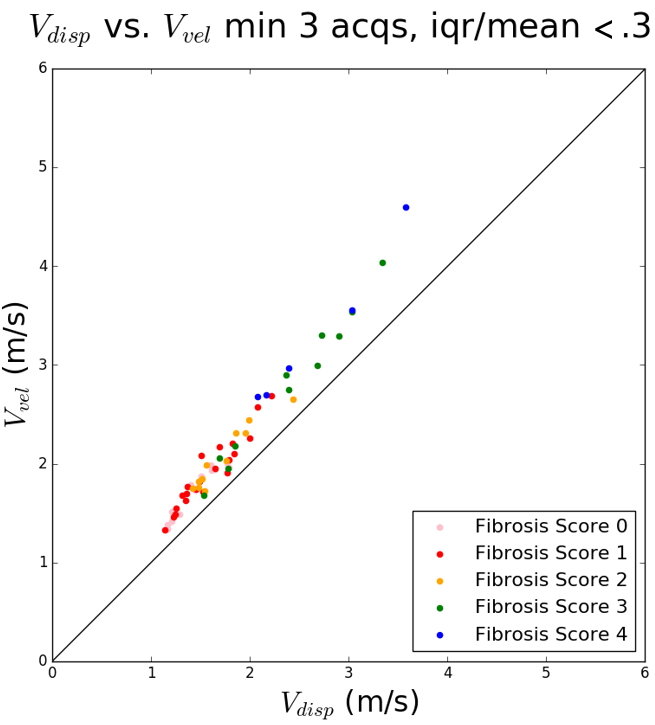
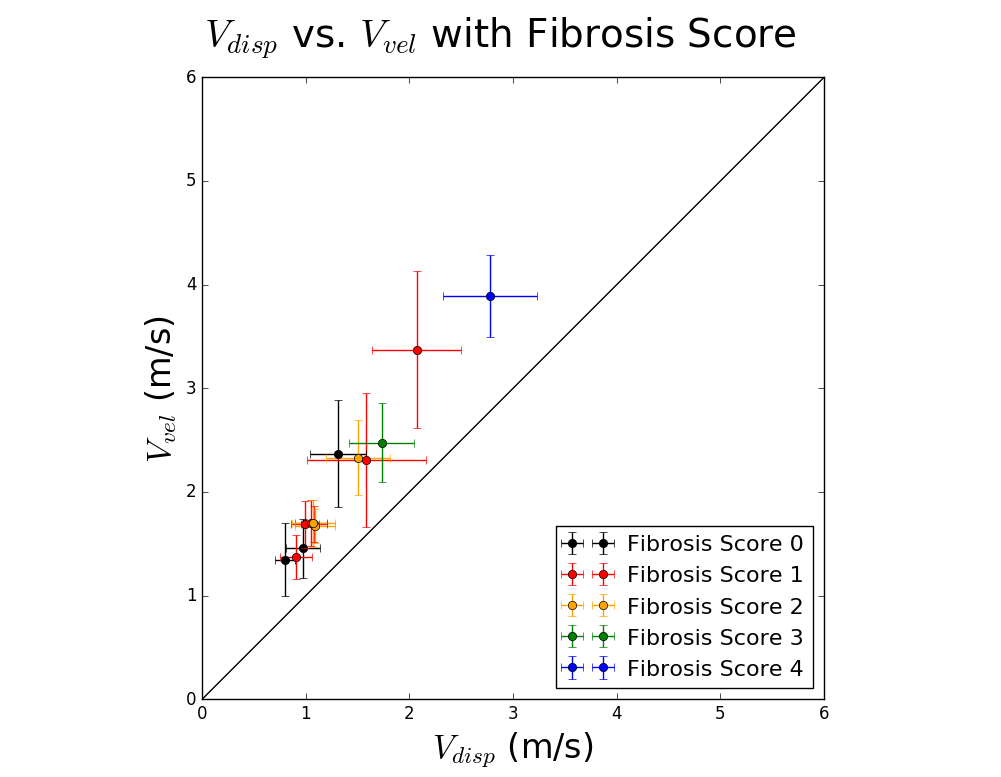


Figure : Measurements demonstrating the longitudinal stability of the Phase II phantoms using the shear wave phase velocity at 200 Hz and the linear slope of the shear wave phase velocity centered at 200 Hz as representative metrics.



# Discussion

The results of this study will allow us to greatly refine the plan for a Phase 2 interlaboratory study of SWS estimation. For example, our current plan is to change the phantom design to provide scanning windows on perpendicular surfaces. This will allow investigation of any potential depth-dependent material properties. A single phantom of each desired stiffness will be circulated among sites involved in the study. In that way, all sites are measuring the same phantom. An identical phantom will be kept at the manufacturing site and its properties will be monitored during the serial measurements at contributing sites. A smaller number of sites and participants can be involved based on the findings of the current study. More suggests for improvements on this study will likely result from further analysis of the existing data.

# Conclusions

There is a statistically significant difference in SWS estimates among systems and with depth into the phantom (demonstrated with all imaging systems). No statistically significant differences were found among appraisers using the same (or equivalent) systems or sites using equivalent systems. These are very encouraging results in our quest for equivalent SWS estimates among commercial systems.

## Acknowledgements

We gratefully acknowledge the support of CIRS, Inc. for providing the phantoms used in this study. We are also grateful to the RSNA for covering the costs of shipping the phantoms to the individual sites. The QIBA effort is funded in part by the RSNA and a contract with the NIBIB (HHSN268201000050C). 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 FDA.

# References

# Appendix A: Phase I Phantom Calibration Values

|  |  |  |  |
| --- | --- | --- | --- |
| **Phantom** | **Focus (mm)** | **Mean SWS (m/s)** | **STD SWS (m/s)** |
| **E1786-1** | 40 | 0.931 | 0.005 |
|  | 60 | 0.887 | 0.008 |
|  | 80 | 0.872 | 0.008 |
| **E1786-2** | 40 | 0.959 | 0.006 |
|  | 60 | 0.919 | 0.008 |
|  | 80 | 0.903 | 0.010 |
| **E1786-3** | 40 | 0.952 | 0.006 |
|  | 60 | 0.910 | 0.008 |
|  | 80 | 0.900 | 0.011 |
| **E1786-4** | 40 | 0.942 | 0.008 |
|  | 60 | 0.902 | 0.010 |
|  | 80 | 0.897 | 0.014 |
| **E1786-5** | 40 | 0.949 | 0.006 |
|  | 60 | 0.903 | 0.007 |
|  | 80 | 0.892 | 0.013 |
| **E1786-6** | 40 | 0.936 | 0.015 |
|  | 60 | 0.892 | 0.012 |
|  | 80 | 0.884 | 0.018 |
| **E1786-7** | 40 | 0.895 | 0.007 |
|  | 60 | 0.852 | 0.006 |
|  | 80 | 0.843 | 0.009 |
| **E1786-9** | 40 | 0.928 | 0.008 |
|  | 60 | 0.884 | 0.007 |
|  | 80 | 0.867 | 0.012 |
| **E1786-10** | 40 | 0.964 | 0.007 |
|  | 60 | 0.918 | 0.009 |
|  | 80 | 0.912 | 0.037 |
| **E1786-11** | 40 | 0.941 | 0.011 |
|  | 60 | 0.881 | 0.017 |
|  | 80 | 0.887 | 0.010 |
| **E1787-1** | 40 | 2.115 | 0.009 |
|  | 60 | 2.041 | 0.010 |
|  | 80 | 2.021 | 0.017 |
| **E1787-2** | 40 | 2.096 | 0.010 |
|  | 60 | 2.016 | 0.011 |
|  | 80 | 1.999 | 0.014 |
| **E1787-3** | 40 | 2.031 | 0.012 |
|  | 60 | 1.961 | 0.008 |
|  | 80 | 1.942 | 0.014 |
| **E1787-4** | 40 | 2.051 | 0.007 |
|  | 60 | 1.973 | 0.007 |
|  | 80 | 1.955 | 0.010 |
| **E1787-5** | 40 | 2.082 | 0.007 |
|  | 60 | 2.009 | 0.009 |
|  | 80 | 1.992 | 0.009 |
| **E1787-6** | 40 | 2.051 | 0.007 |
|  | 60 | 1.972 | 0.011 |
|  | 80 | 1.954 | 0.013 |
| **E1787-7** | 40 | 2.114 | 0.009 |
|  | 60 | 2.043 | 0.007 |
|  | 80 | 2.007 | 0.023 |
| **E1787-8** | 40 | 2.118 | 0.009 |
|  | 60 | 2.050 | 0.015 |
|  | 80 | 2.033 | 0.014 |
| **E1787-9** | 40 | 2.085 | 0.008 |
|  | 60 | 2.010 | 0.005 |
|  | 80 | 1.991 | 0.008 |
| **E1787-10** | 40 | 2.064 | 0.007 |
|  | 60 | 1.993 | 0.007 |
|  | 80 | 1.969 | 0.018 |
| **E1788-1** | 40 | 1.073 | 0.006 |
|  | 60 | 1.028 | 0.017 |
|  | 80 | 1.028 | 0.050 |
| **E1788-2** | 40 | 2.136 | 0.008 |
|  | 60 | 2.059 | 0.011 |
|  | 80 | 2.024 | 0.028 |