



Shear wave elastography using sound touch elastography and supersonic shear imaging for liver measurements: a comparative study

Hongjin Xiang, Wenwu Ling, Lin Ma, Lulu Yang, Tang Lin, Yan Luo

Department of Ultrasound, West China Hospital of Sichuan University, Chengdu, China

Contributions: (I) Conception and design: Y Luo, H Xiang; (II) Administrative support: Y Luo; (III) Provision of study materials or patients: W Ling, L Ma; (IV) Collection and assembly of data: T Lin; (V) Data analysis and interpretation: H Xiang, L Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yan Luo. Department of Ultrasound, West China Hospital of Sichuan University, 37 Guoxue Lane, Chengdu 610041, China. Email: yanluo@scu.edu.cn.

Background: Measurements of liver stiffness obtained with 2-dimensional shear wave elastography (2D-SWE) have been widely used to clinically assess liver fibrosis. However, differences between different 2D-SWE systems can lead to confusion when interpreting measurements. This study investigated the variability between a recently released sound touch elastography (STE) system and a supersonic shear imaging (SSI) system and assessed the degree of intersystem discrepancy using the different liver stiffness value (LSV) thresholds recommended by the Society of Radiologists in Ultrasound (SRU) for assessing liver fibrosis.

Methods: A total of 4,152 patients who had undergone STE and SSI on the same day were enrolled in this retrospective study. First, intrasystem agreement for STE and SSI was assessed. Then, intraclass correlation coefficients (ICCs) and the Bland-Altman method were used to assess intersystem variability for all cases, classified according to the thresholds recommended by the SRU. The effects of age, gender, and body-mass index (BMI) were evaluated using multivariate linear regression analysis and attributive intervals were computed for STE and SSI at each of the different thresholds.

Results: The ICCs for STE and SSI intrasystem agreement were 0.94 [95% confidence interval (CI): 0.937–0.943; $P < 0.001$] and 0.984 (95% CI: 0.984–0.985; $P < 0.001$), respectively. The 95% limit of agreement (LOA) for all cases ranged from –6.96 to 7.44 kPa. The 95% LOA increased as the threshold values rose, and intersystem variability was obvious, even at the smallest threshold (the 95% LOA at values ≤ 5 kPa was –0.85 to 2.08 kPa, while that at values > 17 kPa was –20.81 to 14.71 kPa). The adjusted R^2 for age, gender, and BMI was only 0.018 (all P value < 0.05).

Conclusions: There was clear variability between STE and SSI, in contrast with some previous studies with small sample sizes, and consistent with others. Intersystem variability increased with the elevation of the LSV thresholds recommended by the SRU. Gender and BMI had little effect on intersystem variability. Future research could compare STE and SSI in different liver diseases, assessing the feasibility of the SRU-recommended thresholds in proven pathologies and evaluating the test-retest repeatability.

Keywords: Sound touch elastography (STE); supersonic shear imaging (SSI); intersystem variability; rule of four

Submitted Sep 01, 2021. Accepted for publication Jan 19, 2022.

doi: 10.21037/qims-21-863

View this article at: <https://dx.doi.org/10.21037/qims-21-863>

Introduction

Cirrhosis is one of the top 20 causes of disability-adjusted life years and years of life lost, accounting for 1.6% and 2.1% of the worldwide burden, respectively (1). Hepatic fibrosis is a common pathway to liver cirrhosis (2). Therefore, accurate assessment of the stage of hepatic fibrosis is crucial, especially as treatments aimed at reversing fibrosis are being developed (3). Although liver biopsy is the reference standard for evaluating hepatic fibrosis, its invasiveness, sampling error rate, and subjective interpretations mean that it is unsuitable as a routine screening tool and longitudinal monitoring tool for hepatic fibrosis (4).

With the advantages of safety, good reproducibility, and good applicability, 2-dimensional shear wave elastography (2D-SWE), an acoustic radiation force impulse (ARFI) technique, is recommended for assessing hepatic fibrosis in routine clinical practice (5,6). It is the combination of a radiation force induced in tissues by focused ultrasonic beams and a very high frame rate ultrasound imaging sequence capable of measuring the speed of resulting shear waves (5). The shear wave speed (SWS), expressed in m/s, can be converted to Young's modulus, expressed in kPa, and physical measurement of tissue stiffness can be formulated as follows:

$$\text{Young's modulus} = 3\rho\text{SWS}^2 \quad [1]$$

where ρ is the tissue density assumed to be 1 g/mL. Liver stiffness value (LSV), expressed in kPa, can thus be used to assess the degree of hepatic fibrosis with 2D-SWE technology (6,7).

There are various 2D-SWE systems available from different vendors (5,8). It is generally recommended that LSV thresholds for fibrosis staging and the follow-up of liver stiffness should be system-specific (6-8). Given the heavy burden of liver disease (1), the wide variety of thresholds recommended for different 2D-SWE systems can present considerable confusion to hepatologists, especially when the system used for initial screening is not available for follow-up. For the convenience of clinical application, the consensus of the Society of Radiologists in Ultrasound (SRU) (8) has recommended a vendor-neutral "rule of four", providing 5 thresholds (≤ 5 , < 9 , $9-13$, > 13 , and > 17 kPa) for the interpretation of LSV measured by different systems using the ARFI technique. The SRU also recommends that values < 7 kPa suggest a high probability that significant fibrosis can be ruled out. However, the feasibility of these SRU-recommended thresholds needs to be validated.

The differences between various system measurements increase as liver stiffness increases (7-9). Ferraioli *et al.* reported that the mean difference in LSVs between 2D-SWE systems was 0.55 kPa in patients with little or no fibrosis (< 15 kPa), increasing to 2.43 kPa in patients at a more advanced disease stage (9). In contrast, Ryu *et al.* reported poor intersystem agreement between 2D-SWE systems (10). From these studies, it is clear that there is no definitive agreement about the degree of intersystem agreement or variability between 2D-SWE systems. If intersystem variability were minimal at relatively low ranges of LSV, a reference value representing normal stiffness might be applied to different 2D-SWE systems, facilitating primary screening for liver fibrosis in routine clinical practice.

The SuperSonic Imagine (Aixplorer, Aix-en-Provence, France) supersonic shear imaging (SSI) system has been used extensively for assessing hepatic fibrosis (7), whereas the sound touch elastography (STE) system (Resona 7, Mindray, Shenzhen, China) is a newly available 2D-SWE system that has been tested in only a few studies (11,12). To date, no study with a large sample size has assessed the variability of measurements between STE and SSI with different LSV thresholds.

This study, involving a large sample size, investigated the variability between a newly released STE 2D-SWE system and a well-validated SSI system and assessed the degree of variability at different SRU-recommended LSV thresholds. We present the following article in accordance with the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-21-863/rc>).

Methods

Study participants

This retrospective study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). It was approved by the Ethics Committee of the West China Hospital, and written informed consent was exempted on account of its retrospective nature. Between May 2019 and April 2020, 4,229 consecutive patients visited the Ultrasound Department of the West China Hospital with previously undiagnosed liver diseases. These patients underwent liver stiffness measurement (LSM) of the background liver parenchyma using STE and SSI on the same day. For these patients, the ratio of the interquartile

range of LSVs to the median (IQR/M) for SSI was calculated artificially, while an IQR/M $\leq 30\%$ had already been used for quality control upon the final acquisition of the STE data. The exclusion criteria for this study were as follows: (I) patients who had undocumented demographic information (age, gender, height, and weight) on the day of their LSM (n=12); (II) patients who had an inadequate LSM with either system (i.e., IQR/M more than 30%, n=65). Thus, 4,152 cases were finally enrolled in this study.

LSM

Patients were instructed to fast for at least 8 hours and rest for at least 10 minutes before examination. Then, either of 2 ultrasound physicians both with more than 8 years of experience in hepatic ultrasonography performed the examinations, completed elastography with the 2 systems for an individual patient. The LSM protocol followed the latest European Federation for Ultrasound in Medicine and Biology (EFSUMB) guidelines (7) as previously reported (13). A B-mode intercostal liver scan was first performed, clearly visualizing the liver capsule, whereby patients were instructed to lie in a supine position with the right arm in maximal extension and take a normal breath, holding it for few seconds at mid expiration, and measurements of the right liver lobe were obtained through the appropriate intercostal location.

For each system, an elasticity acquisition box of 4 cm \times 3 cm was placed at the desired location, 1–6 cm beneath the liver capsule in the middle line of the elastogram, free from large vessels, the gallbladder, and artifacts. In the relatively homogeneous areas of the acquisition box, a circular analysis box with a diameter of 15 mm was placed to measure stiffness. The average value (expressed as Young's modulus in kPa) obtained by the analysis box was recorded as the LSV of each acquisition. A valid measurement was acquired when it met the quality control requirements of each system. The sonographers performed 3 consecutive acquisitions for both STE and SSI, and the median value of the 3 acquisitions was taken as the representative LSV for each case.

Statistical analysis

Continuous values were presented as means and standard deviations, and categorical results were summarized as counts and percentages. The IQR/M was calculated for all examinations. The median-centered coefficient of variation (COV), which is the result of expressing the root mean

squares of deviation from the median as a percentage of the median, was calculated for each system.

Intraclass correlation coefficients (ICCs) from a one-way random model were used to assess intrasystem agreement for every examination. Once the median value of the 3 acquisitions was recorded as the representative LSV of each case, ICCs from a 2-way mixed model were calculated, and Bland-Altman and Bradley-Blackwood tests were performed to assess intersystem agreement between STE and SSI for the following groups: (I) all cases and (II) cases whose average of LSV measured by STE (LSV_{STE}) and LSV measured by SSI (LSV_{SSI}) fell within the LSV thresholds recommended by the SRU. Spearman correlation coefficients between the 2 systems were also calculated. If age, gender, and body-mass index (BMI) exhibited no collinearity, multivariate linear regression analysis was performed to investigate whether these 3 factors affected the absolute value of the difference between the systems ($LSV_{STE} - LSV_{SSI}$). To assess the point at which the LSV_{STE} and LSV_{SSI} for individual patients fell within the same SRU-recommended thresholds, sets were computed and compared for each threshold and system.

The Bland-Altman method was used to plot the difference (y-axis, $LSV_{STE} - LSV_{SSI}$) against the average [x-axis, $(LSV_{STE} + LSV_{SSI})/2$] for parallel statistics (14). Based on the regression of the difference on the average, a Bradley-Blackwood test assessed whether a null slope and a null intercept simultaneously existed, and if they did, the mean difference [i.e., the mean value of $(LSV_{STE} - LSV_{SSI})$] was 0 (15). The 95% limit of agreement (LOA) of the difference was calculated as the mean $\pm 1.96 \times$ standard deviations of $(LSV_{STE} - LSV_{SSI})$. As mentioned above, higher variability was found between the 2 systems in stiffer livers. To reduce the influence of individual variations contributed by the absolute value of LSV, the normalized difference (% difference) was defined as $(LSV_{STE} - LSV_{SSI}) / [(LSV_{STE} + LSV_{SSI})/2]$, and the 95% LOA of the normalized difference was also calculated as the mean $\pm 1.96 \times$ standard deviation of the % difference.

The Bland-Altman method (14) suggests a good agreement when a narrow 95% LOA shows, and a minority of cases fall outside the 95% LOA, based on a nonsignificant Bradley-Blackwood test. Agreement evaluated by ICCs was classified as excellent (ICC >0.75), good (ICC =0.60–0.75), fair (ICC =0.40–0.59), or poor (ICC ≤ 0.40) if the ICC was statistically significant (16). The Bonferroni method ($\alpha=0.05/k$) was used for the analyses among SRU-recommended thresholds, and k was the number of thresholds (k=6),

Table 1 Patient age, gender, and BMI

Parameter	Mean \pm standard deviation	Distribution	
		Classification	Percentage
Age (years)	50.2 \pm 12.63	11–14	0.12% (5/4,152)
		15–17	0.48% (19/4,152)
		18–65	88.0% (3,652/4,152)
		66–90	11.4% (476/4,152)
Gender	–	Female	39.8% (1,651/4,152)
		Male	60.2% (2,501/4,152)
BMI (kg/m ²)	23.1 \pm 3.1	<25.0	74.3% (3,083/4,152)
		\geq 25.0	25.7% (1,069/4,152)

Numbers in parentheses are raw data. BMI, body mass index.

Table 2 Summary of liver stiffness measurements with STE and SSI

System	LSVs (kPa) [†]	IQR/M (%) [†]	Median-centered COV (%)	ICC for intrasystem agreement [‡]
STE (Mindray)	9.93 \pm 5.21	8.41 \pm 5.36	65.2	0.94 (0.937–0.943)
SSI (Aixplorer)	9.69 \pm 6.18	5.50 \pm 3.62	86.5	0.984 (0.984–0.985)

[†], data are expressed as means \pm standard deviations; [‡], 95% confidence intervals of ICC are in parentheses. STE, sound touch elastography; SSI, supersonic shear imaging; LSV, liver stiffness value; IQR, interquartile range; M, median; COV, coefficient of variation; ICC, intraclass correlation coefficient.

while the analyses for all cases adopted $\alpha=0.05$. Statistical significance was indicated when P was less than α indicated statistical significance, and 95% confidence intervals (CIs) were presented, if applicable. A Bland-Altman plot, scatter plot between LSV_{STE} and LSV_{SSI} from the individual patient, and box plot between the absolute value of the difference between STE and SSI ($|LSV_{STE} - LSV_{SSI}|$) and age, gender, and BMI were drawn up using GraphPad Prism 8 Software (GraphPad Software, Inc., San Diego, CA, USA). Statistical tests were performed using SPSS version 22.0 for Windows (SPSS Inc., Chicago, IL, USA).

Results

Baseline characteristics

There were 4,152 participants in total, including 1,650 (39.8%) females and 2,501 (60.2%) males. The mean age of the participants was 50.20 \pm 12.63 years, with pediatric cases (<18 years old) accounting for 0.6% of the study population. The mean height, weight, and BMI of the patients were 164 \pm 8 cm, 62.1 \pm 10.8 kg, and 23.1 \pm 3.1 kg/m², respectively (Table 1). The 2 physicians with more than 8

years of experience in hepatic ultrasonography completed 41.5% and 58.5% of the whole cohort, respectively. The LSVs measured by STE and SSI are summarized in Table 2. The IQR/M was 8.41% \pm 5.36% for STE and 5.50% \pm 3.62% for SSI. The ICCs for STE and SSI intrasystem agreement were 0.94 (95% CI: 0.937–0.943, $P<0.001$) and 0.984 (95% CI: 0.984–0.985, $P<0.001$), respectively, and the median-centered COV for SSI was slightly higher than for STE (86.5% vs. 65.2%), indicating excellent intrasystem agreement in both systems, with a higher intrasystem agreement in STE.

Comparison of STE and SSI for all cases

To illustrate the relationship between the 2 systems, LSV_{STE} was plotted against LSV_{SSI} (Figure 1A), and most points made up an area. The border of this area was outlined artificially using 2 lines (upper, $y=x+6.2$ in green, and lower, $y=0.5x+0.9$ in red). If x (LSV_{SSI}) was fixed, y (LSV_{STE}) received a value range, suggesting the differences between STE and SSI increased as liver stiffness increased. Although a strong Spearman correlation coefficient (0.822, $P<0.001$)

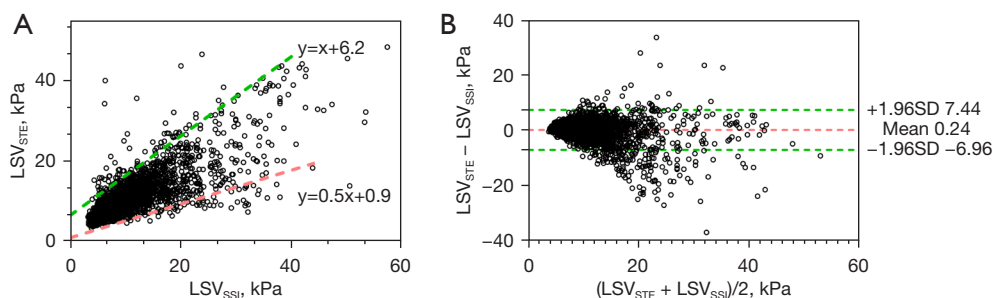


Figure 1 Comparison of STE and SSI for all cases. (A) scatter plot between LSV_{STE} and LSV_{SSI} , where most points made up an area. The border of this area was outlined artificially using 2 lines (upper, $y=x+6.2$ in green, and lower, $y=0.5x+0.9$ in red); (B) Bland-Altman plot of the overall subjects, where the green dashed line and the red dashed line represent 95% LOA and mean difference, respectively. LSV_{STE} and LSV_{SSI} are the LSVs measured for each participant by STE and SSI, respectively. LSV, liver stiffness value; STE, sound touch elastography; SSI, supersonic shear imaging; 95% LOA, 95% limit of agreement; SD, standard deviation.

Table 3 Summary of ICCs and Spearman correlation coefficients between STE and SSI

Threshold (kPa)	N	ρ	ICC for intersystem agreement [†]
Totality	4,152	0.822*	0.791 (0.781 to 0.804)**
$\leq 5^\ddagger$	185	-0.267*	-0.153 (-0.334 to 0.062)**
$< 9^\ddagger$	2,473	0.541*	0.402 (0.218 to 0.537)**
9–13 [†]	897	-0.152*	-0.263 (-0.329 to -0.196)
$> 13^\ddagger$	782	0.268*	0.497 (0.426 to 0.560)**
$> 17^\ddagger$	356	0.124	0.320 (0.208 to 0.420)**
$< 7^\ddagger$	1,521	0.285*	0.174 (0.015 to 0.311)**

[†], thresholds recommended by the SRU. The first five thresholds refer to “rule of four”, and < 7 kPa could help rule out significant fibrosis. [‡], 95% confidence intervals of ICC are in parentheses; *, $P < 0.001$ for Spearman correlation coefficient; **, $P < 0.001$ for ICC. ICC, intraclass correlation coefficient; STE, sound touch elastography; SSI, supersonic shear imaging; ρ , Spearman correlation coefficient; SRU, Society of Radiologists in Ultrasound.

and an excellent ICC (0.791, $P < 0.001$) was found between STE and SSI (Table 3), the significant Bradley-Blackwood test ($P < 0.001$; Table 4) and 95% LOA (-6.96 to 7.44 kPa; Table 4, Figure 1B) suggested a lack of agreement between both systems.

Comparison of STE and SSI at different threshold values

Statistically significant ICCs ($P < 0.001$) for each threshold value were graded as fair or poor (Table 3). One exception was the ICC for values between 9 and 13 kPa, which was

not statistically significant ($P < 0.99$). The results of the Bradley-Blackwood tests showed a paucity of agreement at each threshold ($P < 0.001$) (Table 4). The range of 95% LOA of difference and % difference indicated the intersystem variability increased as the threshold values increased (Figure 2). At the lowest threshold (≤ 5 kPa), 95% LOA of difference and % difference were -0.85 to 2.08 kPa and -18.0% to 44.4%, respectively, suggesting there was also variability between STE and SSI at this value.

Relationship between different LSVs and age, gender, and BMI

Multivariate linear regression analysis showed that the adjusted R^2 was only 0.018, despite P values for the slope of age, gender, and BMI being < 0.001 , 0.048, and < 0.001 , respectively. Our analysis also showed that age, gender, and BMI hardly affected the absolute value of the difference between the systems ($LSV_{STE} - LSV_{SSI}$) (Figure 3).

SRU-recommended threshold intervals with STE and SSI

Cases were assigned to an SRU-recommended threshold interval for each system, according to the LSVs measured by STE and SSI (Figure 4). The STE sets (green) included cases whose LSV_{STE} was within the limits of each SRU-recommended threshold, and the SSI sets (red) were similarly defined according to cases’ LSV_{SSI} . The intersections (blue) show cases whose LSV_{STE} and LSV_{SSI} fell within the same SRU-recommended threshold. There were 1,107 cases in the intersection of the < 7 kPa threshold sets, accounting for 87% (1,107 of 1,276) of the STE set

Table 4 Summary of the Bland-Altman and Bradley-Blackwood tests results between STE and SSI

Threshold (kPa)	Bland-Altman			Bradley-Blackwood	
	Mean difference (kPa)	95% LOA (kPa)	Cases exceeding 95% LOA [‡]	F	P value*
Totality	0.24	-6.96 to 7.44	5.1% (211/4,152)	184.424	<0.001
≤5 [†]	0.62	-0.85 to 2.08	4.3% (8/185)	63.307	<0.001
<9 [†]	0.78	-2.00 to 3.57	6.2% (154/2,473)	376.312	<0.001
9–13 [†]	0.56	-5.23 to 6.34	6.2% (56/897)	18.122	<0.001
>13 [†]	-1.83	-15.70 to 12.03	5.9% (46/782)	33.137	<0.001
>17 [†]	-3.05	-20.81 to 14.71	5.9% (21/356)	20.162	<0.001
<7 [†]	0.81	-1.28 to 2.90	5.6% (85/1,521)	441.596	<0.001

[†], thresholds recommended by the SRU. The first five thresholds refer to “rule of four”, and <7 kPa could help rule out significant fibrosis; [‡], numbers in parentheses are raw data; *, all of the Bradley-Blackwood tests were significant (P<0.001). STE, sound touch elastography; SSI, supersonic shear imaging; 95% LOA, 95% limit of agreement; SRU, Society of Radiologists in Ultrasound.

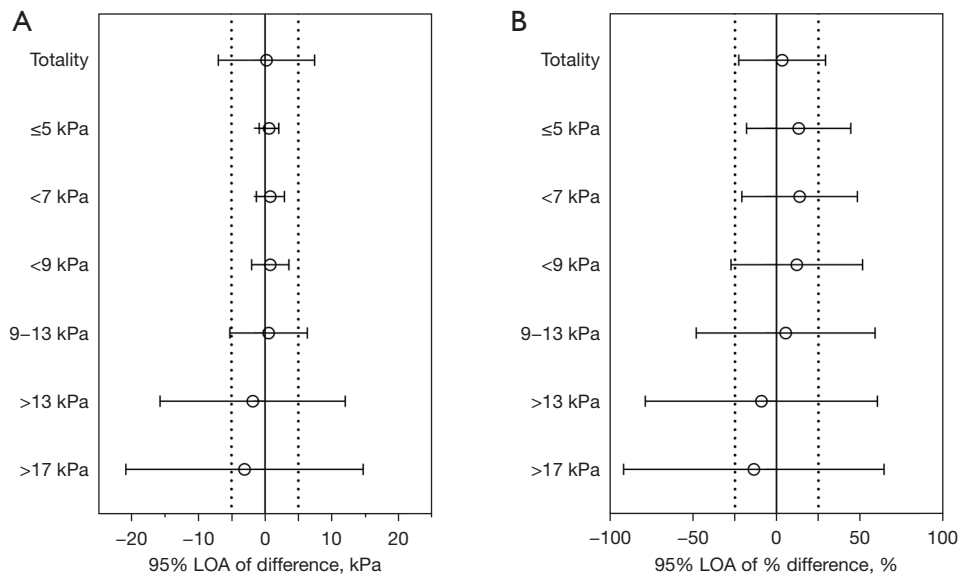


Figure 2 95% LOAs of difference (A) and % difference (B) for various SRU-recommended thresholds. Difference = $LSV_{STE} - LSV_{SSI}$, and % difference = $(LSV_{STE} - LSV_{SSI}) / [(LSV_{STE} + LSV_{SSI}) / 2]$. 95% LOA, 95% limit of agreement; SRU, Society of Radiologists in Ultrasound; LSV_{STE} , liver stiffness values measured by STE; LSV_{SSI} , liver stiffness values measured by SSI; STE, sound touch elastography; SSI, supersonic shear imaging.

and 65% (1,107 of 1,767) of the SSI set. The intersection of the >13 kPa sets included 71% (221 of 313) of the STE set and 68% (221 of 417) of the SSI set.

Discussion

The different cut-off values for various 2D-SWE systems can present a substantial burden of confusion

for hepatologists in clinical practice. The SRU has recommended a vendor-neutral “rule of four”, which gives 5 LSV thresholds for interpreting liver stiffness measured by different systems (8). Previous studies have compared different 2D-SWE systems with small sample sizes, but intersystem agreement has remained unclear (9,10,17-19). Of several available 2D-SWE systems, the SSI was recognized by the latest

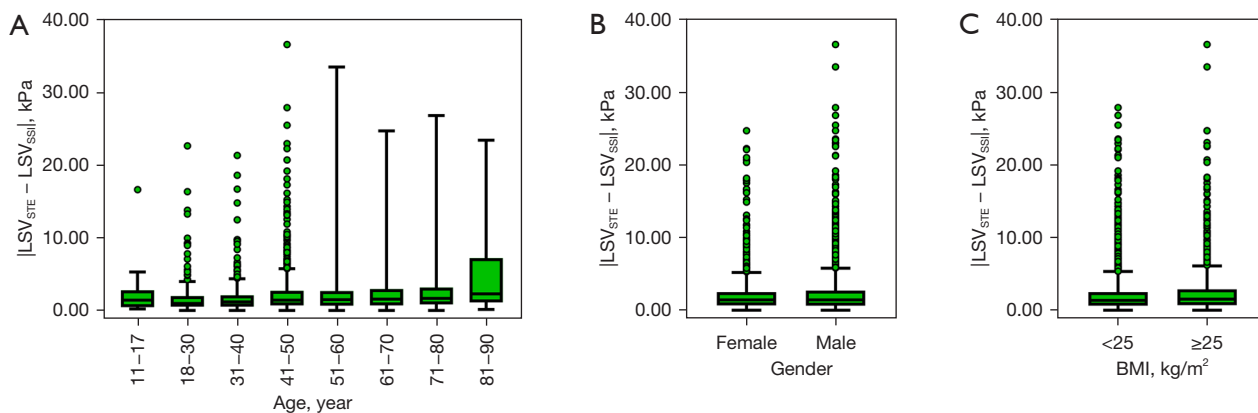


Figure 3 Box plots showing the absolute value of the difference between the systems ($LSV_{STE} - LSV_{SSI}$) among different groups of age (A), gender (B), and BMI (C). Boxes represent the interquartile range (25% to 75%), lines within the boxes represent the median values. The vertical line extends from minimum to maximum value, excluding outliers. LSV_{STE} , liver stiffness values measured by STE; LSV_{SSI} , liver stiffness values measured by SSI; STE, sound touch elastography; SSI, supersonic shear imaging; BMI, body mass index.

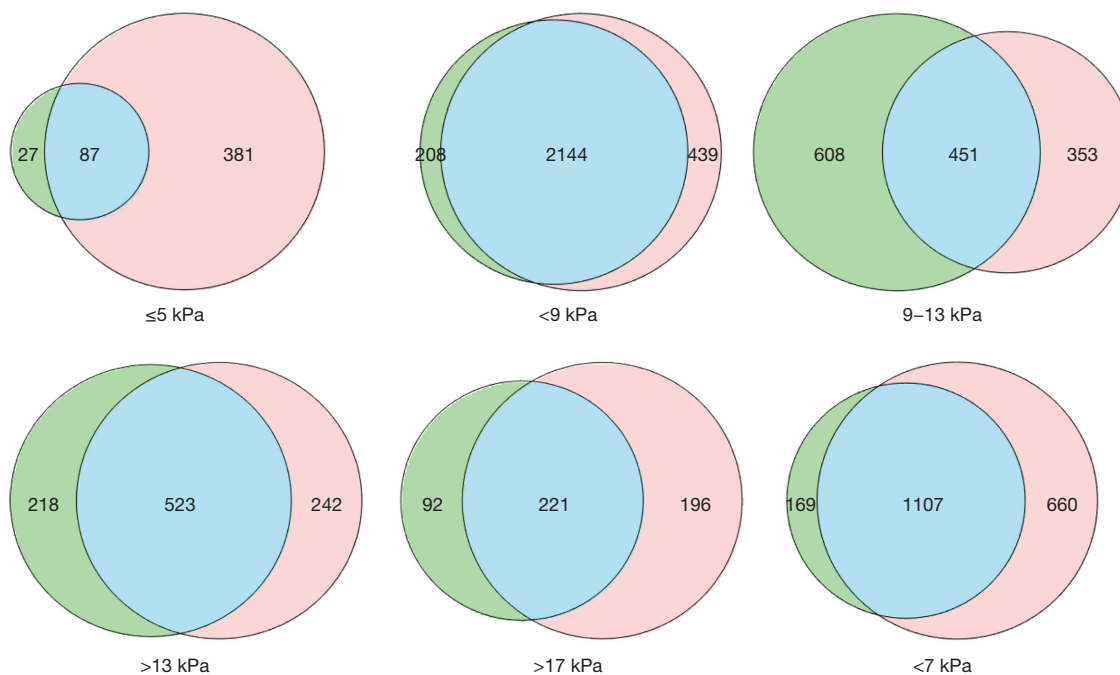


Figure 4 STE and SSI value sets. The green circles represent the STE sets, composed of cases whose LSV_{STE} fell within the limits of each threshold. The red circles represent the SSI set, made up of cases whose LSV_{SSI} fell within the limits of the threshold. The blue areas represent intersections between the sets, where the LSV_{STE} and LSV_{SSI} fell within the same SRU-recommended threshold. STE, sound touch elastography; SSI, supersonic imaging; LSV_{STE} , liver stiffness values measured by STE; LSV_{SSI} , liver stiffness values measured by SSI.

EFSUMB guidelines (7), whereas the recently released STE technique (Resona 7, Mindray, Shenzhen, China) has not yet been widely validated. Therefore, we evaluated intersystem variability between these 2 systems using a large

sample size.

In this study, excellent intrasystem agreement was observed in each system with an $IQR/M \leq 30\%$, consistent with previous studies (7,11,17,20). Meanwhile, the overall

ICC for intersystem agreement between STE and SSI also indicated excellent agreement. In contrast, Bland-Altman analysis showed low intersystem agreement. The discrepancy between the ICC and Bland-Altman analysis results was presumably due to intercase variations, to which ICC is susceptible, but Bland-Altman is not (16,21). The wide range of LSVs caused an overestimation in the ICCs. Therefore, this study placed greater emphasis on the Bland-Altman analysis.

The 95% LOA for STE and SSI (-6.96 to 7.44 kPa) derived from the Bland-Altman analysis suggested obvious intersystem variability between STE and SSI, analogous to the study reported by Ryu *et al.* (10), where significant intersystem variability was found between 2D-SWE systems, even when they were from the same manufacturer (LOGIQ E9 and LOGIQ S8, GE Healthcare, Chicago, IL, USA). However, good intersystem agreement between 2D-SWE devices has been found in some studies, including one performing the pair comparison of 6 devices and including SSI and STE techniques, with a sample of 24 (18), another comparing the Aixplorer and Aplio 500 with a sample size of 26 (9), and another comparing the Aixplorer and Aplio 500 with a sample size of 20 (22). It is presumed that large samples have greater heterogeneity than small, causing intersystem agreement to decrease.

As the SRU-recommended threshold values increased, so too did the intersystem variability between STE and SSI, consistent with a previous study in which the difference between STE and SSI increased with the development of the fibrosis stage (20). At the smallest value (≤ 5 kPa), the mean difference was 0.62 kPa with 95% LOA of -0.85 to 2.08 kPa and a normalized difference of -18.0% to 44.4% , suggesting caution should be used when using a specific reference value, even for people with low liver stiffness.

In this study, the difference between STE and SSI were hardly affected by age, gender, and BMI, because a minimal adjusted R^2 (0.018) of the multivariate linear regression was observed. The number of pediatric cases ($n=24$) in this study might have been a statistical limitation as 1 study has reported that LSVs are lower in pediatric patients than in adults (23). As stated in the SRU consensus, the source of variability between SWE systems should be studied in the future (8).

The SRU consensus recommends a neutral-vendor “rule of four”, which includes 5 LSV thresholds to facilitate the interpretation of LSVs from different ARFI systems. To assess the feasibility of this “rule of four”, it was important to assess LSVs that were at the same threshold

when measured by different systems rather than their interchangeability. In this study, there were relatively large intersections between the STE sets and SSI sets at SRU-recommended thresholds, except for the lowest threshold of ≤ 5 kPa, where the intersection only included 19% of the SSI set. Further studies with pathologic results are necessary to validate the feasibility of the SRU-recommended “rule of four”.

The strengths of this study were as follows. First, the sample size of this study was 4,152 cases and included a wide range of ages, BMIs, etiologies, and LSVs. Such a sample represents a fair simulation of the first consultations with patients with unknown liver diseases and has not been found in previous studies. Second, we assessed the degree of variability between 2 different 2D-SWE systems with different LSVs and concentrated on the smallest interval (≤ 5 kPa) to investigate whether a single reference value could be used for different 2D-SWE systems in patients with low liver stiffness (8). Third, we compared a newer STE system with an SSI system for which reliability has been accredited in the EFSUMB guidelines (7). Last, ICC and Bland-Altman analyses were simultaneously performed to evaluate intersystem agreement, allowing cross-validation for the large sample size.

This study also had limitations. First, it was retrospective without pathologic results, hindering the comparison between STE and SSI according to different etiologies. It has been demonstrated that cut-off values for both systems vary with the etiology (11,12,24). However, the combination of various etiologies might be a fair simulation of the initial consultations with patients with unknown liver disease. In China, a hepatologist would instruct such patients to undergo various serology and imaging examinations, including ultrasound elastography to evaluate liver stiffness, and the interpretation of an LSV would usually be given by a sonographer unaware of the results of other examinations. Moreover, when the comparison was performed for values ≤ 5 kPa, intersystem variability derived from different etiologies might be disregarded. Second, the representative LSVs were obtained from 3 measurements. However, we believe that the number of measurements has little impact on the results of this study since the IQR/M $\leq 30\%$ was used for quality control, and 3 acquisitions met the lower limit of the number of measurements recommended in the EFSUMB guidelines (7,25). Third, some operator-related factors, including test-retest repeatability of the same observer for each system, interobserver variability, and the different locations of ROIs (measurement depth),

were not considered in this study. Despite Gilligan *et al.* finding excellent test-retest repeatability of an experienced observer for the 2 systems (the ICCs for STE and SSI were 0.95 and 0.94, respectively), the test-retest repeatability needs to be validated in the future (18). Many studies (11,19,20,26,27) have reported excellent agreement between expert operators, so the result of the present study may be minimally affected by interobserver variability. Previous studies have demonstrated a system-dependent performance at different measurement depths, and different systems have shown good agreement at a depth of 5 cm (17,28). A measurement depth of 4–5 cm will be used for future research, as recommended by the EFSUMB guidelines (7). Lastly, this study did not standardize some patient-related factors, including different fasting and rest times before LSM and different breath cooperation among patients.

Conclusions

Based on the large sample size, significant variability was observed between STE and SSI, in contrast to some previous, small sample-size studies. Intersystem variability increased as the LSV thresholds recommended by the SRU consensus increased. The variability between STE and SSI cannot be ignored, even in the lowest threshold (≤ 5 kPa). Intersystem variability between STE and SSI was little affected by gender or BMI. Future investigation should include a comparison of the 2 systems with different etiologies, based on the standardization of at least 5 measurements, ROI placement, and breath maneuvers, to assess the feasibility of the SRU-recommended “rule of four” with proven pathologies and investigate reproducibility between 2 systems from the same vendor and the test-retest repeatability of the same operator on the same machine.

Acknowledgments

Funding: This research was partially supported by the National Natural Science Foundation of China (No. 82071940).

Footnote

Reporting Checklist: The authors have completed the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-21-863/rc>

[com/article/view/10.21037/qims-21-863/rc](https://qims.amegroups.com/article/view/10.21037/qims-21-863/rc)

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure forms (available at <https://qims.amegroups.com/article/view/10.21037/qims-21-863/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This retrospective study was approved by the Ethics Committee of the West China Hospital, and written informed consent was exempted.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Asrani SK, Devarbhavi H, Eaton J, Kamath PS. Burden of liver diseases in the world. *J Hepatol* 2019;70:151-71.
2. Shipley LC, Axley PD, Singal A. Liver fibrosis: a clinical update. *Eur Med J Hepatol* 2019;7:105-17.
3. Eslam M, Sarin SK, Wong VW, Fan JG, Kawaguchi T, Ahn SH, et al. The Asian Pacific Association for the Study of the Liver clinical practice guidelines for the diagnosis and management of metabolic associated fatty liver disease. *Hepatol Int* 2020;14:889-919.
4. Tapper EB, Lok AS. Use of liver imaging and biopsy in clinical practice. *N Engl J Med* 2017;377:756-68.
5. Asociacion Latinoamericana para el Estudio del Hgado. EASL-ALEH Clinical Practice Guidelines: non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol* 2015;63:237-64.
6. Ferraioli G, Wong VW, Castera L, Berzigotti A, Sporea I, Dietrich CF, Choi BI, Wilson SR, Kudo M, Barr RG. Liver ultrasound elastography: an update to the world federation for ultrasound in medicine and biology guidelines and recommendations. *Ultrasound Med Biol*

- 2018;44:2419-40.
7. Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, Cosgrove D, Ferraioli G, Friedrich-Rust M, Gilja OH, Goertz RS, Karlas T, de Knecht R, de Ledinghen V, Piscaglia F, Procopet B, Saftoiu A, Sidhu PS, Sporea I, Thiele M. EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version). *Ultraschall Med* 2017;38:e16-47.
 8. Barr RG, Wilson SR, Rubens D, Garcia-Tsao G, Ferraioli G. Update to the society of radiologists in ultrasound liver elastography consensus statement. *Radiology* 2020;296:263-74.
 9. Ferraioli G, De Silvestri A, Lissandrin R, Maiocchi L, Tinelli C, Filice C, Barr RG. Evaluation of inter-system variability in liver stiffness measurements. *Ultraschall Med* 2019;40:64-75.
 10. Ryu H, Ahn SJ, Yoon JH, Lee JM. Reproducibility of liver stiffness measurements made with two different 2-dimensional shear wave elastography systems using the comb-push technique. *Ultrasonography* 2019;38:246-54.
 11. Wang J, Wu M, Linghu R, Chang J, Wu M, Feng C, et al. Usefulness of new shear wave elastography technique for noninvasive assessment of liver fibrosis in patients with chronic hepatitis B: a prospective multicenter study. *Ultraschall Med* 2021. [Epub ahead of print]. doi: 10.1055/a-1376-6734.
 12. Yang L, Ling W, He D, Lu C, Ma L, Tang L, Luo Y, Chen S. Shear wave-based sound touch elastography in liver fibrosis assessment for patients with autoimmune liver diseases. *Quant Imaging Med Surg* 2021;11:1532-42.
 13. Yang L, Li J, Ma L, Xiang H, He D, Lu C, Tang L, Luo Y, Chen S. Noninvasive assessment of liver fibrosis in chronic hepatitis B carriers with sound touch elastography: study of surgical pathology specimens. *Expert Rev Med Devices* 2020;17:845-53.
 14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
 15. Bradley EL, Blackwood LG. Comparing paired data: a simultaneous test for means and variances. *Am Stat* 1989;43:234-5.
 16. Fleiss JL, Shrout PE. Approximate interval estimation for a certain intraclass correlation coefficient. *Psychometrika* 1978;43:259-62.
 17. Gress VS, Glawion EN, Schmidberger J, Kratzer W. Comparison of liver shear wave elastography measurements using Siemens Acuson S3000, GE LOGIQ E9, Philips EPIQ7 and Toshiba Aplio 500 (Software Versions 5.0 and 6.0) in Healthy Volunteers. *Ultraschall Med* 2019;40:504-12.
 18. Gilligan LA, Trout AT, Bennett P, Dillman JR. Repeatability and agreement of shear wave speed measurements in phantoms and human livers across 6 ultrasound 2-Dimensional shear wave elastography systems. *Invest Radiol* 2020;55:191-9.
 19. Hwang J, Yoon HM, Jung AY, Lee JS, Cho YA. Comparison of 2-dimensional shear wave elastographic measurements using ElastQ imaging and supersonic shear imaging: phantom study and clinical pilot study. *J Ultrasound Med* 2020;39:311-21.
 20. Gatos I, Drazinos P, Yarmenitis S, Theotokas I, Zoumpoulis PS. Comparison of sound touch elastography, shear wave elastography and vibration-controlled transient elastography in chronic liver disease assessment using liver biopsy as the "reference standard". *Ultrasound Med Biol* 2020;46:959-71.
 21. Bartko JJ. Measures of agreement: a single procedure. *Stat Med* 1994;13:737-45.
 22. Simkin P, Rattansingh A, Liu K, Hudson JM, Atri M, Jang HJ, Kim TK, Khalili K. Reproducibility of 2 liver 2-dimensional shear wave elastographic techniques in the fasting and postprandial states. *J Ultrasound Med* 2019;38:1739-45.
 23. Etchell E, Jugé L, Hatt A, Sinkus R, Bilston LE. Liver stiffness values are lower in pediatric subjects than in adults and increase with age: a multifrequency mr elastography study. *Radiology* 2017;283:222-30.
 24. Gao Y, Zheng J, Liang P, Tong M, Wang J, Wu C, et al. Liver fibrosis with two-dimensional us shear-wave elastography in participants with chronic hepatitis B: a prospective multicenter study. *Radiology* 2018;289:407-15.
 25. Roccarina D, Iogna Prat L, Buzzetti E, Guerrero Misas M, Aricó FM, Saffioti F, Rosselli M, Pinzani M, Marshall A, Thorburn D, Tsochatzis E. Establishing reliability criteria for liver ElastPQ shear wave elastography (ElastPQ-SWE): comparison between 10, 5 and 3 measurements. *Ultraschall Med* 2021;42:204-13.
 26. Kishimoto R, Kikuchi K, Koyama A, Kershaw J, Omatsu T, Tachibana Y, Suga M, Obata T. Intra- and inter-operator reproducibility of US point shear-wave elastography in various organs: evaluation in phantoms and healthy volunteers. *Eur Radiol* 2019;29:5999-6008.
 27. Lim Z, Whitaker T, DeColle K, Barrett K, Harlton C, Paskar L, Low G. Interobserver and intraobserver

- reliability of hepatic shear wave elastography and the influence of fasted versus nonfasted states in healthy volunteers. *J Ultrasound Med* 2021;40:259-67.
28. Potthoff A, Attia D, Pischke S, Kirschner J, Mederacke I, Wedemeyer H, Manns MP, Gebel MJ, Rifai K. Influence

of different frequencies and insertion depths on the diagnostic accuracy of liver elastography by acoustic radiation force impulse imaging (ARFI). *Eur J Radiol* 2013;82:1207-12.

Cite this article as: Xiang H, Ling W, Ma L, Yang L, Lin T, Luo Y. Shear wave elastography using sound touch elastography and supersonic shear imaging for liver measurements: a comparative study. *Quant Imaging Med Surg* 2022;12(5):2855-2865. doi: 10.21037/qims-21-863