Radiology

Reducing the Number of Measurements in Liver Point Shear-Wave Elastography:
Factors that Influence the Number and Reliability of Measurements in Assessment of Liver Fibrosis in Clinical Practice[1]

Cheng Fang, MBBS, FRCR
Ounali S. Jaffer, MRCP, FRCR
Gibran T. Yusuf, MBBS, FRCR
Eleni Konstantatou, MD, MSc
Daniel J. Quinlan, MBBS
Kosh Agarwal, MD, FRCP
Alberto Quaglia, MD, PhD, FRCPath
Paul S. Sidhu, MRCP, FRCR

Purpose: To identify the minimum number of measurements required for the noninvasive assessment of liver fibrosis by using point shear-wave elastography (pSWE) and determine whether the use of a reliability indicator such as interquartile range [IQR]–to–median ratio will affect diagnostic performance.

Materials and Methods: Ten liver shear-wave velocity (SWV) measurements by pSWE were obtained in 232 participants. Interclass correlation coefficients (ICC) between the median of the first two through the first nine measurements and all 10 measurements were calculated; the minimum number of measurements with ICC greater than 0.95 versus all 10 measurements was determined. The diagnostic performance of the minimum number of measurements and 10 measurements in identifying significant (Ishak stage, ≥3) and severe fibrosis or cirrhosis (Ishak stage, ≥5) was compared by using areas under the receiver operating characteristic curve. These were compared between measurements that demonstrated higher or lower reliability (IQR-to-median ratio of ≤ 30% and IQR-to-median ratio of > 30%, respectively).

Results: Compared with 10 measurements, a minimum of six SWV measurements was required. The overall area under the curve for diagnosing significant (areas under the receiver operating characteristic curve, 0.828 vs 0.839; *P* = .487) and severe fibrosis or cirrhosis (0.953 vs 0.969, respectively; *P* = .145) did not differ according to number of measurements (six vs 10); a median of six measurements resulted in only limited disagreement (nine of 232 [3.9%]) versus histologic evaluation. When using 10 measurements, higher reliability measurements showed a lower percentage of discordance between pSWE and significant fibrosis and severe fibrosis or cirrhosis (22 [14.7%] and three [2.0%] of 150 cases, respectively) compared with lower reliability measurements (26 [31.7%] and eight [9.8%] of 82 cases, respectively). Significant fibrosis was an independent predictor for lower reliability (hazard ratio, 2.22; *P* < .020).

Conclusion: A limited number of SWV measurements (median six vs median 10) were required for the assessment of liver fibrosis by using pSWE. The number of measurements had less influence on the diagnostic accuracy compared with lower reliability measurements.

©RSNA, 2018

Online supplemental material is available for this article.

Radiology

Liver biopsy is a so-called reference standard in the evaluation of the degree of liver fibrosis in patients with chronic liver disease (1). However, noninvasive methods for assessment of liver fibrosis including shear-wave elastography (SWE) have gained increasing acceptance, obviating the need for a liver biopsy, a diagnostic procedure associated with substantial morbidity and mortality (2,3). SWE involves mechanical excitation of tissue with a short duration of an acoustic push pulse, which generates localized displacement in tissue resulting in shear-wave propagation. The speed of shear-wave propagation within the tissue measured in meters per second is related to the stiffness of the tissue (4,5). Previous studies (6–13) demonstrated that point SWE (pSWE) is a reliable and accurate technique in assessing the stage of liver fibrosis in both healthy volunteers and patients with chronic liver disease. This noninvasive diagnostic method is included in current clinical guidelines for the initial assessment of liver fibrosis or cirrhosis, with liver biopsy reserved for patients in whom there is uncertainty about potential additional etiologic causes of liver disease (14,15).

The majority of previous studies that evaluated pSWE to stage liver fibrosis used 10 valid measurements and reported median or mean values, with some studies using a median of five (16), six (17,18), 12 (19), or 20 (20) measurements. Unlike transient elastography where 10 measurements are recommended, limited information is provided in pSWE manufacturer guidelines about the number of repeated

measurements and quality measures required. Recent guidelines from the European Federation of Societies for Ultrasound in Medicine and Biology suggest that at least 10 measurements should be used for liver fibrosis measurements and discuss the potential role of the interquartile range (IQR)-to-median ratio as a reliability indicator (21). However, the practice of acquiring 10 measurements appears to be a common convention rather than determined by systematic evaluation.

The purpose of this study was to identify the minimum number of measurements required for the noninvasive assessment of liver fibrosis by using pSWE and to determine whether the use of a reliability indicator such as IQR-to-median ratio will affect diagnostic performance.

Materials and Methods

This retrospective analysis was performed with local ethics approval for several prospective SWE studies performed at our institution (12). Written informed consent was obtained from all patients and healthy volunteers who participated in each study.

Study Population

Between October 2009 and July 2015, 232 patients and healthy volunteers who underwent pSWE were retrospectively reviewed. pSWE measurements from 11 healthy volunteers were used in a previous study (12); this previous study focused on the reproducibility of two-dimensional SWE by comparing it with pSWE. There was no overlap with our analysis.

The patient cohort, including 221 patients suspected of having chronic liver disease who underwent pSWE to evaluate the degree of liver fibrosis, has not been previously reported. The demographic characteristics of the patient and healthy volunteers are listed in Table 1. The patients attended the Radiology Department Day Case Unit to undergo liver biopsy to assess their underlying disease; the pSWE examination was performed on the same day. The causes of liver disease included hepatitis

B, C, and combined B/C/D; alcoholic liver disease; nonalcoholic fatty liver disease; and autoimmune and inflammatory disease. Aspartate aminotransferase, alanine aminotransferase, γ glutamyl transferase, total bilirubin, platelet count, international normalized ratio, and alkaline phosphatase were measured.

pSWE Examination

All patients and healthy volunteers fasted overnight and were scanned in a supine position with arms extended over their head, in suspended respiration, in accordance with guidelines (21). The pSWE examination was performed by trained radiologists (C.F., O.S.J., G.T.Y., E.K.) who had at least 2 years of ultrasonographic (US) experience and training in the pSWE technique (each of them performed at least 200 previous liver elastography examinations per the European Federation of Societies for Ultrasound in Medicine and Biology competency requirement). pSWE was performed with a quantitative software technique (Virtual Touch Quantification; Siemens Healthcare, Mountain View, Calif) by using a US imaging system (Acuson S2000/S3000; Siemens), equipped with a convex broadband (6C1, 1.5 MHz–6 MHz) transducer. A region-of-interest box was

Implication for Patient Care

- Our results demonstrate that six measurements have excellent reliability compared with 10 measurements when conducting point shear-wave elastography evaluation of liver fibrosis and results in limited disagreement (3.9% of cases) verus histologic evaluation, potentially allowing less time required to acquire measurements.

Table 1

Demographic Characteristics of the Study Patients and Healthy Volunteers

Parameter	Result
Age (y)*	42 (18–70)
Sex (no. of patients)	
Female	85 (36.6)
Male	147 (63.4)
Diagnosis (no. of participants)	
Healthy volunteers	11 (4.7)
Chronic hepatitis B infection	111 (47.8)
Chronic hepatitis C infection	65 (28.0)
Chronic hepatitis B/C/D coinfections	3 (1.3)
NAFLD/ALD	16 (6.9)
Autoimmune	10 (4.3)
Inflammatory (PSC/UC)	7 (3.0)
Other/unknown etiologic cause	9 (3.9)
Steatosis (no. of incidences)	
None	96 (41.4)
Mild	66 (28.4)
Moderate	19 (8.2)
Marked	5 (2.2)
Not available including 11 health volunteers	46 (19.8)
Ishak histologic fibrosis stage (no. of patients)	
0	29 (12.5)
1	62 (26.7)
2	39 (16.8)
3	36 (15.5)
4	22 (9.5)
5	12 (5.2)
6	32 (13.8)
AST (U/L)	36 (15–1413)
ALT (U/L)	43 (10–341)
GGT (U/L)	35 (6–1391)
ALP (U/L)	77.5 (16–584)
Total bilirubin (μmol/L)	10 (4–122)
INR	1.03 (1–2)
Platelets (×10^9/L)	209 (37–661)

Note.—Median variables had nonnormal distribution. Unless otherwise indicated, data in parentheses are percentages. ALD = alcoholic liver disease, ALP = alkaline phosphatase, ALT = alanine aminotransferase aminotransferase, AST = aspartate, GGT = γ glutamyl transferase, INR = international normalized ratio, NAFLD = nonalcoholic fatty liver disease, PSC = primary sclerosing cholangitis, UC = ulcerative colitis.

*Data are median; data in parentheses are range.

placed in the right lobe of the liver at a depth between 2 and 6 cm, avoiding biliary ducts and large vessels by using an intercostal scanning approach (Fig 1). The tissue elasticity analysis was performed by using software (Virtual Touch Quantification; Siemens) that quantified liver stiffness as shear-wave velocity (SWV) in meters per second during brief breath-holding in a neutral position (≤5 seconds). Ten consecutive valid SWV measurements were recorded without discarding any results. These measurements were divided into those that had an IQR-to-median ratio of greater than 30% and 30% or less; the latter were considered to indicate reliable results. Failure of pSWE measurement was defined as the inability to obtain a valid measurement (displayed on the screen as X.XX) after 10 attempts.

Histologic Evaluation

Core liver biopsy samples were taken from the right lobe of the liver by using an 18-gauge needle (Bard Marquee Disposable Core Biopsy Instrument; Bard Biopsy Systems, Tempe, Ariz) and prepared per standard departmental protocol. Biopsy fragments of 15 mm or greater were considered appropriate for pathologic evaluation. Biopsies were examined by a single hepatobiliary pathologist (A.Q., with >20 years of experience) who was blinded to the SWV measurements. Ishak fibrosis staging system was used to evaluate pathologic liver fibrosis and inflammatory activity (Table E1 [online]) (22). Significant fibrosis was defined as Ishak stage 3 and above, and severe fibrosis or cirrhosis was defined as Ishak stage 5 and above. The fibrosis stages of healthy volunteers were Ishak stage 0. The degree of steatosis was classified by visual estimation of the biopsy sample according to the percentage of hepatocytes with fatty changes: none (0% of hepatocytes), mild (≤33%), moderate (33%–66%), or marked (>66%) (23). The presence of steatosis is assumed to be none or mild for the healthy volunteers. Steatosis was grouped into no or mild steatosis and moderate or marked steatosis for the statistical analysis.

Statistical Analysis

Statistical analysis was performed by using software (SPSS for Windows version 22, IBM, Chicago, Ill; and MedCalc for windows, v17.6, MedCalc Software, Ostend, Belgium). To determine the minimum number of SWV measurements required, interclass correlation coefficients (ICCs) and 95% confidence intervals were obtained between subsets of median values from the initial second to ninth measurement, and the reference median value from 10 measurements. In addition, Wilcoxon signed-rank tests were used to determine whether there was significant difference in accuracy between the median of the first five and the median of the last five measurements. The minimum number of measurements with an ICC threshold of greater than 0.95 (ie, an error of less than 5%) versus that from all 10 measurements was determined. Bland-Altman 95% confidence intervals were calculated to assess the agreement between the median value of the minimum number of measurements and the median of 10 measurements. Any difference in continuous variables between patients with an IQR-to-median ratio of 30% or less or an IQR-to-median ratio of greater than 30% was analyzed by using Mann-Whitney U test (ie, nonparametric data) or t test (ie, parametric data). The χ^2 test or Fisher exact test was used for categorical variables. All variables that were significantly associated with IQR-to-median ratio greater than 30% at univariate analysis were analyzed by using multivariable logistic regression analysis. Correlation between the SWV measurements and fibrosis stages was assessed by using Spearman rank correlation coefficient (r). The diagnostic performance of the median value of the minimum number of pSWE measurements and the reference median value were assessed by receiver operator characteristic curves. Differences between the areas under the receiver operating characteristic curve were compared by using a Delong test (24). The optimal cutoff value was determined by using the Youden index (25). Subgroup diagnostic performance analyses were performed by using participants with measurements
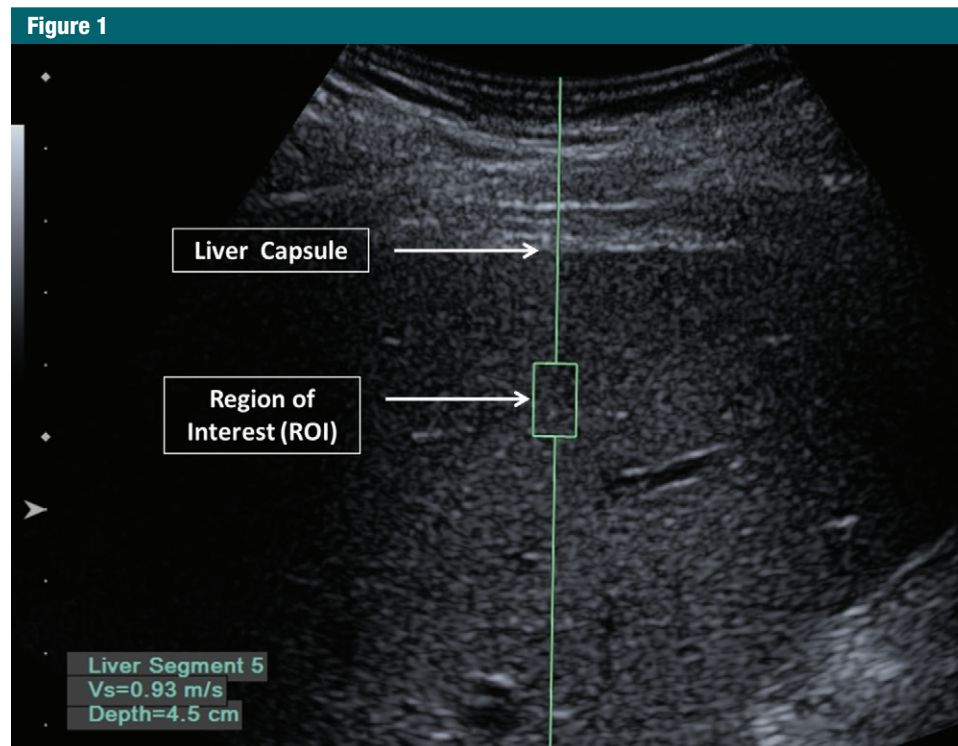
Figure 1

Figure 1: Point shear-wave elastographic examination of the liver performed by using an S2000 US machine (Siemens, Healthcare). The liver capsule and region of interest are shown.

that demonstrated an IQR-to-median ratio greater than 30% and IQR-to-median ratio of 30% or less. A two-sided P value of less than .05 was considered to indicate statistical significance.

Results

Study Participant Characteristics

The characteristics of the study patients and healthy volunteers are listed in Table 1. Histologic grading of liver fibrosis was available from 221 patients. The 11 healthy volunteers did not undergo liver biopsy. SWV measurements were obtained in all 232 patients and healthy volunteers who underwent pSWE. SWV measurements with lower reliability (IQR-to-median ratio, >30%) were observed in 82 of the enrolled participants (35.3%).

ICCs between Median Measurements

In the overall cohort, the ICC between a median of two to a median of nine measurements and a median of 10 measurements ranged from 0.901 (95% confidence interval: 0.873, 0.922) to 0.995 (95% confidence interval: 0.993, 0.996) (Table 2). The ICC between the median of the first six measurements and median of all 10 measurements was 0.966 (95% confidence interval: 0.957, 0.974), which is the minimum number of measurements that demonstrate an ICC above the cutoff value of 0.95. The Bland-Altman analysis shows that the mean difference between median of six and median of 10 measurements is −0.002 m/sec (95% confidence interval: −0.33, 0.32) with no statistically significant systematic bias (P = .82) (Fig 2).

The ICC for the 150 patients and healthy volunteers with IQR-to-median ratio 30% or less ranged from 0.939 (95% confidence interval: 0.916, 0.955; two repeated measurements) to 0.998 (95% confidence interval: 0.997, 0.999; nine repeated measurements), and the ICC between the median of the first three repeated measurements and 10 measurements demonstrated an ICC value above 0.95. In the cohort of patients and volunteers with

measurements that demonstrated IQR-to-median ratio of greater than 30%, the range of ICC was between 0.859 (95% confidence interval: 0.790, 0.907; two measurements) and 0.991 (95% confidence interval: 0.986, 0.994; nine measurements), and the median of the first eight measurements and 10 measurements demonstrated an ICC value above 0.95. There was no significant difference between the ICC by using the median of the first five measurements (ie, measurements from one to five) versus the median of the second five measurements (ie, measurements from six to 10) (P = .562). In the subgroup analysis of patients with no or mild steatosis (n = 162), or moderate or marked steatosis (n = 24), the number minimum measurements required was eight for both subgroups.

Comparison of Performance for the Fibrosis Assessment by Using Median of Different Number of Measurements

The median SWVs increased with higher Ishak stage of fibrosis (1.176 m/sec

Table 2

Interclass Correlation Coefficients between the Median Shear-Wave Velocity Value of Repeated Shear-Wave Elastography Measurement and Median of 10 Measurements

No. of Repetitions	All Participants (n = 232)	Participants with High Reliability (n = 150)	Participants with Low Reliability (n = 82)	Participants with No or Mild Steatosis (n = 162)	Participants with Moderate or Severe Steatosis (n = 24)
2	0.901 (0.873, 0.922)	0.939 (0.916, 0.955)	0.859 (0.790, 0.907)	0.853 (0.804, 0.890)	0.755 (0.513, 0.886)
3	0.903 (0.876, 0.924)	0.973* (0.963, 0.980)	0.837 (0.758, 0.891)	0.865 (0.820, 0.899)	0.839 (0.623, 0.931)
4	0.942 (0.926, 0.955)	0.984 (0.978, 0.989)	0.901 (0.850, 0.935)	0.882 (0.843, 0.912)	0.898 (0.781, 0.954)
5	0.944 (0.928, 0.956)	0.988 (0.983, 0.991)	0.902 (0.852, 0.935)	0.883 (0.844, 0.913)	0.934 (0.855, 0.971)
6	0.966* (0.957, 0.974)	0.991 (0.988, 0.994)	0.941 (0.911, 0.962)	0.925 (0.899, 0.944)	0.940 (0.868, 0.973)
7	0.974 (0.966, 0.980)	0.994 (0.991, 0.995)	0.954 (0.930, 0.970)	0.953 (0.936, 0.965)	0.947 (0.883, 0.977)
8	0.989 (0.986, 0.992)	0.996 (0.994, 0.997)	0.982* (0.973, 0.988)	0.978* (0.970, 0.984)	0.990* (0.978, 0.996)
9	0.995 (0.993, 0.996)	0.998 (0.997, 0.999)	0.991 (0.986, 0.994)	0.994 (0.991, 0.995)	0.980 (0.954, 0.991)

Note.—Data in parentheses are 95% confidence intervals. High reliability is an interquartile range–to–median ratio of 30% or less; low reliability is an interquartile range–to–median ratio of greater than 30%.

* Represent intraclass correlation coefficient of 0.95 or greater between the number of repetitions and median value of 10 measurements. Steatosis data were available in 186 participants.

[stage 0], 1.133 m/sec [stage 1], 1.191 m/sec [stage 2], 1.358 m/sec [stage 3], 1.594 m/sec [stage 4], 2.145 m/sec [stage 5], and 2.650 m/sec [stage 6]), with a significant difference between the different stages of fibrosis ($P \leq$ .001).

The median of 10 SWV measurements differentiated significant fibrosis (Ishak stage 0–2 vs 3–6) and severe fibrosis or cirrhosis (Ishak stage 0–4 vs 5–6) with an area under the receiver operating characteristic curve of 0.839 (95% confidence interval: 0.786, 0.884) and 0.969 (95% confidence interval: 0.937, 0.987), respectively. The median of six SWV measurements differentiated significant fibrosis (Ishak stage 0–2 vs 3–6) and severe fibrosis or cirrhosis (Ishak stage 0–4 vs 5–6), with an area under the receiver operating characteristic curve of 0.828 (95% confidence interval: 0.773, 0.874) and 0.953 (95% confidence interval: 0.918, 0.977), respectively. There was no statistically significant difference in the area under the receiver operating characteristic curve between the median of six SWV and 10 SWV measurements in diagnosing both significant fibrosis (P = .487) and severe fibrosis or cirrhosis (P = .145). Similarly, there was no statistically significant difference in differentiating significant or severe fibrosis between a median of six and median of 10 measurements in
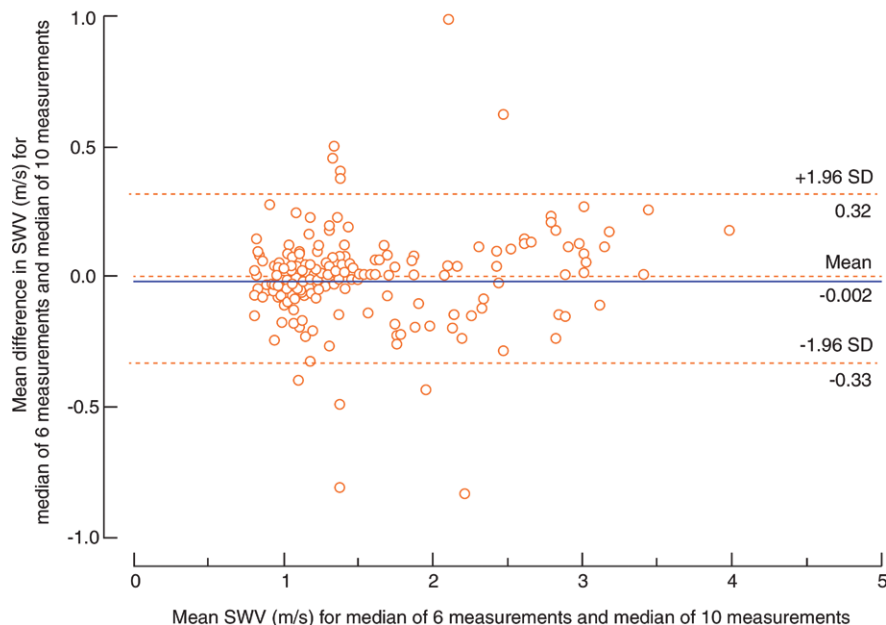
Figure 2

Figure 2: Bland-Altman plot of the differences in the shear-wave velocity (SWV) between median of six and median of 10 measurements. The solid line represents the mean of the difference in SWV; the dashed lines represent the 95% upper and lower limits of agreement. SD = standard deviation.

the subgroups with an IQR-to-median ratio greater than 30% and an IQR-to-median ratio of 30% or less (Table 3).

The number of patients with discordance between pSWE and histologic results that classified significant fibrosis (Ishak stage ≥3) was 48 (20.7%) with a median of 10 measurements and 51

(22.0%) with a median of six measurements (Table 4). Of the additional three patients with discordance, one patient had an IQR-to-median ratio greater than 30%. The number of patients with discordance between pSWE and histologic results for classification of severe fibrosis or cirrhosis (Ishak stage,

Table 3

Diagnostic Performance of Point Shear-Wave Elastography

Parameter	Overall			IQR-to-Median Ratio, ≤30%			IQR-to-Median Ratio, >30%		
	Median 6 pSWE	Median 10 pSWE	P Value	Median 6 pSWE	Median 10 pSWE	P Value	Median 6 pSWE	Median 10 pSWE	P Value
Significant fibrosis			.487			>.999			.463
AUROC	0.828 (0.773, 0.874)	0.839 (0.786, 0.884)		0.855 (0.788, 0.907)	0.855 (0.788, 0.907)		0.798 (0.695, 0.879)	0.822 (0.722, 0.898)	
Sensitivity (%)	76.70	65.05		79.25	77.36		82.00	68.00	
Specificity (%)	78.29	90.70		81.44	82.47		68.75	84.37	
Severe fibrosis/cirrhosis			.145			.202			.162
AUROC	0.953 (0.918, 0.977)	0.969 (0.937, 0.987)		0.980 (0.943, 0.996)	0.984 (0.949, 0.997)		0.920 (0.839, 0.969)	0.943 (0.868, 0.982)	
Sensitivity (%)	95.45	95.45		95.24	95.24		95.65	95.65	
Specificity (%)	90.43	95.74		93.80	98.45		83.05	89.83	

Note.—Data in parentheses are 95% confidence interval. Significant fibrosis refers to Ishak stage 3 or greater; severe fibrosis or cirrhosis refers to Ishak stage 5 or greater. AUROC = area under the receiver operating characteristic curve, IQR = interquartile range, pSWE = point shear-wave elastography.

≥5) was 11 (4.7%) with a median of 10 measurements and 17 (7.3%) with six measurements. Of the additional six patients with discordance for the classification of significant fibrosis or cirrhosis, four patients had an IQR-to-median ratio greater than 30% (Table 4). When a median of 10 measurements were used, the percentage of patients who demonstrated discordance between the pSWE and histologic analysis increased from 14.7% (IQR-to-median ratio, ≤30%) to 31.7% (IQR-to-median ratio, >30%) and 2.0% (IQR-to-median ratio, ≤30%) to 9.8% (IQR-to-median ratio, >30%) in differentiating significant fibrosis and severe fibrosis or cirrhosis, respectively.

Predictors of Lower Reliability

In the univariate analysis, advanced age (P = .038), significant fibrosis (P < .001) and moderate or marked steatosis (P = .024) were associated with IQR-to-median ratio greater than 30%, whereas sex (P = .553), total bilirubin (P = .236), aspartate aminotransferase (P = .458), alanine aminotransferase (P = .064), γ glutamyl transferase (P = .048), alkaline phosphatase (P = .099), international normalized ratio (P = .461), and platelet count (P = .393) were not significantly different between patients with IQR-to-median ratio of 30% or less and of greater than 30%. In the multivariable logistic regression analysis, only significant fibrosis (B, 0.795; standard error, 0.343; hazard ratio, 2.22; P < .020) was an independent predictor of lower reliability (IQR-to-median ratio, >30%) after adjusting for age and steatosis.

Discussion

In our study, a median of six SWV measurements was the minimum number required that resulted in an error rate of less than 5% compared with a median of 10 measurements. In addition, the 95% agreement limit from the Bland Altman analysis was also comparable to the 95% agreement limit from interobserver variability previously described (26). The area under the receiver operating characteristic curve

Radiology

Table 4

Discordance between Point Shear-Wave Elastography and Histologic Analysis in the Overall Cohort and Subgroups

Parameter	Significant Fibrosis		Severe Fibrosis or Cirrhosis	
	Median of 6 pSWE	Median of 10 pSWE	Median of 6 pSWE	Median of 10 pSWE
IQR-to-median ratio, ≤30% (n = 150)	24 (16.0)	22 (14.7)	5 (3.3)	3 (2.0)
IQR-to-median ratio, >30% (n = 82)	27 (32.9)	26 (31.7)	12 (14.6)	8 (9.8)
Overall (n = 232)	51 (22.0)	48 (20.7)	17 (7.3)	11 (4.7)

Note.—Data are number of patients; data in parentheses are percentages. Significant fibrosis refers to Ishak stage 3 or greater; severe fibrosis or cirrhosis refers to Ishak stage 5 or greater. IQR = interquartile range, pSWE = point shear-wave elastography.

for differentiating significant fibrosis and severe fibrosis or cirrhosis by using the median of six measurements was slightly lower than the area under the receiver operating characteristic curve when the median of 10 measurements was used, but the difference was not statistically significant. The area under the receiver operating characteristic curve value by using the median of six measurements was within the range of published data (9,27). In our cohort of 232 participants, the use of six measurements instead of 10 only misclassified three patients with significant fibrosis and six patients with severe fibrosis or cirrhosis.

An increase in the number of measurements (ie, from six to 10) resulted in greater reduction in the percentage of discordant patients with severe fibrosis or cirrhosis (7.3% from six measurements vs 4.7% from 10 measurements) than for those with significant fibrosis (22.0% from six measurements vs 20.7% from 10 measurements).

Previous studies have used the median or the mean of 10 measurements as routine practice. Our data suggest that acquiring fewer measurements will have a limited effect on the diagnostic performance of pSWE with only limited disagreement (nine of 232 [3.9%] cases) versus histologic evaluation, potentially saving clinic time. Yoon et al (18) also suggested that the optimal minimum number of SWE measurements required was six in their cohort of 86 patients; they used the

same approach by comparing the ICC for the median of the initial two to nine measurements with the median of 10 measurements. Other studies (28,29) that compared pSWE (ElastPQ, Philips, Endihoven; and Virtual Touch Quantification) with transient elastography (Fibroscan, Paris, France) showed that five measurements could be used instead of 10 without a significant effect on diagnostic performance. Similar findings have been observed with use of transient elastography where the diagnostic accuracy of detecting cirrhosis was not affected by using five valid measurements instead of 10 with histologic evaluation as the comparison (30).

IQR-to-median ratio is a method for measuring data variability, and is recommended as a reliability indicator for transient elastography by the manufacturer (Echosens, Paris, France) (21). The importance of IQR-to-median ratio as a quality control measure for pSWE was observed in several studies (31,32). Increased accuracy for classifying fibrosis was noted when measurements that showed lower reliability (IQR-to-median ratio, >30%) were excluded (29,31–33). Our results showed that use of higher reliability measurements (IQR-to-median ratio, ≤30%) can further reduce the number of measurement required from six to three by maintaining an ICC greater than 0.95 with 10 measurements. Patients with lower reliability measurements (IQR-to-median ratio, >30%) showed a 2.2-fold and 4.9-fold increase in discordance

rate with histologic evaluation for differentiating between significant and severe fibrosis or cirrhosis, respectively, compared with patients with higher reliability measurements (IQR-to-median ratio, ≤30%). However, the reduction of measurements from 10 to six resulted in 1.1-fold and 1.7-fold increase in the discordance rate with histologic evaluation in differentiating significant and severe fibrosis, respectively. In our study, the percentage of discordance with histologic evaluation were more affected by IQR-to-median ratio than the number of measurements.

IQR-to-median ratio can only be calculated retrospectively because, to our knowledge, no reliability indicator is available at the time of acquiring the measurement. In view of the predefined threshold of a less than 5% measurement error, the number of measurements required compared with 10 measurements will depend on the proportion of measurements with an unreliable result (ie, IQR-to-median ratio of >30%). The presence of significant fibrosis was shown to be an independent predictor for an unreliable result contrary to previous evidence showing that cirrhosis does not influence the rate of reliable pSWE measurements (34). Unlike findings from a previous study by Bota et al (34), age and sex were not independent predictors for unreliable results. This may be because of the proportion of unreliable results in the previous publication also included measurement with success rate less than 60% and the proportion of patients with an unreliable result (6.4%) was significantly lower than the current study (35.3% of participants with 10 measurements). This is comparable to the 42.4% figure published by Ferraioli et al (29) in a cohort of 255 patients by using pSWE with the ElastPQ technique (Philips, Eindhoven). However, despite over a third of participants who were considered to be unreliable, our results demonstrated that it is still possible to assess liver fibrosis with acceptable accuracy (area under the receiver operating characteristic curve for significant fibrosis, 0.822; area under the receiver operating characteristic curve

for severe fibrosis, 0.943) in these patients and it is not necessary to exclude measurements with IQR-to-median ratio greater than 30%.

We did not identify any predictors for lower reliability (IQR-to-median ratio, >30%) that could be assessed prior to obtaining the pSWE measurements because the existence of significant fibrosis was unknown at the time of imaging. Recently, other pSWE techniques by using the Acoustic Radiation Force Impulse principle (35–37) were introduced in which a reliability indicator was displayed at the point of measurement, for example the SWV measurement efficacy rate. This potentially allows selective reduction of the number of measurements required for the noninvasive assessment of liver fibrosis at the point of imaging.

Despite the large number of pSWE studies reported, few have evaluated the number of measurements required and their effect on diagnostic accuracy. The strength of our study compared with previous studies is that the minimum number of measurement was not only determined by the ICC against the conventional practice of the median of 10 measurements but also against liver biopsy results; in previous studies (18,29,38), diagnostic performances were compared with transient elastography and the minimum number of measurements required were only determined by ICC or correlation against a fixed number of measurements.

This study has limitations. First, it was a single center study and the data were analyzed retrospectively. Therefore, certain parameters such as body mass index were not recorded at the time of study. Second, the study population contained a mixed etiologic cause and the optimum cut-off values for assessing fibrosis may vary, depending upon the underlying cause. Finally, the use of Ishak fibrosis stage from liver biopsy, although it provides a reference standard when the diagnostic accuracy of pSWE is assessed, is subject to sampling error because of heterogeneity of the disease process in cirrhosis.

In summary, the reliability of pSWE measurements has a greater effect on diagnostic accuracy than the number of measurements. Although significant fibrosis is associated with lower reliability, fewer than 10 measurements can accurately assess liver stiffness by using pSWE.

References

1. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD; American Association for the Study of Liver Diseases. Liver biopsy. Hepatology 2009;49(3):1017–1044.

2. Piccinino F, Sagnelli E, Pasquale G, Giusti G. Complications following percutaneous liver biopsy. A multicentre retrospective study on 68,276 biopsies. J Hepatol 1986;2(2):165–173.

3. McGill DB, Rakela J, Zinsmeister AR, Ott BJ. A 21-year experience with major hemorrhage after percutaneous liver biopsy. Gastroenterology 1990;99(5):1396–1400.

4. Zhai L, Palmeri ML, Bouchard RR, Nightingale RW, Nightingale KR. An integrated indenter-ARFI imaging system for tissue stiffness quantification. Ultrason Imaging 2008;30(2):95–111.

5. Nightingale K, Soo MS, Nightingale R, Trahey G. Acoustic radiation force impulse imaging: in vivo demonstration of clinical feasibility. Ultrasound Med Biol 2002;28(2):227–235.

6. Guzmán-Aroca F, Reus M, Berná-Serna JD, et al. Reproducibility of shear wave velocity measurements by acoustic radiation force impulse imaging of the liver: a study in healthy volunteers. J Ultrasound Med 2011;30(7):975–979.

7. Colombo S, Buonocore M, Del Poggio A, et al. Head-to-head comparison of transient elastography (TE), real-time tissue elastography (RTE), and acoustic radiation force impulse (ARFI) imaging in the diagnosis of liver fibrosis. J Gastroenterol 2012;47(4):461–469.

8. Ferraioli G, Tinelli C, Lissandrin R, et al. Point shear wave elastography method for assessing liver stiffness. World J Gastroenterol 2014;20(16):4787–4796.

9. Friedrich-Rust M, Wunder K, Kriener S, et al. Liver fibrosis in viral hepatitis: noninvasive assessment with acoustic radiation force impulse imaging versus transient elastography. Radiology 2009;252(2):595–604.

10. Bota S, Sporea I, Sirli R, Popescu A, Danila M, Costachescu D. Intra- and interoperator reproducibility of acoustic radiation force impulse (ARFI) elastography--preliminary results. Ultrasound Med Biol 2012;38(7):1103–1108.

11. D'Onofrio M, Gallotti A, Mucelli RP. Tissue quantification with acoustic radiation force impulse imaging: Measurement repeatability and normal values in the healthy liver. AJR Am J Roentgenol 2010;195(1):132–136.

12. Fang C, Konstantatou E, Romanos O, Yusuf GT, Quinlan DJ, Sidhu PS. Reproducibility of 2-Dimensional Shear Wave Elastography Assessment of the Liver: A Direct Comparison With Point Shear Wave Elastography in Healthy Volunteers. J Ultrasound Med 2017;36(8):1563–1569.

13. Jaffer OS, Lung PFC, Bosanac D, et al. Acoustic radiation force impulse quantification: repeatability of measurements in selected liver segments and influence of age, body mass index and liver capsule-to-box distance. Br J Radiol 2012;85(1018):e858–e863.

14. European Association for the Study of the Liver. EASL recommendations on treatment of hepatitis C 2014. J Hepatol 2014;61(2):373–395.

15. European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu; European Association for the Study of the Liver. EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. J Hepatol 2017;67(2):370–398.

16. Takahashi H, Ono N, Eguchi Y, et al. Evaluation of acoustic radiation force impulse elastography for fibrosis staging of chronic liver disease: a pilot study. Liver Int 2010;30(4):538–545.

17. Piscaglia F, Salvatore V, Di Donato R, et al. Accuracy of VirtualTouch Acoustic Radiation

Force Impulse (ARFI) imaging for the diagnosis of cirrhosis during liver ultrasonography. Ultraschall Med 2011;32(2):167–175.

18. Yoon JH, Lee JM, Han JK, Choi BI. Shear wave elastography for liver stiffness measurement in clinical sonographic examinations: evaluation of intraobserver reproducibility, technical failure, and unreliable stiffness measurements. J Ultrasound Med 2014;33(3):437–447.

19. Fierbinteanu-Braticevici C, Andronescu D, Usvat R, Cretoiu D, Baicus C, Marinoschi G. Acoustic radiation force imaging sonoelastography for noninvasive staging of liver fibrosis. World J Gastroenterol 2009;15(44):5525–5532.

20. Rizzo L, Calvaruso V, Cacopardo B, et al. Comparison of transient elastography and acoustic radiation force impulse for noninvasive staging of liver fibrosis in patients with chronic hepatitis C. Am J Gastroenterol 2011;106(12):2112–2120.

21. Dietrich CF, Bamber J, Berzigotti A, et al. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). Ultraschall Med 2017;38(4):e16–e47.

22. Ishak K, Baptista A, Bianchi L, et al. Histological grading and staging of chronic hepatitis. J Hepatol 1995;22(6):696–699.

23. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol 1999;94(9):2467–2474.

24. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.

25. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3(1):32–35.

26. Mancini M, Salomone Megna A, Ragucci M, et al. Reproducibility of shear wave elastography (SWE) in patients with chronic liver disease. PLoS One 2017;12(10):e0185391.

27. Sporea I, Bota S, Peck-Radosavljevic M, et al. Acoustic Radiation Force Impulse elastography for fibrosis evaluation in patients with chronic hepatitis C: an international multicenter study. Eur J Radiol 2012;81(12):4112–4118.

28. Choi SH, Jeong WK, Kim Y, et al. How many times should we repeat measuring liver stiffness using shear wave elastography?: 5-repetition versus 10-repetition protocols. Ultrasonics 2016;72:158–164.

29. Ferraioli G, Maiocchi L, Lissandrin R, et al. Accuracy of the ElastPQ Technique for the Assessment of Liver Fibrosis in Patients with Chronic Hepatitis C: a "Real Life" Single Center Study. J Gastrointestin Liver Dis 2016;25(3):331–335.

30. Kettaneh A, Marcellin P, Douvin C, et al. Features associated with success rate and performance of FibroScan measurements for the diagnosis of cirrhosis in HCV patients: a prospective study of 935 patients. J Hepatol 2007;46(4):628–634.

31. Bota S, Sporea I, Sirli R, Popescu A, Jurchis A. Factors which influence the accuracy of acoustic radiation force impulse (ARFI) elastography for the diagnosis of liver fibro-

sis in patients with chronic hepatitis C. Ultrasound Med Biol 2013;39(3):407–412.

32. Bota S, Sporea I, Sirli R, Popescu A, Dănilă M, Sendroiu M. Factors that influence the correlation of acoustic radiation force impulse (ARFI), elastography with liver fibrosis. Med Ultrason 2011;13(2):135–140.

33. Goertz RS, Sturm J, Pfeifer L, et al. ARFI cut-off values and significance of standard deviation for liver fibrosis staging in patients with chronic liver disease. Ann Hepatol 2013;12(6):935–941.

34. Bota S, Sporea I, Sirli R, et al. Factors associated with the impossibility to obtain reliable liver stiffness measurements by means of Acoustic Radiation Force Impulse (ARFI) elastography--analysis of a cohort of 1,031 subjects. Eur J Radiol 2014;83(2):268–272.

35. Dietrich CF, Dong Y. Shear wave elastography with a new reliability indicator. J Ultrason 2016;16(66):281–287.

36. Yada N, Sakurai T, Minami T, et al. A Newly Developed Shear Wave Elastography Modality: With a Unique Reliability Index. Oncology 2015;89(Suppl 2):53–59.

37. Jeong WK. Liver stiffness measurement using S-Shearwave: initial experience. White Paper, CS201505-S-Shearwave, Samsung Medison, 2015. http://im2.medias.no/wp-content/uploads/2017/03/49_CS-S-Shearwave_150528.pdf. Accessed July 12, 2017.

38. Sporea I, Grădinaru-Taşcău O, Bota S, et al. How many measurements are needed for liver stiffness assessment by 2D-Shear Wave Elastography (2D-SWE) and which value should be used: the mean or median? Med Ultrason 2013;15(4):268–272.