Received: 27 April 2019 | Revised: 14 July 2019 | Accepted: 4 August 2019

DOI: 10.1002/jcu.22771

RESEARCH ARTICLE

WILEY

Interobserver agreement and diagnostic accuracy of shearwave elastography for the staging of hepatitis C virus-associated liver fibrosis

Pál N. Kaposi MD, PhD[1] | Zsuzsanna Unger MD, PhD[1] | Bence Fejér MD[1] | András Kucsa MD[1] | Ambrus Tóth MD[1] | Anikó Folhoffer MD[2] | Ferenc Szalay MD, DSc[2] | Viktor Bérczi MD, DSc[1]

[1]Department of Radiology, Faculty of Medicine, Semmelweis University, Budapest, Hungary

[2]1st Department of Internal Medicine, Faculty of Medicine, Semmelweis University, Budapest, Hungary

Correspondence
Pál Novák Kaposi, MD, PhD, Department of Radiology, Semmelweis University, Üllői út 78/a, Budapest, Hungary.
Email: kaposipal@gmail.com

Funding information
Hungarian Academy of Sciences

Abstract

Purpose: Our study aimed to evaluate the technical success rate, interobserver reproducibility, and accuracy of shearwave elastography (SWE) in the staging of hepatitis C virus (HCV)-associated liver fibrosis.

Methods: A total of 10 healthy controls and 49 patients with chronic liver disease were enrolled prospectively. Two examiners performed point shearwave elastography (pSWE) and two-dimensional shearwave elastography (2D-SWE) measurements with an RS85A ultrasound scanner using the S-Shearwave application (Samsung Medison, Hongcheon, Korea). The performance of S-Shearwave in the staging (METAVIR F0-F4) of liver fibrosis was compared with prior transient elastography (TE) with receiver operating characteristic (ROC) curve analysis.

Results: The interobserver reproducibility was excellent with pSWE (ICC = 0.92, 95% CI: 0.86-0.95, $P < .001$). A very good agreement was found between pSWE and TE measurements (ICC =0.85, 95% CI: 0.78-0.89, $P < .001$). The ROC analysis determined the optimal cut-off values of pSWE for the staging of chronic hepatitis C-associated fibrosis (F2, 1.46 m/s; F3, 1.63 m/s; F4, 1.95 m/s). Both observers achieved excellent diagnostic accuracy (AUROC: 94% vs 97%) in the detection of significant (≥F2) liver fibrosis.

Conclusion: The interobserver agreement is excellent with S-Shearwave pSWE, and observers can diagnose significant liver fibrosis with a comparable accuracy to TE.

KEYWORDS

chronic hepatitis C, HCV, liver fibrosis, shearwave elastography

1 | INTRODUCTION

Fibrosis and architectural distortion is a common endpoint of chronic liver disease (CLD) of various etiologies. Quantification of liver fibrosis is essential for therapeutic decision-making in these patients. According to current clinical practice guidelines, antiviral treatment is recommended when significant grade fibrosis is detected. After the diagnosis of cirrhosis has been established, regular screenings for complications and surveillance for hepatocellular carcinoma (HCC) must be initiated.[1] Traditionally, liver biopsy has been considered the "gold standard" reference method

for the diagnosis of liver fibrosis. However, low patient compliance, prolonged procedure time, and the risk of complications are limiting its use in large populations.[2] Several noninvasive tests have been recently developed to overcome the limitations of liver biopsies.

The shearwave elastography (SWE) techniques are designed to measure liver stiffness (LS), which has been established as a surrogate marker of liver fibrosis.[3] In many institutions, transient elastography (TE) has become the standard method for noninvasive measurement of liver fibrosis. In recent guidelines on the treatment of HCV infection, noninvasive elastography techniques are recommended for the initial assessment of liver, with liver biopsy reserved for cases where there is uncertainty or potential additional etiologies.[4] Point shearwave elastography (pSWE) and two-dimensional shearwave elastography (2D-SWE) combine SWE with conventional B-mode sonograms.

Some studies have already demonstrated that reliability and diagnostic performance of SWE are identical and occasionally even superior to TE for the assessment of liver fibrosis.[5] The S-Shearwave elastography application has recently been made available on Samsung ultrasound (US) scanners.[6] In the present study, we have assessed interobserver reproducibility of S-Shearwave measurements in a group of healthy control subjects and patients with various grades of liver fibrosis.

2 | PATIENTS AND METHODS

2.1 | Study population

The present study has been approved by the institutional ethics committee of our university, and written informed consent was obtained from all subjects according to the Declaration of Helsinki. Between October 2016 and June 2017, we prospectively enrolled 59 adult subjects, which comprised two cohorts: a group of 10 healthy volunteers without known liver disease and a group of 49 CLD patients, who were referred for a regular follow-up from the outpatient clinic. The most common etiology in the patient cohort was chronic hepatitis C with 33 cases. Clinical and demographic data of the study subjects are summarized in Table 1.

2.2 | Liver US and SWE

The Samsung RS80A Prestige US scanner (Samsung Medison, Hongcheon, Korea) equipped with the CA1-7A convex probe was used for all examinations. All pSWE measurements were performed with the S-Shearwave v3.0 application, following the manufacturer's recommended protocol (Figure 1). Briefly, subjects were fasting for at least 4 hours before the examination. The right lobe of the liver was visualized through an intercostal view. The operator positioned the pSWE measurement box near the centerline of the ultrasound beam, at least 1.5 cm from the liver capsule but not deeper than 6 cm from the skin surface. LS measurements were taken during a breath held at mid-inspiration. The median LS value was reported in m/s. 2D-SWE measurements were performed with the S-Shearwave Imaging application in a subset of the patients. Briefly, a color-coded elastogram of the liver parenchyma was

TABLE 1 Summary of demographic and clinical characteristics of the healthy control and patient groups

	Controls, *n* = 10	Patients, *n* = 49
Sex		
Male	5 (50%)	22 (45%)
Female	5 (50%)	27 (55%)
Age[a]	34 (26-43)	62 (58-67)
BMI[a]	22.9 (22-29)	25.5 (23-28)
Etiology		
Chronic hepatitis C	N/A	34 (69%)
Chronic hepatitis B	N/A	2 (4%)
PSC-PBC	N/A	3 (6%)
NASH	N/A	2 (4%)
Hemochromatosis	N/A	1 (2%)
Alcoholic liver disease	N/A	2 (4%)
Unknown etiology	N/A	5 (10%)
Fibrosis grade[b]		
≤F1	N/A	10 (20%)
F2	N/A	8 (16%)
F3	N/A	9 (18%)
F4	N/A	22 (45%)

Abbreviations: BMI, body mass index; N/A, nonapplicable; NASH, nonalcoholic steatohepatitis; PSC-PBC, primary sclerosing cholangitis—primary biliary cirrhosis; and TE, transient elastography.
[a]Age (in years) and BMI (in kg/m^2) values are presented as the median and the interquartile range.
[b]Fibrosis grade was determined with TE as the reference, with cut-off values as follows: F2 ≥ 7.1 kPa; F3 ≥ 9.5 kPa; F4 ≥ 12.5 kPa[8].

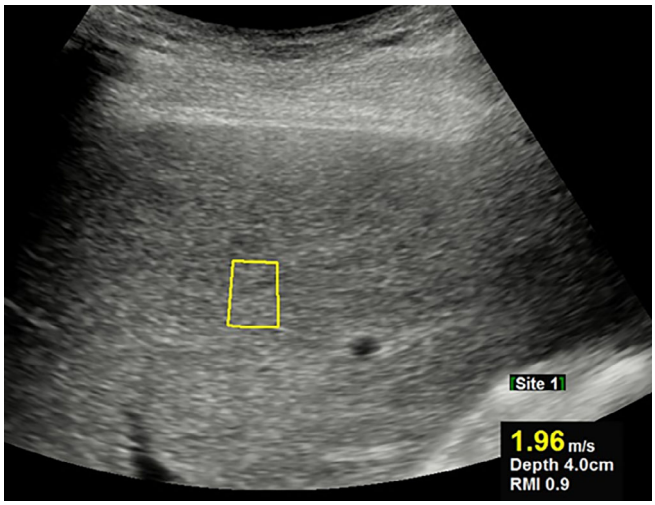

FIGURE 1 RMI is used to select reproducible liver stiffness measurements during S-Shearwave elastography. During pSWE, the measurement box is freely positioned by the operator on the B-mode grayscale image; only stiffness values with a good reliable measure index (RMI ≥0.4) are considered reliable. pSWE, point shearwave elastography; RMI, reliable measure index

obtained with a convex probe, and the average LS was recorded in circular ROIs, which were approximately 1 cm in diameter (Figure 2). The ROI was positioned where the reliable measure index (RMI) was the highest based on a color-coded map (Figure 3).

2.3 | Definition of technical success and reliable measurement

From each subject, at least 10 LS data were collected during pSWE. The RMI was automatically assigned by S-Shearwave to all measurements on a scale from 0.0 to 1.0, with 0.1 increments. Based on the manufacturer's recommendation, only data points with RMI values ≥0.4 were considered reproducible. Measurements were only technically successful if at least five valid data points could be collected. The interquartile range (IQR) of the valid data points also should remain less than 30% of the median (med.), and measurements with IQR/med. ≥30% were considered unreliable.

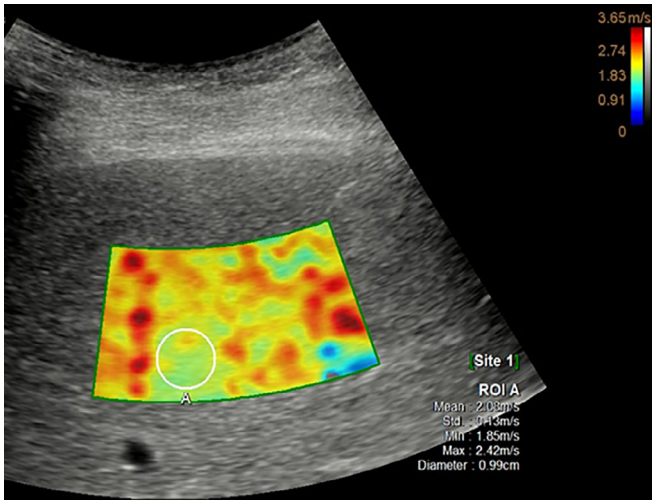

FIGURE 2 Example on a 2D-SWE measurement with the S-Shearwave imaging application. A color-coded map of liver elasticity is obtained, and the average stiffness value is measured in the circular region of interest selected by the operator. 2D-SWE, two-dimensional shearwave elastography

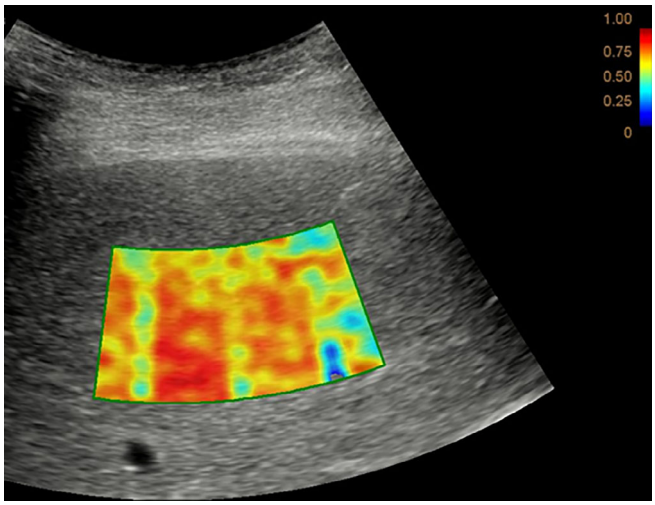

FIGURE 3 A color-coded map of the RMI is used for the ROI selection during 2D-SWE. The measurement ROI (as shown in Figure 2) is preferably placed by the operator at an area where the signal-to-noise ratio is the highest (red color) according to the RMI map. 2D-SWE, two-dimensional shearwave elastography; RMI, reliable measure index; ROI, region of interest

2.4 | Interobserver comparison

In all subjects, pSWE was performed by two different observers on the same occasion. In 18 subjects, a 2D-SWE measurement of the LS was also performed by one of the examiners. The observers were blinded to each other's measurements and the patient's prior TE results.

2.5 | Statistical analysis

The statistical analysis was completed with the R x64 v3.4.1 statistical package (www.r-project.org). The limit of statistical significance was set at $P < .05$. The intraclass correlation coefficient (ICC) was calculated for LS values, and the Cohen's kappa statistic was calculated for agreement on the stages of liver fibrosis (METAVIR F0-F4), where prior TE results were used as a reference. Good interobserver agreement was assumed in these tests when coefficient values were ≥0.75. Diagnostic performance and cut-off LS values of significant liver fibrosis were calculated from measurements from both observers using receiver operating characteristic (ROC) curve analysis.[7]

3 | RESULTS

3.1 | Comparison of the control and patient groups

The multivariate analysis of variance test was used to compare demographic and physiologic variables between the control and patient groups. There was no significant difference in average weight (78 kg, 95% confidence interval [CI]: 72-83 vs 69 kg, 95% CI: 59-80) or height (168 cm, 95% CI: 164-171 vs 173 cm, 95% CI: 163-183) between patients and control subjects, respectively. The median age (62 years, interquartile range: 58-67 vs 34 years, range: 26-43; $P < .001$) and BMI (25.5 kg/cm^2, range: 23-28 vs 22.9 kg/cm^2, range: 21.9-24; $P < .05$) were both significantly higher in the patient group (Table 1). A multiple linear regression analysis of the demographic variables found that the BMI ($\beta = .07$, 95% CI: 0.04-0.09, $P < .001$) and the measurements' depth ($\beta = .25$, 95% CI: 0.08-0.41, $P < .01$) are significant positive predictors of LS ($R^2 = 0.33$, $F[2, 85] = 19.56$, $P < .001$). Two of the enrolled subjects had ascites on US imaging. In two patients, a circumscribed lesion was detected in the liver, and a follow-up imaging study confirmed the diagnosis of HCC.

3.2 | Technical success rate and reliability of S-Shearwave measurements

The overall technical success rate with pSWE was 95% (112/118), considering all cases evaluated by the two examiners. In 4 patients

(6.7%, 4/59), either one (2/59) or both examiners (2/59) failed to collect at least five valid data points; these were excluded from further analysis (Table 2). In 2 cases, the cause of the technical failure was the patient's low tolerance of breath holds. The BMI was ≥30 in 3 out of 4 failed cases compared with 6 among the remaining subjects with successful measurements (χ^2 = 11.8, P < .05). The 2D-SWE was technically successful in all 16 examined patients.

The two examiners achieved a success rate of 80% (1104/1379) in collecting reliable (RMI ≥0.4) data points from the 59 subjects. The square of measurement depth from the skin surface (β = −.076 95% CI: −0.14–0.014) was a significant predictor of low RMI in a quadratic regression model (R^2 = 0.43, (F[2, 81] = 30.58, P < .001) (Figure 4).

3.3 | Validation of interobserver reliability and inter-method reproducibility

Both observers recorded significantly (P < .001) increased median LS in the patients (observer 1:1.9 m/s, IQR: 1.56-2.26 m/s; observer 2: 1.96 m/s, IQR: 1.54-2.37 m/s) compared to healthy controls (observer 1: 1.29 m/s, range: 1.27-1.37 m/s; observer 2: 1.36 m/s, range: 1.25-1.42 m/s).

TABLE 2 Interobserver comparison of liver stiffness, technical success rate, and diagnostic performance for S-Shearwave measurements

	Observer 1	Observer 2	Significance*
Measurement depth (in cm)[a]			
Controls	4.19 ± 0.51	3.92 ± 0.49	P = .24
Patients	4.54 ± 0.74	4.61 ± 0.7	P = .65
LS (in m/s)[a]			
Controls	1.33 ± 0.12	1.34 ± 0.12	P = .75
Patients	2.02 ± 0.60	2.01 ± 0.61	P = .90
Technical success rate	95% (56/59)	95% (56/59)	N/A
Reliable measurements[b]	98% (58/59)	88% (52/59)	P = .03
IQR/med.	12.9% ± 0.06%	17.4% ± 0.11%	P = .007
AUROC for ≥F2 fibrosis[c]	0.97(0.92-1)	0.94 (0.86-1)	P = .1
Cut-off[d]	1.5	1.45	N/A
Sensitivity	96.3%	88.9%	N/A
Specificity	93.8%	81.2%	N/A
PPV	95.5%	88.9%	N/A
NPV	96.3%	81.3%	N/A

Abbreviations: AUROC, area under receiver operating characteristic curve; CI, confidence interval; IQR, interquartile range; LS, liver stiffness; med., median; N/A, nonapplicable; NPV, negative predictive value; PPV, positive predictive value.

*Cut-off for statistical significance at P < .05.

[a]The depth of measurement and LS are shown as mean ± SD.

[b]IQR/med. ≤30%.

[c]ROC curve calculated for ≥F2 grade HCV-related fibrosis, AUROC, and 95% CI.

[d]Cut-off value (m/s).

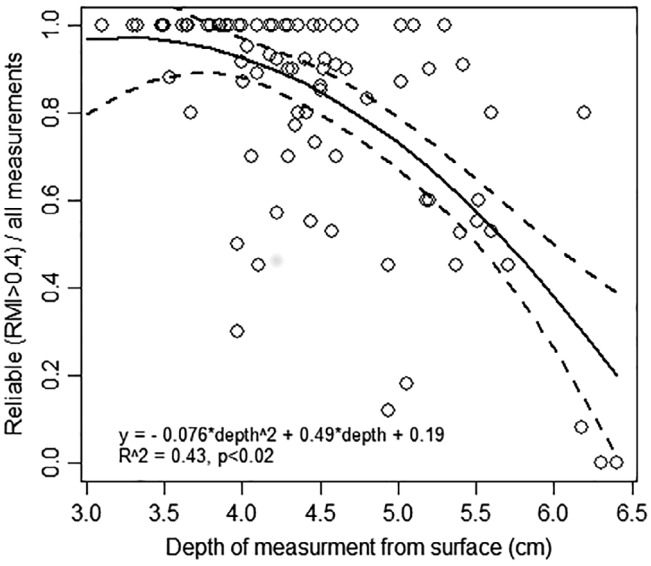

FIGURE 4 The ratio of reliable to all LS measurements shows an inverse relationship with the depth of the measurement. A downward-sloping quadratic regression line can be fitted on the scatter plot, which indicates an exponential decrease in the ratio of reliable measurement by increasing the depth of the measurement box from the skin surface. The dashed lines represent the 95% CI. CI, confidence interval; LS, liver stiffness

The Bland-Altman plot did not indicate significant measurement bias (Figure 5). A Cohen's kappa analysis (kappa = 0.76; 95% CI: 0.67-0.85, P < .001) showed good reproducibility of the fibrosis grade as well. The ICC test found excellent agreement between LS values measured by the two examiners (ICC = 0.92, 95% CI: 0.86-0.95, P < .001). We also found excellent agreement between pSWE and 2D-SWE values as measured by one of the observers (ICC = 0.91, 95% CI: 0.79-0.94, P < .001) in 18 subjects.

After the LS values from pSWE were transformed into kPa units, pSWE was directly compared with TE. Good to excellent reproducibility was found between the methods when the results from the two observers were compared separately (ICC = 0.91, 95% CI: 0.84-0.95, P < .001 or ICC = 0.79, 95% CI: 0.65-0.87, P < .001) (Figure 6). A high degree of co-linearity was also found with TE in the case of both the first (rho = 0.91, 95% CI: 0.82-0.95, P < .001) and the second examiner (rho = 0.8, 95% CI: 0.68-0.88, P < .001) (Figure 7).

3.4 | Staging of HCV-associated liver fibrosis and cut-off LS values

The pSWE measurements from 33 patients with chronic HCV hepatitis and the control group were pooled together to create a prediction set of 86 LS values. The grading of liver fibrosis was based on previous TE with cut-offs at F2 ≥ 7.1 kPa, F3 ≥ 9.5 kPa, and F4 ≥ 12.5 kPa,

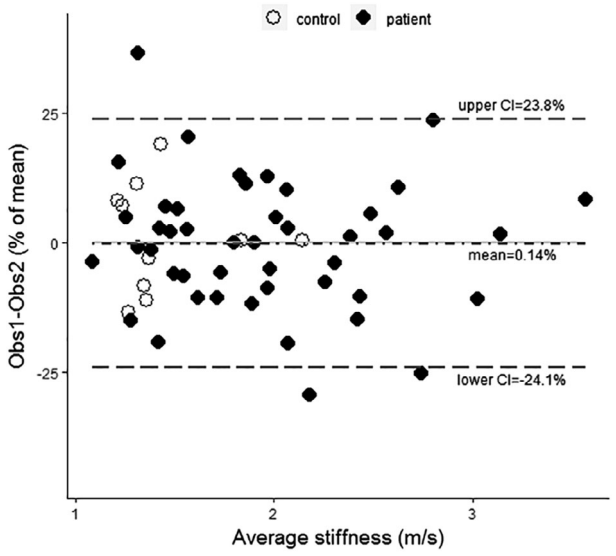

FIGURE 5 Analysis of the interobserver agreement with S-Shearwave elastography. A Bland-Altman plot shows that there is no systemic bias between pSWE measurements of the two independent observers. The mean of LS differences (dash-dotted line) was not significantly different from the baseline (solid line). Dashed lines represent upper and lower limits of 95% interobserver agreement. LS, liver stiffness; pSWE, point shearwave elastography

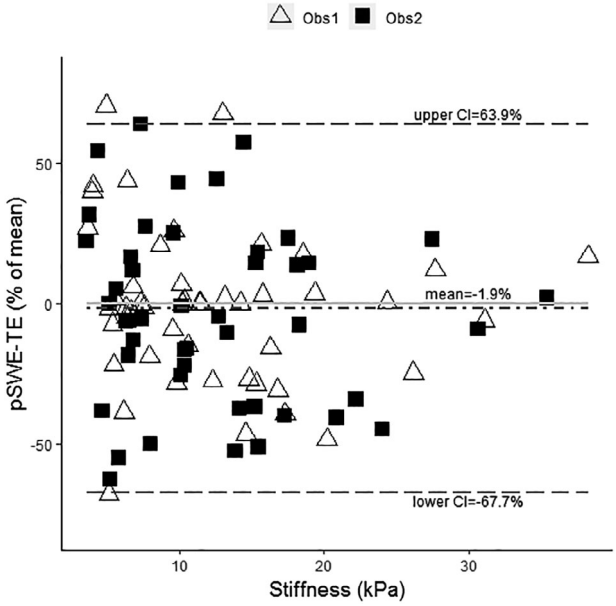

FIGURE 6 A good intermethod agreement is found between S-Shearwave and transient elastography. The Bland-Altman plot shows good agreement between the measurements of both observers and the TE results. Notably, the disagreement between the methods is relatively greater in the low LS range. The stiffness values from pSWE were converted to kPa units before the comparison. LS, liver stiffness; pSWE, point shearwave elastography; TE, transient elastography

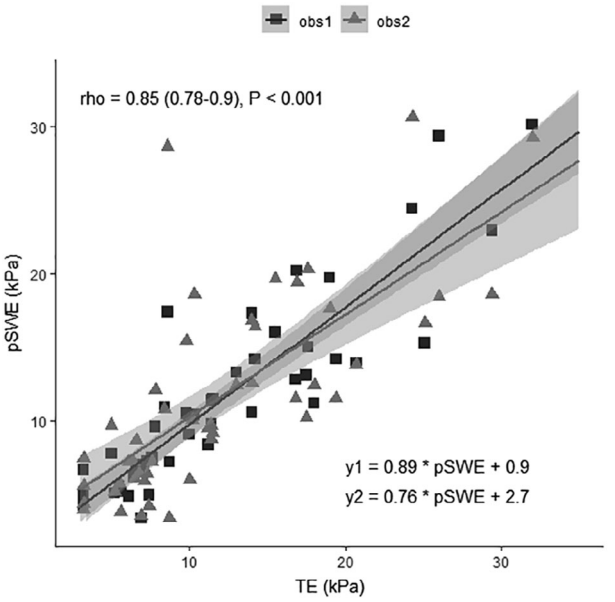

FIGURE 7 S-Shearwave measurements show very good correlation with transient elastography. A high degree of positive linear correlation was found between pSWE and TE (Spearman's rho = 0.85). The solid lines represent the predicted TE values for each observer, while the gray-shaded margins correspond to 95% confidence intervals (CI). CI, confidence interval; pSWE, point shearwave elastography; TE, transient elastography

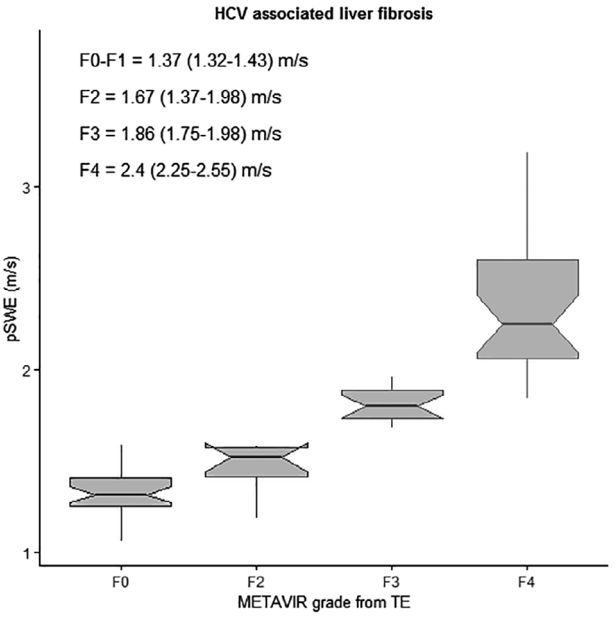

FIGURE 8 Staging of HCV-associated liver fibrosis with S-Shearwave elastography. The boxplots show the range of stiffness values measured with pSWE in the consecutive stages of liver fibrosis. The solid line inside the box represents the median; the whiskers represent the 150% outliers of the interquartile range. HCV, hepatitis C virus; pSWE, point shearwave elastography

respectively.[8] The distribution of LS values measured by pSWE was significantly different between the METAVIR categories in all comparisons ($P < .05$) (Figure 8).

The ROC analysis showed a good classification rate between cases with nonsignificant grade (F0 and F1) and moderate fibrosis (F2) with area under the ROC curve (AUROC = 0.72, 95% CI: 0.5-0.94, $P < .001$)

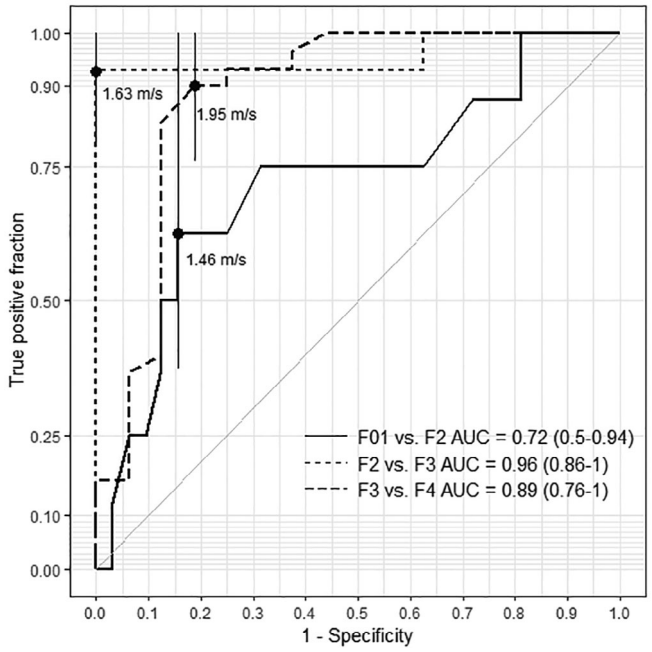

FIGURE 9 ROC analysis of the diagnostic accuracy of S-Shearwave measurements. The ROC curves show the diagnostic performance of pSWE for increasing grades of liver fibrosis (solid, dashed, and long-dashed lines) in chronic hepatitis C virus infection. The accuracy of the method is better for higher grades of fibrosis as indicated by the AUC of the curves. The dots mark the optimal cut-off value with the highest Youden index on each curve, and whiskers represent the 95% CI of the curve. AUC, area under the ROC curve; pSWE, point shearwave elastography; ROC, receiver operating characteristic

being significantly different from the line of no discrimination (Figure 9). The correct classification rate was even better for moderate (F2) vs severe fibrosis (AUROC = 0.96, 95% CI: 0.86-1, $P < .001$) and between severe fibrosis (F3) and cirrhosis (F4; AUROC = 0.89, 95% CI: 0.76-1, $P < .001$). In a three-way ROC analysis, the subjects were divided into three classes (no-significant fibrosis [F0 or F1], significant fibrosis [F2 or F3], and cirrhosis) and the volume under the surface (VUS = 0.82, 95% CI: 0.71-0.93, $P < .001$) indicated a very good diagnostic performance. The cut-off LS values with the highest Youden index were also calculated for ≥F2 (1.46 m/s; specificity (sp) = 84%; sensitivity (se) = 62%; negative predictive value (npv) = 90%; positive predictive value (ppv) = 50%), ≥F3 (1.63 m/s; sp = 100%; se = 93%; npv = 89%; ppv = 100%), and F4 (1.95 m/s; sp = 81%; se = 90%; npv = 81%; ppv = 90%) grade liver fibrosis.

Both observers achieved excellent prediction rate of ≥F2 liver fibrosis (AUROC = 0.97, 0.92-1 vs 0.94 95% CI: 0.86-1), while comparison of the ROC curves did not find a significant difference ($P = 0.1$). The cut-off LS value for the first observer was nearly identical to that of the second observer (1.5 m/s vs 1.45 m/s) (Table 2). The post hoc power calculation indicated that F0-F1 vs F2 classification had greater than 20% type II error rate, otherwise the size of the cohort was sufficient to achieve a minimum of 80% power in all of the ROC classifications.

4 | DISCUSSION

Morphological imaging techniques, such as B-mode US and computed tomography, have low sensitivity for diagnosing liver fibrosis. These modalities can only diagnose the cirrhotic stage when overt signs of architectural distortion and portal hypertension are present.[9] Other methods such as magnetic resonance elastography are only applicable to preselected patient groups.[10]

According to recent guidelines, noninvasive markers including elastography are recommended for the initial assessment of liver fibrosis in both HBV- and HCV-associated liver diseases, while a liver biopsy is only preferred in selected cases, where there is uncertainty or potential additional etiologies.[4,11] SWE, including pSWE and 2D-SWE, are relatively new techniques, which can be performed on any patient during a routine liver US examination.[12] In our opinion, SWE arguably has the potential to become a universally accepted method for the assessment and follow-up of liver fibrosis.

Therefore, SWE can be easily integrated into HCC surveillance protocols. Notably, in two cases, a tumor was also found during the elastography scan, while in an additional case the LS measurement was successful even in the presence of a significant amount of ascites.

Although the operating principle is similar for all SWE systems, the diagnostic accuracy and cut-off values may be different based on the manufacturer.[13] Considering the potentially large number of operators, validation of these applications is essential to maintain diagnostic performance. There was no technically failed measurements in the control group, which highlights the straightforward operation of S-Shearwave.

In agreement with current guidelines, we acquired at least 10 data points in each subject.[3] Each of the data points was tagged with a signal-to-noise indicator RMI. This is calculated from the wave equation residuals and shear wave magnitude by the S-Shearwave application.[14] The decreased signal-to-noise ratio results in low RMI reading. In our experience, RMI has been useful to identify a subset of cases where pSWE is technically challenging, and additional data points need to be collected for reliable measurement. The cause of technical failure in two cases was that elderly patients did not tolerate breath holds. Previous studies have found similar technical success rate with other pSWE scanners.[15,16] A prior study reported one failed (3%) and two unreliable measurements (6%) in 33 subjects examined with S-Shearwave.[6] The authors also claimed that obesity was the primary cause of technical failure in these cases. The BMI was >30 in nine of our patients; among them three had failed pSWE measurements. Measurement depth also proved to be a significant negative predictor of low RMI reading and technical failure in a linear quadratic model. Thus, we have demonstrated in vivo that the reliability of LS measurements exponentially decreases with increasing distance of ROI from the skin surface. These findings are also in agreement with phantom studies, which demonstrated higher coefficient variations with increasing probe-to-ROI distance.[17]

The intra- and interobserver reproducibility values of various pSWE scanners have been extensively validated in previous studies.[6,15] We found excellent interobserver agreement (ICC = 0.92), which surpassed other platforms tested on healthy subjects (ICC = 0.86) or liver patients (ICC = 0.84

and ICC = 0.798).[18,19] One previous publication has also reported an excellent interobserver agreement (ICC > 0.9) in a smaller patient group with the S-Shearwave application.[6] The intermethod agreement between pSWE and 2D-SWE (ICC = 0.91) was also excellent, although the relatively low number of subjects limits the power of the analysis.

TE is accepted as a surrogate of liver biopsy for the diagnosis of liver fibrosis according to the international guidelines on the diagnosis and treatment of chronic hepatitis.[1-3] Therefore it is quite reassuring that S-Shearwave elastography showed a high degree of correlation (rho = 0.85) and very good agreement on LS values (ICC = 0.86) with TE measurements.

In our study, S-Shearwave demonstrated excellent diagnostic accuracy in the detection of clinically significant (≥F2 grade) liver fibrosis in the hands of both operators. The high AUROC values of 0.94 and 0.97 are well comparable to other reports assessing pSWE. S-Shearwave also showed very good performance in the detection of liver cirrhosis (AUROC 0.89). The three-way ROC analysis performs simultaneous detection and classification of the cases.[20] pSWE achieved a very good classification rate (VUS = 0.82) when nonsignificant fibrosis, fibrosis, and cirrhosis were simultaneously compared in a multiclass decision model. These findings suggest that S-Shearwave can be an excellent tool for screening of liver fibrosis. A post hoc power calculation was performed and it showed that our cohort size was sufficient to achieve a less than 0.2 type II error rate in all of the classification steps with one exception. The ROC analysis was underpowered only for the F0-1 vs F2 prediction. In our opinion, it does not change the basic interpretation of our findings as cases with significant fibrosis (≥F2) could be classified by both observers with a minimum power of 80%. Moreover, all of the ROC curves were also statistically significant ($P < .05$).

Our study has several limitations. Firstly, fibrosis grade in the patient group was estimated by TE, and liver biopsy was not available for comparison with pSWE. Our study design is still sound, as TE is a well-established surrogate of liver biopsy for grading fibrosis according to current clinical protocols. Patients could also be spared from unnecessary risk of complication from a liver biopsy. Secondly, the number of subjects was relatively small compared to larger multicenter studies. Also, high grades of fibrosis (F3 and F4) were overrepresented among the patients. Therefore, investigation of larger cohorts of patients with a more uniform clinical background is necessary to formulate comprehensive diagnostic recommendations for S-Shearwave.

In conclusion, S-Shearwave elastography utilizes a user-friendly quality index the RMI for the selection of reliable LS measurements. Both interobserver agreement intermethod reproducibility was excellent with S-Shearwave, and the method showed very good overall performance in the staging of liver fibrosis.

ACKNOWLEDGMENTS

Pal Novák Kaposi, M.D., Ph.D., was supported by the János Bolyai research scholarship of the Hungarian Academy of Sciences.

ORCID

Pál N. Kaposi https://orcid.org/0000-0002-7150-3495

REFERENCES

1. EASL-ALEH. Clinical practice guidelines: non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol*. 2015; 63:237.
2. Barr RG, Ferraioli G, Palmeri ML, et al. Elastography assessment of liver fibrosis: Society of Radiologists in Ultrasound Consensus Conference Statement. *Radiology*. 2015;276:845-861.
3. 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:e16.
4. Pawlotsky JM, Negro F, Aghemo A, et al. EASL Recommendations on treatment of hepatitis C. *J Hepatol*. 2018;69:461-511.
5. Bamber J, Cosgrove D, Dietrich CF, et al. EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. Part 1: basic principles and technology. *Ultraschall Med*. 2013;34(2): 169-184.
6. Ahn SJ, Lee JM, Chang W, et al. Prospective validation of intra- and interobserver reproducibility of a new point shear wave elastographic technique for assessing liver stiffness in patients with chronic liver disease. *Korean J Radiol*. 2017;18(6):926-935.
7. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
8. Castera L, Vergniol J, Foucher J, et al. Prospective comparison of transient elastography, fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology*. 2005;128(2): 343-350.
9. Mueller S, Sandrin L. Liver stiffness: a novel parameter for the diagnosis of liver disease. *Hepat Med*. 2010;2:49-67.
10. Venkatesh SK, Yin M, Ehman RL. Magnetic resonance elastography of liver: technique, analysis, and clinical applications. *J Magn Reson Imaging*. 2013;37(3):544-555.
11. EASL. 2017 clinical practice guidelines on the management of hepatitis B virus infection. *J Hepatol*. 2017;67(2):370-398.
12. 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.
13. Mulabecirovic A, Vesterhus M, Gilja OH, Havre RF. In vitro comparison of five different elastography systems for clinical applications, using strain and shear wave technology. *Ultrasound Med Biol*. 2016;42 (11):2572-2588.
14. Kiwan C, Donggeon K, Zaegyoo H, Hyoung-Ki L. A reliability index of shear wave speed measurement for shear wave elastography. Paper presented at: 2015 IEEE International Ultrasonics Symposium (IUS); 21–24 Oct. 2015.
15. Bota S, Herkner H, Sporea I, et al. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. *Liver Int*. 2013;33(8):1138-1147.
16. Yoo H, Lee JM, Yoon JH, Lee DH, Chang W, Han JK. Prospective comparison of liver stiffness measurements between two point shear wave elastography methods: virtual touch quantification and elastography point quantification. *Korean J Radiol*. 2016;17(5):750-757.
17. Kishimoto R, Suga M, Koyama A, et al. Measuring shear-wave speed with point shear-wave elastography and MR elastography: a phantom study. *BMJ Open*. 2017;7(1):e013925.
18. 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.
19. Wang CZ, Zheng J, Huang ZP, et al. Influence of measurement depth on the stiffness assessment of healthy liver with real-time shear wave elastography. *Ultrasound Med Biol*. 2014;40(3):461-469.
20. He X, Frey EC. The meaning and use of the volume under a three-class ROC surface (VUS). *IEEE Trans Med Imaging*. 2008;27(5): 577-588.

How to cite this article: Kaposi PN, Unger Z, Fejér B, et al. Interobserver agreement and diagnostic accuracy of shearwave elastography for the staging of hepatitis C virus-associated liver fibrosis. *J Clin Ultrasound*. 2019;1–8. https://doi.org/10.1002/jcu.22771