



# Reproducibility of liver stiffness measurements made with two different 2-dimensional shear wave elastography systems using the comb-push technique

ULTRA  
SONO  
GRAPHY

Hwaseong Ryu<sup>1</sup>, Su Joa Ahn<sup>2</sup>, Jeong Hee Yoon<sup>3</sup>, Jeong Min Lee<sup>3,4,5</sup>

<sup>1</sup>Department of Radiology, Pusan National University Yangsan Hospital, Yangsan;

<sup>2</sup>Department of Radiology, Gachon University Gil Medical Center, Incheon; <sup>3</sup>Department of Radiology, Seoul National University Hospital, Seoul; <sup>4</sup>Department of Radiology, Seoul National University College of Medicine, Seoul; <sup>5</sup>Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul, Korea

**Purpose:** The purpose of this study was to retrospectively compare the technical success and reliability of the measurements made using two 2-dimensional (2D) shear wave elastography (SWE) systems using the comb-push technique from the same manufacturer and to assess the intersystem reproducibility of the resultant liver stiffness (LS) measurements.

**Methods:** Ninety-four patients with suspected chronic liver diseases were included in this retrospective study. LS measurements were obtained using two 2D-SWE systems (LOGIQ E9 and LOGIQ S8) from the same manufacturer, with transient elastography (TE) serving as the reference standard, on the same day. The technical success rates and reliability of the measurements of the two 2D-SWE systems were compared. LS values measured using the two 2D-SWE systems and TE were correlated using Spearman correlation coefficients and 95% Bland-Altman limits of agreement. Thereafter, Bland-Altman limits of agreement and intraclass correlation coefficients (ICCs) were used to analyze the intersystem reproducibility of LS measurements.

**Results:** The two 2D-SWE systems showed similar technical success rates (98.9% for both) and reliability of LS measurements (92.3% for the LOGIQ E9, 91.2% for the LOGIQ S8;  $P=0.185$ ). Despite the excellent correlation ( $ICC=0.92$ ), the mean LS measurements obtained by the two 2D-SWE systems were significantly different (LOGIQ E9,  $6.57\pm 2.33$  kPa; LOGIQ S8,  $6.90\pm 6.64$  kPa;  $P=0.018$ ).

**Conclusion:** Significant intersystem variability was observed in the LS measurements made using the two 2D-SWE systems. Therefore, even 2D-SWE systems from the same manufacturer should not be used interchangeably in longitudinal follow-up.

**Keywords:** Ultrasonography; Sonoelastography; Reproducibility of results; Liver cirrhosis

## ORIGINAL ARTICLE

<https://doi.org/10.14366/usg.18046>  
pISSN: 2288-5919 • eISSN: 2288-5943  
Ultrasonography 2019;38:246-254

Received: August 12, 2018  
Revised: November 21, 2018  
Accepted: November 23, 2018

**Correspondence to:**  
Jeong Min Lee, MD, Department of Radiology, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea

Tel. +82-2-2072-3154  
Fax. +82-2-743-6385  
E-mail: jmsh@snu.ac.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2019 Korean Society of Ultrasound in Medicine (KSUM)



**How to cite this article:**  
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 Jul;38(3):246-254.

## Introduction

Chronic liver disease (CLD) is a well-known worldwide problem, and common etiologies of CLD include viral hepatitis (hepatitis B or C), alcohol abuse, and non-alcoholic fatty liver disease. Regardless of the cause of CLD, it is important to determine the stage of liver fibrosis in order to identify the most suitable treatment for preventing progression to liver cirrhosis [1–3]. So far, liver biopsy has been the gold standard for evaluating liver fibrosis. However, it has several limitations, such as sampling variability, complications due to invasiveness, and intraobserver or interobserver variability in fibrosis staging [4–8].

As an alternative approach to liver biopsy for hepatic fibrosis staging, various ultrasound-based shear wave elastography (SWE) techniques have been developed, including transient elastography (TE) (FibroScan, Echosens, Paris, France) using a mechanical push, point SWE (p-SWE), and 2-dimensional SWE (2D-SWE) [9–12]. Currently, many p-SWE and 2D-SWE techniques are commercially available, and they share the principle of using an ultrasound-induced acoustic radiation force impulse to generate shear waves [13,14]. In clinical practice, both the diagnostic performance of different SWE systems for fibrosis staging and their reproducibility are important issues for detecting severe fibrosis and longitudinal monitoring of the therapeutic response [15,16]. Several previous studies on the reproducibility of various SWE techniques have recommended that liver stiffness (LS) measurements obtained using different techniques should not be used interchangeably [11,17,18]. However, intersystem reproducibility using the same SWE technique has not yet been well evaluated.

The purpose of this study was to evaluate the intersystem reproducibility of LS measurements with 2D-SWE using the comb-push technique for the evaluation of hepatic fibrosis, with TE as the reference method.

## Materials and Methods

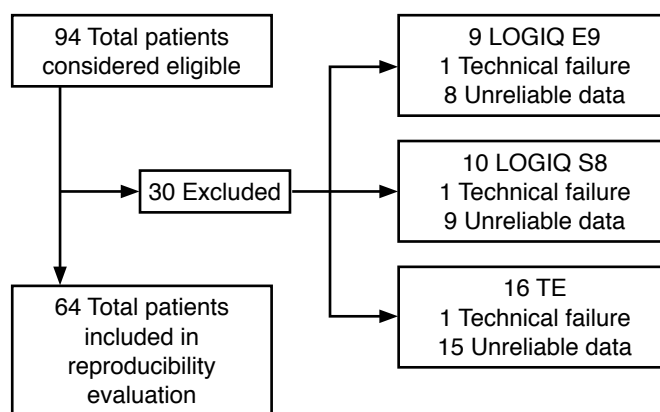
This study was performed retrospectively with approval from our institutional review board, and the requirement for informed consent was waived.

### Study Population

Between May and September 2017, 94 patients (39 men, 55 women; mean age, 57.5±12.1 years; age range, 25–80 years; median body mass index [BMI], 24.9±3.59 kg/m<sup>2</sup>; BMI range, 18.5–24.9 kg/m<sup>2</sup>) with suspected CLD who were referred to the Department of Radiology for fibrosis screening were included in this study (Fig. 1). No patients had ascites. Patients younger than

18 years and patients who had undergone right hepatectomy were not included in this study. During the study period, a LOGIQ S8 ultrasound (US) platform (GE Healthcare, Wauwatosa, WI, USA) equipped with TE and a LOGIQ E9 US platform using 2D-SWE were tested to analyze the intersystem reproducibility of LS measurements. After explaining the use of two platforms during a clinical routine liver US examination, 2D-SWE exams and TE were performed with verbal agreement from the patients.

The etiologies of CLD in the study participants were chronic



**Fig. 1.** Study design. Unreliable data were obtained from both the LOGIQ E9 and LOGIQ S8 in one patient who experienced technical failure with transient elastography (TE) (n=1), both the LOGIQ E9 and LOGIQ S8 (n=2), and both the LOGIQ E9 and TE (n=1).

**Table 1.** Patient characteristics

Parameter	Value
Age (yr)	57.5±12.1 (25 to 80)
Sex	
Male	39 (41.5)
Female	55 (58.5)
Etiology of liver disease	
Chronic hepatitis B	62 (66.0)
Chronic hepatitis C	10 (10.6)
Chronic non-viral hepatitis (NAFLD, alcoholic, PBC)	18 (19.1)
None	4 (4.3)
Fibrosis grade <sup>a)</sup>	
<F2	51 (79.7)
F2	4 (6.2)
F3	3 (4.7)
F4	6 (9.4)

Values are presented as mean±SD (range) or number (%).

NAFLD, nonalcoholic fatty liver disease; PBC, primary biliary cirrhosis; SD, standard deviation.

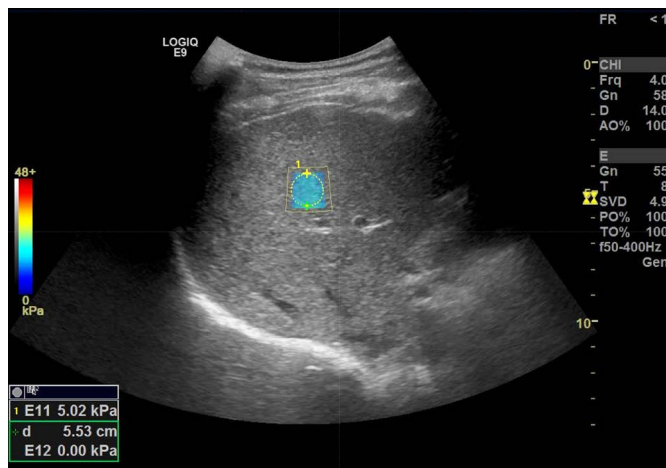
<sup>a)</sup>Distribution of liver fibrosis using transient elastography cut-off values.

hepatitis B (n=62, 66.0%); chronic hepatitis C (n=10, 10.6%); chronic non-viral hepatitis, such as non-alcoholic steatohepatitis, alcoholic liver disease, and primary biliary cirrhosis (n=18, 19.1%); and others (n=4, 4.3%) (Table 1).

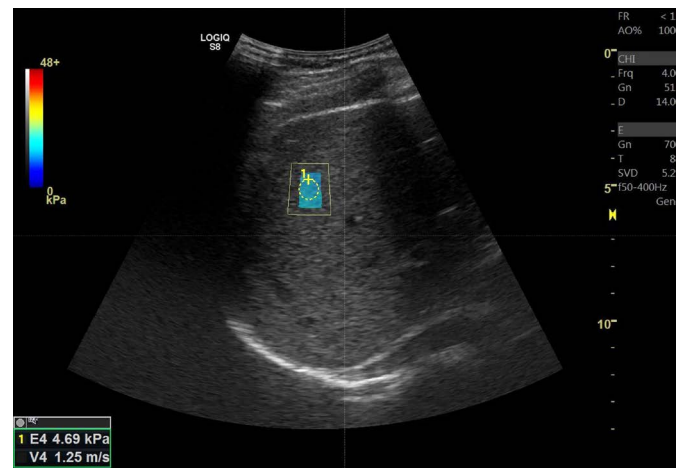
In addition, we classified the degree of liver fibrosis in patients based on LS values measured by TE, since TE is the best-validated method for evaluating liver fibrosis [19–21]. We used the LS cutoff values for TE proposed in the latest meta-analysis [20]: 7.9 kPa for moderate fibrosis ( $F \geq 2$ ), 8.8 kPa for severe fibrosis ( $F \geq 3$ ), and 11.7 kPa for liver cirrhosis ( $F = 4$ ). In the 64 patients who had reliable LS measurements from all examinations, the most common fibrosis stage was mild (F1) or no liver fibrosis (F0) (51 of 64, 79.7%), followed by liver cirrhosis (F4) (6 of 64, 9.4%), moderate (F2) liver fibrosis (4 of 64, 6.2%), and severe fibrosis (F3) (3 of 64, 4.7%) (Table 1).

**SWE Examinations**

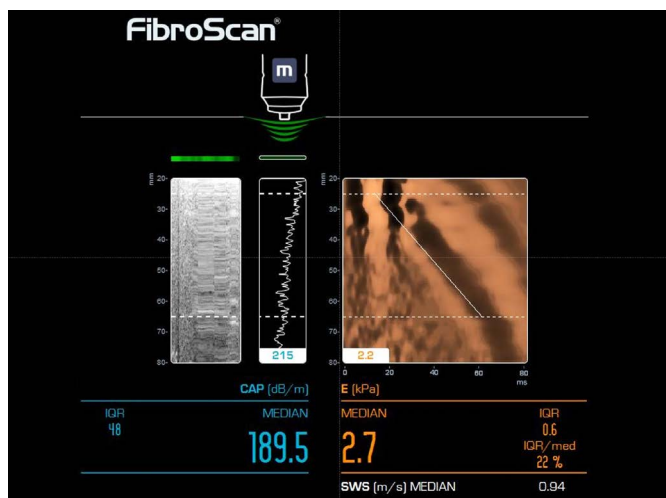
All patients underwent US examinations after fasting for more than 6 hours. All examinations were performed with LOGIQ E9 and LOGIQ S8 US systems (GE Healthcare), and both systems were equipped with convex broadband (C1-6; 1.5–6 MHz; central frequency, 4 MHz). All US examinations were performed by one radiologist, who had 6 years of experience in US-based elastography (including 2D-SWE) and TE (>200 examinations), as well as 20 years of experience with abdominal US examinations. At first, conventional B-mode sonography was performed using a 4-MHz convex probe. LS measurements were then made using the same probe with an intercostal approach while patients maintained a supine position or right anterior oblique position, with the right arm in maximum abduction during the SWE examinations. For 2D-SWE using both systems, a 1×1-cm<sup>2</sup> region of interest was placed in the right anterior segment of the liver, avoiding large vessels and areas



A



B



C

**Fig. 2.** Liver stiffness (LS) measurements of the two different 2-dimensional shear wave elastography systems and transient elastography (TE): LOGIQ E9 (A), LOGIQ S8 (B), and TE (C).

A, B. For LS measurements, the regions of interest were placed in the right anterior segment of the liver, avoiding vascular structures. The blue color in the SWE boxes is related to the speed of the shear waves. C. The slope of the line at the right panel of TE measurements indicates shear wave speed.

with artifacts, 1.5–2.0 cm away from the Glisson capsule, and less than 6 cm deep from the capsule (Fig. 2). Finally, LS measurements were made with TE (FibroScan Echosens) using the LOGIQ S8 platform on each patient. All TE examinations were performed by the same operator who performed the 2D-SWE examination. The operator, who had previously performed more than 200 TE examinations, performed all FibroScan examinations according to the manufacturer's recommendations. The tip of the transducer probe (M+ probe or XL+ probe when prompted by the automatic probe selection tool) was placed on the skin between the ribs over the right lobe of the liver, and a valid LS measurement was obtained under the guidance of M-mode monographic images [12,13,22]. During LS measurement, patients were instructed to hold breathing while avoiding deep inspiration or expiration. At least 10 valid measurements of every patient were made using each method of SWE. To reduce recall bias, the summary of the serial measurements of each system was not made available to the operator until all examinations were completed [18].

#### Definition of Technical Failure and Reliable (or Unreliable) Measurements

Technical failure of SWE was defined as the failure to acquire 10 valid measurements after at least 15 trials [18]. If the ratio of the interquartile range to the median LS was greater than 30%, the result was regarded as an unreliable measurement [23].

#### Statistical Analysis

Continuous data are presented as means and ranges, and categorical data as counts and percentages. The Friedman test was used to compare the rates of technical failure and unreliable measurements between the two 2D-SWE imaging systems and TE. For comparison of the fibrosis stages between patients with reliable and unreliable LS measurements, the Mann-Whitney test was used. The Wilcoxon signed-rank test was used to compare the LS measurements obtained by the two 2D-SWE imaging systems. The Spearman correlation coefficient and 2-way mixed model intra-class correlation coefficients (ICCs) with 95% confidence intervals (CIs) were obtained to evaluate the agreement between the two 2D-SWE systems and TE. Correlation coefficients were classified using the following definitions: 0–0.19, very weak; 0.2–0.39, weak; 0.40–0.59, moderate; 0.60–0.79, strong; and 0.80–1.0, very strong [24]. Agreement based on ICCs was classified using the following definitions: 0–0.39, poor; 0.40–0.59, fair; 0.60–0.74, good; and 0.75–1.0, excellent [25]. Bland-Altman analysis was used to evaluate method-related variations using the mean value between the two 2D-SWE systems and TE. The 95% limits of agreement, as well as the coefficient of reproducibility ( $CR=1.96 \times \text{standard}$

deviation of bias), were determined to evaluate the intersystem variability in the liver LS measurements. The coefficient of variation (CV) of the LS values between the two 2D-SWE systems was also calculated. All statistical analyses were performed using commercially available software programs version 23 (IBM Corp., Armonk, NY, USA; or MedCalc version 16, MedCalc Software, Mariakerke, Belgium). P-values less than 0.05 were considered to indicate statistical significance.

## Results

All 94 patients were included in the assessment of the technical success rate and reliability of the measurements of the two 2D-SWE systems and TE. However, the reproducibility and performance of the two 2D-SWE systems were only evaluated in the 64 patients who had reliable LS values from both 2D-SWE systems and TE. A total of 30 patients were excluded due to technical failure ( $n=3$ ) or an unreliable result from any technique ( $n=27$ ) (Fig. 1).

#### Technical Failure and Unreliable Measurement Rates

##### *Technical failure of LS measurements*

Among the 94 patients, a set of 10 LS measurements was made successfully in 91 patients (91 of 94, 97%) by both 2D-SWE systems and TE (Fig. 1). One case of LS measurement failure occurred with each of the LOGIQ E9, LOGIQ S8, and TE in different patients (1 of 94; 1.1% for each). There was no significant difference in the technical success rate between the two 2D-SWE systems and TE ( $P>0.990$ ). Regarding the elastography method, the 2D-SWE systems and TE had the same failure rate (1 of 94, 1.1%).

##### *Unreliable LS measurements*

Among patients with technically successful LS measurements, unreliable LS measurements were obtained in 8.6% (8 of 93) of the patients using the LOGIQ E9, 9.7% (9 of 93) using the LOGIQ S8, and 16.1% (15 of 93) using TE. Two patients failed to demonstrate a reliable measurement with both the LOGIQ E9 and LOGIQ S8 and one patient failed to show a reliable measurement with both the LOGIQ E9 and TE. There was no significant difference in the unreliable measurement rate among the two 2D-SWE systems and TE ( $P>0.990$ ). Among the 76 patients with reliable LS measurements by TE, there was also no significant difference in fibrosis stage between patients with reliable LS measurements and those with unreliable LS measurements ( $P=0.475$ ).

#### Overall Correlations and Intersystem Reproducibility of LS Values across 2D-SWE Systems

In the 64 patients who had reliable LS measurements from the two

2D-SWE systems and TE, the mean LS values obtained by the two 2D-SWE systems were significantly different (LOGIQ E9,  $6.57 \pm 2.33$  kPa; LOGIQ S8,  $6.90 \pm 6.64$  kPa;  $P=0.018$ ) (Table 2). In a subgroup analysis according to fibrosis grade by TE, a significant difference was found in the mean LS values between the LOGIQ E9 and LOGIQ S8 in F1 patients ( $P=0.037$ ), while no significant difference was found in

**Table 2.** Mean LS values obtained by the 2D-SWE systems and TE

	Mean LS value (kPa)	P-value <sup>a)</sup>
E9	$6.57 \pm 2.33$ (3.36 to 14.73)	0.018
S8	$7.00 \pm 2.27$ (4.37 to 15.85)	
TE	$6.90 \pm 6.64$ (2.10 to 47.90)	

LS, liver stiffness; 2D-SWE, 2-dimensional shear wave elastography; TE, transient elastography; E9, LOGIQ E9; S8, LOGIQ S8.

<sup>a)</sup>Comparison of the LOGIQ E9 and LOGIQ S8.

the mean LS values in  $\geq F2$  patients ( $P=0.131-0.223$ ) (Table 3).

The LS values for the two 2D-SWE systems demonstrated a very strong positive correlation ( $r=0.86$ ,  $P<0.001$ ). The correlations of

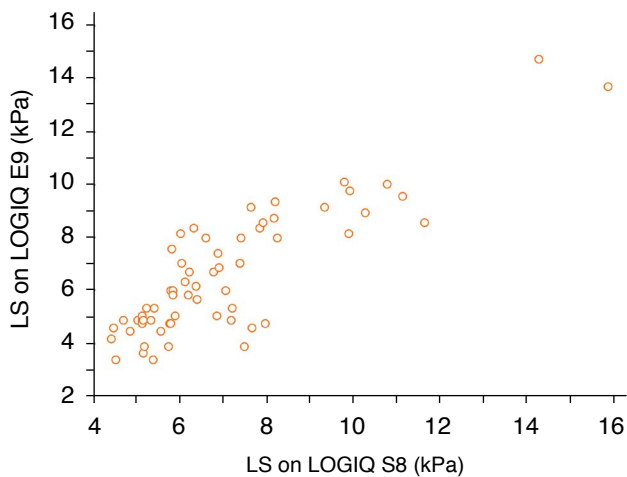
**Table 3.** Comparison of the mean LS values obtained by each 2D-SWE system and TE in subgroups of patients

	LS value			P-value <sup>a)</sup>
	E9	S8	TE	
F<2	$5.78 \pm 1.65$	$6.18 \pm 1.20$	$4.77 \pm 1.38$	0.045
F $\geq 2$	$9.35 \pm 2.29$	$9.90 \pm 2.75$	$14.40 \pm 11.24$	0.223
F $\geq 3$	$9.98 \pm 2.59$	$10.86 \pm 2.86$	$18.04 \pm 12.74$	0.161
F=4	$10.53 \pm 3.09$	$11.72 \pm 3.15$	$22.03 \pm 14.22$	0.131

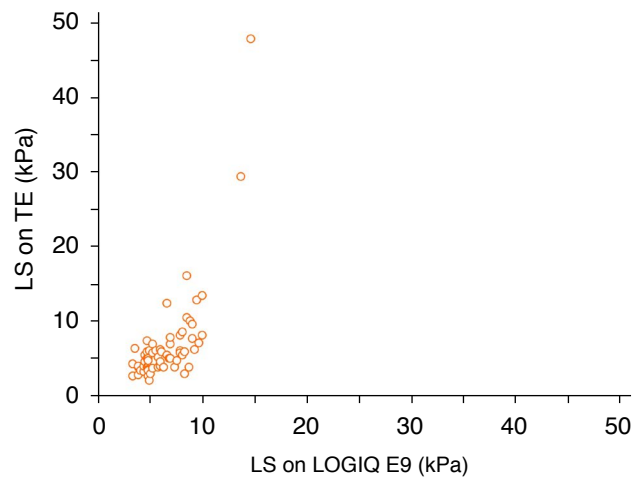
Values are presented as mean $\pm$ SD.

LS, liver stiffness; 2D-SWE, 2-dimensional shear wave elastography; TE, transient elastography; E9, LOGIQ E9; S8, LOGIQ S8; SD, standard deviation.

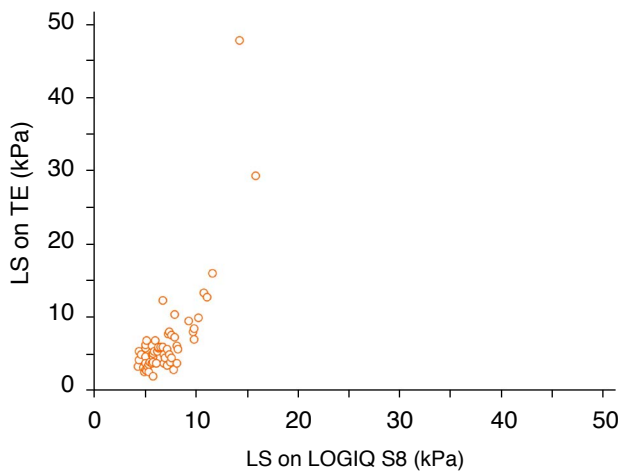
<sup>a)</sup>Comparison of the LOGIQ E9 and LOGIQ S8.



A



B



C

**Fig. 3.** Correlations between liver stiffness (LS) measurements obtained with two different 2-dimensional shear wave elastography systems and transient elastography (TE).

Scatter diagrams (A, LOGIQ E9 vs. LOGIQ S8; B, LOGIQ E9 vs. TE; and C, LOGIQ S8 vs. TE) show strong correlations in all combinations of the three systems used in this study.

the LS values between the two 2D-SWE systems and TE were also strong (Fig. 3). The ICC for the LS measurements between the LOGIQ E9 and LOGIQ S8 was 0.92, indicating excellent agreement (95% CI, 0.85 to 0.95). The CV between the two 2D-SWE systems was 15.9% (Table 4).

The Bland-Altman plots for reproducibility between TE and the two 2D-SWE systems showed a tendency for larger LS values to be obtained by the two 2D-SWE systems than by TE in patients with

a lower fibrosis grade. However, in patients with a higher fibrosis grade, smaller LS values were obtained by the two 2D-SWE systems than by TE (Fig. 4).

**Performance of the Three SWE Techniques in Detecting Significant Fibrosis (F $\geq$ 2)**

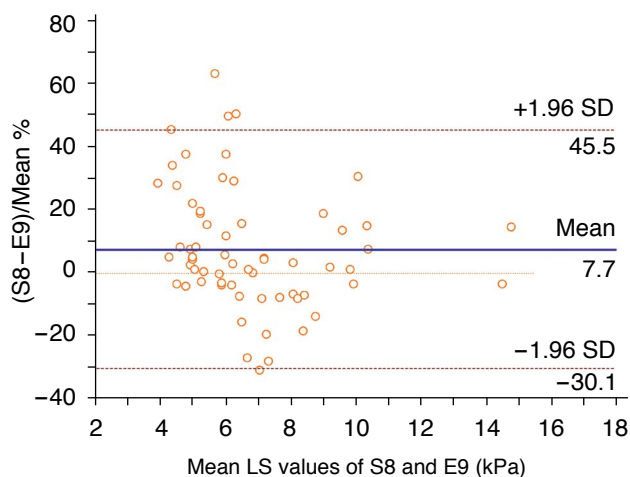
Using the LS values of TE as the reference standard, the LOGIQ E9 showed an area under the receiver operating characteristic curve

**Table 4.** Correlation of LS values between the 2D-SWE systems

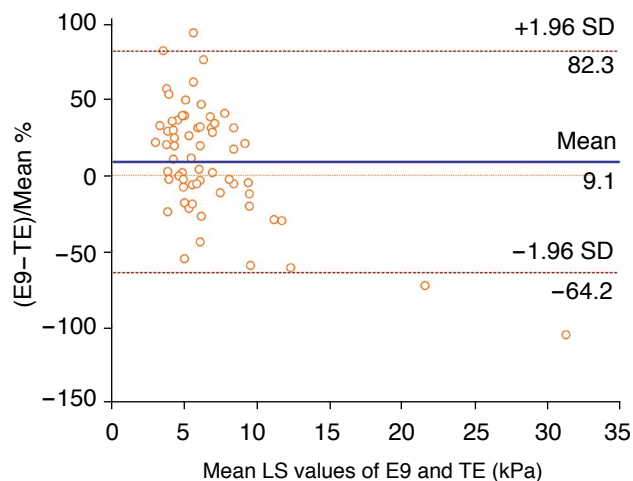
	r	ICC	CV (%)	CR (%)	Mean bias (%)	BALA
Correlation of LS value	0.86 (0.77 to 0.91)	0.92 (0.85 to 0.95)	15.93 (12.92 to 19.03)	2.54 (2.17 to 3.08)	7.66 (2.80 to 12.51)	- (-30.14 to 45.45)

The numbers in parentheses represent the 95% confidence interval.

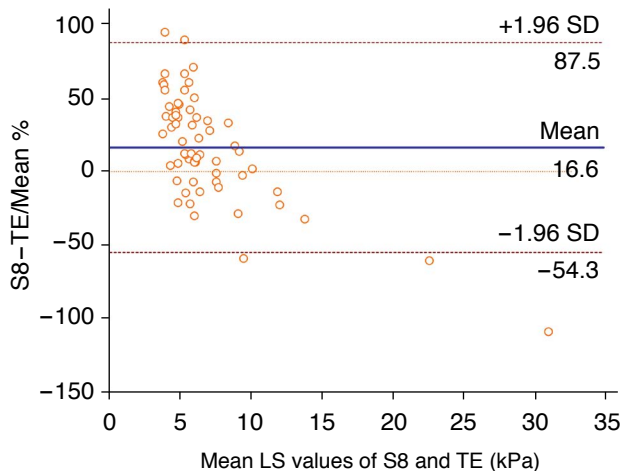
LS, liver stiffness; 2D-SWE, 2-dimensional shear wave elastography; ICC, intra-class correlation; CV, coefficient of variation (%); CR, coefficient of reproducibility (%); BALA, Bland-Altman limits of agreement (%).



A



B

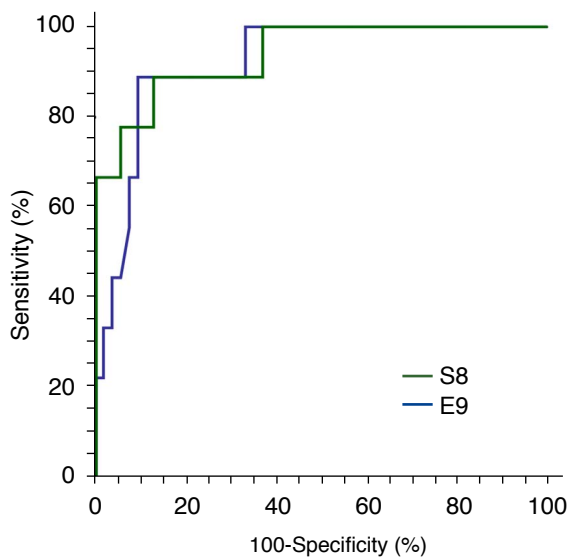


C

**Fig. 4.** Bland-Altman plot comparing liver stiffness (LS) values obtained from two different 2-dimensional shear wave elastography systems and transient elastography (TE).

The solid blue line in the middle represents the mean LS values obtained from each pair of the three systems (A, LOGIQ E9 vs. LOGIQ S8; B, LOGIQ E9 vs. TE; and C, LOGIQ S8 vs. TE), and the dotted purple line define  $\pm 1.96$  standard deviations (SD).





**Fig. 5.** Comparison of receiver operating characteristic (ROC) curves for the two different 2-dimensional shear wave elastography (2D-SWE) systems for the detection of significant fibrosis ( $F \geq 2$ ). The blue line delineates the ROC curve of the LOGIQ E9 with an area under the ROC curve (AUROC) of 0.921 ( $P < 0.001$ ), and the green line delineates the ROC curve of the LOGIQ S8, with an AUROC of 0.938 ( $P < 0.001$ ). In the pairwise ROC curve comparison, the AUROCs of the two 2D-SWE systems were not significantly different ( $P = 0.436$ ).

(AUROC) of 0.92 (95% CI, 0.83 to 0.97) using a cut-off value of  $>8.33$  kPa for the diagnosis of significant fibrosis, with a positive predictive value (PPV) of 61.5% and a negative predictive value (NPV) of 98.0%. The LOGIQ S8 demonstrated an AUROC of 0.94 (95% CI, 0.85 to 0.98) using a cut-off value of  $>7.8$  kPa for the diagnosis of significant fibrosis, with a PPV of 53.3% and an NPV of 97.9%. In the pairwise receiver operating characteristic curve comparison, the AUROCs of the two 2D-SWE techniques for the prediction of significant fibrosis were not significantly different ( $P = 0.436$ ) (Fig. 5).

## Discussion

In this study, we tested the intersystem agreement of the comb-push 2D-SWE technique using two clinical US platforms from the same manufacturer. In recent years, various SWE techniques are gaining wide acceptance among clinicians as a noninvasive tool for the detection and precise staging of fibrosis and cirrhosis, which are very important for the timely initiation of appropriate therapeutic regimens [14]. Furthermore, the SWE technique could be widely used for monitoring liver fibrosis after the application of antiviral agents [26,27]. Therefore, for the longitudinal assessment or monitoring

of liver fibrosis, the interchangeability of LS measurements made using different SWE techniques could be of significant value. In daily practice, however, more than one SWE method or US platform can be used for the longitudinal assessment of liver fibrosis during the management of CLD over time [18]. Therefore, an evaluation of the intersystem agreement of LS values would make a major contribution to this resolving this issue. Importantly, we found that there was a significant difference in mean LS values measured by the two US platforms ( $P = 0.018$ ) despite good to excellent correlations. Based on our study results, absolute LS measurements from multiple systems, even made with the same technique using systems from the same factory, should not be used interchangeably during the follow-up or monitoring of liver fibrosis.

The two SWE systems use the same SWE software and beamforming technique. Therefore, the difference in mean LS values measured by the same 2D-SWE technique using two different platforms could be attributed to machine-specific factors [28]. Indeed, according to an interlaboratory study comparing shear wave velocities obtained with four different machines (Fibroscan, Philips iU22, ACUSON S2000, and Aixplorer) using elastic phantoms by the Ultrasound Shear Wave Speed technical committee of the Quantitative Imaging Biomarker Alliance of the Radiological Society of North America, there was a statistically significant difference in shear wave speed estimates among systems [13,28]. The differences in LS measurements across different US systems could be related to variations across manufacturers in aspects of the hardware, such as the frequency of the transducer and beam-forming technology, or the algorithm used to measure shear wave velocity [28]. Moreover, several previous studies also reported that a significant difference in LS measurements or shear wave velocity measurements was consistently observed across different SWE techniques [11,18,29,30].

In addition, the two 2D-SWE systems showed excellent technical success and reliable LS measurement rates, similar to those of TE. Our study results correspond well with the results of the previous study by Bende et al. [31], who demonstrated a similar rate of reliable LS measurements for TE and 2D-SWE (LOGIQ E9). We also found that for the two 2D-SWE systems, the AUROCs with the optimal cut-off values (7.8–8.3 kPa) were not significantly different for the detection of significant fibrosis ( $P = 0.436$ ), when using the cut-off values of TE as the reference standard.

One of the limitations in our study was the relatively small size of the study population, with an uneven distribution of liver fibrosis grades. However, we believe that the sample size was large enough to estimate technical success, the reliability of LS measurements, and intersystem agreement in LS measurements. Second, there was no histological diagnosis of fibrosis staging in the patients. For the reference method, we used TE, which has been well validated

in previous studies. Furthermore, the primary goal of our study was to evaluate intersystem variability in the measurement of LS. Another important limitation is that a single radiologist performed all examinations with the two 2D-SWE systems and TE. However, we did not allow the operator to see the summary of the serial measurements of each system until all measurements made using the three systems were completed. We think that recall bias was reduced to a reasonable level. Finally, our study was performed retrospectively. Further large prospective studies with histologic confirmation are needed.

In conclusion, the two 2D-SWE systems showed a similar technical success rate and reliability of LS measurements, but there was significant intersystem variability. Therefore, even 2D-SWE systems from the same manufacturer should not be used interchangeably in longitudinal follow-up.

ORCID: Hwaseong Ryu: <https://orcid.org/0000-0003-3143-3733>; Su Joa Ahn: <https://orcid.org/0000-0001-9026-0399>; Jeong Hee Yoon: <https://orcid.org/0000-0002-9925-9973>; Jeong Min Lee: <https://orcid.org/0000-0003-0561-8777>

### Conflict of Interest

No potential conflict of interest relevant to this article was reported.

## References

- Friedman SL. Mechanisms of hepatic fibrogenesis. *Gastroenterology* 2008;134:1655-1669.
- Lefton HB, Rosa A, Cohen M. Diagnosis and epidemiology of cirrhosis. *Med Clin North Am* 2009;93:787-799.
- Scaglione S, Kliethermes S, Cao G, Shoham D, Durazo R, Luke A, et al. The epidemiology of cirrhosis in the United States: a population-based study. *J Clin Gastroenterol* 2015;49:690-696.
- Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;38:1449-1457.
- Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. *Hepatology* 1994;20:15-20.
- Gronbaek K, Christensen PB, Hamilton-Dutoit S, Federspiel BH, Hage E, Jensen OJ, et al. Interobserver variation in interpretation of serial liver biopsies from patients with chronic hepatitis C. *J Viral Hepat* 2002;9:443-449.
- Al Knawy B, Shiffman M. Percutaneous liver biopsy in clinical practice. *Liver Int* 2007;27:1166-1173.
- Piccinino F, Sagnelli E, Pasquale G, Giusti G. Complications following percutaneous liver biopsy: a multicentre retrospective study on 68,276 biopsies. *J Hepatol* 1986;2:165-173.
- Ophir J, Cespedes I, Ponnekanti H, Yazdi Y, Li X. Elastography: a quantitative method for imaging the elasticity of biological tissues. *Ultrason Imaging* 1991;13:111-134.
- Jeong WK, Lim HK, Lee HK, Jo JM, Kim Y. Principles and clinical application of ultrasound elastography for diffuse liver disease. *Ultrasonography* 2014;33:149-160.
- 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;36:1563-1569.
- Foucher J, Chanteloup E, Vergniol J, Castera L, Le Bail B, Adhoute X, et al. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. *Gut* 2006;55:403-408.
- Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, et al. EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, uUpdate 2017 (long version). *Ultraschall Med* 2017;38:e16-e47.
- Piscaglia F, Salvatore V, Mulazzani L, Cantisani V, Schiavone C. Ultrasound shear wave elastography for liver disease: a critical appraisal of the many actors on the stage. *Ultraschall Med* 2016;37:1-5.
- Mancini M, Salomone Megna A, Ragucci M, De Luca M, Marino Marsilia G, Nardone G, et al. Reproducibility of shear wave elastography (SWE) in patients with chronic liver disease. *PLoS One* 2017;12:e0185391.
- Tang A, Cloutier G, Szeverenyi NM, Sirlin CB. Ultrasound elastography and MR elastography for assessing liver fibrosis: part 2, diagnostic performance, confounders, and future directions. *AJR Am J Roentgenol* 2015;205:33-40.
- Barr RG. Shear wave liver elastography. *Abdom Radiol (NY)* 2018;43:800-807.
- Ahn SJ, Lee JM, Chang W, Lee SM, Kang HJ, Yang H, 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;18:926-935.
- Wong VW, Vergniol J, Wong GL, Foucher J, Chan HL, Le Bail B, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology* 2010;51:454-462.
- Chon YE, Choi EH, Song KJ, Park JY, Kim DY, Han KH, et al. Performance of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B: a meta-analysis. *PLoS One* 2012;7:e44930.
- Friedrich-Rust M, Ong MF, Martens S, Sarrazin C, Bojunga J, Zeuzem S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. *Gastroenterology* 2008;134:960-974.
- Fraquelli M, Rigamonti C, Casazza G, Conte D, Donato MF, Ronchi G, et al. Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut*



- 2007;56:968-973.
23. Castera L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* 2010;51:828-835.
  24. Evans JD. *Straightforward statistics for the behavioral sciences*. Pacific Grove: Brooks/Cole Publishing Co., 1996.
  25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.
  26. Korda D, Lenard ZM, Gerlei Z, Jakab Z, Haboub-Sandil A, Wagner L, et al. Shear-wave elastography for the assessment of liver fibrosis in liver transplant recipients treated for hepatitis C virus recurrence. *Eur J Gastroenterol Hepatol* 2018;30:27-32.
  27. Wang JH, Changchien CS, Hung CH, Tung WC, Kee KM, Chen CH, et al. Liver stiffness decrease after effective antiviral therapy in patients with chronic hepatitis C: longitudinal study using FibroScan. *J Gastroenterol Hepatol* 2010;25:964-969.
  28. Barr RG, Ferraioli G, Palmeri ML, Goodman ZD, Garcia-Tsao G, Rubin J, et al. Elastography assessment of liver fibrosis: Society of Radiologists in Ultrasound Consensus Conference Statement. *Radiology* 2015;276:845-861.
  29. Sigrist RMS, El Kaffas A, Jeffrey RB, Rosenberg J, Willmann JK. Intra-individual comparison between 2-D shear wave elastography (GE System) and Virtual Touch tissue quantification (Siemens System) in grading liver fibrosis. *Ultrasound Med Biol* 2017;43:2774-2782.
  30. Zeng J, Zheng J, Huang Z, Chen S, Liu J, Wu T, et al. Comparison of 2-D shear wave elastography and transient elastography for assessing liver fibrosis in chronic hepatitis B. *Ultrasound Med Biol* 2017;43:1563-1570.
  31. Bende F, Sporea I, Sirlu R, Popescu A, Mare R, Miutescu B, et al. Performance of 2D-SWE.GE for predicting different stages of liver fibrosis, using transient elastography as the reference method. *Med Ultrason* 2017;19:143-149.