Assessment of liver fibrosis using two-dimensional shear wave elastography: a prospective study of intra-and inter-observer repeatability, and comparison with point shear wave elastography

Abbreviated title: Repeatability of two-dimensional shear wave elastography
Type of manuscript: Original Article

Jeongin Yoo, MD¹; Jeong Min Lee, MD¹; Ijin Joo, MD¹; Jeong Hee Yoon, MD¹

¹Department of Radiology, Seoul National University Hospital, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea

Disclosures of Conflicts of Interest: All authors disclosed no relevant relationships.
Funding: This work was supported by Samsung Healthcare (Seoul, Korea).

Correspondence to:
Jeong Min Lee, M.D.
Department of Radiology and Institute of Radiation Medicine, Seoul National University College of Medicine, Daehak-ro 101, Jongno-gu, Seoul 03080, Korea
Telephone: 82-2-2072-3154; Fax: 82-2-743-6385 E-mail: jmlshy2000@gmail.com
Assessment of liver fibrosis using two-dimensional shear wave elastography: a prospective study of intra-and inter-observer repeatability, and comparison with point shear wave elastography

Type of manuscript: Original Article
Abstract

Purpose: To prospectively investigate the intra- and inter-observer repeatability of a new two-dimensional (2D) shear wave elastography (SWE) technique (S-Shearwave) for assessment of liver fibrosis in patients with chronic liver disease, and to compare liver stiffness measurements (LSMs) using 2D-SWE with point SWE (pSWE).

Methods: This prospective study was approved by our institutional review board and informed consent was obtained from all patients. Fifty-three patients with chronic liver disease were randomly allocated to either group 1 (for intra-observer repeatability [n=33]) or group 2 (for inter-observer repeatability [n=20]). In group 1, two sessions of 2D-SWE and one session of pSWE were performed by one radiologist. In group 2, one session of 2D-SWE and one session of pSWE was performed by the aforementioned radiologist, and the second session of 2D-SWE by another radiologist, were performed. Intra- and inter-observer repeatability was assessed using intraclass correlation coefficient (ICC). LSMs using 2D-SWE were compared and correlated with those of pSWE using the paired t-test and Pearson’s correlation coefficient, respectively.

Results: LSMs using 2D-SWE demonstrated excellent intra- and inter-observer repeatability (ICC, 0.997 [95% confidence interval, 0.994–0.999]) and 0.995 [0.988–0.998], respectively). LSMs using 2D-SWE were significantly different from those using pSWE (2.1±0.6 m/s versus 1.9±0.6 m/s; P<0.001), although significant correlation was found between LSMs using two techniques (rho=0.836; P<0.001).

Conclusion: S-Shearwave demonstrated excellent intra- and inter-observer repeatability, and strong correlation with pSWE for assessment of liver stiffness. However, because of a significant difference between LSMs assessed using 2D-SWE and pSWE, these methods should not be used interchangeably.

Keywords: Ultrasonography; Elasticity Imaging Techniques; Liver Cirrhosis
INTRODUCTION

Liver fibrosis is the final common pathway of most of chronic liver diseases that progress to cirrhosis and hepatocellular carcinoma [1,2]. Staging the severity of liver fibrosis is central to treatment planning and disease prognosis [3,4]. Therefore, it is important to accurately assess the degree of fibrosis in the management of patients with chronic hepatitis. The current gold standard for evaluation of liver fibrosis is histopathological examination of biopsy samples. However, liver biopsy has several important limitations, including invasiveness and sampling error; thus, there is a need for non-invasive methods that accurately assess the state of the entire liver [5]. Accordingly, recent studies have investigated non-invasive measurement methods using blood tests [6,7] or imaging techniques such as ultrasound (US)-based elastography or magnetic resonance (MR) elastography [8-10]. Owing to previous studies that reported successful fibrosis staging using MR elastography, it is now widely accepted as a useful tool [11]; however, it is expensive and not widely accessible.

US-based elastography is non-invasive, can provide gray-scale images of the liver, has no radiation hazard, and can quantitatively evaluate the severity of fibrosis, unlike conventional US, which often relies on subjective judgement. US-based elastographic methods are divided into strain elastography and shear-wave elastography (SWE) techniques [12,13]. The latter includes transient elastography (TE), point SWE (pSWE), and real-time two-dimensional (2D) SWE (2D-SWE). Among these, TE (Fibroscan, Echosense, Paris, France) is the most widely used and has been validated in several meta-analyses [14,15]. In addition, several meta-analyses have reported good diagnostic accuracy and reliability for pSWE [16,17]. However, pSWE measures only a fixed area, approximately 5 mm × 10 mm, without displaying a color image in a region-of-interest (ROI), which can result in low spatial resolution and neglect of inhomogeneous liver fibrosis, similar to TE [18]. In contrast, more recently developed 2D-SWE techniques enable real-time measurements of liver stiffness in a larger ROI, which result in significantly larger sampling volume compared
with TE or pSWE. To date, several studies have reported that 2D-SWE techniques of several manufacturers also provide good diagnostic accuracy for liver fibrosis [19-22].

S-Shearwave imaging (S-SWI, Samsung Healthcare, Seoul, Korea) is a real-time, 2D-SWE technique with a large ROI (3 cm × 4 cm) based on an acoustic radiation force impulse (ARFI) push method and has been introduced into clinical practice. However, to date, there have been no studies evaluating the repeatability of this new 2D-SWE technique to assess fibrosis, which is one of the key important parameters in longitudinal follow-up of patients with chronic liver disease and/or the evaluation of treatment response [4]. Considering that patients with chronic liver disease undergo screening US at intervals of 3 to 6 months, SWE can be widely used to evaluate liver fibrosis, or can be used as an add-on examination to conventional liver US examination, but only if there is little variation among different operators and US systems.

The purpose of this prospective study, therefore, was to evaluate the intra- and inter-observer repeatability of S-SWI in liver stiffness measurement (LSM) in patients with chronic liver disease, and to compare shear-wave speed (SWS) obtained using 2D-SWE with that using pSWE (Virtual Touch Quantification [VTQ], Siemens Healthcare, Erlangen, Germany) techniques.
MATERIALS AND METHODS

This prospective study was approved by our institutional review board and informed consent was obtained from all patients.

Patients

Fifty-six patients who met the following eligible criteria were enrolled from April to July 2018: 1) diagnosed with chronic liver disease based on clinical findings or laboratory tests, including chronic hepatitis B or C, alcoholic liver disease, or nonalcoholic fatty liver disease; 2) > 18 years of age; and 3) scheduled to undergo liver US examination at our radiology department. Patients who were not able to hold their breath for > 3 seconds during the US examination (n=2) and those who had undergone right hepatectomy (n=1) according to European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines, which recommend the avoidance of LSM in the left liver lobe [12] were excluded. Fifty-three patients were randomly allocated to one of two groups: group 1 for intra-observer repeatability (n=33); and group 2 for inter-observer repeatability (n=20) (Table 1).

Demographics of patients including sex, age, and body mass index (BMI) were recorded on the basis of the electronic medical chart.

Gray-scale imaging and liver stiffness measurements

Patients underwent gray-scale imaging, S-SWI, and pSWE on the same day. Gray-scale US and two sessions of S-SWI were performed using a clinical US system (RS-85A, Samsung Healthcare, Seoul, Korea) and a convex probe. Thereafter, pSWE was performed using the Siemens Acuson S2000 Virtual Touch US system (Siemens Healthcare, Erlangen, Germany) equipped with a convex probe. All patients fasted for at least 6 hours before the examination, and were positioned supine with the right arm maximally abducted above the head to stretch the intercostal muscles. In group 1, two sessions of S-SWI and one session of pSWE were performed by one board-certified radiologist (J.M.L. with 20 years’ experience in abdominal imaging). In order to evaluate intraobserver repeatability, the radiologist performed two sessions of S-SWI measurements within 1-hour interval on each patient, after
having changes of posture. In group 2, one session of S-SWI and one session of pSWE by
one radiologist, and the second session of S-SWI by the other board-certified radiologist (J.Y.
with 5 years' experience) were performed in each patient.

During gray-scale imaging, the radiologist decided where to measure the LS,
avoiding focal liver lesions and hepatic vasculatures. Each session of pSWE and S-SWI
consisted of 10 measurements recorded at an interval of 2–5 min, and a positional change
was required between sessions. A 2D-SWE map was obtained by placing a 2 × 3 cm²
sample box overlaid on a gray-scale image in the right lobe of the liver via an intercostal
approach while patients held their breath (Fig. 1). When the radiologist placed 2–4 circular
ROIs (1 cm in diameter) in the sample box, the LS (kPa) was automatically displayed with
the reliability measurement index (RMI). RMI is a performance index, which is calculated as
the weighted sum of the residual of the wave equation and the magnitude of the shear wave
[23]. According to the manufacturer of the device, the operators attempted to obtain 10
measurements with RMI > 0.4. After obtaining 10 measurements per session, the median
and interquartile range (IQR [the difference between the 75th and 25th percentiles]) divided
by the median LS (kPa) and shear-wave speed (m/s) each were calculated and shown in a
table. It took approximately 5–7 seconds to obtain 2–4 measurements in each sample box.
The median LS values in kPa and IQR/median ratio of the 10 measurements for each
session were used for analysis. However, for evaluation of agreement between LS
measurements using VTQ and S-Shearwave, the median values of shear-wave speed of
both techniques were compared.

For assessment of the applicability of 2D-SWE, technical success rate and reliable
measurement rate were evaluated. Technical failure was defined as failure to acquire a color
map in > 50% of the sampling area for all acquisitions [24]. A reliable measurement was
defined as a measurement in which the IQR/median LS of 10 measurements was < 30%.
The applicability rate was calculated as the ratio of examinations that demonstrated
technical success and reliable measurement. Liver stiffness values with technical failure or
unreliable measurements were included only in the assessment of applicability rate and excluded in the evaluation of intra- and inter-observer repeatability and correlation between 2D-SWE and pSWE measurements.

**Statistical analysis**

The Wilcoxon signed-rank test was used to compare LS values obtained in the first and second sessions, and using 2D-SWE and pSWE techniques. Intra-observer (group 1) and inter-observer (group 2) repeatability of 2D-SWE were assessed using intraclass correlation coefficients (ICCs), Bland-Altman test, and coefficients of variation. ICC estimates and their 95% confidence intervals (CIs) were calculated based on a mean rating, absolute agreement, two-way mixed-effects model. Based on the 95% CI of the ICC estimate, values < 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and > 0.90 were indicative of poor, moderate, good, and excellent reliability, respectively. Bland-Altman analysis tested the relationship of the difference between the first and second sessions of 2D-SWE measurements in each group plotted against their mean, showing mean value of the difference and limits of agreement (LOA) of two series of data. The coefficient of variation of intra-observer and inter-observer repeatability was calculated, which is the standard deviation (SD) divided by the mean value. LSM using 2D-SWE in the first session was compared and correlated with that of pSWE using a paired t-test and Pearson's correlation coefficient, respectively. All statistical analyses were performed using IBM SPSS version 25.0 (IBM Corporation, Armonk, NY, USA) for Windows (Microsoft Corporation, Redmond, WA, USA) and MedCalc version 15.2 (MedCalc Software, Ostend, Belgium; http://www.medcalc.org; 2015); \( P < 0.05 \) was considered indicative of a statistical significance.
RESULTS

The applicability rate of 2D-SWE and pSWE

In LSM using 2D-SWE, no technical failure was observed in both groups 1 and 2. There were 3 patients with unreliable measurements in group 1 (3/33 [9.1%]); an unreliable measurement occurred in the first session in 1 patient; 2 patients exhibited unreliable measurement results in the second session. All three patients had BMI > 25 kg/m². In group 2, no patient exhibited an unreliable measurement result. The applicability rate of 2D-SWE was 90.9% (30/33) in group 1 and 100% (20/20) in group 2. In the pSWE session, no technical failure occurred, and only one patient in group 2 exhibited an unreliable measurement result. pSWE demonstrated an applicability rate of 100% (33/33) in group 1 and 95% (19/20) in group 2. No significant differences were found in applicability rates between the two methods in both group 1 (P=0.24) and 2 (P=1.00).

Intra-observer and inter-observer variability of LSM

In group 1, mean liver stiffness values were 13.2±8.5 kPa (range, 4.5–37.0 kPa) in the first session and 13.5±8.8 kPa (range, 4.3–38.0 kPa) in the second session. For intra-observer repeatability of 2D-SWE in group 1, the ICC was 0.997 (95% CI 0.994–0.999), which indicated excellent reliability. According to the Bland-Altman test, the mean difference between the 2 sessions was -0.2±1.8 kPa. The 95% upper and lower LOA were 1.5 and -2.0 kPa, respectively (Fig. 2). The coefficient of variation was 6.0% (95% CI 4.4–7.7).

In group 2, mean liver stiffness values were 12.1±6.9 kPa in the first session and 12.0±6.4 kPa in the second session, respectively. For inter-observer repeatability in group 2, the ICC was 0.995 (95% CI 0.988–0.998), which indicated excellent reliability. The 95% Bland-Altman limit of agreement was 0.1±1.8 kPa. The 95% upper and lower LOA were 2.0 and -1.7 kPa, respectively (Fig. 3). The coefficient of variation was 4.1% (95% CI 2.7%–5.5%).

Correlation of LSM between pSWE and 2D-SWE

In 49 patients who had reliable 2D-SWE and pSWE measurements, the LS values
of the two techniques demonstrated significant correlation (rho=0.836; P<0.001) (Figure).

The mean LS values for 2D-SWE were significantly different from those for pSWE: 2.1±0.6 m/s versus 1.9±0.6 m/s, P<0.001. The 95% Bland-Altman limit of agreement between LS measurements using VTQ and S-Shearwave was 34.7% of the mean (Fig. 4).
DISCUSSION

In our prospective study, both intra- and inter-observer repeatability of 2D-SWE obtained using the S-SWI technique in patients with chronic liver disease were excellent (ICC 0.997 and 0.995, respectively). In addition, both 2D-SWE and pSWE demonstrated a high applicability rate without a significant difference between them. LS values for 2D-SWE demonstrated good correlation with those for VTQ. The main clinical indication for SWE is fibrosis staging of chronic liver disease, with a primary objective of determining the presence or absence of advanced fibrosis [25]. In recent years, SWE has increasingly been used for the evaluation of significant fibrosis, or liver cirrhosis or portal hypertension during patient management, and also for evaluation of treatment response to antiviral treatments [26,27].

In this regard, the high reproducibility of SWE techniques is essential to its clinical use. In our study, the values measured using S-Shearwave were not interchangeable with those measured using VTQ. This result was consistent with a previous study involving phantoms by the Ultrasound Shear Wave Speed technical committee of the Radiological Society of North America Quantitative Imaging Biomarker Alliance, in which the differences in measurements between machines and observers varies by as much as 12% [25]. Considering these results, we believe that S-Shearwave can be widely used for the evaluation of liver fibrosis in patients with chronic liver disease with validated repeatability.

There have been several studies investigating the repeatability of 2D-SWE techniques in terms of intra- and inter-observer variability, most of which were performed using supersonic shear imaging (SSI). For SSI, intra-observer repeatability in patients with liver fibrosis was excellent (ICC 0.90–0.95) [28-30], and ICCs for inter-observer repeatability ranged from 0.83 to 0.94, which may be inferior to pSWE using VTQ [31,32]. Our study demonstrated higher ICCs for both intra- and inter-observer agreement of S-Shearwave than those of SSI in published studies, although a direct comparison between S-Shearwave and SSI was not performed in our study. Furthermore, while a previous study suggested that operator experience may have a role in reliable measurement, our results yielded excellent
inter-observer agreement between two radiologists with different levels of experience. These discrepancies may be explained by RMI implemented in S-Shearwave, which enabled the operators in our study to filter out unreliable measurements, thus resulting in performance improvement of SWE. Our result is concordant with a previous study that reported a strong correlation between high RMI values and reproducible measurements [23]. However, although the ICC was high, the Bland-Altman analysis demonstrated that the mean differences in group I and II (95% Bland-Altman limit of agreement) were -0.2±1.8 kPa and 0.1±1.8 kPa, respectively. Therefore, 1~2 kPa discrepancy in LS measurements by the same operator or different operator could occur but considering that the cutoff value for diagnosing significant fibrosis (F ≥ 2) was proposed as > 7.1kPa [20], this range of discrepancy would be clinically acceptable.

A few studies have focused on the direct comparison of diagnostic capabilities of different SWE techniques from various manufacturers, most of which suggested that different SWE techniques should not be used interchangeably [31,33]. To address this issue on direct comparison of SWS measurements in the liver, the Radiological Society of North America Quantitative Imaging Biomarker Alliance conducted a phantom study using various commercial SWE systems and found a statistically significant difference in SWS estimates among systems and with depth into the phantom [34]. This phantom study concluded that there are several sources of bias and variance that can be addressed to improve the consistency of measurements [34]. The variability of measured SWS among different SWE technologies may occur due to shear-wave vibration frequency and bandwidth, as well as the software used to calculate relative shear-wave arrival time and speed [12]. Our results in human patients are in good agreement with the aforementioned studies in that a significant difference was found between SWS measured using 2D-SWE and pSWE. Further studies to specify the source(s) of errors to enable interchangeable use of different SWE techniques in clinical practice are warranted.

Both American and European guidelines for the management of patients with
hepatitis C virus infection have recommended evaluating the degree of hepatic fibrosis to
assess the urgency of treatment and, according to a recent European guideline, non-
invasive methods should be used instead of liver biopsy to assess liver disease severity
before therapy [35,36]. Moreover, while TE has been accepted as a noninvasive test for the
assessment of liver fibrosis in previous guidelines, recent European recommendations
included ARFI (VTQ) and Aixplorer (SuperSonic Imagine, Aix-en-Provence, France) as non-
invasive markers and suggested cut-offs for each SWE system for the prediction of liver
fibrosis stages [36]. Although 2D-SWE has only been recently validated, and there have
been only a limited number of studies investigating the diagnostic capability of 2D-SWE
techniques other than SSI, it has several advantages over well-established techniques such
as TE or pSWE. First, 2D-SWI is derived from the characteristic broadband (60-600-Hz)
pulse using ARFI [37], while TE applies 50-Hz push by an external vibrator [38]. A previous
study reported that stiffness imaging using a broadband pulse provided a more discriminant
parameter for fibrosis evaluation [39]. Second, both pSWE and 2D-SWE can be performed
with conventional gray-scale US, which demonstrates hepatic parenchymal echogenicity as
well as a focal liver lesion. Third, 2D-SWE has a larger sample volume of liver parenchyma
and displays color-coded elasticity maps, enabling more opportunities for valid
measurements compared with pSWE.

Our study had several limitations, the first of which was that intra- and inter-observer
repeatability were evaluated in different study patients. Second, there was a relatively short
interval between the two 2D-SWE sessions, and patients with a relatively small body habitus
(mean BMI < 25 kg/m² in both groups) were enrolled, which may have resulted in an
overestimation of repeatability compared with repeatability on different days or in obese
patients. Third, the diagnostic performance of S-Shearwave could not be assessed because
biopsies were not performed. However, our study aimed to evaluate the repeatability—rather
than the diagnostic performance—of 2D-SWE for LSM. Further studies investigating
diagnostic performance using pathological results as the reference standard are, therefore,
In conclusion, S-Shearwave demonstrated excellent intra- and inter-observer repeatability. Although a significant correlation was found in LSMs between 2D-SWE and pSWE, liver stiffness values obtained using 2D-SWE were significantly higher than those with pSWE, suggesting that these techniques should not be used interchangeably.
References


Table 1. Patient characteristics

<table>
<thead>
<tr>
<th></th>
<th>Group 1 (n=33)</th>
<th>Group 2 (n=20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>65.3±9.1 years</td>
<td>66.1±9.3 years</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>26 (78.8%)</td>
<td>16 (80%)</td>
</tr>
<tr>
<td>Female</td>
<td>7 (21.2%)</td>
<td>4 (20%)</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td>24.8 ± 2.1 kg/m²</td>
<td>23.1±3.0 kg/m²</td>
</tr>
<tr>
<td><strong>Causes of liver disease</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>25 (75.8%)</td>
<td>13 (65%)</td>
</tr>
<tr>
<td>Hepatitis C</td>
<td>4 (12.1%)</td>
<td>5 (25%)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>1 (0.3%)</td>
<td>1 (5%)</td>
</tr>
<tr>
<td>Hepatitis B &amp; alcohol</td>
<td>1 (0.3%)</td>
<td>0</td>
</tr>
<tr>
<td>NAFLD</td>
<td>1 (0.3%)</td>
<td>0</td>
</tr>
<tr>
<td>Non-B Non-C hepatitis</td>
<td>1 (0.3%)</td>
<td>1 (5%)</td>
</tr>
<tr>
<td><strong>No. of patients with liver cirrhosis</strong></td>
<td>28 (84.8%)</td>
<td>16 (80%)</td>
</tr>
<tr>
<td><strong>Laboratory findings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albumin (mg/dL)</td>
<td>3.8 ± 0.6</td>
<td>3.9 ± 0.3</td>
</tr>
<tr>
<td>AST (IU/L)</td>
<td>32 ± 14</td>
<td>33 ± 15</td>
</tr>
<tr>
<td>ALT (IU/L)</td>
<td>25 ± 12</td>
<td>26 ± 10</td>
</tr>
<tr>
<td>Platelet (K/mm³)</td>
<td>125 ± 59</td>
<td>118 ± 52</td>
</tr>
<tr>
<td>Prothrombin time (INR)</td>
<td>1.07 ± 0.12</td>
<td>1.05 ± 0.09</td>
</tr>
</tbody>
</table>

Note. – Data are mean ± standard deviation.

BMI, body mass index; NAFLD, nonalcoholic fatty liver disease; AST, aspartate aminotransferase; ALT, alanine aminotransferase; INR, international normalized ratio.

*All patients were clinically diagnosed with liver cirrhosis.
**Figure legends**

**Fig. 1.** Liver stiffness measurement using S-Shearwave (Samsung Healthcare, Seoul, Korea) (A) and Virtual Touch Quantification (Siemens Healthcare, Erlangen, Germany) (B).

A. A $2 \times 3$ cm$^2$ square sample box was placed in the right lobe of the liver on a grayscale image. When 2–4 circular regions of interest (1 cm in diameter) were placed in the color-coded sample box, the liver stiffness (kPa) was automatically displayed with the reliability measurement index.

B. A $1 \times 0.6$ cm$^2$ measurement box was placed in the right lobe of the liver on a grayscale image and the liver stiffness (m/s) was displayed.

**Fig. 2.** Bland-Altman plot of differences in liver stiffness (LS) values in each session for group 1 in intra-observer repeatability analysis. The solid line represents the mean of the difference in the median LS values measured in each session; the dashed lines define the limits of agreement. The mean difference was $-0.2 \pm 1.8$ kPa. The 95% upper and lower limits of agreement were 1.5 and $-2.0$ kPa, respectively.

**Fig. 3.** Bland-Altman plot of differences in liver stiffness (LS) values in each session for group 2 in inter-observer repeatability analysis. The solid line represents the mean of the difference in median LS values measured by two radiologists; the dashed lines define the limits of agreement. The mean difference was $0.1 \pm 1.8$ kPa. The 95% upper and lower limits of agreement were 2.0 kPa and $-1.7$ kPa, respectively.

**Fig. 4.** Bland-Altman plot revealing agreement in Virtual Touch Quantification (VTQ; Siemens Healthcare, Erlangen, Germany) and S-Shearwave (Samsung Healthcare, Seoul, Korea) liver stiffness (LS) measurements, expressed as percentages of the values on the axis versus the mean of the two measurements. The blue line represents the mean of the difference; the green line indicates the confidence interval limits for the mean; the dashed
lines define the limits of agreement. The 95% Bland-Altman limit of agreement between LS measurements by VTQ and S-Shearwave was 34.7% of the mean.
Fig. 1. Liver stiffness measurement using S-Shearwave (Samsung Healthcare, Seoul, Korea) (A) and Virtual Touch Quantification (Siemens Healthcare, Erlangen, Germany) (B). A 2 × 3 cm² square sample box was placed in the right lobe of the liver on a gray-scale image. When 2–4 circular regions of interest (1 cm in diameter) were placed in the color-coded sample box, the liver stiffness (kPa) was automatically displayed with the reliability measurement index.
Fig. 1B

B. A 1 × 0.6 cm² measurement box was placed in the right lobe of the liver on a gray-scale image and the liver stiffness (m/s) was displayed.
Fig. 2 Bland-Altman plot of differences in liver stiffness (LS) values in each session for group 1 in intra-observer repeatability analysis. The solid line represents the mean of the difference in the median LS values measured in each session; the dashed lines define the limits of agreement. The mean difference was -0.2 ± 1.8 kPa. The 95% upper and lower limits of agreement were 1.5 and -2.0 kPa, respectively.
Fig. 3

Fig. 3. Bland-Altman plot of differences in liver stiffness (LS) values in each session for group 2 in inter-observer repeatability analysis. The solid line represents the mean of the difference in median LS values measured by two radiologists; the dashed lines define the limits of agreement. The mean difference was 0.1±1.8 kPa. The 95% upper and lower limits of agreement were 2.0 kPa and -1.7 kPa, respectively.
Fig. 4. Bland-Altman plot revealing agreement in Virtual Touch Quantification (VTQ; Siemens Healthcare, Erlangen, Germany) and S-Shearwave (Samsung Healthcare, Seoul, Korea) liver stiffness (LS) measurements, expressed as percentages of the values on the axis versus the mean of the two measurements. The blue line represents the mean of the difference; the green line indicates the confidence interval limits for the mean; the dashed lines define the limits of agreement. The 95% Bland-Altman limit of agreement between LS measurements by VTQ and S-Shearwave was 34.7% of the mean.