



Test-retest repeatability of ultrasonographic shear wave elastography in a rat liver fibrosis model: toward a quantitative biomarker for preclinical trials

ULTRASONOGRAPHY

Youngbin Shin^{1,2}, Jimi Huh^{1,3}, Su Jung Ham⁴, Young Chul Cho⁴, Yoonseok Choi⁴, Dong-Cheol Woo⁴, Jeongjin Lee², Kyung Won Kim^{1,4}

¹Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul; ²School of Computer Science and Engineering, Soongsil University, Seoul; ³Department of Radiology, Ajou University Hospital, Ajou University School of Medicine and Graduate School of Medicine, Suwon; ⁴Bioimaging Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea

Purpose: This study evaluated the test-retest repeatability and measurement variability of ultrasonographic shear wave elastography (SWE) for liver stiffness in a rat liver fibrosis model.

Methods: In 31 Sprague-Dawley rats divided into three groups (high-dose, low-dose, and control), liver fibrosis was induced by intraperitoneal administration of thioacetamide for 8 weeks. A dedicated radiographer performed SWE to measure liver stiffness in kilopascals in two sessions at a 3-day interval. We calculated correlations between liver stiffness and histopathologic results, measurement variability in each session using coefficients of variation (CoVs) and interquartile/median (IQR/M), and test-retest repeatability between both sessions using the repeatability coefficient.

Results: Different levels of liver fibrosis in each group were successfully induced in the animal model. The mean liver stiffness values were 8.88 ± 1.48 kPa in the control group, 11.62 ± 1.70 kPa in the low-dose group, and 11.91 ± 1.73 kPa in the high-dose group. The correlation between collagen areas and liver stiffness values was moderate ($r=0.6$). In all groups, the second session yielded lower CoVs (i.e., more reliable results) for liver stiffness than the first session, suggesting a training effect for the operator. The mean IQR/M values were also lower in the second session than in the first session, which had four outliers (0.21 vs. 0.12, $P < 0.001$). The test-retest repeatability coefficient was 3.75 kPa and decreased to 2.82 kPa after removing the four outliers.

Conclusion: The use of ultrasonographic SWE was confirmed to be feasible and repeatable for evaluating liver fibrosis in preclinical trials. Operator training might reduce variability in liver stiffness measurements.

Keywords: Shear wave elastography; Ultrasound; Liver fibrosis; Liver stiffness; Repeatability

ORIGINAL ARTICLE

<https://doi.org/10.14366/usg.19088>
pISSN: 2288-5919 • eISSN: 2288-5943
Ultrasonography 2021;40:126-135

Received: December 16, 2019

Revised: April 20, 2020

Accepted: April 23, 2020

Correspondence to:

Jimi Huh, MD, PhD, Department of Radiology, Ajou University Hospital, Ajou University School of Medicine and Graduate School of Medicine, 164 World cup-ro, Yeongtong-gu, Suwon 16499, Korea

Tel. +82-31-219-4369

Fax. +82-31-219-5862

E-mail: jimihuh.rad@gmail.com

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 © 2021 Korean Society of Ultrasound in Medicine (KSUM)



How to cite this article:

Shin Y, Huh J, Ham SJ, Cho YC, Choi Y, Woo DC, et al. Test-retest repeatability of ultrasonographic shear wave elastography in a rat liver fibrosis model: toward a quantitative biomarker for preclinical trials. *Ultrasonography*. 2021 Jan;40(1):126-135.

Introduction

Liver shear wave elastography (SWE) is a rapidly developing ultrasonographic technique that evaluates liver stiffness through shear wave speed measurements in liver tissue of patients with liver fibrosis or cirrhosis [1–3]. Liver SWE has been increasingly used as a quantitative biomarker for evaluating new drugs or treatments, such as anti-viral agents and anti-fibrotic agents [4]. In the clinical field, the physics and manufacturing techniques, measurement variability, quality criteria, and procedures for performing elastography have been extensively investigated, and have now become standardized through the emergence of an international consensus [5–7]. As such, the World Federation for Ultrasound in Medicine and Biology (WFUMB) issued guidelines on the use of SWE for the assessment of liver diseases in 2015 and 2018 [8,9].

In preclinical trials, the use of liver elastography has also become increasingly common as part of accelerated drug development efforts for anti-fibrotic agents for chronic liver disease and anti-viral agents for chronic hepatitis [4]. Non-invasive monitoring tools in mouse or rat models are very useful for assessing the treatment response to new drug candidates. In this regard, ultrasonographic SWE in preclinical trials might be a very powerful tool, as long as its technical feasibility is validated and repeatability is ensured. Currently, clinical ultrasonographic machines are readily available for SWE in a rat liver fibrosis model, thereby facilitating the use of SWE in preclinical trials. In contrast, other liver elastographic techniques such as magnetic resonance elastography (MRE) are very rarely used in preclinical models, because MRE requires a customized animal coil and electromechanical driver [10].

With regard to technical feasibility, several prior studies have investigated SWE techniques in rat liver fibrosis models and reported that SWE biomarkers reflected pathologic processes and outcomes [11,12]. However, the methods slightly differed across studies, and a standardized method of performing SWE has not been established in rat liver models. In addition, only sparse reports have been published on the repeatability of SWE in rat models [13], even though the repeatability of SWE in phantom studies and clinical studies has been thoroughly established [14–17]. Before using SWE as a quantitative biomarker for new drug development or for assessing treatment response, the repeatability of measurements must be ensured.

From this perspective, we performed an animal study to establish a standardized operating procedure (SOP) to perform ultrasonographic SWE and evaluated the test-retest repeatability of SWE in a rat liver fibrosis model.

Materials and Methods

Animal Model

All experiments conducted as part of this study were approved by our institutional animal care and use committee. A drug-induced chronic liver injury model was adopted to generate liver fibrosis. All Sprague-Dawley rats (male, 8 weeks old, weighing 270–280 g) were obtained from Orient Bio (Seoul, Korea) and maintained under specific-pathogen-free conditions.

Thioacetamide (TAA; Sigma-Aldrich Co., St. Louis, MO, USA), which is a hepatotoxic agent that causes centrilobular necrosis, was used to model drug-induced liver injury [4]. To minimize any potential selection bias on the part of the researcher, 34 rats were randomly assigned to three groups using a computerized random number generator (<https://randomizer.org>). The assignment resulted in eight rats in the control group, 11 rats in the low-dose group, and 15 rats in the high-dose group.

Intermittent intraperitoneal injections of TAA or saline were administered 3 times per week for 8 weeks. The doses administered were 200 mg/kg of TAA in the high-dose group, 150 mg/kg of TAA in the low-dose group, and 500 μ L of saline in the control group [18–20]. The TAA dose was determined by a preliminary experiment (results not shown) using 11 rats administered 0 (n=1), 100 (n=3), 150 (n=2), 200 (n=3), and 250 mg/kg (n=2) of TAA. A dose of 100 mg/kg did not induce liver fibrosis consistently, and a dose of 250 mg/kg resulted in death of all rats during the TAA medication period. Based on those results, doses of 150 mg/kg and 200 mg/kg were selected for this experiment.

After 8 weeks of TAA medication, SWE examinations were performed in two measurement sessions at a 3-day interval to evaluate the test-retest repeatability of the entire process of liver stiffness measurements, including animal preparation, anesthesia, and SWE examinations.

Histopathologic Evaluation as Reference Standards

After the second SWE examination, the animals were euthanized in a carbon dioxide chamber. We performed *en bloc* resection of the liver and then sliced it cross-sectionally at 5-mm intervals. The excised tissues were then fixed in 10% formalin and embedded in paraffin blocks. Hematoxylin and eosin (H&E) staining was performed for a microscopic evaluation of the liver parenchyma. We requested the Experimental Animal Histopathology Laboratory in our institution to conduct a histopathologic examination of the liver. Liver fibrosis was analyzed based on the presence of portal fibrosis (fibrous portal expansion), periportal fibrosis (periportal fibrosis with short septa extending into the lobules or porto-portal septa), septal fibrosis (fibrous septa reaching the adjacent portal tracts and

terminal hepatic venule), or cirrhosis (diffuse nodular formation). Liver inflammation was assessed based on the presence of inflammatory cells (neutrophils and lymphocytes) in the lobules. Liver steatosis was evaluated based on the presence of lipid accumulation in the hepatocyte vacuoles.

The Masson trichrome stain was performed to evaluate the extent of liver fibrosis using a commercially available kit (Sigma-Aldrich Korea, Seoul, Korea), which stains collagen blue. The area of collagen was quantified with ImageJ software (National Institutes of Health, Bethesda, MD, USA) using the following steps: (1) five representative hotspots were determined at a lower magnification ($\times 40$), (2) those areas were captured and digitized for morphometric analysis, and (3) areas of collagen were selected using the colorimetric threshold of the blue color [21,22]. If ImageJ did not automatically select areas of collagen, we adjusted the areas manually using the H&E stain as a reference [23]. The area of collagen as a percentage (area of blue pixels/area of selected rectangle) from the five hotspots was averaged and used for statistical analysis.

Shear Wave Elastography

Two-dimensional SWE was acquired with an Aplio 500 Platinum ultrasound machine (Canon Medical Systems Corp., Tokyo, Japan) using a linear probe (14L5 transducer, 5.0–14.0 MHz). SWE was performed by a single operator (Y.C.C.), a radiographer with 2 years of experience in SWE, TE, and liver Doppler ultrasonography in our clinical unit and 3 years of experience in animal ultrasonographic imaging at our animal imaging center. A supervising radiologist (J.H.) with 2 years of experience in SWE and 10 years of experience in abdominal ultrasonography checked the measurement process.

During the preliminary experiments, the SOPs were established mainly based on the WFUMB guidelines [8,9], and the operator received training in the SOPs. The SOPs are presented in detail in Supplementary Data 1.

The rats were fasted for 4 hours before testing. Under anesthesia with isoflurane, each rat was positioned in the supine position with both anterior limbs abducted. After shaving the upper abdomen, the transducer was gently applied with a large amount of sonographic gel warmed to 37°C to achieve a good acoustic window, while avoiding stiffness artifacts radiating from the contact area, hand motion, and pressure. Measurements were taken through a subcostal approach at the location of the best acoustic window of the left hepatic lobe. If possible, the transducer was held perpendicular to the liver capsule. The operator measured the liver stiffness at a depth of approximately 1 cm from the probe surface in all animals.

In each SWE examination, liver stiffness measurements in kilopascals (kPa) were made eight times in the same location by placing a sampling box (i.e., a field of view), where a color map was displayed. For each measurement, the radiographer drew several (generally three or four) regions of interest (ROIs) with a diameter of 0.3 cm away from the large vessel and bile duct. Then, the radiographer and supervising radiologist chose an ROI that was placed in the area of greatest shear wave uniformity, demonstrating a relatively homogeneous color on the liver stiffness map and speed map (Fig. 1). Of the eight measurement values, we selected the median value from each SWE examination as the liver stiffness value.

During measurements of liver stiffness, we did not consider the

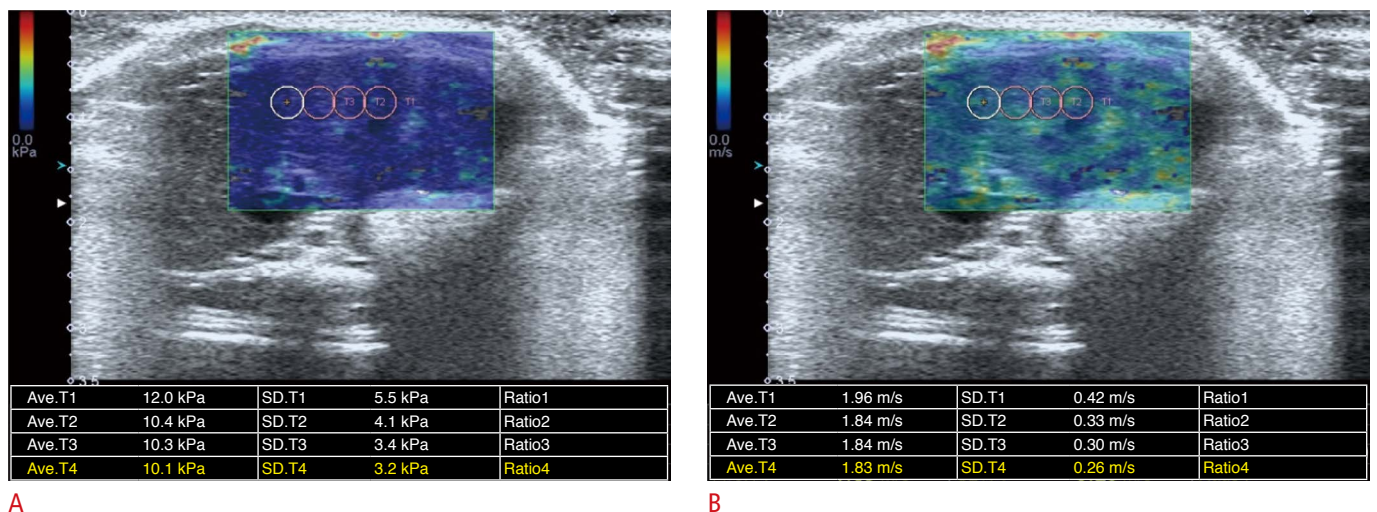


Fig. 1. Measurement of shear wave elastography. Regions of interest were positioned in an area of homogeneous color on an elasticity map (A) and a speed map (B). A. Liver stiffness map shows the distribution of stiffness (kPa) in the liver. B. Speed map shows the distribution of shear wave velocities in the liver.

interquartile/median (IQR/M) value as a quality check criterion. After finishing our experiment, we manually calculated the IQR/M values of the eight measurements in each SWE examination.

Statistical Analysis

The mean values and standard deviation (SD) were determined for all SWE examinations for each group in each measurement session. Repeated-measures analysis of variance (RMANOVA) was used to compare the mean values of liver stiffness (kPa) between measurement sessions and across groups. Other quantitative values were compared using the Student t-test, Fisher exact test, or one-way analysis of variance, with *post-hoc* analysis using the Tukey-Kramer method. Pearson correlation coefficients were calculated to evaluate correlations between liver stiffness (kPa) and collagen area. A $P < 0.05$ was considered to indicate statistical significance. MedCalc version 17.7.2 (MedCalc Software bvba, Ostend, Belgium) was used for all analyses.

The measurement variability of liver stiffness between the first session and the second session was evaluated as follows: (1) In each group, the coefficient of variation (CoV) was calculated by dividing the SD by the mean liver stiffness value, and the CoV was compared between sessions using the Levene test. (2) In each SWE examination, the IQR/M value was calculated, and the mean IQR/M values of all SWE examinations were compared between sessions using the Student t-test [24].

To determine test-retest repeatability, we calculated the repeatability coefficient and within-subject coefficient of variation (WSCV), which are statistical methods recommended by the Quantitative Imaging Biomarker Alliance (QIBA), using a web-based calculator (available at <http://datasharing.aim-aicro.com/reliability>) [25–27]. Test-retest repeatability was also evaluated using Bland-Altman plots with a mean relative difference (percentage) and 95% limits of agreement (LOA) [28]. The agreement between measurements from the first and second sessions was visually assessed using Bland-Altman plots, in which the relative difference between the two measurements was plotted against the average value. The 95% LOA is the range of values within ± 1.96 SD of the mean relative difference.

Results

Histopathologic Characteristics of the Animal Models

Among the 34 rats, three (two in the low-dose group and one in the high-dose group) died during the 8 weeks of TAA injection. Finally, eight rats in the control group, nine in the low-dose group, and 14 in the high-dose group were included in this study. Signs of toxicity, such as ruffled fur, anorexia, cachexia, skin tenting, skin ulcerations,

or toxic death were not seen in any of the surviving rats.

All rats in the control group showed normal histologic findings without fibrosis, inflammation, or steatosis. Based on the histopathologic results, liver fibrosis with abundant collagen deposits was successfully induced in all rats in both the low-dose and high-dose groups (Fig. 2A). The areas of collagen were most extensive in the liver histopathologic specimens of the high-dose group ($24.86\% \pm 4.55\%$), followed by the low-dose group ($16.01\% \pm 3.25\%$) and the control group ($6.27\% \pm 2.10\%$), as presented in Fig. 2B. *Post-hoc* multiple comparison analysis revealed that all pairs were significantly different ($P < 0.05$, Tukey-Kramer test).

SWE per Group

The mean values of the first and second liver stiffness measurements were 8.88 ± 1.48 kPa in the control group, 11.62 ± 1.70 kPa in the low-dose group, and 11.91 ± 1.73 kPa in the high-dose group. Liver stiffness differed significantly between groups ($P < 0.001$; between-subjects effects; RMANOVA). A *post-hoc* analysis showed that liver stiffness significantly differed between the control group and low-dose group ($P < 0.001$), as well as between the control group and high-dose group ($P < 0.001$), but did not differ between the low-dose and high-dose groups.

Regarding the correlation between liver stiffness and liver fibrosis, a significant positive correlation was found between mean liver stiffness measurements and the areas of collagen in the liver specimens ($r = 0.6$, $P = 0.005$) (Fig. 3).

SWE per Session

The box-and-whiskers plots of liver stiffness measured at the first and second sessions are presented in Fig. 4A. When comparing the first and second measurement sessions, no significant difference was found in the mean values between the first and second measurements in all groups ($P > 0.05$; within-subject effects; RMANOVA).

However, as presented in Table 1, the second measurement session yielded more reliable results than the first measurement session, as follows: (1) the CoV, a measure of variability in a group, was higher in the first session than the second session ($P < 0.001$ in all groups), and (2) the mean values of IQR/M, a measure of variability in an SWE examination, were also higher in the first session than the second session (0.21 ± 0.12 vs. 0.12 ± 0.08 , $P < 0.001$). In addition, four outliers were present in the first session, but not in the second (Fig. 4B).

Repeatability Coefficient

The repeatability coefficient of liver stiffness between the first and

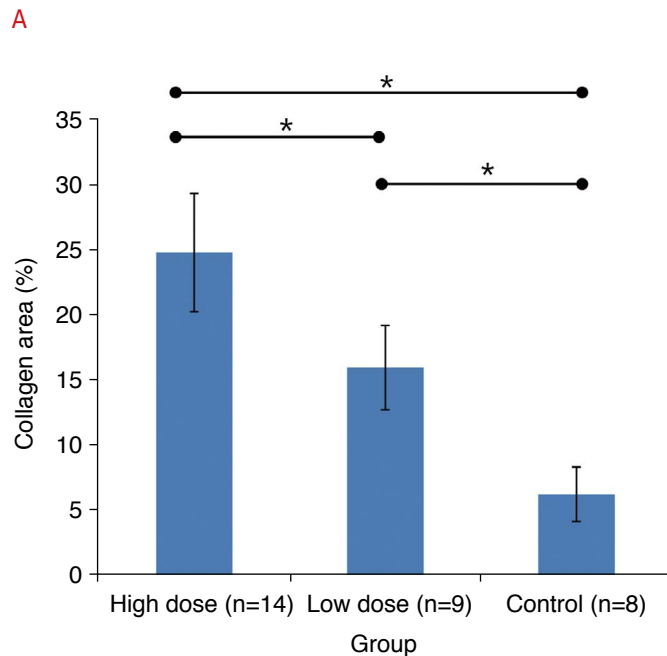
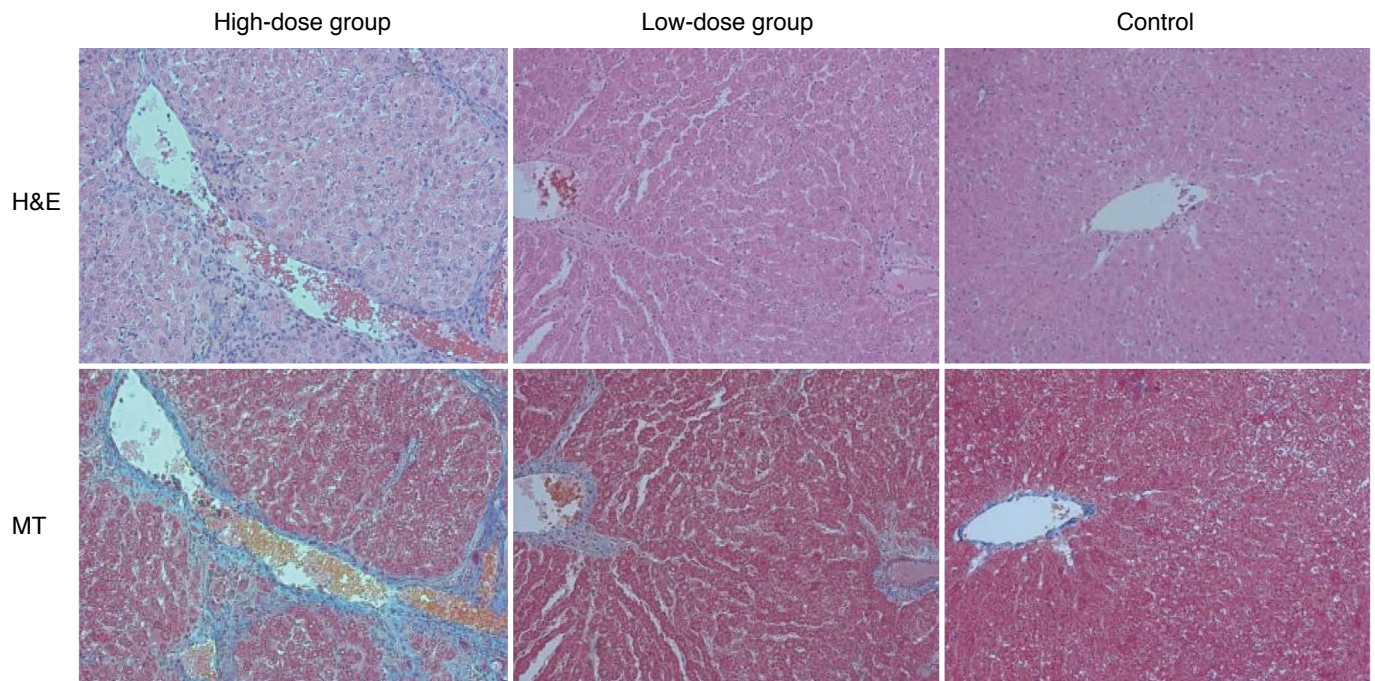


Fig. 2. Histopathology of the liver parenchyma.

A. In hematoxylin and eosin (H&E) stain ($\times 200$), liver inflammation is assessed based on the presence of inflammation cells in the hepatic lobules. In Masson's trichrome (MT) stain ($\times 200$), liver fibrosis is analyzed based on the presence of collagen deposition with blue color. **B.** The graph shows the collagen area was the highest in the high-dose group, followed by low-dose group and control group. * $P < 0.05$.

B

second measurement sessions was 3.75 kPa (95% confidence interval, 3.01 to 4.99 kPa; WSCV, 12.26%). In the Bland-Altman analysis, the 95% LOA was -3.05 to 4.22 kPa. When the four outliers were removed, the repeatability coefficient decreased to 2.82 kPa (95% confidence interval, 2.24 to 3.85 kPa; WSCV, 9.29%) and the Bland-Altman 95% LOA also became narrower (-2.12 to 3.21 kPa) (Fig. 5).

Discussion

In our experiment, we evaluated short-term test-retest repeatability at a 3-day interval in a rat liver fibrosis model, in order to provide repeatability data for preclinical trials for anti-fibrotic agents. The repeatability coefficient of liver stiffness between the first and second measurement sessions was 3.75 kPa (WSCV, 12.26%) which decreased to 2.82 kPa (WSCV, 9.29%) after removing four outliers

Table 1. Liver stiffness measurements and variability between sessions

	First session	Second session	P-value
Liver stiffness values in each group, mean±SD (kPa)			
Control group (n=8)	9.03±2.02	8.72±0.75	0.690 ^{a)}
Low-dose group (n=9)	11.97±2.03	11.26±1.32	0.392 ^{a)}
High-dose group (n=14)	12.24±2.17	11.58±1.12	0.321 ^{a)}
CoV of liver stiffness values in each group			
Control group (n=8)	0.22	0.08	<0.001 ^{b)}
Low-dose group (n=9)	0.16	0.11	<0.001 ^{b)}
High-dose group (n=14)	0.17	0.09	<0.001 ^{b)}
IQR/M of 8 measurements in each SWE examination (n=31)			
Mean±SD	0.21±0.12	0.12±0.08	<0.001 ^{a)}
Examinations with IQR/M <0.3 ^{d)} , n (%)	27 (87.1)	31 (100)	0.113 ^{c)}

CoV, coefficient of variation; IQR/M, interquartile/median; SWE, shear wave elastography.

^{a)}Student t-test. ^{b)}Levene test. ^{c)}Fisher exact test. ^{d)}The 2018 World Federation for Ultrasound in Medicine and Biology guideline recommends using an IQR/M <0.3 as a measure of quality [9].

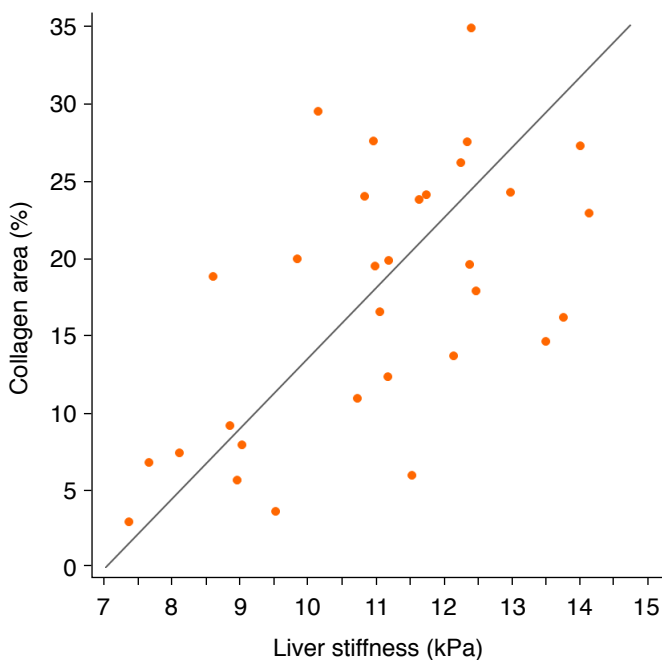


Fig. 3. Correlation between liver stiffness (kPa) and the area of collagen (%) in liver specimens.

from the first measurement session.

Extensive research has been performed in clinical practice to obtain accurate and reliable SWE data, revealing that many technical factors, including measurement depth, ROI location and selection, number of measurements, and patient position and respiration, can affect the results of ultrasound elastography [29]. The WFUMB guidelines advised that variability can be decreased by using SOPs and implementing a learning curve through suitable training and

experience [8,9].

In our preclinical study, we standardized the SWE method as follows: the measurement depth was approximately 1 cm below the probe surface, ROIs were selected from areas with homogeneous color on the speed map of the left liver, and measurements were performed at least eight times. In rats, the subcostal approach to the left lobe is the best way to achieve a good sonic window of the liver. However, the subcostal approach is sensitive to compression of the liver by the probe [30]. Thus, we placed the ultrasonographic probe very carefully to avoid pressure.

Despite these efforts, there were four outliers in the first measurement session. When we reviewed the outliers, the ROIs had been placed in areas where the speed maps showed a heterogeneous color map with several non-filling pixels, indicative of invalid shear wave characteristics. We learned that making measurements in areas with visible homogeneity within the ROIs on images with a good sonic window may improve artifactual and technical sources of variation [31]. In the second session, the radiographer’s skill might have improved in terms of obtaining good sonic windows, selecting sampling boxes, and placing ROIs in the areas of greatest shear wave uniformity. Based on our experience, the training effect might be summarized as adherence to the protocol [32].

In our study, the test-retest repeatability of liver stiffness measurements on SWE was generally good because a well-trained operator measured the same rats twice at a 3-day interval following a trial-specific standardized protocol. The repeatability might be worse in routine preclinical trials that monitor drug efficacy once every few weeks or months. Training might be necessary before each

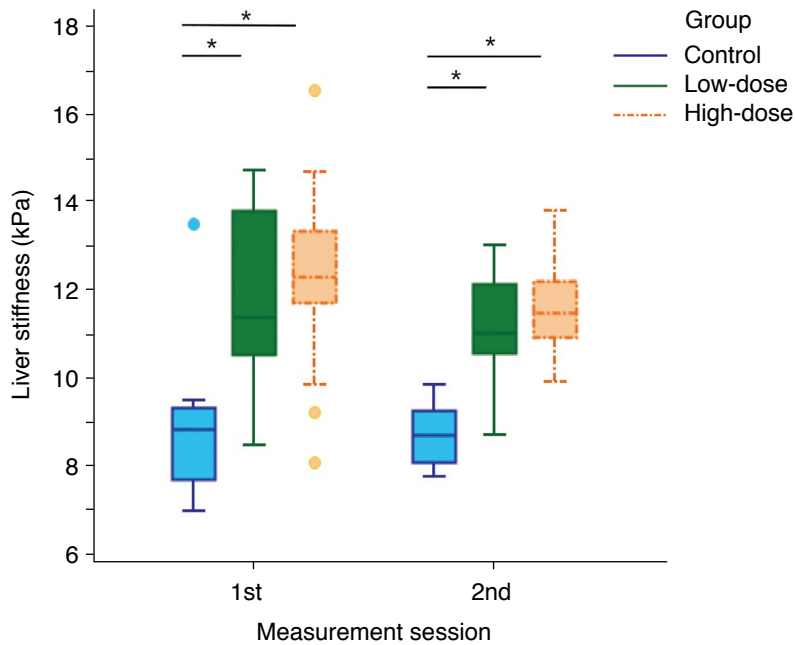
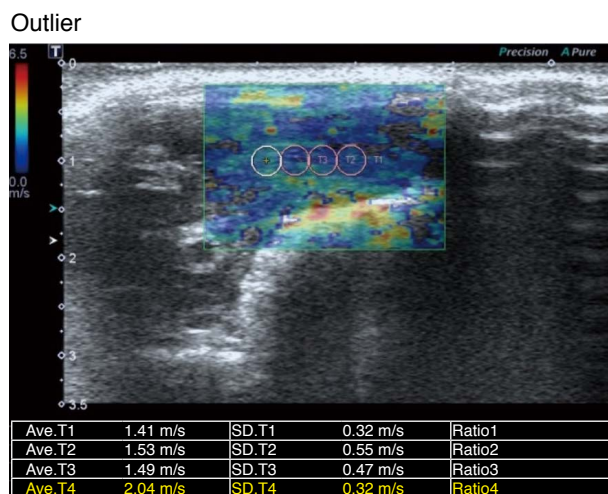
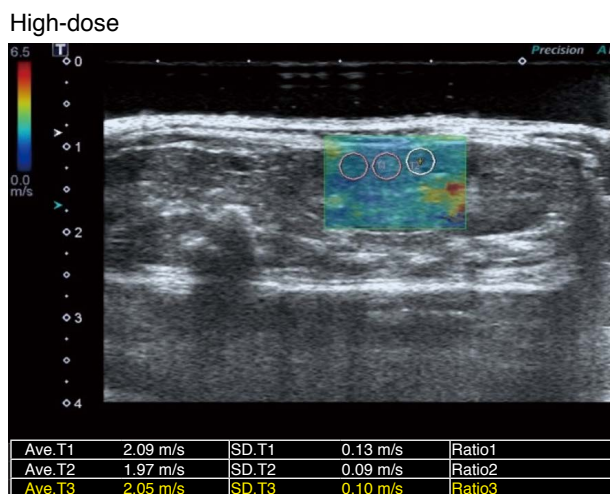
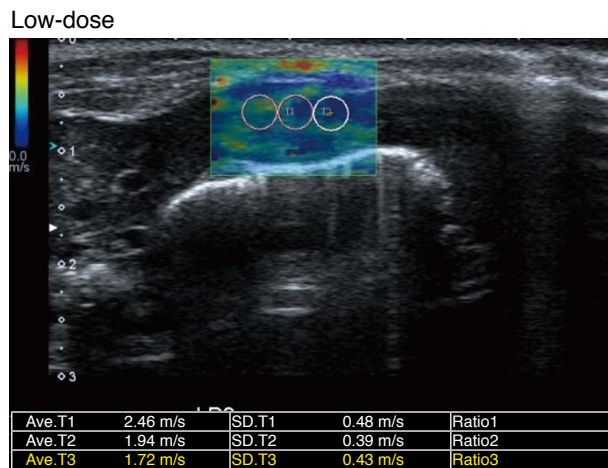
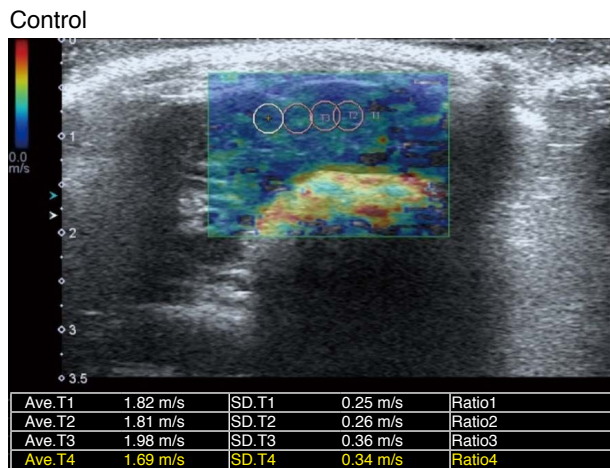


Fig. 4. Liver stiffness of each measurement session. **A.** Box and whiskers plots show liver stiffness measured in the first and second sessions. The interquartile ranges were higher in the first session than in the second session in all groups. The outliers in the first measurement session are presented as round dots. *P<0.05. **B.** Shear wave speed maps demonstrate sampling boxes where color maps are displayed in the control, low-dose, and high-dose groups and the outliers.

A



B

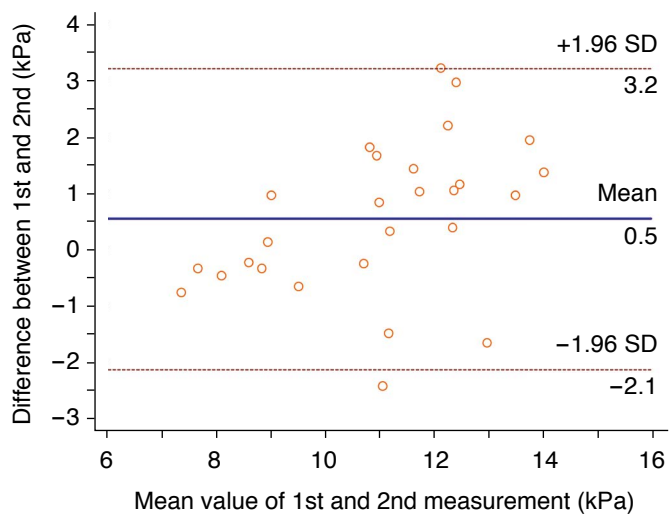


Fig. 5. Bland-Altman plot to evaluate the repeatability of liver stiffness measurements. SD, standard deviation.

measurement session to minimize intra-operator variability.

In our experiment, the areas of collagen observed in histopathologic examinations were significantly different between the high-dose group ($24.86\% \pm 4.55\%$) and the low-dose group ($16.01\% \pm 3.25\%$), whereas liver stiffness measurements were not significantly different between the high-dose group and low-dose group. We postulate that quantitative histologic measurements of collagen area might be more sensitive than SWE as a way to quantify liver fibrosis. The histologic measurements used five representative hotspots with substantial areas of Masson trichrome staining, whereas SWE measurements were made in areas with shear wave uniformity. Discrepancies between histologically observed liver fibrosis and liver stiffness measurements have been also reported in prior research. Georges et al. [33] reported that the liver stiffness increased marginally in the late treatment period starting on the 11th day after the start of fibrosis-inducing treatment, while the amount of liver fibrosis grew substantially. Indeed, the degree of correlation between areas of collagen and liver stiffness on SWE was moderate ($r=0.6$) in our study, similarly to prior animal studies [34,35]. Wang et al. [34] reported a correlation of $r^2=0.43$ (equivalent to $r=0.65$) and attributed the discrepancies to differences in the liver locations where the histologic analysis and SWE were performed.

There are several limitations in our study. First, we did not measure long-term repeatability, because the liver fibrosis may change in the long term. Second, our test-retest repeatability results at a 3-day interval might be underestimated compared to those of same-day repeatability. The reason why we chose a 3-day interval between the two measurement sessions was that we aimed to include the entire

measurement process, including animal preparation, anesthesia, and SWE examination. If we had evaluated same-day repeatability while maintaining anesthesia, we could not have evaluated the impact of animal preparation and anesthesia. Third, the severity of liver fibrosis may change during a 3-day period. However, we believe that such changes were likely very minimal, because we discontinued administration of TAA during the 3-day interval. In addition, we believe that a 3-day interval is appropriate to minimize futile animal loss. In studies where experiments are conducted at 1-day or 2-day intervals, the rats might be distressed due to repeated anesthesia in such a short time. Lastly, our study only evaluated intra-reader variability with an experienced radiographer. However, we have provided data on inter-reader variability obtained from our preliminary experiment in Supplementary Data 2.

In conclusion, in a preclinical trial with an animal liver fibrosis model, the feasibility of using ultrasonographic SWE to evaluate histopathologic liver fibrosis in a non-invasive and repeatable manner was confirmed. The operator's experience and training in how to adhere to the SOPs might contribute to reducing the variability of liver stiffness measurements on SWE.

ORCID: Youngbin Shin: <https://orcid.org/0000-0003-2753-9586>; Jimi Huh: <https://orcid.org/0000-0002-8832-6165>; Su Jung Ham: <https://orcid.org/0000-0003-1455-9262>; Young Chul Cho: <https://orcid.org/0000-0001-5469-5467>; Yoonseok Choi: <https://orcid.org/0000-0001-9291-1714>; Dong-Cheol Woo: <https://orcid.org/0000-0001-8202-015X>; Jeongjin Lee: <https://orcid.org/0000-0002-4151-6922>; Kyung Won Kim: <https://orcid.org/0000-0002-1532-5970>

Author Contributions

Conceptualization: Huh J, Kim KW. Data acquisition: Ham SJ, Cho YC, Choi Y. Data analysis or interpretation: Shin Y, Woo DC, Lee J. Drafting of the manuscript: Shin Y, Huh J, Kim KW. Critical revision of the manuscript: Choi Y, Woo DC, Lee J. Approval of the final version of the manuscript: all authors.

Conflict of Interest

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

Acknowledgments

This study was supported by a grant from the National Research Foundation of Korea (2019R111A1A01060744). We thank the Experimental Animal Histopathology Laboratory core facility (pathologist: Dr. Woo Chan Son) at Asan Medical Center for assisting with the histopathologic analysis.

Supplementary Material

Supplementary Data 1. Standard operating procedure (English-

translated version) (<https://doi.org/10.14366/usg.19088>).

Supplementary Data 2. Shear wave elastography (SWE) results of the preliminary study (<https://doi.org/10.14366/usg.19088>).

References

- Sugimoto K, Moriyasu F, Oshiro H, Takeuchi H, Yoshimasu Y, Kasai Y, et al. Clinical utilization of shear wave dispersion imaging in diffuse liver disease. *Ultrasonography* 2020;39:3-10.
- Hong EK, Choi YH, Cheon JE, Kim WS, Kim IO, Kang SY. Accurate measurements of liver stiffness using shear wave elastography in children and young adults and the role of the stability index. *Ultrasonography* 2018;37:226-232.
- 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.
- Chen YW, Tsai MY, Pan HB, Tseng HH, Hung YT, Chou CP. Gadoteric acid-enhanced MRI and sonoelastography: non-invasive assessments of chemoprevention of liver fibrosis in thioacetamide-induced rats with Sho-Saiko-To. *PLoS One* 2014;9:e114756.
- You MW, Kim KW, Pyo J, Huh J, Kim HJ, Lee SJ, et al. A meta-analysis for the diagnostic performance of transient elastography for clinically significant portal hypertension. *Ultrasound Med Biol* 2017;43:59-68.
- Kim DW, Suh CH, Kim KW, Pyo J, Park C, Jung SC. Technical performance of two-dimensional shear wave elastography for measuring liver stiffness: a systematic review and meta-analysis. *Korean J Radiol* 2019;20:880-893.
- Suh CH, Kim KW, Park SH, Lee SS, Kim HS, Tirumani SH, et al. Shear wave elastography as a quantitative biomarker of clinically significant portal hypertension: a systematic review and meta-analysis. *AJR Am J Roentgenol* 2018;210:W185-W195.
- Ferraioli G, Filice C, Castera L, Choi BI, Sporea I, Wilson SR, et al. WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 3: liver. *Ultrasound Med Biol* 2015;41:1161-1179.
- Ferraioli G, Wong VW, Castera L, Berzigotti A, Sporea I, Dietrich CF, et al. Liver ultrasound elastography: an update to the World Federation for Ultrasound in Medicine and Biology guidelines and recommendations. *Ultrasound Med Biol* 2018;44:2419-2440.
- Bayly PV, Garbow JR. Pre-clinical MR elastography: principles, techniques, and applications. *J Magn Reson* 2018;291:73-83.
- Huh J, Ham SJ, Cho YC, Park B, Kim B, Woo CW, et al. Gadoteric acid-enhanced dynamic contrast-enhanced MRI for evaluation of liver function and liver fibrosis in preclinical trials. *BMC Med Imaging* 2019;19:89.
- Sugimoto K, Moriyasu F, Oshiro H, Takeuchi H, Yoshimasu Y, Kasai Y, et al. Viscoelasticity measurement in rat livers using shear-wave US elastography. *Ultrasound Med Biol* 2018;44:2018-2024.
- Li J, Liu H, Zhang C, Yang S, Wang Y, Chen W, et al. Native T1 mapping compared to ultrasound elastography for staging and monitoring liver fibrosis: an animal study of repeatability, reproducibility, and accuracy. *Eur Radiol* 2020;30:337-345.
- Hollerith K, Gassmann B, Wagenpfeil S, Kemmner S, Heemann U, Stock KF. Does standoff material affect acoustic radiation force impulse elastography? A preclinical study of a modified elastography phantom. *Ultrasonography* 2018;37:140-148.
- Yoo J, Lee JM, Joo I, Lee DH, Yoon JH, Kang HJ, et al. Reproducibility of ultrasound attenuation imaging for the noninvasive evaluation of hepatic steatosis. *Ultrasonography* 2020;39:121-129.
- Yoo J, Lee JM, Joo I, Yoon JH. Assessment of liver fibrosis using 2-dimensional shear wave elastography: a prospective study of intra- and inter-observer repeatability and comparison with point shear wave elastography. *Ultrasonography* 2020;39:52-59.
- Ryu H, Ahn SJ, Yoon JH, Lee JM. Inter-platform reproducibility of liver stiffness measured with two different point shear wave elastography techniques and 2-dimensional shear wave elastography using the comb-push technique. *Ultrasonography* 2019;38:345-354.
- Park SY, Shin HW, Lee KB, Lee MJ, Jang JJ. Differential expression of matrix metalloproteinases and tissue inhibitors of metalloproteinases in thioacetamide-induced chronic liver injury. *J Korean Med Sci* 2010;25:570-576.
- Kim YO, Popov Y, Schuppan D. Optimized mouse models for liver fibrosis. *Methods Mol Biol* 2017;1559:279-296.
- Liedtke C, Luedde T, Sauerbruch T, Scholten D, Streetz K, Tacke F, et al. Experimental liver fibrosis research: update on animal models, legal issues and translational aspects. *Fibrogenesis Tissue Repair* 2013;6:19.
- Collins TJ. ImageJ for microscopy. *Biotechniques* 2007;43(1 Suppl):25-30.
- Chen P, Kakan X, Wang S, Dong W, Jia A, Cai C, et al. Deletion of clock gene *Per2* exacerbates cholestatic liver injury and fibrosis in mice. *Exp Toxicol Pathol* 2013;65:427-432.
- Ham SJ, Choi Y, Lee SI, Kim J, Kim YI, Chung JW, et al. Enhanced efficacy of radiofrequency ablation for hepatocellular carcinoma using a novel vascular disrupting agent, CKD-516. *Hepatol Int* 2017;11:446-451.
- Lim S, Kim SH, Kim Y, Cho YS, Kim TY, Jeong WK, et al. Coefficient of variance as quality criterion for evaluation of advanced hepatic fibrosis using 2D shear-wave elastography. *J Ultrasound Med* 2018;37:355-362.
- Park JE, Han K, Sung YS, Chung MS, Koo HJ, Yoon HM, et al. Selection and reporting of statistical methods to assess reliability of a diagnostic test: conformity to recommended methods in a peer-reviewed journal. *Korean J Radiol* 2017;18:888-897.

26. Barnhart HX, Barboriak DP. Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl Oncol* 2009;2:231-235.
27. Kessler LG, Barnhart HX, Buckler AJ, Choudhury KR, Kondratovich MV, Toledano A, et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Stat Methods Med Res* 2015;24:9-26.
28. Krajewski KM, Nishino M, Franchetti Y, Ramaiya NH, Van den Abbeele AD, Choueiri TK. Intraobserver and interobserver variability in computed tomography size and attenuation measurements in patients with renal cell carcinoma receiving antiangiogenic therapy: implications for alternative response criteria. *Cancer* 2014;120:711-721.
29. Park SH, Kim SY, Suh CH, Lee SS, Kim KW, Lee SJ, et al. What we need to know when performing and interpreting US elastography. *Clin Mol Hepatol* 2016;22:406-414.
30. Karlas T, Pfrepper C, Wiegand J, Wittekind C, Neuschulz M, Mossner J, et al. Acoustic radiation force impulse imaging (ARFI) for non-invasive detection of liver fibrosis: examination standards and evaluation of interlobe differences in healthy subjects and chronic liver disease. *Scand J Gastroenterol* 2011;46:1458-1467.
31. Nadebaum DP, Nicoll AJ, Sood S, Gorelik A, Gibson RN. Variability of liver shear wave measurements using a new ultrasound elastographic technique. *J Ultrasound Med* 2018;37:647-656.
32. Maruyama H, Kobayashi K, Kiyono S, Sekimoto T, Kanda T, Yokosuka O. Two-dimensional shear wave elastography with propagation-based reliability assessment for grading hepatic fibrosis and portal hypertension. *J Hepatobiliary Pancreat Sci* 2016;23:595-602.
33. Georges PC, Hui JJ, Gombos Z, McCormick ME, Wang AY, Uemura M, et al. Increased stiffness of the rat liver precedes matrix deposition: implications for fibrosis. *Am J Physiol Gastrointest Liver Physiol* 2007;293:G1147-G1154.
34. Wang MH, Palmeri ML, Guy CD, Yang L, Hedlund LW, Diehl AM, et al. In vivo quantification of liver stiffness in a rat model of hepatic fibrosis with acoustic radiation force. *Ultrasound Med Biol* 2009;35:1709-1721.
35. Chen X, Shen Y, Zheng Y, Lin H, Guo Y, Zhu Y, et al. Quantification of liver viscoelasticity with acoustic radiation force: a study of hepatic fibrosis in a rat model. *Ultrasound Med Biol* 2013;39:2091-2102.