

83511_Auto_Edited.docx

Name of Journal: *World Journal of Gastroenterology*

Manuscript NO: 83511

Manuscript Type: ORIGINAL ARTICLE

Observational Study

Artifacts in two-dimensional shear wave elastography of liver

Wang HP *et al.* Artifacts in liver elastography

Hui-Peng Wang, Peng-Chao Zheng, Xue-Mei Wang, Liang Sang

Abstract

BACKGROUND

Artifacts are common when using two-dimensional shear wave elastography (2-D SWE) to measure liver stiffness (LS), but they are poorly recognized.

AIM

To investigate the presence and influence of artifacts in 2-D SWE of liver.

METHODS

We included 158 patients with chronic liver disease, who underwent 2-D SWE examination by a novice and an expert. A cross line at the center of the elastogram was drawn and was divided it into four locations: top-left, top-right, bottom-left, and bottom-right. The occurrence frequency of artifacts in different locations was compared. The influence of artifacts on the LS measurements was evaluated by comparing the elastogram with the most artifacts (EMA) and the elastogram with the least artifacts (ELA).

RESULTS

The percentage of elastograms with artifacts in the novice (51.7%) was significantly higher than that of the expert (19.6%) ($P < 0.001$). It was found that both operators had the highest frequency of artifacts at bottom-left, followed by top-left and bottom-right, and top-right had the lowest frequency. The LS values (LSVs) and standard deviation values of EMAs were significantly higher than those of ELAs for both operators. An intraclass correlation coefficient value of 0.96 was found in the LSVs of EMAs of the two operators, and it increased to 0.98 when the LSVs of the ELAs were used. Both operators had lower stability index values for EMAs than ELAs, but the difference was only statistically significant for the novice.

CONCLUSION

Artifacts are common when using 2-D SWE to measure LS, especially for the novice. Artifacts may lead to the overestimation of LS and reduce the repeatability and reliability of LS measurements.

Key Words: Ultrasound; Elastography; Artifact; Liver

Wang HP, Zheng PC, Wang XM, Sang L. Artifacts in two-dimensional shear wave elastography of liver. *World J Gastroenterol* 2023; In press

Core Tip: Artifacts are common when using two-dimensional shear wave elastography (2-D SWE) to measure liver stiffness (LS), especially for the novice. We investigated the presence and influence of artifacts in 2-D SWE of liver. Our results showed artifacts were more likely to occur in the bottom-left corner of the elastogram. Artifacts may lead to the overestimation of LS and reduce the repeatability and reliability of LS measurements. For the elastograms with artifacts, we should place the Q-Box away from the artifacts.

¹ **INTRODUCTION**

Chronic liver disease is a growing problem worldwide. The main causes of chronic liver disease include hepatitis virus infection, ³alcoholic liver disease, and non-alcoholic fatty liver disease^[1]. It mainly causes diffuse liver fibrosis, which in turn leads to liver cirrhosis. Some of them eventually develop hepatocellular carcinoma, portal hypertension, and hepatic encephalopathy^[2,3]. ¹Accurate assessment of liver fibrosis is important for treatment prioritization, surveillance, and determination of prognosis^[4]. Moreover, liver biopsy allows the assessment of the degree of fibrosis^[5]. However, liver biopsy is an expensive and invasive diagnostic tool. Its main complications are bleeding and pain^[6,7], which limit its clinical application.

Recently, the application of ultrasound elastography in the diagnosis of non-invasive assessment of liver fibrosis has developed rapidly^[8]. US elastography is mainly

classified into two major types: ¹Strain elastography and shear wave elastography^[9,10]. Two-dimensional shear wave elastography (2-D SWE) is a type of shear wave elastography that uses acoustic radiation force to create shear waves. The velocity of the ³shear wave can be used to calculate the tissue stiffness by the formula $E = 3\rho c^2$, where E is tissue elasticity (Young's modulus, kPa), ρ is tissue density (kg/m³), and c is shear wave velocity (m/s). The 2-D SWE is based on the quantification of the propagation speed of shear waves in the liver to create an elastogram. ¹The elastogram is displayed using a color-coded map superimposed on a conventional B-mode image, where different colors represent different stiffness, allowing an assessment of homogeneity^[10].

⁸It has been reported that 2-D SWE has shown sufficient accuracy in evaluating the degree of liver fibrosis^[11-13]. However, there was significant heterogeneity in the results of these studies. This heterogeneity may be caused by different patient populations, research designs and equipment used^[14]. Another important reason may be that the presence of artifacts leads to inaccurate liver stiffness (LS) measurements. Bruce *et al*^[15] reported that 2-D SWE artifacts resulted in a significant variability in the assessed LS.

Although 2-D SWE artifacts of the liver are common in clinical practice, they are poorly recognized, and there is even no clear definition. To the best of our knowledge, only a few review articles have been published^[15,16]. Therefore, ²the purpose of this study was to investigate the presence and effects of artifacts in 2-D SWE of the liver. This is important to avoid artifacts and improve diagnostic performance in future operations.

MATERIALS AND METHODS

Patient selection

This prospective study was approved by the institutional ethical review board of our hospital. All patients signed a written informed consent document to participate in the study. We included 158 consecutive patients with chronic liver disease, who underwent 2-D SWE examination in our department. The study was conducted according to the principles reported in the Declaration of Helsinki and approved by the authors'

institutional review board. The exclusion criterion was that no valid measurement was obtained by either operator. Seven patients were excluded because the novice operator did not obtain any valid measurements after five consecutive measurements. The baseline characteristics of the patients were presented in Table 1.

2-D SWE examination

LS measurements were performed with an Aixplorer US system (SuperSonic Imagine, Aix-en-Provence, France) with a convex probe (SC6-1, 1-6 MHz). Patients fasted for more than 6 h and were examined in the supine position with the right arm in maximal abduction. The right anterior lobe of the liver was examined by intercostal scanning, and the SWE mode was started with neutral breathing during breath-holding. The upper limit of the color-coding scale was set to 70 kPa. The sampling frame was approximately 2.5 cm × 3.5 cm, placed at least 1 cm below the liver capsule, avoiding the large vascular structures. Image acquisition was performed after the elastography image was stable for 3-5 s. The quantitative analysis system (Q-Box) was then activated and placed at the center of the sampling frame. The Q-Box was 2 cm in diameter and the measurement depth was 3-5 cm. The LS measurement was considered invalid if there was no color-coding or the coded area was smaller than the Q-Box size^[17]. When the area of color-coding is larger than the Q-Box size, the LS measurement was considered valid even if there are artifacts within it.

Each patient was continuously measured five times by an expert and a novice, respectively. The operators performed consecutive LS measurements in a randomized blinded manner. The median value of all valid measurements performed by the two operators represents the LS value (LSV) of the subject and was used for the correlation analysis with artifacts. The expert operator had 9 years of experience in the 2-D SWE examinations and had successfully performed approximately 15000 2D SWE examinations. The novice operator was trained by an expert operator and successfully performed 50 2-D SWE examinations.

Analysis of elastogram for artifacts

Artifacts were defined as the mottled area in the elastograms, and the area of the artifacts was measured using a tracing instrument attached to the device. We can manually trace the edge of the artifacts and automatically display the area and perimeter of the artifacts (Figure 1A). We drew a cross line at the center of the elastogram and divided it into four locations: top-left, top-right, bottom-left, and bottom-right. The location of the artifacts in each elastogram was recorded. The elastogram with the most artifacts (EMA) and the elastogram with the least artifacts (ELA) in each patient measured by the two operators were found by all authors. For the elastograms with artifacts, the Q-Box was placed in the center of the sampling frame (Figure 1B) and away from the artifacts for measurements (Figure 1C). The influence of artifacts on LS measurement was evaluated by comparing the differences in LSVs, standard deviation (SD) values and stability index (SI) values.

Statistical analysis

All quantitative data are expressed as mean \pm SD (range), and qualitative variables are expressed as numbers (percentages). The Shapiro-Wilk test was used to test whether the numeric variables were normally distributed. Non-parametric tests with the Kruskal-Wallis method were used to compare the difference in numeric variables with a non-normal distribution. Differences between numeric variables with a normal distribution were assessed using a parametric test (*t*-test). The χ^2 -test was used to compare the proportions expressed as percentages. Interobserver repeatability was evaluated using intraclass correlation coefficient (ICC). Relationships between various parameters were examined using Pearson's correlation test. Statistical significance was set at $P < 0.05$, and all *P*-values were two-sided. Statistical analysis was performed using MedCalc software (MedCalc Software, version 17.4, Ostend, Belgium).

RESULTS

Among the 158 patients, 151 patients with valid measurements obtained by both operators were enrolled in this study. In theory, each operator should obtain 755 ($151 \times 5 = 755$) elastography images. However, in the examination of 12 patients by the two operators, 35 elastography images were invalid and excluded. To ensure that the two operators had the same number of valid elastograms for each patient, valid measurements corresponding to the 35 invalid measurements were also excluded. Therefore, 720 elastography images from each operator were included (Figure 2).

The presence of artifacts

For the expert operator, the percentage of elastograms with artifacts was 19.6% (141/720), and the area of artifacts was $0.92 \pm 0.68 \text{ cm}^2$. For the novice operator, the percentage of elastograms with artifacts was 51.7% (372/720), and the area of artifacts was $1.36 \pm 0.87 \text{ cm}^2$. The percentage of elastograms with artifacts and the area of artifacts in the novice were significantly higher than those in the expert, and the difference were both statistically significant (both $P < 0.001$). We counted all the artifacts according to their locations, and the results are shown in Table 2. There were no significant differences in the frequency of the occurrence of artifacts between the two operators at the same location (all $P > 0.05$). Comparing the occurrence frequency of artifacts in all locations of the two operators, it was found that both operators had the highest frequency of bottom-left, followed by top-left and bottom-right, and top-right had the lowest frequency. No statistical difference was found between the frequency of top-left and bottom-right ($P > 0.05$), but the frequency among other locations was statistically different (all $P < 0.001$) (Table 3).

Influence of artifacts on LS measurements

The LSVs of EMAs were higher than those of ELAs for both operators, and the differences were statistically significant (both $P < 0.001$). There was a significant difference in the LSVs of the EMAs between the two operators ($P = 0.006$). However,

there was no statistically significant difference in the LSVs of the ELAs between the two operators ($P = 0.051$) (Table 4).

The ICC values and 95% CIs were calculated by comparing the LSVs of the EMAs and ELAs of the two operators. An ICC value of 0.96 (95% CI: 0.94-0.98) was found in the LSVs of EMAs, and it increased to 0.98 (95% CI: 0.97-0.99) when the LSVs of the ELAs were used. The SD values of EMAs were higher than those of ELAs for both operators, and the differences were statistically significant (both $P < 0.001$). The SI values of the EMAs were lower than those of the ELAs for both operators. The difference was only statistically significant for the novice ($P = 0.002$), but not for the expert ($P = 0.135$) (Table 5).

For the elastograms with artifacts, the LSVs and SD values of the Q-Box placed in the center of the sampling frame were higher than those of the Q-Box placed away from the artifacts. The SI values of the Q-Box placed in the center of the sampling frame were lower than those of the Q-Box placed away from the artifacts. There were significant differences in LSVs, SD values and SI values between the Q-Box placed in the center of the sampling frame and away from the artifacts for both operators (all $P < 0.05$) (Table 6).

Patient characteristics and artifacts

The total number of elastograms with artifacts measured by the two operators was 513 (141 by the expert, 372 by the novice). The number of elastograms with artifacts in male subjects was 238 (46.4%), and that in female subjects was 275 (53.6%). There was no significant difference between the male and female subjects ($P = 0.378$). Pearson's correlation test showed that there was no significant linear correlation between age and the number of elastograms with artifacts ($r = 0.21$, $P = 0.126$). In the entire cohort, Pearson's correlation test showed that there was a positive correlation between LSV, body mass index (BMI), subcutaneous fat thickness and the number of elastograms with artifacts ($r = 0.47$, $P = 0.001$; $r = 0.41$, $P = 0.002$; and $r = 0.42$, $P = 0.002$, respectively).

DISCUSSION

When using 2-D SWE to measure LS in clinical practice, artifacts are commonly observed in elastograms^[18]. It is difficult for some subjects to obtain satisfactory elastograms, such as obesity, poor acoustic window and inability of the subjects to hold their breath. Despite our best efforts to avoid artifacts, even operators with 9 years of operating experience still have a certain percentage of artifacts. In this study, we compared the difference in the frequency of occurrence artifacts between two different experienced operators. The results showed that the percentage of elastograms with artifacts and the area of artifacts in the novice were significantly higher than that of the expert. This may be because the expert operator can obtain high-quality B-mode imaging, which is required for accurately tracking shear waves^[18]. Previous studies have shown that experts have better repeatability and reliability in measuring LS, which may have an important relationship with the fact that there were few artifacts in their elastograms^[19,20]. Therefore, some studies have suggested that novices should perform at least 300 abdominal US scans or more than 50 supervised 2-D SWE examinations; however, this may not be sufficient^[19,21]. A learning curve has been observed for 2-D SWE, a proportion of operator error would decrease over time^[22].

We divided the elastogram into four locations and calculated the frequency of occurrence of artifacts at each location. The occurrence frequency of artifacts is arranged in descending order: bottom-left, top-left, bottom-right, and top-right. The two operators in this study had the same results, indicating that this difference may have certain regularity. The reason for this result may be that the aerated lung leads to a shadowing artifact on the left side of the B-mode image, which makes it impossible to form a well-defined push beam in this area^[15,23]. On the other hand, to avoid liver capsule reverberation artifacts, the depth of the sampling frame has increased, especially in obese or overweight patients. When the depth exceeds the penetration limit, attenuation artifacts and larger vessels may have more pulsatile artifacts at the bottom of the sampling frame^[16,23]. We found the same phenomenon on another 2-D SWE ultrasound system (Aplio500, Canon, Tochigi, Japan). We found that artifacts were

more likely to occur in the bottom-left corner of the elastogram, where distortion waves were noted in the propagation map of the corresponding site. The distribution of artifacts may also be applicable to other devices of 2-D SWE technology, because they have the same imaging principles.

Usually a color-coding scale of up to 30 kPa is sufficient, but in this study the upper limit of the color-coding scale was set to 70 kPa. The reason is that some patients have an LSV greater than 30 kPa, and a lower color-coding scale setting will make the elastogram appear only in red. At this time, it is impossible to distinguish whether there is an artifact or not. Although the color-coding scale was set to 70 kPa may ignore tiny artifacts, it is easier to show obvious artifacts.

The presence of artifacts affects the assessment of LS, but there is no detailed research report yet. This study showed that the LSVs of the EMAs were higher than those of the ELAs. This indicates that artifacts may lead to the overestimation of LS. This study compared the differences between the two operators in the LSVs of EMAs and ELAs. The results showed that in either the EMAs or ELAs, the LSVs of the novice were higher than that of the expert, which may be due to the higher proportion of artifacts in the elastograms measured by the novice. The ICC value between the two operators calculated with the LSVs of the EMAs was lower than that calculated with the LSVs of ELAs. This shows that artifacts can reduce inter-observer repeatability.

Although the degree of liver fibrosis in chronic liver disease will be slightly different, the color-coded LSmapping image will hardly show obvious mottled area. These mottled areas are considered as artifacts and belong to noise. Some studies use signal-to-noise ratio as the standard to evaluate image quality^[24,25]. The new software version of the device provides SD and SI as indicators to evaluate the reliability of LS measurement^[26-28]. The SD reflects the homogeneity of LSVs in the measurement area of the Q-Box. The higher the SD values, the greater heterogeneity of the LSVs in the measurement area. Thiele *et al*^[29] reported that the diagnostic accuracy for cirrhosis by 2D SWE increased at SD < 1.75 kPa. The SI is an indicator of temporal stability of the measurement area, and the manufacturer recommends that a reliable LS measurement

should have a SI greater than 90%. Our study showed that the SD values of the EMAs were much higher than those of the ELAs, which indicated that artifacts made the elastograms heterogeneous. The SI values of the EMAs were lower than those of the ELAs, which showed that artifacts may reduce the temporal stability of the elastograms. In short, artifacts can reduce the reliability of the LS measurements. For the elastograms with artifacts, we found that placing the Q-Box away from the artifacts can obtain more reliable LS measurements than placing it in the center of the sampling frame (generally the default measurement position of the equipment).

Furthermore, we investigated the relationship between patient characteristics and the occurrence of artifacts. We found that the occurrence of artifacts had no significant relationship with sex or age. However, we found that patients' BMI, subcutaneous fat thickness and LSVs were positively correlated with the occurrence of artifacts. Higher BMI and subcutaneous fat thickness usually indicate overweight or obesity with a thicker abdominal wall. Artifacts are prone to occur when measuring LS in overweight or obese subjects due to the combined effects of attenuation artifacts, reverberation artifacts, and vessels^[16,30]. Previous studies have also shown that a high BMI is the main reason for measurement failure and unreliable assessment^[17,31,32]. Patients with liver cirrhosis usually have higher LSVs, and they often have artifacts because of their shrunken liver volumes and poor sonic window. Other studies have demonstrated that severe liver fibrosis is a risk factor for unreliable LS measurements^[17,33].

⁷ To the best of our knowledge, this is the first prospective study to analyze artifacts in 2-D SWE of the liver. This study analyzed the predilection sites and people for artifacts, and explored the effects of artifacts on LS measurements. Knowledge of the artifacts is essential to improve operation technology to obtain high-quality images. It is very important to obtain accurate measurements in an attempt to optimize its performance and application value. ² In addition, knowledge from this and other studies on artifacts can be used to investigate how training and education could reduce the occurrence of artifacts. ² Hopefully, engineers and researchers can improve the product design, provide

quality indicators and other ways to avoid the acquisition of improper data due to artifacts.

Our study had several limitations. First, artifacts may be ignored when the color changes are inconspicuous. Second, only one device was tested in this study. Third, this study did not analyze the causes of artifacts, because it is sometimes difficult to accurately determine. Finally, we analyzed only a small sample of data from two operators. Therefore, a larger sample study involving more operators and devices needs to be conducted in future.

CONCLUSION

In conclusion, artifacts are common when using 2-D SWE to measure LS, especially for the novice. Artifacts may lead to the overestimation of LS and reduce the repeatability and reliability of LS measurements. For the elastograms with artifacts, we should place the Q-Box away from the artifacts.

9%

SIMILARITY INDEX

PRIMARY SOURCES

- | | | |
|----------|--|-----------------|
| 1 | Huipeng Wang, Pengchao Zheng, Xuemei Wang, Liang Sang. " Effect of size on liver stiffness measurement by shear wave elastography ", Journal of Clinical Ultrasound, 2021
<small>Crossref</small> | 91 words — 2% |
| <hr/> | | |
| 2 | coek.info
<small>Internet</small> | 59 words — 2% |
| <hr/> | | |
| 3 | www.wjgnet.com
<small>Internet</small> | 47 words — 1% |
| <hr/> | | |
| 4 | www.ics.org
<small>Internet</small> | 42 words — 1% |
| <hr/> | | |
| 5 | Wei Zheng, Zhong-guo Zhou, Chong-hei Wong, Xiao-qing Pei, Shu-lian Zhuang, Qing Li, Min-Shan Chen, An-hua Li, Fu-jun Zhang. "Evaluation of liver parenchyma stiffness in patients with liver tumours: optimal strategy for shear wave elastography", European Radiology, 2018
<small>Crossref</small> | 15 words — < 1% |
| <hr/> | | |
| 6 | repository.supsi.ch
<small>Internet</small> | 12 words — < 1% |
| <hr/> | | |
| 7 | www.frontiersin.org
<small>Internet</small> | 12 words — < 1% |

-
- 8 Jie Zeng, Jian Zheng, Zeping Huang, Shigao Chen, Jing Liu, Tao Wu, Rongqin Zheng, Mingde Lu. "Comparison of 2-D Shear Wave Elastography and Transient Elastography for Assessing Liver Fibrosis in Chronic Hepatitis B", Ultrasound in Medicine & Biology, 2017
Crossref 11 words — < 1%
-
- 9 forums.duelyst.com
Internet 11 words — < 1%
-
- 10 trepo.tuni.fi
Internet 11 words — < 1%
-
- 11 www.maedica.ro
Internet 11 words — < 1%
-
- 12 www.researchgate.net
Internet 11 words — < 1%
-

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES

EXCLUDE MATCHES

< 10 WORDS

< 10 WORDS