

Original Article

Ultrasound evaluation of liver stiffness: accuracy of ultrasound imaging for the prediction of liver cirrhosis as evaluated using a liver stiffness measurement

Momoe Endo^{1)*}, Yoko Soroida^{1,3)*}, Masaya Sato^{1,2)}, Tamaki Kobayashi¹⁾, Hiromi Hikita¹⁾, Mamiko Sato¹⁾, Hiroaki Gotoh¹⁾, Tomomi Iwai¹⁾, Shinji Sone¹⁾, Tetsuo Sasano³⁾, Yuki Sumi³⁾, Kazuhiko Koike²⁾, Yutaka Yatomi¹⁾ and Hitoshi Ikeda¹⁾

1) Department of Clinical Laboratory Medicine and

2) Department of Gastroenterology, Graduate School of Medicine, The University of Tokyo and

3) Department of Biofunctional Informatics, Tokyo Medical and Dental University School of Health Care Sciences

* These authors contributed equally to this work

Background and Aims: Because of the low penetration rate of transient elastography (TE) or its limitations in patients with obesity, narrow intercostal spaces, or ascites, the physical appearance of the liver as visualized using ultrasonography (US) is still thought to provide important information for the prediction of liver fibrosis. We examined the accuracy of various US signs when assessing the presence of liver cirrhosis, compared with TE.

Methods: We enrolled 189 patients who had undergone both conventional US and TE examinations. We then assessed the associations between US parameters of the liver (surface, edge, and parenchymal texture) or the US score (sum of each parameter score), and the presence of liver cirrhosis as determined based on a liver stiffness measurement (LSM) of >15 .

Results: A significant increase in the LSM was observed according to the liver surface score ($P < 0.001$), liver edge score ($P < 0.001$), parenchymal texture score ($P < 0.001$), and US score ($P < 0.001$). The areas under the curves (AUROC) for the prediction of an LSM >15 for the liver surface, liver edge, parenchymal texture, and the US score

were 0.859, 0.768, 0.837, and 0.902, respectively. The AUROC of the US score was higher than that of the APRI score (0.823) or the FIB-4 index (0.804). Using an optimal cut-off value of 3.5, the sensitivity and specificity of the US score were 0.815 and 0.858, respectively.

Conclusions: The US score was clinically useful for the diagnosis of an LSM >15 . The US score can be used as a substitute for TE data in patients with obesity, narrow intercostal spaces, or ascites or in hospitals where TE is unavailable.

Key Words: liver stiffness, US score, ultrasound imaging, liver cirrhosis

Introduction

Hepatic fibrosis occurs as a common response to chronic liver injury and is a pivotal factor influencing the selection of therapeutic options and in determining the prognosis of patients with chronic liver disease^{1, 2}. Although a liver biopsy is considered the gold standard for the diagnosis of fibrosis characterized by morphological damage, this procedure is associated with a significant morbidity or mortality³. Other limitations of liver biopsy are sampling errors and inter- or intraobserver variability. Liver biopsy has also been shown to yield false-negative results in nearly 20%–30% of cases^{4–6}.

Recently, noninvasive liver stiffness measurement (LSM) using transient elastography (TE) has been reported to be well correlated with histologically

Corresponding Author: Masaya Sato, M.D., Ph.D.
Department of Clinical Laboratory, The University of Tokyo Hospital,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.
Tel: +81-3-3815-5411 Fax: +81-3-3814-0021
E-mail: satoma-int@h.u-tokyo.ac.jp
Received May 27 : Accepted September 8, 2017

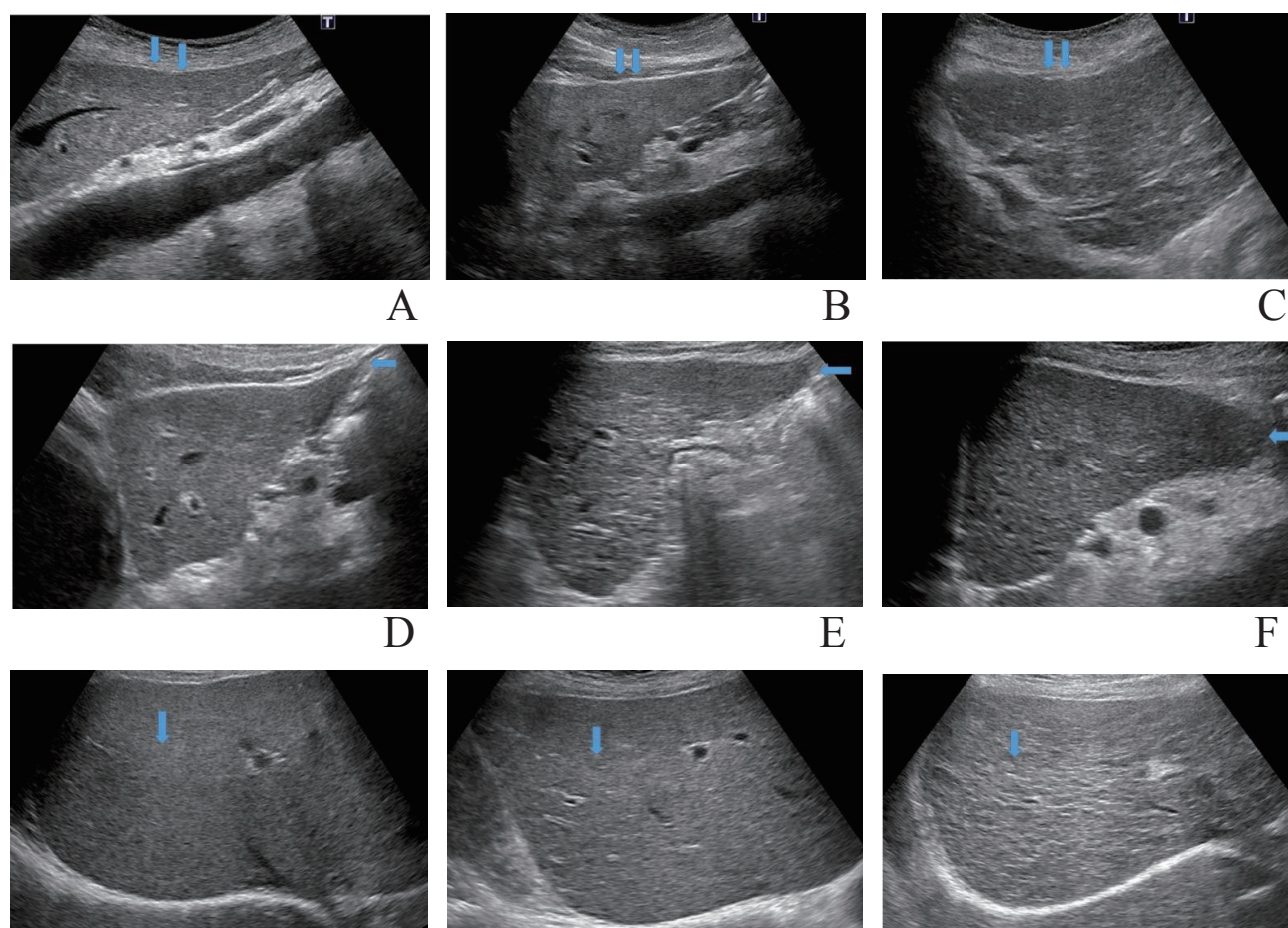


Figure 1. Ultrasound features of the liver.

(A)–(C) Surface: (A) smooth surface, (B) mildly irregular surface, and (C) highly irregular surface. (D)–(F) Edge: (D) sharp edge, (E) mildly blunted edge, and (F) highly blunted edge. (G)–(I) Parenchymal texture: (G) smooth texture, (H) mildly coarse texture, and (I) highly coarse texture.

assessed liver fibrosis stages^{7, 8} and has also been shown to be an accurate predictor of the development of hepatocellular carcinoma (HCC) in patients with chronic liver disease^{9, 10}. Although the excellent performance of TE for predicting the histological stage of liver fibrosis has been well known, TE is not available in every hospital or clinic. In addition, LSM is difficult in patients with obesity or narrow intercostal spaces and impossible in patients with ascites¹¹. Therefore, the physical appearance of the liver as evaluated using ultrasonography (US) is still thought to provide important information for the prediction of liver fibrosis. Thus, we wondered if the US findings could be a substitute for TE.

The aim of the present study was to examine the accuracy of various US signs when assessing the presence of liver cirrhosis, compared with TE.

Patients and Methods

Patients

We enrolled consecutive patients visiting the liver clinic who underwent TE between July 1 and August 31, 2016, at the Department of Clinical Laboratory, the University of Tokyo Hospital. TE was used to assess liver fibrosis, and an LSM > 15 was classified as bridging fibrosis or cirrhosis (F stage 3–4)^{8, 12, 13}. Since LSM was reported to be overestimated in patients with hepatic congestion because of heart disease^{14, 15}, we excluded patients with congestive heart disease from the primary analysis. Also, we excluded patients who failed to undergo an LSM measurement because of obesity, narrow intercostal spaces, or ascites.

The present study was performed in accordance with the ethical guidelines of the Declaration of Helsinki. The study design was included in a comprehensive protocol

for retrospective studies at the University of Tokyo Hospital, Department of Clinical Laboratory Medicine (Tokyo, Japan) and was approved by the University of Tokyo Medical Research Center Ethics Committee (approval number, 3683).

Ultrasound investigation

The US evaluation was performed before TE on the same day by a single specifically trained technician (Y.S). All the ultrasound examinations were performed using a Toshiba Aplio 300 or Aplio 500 with a 3.5–5 MHz transducer. LSM was performed using Fibroscan touch 502 (Echosens, Paris, France)¹⁶. The parameters for the US evaluation were (1) liver surface, scored 0 for smooth (Figure 1A), 1 for mildly irregular (Figure 1B), and 2 for highly irregular (Figure 1C); (2) liver edge, scored 0 for sharp (Figure 1D), 1 for mildly blunted (Figure 1E), and 2 for highly blunted (Figure 1F); and (3) liver parenchymal texture, scored 0 for smooth (Figure 1G), 1 for mildly coarse (Figure 1H), and 2 for highly coarse (Figure 1I). The sum of these three parameters was defined as the US score.

Statistical analysis

Continuous variables were expressed as the medians with the first and third quartiles, while categorical variables were expressed as the frequencies (%) unless otherwise noted. Continuous data were evaluated using the Student *t*-test. We assessed the relationships between each US parameter or the US score and LSM using the Jonckheere-Terpstra trend test. The predictive accuracy of the parameters for the presence of liver cirrhosis was assessed using a receiver operating characteristic (ROC) curve analysis. The area under the curve (AUC) was evaluated as the ability to predict the presence of liver cirrhosis. Comparison of the AUC values was carried out using the Delong test¹⁷. The Delong test was undertaken using R software version 3.4.0 (<http://www.r-project.org>). All other statistical analyses were performed using SPSS software (ver. 19.0; IBM Corp., Armonk, NY, USA). The threshold of *P* values for significance was set at <0.05.

Results

Patient characteristics

Finally, we enrolled 189 patients. Table 1 shows the clinical characteristics of the patients. Twenty-seven patients (14.3%) were diagnosed as having liver cirrhosis (LSM > 15). Approximately 50% were male, and the median age was 62 years. Twenty patients were positive for HBs antigen, and 64 patients were positive for HCV antibody.

US parameters or the US score and LSM

Figure 2 shows an association between each US parameter and the LSM (Figure 2A: liver surface, Figure 2B: liver edge, Figure 2C: liver parenchymal texture). The median LSM in patients with US parameters were 4.8, 7.8, and 16.9 for the liver surface scores; 4.8, 6.3, and 10.2 for the liver edge scores; and 4.7, 7.0, and 17.8 for the parenchymal texture scores, respectively. A Jonckheere-Terpstra trend test showed a significant increasing trend for LSM according to the liver surface score ($P < 0.001$), liver edge score ($P < 0.001$), and parenchymal texture score ($P < 0.001$). We also assessed the association between the US score (sum of each parameter score) and the LSM. We merged patients with a US score of 5 or more into one category, since the number of patients with a US score of 6 was relatively small. As shown in Figure 2D, a significant rising trend with an increment in the US score was observed ($P < 0.001$).

Table 1. Patient Characteristics (n = 189)

Parameter*	Values
Cirrhosis (LSM > 15)	
Present	27 (14.3)
Absent	162 (85.7)
LSM	5.9 (4.4–10.0)
Sex, n (%)	
Female	91 (48.1)
Male	98 (51.9)
Age (years)	62 (51–72)
Virus, n (%)	
HBs-Ag (+)	20 (10.6)
HCV-Ab (+)	64 (33.9)
AST	28 (22–41)
ALT	25 (16–54)
Platelet Count	19.0 (14.1–24.2)
Albumin	4.2 (4.0–4.4)
GGTP	33.5 (19.0–76.8)
Total Bilirubin	0.8 (0.7–1.1)

* Continuous variables are expressed as medians with the first and third quartiles, while categorical variables are expressed as percentages (%).

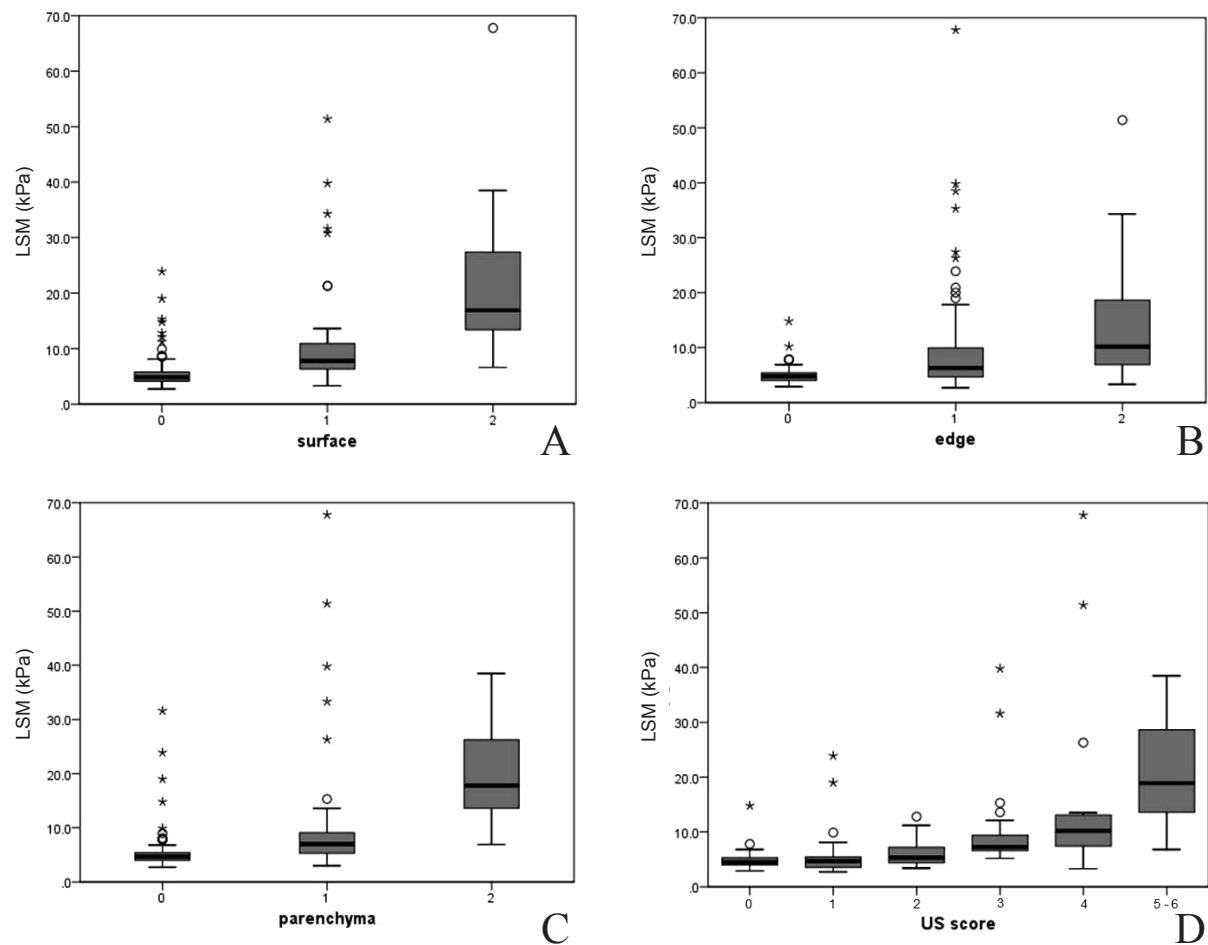


Figure 2. Association between each ultrasound (US) parameter or US score and liver stiffness measurement. (Figure 2A: liver surface, Figure 2B: liver edge, Figure 2C: liver parenchymal texture, Figure 2D: US score)

Table 2. AUC, optimum cut-off level, and the sensitivity and specificity of this cut-off				
	AUC	Optimal cut-off	Sensitivity	Specificity
US score	0.902	3.5	0.815	0.858
Surface	0.859	1.5	0.63	0.944
Edge	0.768	1.5	0.556	0.821
Parenchymal texture	0.837	1.5	0.667	0.944

Receiver -operating characteristic curve analysis of US parameters or the US score for the presence of liver cirrhosis

The AUC, the optimum cut-off level, and the sensitivity or specificity using the optimum cut-off for the liver surface, liver edge, parenchymal texture, and US score are shown in Table 2. The AUCs for the prediction of an LSM > 15 for the liver surface, liver edge, parenchymal texture, and the US score were 0.859, 0.768, 0.837,

and 0.902, respectively. Using the optimal cut-off value of 3.5, the sensitivity and specificity of the US score were 0.815 and 0.858, respectively. We then compared the characteristics of patients with US score less than optimal cut-off value (0-3) in 139 non-cirrhrotic (true negative) patients and 5 cirrhrotic patients (false negative); or patients with US score more than optimal cut-off value (4-6) in 22 cirrhrotic (true positive) patients and 23 non-cirrhrotic patients (true negative). In patients

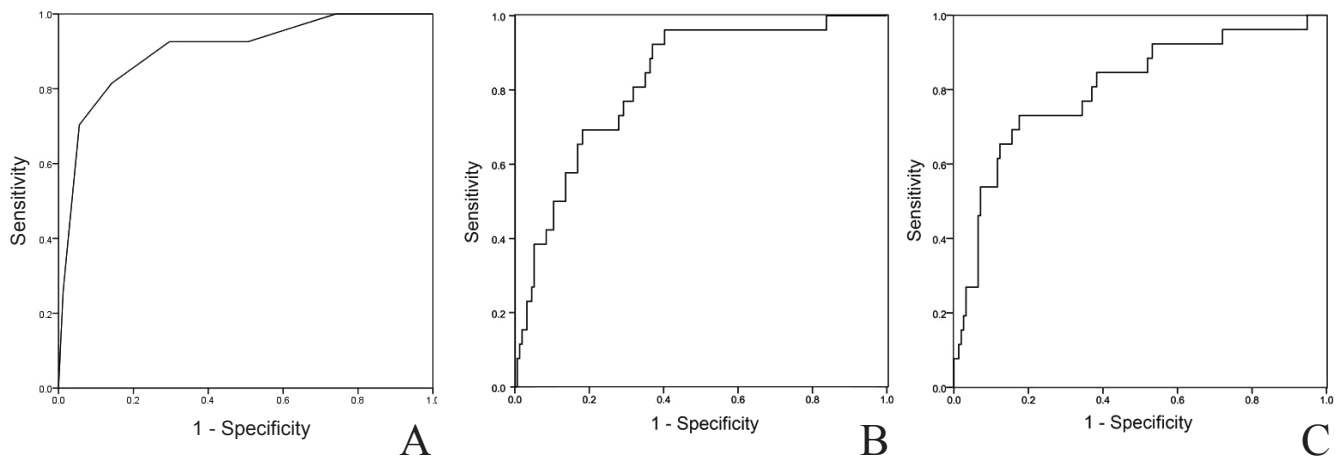
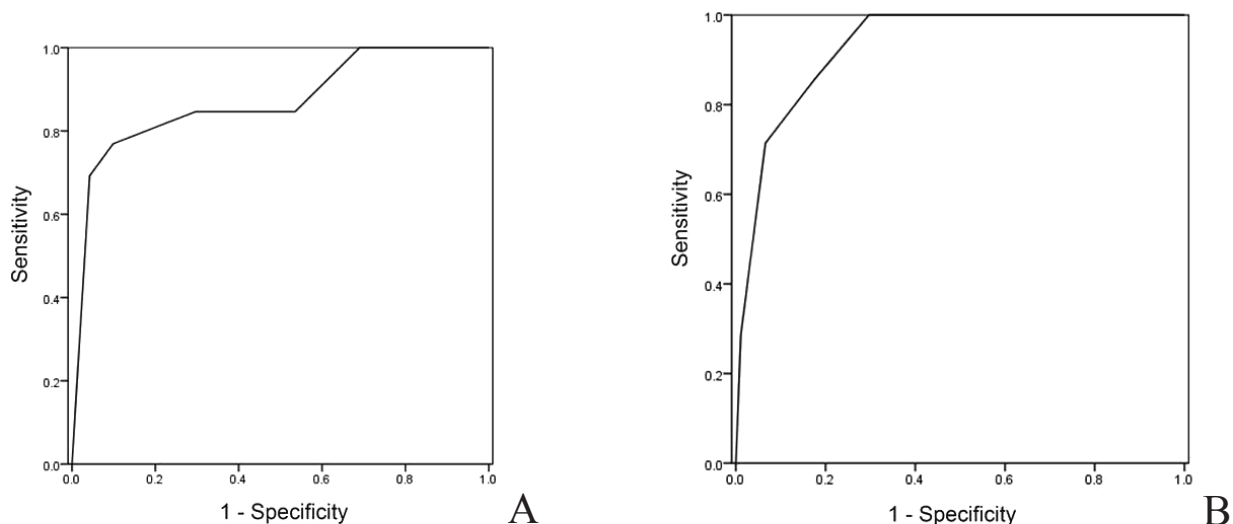


Figure 3. Receiver-operating characteristic curve for the prediction of liver cirrhosis.

(A) US score: the area under the curve was 0.902. (B) APRI: the area under the curve was 0.823. (C) FIB-4 index: the area under the curve was 0.804.



Supplementary Figure 1. Receiver -operating characteristic curve for the prediction of liver cirrhosis of the US score in (A) patients with a viral etiology and (B) patients without a viral etiology.

with US score 0-3, mean liver edge score in correctly diagnosed non-cirrhotic patients was significantly lower than that in cirrhotic patients (0.67 vs. 1.40, $p = 0.01$). In patients with US score 4-6, mean liver surface score and parenchymal texture score in correctly diagnosed cirrhotic patients was significantly higher than that in non-cirrhotic patients (liver surface score: 1.77 vs. 1.39, $p = 0.009$; parenchymal texture score: 1.82 vs. 1.35, $p = 0.001$) (data not shown).

Next, we compared the US score to the aspartate aminotransferase to platelet ratio index (APRI)¹⁸ or the Fibrosis 4 (FIB-4) index¹⁹, which are the most extensively used and well-validated non-invasive markers for the assessment of liver cirrhosis. The ROC curves for the

prediction of an LSM > 15 were plotted for the US score, APRI, and FIB-4 index (Figure 3). The AUCs for the prediction of the presence of an LSM > 15 of the APRI and FIB-4 index were 0.823 and 0.804, respectively. The difference of AUCs for the prediction was significantly different between the US score and FIB-4 index ($p = 0.01$), and quasi-significant between the US score and APRI ($p = 0.08$).

Additionally, we evaluated the AUCs of the US score separately in patients with a viral etiology and those with a non-viral etiology. The AUCs for the prediction of a US score for an LSM > 15 in the viral and non-viral etiology groups were 0.871 and 0.931, respectively (Supplementary Figure 1).

Discussion

The diagnosis of cirrhosis in patients with chronic hepatitis is of therapeutic and prognostic importance. Although the excellent performance of TE for predicting the histological stage of liver fibrosis is well known^{7,8}, the penetration of TE is not sufficient because of problems associated with its cost or a shortage of manpower, and LSM is difficult to perform in patients with obesity or narrow intercostal spaces, and impossible in patients with ascites¹¹. In the present study, we assessed the accuracy of various US signs when assessing the presence of an LSM > 15.

The evaluation of fibrosis stage based on US signs such as liver size, liver morphology, spleen size, diameter of the portal vein, or diameter of the splenic vein has been investigated in previous studies²⁰⁻²⁶. However, these reports were written over ten years ago, and recent advances in US technology have improved the diagnostic accuracy. Also, whether US signs can be substituted for TE findings has not been investigated. Therefore, we conducted the present study to examine the accuracy of US signs when assessing the presence of liver cirrhosis, compared with TE.

In the current study, we examined the diagnostic accuracies of the liver surface, edge, and parenchymal texture for the prediction of cirrhosis (LSM > 15). The liver edge was shown to be less specific for evaluating cirrhosis (AUC 0.768), compared with the liver surface or parenchymal texture. Nishimura et al. assessed the accuracies of these parameters for the prediction of cirrhosis using a correlation coefficient²⁰, and the liver edge was also shown to be less specific. They reported that a mildly blunted (score 1) or a highly blunted edge (score 2) was frequently observed during the early fibrosis stage, and this finding was thought to be a probable cause of the lower correlation coefficient. In line with this study, a large number of patients without cirrhosis were categorized as having a liver edge score of 1 or 2 in the current study (data not shown).

The liver surface, which has been most commonly utilized as the sole indicator for the diagnosis of cirrhosis^{25, 27}, showed the highest predictive accuracy (AUC 0.859) of the three US parameters examined in the present study. In our study, the specificity and sensitivity of the liver surface were 0.944 and 0.630, respectively. A previous study reported by Colli et al. in 2003²³ also showed that the liver surface was highly specific but was not highly sensitive. In this study, the sensitivity of liver surface nodularity for the prediction of severe liver fibrosis was no more than 0.540. Advances in ultrasound

technology may be one of the probable causes of the higher sensitivity of the liver surface in the present study, compared with the previous study.

Little is known about the predictive accuracy of US evaluations of the fibrosis stage based on comparisons with other predictive parameters. We also compared the US score to the APRI¹⁸ and FIB-4 index¹⁹, which are the most extensively used and well-validated non-invasive markers for the assessment of liver cirrhosis. The prediction of the presence of liver cirrhosis (LSM > 15) using the AUC of the US score, which was the sum of three US parameters, was 0.902, which was higher than those of the APRI or FIB-4 index. The AUCs continued to be high in patients with and those without viral etiologies. Mean liver edge score in cirrhotic patients with US score 0-3 (false negative cases) was significantly higher than those of correctly diagnosed non-cirrhotic patients (true negative cases). Also, mean liver surface or parenchymal texture score in non-cirrhotic patients with US score 4-6 (false positive cases) were significantly lower than those in correctly diagnosed cirrhotic patients (true positive cases). This result may indicate that we can potentially separate true negative (positive) cases and false negative (positive) cases by focusing attention on specific US features. By considering this fact, observer may improve predictive accuracy of US score.

Although the current study aimed to examine the accuracy of various US signs in assessing the presence of liver cirrhosis diagnosed by TE, the absence of histological findings should be mentioned as a study limitation. While LSM measured using TE has been reported to be well correlated with the histologically assessed liver fibrosis stage^{7, 8}, LSM was shown to misestimate the fibrosis status in patients with particular conditions, such as hepatic congestion because of heart disease¹⁴. To minimize this limitation, we excluded patients with congestive heart disease from the present study. Another limitation is a lack of objectivity of US score. US is subjective and evaluation accuracy is depend on a skill of observer. Therefore, it requires considerable skill and experience to form a reliable diagnosis. Further study with multiple observers in different learning stages is needed to confirm influence of observer skill on US score. Also, the machine performance of the US device is also related to the predictive accuracy of the US parameters.

In conclusion, a US score composed of three US parameters (liver surface, edge and parenchymal texture) was clinically useful for the diagnosis of an LSM > 15. The US score could be used as a substitute for TE in patients with obesity, narrow intercostal spaces, or

ascites; and in hospitals where TE is not available.

Conflict of interest

None of the authors have any conflicts of interest

References

1. Yang C, Zeisberg M, Mosterman B, Sudhakar A, Yerramalla U, Holthaus K, et al. Liver fibrosis: insights into migration of hepatic stellate cells in response to extracellular matrix and growth factors. *Gastroenterology*. 2003; 124(1):147–59.
2. Botta F, Giannini E, Romagnoli P, Fasoli A, Malfatti F, Chiarbonello B, et al. MELD scoring system is useful for predicting prognosis in patients with liver cirrhosis and is correlated with residual liver function: a European study. *Gut*. 2003; 52(1):134–9.
3. Garcia-Tsao G, Boyer JL. Outpatient liver biopsy: how safe is it? *Annals of internal medicine*. 1993; 118(2):150–3.
4. Pagliaro L, Rinaldi F, Craxi A, Di Piazza S, Filippazzo G, Gatto G, et al. Percutaneous blind biopsy versus laparoscopy with guided biopsy in diagnosis of cirrhosis. A prospective, randomized trial. *Digestive diseases and sciences*. 1983; 28(1):39–43.
5. Nord HJ. Biopsy diagnosis of cirrhosis: blind percutaneous versus guided direct vision techniques—a review. *Gastrointestinal endoscopy*. 1982; 28(2):102–4.
6. Poniachik J, Bernstein DE, Reddy KR, Jeffers LJ, Coelho-Little ME, Civantos F, et al. The role of laparoscopy in the diagnosis of cirrhosis. *Gastrointestinal endoscopy*. 1996; 43(6):568–71.
7. Castera L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology*. 2005; 128(2):343–50.
8. 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(3):403–8.
9. Jung KS, Kim SU, Ahn SH, Park YN, Kim DY, Park JY, et al. Risk assessment of hepatitis B virus-related hepatocellular carcinoma development using liver stiffness measurement (FibroScan). *Hepatology (Baltimore, Md)*. 2011; 53(3):885–94.
10. Masuzaki R, Tateishi R, Yoshida H, Goto E, Sato T, Ohki T, et al. Prospective risk assessment for hepatocellular carcinoma development in patients with chronic hepatitis C by transient elastography. *Hepatology (Baltimore, Md)*. 2009; 49(6):1954–61.
11. 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 (Baltimore, Md)*. 2010; 51(3):828–35.
12. Castera L, Le Bail B, Roudot-Thoraval F, Bernard PH, Foucher J, Merrouche W, et al. Early detection in routine clinical practice of cirrhosis and oesophageal varices in chronic hepatitis C: comparison of transient elastography (FibroScan) with standard laboratory tests and non-invasive scores. *Journal of hepatology*. 2009; 50(1):59–68.
13. Coco B, Oliveri F, Maina AM, Ciccorossi P, Sacco R, Colombatto P, et al. Transient elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. *Journal of viral hepatitis*. 2007; 14(5):360–9.
14. Chon YE, Kim SU, Park JY, Kim DY, Ahn SH, Han KH, et al. Dynamics of the liver stiffness value using transient elastography during the perioperative period in patients with valvular heart disease. *PLoS one*. 2014; 9(3):e92795.
15. Millonig G, Friedrich S, Adolf S, Fonouni H, Golriz M, Mehrabi A, et al. Liver stiffness is directly influenced by central venous pressure. *Journal of hepatology*. 2010; 52(2):206–10.
16. Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, Mal F, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound in medicine & biology*. 2003; 29(12):1705–13.
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988; 44(3):837–45.
18. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology (Baltimore, Md)*. 2003; 38(2):518–26.
19. Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, Dhalluin-Venier V, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology (Baltimore, Md)*. 2007; 46(1):32–6.
20. Nishiura T, Watanabe H, Ito M, Matsuoka Y, Yano K, Daikoku M, et al. Ultrasound evaluation of the fibrosis stage in chronic liver disease by the simultaneous use of low and high frequency probes. *The British journal of radiology*. 2005; 78(927):189–97.
21. Lu LG, Zeng MD, Wan MB, Li CZ, Mao YM, Li JQ, et al. Grading and staging of hepatic fibrosis, and its relationship with noninvasive diagnostic parameters. *World journal of gastroenterology*. 2003; 9(11):2574–8.
22. Hung CH, Lu SN, Wang JH, Lee CM, Chen TM, Tung HD, et al. Correlation between ultrasonographic and pathologic diagnoses of hepatitis B and C virus-related cirrhosis. *Journal of gastroenterology*. 2003; 38(2):153–7.
23. Colli A, Fraquelli M, Andreoletti M, Marino B, Zuccoli E, Conte D. Severe liver fibrosis or cirrhosis: accuracy of US for detection—analysis of 300 cases. *Radiology*. 2003; 227(1):89–94.
24. Shen L, Li JQ, Zeng MD, Lu LG, Fan ST, Bao H. Correlation between ultrasonographic and pathologic diagnosis of liver fibrosis due to chronic virus hepatitis. *World journal of gastroenterology*. 2006; 12(8):1292–5.

25. Gaiani S, Gramantieri L, Venturoli N, Piscaglia F, Siringo S, D'Errico A, et al. What is the criterion for differentiating chronic hepatitis from compensated cirrhosis? A prospective study comparing ultrasonography and percutaneous liver biopsy. *Journal of hepatology*. 1997; 27(6):979-85.
26. Kutcher R, Smith GS, Sen F, Gelman SF, Mitsudo S, Thung SN, et al. Comparison of sonograms and liver histologic findings in patients with chronic hepatitis C virus infection. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine*. 1998; 17(5):321-5.
27. Ferral H, Male R, Cardiel M, Munoz L, Quiroz y Ferrari F. Cirrhosis: diagnosis by liver surface analysis with high-frequency ultrasound. *Gastrointestinal radiology*. 1992; 17(1):74-8.