Systematic Review and Meta-Analysis

Medicine®

OPEN

Feasibility and Efficacy of Transient Elastography using the XL probe to diagnose liver fibrosis and cirrhosis

A meta-analysis

Bingqing Xia, MD[a,b], Fengyan Wang, MD[c], Mireen Friedrich-Rust, MD[d], Fang Zhou, MD[e], Jingyu Zhu, MD[f], Hua Yang, MD[g], Weishan Ruan, MD[a], Zhirong Zeng, MD[b,*]

Abstract

Background: Transient elastography (TE) has been validated as an effective noninvasive tool for the assessment of liver fibrosis. The XL probe is a new probe that was initially designed for use in patients with obesity. A meta-analysis was performed to assess the feasibility and efficacy of TE using the XL probe.

Methods: In September 2016, we systematically searched the PubMed and Science Direct search engines. The feasibility of TE was evaluated based on the failure rate and the results of the unreliable liver stiffness measurement (LSM). The efficacy of TE was measured using sensitivity, specificity, and summary receiver-operating characteristic as measures/indices assessed in different stages of fibrosis. Heterogeneity was measured using the chi-squared test and the Q-statistic. We used the 95% confidence interval (95% CI) as an effect measure.

Results: We included 8 studies in the meta-analysis. When the XL was compared to the M probe, the former showed a lower risk of failure rate [relative risk (RR) 0.24, 95% CI 0.14–0.38]. In patients with a body mass index ≥30 kg/m^2, the XL probe showed a statistically significantly lower risk of failure rate (RR 0.16, 95% CI 0.08–0.32) but no significant improvement (RR 0.76, 95% CI 0.50–1.16) in the unreliable LSM result. In patients showing liver fibrosis stage ≥F2, the XL probe showed a sensitivity of 0.56 (95% CI 0.39–0.72), specificity of 0.71 (95% CI 0.61–0.79), and an area under the curve (AUC) of 0.71. The results observed in patients with liver fibrosis stage F4 were more promising with a sensitivity of 0.84 (95% CI 0.76–0.90), specificity of 0.78 (95% CI 0.70–0.84), and an AUC of 0.88.

Conclusion: TE using the XL probe demonstrates significant diagnostic utility in patients with liver fibrosis and is likely to be more reliable than the M probe in patients with obesity. Large prospective multicenter studies are, however, necessary to establish the new cut-off values to be used for the XL probe in patients with obesity.

Abbreviations: AUC = area under the curve, AUROC = area under the receiver operating characteristic curve, BMI = body mass index, CI = confidence interval, LSM = liver stiffness measurement, RR = relative risk, cROC = summary receiver-operating characteristic, TE = transient elastography, US = ultrasound.

Keywords: liver fibrosis, obese, transient elastography, XL probe

1. Introduction

Liver fibrosis is a final common pathway for all causes of various liver injuries. Progressive liver fibrosis and cirrhosis lead to multiple complications.[1] An accurate estimation of the degree of liver fibrosis is crucial for prognostication and clinical decision making.[2,3] To date, a liver biopsy has been used as a primary diagnostic tool in this context. However, its invasiveness, high costs, variability in interpretation, sampling error, and difficulty with continued observation and monitoring of fibrosis limit the use of a liver biopsy.[4,5]

Because of these limitations, transient elastography (TE) has been introduced as a useful diagnostic aid in clinical settings. This noninvasive tool was developed for staging liver fibrosis by measuring the mechanical shear wave propagation through the hepatic parenchyma. Advantages of TE include simplicity of use, its safety, rapid results, high compliance, and relatively lower costs in both inpatient and outpatient settings. Recently, several studies have evaluated the diagnostic value of TE in a number of liver diseases.[6–9]

Despite its utility in the management of liver disease, TE is associated with a few limitations. We observed that in 5% of the patients studied, TE failed to detect liver fibrosis, whereas 15% showed unreliable results. Furthermore, the presence of obesity acts as a primary limiting factor to its usefulness.[10–12] Therefore, the XL probe was designed as a new Fibroscan probe for use in patients diagnosed with obesity.[13] The M and XL probes differ in the following respects: central ultrasound (US) frequency 3.5 vs 2.5 MHz, respectively; US transducer focal length 35 vs 50 mm,

Editor: Ludovico Abenavoli.

The authors report no conflicts of interest.

[a] Department of Gastroenterology, Zhongshan People's Hospital, Zhongshan, Guangdong, [b] Department of Gastroenterology, The First Affiliated Hospital, Sun Yan-sen University, Guangzhou, [c] Guangdong Institute of Respiratory Disease, Guangdong, China, [d] Department of Internal Medicine, Goethe-University Hospital, Frankfurt, Germany, [e] Department of Radiology, Nanfang Hospital, Southern Medical University, Guangzhou, [f] Department of Anesthesiology, Zhongshan People's Hospital, Zhongshan, [g] School of Stomatology and Medicine, Foshan University, Foshan, Guangdong, China.

[*] Correspondence: Zhirong Zeng, Department of Gastroenterology, The First Affiliated Hospital, Sun Yan-sen University, Guangzhou 510515, China (e-mail: zengzhirong@vip.163.com).

Medicine (2018) 97:39(e11816)

Received: 31 March 2017 / Accepted: 16 July 2018

http://dx.doi.org/10.1097/MD.0000000000011816

respectively; vibration amplitude (peak to peak) 2 vs 3 mm, respectively; external diameter of the tip of the probe 9 vs 12 mm, respectively; and depth of measurement 25 to 65 vs 35 to 75 mm, respectively. Encouraging results were obtained with the XL probe compared to those obtained with the use of the classic probe.[14]

A meta-analysis was performed to assess the overall performance of TE using the XL probe for the diagnosis of liver fibrosis.

2. Materials and methods

All the data were obtained from previously published studies, and no patient consent and ethical approval were required for this study.

2.1. Search strategy and selection criteria

The PubMed and Science Direct search engines were systematically searched in September 2016 without any time restriction. The search terms we used were "controlled attenuation parameter" (All Fields) OR "elasticity imaging techniques" (MeSH Terms) OR "elasticity imaging techniques" (All Fields) OR "elastography" (All Fields) OR "transient elastography" (All Fields) OR "ultrasound elastography "(All Fields) OR "elasticity imaging technique" (All Fields) OR "tissue elasticity imaging" (All Fields) OR "resonance elastography" (All Fields) OR "Fibroscan" (All Fields) AND "obesity" (MeSH Terms) OR "obesity" (All Fields) OR "adiposity" (MeSH Terms) OR "adiposity" (All Fields) OR "corpulence" (All Fields) OR "body weight" (All Fields) OR "overweight" (MeSH Terms) OR "overweight" (All Fields) OR "abdominal diameter" (All Fields) OR "abdominal height" (All Fields) OR "abdominal circumference" (All Fields) OR "abdominal perimeter" (All Fields) OR "girth" (All Fields) OR "paunch" (All Fields) OR "abdominal girth" (All Fields) OR "BMI" (All Fields) OR "body mass index" (All Fields).

Studies chosen were those that evaluated TE using the XL probe and a liver biopsy was used as a reference standard. Only reports in English were included. Exclusion criteria were studies that did not evaluate TE; studies that evaluated other organs and their characteristics such as muscle stiffness/breast lesions/subcutaneous adipose tissue/myocardial elasticity/ovarian tissue/salivary glands; studies wherein the XL probe was not used; those that did not report data on true and false positivity, true and false negativity, sensitivity, or specificity for any stage of fibrosis stage; enrolled the same study cohort with another published article; those that were reviews, correspondence/letters, or editorials that did not report results of their own; animal studies; studies focused on special population groups including organ transplant recipients, pregnant women, and children; and studies aimed at other aspects such as epidemiology and/or mechanism.

2.2. Data extraction procedure

Data extraction was performed by 2 reviewers (BX and FW). We manually scanned references of all included studies to identify additional relevant publications. Disagreements were resolved by discussion and analysis of the data. Authors were contacted for inadequate information and a few responded to our queries. The study selection process has been shown in Figure 1.

The following data were extracted: first author; year of publication; study design; characteristics of the study population (number of patients included, age, and sex); fibrosis stage detected by a liver biopsy; sampling time (interval between TE and the liver biopsy); and the sensitivity, specificity, failure rate, and the unreliable liver stiffness measurement (LSM) of TE. When the study cohorts were too small to calculate the cut-off value, the diagnostic values of 7.1 kPa for $F>2$, 9.5 kPa for $F>3$, and 12.5 kPa for $F=4$, respectively were used as diagnostic cut-off values for TE based on results of previously published cohort studies.[15]

2.3. Statistical methods

Because of the variability in fibrosis staging systems based on variability in the histopathological findings, the overall calcula-

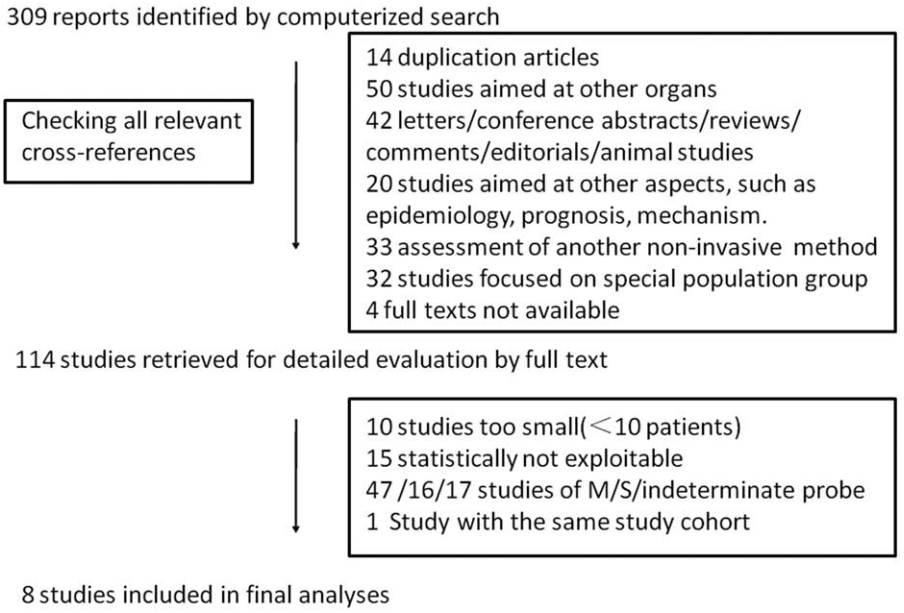

Figure 1. Flow chart of the study selection process for the meta-analysis.

tion was obtained by pooling scoring systems using scores ranging from 0 to 4 for fibrosis staging. Patients were divided into 3 groups: histopathologically documented fibrosis stage ≥2 ($F \geq 2$), histopathologically documented fibrosis stage ≥3 ($F \geq 3$), and those showing definitive cirrhosis ($F = 4$).

The feasibility of TE was evaluated by assessing the failure rate and the unreliable LSM result. Successful LSM examination results were obtained using a total of 10 valid measurements with each probe. Unreliable results were defined as <10 valid measurements or an interquartile range-to-LSM ratio >30%.[16] Intention-to-treat analysis was used for dichotomous data. The odds ratio between the experimental and control groups with 95% confidence intervals (95% CIs) was used as the evaluation index, whereas a 95% CI that does not include was considered statistically significant. The interaction between sensitivity and specificity was assessed using summary receiver-operating characteristic (sROC) analysis. The diagnostic value was analyzed using the Q-statistic and the area under the curve (AUC).

The Mantel-Haenszel method was used with the fixed-effect model to analyze data. Heterogeneity was quantified using the x^2 and I^2 tests. The P value was calculated for pooled estimates, and a random-effects model was chosen for a P value <.05 with the I^2 value ≥50%. This model assumed the occurrence/presence of heterogeneity of methodological or clinical features existing among studies. Therefore, the result from the random-effects model is overall more conservative than the fixed-effects model. Statistical analysis was performed using the statistical software packages Stata v14, and the Review Manager v5.3. Study quality and the risk of bias were assessed using the GRADE system.[17]

3. Results

Based on the inclusion and exclusion criteria, 8 studies involving 1310 patients were included in the meta-analysis.[14,18–24] All patients underwent a liver biopsy as the reference except the study.[19] Five studies reported data regarding failure rates using the XL probe.[14,18,20,21,23] LSM failures occurred in 22 (1.99%) of 1106 cases. Seven studies involving 1721 patients reported data on unreliable LSM results using the XL probe.[14,18,20–24] Unreliable LSM occurred in 232 (13.48%) patients. Seven studies[18–24] showed the efficacy of the XL probe in detection of liver fibrosis stage ≥F3. Six studies[18–20,21,23,24] detected liver fibrosis stage ≥F2 and stage ≥F4.

Patient characteristics and study results varied between studies and are shown in Tables 1 and 2. The study selection process has been shown in Figure 1.

3.1. Failure and unreliable liver stiffness measurement

Four studies reported the failure rates associated with the use of an M and XL probe. Among these 857 patients,[14,18,20,23] LSM failures using the M probe occurred in 87 (10.15%) patients. When the XL was compared to the M probe, the former showed a lower risk of failure rate [relative risk (RR) 0.24, 95% CI 0.14–0.38, $I^2 = 0\%$, $P < .00001$]. In patients with a body mass index (BMI) ≥30 kg/m^2, the XL probe showed a significantly lower risk of failure rate (RR 0.16, 95% CI 0.08–0.32, $I^2 = 0\%$, $P < .00001$). Five studies[14,18,20,22,23] involving 1214 patients reported data regarding unreliable LSM results using the M probe. Unreliable LSM occurred in 253 (20.84%) patients. When the M probe was compared to the XL probe, the latter did not show a significantly lower LSM rate (RR 0.63, 95% CI, 0.30–1.32, $I^2 = 92\%$, $P = .22$). In patients with a BMI ≥30 kg/m^2, the XL probe did not show significant improvement in terms of detection rate (RR 0.76, 95% CI 0.50–1.16, $I^2 = 0\%$, $P = .20$) (Table 3).

3.2. Diagnosis of significance with XL probe

In patients with liver fibrosis stage ≥F2, the XL probe showed a sensitivity of 0.56 (95% CI 0.39–0.72), specificity of 0.71 (95% CI 0.61–0.79), and an AUC of 0.71 (Fig. 2A). In patients with liver fibrosis stage ≥F3, the XL probe showed a sensitivity of 0.66

Table 1

Baseline of characteristics of included studies.

First author	Country	Study design	No. for analysis	Etiology of underlying CLD
Ludmila Gerber, 2015[19]	German	Prospective single center double blind	31	NR
Masato Yoneda, 2015[24]	America	Prospective single center	124	HCV 102 HBV 4 NAFLD 8 AIH 2 PBC 5 PSC 2 Alcohol 1
Christophe Cassinotto, 2013[20]	France	Prospective double blind single center	260	Viral hepatitis 136 alcoholic or NAFLD 113 some other disease 72
Grace Lai-Hung Wong, 2013[22]	China	Prospective, multicenter double blind	149	HCV 14 HBV 7 NAFLD 87 Alcohol 9 Others 30
Mireen Friedrich-Rusta, 2012[18]	German	Prospective single center double blind	57	NAFLD or NASH
Robert P. Myers, 2012[21]	Canada	Prospective multicenter	210	HCV 69 HBV 25 NAFLD 116
Victor de Lédinghen, 2012[14]	France China	Prospective multicenter double blind	286	HCV 62 HBV 31 NAFLD 120 Alcohol 15
Vincent Wai-Sun Wong, 2012[23]	China	Prospective multicenter cohort double blind	193	NAFLD 193

First author	Age, y	Sex (female)	BMI, kg/m^2	Skin-capsular distance, mm
Ludmila Gerber, 2015[19]	Mean ± SD 53 ± 10	11 (35%)	>28 were enrolled Mean ± SD 32.6 ± 4.1 >25 were enrolled mean ± SD (95% CI)	≥25 were enrolled ≥25 were enrolled
Masato Yoneda, 2015[24]	Mean ± SD (95%CI) 57 ± 12 (56–59)	121 (47%)	30.1 ± 4.1 (29.6–30.6) 25–28 90 28–30 79 >30 89	
Christophe Cassinotto, 2013[20]	Mean ± SD 54.4 ± 14	128 (40%)	<25 134 25–29.9 101 >30 86	NR
Grace Lai-Hung Wong, 2013[22]	Mean ± SD 54 ± 12	57 (39%)	>25 were enrolled mean ± SD 29.5 ± 4.1	NR
Mireen Friedrich-Rusta, 2012[18]	Mean ± SD 45 ± 14 median (range) 45 (21–71)	27 (47%)	Mean ± SD 28 ± 5.5 median (range) 27.8 (18–43)	Mean ± SD 25 ± 7 median (range) 24 (15–45)
Robert P. Myers, 2012[21]	Median (IQR) 50 (43–56)	72 (34%)	≥28 were enrolled ≥40 31 median (IQR) 33 (30–36)	Median (IQR) 22 (20–25) >35 (3.9%)
Victor de Lédinghen, 2012[14]	Median ± SD 52.9 ± 12.7	131 (46%)	Median (range) 25.9 (16–51) <25 112 25–30 109 ≥30 65	NR
Vincent Wai-Sun Wong, 2012[23]	Mean ± SD 52 ± 11	83 (43%)	Mean ± SD 28.9 ± 4.8 <25 14 25–30 111 ≥30 68	NR

AIH = autoimmune hepatitis, BMI = body mass index, CI = confidence interval, CLD = chronic liver disease, HBV = hepatitis B virus, HCV = hepatitis C virus, IQR = interquartile range, NAFLD = nonalcoholic fatty liver disease, NASH = nonalcoholic steatohepatitis, NR = not reported, PBC = primary biliary cirrhosis, PSC = primary sclerosing cholangitis.

Table 2

Overview of the study design variables.

First author	Liver biopsy staging system	Fibrosis stage on liver biopsy				
		F=0	*F*=0–1	*F*=2	*F*=3	*F*=4
Ludmila Gerber, 2015[19]	Metavir	31	104	69	44	73
Masato Yoneda, 2015[24]	Metavir (patients with viral hepatitis) Kleiner	NR	148	56	48	34
Christophe Cassinotto, 2013[20]	Metavir (viral hepatitis) Kleiner (NAFLD)	NR	105	44	36	25
Grace Lai-Hung Wong, 2013[22]	Metavir	19	42	32	28	26
Mireen Friedrich-Rusta, 2012[18]	Kleiner	NR	107	29	32	25
Robert P. Myers, 2012[21]	Metavir	10	27	30	16	36
Victor de Lédinghen, 2012[14]	Kleiner	21	20	5	9	2
Vincent Wai-Sun Wong, 2012[23]	Metavir	6	9	4	3	9

First author	Mean or median length of live biopsy, mm	Interval between TE and liver biopsy
Ludmila Gerber, 2015[19]	<10 mm were excluded	1 mo
Masato Yoneda, 2015[24]	<11 mm were excluded	1 wk
Christophe Cassinotto, 2013[20]	Median (IQR) 28 (23–32)	6 mo
Grace Lai-Hung Wong, 2013[22]	<11 mm were excluded mean ± SD 26 ± 7	1 wk
Mireen Friedrich-Rusta, 2012[18]	Mean ± SD 24 ± 6 <15 mm were excluded	1 day
Robert P. Myers, 2012[21]	NR	NR
Victor de Lédinghen, 2012[14]	Median (range)22 (10–60) mean ± SD 22.9 ± 9.5	Median (range) 3.0 (0–17) m<LBREAK"/>–>mean ± SD 4.3 ± 4.0 mo
Vincent Wai-Sun Wong, 2012[23]	<15 mm were excluded	1 mo

IQR = interquartile range, NAFLD = nonalcoholic fatty liver disease, NR = not reported, TE = transient elastography.

Table 3

Failure and unreliable liver stiffness measurement.

		Pooled odds ratio (95% CI)	Heterogeneity I^2%	Effects model
Failure rate	Overall	0.24 (0.14–0.38)	0% *P*<.00001	Fixed effect model
	BMI < 30 kg/m^2	0.20 (0.07–0.58)	0% *P*=.003	Fixed effect model
	BMI> = 30 kg/m^2	0.16 (0.08–0.32)	0% *P*<.00001	Fixed effect model
Unreliable LSM rate	Overall	0.63 (0.30–1.32)	92% *P*=.22	Fixed effect model
	BMI < 30 kg/m^2	1.32 (0.96–1.82)	50% *P*=.09	Fixed effect model
	BMI> = 30 kg/m^2	0.76 (0.50–1.16)	0% *P*=0.20	Fixed effect model

BMI = body mass index, LSM = liver stiffness measurement, NR = not reported.

(95% CI 0.47–0.81), a specificity of 0.82 (95% CI 0.74–0.88), and an AUC of 0.83 (Fig. 2B). In patients with liver fibrosis stage = F4 the results were observed to be more promising—the XL probe showed a sensitivity of 0.84 (95% CI 0.76–0.90), a specificity of 0.78 (95% CI 0.70–0.84), and an AUC of 0.88 (Fig. 2C).

Three studies[20,22,23] involving 573 patients evaluated liver fibrosis using the XL and M probe. The pooled sensitivity and specificity could not be calculated using the Stata v14 software due to the lack of data. A comparison of the sROC drawn using the Review Manager v5.3. software has been shown in Figure 3 (Table 4).

4. Discussion

Usually, progression of chronic liver disease is asymptomatic; however, patients tend to present with complications during advanced stages of the disease. TE, the first developed US-based elastography method, has been validated for assessment of liver fibrosis.[25,26] It is cheap, safe, easy to operate, and is associated with fewer adverse effects. Inability to measure or unreliable LSM results are, however, the primary limitations of TE particularly in those diagnosed with obesity. In 2011, a systematic analysis was performed to include 960 country-years and 9.1 million participants. The study showed that the mean BMI was observed to have increased worldwide by 0.4 kg/m^2 per decade between 1980 and 2008 suggesting that 1.46 billion adults (1.41–1.51 billion) worldwide were noted to have a BMI ≥25 kg/m^2. Among these, 297 million women (280–315 million) and 205 million men (193–217 million) were diagnosed with obesity.[27] Thus, it was important to overcome the drawbacks of TE as a diagnostic tool.

To overcome these limitations, a new probe was introduced in the clinical setting. The XL probe uses a lower frequency, a deeper focal length, a more sensitive ultrasonic transducer, a larger vibration amplitude, and a greater depth of measurements below

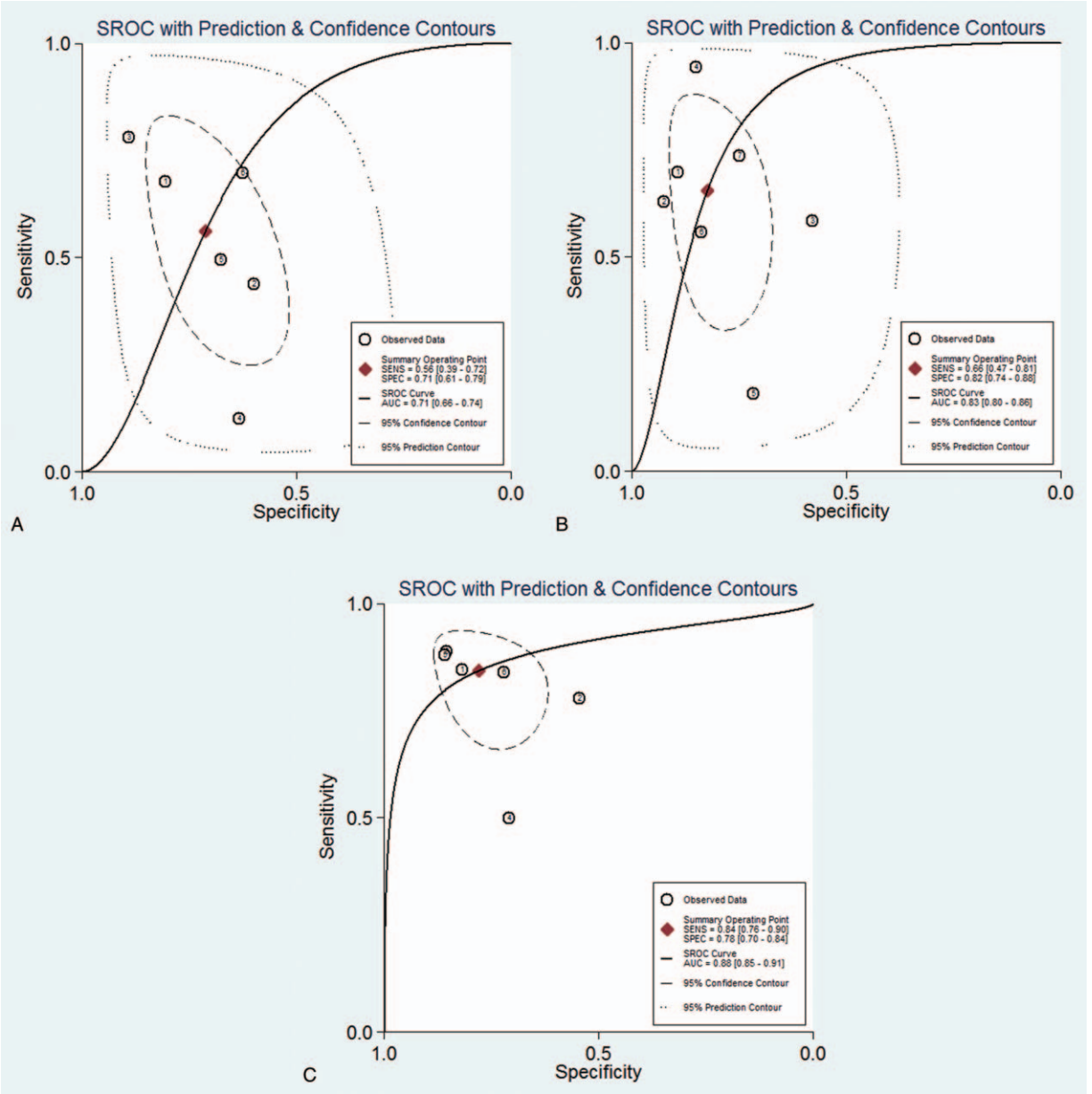

Figure 2. Summary ROC (sROC) plots with 95% CIs and 95% predictive ellipses for transient elastography (TE) with XL probe for the detection of liver fibrosis in the liver fibrosis stage ≥F2 (2a) (fixed effect model n = 613), liver fibrosis stage ≥F3(2b) (random effect model n = 383), liver fibrosis stage = F4(2c) (fixed effect model n = 230). The confidence region consists of the most probable values of true summary sensitivity and specificity and indicates the precision of the summary points. The prediction region predicts the true sensitivity and specificity of a future study, and the size of this region reflects the variation among studies. Individual study estimates are represented by crosses. AUC = area under the curve.

the skin surface, and is designed for the measurement of depths ≥35 mm below the skin, whereas the detection of the M probe begins 25 mm below the skin. This suggests a greater efficacy of the XL probe in patients with obesity, thereby leading to the conclusion that higher the BMI, greater the benefit obtained with the use of the new XL probe. The results of the present study showed a significantly decreased risk of TE failures with the use of the XL probe—the promising results observed with use of this probe in patients with a BMI ≥30 kg/m^2 supported this presumption. The unreliable result rate, however, showed no significant difference between the 2 probes that were assessed, suggesting that even though TE using the XL probe was associated with a higher success rate, the consistency of results obtained with the probe needs improvement, which ought to be the focus of future studies.

The AUC were used to analyze the diagnostic value in the present study. A test showing 100% sensitivity and specificity would have an AUC of 1.0, whereas a test that was equally likely to diagnose a positive result as either positive or negative would have an AUC of 0.5. Therefore, an AUC closer to 1.0 indicates better diagnostic accuracy.[28] The area under the receiver operating characteristic curve (AUROC) for $F \geq 2$ was 0.71 (95% CI 0.66–0.74) with a sensitivity of 0.56 (95% CI 0.39–0.72). The diagnostic value was observed to be acceptable but not satisfactory. Histopathologically proven fibrosis stage ≥2 disease indicated a progression of liver disease in patients who therefore demonstrated an increased risk of

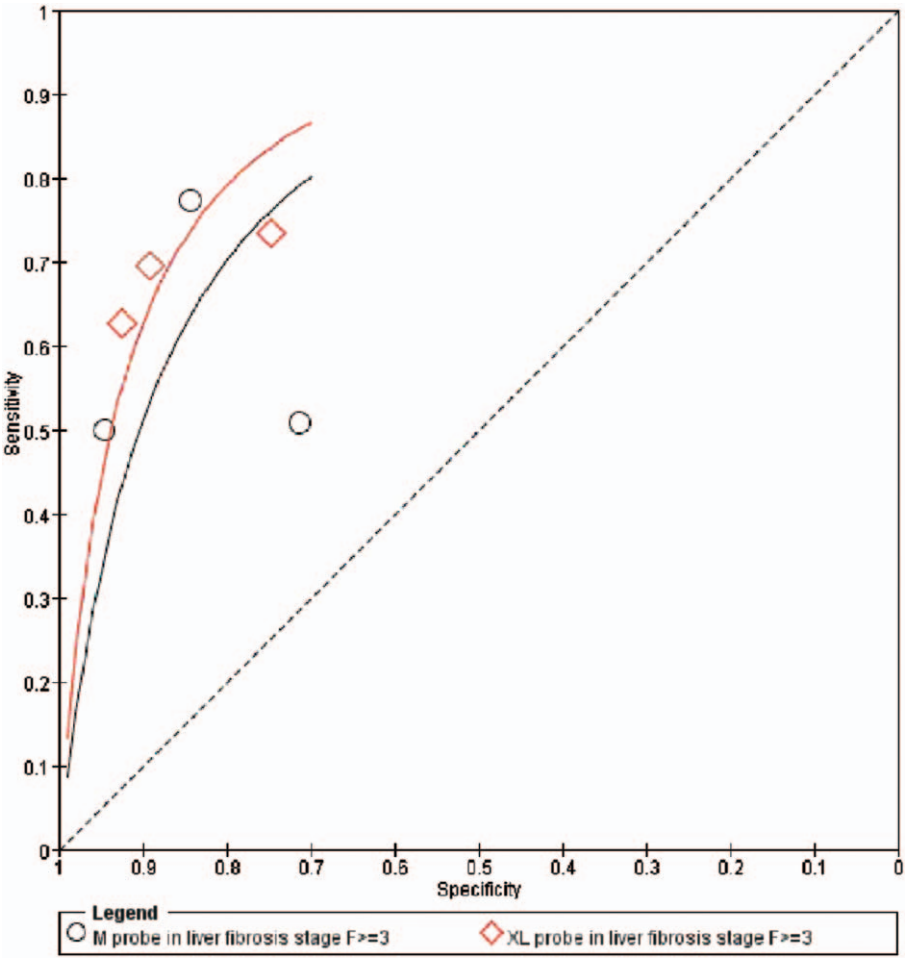

Figure 3. The comparison of summary ROC (sROC) plots with 95% CIs and 95% predictive ellipses for transient elastography (TE) between XL probe and M probe.

developing cirrhosis. Therefore, compared to patients with no or mild fibrosis (F0/1), treatment is strongly indicated in patients with fibrosis stage F2.[29–34] Thus, TE alone (with a relatively low sensitivity) may not be sufficient in clinical practice for the large-scale screening of liver fibrosis. Further studies are, however, necessary to analyze the diagnostic value of TE concomitant with the use of characteristic clinical features and other noninvasive diagnostic methods.

With the progression of liver fibrosis, the AUC of the detection of TE is noted to increase. Our data showed that TE using the XL probe demonstrated considerable/marked diagnostic accuracy, particularly in differentiating cirrhosis vs no cirrhosis with a mean AUROC of 0.88 (95% CI 0.85–0.91) suggesting that in clinical practice, although other clinical features and diagnostic results may be indeterminate, TE may be sufficient to confirm the presence of cirrhosis. Five out of 8 studies excluded patients with a BMI of $<25\,kg/m^2$ suggesting that these studies primarily focused on patients with obesity and that obesity was not a drawback/impediment to the evaluation.

Table 4

Area under the receiver-operating characteristics curves for liver stiffness measurement using XL probe for the diagnosis of significant fibrosis, advanced fibrosis, and cirrhosis.

	Sensitivity	Specificity	AUROC (95% CI)	Heterogeneity I^2%	Effects model
≥F2	0.56 (0.39–0.72)	0.71 (0.61–0.79)	0.71 (0.66–0.74)	54.95% *P*=.058	Fixed effect model
≥F3	0.66 (0.47–0.81)	0.82 (0.74–0.88)	0.83 (0.80–0.86)	83.95% *P*=.001	Random effect model
=F4	0.84 (0.76–0.90)	0.78 (0.70–0.84)	0.88 (0.85–0.91)	0% *P*=.413	Fixed effect model

AUROC = area under the receiver operating characteristic curve, CI = confidence interval.

A few researchers reckon that TE cannot replace a liver biopsy. Compared to TE, a liver biopsy provides additional information regarding the severity of the steatosis and necroinflammatory activity. In a few patients, it might additionally indicate the etiology of the condition. Moreover, in patients with acute hepatitis or an exacerbation of hepatitis, the results of TE may significantly overestimate the stage of liver fibrosis during an alanine aminotransferase flare.[35] The role of a liver biopsy, however, remains controversial. Except for the significant intra- and interobserver variability and sampling errors,[5,36,37] a study performed by Mehta et al[36] demonstrated that the error in liver biopsy results precludes its use as the "criterion standard." Therefore, an alternative "criterion standard" is needed for assessment of liver fibrosis. It is important to consider the possibility that the efficacy/diagnostic performance of TE might be underestimated if a liver biopsy is used as a reference method.

Limitations of our study include no significant heterogeneity was detected in the present study; however, our data were not sufficient to analyze potential coherence factors (length of liver biopsy specimen/time interval between TE and liver biopsy/different thresholds of fibrosis stages noted using TE). In line with a previous study, we observed that the median liver stiffness measured using the M probe was significantly higher than that measured using the XL probe (7.7 vs 7.0 kPa, respectively).[38] In patients with obesity, the nonhepatic tissue between the skin and the liver capsule may interfere with/modify the result obtained using the M probe. In such instances, the XL probe may not use the existing cut-offs defined for the use of the M probe. Two previous studies[18,19] in this context enrolled a very small number of patients to accurately define new cut-offs for the XL probe. There is an urgent need to perform large-scale population-based trials to determine cut-off values for the staging of liver fibrosis. A relatively small number of patients were involved in the subgroup study. Thus, a comparison of diagnostic accuracy between the M and the XL probe is not available. The graph showing a comparison of the sROC in liver fibrosis stage $F \geq 3$ suggested that the XL probe demonstrated a higher diagnostic value than the M probe. Pooled analysis of studies involving only patients with obesity is not available. Further studies and the relevant results are warranted as conclusive evidence in this regard.

5. Conclusion

In summary, TE using the XL probe demonstrated significant diagnostic utility in detecting liver fibrosis and is likely to be a more reliable diagnostic tool than the M probe in patients diagnosed with obesity. Large-scale prospective multicenter studies are necessary to establish new cut-off values for the XL probe in patients diagnosed with obesity.

Acknowledgments

The authors thank Wolters Kluwer Author Services for improving the language of the article.

Author contributions

Bingqing Xia and Fengyan Wang searched the databases, extracted the data, and drafted the manuscript. Fang Zhou, and Hua Yang analyzed the data. Weishan Ruan and Zhirong Zeng conceived and designed the study and edited this article. Mireen Friedrich-Rust generously shared the data from her original study and answered our questions in a patient manner. Bingqing Xia and Fengyan Wang contributed equally to this work.

Data curation: Bingqing Xia, Fengyan Wang.
Formal analysis: Bingqing Xia, Fengyan Wang.
Investigation: Zhirong Zeng.
Methodology: Zhirong Zeng.
Project administration: Weishan Ruan, Zhirong Zeng.
Resources: Mireen Friedrich-Rust, Hua Yang.
Software: Fang Zhou, Hua Yang.
Supervision: Zhirong Zeng.
Validation: Jingyu Zhu.
Writing – original draft: Bingqing Xia.
Writing – review and editing: Bingqing Xia.

References

[1] Schuppan D, Afdhal NH. Liver cirrhosis. Lancet 2008;371:838–51.
[2] Nguyen-Khac E. Results and place of Fibroscan in the non-invasive diagnosis of hepatic fibrosis. Rev Med Interne 2007;28:94–102.
[3] De Franchis R, Dell'Era A. Non-invasive diagnosis of cirrhosis and the natural history of its complications. Best Pract Res Clin Gastroenterol 2007;21:3–18.
[4] Cadranel JF, Rufat P, Degos F. Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AFEF). Hepatology 2000;32:477–81.
[5] Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology 2003;38:1449–57.
[6] Chon YE, Choi EH, Song KJ, et al. Performance of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B: a meta-analysis. PLoS One 2012;7:e44930.
[7] Friedrich-Rust M, Ong M-F, Martens S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology 2008;134:960–74.
[8] De Ledinghen V, Vergniol J. Transient elastography for the diagnosis of liver fibrosis. Expert Rev Med Devices 2010;7:811–23.
[9] Fraquelli M, Rigamonti C, Casazza G, et al. Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. Gut 2007;56:968–73.
[10] Wong GL, Wong VW, Chim AM, et al. Factors associated with unreliable liver stiffness measurement and its failure with transient elastography in the Chinese population. J Gastroenterol Hepatol 2011;26:300–5.
[11] Castera L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology 2010;51:828–35.
[12] Sirli R, Sporea I, Bota S, et al. Factors influencing reliability of liver stiffness measurements using transient elastography (M-probe)-monocentric experience. Eur J Radiol 2013;82:e313–6.
[13] De Ledinghen V, Vegniol J, Foucher J, et al. Feasibility of liver transient elastography with FibroScan using a new probe for obese patients. Liver Int 2010;30:1043–8.
[14] De Ledinghen V, Wong VW, Vergniol J, et al. Diagnosis of liver fibrosis and cirrhosis using liver stiffness measurement: comparison between M and XL probe of FibroScan(R). J Hepatol 2012;56:833–9.
[15] Castera L, Vergniol J, Foucher J, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. Gastroenterology 2005;128:343–50.
[16] Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. J Hepatol 2008;48:835–47.
[17] Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews for Interventions, Version 5.1.0. 2011;The Cochrane Collaboration,
[18] Friedrich-Rust M, Romen D, Vermehren J, et al. Acoustic radiation force impulse-imaging and transient elastography for non-invasive assessment of liver fibrosis and steatosis in NAFLD. Eur J Radiol 2012;81:e325–31.
[19] Gerber L, Kasper D, Fitting D, et al. Assessment of liver fibrosis with 2-D shear wave elastography in comparison to transient elastography and acoustic radiation force impulse imaging in patients with chronic liver disease. Ultrasound Med Biol 2015;41:2350–9.
[20] Cassinotto C, Lapuyade B, Aït-Ali A, et al. Liver fibrosis: noninvasive assessment with acoustic radiation force impulse elastography—comparison

with FibroScan M and XL probes and FibroTest in patients with chronic liver disease. Radiology 2013;269:283–92.
[21] Myers RP, Pomier-Layrargues G, Kirsch R, et al. Discordance in fibrosis staging between liver biopsy and transient elastography using the FibroScan XL probe. J Hepatol 2012;56:564–70.
[22] Wong GL, Vergniol J, Lo P, et al. Non-invasive assessment of liver fibrosis with transient elastography (FibroScan(R)): applying the cut-offs of M probe to XL probe. Ann Hepatol 2013;12:570–80.
[23] Wong VW, Vergniol J, Wong GL, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol 2012;107:1862–71.
[24] Yoneda M, Thomas E, Sclair SN, et al. Supersonic shear imaging and transient elastography with the XL probe accurately detect fibrosis in overweight or obese patients with chronic liver disease. Clin Gastroenterol Hepatol 2015;13:1502.e5–9.e5.
[25] European Association for Study of LiverEASL clinical practice guidelines: management of hepatitis C virus infection. J Hepatol 2014;60:392–420.
[26] European Association for the Study of the LiverEASL clinical practice guidelines: management of chronic hepatitis B virus infection. J Hepatol 2012;57:167–85.
[27] Finucane MM, Stevens GA, Cowan M, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. Lancet 2011;377:557–67.
[28] Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285–93.
[29] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology 2012;55:2005–23.
[30] Lok AS, McMahon BJ. Chronic hepatitis B. Hepatology 2007;45:507–39.
[31] Veldt BJ, Heathcote EJ, Wedemeyer H, et al. Sustained virologic response and clinical outcomes in patients with chronic hepatitis C and advanced fibrosis. Ann Intern Med 2007;147:677–84.
[32] Verlinden W, Francque S, Michielsen P, et al. Successful antiviral treatment of chronic hepatitis C leads to a rapid decline of liver stiffness without an early effect on spleen stiffness. Hepatology 2016;64:1809–10.
[33] Kose S, Tatar B, Gül S, et al. The effect of long-term entecavir therapy on liver histopathology in patients with chronic viral hepatitis B. Acta Clin Belg 2016;71:244–9.
[34] Jung CH, Um SH, Kim TH, et al. Treatment response and long-term outcome of peginterferon alpha and ribavirin therapy in Korean patients with chronic hepatitis C. Gut Liver 2016;10:808–17.
[35] Coco B, Oliveri F, Maina AM, et al. Transient elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. J Viral Hepat 2007;14:360–9.
[36] Mehta SH, Lau B, Afdhal NH, et al. Exceeding the limits of liver histology markers. J Hepatol 2009;50:36–41.
[37] Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol 2002;97:2614–8.
[38] Friedrich-Rust M, Hadji-Hosseini H, Kriener S, et al. Transient elastography with a new probe for obese patients for non-invasive staging of non-alcoholic steatohepatitis. Eur Radiol 2010;20:2390–6.