

Direct comparison of the FibroScan XL and M probes for assessment of liver fibrosis in obese and nonobese patients

Esteban Durango^{1,*}
Christian Dietrich^{1,*}
Helmut Karl Seitz¹
Cornelia Ursula Kunz²
Gilles T Pomier-Layrargues³
Andres Duarte-Rojo⁴
Melanie Beaton⁵
Magdy Elkhashab⁶
Robert P Myers⁷
Sebastian Mueller^{1,3}

¹Department of Medicine and Center for Alcohol Research, Liver Disease and Nutrition, Salem Medical Center,

²Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany; ³Liver Unit, Centre Hospitalier de l'Université de Montréal, Hôpital Saint-Luc, Montréal, Quebec, ⁴Toronto Western Hospital Liver Centre, Toronto, Ontario; ⁵Multi-Organ Transplant Unit, University of Western Ontario, London, Ontario; ⁶The Toronto Liver Centre, Toronto, Ontario; ⁷Liver Unit, Division of Gastroenterology and Hepatology, Department of Medicine, University of Calgary, Calgary, Alberta, Canada

*These authors contributed equally to this research

Correspondence: Sebastian Mueller
Department of Internal Medicine,
Salem Medical Center, University
of Heidelberg, Zeppelinstrasse 11–33,
69121 Heidelberg, Germany
Tel +62 2148 3210
Fax +62 2148 4494
Email sebastian.mueller@urz.uni-heidelberg.de

Background: A novel Fibroscan XL probe has recently been introduced and validated for obese patients, and has a diagnostic accuracy comparable with that of the standard M probe. The aim of this study was to analyze and understand the differences between these two probes in nonobese patients, to identify underlying causes for these differences, and to develop a practical algorithm to translate results for the XL probe to those for the M probe.

Methods and results: Both probes were directly compared first in copolymer phantoms of varying stiffness (4.8, 11, and 40 kPa) and then in 371 obese and nonobese patients (body mass index, range 17.2–72.4) from German (n = 129) and Canadian (n = 242) centers. Liver stiffness values for both probes correlated better in phantoms than in patients (r = 0.98 versus 0.82, $P < 0.001$). Significantly more patients could be measured successfully using the XL probe than the M probe (98.4% versus 85.2%, respectively, $P < 0.001$) while the M probe produced a smaller interquartile range (21% versus 32%). Failure of the M probe to measure liver stiffness was not only observed in patients with a high body mass index and long skin-liver capsule distance but also in some nonobese patients (n = 10) due to quenching of the signal from subcutaneous fat tissue. In contrast with the phantoms, the XL probe consistently produced approximately 20% lower liver stiffness values in humans compared with the M probe. A long skin-liver capsule distance and a high degree of steatosis were responsible for this discordance. Adjustment of cutoff values for the XL probe (<5.5, 5.5–7, 7–10, and >10 kPa for F0, F1–2, F3, and F4 fibrosis, respectively) significantly improved agreement between the two probes from r = 0.655 to 0.679.

Conclusion: Liver stiffness can be measured in significantly more obese and nonobese patients using the XL probe than the M probe. However, the XL probe is less accurate and adjusted cutoff values are required.

Keywords: cirrhosis, liver fibrosis, liver stiffness, obesity, steatosis, transient elastography, M probe, XL probe

Introduction

Assessment of liver fibrosis is essential for prediction of prognosis and complications, to decide on treatment strategies, and to monitor disease progression and response to treatment. The introduction of transient elastography (FibroScan) to measure liver stiffness¹ has dramatically improved the noninvasive diagnosis of fibrosis.^{2–6} Liver stiffness is obtained rapidly within 10 minutes, with high reproducibility in more than 95% of patients.^{7–9} Cutoff values have been established that enable allocation to histologic stages of fibrosis with high accuracy. Thus, a cutoff value of 12.5 kPa has been widely accepted for discrimination between liver fibrosis (F1–3) and cirrhosis (F4),

with liver stiffness values below 6 kPa considered to be normal.^{5,6} However, correct disease-specific classification of fibrosis is still under discussion. This is mainly due to the fact that various conditions other than fibrosis are able to increase liver stiffness significantly. The most important disorders are inflammation or liver damage,^{10–12} congestion,¹³ and cholestasis,¹⁴ along with variables such as food intake,¹⁵ mastocytosis,¹⁶ pregnancy,¹⁷ and amyloidosis.^{18,19} Assessment of liver fibrosis in a more complex setting, ie, with coexisting inflammation, is presently a matter of extensive debate. While some authors recommend use of disease-specific cutoff values,²⁰ we and others have proposed algorithms that consider these individual factors or sequential measurement of liver stiffness after potential interventions, such as alcohol detoxification or treatment with diuretics in patients with heart failure.^{5,13,14,21}

One limitation of transient elastography in assessing liver stiffness is the fact that a significant proportion of patients have either measurement failure (about 5%) or so-called unreliable measurements, mainly due to obesity (about 15%).^{20,22} Thus, liver stiffness measurements are increasingly limited in a large proportion of the general population given that obesity is a major health problem, as evidenced by an increase in the mean body mass index (BMI) by 0.4 kg/m² per decade between 1980 and 2008.²³ To cope with these challenges, a novel probe, known as the XL probe, has recently been developed for use in obese patients. This probe differs from the conventional M probe by using a more sensitive ultrasound transducer, a lower frequency, a larger vibration amplitude, a deeper focal length, and a greater depth of measurement.²⁴ The XL probe allows measurement of liver stiffness in significantly more patients, including those who are obese,^{24–27} and comparable diagnostic accuracy has been found for both probes in patients with histologic fibrosis stages F3 and F4.²⁸ So far, the XL probe has been studied exclusively in obese patients, and at least two reports have demonstrated consistently lower liver stiffness values for the XL probe as compared with the M probe.^{24,27}

In this study, we directly compared both probes in liver phantoms and in a large cohort of patients from various centers in an attempt to elucidate these discrepancies further in nonobese patients. We did not attempt to obtain liver histology for nonobese patients without liver disease in this study because this issue has already been addressed elsewhere.^{24–27} Our major aims were to confirm the discrepancies between the two probes in various settings, to identify the underlying causes for these differences, and to define adjusted cutoff

values for the novel XL probe. Because both probes were used at the same ultrasound session within 15 minutes of each other, all the aforementioned conditions modifying liver stiffness were kept as constant as possible. Our findings indicate that the XL probe consistently produces lower liver stiffness values that require adaptation of the cutoff values for interchangeable interpretation.

Materials and methods

Patients

In total, 371 patients (144 women and 227 men) of mean age 50.8 ± 13.1 years from Germany and Canada were consecutively enrolled from 2009 to 2012 (Table 1 and Figure 1A). The Canadian cohort consisted of 242 patients with successful liver stiffness measurement using both probes.²⁶ This cohort included adults aged >18 years with chronic liver disease of any etiology and a BMI > 28 kg/m² who had undergone percutaneous liver biopsy. The German cohort included 129 patients from the Salem Medical Center at the University of Heidelberg. The same inclusion and exclusion criteria were applied, except that a larger proportion of nonobese patients (67.4%) with a BMI ranging from 15.8 to 72.4 kg/m² were included. In addition, the German cohort contained more patients with cirrhosis and alcoholic liver disease (see Table 1). There were no significant differences between the two cohorts with respect to gender, median liver stiffness, and waist circumference. Liver stiffness was measured using both the M probe and the XL probe, and anthropomorphic parameters, such as hip and waist circumference, and ultrasound parameters were obtained at the same time. Morphologic and biological data (eg, skin-liver capsule distance) were assessed before each examination using an ultrasound scanner. The study was approved by the local ethics committees. Feasibility analysis was done for all 371 patients, but correlation between liver stiffness values obtained by the M probe and the XL probe as well as analysis of discordant measurements were performed in 165 patients (see Figure 1A) who had reliable liver stiffness values using both probes (see criteria outlined below) and liver stiffness values <75 kPa (the detection limit of the Fibroscan device).

Measurement of liver stiffness

Liver stiffness was measured by transient elastography, which has already been described in detail elsewhere.¹ The tip of the probe transducer was placed on the skin between the ribs and the level of the right lobe of the liver. A minimum of ten measurements was obtained, and the results are

Table 1 Patient's characteristics

Characteristics	German cohort (n = 129)		Canadian cohort (n = 242)	
	Median or proportion (%)	Range	Median or proportion (%)	Range
Demographics				
Male	72 (55.8%)		117 (48.3%)	
Age years	53	20–85	50	18–74
Anthropometrics				
Weight (kg)	80.0	42.0–174.0	93.7	56.0–149.0
Size (m)	1.72	1.46–1.93	1.70	1.38–2.00
BMI (kg/m ²)	26.9	15.8–72.4	32.0	28.0–52.0
<25	40.3		0.0	
25–29.9	27.1		15.3	
30–40	19.4		76.0	
>40	13.2		8.7	
Liver disease etiology				
Viral	0.0		42.0	
ALD	36.4			
NALD	22.5		46.0	
Other	39.5		12.0	
Waist (cm)	102.0	72–165	108	86–150
Hip (cm)	100.0	82.5–170	104	83–133
Waist/hip ratio	0.98	0.80–1.35	1.04	0.88–1.28
Ultrasound parameters				
Skin-capsula distance without pressure (mm)	19.8	10.9–55		
Skin-capsula distance with pressure (mm)	18.5	9.8–52.6	21.9	10.0–34.6
Liver size middle axillar line (cm)	14.0	8.0–26.0		
Spleen size (cm) transient elastography (SM probe)	10.0	8.0–18.3		
Liver stiffness (kPa)	6.0	2.1–75.0	8.4	2.7–75.0
IQR	0.9	0.0–30.5	1.4	0.0–26.9
Success rate (%)	90.0	10.0–100.0	76.9	0.0–100.0
Transient elastography (XL probe)				
Liver stiffness (kPa)	5.7	2.1–75.0	6.9	2.7–75.0
IQR	1.4	0.0–39.8	1.2	0.0–24.3
Success rate (%)	100.0	10.0–100.0	100.0	45.5–100.0
Blood tests				
AST (IU/L)	75	14–348	43	30–61
ALT (IU/L)	70	5–771	55	36–87
GGT (IU/L)	268	7–1886	51	30–104
AP (IU/L)	116	33–970	81	65–101
Bilirubin (mg/dL)	1.5	0.1–24.7	0.6	0.5–0.9
Quick (%)	100.7	12–120		–
INR	1.02	0.76–2.50	1.00	1.00–1.10
Platelets (×10 ⁹ /L)	217	64–402	207	161–313

Note: $P < 0.005$ indicates a significant difference between the two cohorts.

Abbreviations: BMI, body mass index; IQR, interquartile range; ALD, alcoholic liver disease; NALD, non-alcoholic liver disease; AST, asparagine aminotransferase; ALT, alanine aminotransferase; AP, alkaline phosphatase; GGT, gamma glutamyltransferase; IQR, interquartile range; INR, international normalized ratio.

expressed as median liver stiffness in kPa and the interquartile range. Characteristics of the M and XL probes, respectively, are as follows: central ultrasound frequency 5 mHz versus 3.5 mHz, ultrasound transducer focal length 35 mm versus 50 mm, tip of the external diameter of the probe 9 mm versus 12 mm, vibration amplitude (peak to peak) 2 mm versus 3 mm, and measurement depth 25–65 mm versus

35–75 mm. Examinations were performed by trained and experienced operators. The maximum number of attempts was set at 20 under difficult conditions (eg, severe obesity). Liver stiffness values were divided into three categories, as described recently, ie, failure, no liver stiffness obtained; unreliable liver stiffness values, success rate <60% and/or interquartile range for the M probe >30%; reliable liver

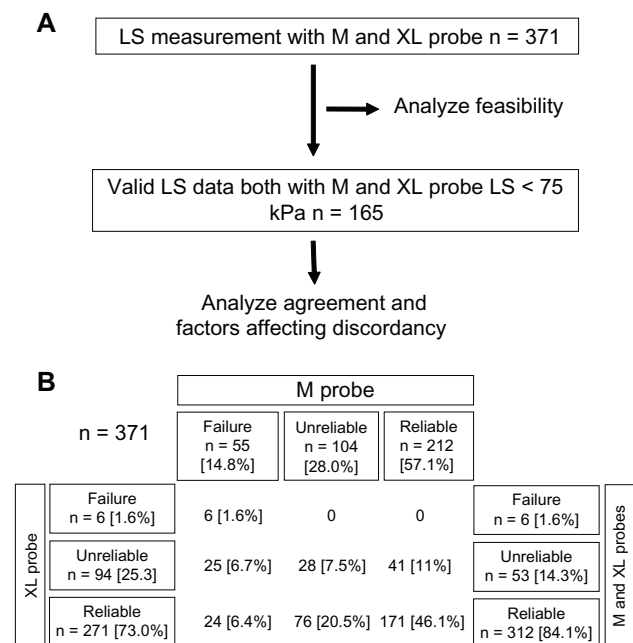


Figure 1 Study design (A) and overall results (B) of comparison between the M and XL probes.

Abbreviation: LS, liver stiffness; n, number.

stiffness values, success rate >60% and interquartile range for the M probe <30%.¹⁶

Degree of steatosis by ultrasound

Every patient in the German cohort underwent regular routine abdominal ultrasound prior to measurement of liver stiffness, and the degree of steatosis was estimated semi-quantitatively based on comparison with the kidney cortex using the following four stages: 0 = no steatosis, 1 = mild steatosis, 2 = pronounced steatosis, and 3 = fatty liver. The mean steatosis grade was 1.1 ± 1.0.

Phantom model

To compare the physical properties of the M and XL probes directly, phantoms mimicking liver tissue were created using a mixture of styrene-ethylene/butylstyrene copolymer and mineral oil, as recently described.²⁹ The phantoms had three different degrees of stiffness (4.8, 11, and 40 kPa) covering a broad range of fibrosis stages (F0, F3, and F4, respectively). A specifically programmed transient elastography mode allowed us to perform these measurements without applying any pressure on the surface of the phantom.

Statistical analysis

Receiver operating curve analyses were performed to define the criteria predicting failure of liver stiffness measurement

with each probe. The cutoff values to predict measurement failure, such as BMI, were calculated from these receiver operating curves using the maximum product of sensitivity and specificity. Correlations between liver stiffness and biological parameters were done using Spearman’s correlation coefficient (regression coefficient *r*, *r*², *P*). All *P* values were two-sided. Differences were considered to be statistically significant at *P* < 0.05. Agreement between measurements obtained using the different probes was assessed by Bland–Altman plot and Passing–Bablok regression analysis.^{30,31} Consistency of agreement between the probes was calculated by dividing the number of patients with an identical fibrosis classification for both probes by the total number of patients. Statistical analysis was performed using PASW Statistics version 18.0.0 (SPSS Inc, Munich, Germany).

Results

Feasibility and contributing factors

Liver stiffness was not measurable in six (1.6%) of the 371 patients using the XL probe and in 55 patients (14.8%) using the M probe, as shown in Figure 1B. None of the six patients with measurement failure using the XL probe could be measured with the M probe. The XL probe also yielded more reliable measurements than the M probe (74% versus 67%). Despite these advantages of the XL probe, additional use of the M probe increased the overall number of reliable tests from 73% to 84%. In addition, the M probe produced a smaller interquartile range (21% versus 32%), which is an advantage in longitudinal follow-up (data not shown). Using receiver operating curve analysis, we then studied the factors that determined measurement failure by transient elastography in more detail (Table 2). Skin-liver capsule distance, hip circumference, waist circumference, and BMI were the best indicators of whether a reliable liver stiffness value could be obtained. Thus, a BMI > 33 kg/m², a waist circumference >114 cm, a skin-liver capsule distance >20 mm, and a body weight >89 kg were used as cutoff values for predicting failure of liver stiffness measurement using the M probe. However, it should be noted that ten patients were not measurable with the M probe despite being nonobese; failure of liver stiffness measurement in these patients was mainly due to existence of a specific type of subcutaneous fat tissue in which propagation of ultrasound waves is poor. Of the six patients not measurable with the XL probe, four had a BMI > 30 kg/m² and the remaining two patients were nonobese. Of note, the maximum BMI that was measurable with the XL probe was 71 kg/m². Taken together, the XL probe enabled more successful and

Table 2 Parameters affecting the ability of the M probe to measure liver stiffness

Variable of test	Area	Standard error	Asymptotic significance	Asymptotic 95% confidence interval	
				Lower limit	Upper limit
Skin-capsula distance without pressure (mm)	0.925	0.026	0.000	0.874	0.977
Skin-capsula distance with pressure (mm)	0.907	0.034	0.000	0.840	0.973
Waist (cm)	0.823	0.062	0.000	0.702	0.944
BMI (kg/m ²)	0.805	0.074	0.000	0.660	0.950
Waist/hip ratio	0.801	0.062	0.000	0.679	0.924
Hip (cm)	0.769	0.078	0.001	0.615	0.922
Weight (kg)	0.729	0.074	0.004	0.585	0.874
Ascites (mm)	0.569	0.084	0.379	0.405	0.734
Size (m)	0.368	0.084	0.095	0.203	0.533

Note: Area under the receiver operating characteristic curve were determined for various morphometric factors that may determine the measurability of liver stiffness by the M probe.

Abbreviation: BMI, body mass index.

reliable measurements independent of BMI status. However, use of both probes increased overall performance.

Liver stiffness values between the two probes were highly correlated in liver phantoms and patients

First, we compared the XL and M probes directly in copolymer phantoms representing three stages of fibrosis, ie, 4.8, 11, and 40 kPa. An excellent correlation was found between the two probes ($r = 0.98$, $P < 0.001$). We next performed correlation analysis in the 165 patients with reliable liver stiffness results for both probes. Figure 2A shows a scatter plot of liver stiffness obtained by the M and XL probes and Passing–Bablok regression. Liver stiffness values were highly correlated, but the correlation was significantly smaller as compared with the phantom data ($r = 0.82$ versus 0.95 , $P < 0.001$). Passing–Bablok regression showed a slope of 0.763 (95% confidence interval 0.718 – 0.808) with an intercept of 0.696 (95% confidence interval 0.334 – 1.042). Thus, the XL probe consistently yielded lower liver stiffness values of around 80%. Separate regression analysis for the German and Canadian cohorts showed no significant difference with regard to the slope, despite the differences with regard to etiology and BMI (not shown). The Bland–Altman plot based on the percent differences (Figure 2B) confirmed that the M probe resulted in higher liver stiffness values, with a mean of 13.1%. This discrepancy between the two probes was noted over the entire liver stiffness range, and Passing–Bablok regression could not identify significant differences between low and high liver stiffness values. In summary, comparison of the M and XL probes show excellent correlation, but a consistently smaller liver stiffness of 80% with the XL probe. The XL

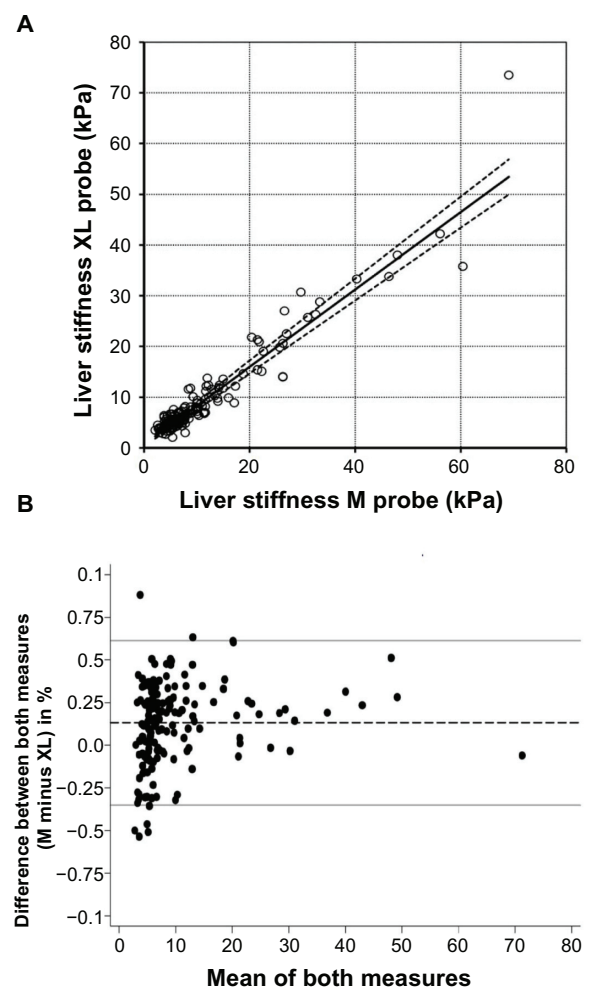


Figure 2 Passing–Bablok regression (A) and Bland–Altman (B) plots comparing the M and XL probes. (A) Significant lower liver stiffness values of around 80% are obtained using the XL probe over the whole measuring range. (B) The consistently lower liver stiffness values obtained using the XL probe were confirmed using the Bland–Altman plot, which shows the percent difference between liver stiffness values for both probes against the mean of both measures. The signal-to noise ratio improves at higher liver stiffness values, but no significant differences are observed between low and high liver stiffness.

probe yields LS values which are consistently lower at about 80% of the values measured with the M probe.

Factors determining inconsistency between the XL and M probes

Using Spearman rank correlation, we first determined the factors associated with the percent difference in values obtained between the XL and M probes. As shown in Table 3A, this percent difference was significantly correlated with the skin-liver capsule distance, BMI, body weight, and waist and hip circumference. Interestingly, the percent difference correlated only with liver stiffness obtained by the M probe but not by the XL probe (Table 3B). No differences were observed between the two probes, when an artificial liver phantom was used, but the differences are observed in real patients. Indeed, only liver stiffness for the M probe correlated well with the difference in skin-liver capsule distance (Table 3B).

Table 3 Morphometric parameters that correlate with (A) the discrepancy between the M and XL probes and (B) liver stiffness as measured by the M and XL probes

Parameter	Percentage difference M versus XL probe		
	Spearman Rho	P	n
(A)			
LS (M probe)	0.388**	2.86E-12	302
Skin-capsula distance (mm)	0.274**	1.20E-06	304
BMI (kg/m ²)	0.268**	2.05E-06	304
Weight (kg)	0.267**	2.34E-06	304
Waist (cm)	0.185**	1.50E-03	292
Steatosis degree ultrasound (0.1.2.3)	0.352**	1.56E-03	78
Hip (cm)	0.173**	2.50E-03	304
Size (cm)	0.126*	2.83E-02	304
LS (XL probe)	-0.116*	4.36E-02	302
Parameter	LS (M probe)	LS (XL probe)	
(B)			
Percentage difference M versus XL probe	0.388**	-0.116*	
BMI (kg/m ²)	0.320**	0.185**	
Hip (cm)	0.278**	0.202**	
Steatosis degree ultrasound (0.1.2.3)	0.505**	0.308**	
Skin-capsula distance (mm)	0.254**		
Weight (kg)	0.251**	0.133*	
Waist (cm)	0.228**	0.135*	
Waist/hip ratio	0.155**	0.149*	
Age	0.125*	0.133*	

Notes: (A) Liver stiffness using the M probe and the liver-skin capsule distance correlate significantly with this discrepancy. *P < 0.05, **P < 0.01. (B) Only liver stiffness measured using the M probe correlated with liver-skin capsule distance. The degree of steatosis as determined semiquantitatively by ultrasound was identified as an independent additional factor. *P < 0.05, **P < 0.01.

Abbreviations: LS, liver stiffness; BMI, body mass index.

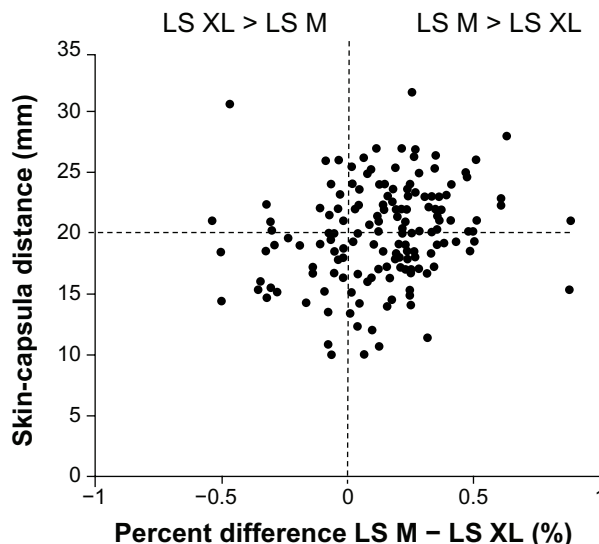


Figure 3 Scatter plot of percent difference of liver stiffness values between the M and XL probes as a function of skin-liver capsule distance. Because higher liver stiffness values are obtained with the M probe, more dots are found in the right panel. The M probe overestimates liver stiffness, especially in obese patients with a difference in skin-liver capsule distance of more than 20 mm (left upper panel).

Abbreviation: LS, liver stiffness.

Figure 3 also shows the percent difference between the two probes as a function of the difference in skin-liver capsule distance, indicating that the M probe overestimates liver stiffness, namely at differences in skin-liver capsule distance greater than 20 mm (see right upper panel). Interestingly, in the German cohort, the degree of steatosis, as estimated semiquantitatively by ultrasound (stages 0–3), also correlated well with the percent discrepancy between the two probes.

We next studied elastographs in detail for patients whose measurements differed at least by 30% between the M and XL probes. Almost four times as many patients were overestimated by the M probe as compared with the XL probe (37 [22.4%] versus 11 [6.6%]). Moreover, mean liver stiffness in the 37 patients with higher values using the M probe than the XL probe differed markedly (13.1 kPa [M probe] versus 8.3 kPa [XL probe]). For 23 cases (62%) in this subgroup, the fibrosis classification changed by one stage, and by two stages in one case. A typical pattern could be identified in the elastographic images of these misclassified patients, ie, in all cases, the M probe included subcutaneous fat in the region of interest measured and/or yielded a scattered, angulated, or broad shear wave, probably due to the effects of diffraction. As shown in a representative elastograph in Figure 4A, the regression algorithm for the Fibroscan device when used for calculation of the resulting liver stiffness values clearly preferred the left rim of the shear wave, with the higher velocity eventually leading to overestimation of liver stiffness. In contrast, the XL probe

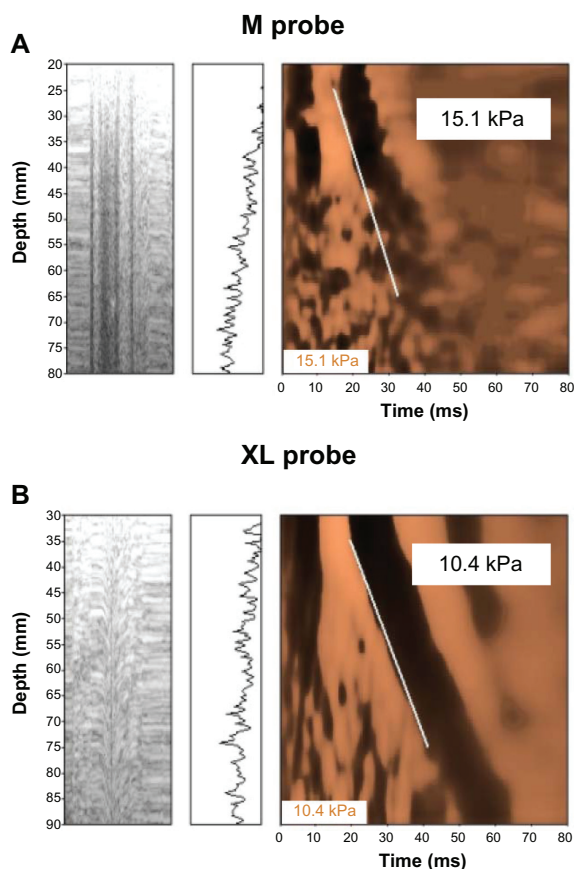


Figure 4 Representative elastograph obtained by the (A) M probe and (B) XL probe for a patient whose liver stiffness measurements differed markedly between the two probes. (A) The M probe yields a scattered broad shear wave, probably because of the effects of diffraction. The regression algorithm for the Fibrosan device clearly preferred the left rim of the shear wave with the higher velocity, eventually leading to an overestimation of liver stiffness. (B) In contrast, in all patients, the XL probe yielded a well defined shear wave and correctly calculated liver stiffness.

yielded a well defined shear wave (Figure 4B) and correctly calculated liver stiffness in all of these patients. In summary, these findings strongly suggest that the M probe consistently overestimates liver stiffness, especially in patients with a large skin-liver capsule distance and in those with pronounced steatosis. The consistently overestimated liver stiffness using the M probe is mainly due to overestimated assessment of shear wave velocity and not a distortion of the shear wave itself.

Adapted cutoff values for the XL probe improve agreement between the two probes

We next analyzed how the consistent discrepancy between the probes affected the classification of fibrosis and whether revised cutoff values for the XL probe could increase the concordance of fibrosis classification between the two probes. Table 4A shows the classification according to F0–4 fibrosis

stages using the conventionally established cutoff values for the M probe (F0, <6 kPa; F1–2, 6–8 kPa; F3, 8–12.5 kPa; F4, >2.5 kPa).⁴ Using these cutoff values, there was agreement between the two probes in 108 patients and disagreement in 57 patients, and by two fibrosis stages in one case. This resulted in a quite good overall agreement of $r = 0.665$ between the two probes with regard to classification of fibrosis (Table 4C, lane 1). We then adjusted the cutoff values for the XL probe using a slope factor of 0.76 as identified in Figure 2A by Passing–Bablok regression analysis. Correction of liver stiffness and XL values yielded a diagonal scatter plot without consistent discrepancy (data not shown). Using these novel cutoff values for classification of fibrosis (Table 4C, lane 2), the number of patients showing agreement increased dramatically from 39 to 50 for advanced fibrosis (stages F3–4). By consistently changing the cutoff values and calculating the agreement in a reiterating procedure (see Table 4C), we eventually identified improved cutoff values for the XL probe (5.5, 5.5–7, 7–10, and >10 kPa for F0, F1–2, F3, and F4 fibrosis stages, respectively). These cutoff values were rounded up for practical reasons. The revised cutoff values improved the overall agreement from 0.655 to 0.679 and agreement for the advanced fibrosis stages from 0.769 to 0.805. The lower panel of Table 4B shows the allocation table for these cutoff values. Taken together, the adapted cutoff values for the XL probe increased the overall agreement between the two probes with regard to classification of fibrosis.

Discussion

In this study, we compared the ability of the novel XL probe and the conventional M probe to measure liver stiffness in a large multicenter cohort of 371 patients. In contrast with previous studies,^{24–28} we included a larger proportion of nonobese patients and liver phantoms, rather than obtaining liver histology. In general, the more powerful XL probe results in significantly fewer measurement failures and more reliable liver stiffness measurements, even in nonobese patients. In our study, the XL probe consistently produced liver stiffness values which were about 80% lower in patients, whereas both probes measured identical liver stiffness values in liver phantoms. Previous^{24,26,27} but not all²⁵ studies have identified a similar relationship. Subanalysis of underlying factors for this discrepancy and direct elastographic analysis showed that the M probe consistently overestimated shear wave velocity under conditions of pronounced steatosis and a skin-liver capsule distance. Finally, we demonstrated in this large and heterogeneous study population that adapted cutoff values improves agreement in classification of fibrosis between the two probes.

Table 4 Classification of fibrosis stage by liver stiffness values using (A) conventional (M probe) and (B) optimized cutoff values (XL probe), and (C) agreement between the two probes for the different cutoff values

(A)					(B)			
M probe	XL probe				XL probe			
LS (kPa)	<6	6–8	8–12.5	>12.5	<5.5	5.5–7	>7–10	>10
<6	61	53	8		51	9	1	
6–8	40	23	16	1	18	15	7	
8–12.5	30	1	13	15	1	6	16	8
>12.5	34			10	24		4	30
	165	77	37	26	25	69	30	28
								38

No	Cut-off values for XL probe (kPa)				Agreement			Discrepancy (>two fibrosis stages)		
	F0	F1–F2	F3	F4	Overall	No fibrosis	No advanced fibrosis	Advanced fibrosis	F0–4	F4
(C)										
1	<6	6–8	8–12.5	>12.5	0.65	0.62	0.87	0.77	1	0
2	<4.8	4.8–6.4	6.4–10	>10	0.66	0.57	0.82	0.77	4	0
3	<6	6–8	8–10	>10	0.60	0.62	0.87	0.77	1	0
4	<6	6–7	7–10	>10	0.66	0.62	0.87	0.81	2	0
5	<5.5	5.5–7	7–10	>10	0.68	0.65	0.87	0.81	1	0

Notes: Cutoff values used for classification of fibrosis stages F0, F1–2, F3, and F4 are (A) <6, 6–8, 8–12.5, and >12.5 kPa and (B) <5.5, 5.5–7, 7–10, and >10 kPa. No further cutoff value was introduced for fibrosis stages F1 and F2, given that resolution in this low range of liver stiffness was very low because of other confounders. (C) Agreement was calculated as the ratio of agreeing liver stiffness values divided by the sum of all liver stiffness values. The left columns describe the chosen cutoff values for fibrosis stages F0–4 derived from Table 4B. The revised optimized cutoff values (lane 5) significantly improve agreement for both low and advanced fibrosis stages. The data were calculated from 165 patients with reliable liver stiffness measurements using both probes.

Abbreviation: LS, liver stiffness.

Notably, these cutoff values are independent of the underlying etiology of liver disease.

There are several potential reasons for the discrepancy in measurements obtained by the M and XL probes. First, there could be a simple technical difference, in that the two probes differ markedly with respect to their energy and ultrasound frequency. However, in our experiments using copolymer phantoms, the M and XL probes yielded identical results over a wide range of stiffness values. Second, liver stiffness could be higher in the subcapsular region of the liver due to the increased matrix proximal to the liver capsule. However, consequent deviations of shear waves in the subcapsular region are typically not observed, and no such liver stiffness changes close to the liver surface have been observed using three-dimensional elastographic approaches, such as magnetic resonance elastography.³² Another reason for overestimation of liver stiffness using the M probe could be the simple fact that the measuring depth starts too early and the measuring range is too short. This has recently been reported by Myers et al²⁶ through direct comparison with histologic fibrosis scores. Indeed, in our study, it can be observed in most elastographs that the initial shear wave is much faster when penetrating the skin-muscle region, causing diffraction artifacts when

entering the liver. When we tried to identify factors that correlated significantly with the discrepancy between the two probes, two major observations were made. First, the percent difference between the two probes depended only on the liver stiffness obtained using the M probe but not using the XL probe. Second, only the liver stiffness obtained using the M probe correlated with skin-liver capsule distance, and overestimation became more likely at a skin-liver capsule distance higher than 20 mm. These findings, in association with the absolute agreement of the two probes in the phantom studies, strongly suggest that the shorter measuring depth of the M probe is critical and affects the liver stiffness value. Of note, the M probe failed to calculate the shear wave velocity correctly, and consequently overestimated liver stiffness in all elastographs from patients who showed a discrepancy greater than 30% between the M and XL probes.

Nevertheless, the skin-liver capsule distance does not explain completely the discrepancy between the M and XL probes. The M probe significantly overestimated liver stiffness when only considering patients with normal BMI < 25 kg/m² and those with a skin-liver capsule distance <25 mm, which is in contrast with earlier findings.²⁶ What else other than skin-liver capsule distance could

explain the persistent discrepancy between measurements obtained using these two probes in nonobese patients? As demonstrated in Table 3B, and to some extent surprisingly, the degree of steatosis obtained by ultrasound also contributed substantially to the discrepancy. This could mean that fatty liver itself could cause scattering artifacts of the shear wave and eventually cause overestimation by the less energetic M probe.

We eventually identified novel cutoff values for the XL probe that enabled consistently improved classification of fibrosis stage and better agreement for all fibrosis stages. Although our data suggest that the XL probe more closely represents true stiffness in the human liver, we opted to adapt the XL cutoff values because those for the M probe are well established and widely used. In addition, we believe that very large histology-based study cohorts would be required to confirm these differences because of the high sampling error with liver biopsies.

What are the implications of our results for transient elastography in daily practice? First of all, in contrast with what is recommended by the manufacturer and proposed elsewhere,²⁵ the two types of probe could be used interchangeably as long as revised cutoff values are used. The XL probe could be used tentatively first-line in obese patients. According to the results of our study, it could also be used successfully in nonobese patients, although the interquartile ranges are generally smaller using the M probe in such patients. Importantly, we strongly recommend use of the same probe once a patient has been measured successfully with one probe in the past, to keep the longitudinal sample error to a minimum. We also suggest that patients should be measured with the alternative probe in cases of measurement failure and also in the event of unreliable measurements. Data from this study indicate that approximately 30% of patients may benefit from this approach, ultimately achieving reliable liver stiffness values. Although the XL probe improves the success rate and reliability of liver stiffness measurement, the availability of both probes improves the overall performance further. Finally, the M probe is still preferred for longitudinal follow-up of individual patients because it consistently produces smaller interquartile ranges.

This study, which was performed in a large multicenter cohort of obese and nonobese patients with variable etiology of liver fibrosis demonstrates that the XL probe is more successful and yields more reliable but consistently lower liver stiffness values. The reasons for this consistent discrepancy are related to the shorter measuring range and penetration depth of the M probe in obese patients and also to the degree of hepatic steatosis present. Given that the cutoff values for

the M probe are well established, the optimized cutoff values for the XL probe presented here should improve agreement between the two probes.

Acknowledgments

This study was supported by the Dietmar Hopp Foundation and the Manfred Lautenschläger Foundation. RPM is supported by salary awards from the Alberta Heritage Foundation for Medical Research (now Alberta Innovates-Health Solutions) and the Canadian Institute for Health Research.

Disclosure

The authors report no potential conflicts in this work.

References

- Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Biol*. 2003;29:1705–1713.
- 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;128:343–350.
- Ganne-Carrie N, Ziol M, de Ledinghen V, et al. Accuracy of liver stiffness measurement for the diagnosis of cirrhosis in patients with chronic liver diseases. *Hepatology*. 2006;44:1511–1517.
- Friedrich-Rust M, Ong MF, Martens S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. *Gastroenterology*. 2008;134:960–974.
- Mueller S, Sandrin L. Liver stiffness: a novel parameter for the diagnosis of liver disease. *Hepat Med*. 2010;2:49–67.
- Castera L, Pinzani M. Biopsy and non-invasive methods for the diagnosis of liver fibrosis: does it take two to tango? *Gut*. 2010;59:861–866.
- Vignier N, Esmat G, Sharkawy AE, et al. Reproducibility of liver stiffness measurements in hepatitis C virus (HCV)-infected patients in Egypt. *J Viral Hepat*. 2011;18:e358–e365.
- Boursier J, Konate A, Gorea G, et al. Reproducibility of liver stiffness measurement by ultrasonographic elastometry. *Clin Gastroenterol Hepatol*. 2008;6:1263–1269.
- Foucher J, Castera L, Bernard PH, et al. Prevalence and factors associated with failure of liver stiffness measurement using FibroScan in a prospective study of 2114 examinations. *Eur J Gastroenterol Hepatol*. 2006;18:411–412.
- Sagir A, Erhardt A, Schmitt M, Haussinger D. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. *Hepatology*. 2008;47:592–595.
- Arena U, Vizzutti F, Corti G, et al. Acute viral hepatitis increases liver stiffness values measured by transient elastography. *Hepatology*. 2008;47:380–384.
- Dechene A, Sowa JP, Gieseler RK, et al. Acute liver failure is associated with elevated liver stiffness and hepatic stellate cell activation. *Hepatology*. 2010;52:1008–1016.
- Millonig G, Friedrich S, Adolf S, et al. Liver stiffness is directly influenced by central venous pressure. *J Hepatol*. 2010;52:206–210.
- Millonig G, Reimann FM, Friedrich S, et al. Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. *Hepatology*. 2008;48:1718–1723.
- Mederacke I, Wursthorn K, Kirschner J, et al. Food intake increases liver stiffness in patients with chronic or resolved hepatitis C virus infection. *Liver Int*. 2009;29:1500–1506.
- Kohlhaas A, Durango E, Millonig G, et al. Transient elastography with the XL probe rapidly identifies patients with non-hepatic ascites. *Hepat Med*. 2012;4:11–18.

17. Kohlhaas A, Millonig G, Schutz F, Seitz HK, Mueller S. Liver stiffness during pregnancy. *J Hepatol.* 2011;54:S134–S135.
18. Lanzi A, Gianstefani A, Mirarchi MG, Pini P, Conti F, Bolondi L. Liver AL amyloidosis as a possible cause of high liver stiffness values. *Eur J Gastroenterol Hepatol.* 2010;22:895–897.
19. Bastard C, Bosisio MR, Chabert M, et al. Transient micro-elastography: a novel non-invasive approach to measure liver stiffness in mice. *World J Gastroenterol.* 2011;17:968–975.
20. 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;26:300–305.
21. Mueller S, Millonig G, Sarovska L, et al. Increased liver stiffness in alcoholic liver disease: differentiating fibrosis from steatohepatitis. *World J Gastroenterol.* 2010;16:966–972.
22. Castera L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology.* 2010;51:828–835.
23. Finucane MM, Stevens GA, Cowan MJ, 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;377:557–567.
24. de Ledinghen V, Vergniol J, Foucher J, El-Hajbi F, Merrouche W, Rigalleau V. Feasibility of liver transient elastography with FibroScan using a new probe for obese patients. *Liver Int.* 2010;30:1043–1048.
25. 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;56:833–839.
26. Myers RP, Pomier-Layrargues G, Kirsch R, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. *Hepatology.* 2012;55:199–208.
27. 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;20:2390–2396.
28. 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;56:564–570.
29. Oudry J, Bastard C, Miette V, Willinger R, Sandrin L. Copolymer-in-oil phantom materials for elastography. *Ultrasound Med Biol.* 2009;35:1185–1197.
30. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem.* 1983;21:709–720.
31. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307–310.
32. Lee VS, Miller FH, Omary RA, et al. Magnetic resonance elastography and biomarkers to assess fibrosis from recurrent hepatitis C in liver transplant recipients. *Transplantation.* 2011;92:581–586.

Hepatic Medicine: Evidence and Research

Publish your work in this journal

Hepatic Medicine: Evidence and Research is an international, peer-reviewed, open access journal covering all aspects of adult and pediatric hepatology in the clinic and laboratory including the following topics: Pathology, pathophysiology of hepatic disease; Investigation and treatment of hepatic disease; Pharmacology of drugs used for

Submit your manuscript here: <http://www.dovepress.com/hepatic-medicine-evidence-and-research-journal>

the treatment of hepatic disease. Issues of patient safety and quality of care will also be considered. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress