Pitfalls of Liver Stiffness Measurement: A 5-Year Prospective Study of 13,369 Examinations

Laurent Castéra,[1,2] Juliette Foucher,[1,2] Pierre-Henri Bernard,[2] Françoise Carvalho,[1] Daniele Allaix,[1] Wassil Merrouche,[1] Patrice Couzigou,[1] and Victor de Lédinghen[1]

Liver stiffness measurement (LSM) based on transient elastography (TE, FibroScan) is gaining in popularity for noninvasive assessment of liver fibrosis. However, LSM has limitations, which have not yet been thoroughly evaluated. We prospectively investigated the frequency and determinants of LSM failure and unreliable results over a 5-year period, based on 13,369 examinations (134,239 shots). LSM failure was defined as zero valid shots, and unreliable examinations were defined as fewer than 10 valid shots, an interquartile range (IQR)/LSM greater than 30%, or a success rate less than 60%. LSM failure occurred in 3.1% of all examinations (4% at first examination [n = 7261]) and was independently associated at first examination with body mass index (BMI) greater than 30 kg/m² (odds ratio [OR], 7.5; 95% confidence interval [CI], 5.6-10.2; *P* = 0.0001), operator experience fewer than 500 examinations (OR 2.5 [1.6-4.0]; *P* = 0.0001); age greater than 52 years (OR 2.3 [1.6-3.2]; *P* = 0.0001), and type 2 diabetes (OR 1.6 [1.1-2.2]; *P* = 0.009). Unreliable results were obtained in a further 15.8% of cases (17% at first examination) and were independently associated at first examination with BMI greater than 30 kg/m² (OR 3.3 [2.8-4.0]; *P* = 0.0001), operator experience fewer than 500 examinations (OR 3.1 [2.4-3.9]; *P* = 0.0001), age greater than 52 years (OR 1.8 [1.6-2.1]; *P* = 0.0001), female sex (OR 1.4 [1.2-1.6], *P* = 0.0001), hypertension (OR 1.3 [1.1-1.5]; *P* = 0.003), and type 2 diabetes (OR 1.2 [1.0-1.5]; *P* = 0.05). When metabolic syndrome and waist circumference were taken into account in a subgroup of 2835 patients, waist circumference was the most important determinant of LSM failure and unreliable results. *Conclusion:* In our experience, liver stiffness measurements are uninterpretable in nearly one in five cases. The principal reasons are obesity, particularly increased waist circumference, and limited operator experience. These results emphasize the need for adequate operator training and for technological improvements in specific patient subpopulations. (HEPATOLOGY 2010;51:828-835.)

Transient elastography (TE) (FibroScan; Echosens, Paris, France) is a novel noninvasive method for measuring liver stiffness.[1] Values range from 2.5 to 75 kPa,[2] and values of approximately 5.5 kPa were recently reported to define normality.[3] Liver stiffness measured by means of TE correlates with hepatic fibrosis stages, both in chronic hepatitis C[4,5] and in other chronic liver diseases.[6-14] The diagnostic accuracy of liver stiffness measurement (LSM) is excellent for cirrhosis,[15-17] and, among noninvasive methods, TE appears to be the most accurate for early detection of cirrhosis.[18] Because TE is a user-friendly technique that can be performed rapidly (less than 5 minutes) at the bedside, with immediate results and high patient acceptance, it is likely to become an important tool in clinical practice in the near future.[19-22]

However, TE has limitations: LSM results may be influenced by acute liver injury (as reflected by ALT flares), with a risk of overestimating liver stiffness,[6,23,24] and also by extrahepatic cholestasis.[25] In addition, LSM is difficult in patients who are obese or who have narrow intercostal spaces, and impossible in patients with ascites (in whom the diagnosis of cirrhosis is obvious). The TE interpreta-

Abbreviations: ALT, alanine aminotransferase; BMI, body mass index; CI, confidence interval; GGT, gamma-glutamyl transpeptidase; IQR, interquartile range; LSM, liver stiffness measurement; NAFLD, nonalcoholic fatty liver disease; OR, odds ratio; SR, success rate; TE, transient elastography; WC, waist circumference.

From the Services d'Hépato-gastroentérologie, Centre Hospitalo-Universitaire (CHU) de Bordeaux, [1]Hôpital Haut Lévêque, 33604 Pessac; and [2]Hôpital Saint André, Bordeaux, France.

Received May 26, 2009; accepted October 27, 2009.

Address reprint requests to: Laurent Castera, M.D., Ph.D., Service d'Hépato-Gastroentérologie, CHU de Bordeaux, Hôpital Haut Lévêque, Avenue Magellan, 33604 Pessac, France. E-mail: laurent.castera@chu-bordeaux.fr; fax: 00 33 5 57 65 64 45.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/hep.23425

Potential conflict of interest: Dr. Castéra is on the speaker's bureau of Echosens.

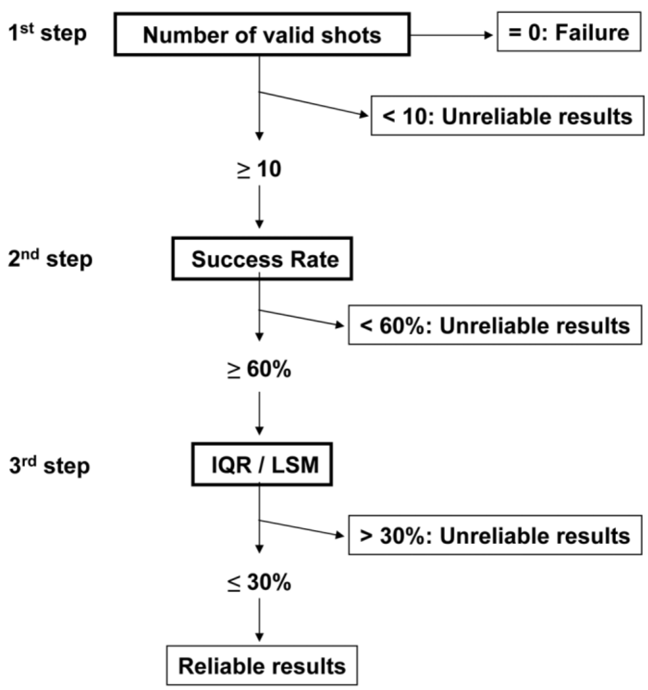

Fig. 1. Manufacturer's criteria for LSM interpretation.

tion software indicates whether each measurement (or "shot") is successful. When a shot is considered unsuccessful, the machine provides no result. The entire LSM procedure is considered to have failed when no value is obtained after 10 shots or more. The manufacturer recommends that successful measurements be validated by using the following three criteria: (1) number of valid shots at least 10; (2) success rate (SR: the ratio of valid shots to the total number of shots) at least 60%; and (3) interquartile range (IQR, reflecting the variability of measurements) less than 30% of the median LSM value (IQR/LSM ≤30%) (Fig. 1).

Here we prospectively investigated the frequency and determinants of LSM failure and unreliable results in clinical practice over a 5-year period, based on 13,369 examinations (134,239 shots) and multivariate analysis.

Patients and Methods

Study Population

Between May 2003 and May 2008, all TE examinations performed in adult patients with suspected chronic liver disease were included. The following clinical and biological parameters were recorded at the time of LSM: age, sex, body mass index (BMI), waist circumference (WC), hypertension (defined as ongoing antihypertensive pharmacological treatment), type 2 diabetes (defined as fasting serum glucose >5.6 mmol/L or ongoing antidiabetic treatment [oral antidiabetic drug or insulin]), hyperlipidemia (defined as triglycerides levels >1.7 mmol/L or ongoing treatment with low-density lipoprotein–lowering drugs), indication for TE, operator experience in TE, and serum aspartate aminotransferase, alanine aminotransferase (ALT), and gamma-glutamyl transpeptidase (GGT) levels.

Because triglycerides and WCs were not available for all patients, the effect of metabolic syndrome was studied specifically in a subgroup of patients (n = 2835 who underwent 6221 examinations) in whom these data were available. Metabolic syndrome was defined according to the criteria proposed for the European population by the International Diabetes Foundation[26]: central obesity (WC ≥80 cm in women or ≥94 cm in men), plus any two of the following conditions: hyperlipidemia, hypertension, and type 2 diabetes.

The study protocol conformed to the ethical guidelines of the 1975 Helsinski declaration, and patients were enrolled after giving their written informed consent.

Liver Stiffness Measurement and Definition of Failure and Unreliable Results.

Measurements were made on the right lobe of the liver, through the intercostal spaces, with the patient lying in dorsal decubitus with the right arm in maximal abduction. The tip of the transducer probe was covered with coupling gel and placed on the skin, between two ribs at the level of the right lobe. The operator, assisted by a time-motion ultrasound image, located a liver portion at least 6 cm thick and free of large vascular structures. The operator then pressed the probe button to begin the measurements ("shots"). The patients were examined by a total of seven operators during the study period, each of whom performed more than 100 examinations.

LSM failure was recorded when no value was obtained after at least 10 shots (valid shots = 0). The results were considered unreliable in the following circumstances: valid shots fewer than 10, SR less than 60%, or IQR/LSM greater than 30% (Fig. 1).

Analysis of Factors Associated with LSM Failure and Unreliable Results

In descriptive analyses, continuous variables were expressed as mean ± standard deviation and categorical variables as absolute figures and percentages. All statistical analyses were performed with LSM failure or unreliable results as the outcome, and with age, sex, BMI, hypertension, type 2 diabetes, indication for TE, operator experience for TE, aspartate aminotransferase, ALT, and GGT as explanatory factors in the total number of examinations. Because some patients underwent several examinations during the 5-year study period, we also took into

Table 1. Factors Associated with LSM Failure in Univariate and Multivariate Analyses

	Univariate			Multivariate		
Parameter	OR	95% CI	*P*	OR	95% CI	*P*
BMI (>30 kg/m^2)	10.5	8.6-12.8	0.0001	8.4	6.6-10.8	0.0001
Operator experience (<500 versus >500 examinations)	2.8	2.1-3.6	0.0001	2.6	1.8-3.9	0.0001
Age (>52 years)	3.5	2.7-4.4	0.0001	2.2	1.6-2.9	0.0001
Type 2 diabetes (yes versus no)	4.0	3.3-5.0	0.0001	2.0	1.5-2.6	0.0001
Hypertension (yes versus no)	3.1	2.5-3.9	0.0001			
Time of examination (first versus others)	2.0	1.6-2.4	0.0001	1.5	1.2-2.0	0.0001
GGT (>3× ULN)	1.5	1.2-1.8	0.001			

n = 13,369 examinations.

account the time of examination (first examination versus others).

A subanalysis was performed in the subgroup of 2835 patients at the time of the first examination, also taking into account as explanatory factors either the different features of the metabolic syndrome (WC, type 2 diabetes, hypertension, and hyperlipidemia) or metabolic syndrome (as previously defined) as a composite factor (present versus absent).

The relationship between each factor and the outcome was investigated by using univariate logistic regression, $P < 0.05$ being considered to denote significant associations. Continuous variables such as age and TE operator experience were dichotomized around the median (52 years and ±500 examinations, respectively), unless a cutoff was considered clinically relevant (BMI > 30 kg/m^2, aspartate aminotransferase > 3 × upper limit of normal value [ULN], ALT > 3 × ULN, and GGT > 3 × ULN).

Factors significantly associated with the outcome in univariate analyses were entered in a multivariate logistic model. Odds ratios were estimated from the model and are given with their 95% confidence intervals (CI). Statistical analyses were performed with SPSS software (Statistical Systems, Kayville, UT).

Results

Characteristics of the Study Population

A total of 14,056 examinations were performed, and 687 examinations were excluded because of missing data. Thus, 13,369 examinations were analyzed (7261 patients, 134,239 shots). Some patients underwent several examinations during the study period: one examination (n = 7261); two (n = 3041); three (n = 1626); four (n = 889), and five (n = 552).

At the time of first liver stiffness measurement, mean age was 51.7 ± 13.7 years, 56.% of patients were males, mean BMI was 24.8 ± 4.7 kg/m^2, 42.6% of patients were overweight (BMI ≥25 kg/m^2), 22.7% were hypertensive, and 11.7% were diabetic. The median LSM value was 6.1 kPa (range, 2.5-75.0). The mean number of valid shots per examination was 10.3 ± 2.7, and the mean success rate was 87.6% ± 19.5%. The indications for TE were chronic hepatitis C in 42.7% of cases; nonalcoholic fatty liver disease (NAFLD) in 13.7%; alcoholic liver disease in 11.4%; chronic hepatitis B in 6.8%; and miscellaneous in 25.4%.

The characteristics of the subgroup of 2835 patients analyzed for the effect of metabolic syndrome at the time of first liver stiffness examination were as follows: mean age, 51.9 ± 13.9 years; male sex 55.6%; mean BMI 25.1 ± 4.8 kg/m^2; overweight (BMI ≥25 kg/m^2) 45.3%; mean WC 88.9 ± 13.8 cm; hypertension 23.6%; type 2 diabetes 14.5%; metabolic syndrome 15.1%; median LSM 5.8 kPa (range, 2.5-75.0); mean number of valid shots per examination 10.2 ± 2.5, and mean success rate 87.8% ± 20.1%. Indications for TE were chronic hepatitis C 32.3%; NAFLD 14.8%; alcoholic liver disease 12.2%; chronic hepatitis B 7.1%; and miscellaneous 33.6%.

Frequency and Determinants of LSM Failure in the Total Population

Overall, LSM failed in 420 cases (3.1%). LSM failed in respectively 4%, 2.4%, 2.2%, 1.2% and 1.2% of patients who underwent one, two, three, four, and five examinations ($P < 0.0001$).

Factors associated with LSM failure in univariate and multivariate analyses are shown in Table 1. When taking into account only the 7261 patients at the time of first examination, LSM failure was independently associated with the following factors: BMI greater than 30 kg/m^2 (OR 7.5, 95% CI 5.6-10.2; $P = 0.0001$), operator experience fewer than 500 examinations (OR 2.5, 95% CI 1.6-4.0; $P = 0.0001$); age greater than 52 years (OR 2.3, 95% CI 1.6-3.2; $P = 0.0001$), and type 2 diabetes (OR 1.6, 95% CI 1.1-2.2; $P = 0.009$).

The LSM failure rate ranged from 1.0% in patients with BMI less than 25 kg/m^2 to 41.7% in patients with BMI >40 kg/m^2 (Fig. 2A) and was not influenced by sex (Fig. 2B).

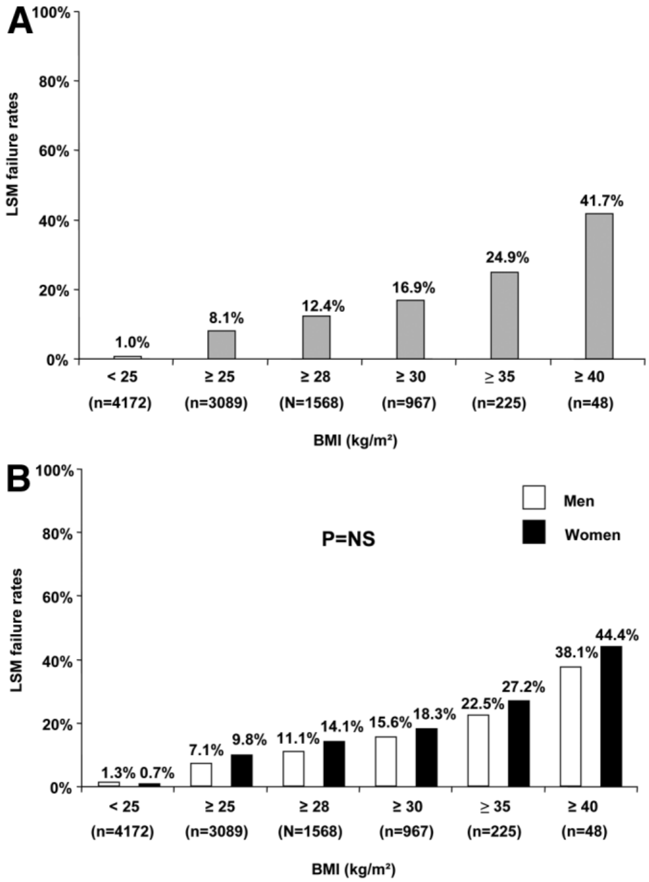

Fig. 2. (A) Liver stiffness measurement failure rates according to BMI in the 7261 patients at the time of the first examination. (B) Liver stiffness measurement failure rates according to BMI and sex in the 7261 patients at the time of the first examination.

The failure rate was 8.3% when the operator had performed fewer than 500 examinations and 3.5% when the operator had performed more than 500 examinations ($P < 0.0001$).

The LSM failure rate ranged from 0.2% in lean young nondiabetic patients (age <52 years and BMI <25 kg/m^2; n = 2439) to 20.9% in elderly obese diabetic patients (age ≥52 years and BMI >30 kg/m^2; n = 201).

Relationship Between LSM Failure and Metabolic Syndrome

LSM failure was observed in 2.6% of cases (2835 patients at the time of first examination). Factors associated with LSM failure in univariate and multivariate analyses when features of the metabolic syndrome were introduced in the model are shown in Table 2. When metabolic syndrome was introduced in the multivariate model as a composite factor, LSM failure was independently associated with age greater than 52 years (OR 2.5, 95% CI 1.4-4.4; $P = 0.002$), GGT greater than 3 × ULN (OR 2.2, 95% CI 1.3-3.6; $P = 0.002$), female sex (OR 2.1, 95% CI 1.3-3.5; $P = 0.002$), and metabolic syndrome (OR 2.1, 95% CI 1.3-3.6; $P = 0.004$).

Frequency and Determinants of Unreliable LSM Results in the Total Population

Among the 12,949 examinations in which LSM did not fail, fewer than 10 valid shots were obtained in 404 cases (3.1%), SR less than 60% in 1055 cases (8.1%), and IQR/LSM greater than 30% in 1189 cases (9.2%). When the three criteria were combined (valid shots <10, SR <60%, or IQR/LSM >30%), unreliable results were obtained in a total of 2046 cases (15.8%).

Unreliable LSM results were observed in, respectively, 17%, 15.2%, 14.5%, 14.3%, and 9.6% of patients who underwent one, two, three, four, and five examinations ($P < 0.0001$).

Factors associated with unreliable LSM results (valid shots <10, SR <60%, or IQR/LSM >30%) in univariate and multivariate analyses are shown in Table 3.

Factors associated with each of the three criteria defining unreliable results are analyzed separately.

Fewer than 10 valid shots were independently associated with BMI greater than 30 kg/m^2 (OR 5.0, 95% CI 3.9-6.2; $P = 0.0001$), age older than 52 years (OR 1.7, 95% CI 1.4-2.2; $P = 0.0001$), and hypertension (OR 1.5, 95% CI 1.2-1.9; $P = 0.001$).

Table 2. Relationship Between LSM Failure and Features of the Metabolic Syndrome in Univariate and Multivariate Analyses

	Univariate			Multivariate		
Parameter	OR	95% CI	*P*	OR	95% CI	*P*
WC (≥80 cm in women or ≥94 cm in men)	25.0	7.8–79.3	0.0001	16.3	5.0–52.8	0.0001
Hypertension (yes versus no)	3.5	2.2–5.5	0.0001			
Age (>52 years)	3.5	2.0–5.9	0.0001			
Type 2 diabetes (yes versus no)	2.7	1.6–4.5	0.0001			
GGT (>3× ULN)	2.2	1.4–3.7	0.001	2.0	1.2–3.4	0.007
Female sex	2.1	1.3–3.5	0.002	1.9	1.1–3.0	0.01

n = 2835 examinations.

Table 3. Factors Associated with Unreliable LSM Results in Univariate and Multivariate Analyses

	Univariate				Multivariate		
Parameter	OR	95% CI	P	Parameter	OR	95% CI	P
Operator experience (<500 versus >500 examinations)	2.7	2.3–3.1	0.0001		3.3	2.7–4.0	0.0001
BMI (>30 kg/m^2)	3.2	2.8–3.6	0.0001		3.1	2.7–3.6	0.0001
Age (>52 years)	2.0	1.8–2.2	0.0001		1.8	1.6–2.0	0.0001
Type 2 diabetes (yes versus no)	1.7	1.5–2.0	0.0001		1.2	1.0–1.4	0.02
Hypertension (yes versus no)	1.9	1.7–2.1	0.0001		1.2	1.1–1.4	0.003
Female sex	1.3	1.1–1.4	0.0001		1.2	1.1–1.3	0.004
Time of examination (first versus others)	1.2	1.1–1.3	0.0001		1.1	1.0–1.2	0.048
ALT (>3× ULN)	0.8	0.7–0.9	0.007		0.8	0.7–0.9	0.042

Valid shots <10, SR <60%, or IQR/LSM >30%.
n = 12949 examinations.

SR less than 60% was independently associated with BMI greater than 30 kg/m^2 (OR 5.5, 95% CI 4.7-6.4; P = 0.0001), age older than 52 years (OR 2.3, 95% CI 2.0-2.7; P = 0.0001), operator experience fewer than 500 examinations (OR 2.1, 95% CI 1.6-2.8; P = 0.0001), type 2 diabetes (OR 1.4, 95% CI 1.1-1.7; P = 0.001), hypertension (OR 1.4, 95% CI 1.2-1.7; P = 0.0001), and female sex (OR 1.2, 95% CI 1.1-1.7, P = 0.004).

IQR/LSM greater than 30% was independently associated with operator experience fewer than 500 examinations (OR 3.7, 95% CI 3.1-4.6; P = 0.0001), age older than 52 years (OR 1.6, 95% CI 1.4-1.8; P = 0.0001), BMI greater than 30 kg/m^2 (OR 1.5, 95% CI 1.2-1.8; P = 0.0001), and first examination (OR 1.1, 95% CI 1.0-1.3; P = 0.04).

When taking into account only the 6968 patients without failure at the time of first examination, unreliable LSM results were independently associated with: BMI greater than 30 kg/m^2 (OR 3.3, 95% CI 2.8-4.0; P = 0.0001), operator experience fewer than 500 examinations (OR 3.1, 95% CI 2.4-3.9; P = 0.0001), age older than 52 years (OR 1.8, 95% CI 1.6-2.1; P = 0.0001), female sex (OR 1.4, 95% CI (1.2-1.6; P = 0.0001), hypertension (OR 1.3, 95% CI 1.1-1.5; P = 0.003), and type 2 diabetes (OR 1.2, 95% CI 1.0-1.5; P = 0.05).

Overall, unreliable results were obtained in 30.5% of cases when the operator had performed fewer than 500 examinations and in 15.6% of cases when the operator had performed more than 500 examinations (P < 0.0001).

Unreliable LSM results ranged from 12.0% in patients without failure and BMI less than 25 kg/m^2 to 53.6% in patients with BMI greater than 40 kg/m^2 (Fig. 3A). For a given BMI, unreliable LSM results were significantly more frequent in women than in men (for example, BMI ≥ 30 kg/m^2: 41.4% and 30.5%, respectively, P = 0.0001; Fig. 3B).

Finally, the frequency of unreliable LSM results ranged from 7.2% in lean young men (age <52 years, BMI <25 kg/m^2, no diabetes or hypertension; n = 1157) to 60.4% in elderly obese diabetic and hypertense women (age ≥52 years, BMI >30 kg/m^2, diabetes and hypertension; n = 62).

Relationship Between Unreliable LSM Results and Metabolic Syndrome

Unreliable LSM results were observed in 15.3% of cases (2760 patients without failure at the time of first

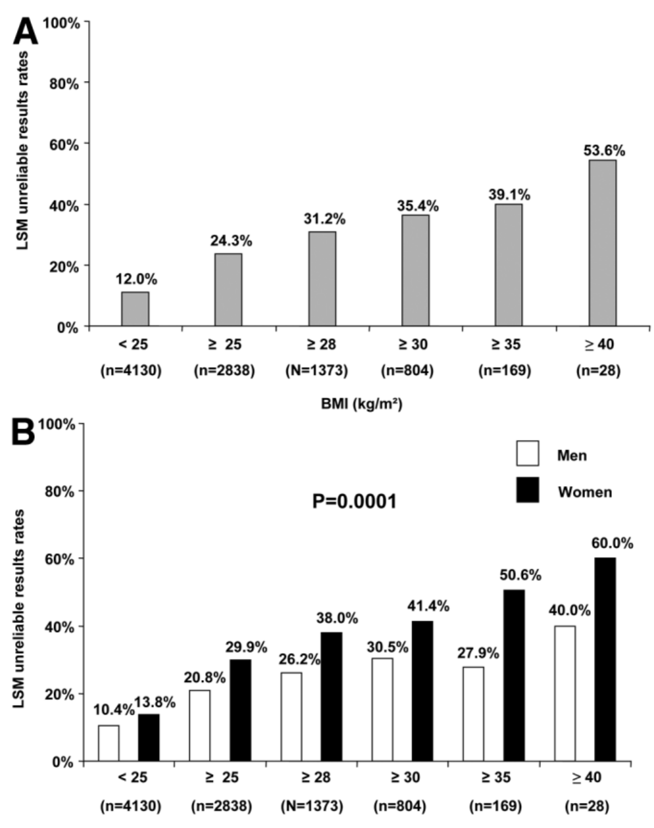

Fig. 3. (A) Rate of unreliable results according to BMI in the 6968 patients without LSM failure at the time of the first examination. (B) Rate of unreliable results according to BMI and sex in the 6968 patients without LSM failure at the time of the first examination.

Table 4. Relationship Between Unreliable LSM Results and Features of the Metabolic Syndrome in Univariate and Multivariate Analyses

	Univariate			Multivariate		
Parameter	OR	95% CI	*P*	OR	95% CI	*P*
WC (’80 cm in women or ≥94 cm in men)	3.9	3.0–4.8	0.0001	3.0	2.3–3.9	0.0001
Operator experience (<500 versus >500 examinations)	2.4	1.6–3.7	0.0001	2.5	1.6–4.0	0.0001
Type 2 diabetes (yes versus no)	2.0	1.6–2.5	0.0001			
Hypertension (yes versus no)	2.1	1.7–2.7	0.0001			
Hyperlipidemia (yes versus no)	1.5	1.2–1.9	0.0001			
Age (>52 years)	2.3	1.8–2.8	0.0001	1.5	1.1–1.9	0.002
Female sex	1.4	1.2–1.8	0.001	1.3	1.0–1.6	0.03

n = 2760 examinations.

examination). Factors associated in univariate and multivariate analyses with unreliable LSM results when features of the metabolic syndrome were introduced in the model are shown in Table 4.

When metabolic syndrome was introduced in the model as a composite factor, unreliable LSM results were independently associated with metabolic syndrome (OR 2.2, 95% CI 1.7-2.8; $P = 0.0001$), operator experience fewer than 500 examination (OR 2.2, 95% CI 1.4-3.5; $P = 0.001$), age older than 52 years (OR 1.8, 95% CI 1.4-2.3; $P = 0.0001$), and female sex (OR 1.4, 95% CI 1.1-1.7; $P = 0.003$).

Discussion

In this prospective 5-year study based on more than 13,000 examinations, LSM results were not interpretable in 18.4% (2466 of 13,369) of cases. Among the three criteria recommended by the manufacturer to define reliable results, failure to comply with IQR/LSM ≤30% was the most frequent reason for unreliable results (9.2%), closely followed by SR ≥60% (8.1%) and valid shots ≥10 (3.1%), respectively. When these three criteria were combined, 15.8% of examinations were uninterpretable. Failure, defined as zero valid shots, was observed in a further 3.1% of cases.

This failure rate is slightly lower than the one we obtained in a previous study (4.5% of 2114 examinations).[27] However, a significant decrease in the failure rate was observed over time in patients who underwent several examinations (4.0% versus 1.2% at the 1st and 5th examinations, respectively, $P < 0.0001$). These results suggest that this improvement might be related to greater operator experience, although we cannot rule out the possibility that patients at high risk of LSM failure, and especially those with a high BMI, tended to be spontaneously excluded from TE examination over time. It is noteworthy, however, that the mean BMI was nearly identical in the two studies (24.8 ± 4.7 and 24.4 ± 4.4).

Similarly, the rate of unreliable LSM results significantly decreased over time in patients who underwent several examinations (17.0% versus 9.6% at the 1st and 5th examinations, respectively, $P < 0.0001$), suggesting again the role of operator experience. When taking into account only the first examinations, limited operator experience and obesity were the principal reasons for LSM failure and unreliable results. It should be stressed that the threshold we used to define operator experience (>500 examinations) is higher than previously reported (>100 examinations) because the current study is based on a very large number of examinations with highly experienced operators (each of them having performed at least 100 examinations).

The importance of operator experience challenges claims[28,29] that LSM requires no learning curve and that a novice can consistently obtain reliable results after a short training period of approximately 50 examinations. However, these claims were based on a small number of examinations and did not take the manufacturer's recommendations into consideration. For instance, a success rate of 30% was considered satisfactory in one study[28] and more than five valid shots in the other,[29] whereas the IQR/LSM less than 30% criterion was not taken into account in either study.[28,29] However, we found that operator experience influenced not only the success rate, as previously reported,[28,29] but also the IQR/LSM ratio, which was recently shown to be critical for LSM accuracy.[30] Our findings therefore emphasize the need for adequate LSM operator training.

Obesity (BMI >30 kg/m^2) was another important determinant of LSM failure and unreliable results. For instance, LSM failure rate ranged from 1% in patients with BMI less than 25 to 41.7% in patients with BMI 40 or higher (Fig. 2A). This influence of BMI might partly account for the wide range of LSM failure rates (2.4%-9.4%) reported in the literature.[2] Similarly, unreliable LSM results rate ranged from 12% in patients with BMI less than 25 to 53.6% in patients with BMI 40 or higher (Fig. 3A).

Apart from obesity, other patient characteristics such as age, female sex, type 2 diabetes, hypertension, and ALT levels influenced LSM reliability. The influence of age and BMI has already been observed in several studies,[28,29,31] but the influence of sex, type 2 diabetes, hypertension, and ALT is a new finding. The absence of flare (ALT $<3 \times$ ULN) was associated with less unreliable LSM results. We have no clear explanation for this finding; however, when taking into account only the first examinations, ALT was no longer a determinant of unreliable results, whereas female sex, type 2 diabetes, and hypertension remained determinants of reliability. Type 2 diabetes and hypertension are, with obesity (defined according to waist circumference and sex) and hyperlipidemia, features of the metabolic syndrome.[26] Interestingly, when metabolic syndrome was specifically studied as a composite factor in a subgroup of patients, it was independently associated with both LSM failure and reliability. However, when the different features of metabolic syndrome were introduced in the model, waist circumference was the most important determinant of LSM failure (Table 2) and reliability (Table 4). Although we need to be cautious because waist circumference was not available in the total population, this is obviously an important and novel finding of the current study. It is consistent with the fact that in patients with large waist circumference, the subcutaneous and prehepatic fat thickness is increased, attenuating both elastic waves and ultrasound and making LSM unreliable or impossible with a regular TE probe that is calibrated for a given distance between the liver and the chest wall. TE probes specifically designed for obese patients should soon be available.

Our findings also may have important implications for current and future applications of TE. Indeed, the increasing prevalence of obesity, insulin resistance, and the metabolic syndrome is changing the face of chronic liver disease; in particular, nearly one third of American adults now have NAFLD.[32] Two TE studies have given promising preliminary results for the prediction of fibrosis stages in patients with NAFLD,[11,12] but this application may be limited by obesity. The LSM failure rates were quite low in these two studies (5% and 4%, respectively), and similar to those reported elsewhere in patients without NAFLD,[27] but these studies involved Japanese patients[12] and children[11] with low mean BMI values (26.6 ± 4.2 and 26 ± 4, respectively). In addition, the study populations were small (97 and 52 subjects, respectively). LSM accuracy and failure rates need to be studied more closely in western adults with NAFLD, but it is already clear that probe refinements will be needed for stiffness measurement in this population.[33]

In conclusion, this study emphasizes the need for rigorous scientific evaluation of all novel procedures such as TE before their widespread uptake. Indeed, our results, based on a very large study population, show that LSM are uninterpretable in nearly one in five cases. Obesity, particularly increased waist circumference, as well as limited operator experience, are the main determinants of LSM failure or unreliable results. As TE is becoming an increasingly popular technique, we believe that such a message is important in routine clinical practice for physicians managing patients with liver diseases and that they should be aware of the pitfalls of LSM. Although TE has been shown to be a reproducible technique,[31] efforts should be made to further standardize its practice. Our findings emphasize the importance of adequate operator training and the need to follow the manufacturer's recommendations for interpreting LSM results, particularly in obese patients. Further technological refinements are required for specific populations.

Acknowledgment: The authors thank Prof. J. M. Pawlotsky for his helpful comments on the manuscript.

References

1. 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 Med Biol 2003;29:1705-1713.
2. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. J Hepatol 2008;48:835-847.
3. Roulot D, Czernichow S, Le Clesiau H, Costes JL, Vergnaud AC, Beaugrand M. Liver stiffness values in apparently healthy subjects: influence of gender and metabolic syndrome. J Hepatol 2008;48:606-613.
4. Ziol M, Handra-Luca A, Kettaneh A, Christidis C, Mal F, Kazemi F, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. HEPATOLOGY 2005;41:48-54.
5. 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;128:343-350.
6. 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. J Viral Hepat 2007;14:360-369.
7. Marcellin P, Ziol M, Bedossa P, Douvin C, Poupon R, de Ledinghen V, et al. Non-invasive assessment of liver fibrosis by stiffness measurement in patients with chronic hepatitis B. Liver Int 2009;29:242-247.
8. de Ledinghen V, Douvin C, Kettaneh A, Ziol M, Roulot D, Marcellin P, et al. Diagnosis of hepatic fibrosis and cirrhosis by transient elastography in HIV/hepatitis C virus-coinfected patients. J Acquir Immune Defic Syndr 2006;41:175-179.
9. Vergara S, Macias J, Rivero A, Gutierrez-Valencia A, Gonzalez-Serrano M, Merino D, et al. The use of transient elastometry for assessing liver fibrosis in patients with HIV and hepatitis C virus coinfection. Clin Infect Dis 2007;45:969-974.
10. Corpechot C, El Naggar A, Poujol-Robert A, Ziol M, Wendum D, Chazouilleres O, et al. Assessment of biliary fibrosis by transient elastography in patients with PBC and PSC. HEPATOLOGY 2006;43:1118-1124.
11. Nobili V, Vizzutti F, Arena U, Abraldes JG, Marra F, Pietrobattista A, et al. Accuracy and reproducibility of transient elastography for the diagnosis of

fibrosis in pediatric nonalcoholic steatohepatitis. HEPATOLOGY 2008;48: 442-448.
12. Yoneda M, Yoneda M, Mawatari H, Fujita K, Endo H, Iida H, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with nonalcoholic fatty liver disease (NAFLD). Dig Liver Dis 2008;40:371-378.
13. Carrion JA, Navasa M, Bosch J, Bruguera M, Gilabert R, Forns X. Transient elastography for diagnosis of advanced fibrosis and portal hypertension in patients with hepatitis C recurrence after liver transplantation. Liver Transpl 2006;12:1791-1798.
14. Rigamonti C, Donato MF, Fraquelli M, Agnelli F, Ronchi G, Casazza G, et al. Transient elastography predicts fibrosis progression in patients with recurrent hepatitis C after liver transplantation. Gut 2008;57:821-827.
15. 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;55:403-408.
16. Ganne-Carrie N, Ziol M, de Ledinghen V, Douvin C, Marcellin P, Castera L, et al. Accuracy of liver stiffness measurement for the diagnosis of cirrhosis in patients with chronic liver diseases. HEPATOLOGY 2006;44:1511-1517.
17. Friedrich-Rust M, Ong MF, Martens S, Sarrazin C, Bojunga J, Zeuzem S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology 2008;134:960-974.
18. 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. J Hepatol 2009;50:59-68.
19. Castera L, Denis J, Babany G, Roudot-Thoraval F. Evolving practices of non-invasive markers of liver fibrosis in patients with chronic hepatitis C in France: time for new guidelines? J Hepatol 2007;46:528-529.
20. Rockey DC, Bissell DM. Noninvasive measures of liver fibrosis. HEPATOLOGY 2006;43:S113-S120.
21. Pinzani M, Vizzutti F, Arena U, Marra F. Technology Insight: noninvasive assessment of liver fibrosis by biochemical scores and elastography. Nat Clin Pract Gastroenterol Hepatol 2008;5:95-106.
22. Castera L. Assessing liver fibrosis. Expert Rev Gastroenterol Hepatol 2008; 2:541-552.
23. Arena U, Vizzutti F, Corti G, Ambu S, Stasi C, Bresci S, et al. Acute viral hepatitis increases liver stiffness values measured by transient elastography. HEPATOLOGY 2008;47:380-384.
24. Sagir A, Erhardt A, Schmitt M, Haussinger D. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. HEPATOLOGY 2007;47:592-595.
25. Millonig G, Reimann FM, Friedrich S, Fonouni H, Mehrabi A, Buchler MW, et al. Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. HEPATOLOGY 2008;48:1718-1723.
26. Alberti KG, Zimmet P, Shaw J. The metabolic syndrome: a new worldwide definition. Lancet 2005;366:1059-1062.
27. Foucher J, Castera L, Bernard PH, Adhoute X, Laharie D, Bertet J, 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.
28. Boursier J, Konate A, Guilluy M, Gorea G, Sawadogo A, Quemener E, et al. Learning curve and interobserver reproducibility evaluation of liver stiffness measurement by transient elastography. Eur J Gastroenterol Hepatol 2008;20:693-701.
29. Kettaneh A, Marcellin P, Douvin C, Poupon R, Ziol M, Beaugrand M, et al. Features associated with success rate and performance of FibroScan measurements for the diagnosis of cirrhosis in HCV patients: a prospective study of 935 patients. J Hepatol 2007;46:628-634.
30. Lucidarme D, Foucher J, Le Bail B, Vergniol J, Castera L, Duburque C, et al. Factors of accuracy of transient elastography (fibroscan) for the diagnosis of liver fibrosis in chronic hepatitis C. HEPATOLOGY 2009:49:1083-1089.
31. Fraquelli M, Rigamonti C, Casazza G, Conte D, Donato MF, Ronchi G, et al. Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. Gut 2007;56:968-973.
32. Torres DM, Harrison SA. Diagnosis and therapy of nonalcoholic steatohepatitis. Gastroenterology 2008;134:1682-1698.
33. Afdhal NH, Curry M. Technology evaluation: a critical step in the clinical utilization of novel diagnostic tests for liver fibrosis. J Hepatol 2007;46: 543-545.