Significant Variations in Elastometry Measurements Made Within Short-term in Patients With Chronic Liver Diseases

Fabio Nascimbeni,[*,‡] Pascal Lebray,[*] Larysa Fedchuk,[*] Claudia P. Oliveira,[§] Mario Reis Alvares-da-Silva,[‖] Anne Varault,[*] Patrick Ingiliz,[*] Yen Ngo,[¶] Mercedes de Torres,[*] Mona Munteanu,[¶] Thierry Poynard,[*] and Vlad Ratziu,[*] for the LIDO Study Group

*Université Pierre et Marie Curie, Assistance Publique Hôpitaux de Paris, Hôpital Pitié Salpêtrière, INSERM U938, CdR Saint-Antoine, Paris, France; ‡Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy; §Department of Gastroenterology, Clinical Division (LIM-07), University of São Paulo School of Medicine, Sao Paulo, Brazil; ‖School of Medicine, Universidade Federal do Rio Grande do Sul, Division of Gastroenterology and Hepatology, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; and ¶BioPredictive, Paris, France

BACKGROUND & AIMS: **Transient elastometry is a noninvasive procedure used to measure fibrosis when patients are diagnosed with liver disease; it might be used to monitor changes over time. We investigated whether there are short-term variations in stiffness measurements that are not attributable to changes in fibrosis by studying patients with stable liver disease.**

METHODS: **We performed a retrospective analysis of 531 paired liver stiffness measurements made by Fibroscan when the study began (LSM1) and at follow-up (LSM2), more than 1 day and less than 1 year apart, from 432 stable (for body mass index, waist circumference, and alcohol consumption), untreated, immunocompetent patients with chronic liver disease (from January 2006 through March 2009). Variations between the first and follow-up measurements were expressed as absolute (LSM2-LSM1, kPa) or relative ([LSM2-LSM1]/LSM1*100) or as changes in fibrosis stage.**

RESULTS: **There was >20% variation in 49.7%, >30% in 34.3%, and >50% in 12.2% of paired measurements; this variation was constant across the spectrum of LSM1 values. The variations produced a 1-fibrosis stage difference in 31.5% of pairs and a ≥2-stage difference in 9.8% of pairs. Patients with LSM1 >7 kPa had increased probability of having a different stage of fibrosis at LSM2, compared with patients with LSM1 <7 kPa. Factors associated with variation included measurements made by 2 different operators or at least 1 non-senior operator, ratios of interquartile range:median values, significant fibrosis (≥7 kPa) at LSM1, baseline body mass index, or a 2-fold difference in level of alanine aminotransferase between measurements. When the analyses were restricted to measurements made by the same operator, the variation was slightly reduced; fibrosis stage differed between measurements for only 34.3% of cases.**

CONCLUSIONS: **Operator-related and patient-related factors produce significant variations in liver stiffness measurements made by transient elastometry, limiting its use in monitoring patients. These variations are unrelated to disease progression. The lowest levels of variation occur in measurements made in patients with no or early-stage fibrosis or by a single experienced operator.**

Keywords: Liver Fibrosis Markers; Cirrhosis; BMI; ALT; Disease Progression.

See editorial on page 780; see similar article on page 772.

Transient elastometry (TE) is a popular method for the noninvasive assessment of liver fibrosis. It measures the velocity of shear waves propagating through the liver, which is directly related to tissue stiffness and hence provides an indication on fibrosis. It has good overall accuracy for the diagnosis of advanced fibrosis/cirrhosis, independent of the etiology.[1–3] TE values are correlated with cirrhosis complications, long-term outcomes, and survival.[4–9] After antiviral treatment, liver stiffness measurement (LSM) values

Abbreviations used in this paper: ALT, alanine aminotransferase; ANOVA, analysis of variance; BMI, body mass index; CI, confidence interval; HBV, hepatitis B virus; HCV, hepatitis C virus; ICC, intraclass correlation coefficient; IQR, interquartile range; IQR/M, interquartile range/median ratio; LSM, liver stiffness measurement; NAFLD, nonalcoholic fatty liver disease; SD, standard deviation; TE, transient elastography.

1542-3565/$36.00
http://dx.doi.org/10.1016/j.cgh.2014.07.037

may improve, suggesting fibrosis regression.[10,11] On the basis of these data, it has been proposed that repeated TE measurements could be useful for the monitoring of patients with chronic liver diseases.[12] This entails that changes in LSM should not be prone to substantial analytical variability but rather reflect true variation in the course of the disease. Remarkably, the variability of LSM unrelated to disease progression/regression is unknown.

In this study we attempted to examine the variability of short-term TE measurements in patients with no or minimal progression/regression of chronic liver disease. We assumed that because of the slow course of liver fibrosis in most chronic liver diseases, an interval of 1 year or less is too short for the occurrence of meaningful progression or regression of fibrosis in untreated patients with stable condition. The assumption is that for TE to be a reliable monitoring tool, LSM values should not fluctuate, or only minimally, independent of the course of the disease, and that only a minority of patients should have changes in LSM within a clinically relevant range.

Methods

Patients

We retrospectively analyzed our database of TE examinations between January 2006 and March 2009 and identified all adult patients who had undergone repeated, successful LSM more than 1 day and less than 1 year apart. All chronic liver diseases were included as long as patients were not receiving specific treatment including antiviral therapy, glitazones, vitamin E, DPP-4 inhibitors, or GLP-1 analogues for nonalcoholic steatohepatitis and endoscopic treatment for chronic cholestatic liver diseases. Only immunocompetent patients (non-transplanted and CD4 count $>300/mm^3$) with stable alcohol consumption ($\pm$20 g/day) between the baseline and follow-up LSM were included. No specific dietary interventions or physical activity programs were implemented. Patients with decompensated cirrhosis, hepatitis flare-ups, acute liver diseases, or unstable clinical conditions were excluded.

Clinical and laboratory parameters at the time of each LSM were age, sex, body mass index (BMI), waist circumference, thoracic fold thickness, alanine aminotransferase (ALT) and SteatoTest, etiology of liver disease, and TE operator.

Liver Stiffness Measurements

LSMs were performed by using FibroScan with an M probe (Echosens, Paris, France) in a typically fasting patient. LSMs were successful if at least 10 shots were valid, the success rate was $\geq$60%, and the interquartile range (IQR) $<$30% of the median LSM (IQR/M). In our department, empirical fibrosis monitoring practices repeated the noninvasive assessment every 6–12 months. In other patients, repeated assessment was performed because of inclusion in clinical trials, observational or therapeutic, on patient request, because of discordant results with other noninvasive serum markers, or the investigator's willingness to check results indicative of advanced fibrosis. A total of 14 operators were active. Operators were defined as senior or non-senior according to their experience (more vs less than 1000 examinations). Variability between baseline and follow-up paired LSMs was defined as absolute (LSM2-LSM1, in kPa), relative [percent change from first measurement (LSM2-LSM1)/LSM1*100], or as a change in fibrosis stage. We considered the following cutoff values of LSM: $<$7 kPa for absent/mild fibrosis (F0-F1), $\geq$7 to $<$9.5 kPa for significant fibrosis (F2), $\geq$9.5 to $<$12.5 kPa for severe fibrosis (F3), $\geq$12.5 to $<$20 kPa for early cirrhosis, and $\geq$20 kPa for advanced cirrhosis. Because most patients were hepatitis C virus (HCV) monoinfected or coinfected, we chose to extend Castera's cutoffs to all liver disease etiologies.[1] Moreover, because cirrhosis is multistage[13] and LSM predicts liver-related complications and death,[5–9] we split the F4 stage into early and advanced cirrhosis with a cutoff of 20 kPa, because this threshold is associated with development of esophageal varices and liver-related complications.[4–7]

Statistical Analysis

In descriptive analyses, continuous variables were expressed as mean $\pm$ standard deviation (SD) or median with IQR and categorical variables as frequency and percentage. LSM variability between baseline and follow-up paired TE was the outcome, and age, sex, etiology, time between TE1 and TE2, different operators, operator experience, maximum IQR/M and success rate between paired readings, and fibrosis class at TE1 were explanatory factors. Baseline BMI, waist circumference, thoracic fold thickness, ALT, and SteatoTest and absolute/relative changes between paired TE1 and TE2 were additional explanatory factors. The most discriminant cutoff values for relative changes were chosen. Comparisons between paired variables at TE1 and TE2 were performed by using paired sample t test or McNemar tests for numerical or categorical variables. Correlations of LSM values and fibrosis classes were assessed with the Spearman rank test and Kendall tau-b test. The Bland–Altman graphic was plotted, and inter-rater agreement was assessed by using kappa index and intraclass correlation coefficient (ICC). Numerical variables were compared by using the Student t test and the Mann–Whitney U test. The χ^2 and Fisher exact tests were used for qualitative data. Two-sided P values $<$.05 were considered statistically significant. Binary logistic regression analyses were used with factors that were significant in univariate analyses. Sensitivity analyses of the weight of a single variable on LSM variability and

subanalyses with etiology-specific LSM thresholds were also performed. For HCV we used 7.1, 9.5, and 12.5,[1] for hepatitis B virus (HBV) we used 7.2, 8.1, and 11,[2] and for nonalcoholic fatty liver disease (NAFLD) we used 5.8, 7.9, and 10.3 kPa.[3] Statistical tests were performed by using SPSS 17.0 software (SPSS, Inc, Chicago, IL).

Results

Study Population

After exclusion of single TE, paired TE performed on the same day or more than 1 year apart, failed or unreliable TE, and patients with specific etiologic treatment, 531 paired TE from 432 patients were analyzed (Supplementary Figure 1). The majority of patients were male (71.3%), and the mean age was 50.6 ± 11.4 years. The causes of liver disease were 43.1% chronic HCV, 18.1% human immunodeficiency virus–HCV coinfection, 16.0% NAFLD, 12.3% chronic HBV, 3.7% alcoholic liver disease, and 6.9% miscellaneous etiologies. The mean time interval between paired TE was 7.8 ± 3.4 months (median [IQR], 8.3 [5.6–10.9]). The same operator performed the baseline and the follow-up readings in 31.8% of cases. Of the 14 operators, 4 were considered senior. These senior operators performed 71.6% of paired TE.

At the time of the first examination, the mean LSM value was 10.2 ± 10.0 kPa (median [IQR], 7.0 [5.6–10.2]); the mean IQR/M was 15.7% ± 6.9%, and the mean success rate was 93.6% ± 10.3%. Half of the measurements were classified as absent/mild fibrosis, 21.1% as significant fibrosis, 12.4% as severe fibrosis, and 7.7% and 8.9% as early and advanced cirrhosis, respectively. Mean LSM, IQR/M, success rate, and frequency of fibrosis classes did not significantly change between the baseline and follow-up readings (Table 1). However, the correlation/agreement between baseline and follow-up LSMs was only moderate (Spearman rank test, $r = 0.674$, $P < .001$; ICC, 0.75; 95% confidence interval [CI], 0.71–0.79; several outliers in the Bland–Altman plot) (Supplementary Figure 2). Also, the correlation between baseline and follow-up fibrosis classes was only fair (Kendall tau-b, 0.612; $P < .001$) (Supplementary Table 1). In addition, inter-rater reliability between paired readings ranged from fair to moderate (k values of 0.38, 0.55, 0.60, and 0.64 for any grade of fibrosis, fibrosis stages F ≥2, fibrosis stages F ≥3, and cirrhosis, respectively).

Overall Changes Between Liver Stiffness Measurements

The mean absolute difference between LSM1 and LSM2 was 3.2 ± 6.1 kPa, equally distributed among positive (LSM1>LSM2) and negative (LSM1<LSM2) changes. A relative discordance >20% occurred in 49.7%, >30% in 34.3%, and >50% in 12.2% of paired readings. The mean absolute difference according to baseline fibrosis class was 1.4 ± 1.7 kPa for absent/mild fibrosis, 2.7 ± 6.9 kPa for significant fibrosis, 3.4 ± 2.8 kPa for severe fibrosis, and 5.2 ± 3.7 kPa and 12.4 ± 12.7 kPa for early and advanced cirrhosis, respectively. The absolute difference was directly correlated with LSM1 (Figure 1*A* and *C*). Conversely, the relative difference was high throughout the spectrum of LSM1 values (Figure 1*B* and *D*), with a mean relative difference between the 2 paired readings as high as 29.4% ± 48.0%. Altogether, these

Table 1. Overall Changes Between Baseline (LSM1) and Follow-up (LSM2) Paired LSMs

	LSM1	LSM2	Mean absolute difference between LSM1 and LSM2	*P* value
Fibroscan parameters				
LSM (*kPa*)	10.2 ± 10.0	9.7 ± 9.3	3.2 ± 6.1	.076
IQR/M (%)	15.7 ± 6.9	15.1 ± 6.4	—	.105
Success rate (%)	93.6 ± 10.3	94.0 ± 9.9	—	.463
Fibrosis classes			—	.458
Absent/mild fibrosis (<7 kPa)	49.9	52.7		
Significant fibrosis (≥7 to <9.5 kPa)	21.1	21.1		
Severe fibrosis (≥9.5 to <12.5 kPa)	12.4	10.9		
Early cirrhosis (≥12.5 to <20 kPa)	7.7	8.1		
Advanced cirrhosis (≥20 kPa)	8.9	7.2		
Senior operator (>1000 exams)	78.5	90.6	—	<.001
Anthropometric and biochemical variables				
Weight (*kg*)	72.0 ± 13.4	71.7 ± 13.0	2.5 ± 3.4	.159
BMI (kg/m^2)	24.4 ± 3.6	24.4 ± 3.6	0.8 ± 1.1	.171
Waist circumference (*cm*)	85.8 ± 11.2	84.8 ± 11.0	4.8 ± 4.9	.053
Thoracic fold thickness (*mm*)	13.2 ± 7.1	13.2 ± 6.9	3.9 ± 4.9	.584
ALT (*U/L*)	71.9 ± 102.9	59.9 ± 45.3	26.1 ± 75.7	.135
SteatoTest	0.36 ± 0.21	0.36 ± 0.21	0.09 ± 0.09	.383

NOTE. All data are expressed as percentage or as mean ± SD.
LSM1, liver stiffness measurement at baseline; LSM2, liver stiffness measurement at follow-up.

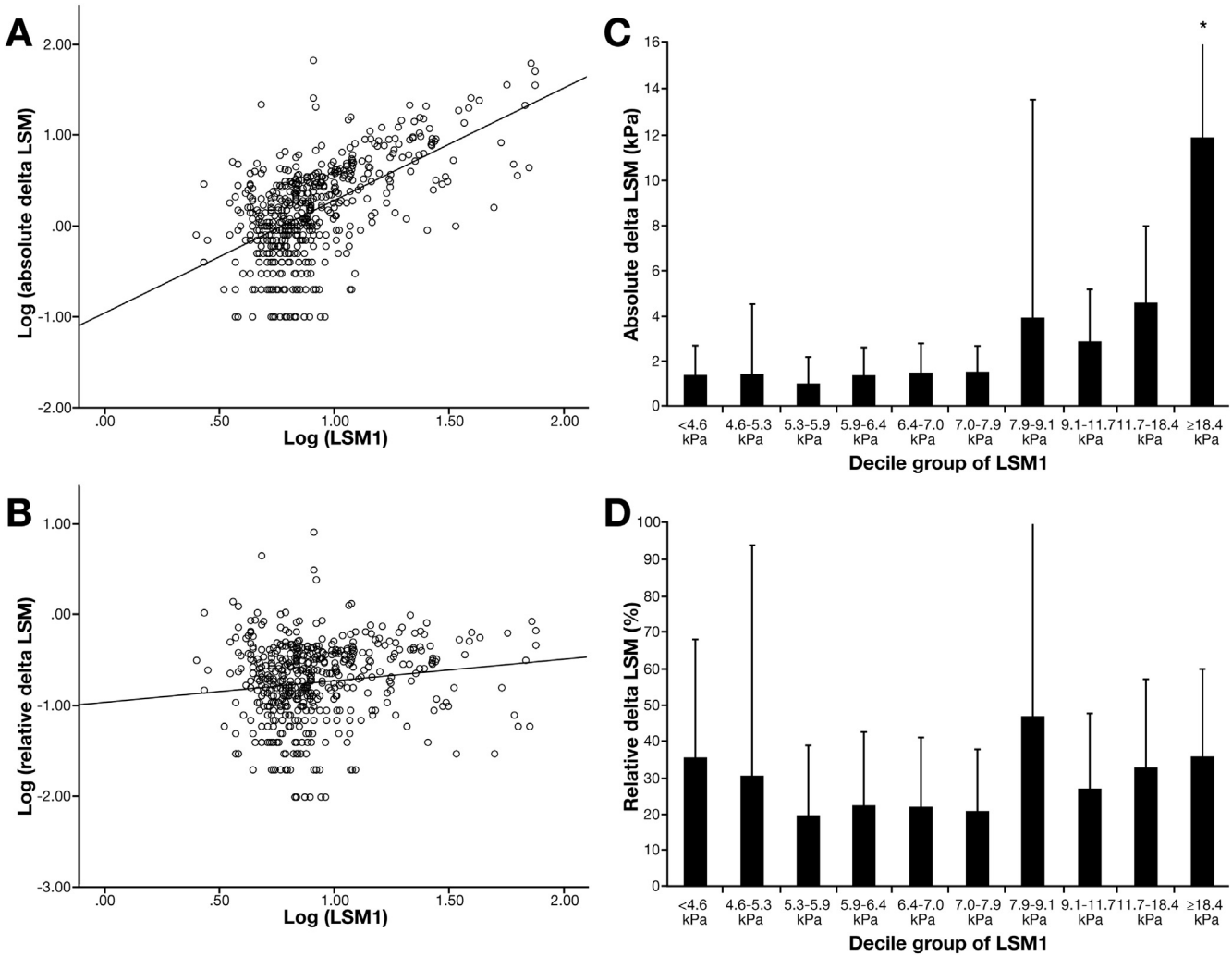


Figure 1. Correlation between baseline liver stiffness measurement (LSM1) and variability between paired LSMs. (*A*) Correlation between logarithmic LSM1 and logarithmic absolute variability [log (LSM2-LSM1)] between paired LSMs ($r = 0.542$; $P < .001$). (*B*) Correlation between logarithmic LSM1 and logarithmic relative variability {log [(LSM2-LSM1)/LSM1*100]} between paired LSMs ($r = 0.150$; $P = .001$). (*C*) Correlation between percentile group (decile) of LSM1 and mean + SD absolute variability (LSM2-LSM1) between paired LSMs. Analysis of variance (ANOVA) *P* for trend < .001. *Tenth decile group significantly differs from other groups according to one-way ANOVA test with post hoc Bonferroni analysis ($P < .001$). (*D*) Correlation between percentile group (decile) of LSM1 and mean + SD relative variability [(LSM2-LSM1)/LSM1*100] between paired LSMs. ANOVA *P* for trend: .234. No significant differences were found between decile groups according to one-way ANOVA test with post hoc Bonferroni analysis.

findings resulted in 1 fibrosis class change in 31.5% of paired examinations and ≥2 classes change in 9.8%. There was a significant positive correlation between LSM1 values and fibrosis class changes ($r = 0.414$, $P < .001$), and the higher the decile of LSM1, the greater the probability of class changes (Figure 2*A*). Patients with significant or severe fibrosis had an alarmingly high rate of ≥1 fibrosis class change on a second TE measurement (Figure 2*B*). Figure 2*A* shows that only patients with LSM1 <6.5 kPa had reduced probability of fibrosis class change on repeated measurements. The best LSM1 cutoff value for predicting at least 1 fibrosis class change was 7.15 kPa, close to the cutoff value for significant fibrosis (Supplementary Figure 3).

The subanalyses according to maximum IQR/M between baseline and follow-up TE displayed an increasing prevalence of discordance and fibrosis class changes for higher IQR/M groups (Figure 3 and Supplementary Material).

Overall, there were no significant modifications between baseline and follow-up TE readings for the mean values of the tested anthropometric and biochemical variables (Table 1).

Factors Associated With Variability Between Baseline and Follow-up Paired Liver Stiffness Measurements

We first assessed the determinants of >20%, >30%, and >50% relative discordance (Supplementary Table 2 and Table 2). Both the presence of 2 different operators

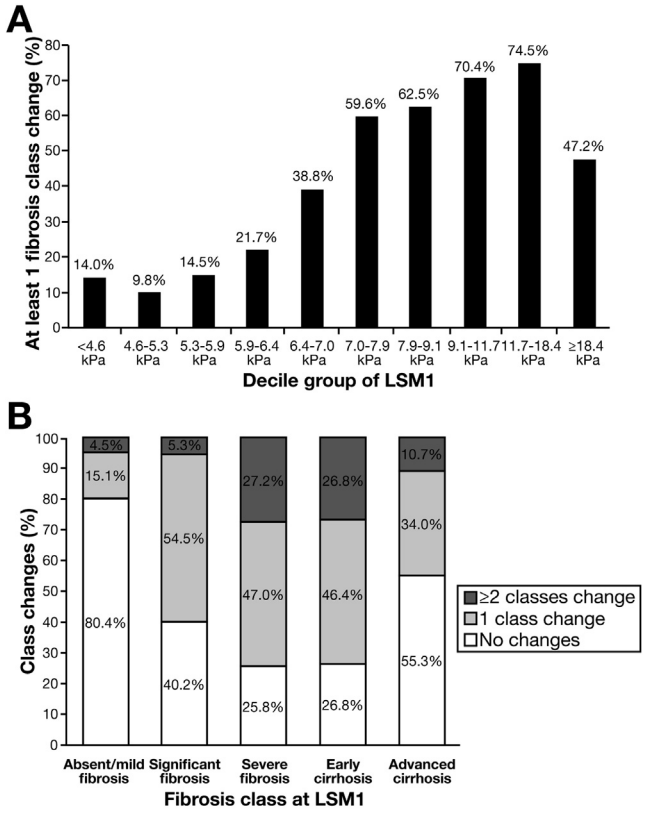


Figure 2. Fibrosis class changes between baseline and follow-up paired LSMs according to baseline liver stiffness measurement (LSM1) decile group and baseline fibrosis class. (*A*) Percentage of at least 1 fibrosis class change between paired LSMs according to baseline decile group of LSM1. The higher the decile group, the greater the probability of at least 1 fibrosis class change between the 2 readings (χ^2 *P* for trend < .001). (*B*) Percentage of class changes according to baseline fibrosis class. χ^2 *P* for trend < .001.

and at least 1 non-senior operator were significantly associated with >20% and >30% discordance; the latter was also a risk factor for >50% discordance. The maximum IQR/M was significantly associated with >20% and >30% discordance, whereas success rate was not. A shorter time interval between TE measurements was associated with a higher discordance rate; this could be confounded by the higher baseline LSM values in paired readings with shorter follow-up and by the level of experience of the operators. With regard to fibrosis stage at the first reading, discordance >20% and >30% were more strongly associated with the presence of at least severe fibrosis, whereas discordance >50% was more significantly associated with cirrhosis. Sex, age, and etiology of chronic liver disease were not significant factors of variability. Although BMI, waist circumference, thoracic fold thickness, ALT, and SteatoTest values were not associated with variability, changes in BMI and waist circumference were significantly associated with >20% discordance. Doubling of ALT was a risk factor for both >20% and >30% discordance.

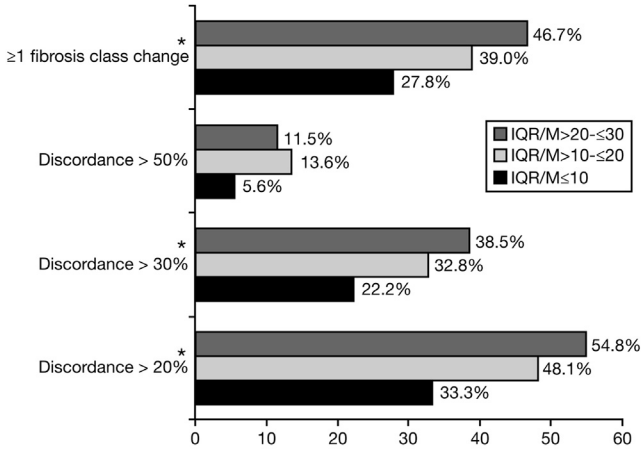


Figure 3. Variability according to maximum IQR/M between paired LSMs. *χ^2 *P* for trend <.05 for discordance >20%, >30%, and ≥1 fibrosis class change.

We next assessed the factors associated with fibrosis class changes (at least 1, and 2 or more) (Supplementary Table 3 and Table 3). In univariate analyses, sex, age, and etiology of liver disease were not determinants of fibrosis class changes. Both at least 1 and ≥2 fibrosis class changes were significantly associated with a less experienced operator. Different operators and maximum IQR/M between paired examinations were risk factors for at least 1 fibrosis class change; the success rate was not. Again, a longer time interval between TE measurements minimized fibrosis class changes. With regard to fibrosis stage at the first reading, at least 1 fibrosis class change was associated with significant fibrosis, whereas 2 or more classes changes were associated with severe fibrosis. The baseline BMI, the doubling of ALT, and relative changes ≥10% in SteatoTest between the 2 TE were significantly associated with at least 1 fibrosis class change. No associations were found between other anthropometric or biochemical variables and ≥2 fibrosis classes changes.

Tables 2 and 3 show the results of multivariate analyses, adjusted for sex, age, and time between examinations. The time interval between LSMs was not predictive of discordance or fibrosis class changes in multivariate analyses.

The results of sensitivity analyses aiming at evaluating the specific effect of time interval, maximum IQR/M, ALT levels, weight changes, and inter-operator variability on paired LSM variability and the subanalyses with etiology-specific LSM thresholds are shown in Supplementary Material and Supplementary Tables 4 and 5.

Discussion

This study demonstrates a substantial variability of repeated, short-term TE measurements in patients with chronic liver diseases. We are confident that these changes in elastometry do not reflect bona fide changes in liver fibrosis, because our selection criteria minimized the risk of fibrosis progression/regression during the study

Table 2. Factors Associated With Relative Variability in Multivariate Analyses

Discordance >20%	Model 1 (age, sex, time, different operators, at least 1 non-senior operator, IQR/M, and severe fibrosis)			Model 1 + relative changes in BMI			Model 1 + relative changes in WC			Model 1 + doubling of ALT		
	OR	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value
Different operators	1.49	0.96–2.29	.074	1.49	0.92–2.44	.108	1.34	0.82–2.20	.245	1.72	0.97–3.03	.064
At least 1 non-senior operator	1.24	0.80–1.92	.338	1.38	0.82–2.34	.226	1.19	0.69–2.05	.533	1.66	0.94–2.93	.083
Maximum IQR/M between LSMs	1.04	1.00–1.07	.030	1.03	0.99–1.07	.104	1.04	1.00–1.08	.038	1.06	1.02–1.10	.007
Severe fibrosis at LSM1	2.58	1.71–3.90	<.001	3.21	1.96–5.24	<.001	2.10	1.29–3.43	.003	3.27	1.86–5.73	<.001
Relative BMI changes ≥7%				2.08	1.14–3.79	.017						
Relative WC changes ≥10%							2.09	1.21–3.59	.008			
Doubling of ALT										3.82	1.17–12.52	.027

Discordance >30%	Model 1 (age, sex, time, different operators, at least 1 non-senior operator, IQR/M, and severe fibrosis)			Model 1 + relative changes in BMI			Model 1 + doubling of ALT		
	OR	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value
Different operators	1.45	0.91–2.31	.123	1.41	0.83–2.38	.204	1.70	0.92–3.14	.094
At least 1 non-senior operator	1.18	0.75–1.84	.477	1.14	0.67–1.95	.619	1.36	0.77–2.41	.291
Maximum IQR/M between LSMs	1.03	0.99–1.06	.093	1.03	0.99–1.06	.195	1.04	1.00–1.09	.054
Severe fibrosis at LSM1	2.35	1.56–3.53	<.001	2.69	1.68–4.31	<.001	2.66	1.56–4.51	<.001
Relative BMI changes ≥7%				1.99	1.11–3.58	.022			
Doubling of ALT							3.27	1.18–9.05	.023

Discordance >50%	Model 1 (age, sex, time, at least 1 non-senior operator, and cirrhosis)		
	OR	95% CI	*P* value
At least 1 non-senior operator	1.82	1.05–3.15	.033
Time interval between LSMs	0.92	0.85–1.00	.046
Cirrhosis at LSM1	2.39	1.29–4.42	.006

NOTE. Significant fibrosis was defined as ≥7 kPa, severe fibrosis as ≥9.5 kPa, and cirrhosis as ≥12.5 kPa.
LSM1, liver stiffness measurement at baseline; WC, waist circumference.

period. Indeed, a time interval of less than 1 year is a very short time compared with the natural history of most chronic liver diseases in which discernible changes in fibrosis take decades to occur. Moreover, the selection of untreated, immunocompetent patients, with stable anthropometric/metabolic parameters and alcohol consumption, further reduced the possibility of meaningful fibrotic disease progression. The significant variability of TE measurements unrelated to disease progression/regression shown in this study highlights the limitations of elastometry used alone as a fibrosis monitoring tool. We also identified some of the determinants of the analytical variability of TE measurements that should incite caution when interpreting repeated measurements.

Although histologic staging is semiquantitative, elastometry-based methods have the theoretical advantage of providing a quantitative assessment of liver fibrosis. Clinicians may be tempted to rely on absolute kPa changes to predict fibrotic disease progression or regression. This study shows that this could be inaccurate, because relative changes in sequential measurements can be >20% in as many as half of patients and >30% in one-third of them. Interestingly, the whole spectrum of elastometry values is prone to such variability, suggesting that changes in tissue stiffness that are due to biological reasons unrelated to fibrosis per se or analytical issues of TE measurement are a general phenomenon occurring at all stages of fibrosis. Therefore, clinicians should not rely on absolute or relative changes in elastometry for patient monitoring. An alternative option is to consider fibrosis class changes. Even by this standard, one-third of sequential measurements incorrectly diagnosed a 1-stage fibrosis change. In contrast, a 2-class change was marginal (<10%) and therefore most probably associated with real disease change.

Importantly, we inquired whether fibrosis class changes that are due to short-term variability of

Table 3. Factors Associated With Fibrosis Class Changes in Multivariate Analyses

	Model 1 (age, sex, time, different operators, at least 1 non-senior operator, IQR/M, and significant fibrosis)			Model 1 + baseline BMI			Model 1 + doubling of ALT			Model 1 + relative SteatoTest changes		
≥1 Fibrosis class change	OR	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value
Different operators	1.62	0.99–2.65	.053	1.44	0.86–2.42	.163	1.56	0.83–2.95	.168	1.65	0.61–4.47	.322
At least 1 non-senior operator	1.09	0.67–1.75	.736	1.13	0.66–1.93	.665	1.91	1.04–3.51	.038	1.24	0.41–3.71	.705
Maximum IQR/M between LSMs	1.03	1.00–1.07	.060	1.02	0.99–1.06	.237	1.03	0.99–1.08	.182	1.05	0.97–1.13	.230
Time interval between LSMs	0.97	0.91–1.03	.296	0.95	0.89–1.02	.156	0.96	0.88–1.05	.324	0.95	0.81–1.12	.538
Significant fibrosis at LSM1	6.56	4.37–9.84	<.001	6.40	4.14–9.88	<.001	6.95	4.18–11.55	<.001	4.20	1.83–9.66	.001
Baseline BMI				1.07	1.01–1.14	.031						
Doubling of ALT							4.24	1.33–13.53	.015			
Relative SteatoTest changes ≥10%										2.83	1.07–7.46	.036

	Model 1 (age, sex, time, at least 1 non-senior operator, and severe fibrosis)		
≥2 Fibrosis classes changes	OR	95% CI	*P* value
At least 1 non-senior operator	2.62	1.42–4.84	.002
Time interval between LSMs	0.96	0.88–1.05	.400
Severe fibrosis at LSM1	5.22	2.74–9.93	<.001

NOTE. Significant fibrosis was defined as ≥7 kPa, severe fibrosis as ≥9.5 kPa, cirrhosis as ≥12.5 kPa.
LSM1, liver stiffness measurement at baseline.

elastometry occurred for a particular range of LSM values. The data show that patients with LSM value <6.5 kPa had significantly lower prevalence of fibrosis class change at a subsequent LSM measurement than patients with LSM >6.5 kPa. Thus, the probability that an increase or a decrease in elastometry values compatible with a 1-class change in fibrosis can be explained by intrinsic variability of elastometry measurements is much higher in patients with significant or advanced fibrosis (baseline LSM >7 kPa with the most validated thresholds). Univariate and multivariate analyses strengthened these conclusions by identifying significant fibrosis on the first elastometry measurement as a factor of substantial short-term variability on a subsequent measurement. These data provide important clarifications on the interpretation of longitudinal follow-up by elastometry.

We identified several procedure-related and patient-related determinants of substantial variability of LSM independent from actual changes in the underlying liver fibrosis. One such factor was operator-dependence. Different operators and/or less experienced operators were major predictors of LSM variability. This was surprising because TE is considered an operator-independent, highly reproducible, and user-friendly method that requires a short learning curve and has an excellent intraobserver and interobserver agreement.[14–16] Previous investigators have challenged this view. Castera et al[17] showed that operator experience strongly impacted LSM failure rate and unreliable results. Poynard et al[18] described an operator effect as a significant cause of discordance between LSM and serum fibrosis markers. Others recently disputed LSM reproducibility, suggesting that interoperator variability could be more common than generally reported.[19,20] Our data also ask for a more balanced view and emphasize the importance of a higher number of examinations for operator experience and the need for clinicians to consider operator experience when interpreting dynamic changes of TE measurements. Restricting the analysis to the same experienced operator somehow reduced the variability, but this was still high enough to impact the clinical relevance of the results. The finding that IQR/M is a robust predictor of TE variability is germane to the finding of operator dependency and is another indication that the quality of the TE examination is critical. Much more than the arbitrarily chosen "success rate," the IQR/M has been shown by others to be a key quality control independently associated with the risk of discordance with histology[21,22] or with interobserver discrepancy.[16] Our data showing that the lower the IQR/M, the lower the variability suggest that careful consideration of this variable is also critical to the correct interpretation of follow-up TE measurements.

Among patient-related factors, BMI is consistently associated with discordant TE readings, even though BMI changes in this population were minimized by the selection criteria. ALT flares, as a surrogate of necroinflammatory activity, have been shown to interfere with

liver stiffness.[23,24] Interestingly, the present data show that even a small increase in ALT levels has a non-negligible impact on the ability of TE to reliably diagnose changes in liver stiffness.

This study has several strengths. It tested a large number of paired TE measurements that were representative of all etiologies of chronic liver diseases, which makes the findings of this report generalizable. Patients were carefully selected for stable fibrotic disease and well-documented for confounding profibrotic factors and confounders of liver stiffness. Several limitations should be mentioned including, first, the retrospective design that calls for a prospective confirmation of our findings, preferably at a predetermined interval between TE readings. Second, there is no histologic documentation that could have confirmed the actual lack of progression/regression of fibrosis during the short study period. Also, we only studied the M probe. Finally, not all patients were fasting; nonetheless, the difference in stiffness between fasting and feeding state is minimal.[25]

In conclusion, there is a strong medical need for reliable, noninvasive methods of fibrosis progression monitoring. This need will only increase in the next years with the advent of efficient but expensive medications for chronic liver diseases. Elastometry has gained widespread acceptance for the baseline assessment of liver fibrosis, but to what extent changes in elastometry reflect changes in fibrosis on follow-up has not been studied. Here we show that there is a large intrinsic variability of repeated elastometry measurements that is not related to changes in fibrosis. In particular, patients with stiffness values higher than 7 kPa and exams by different operators are at high risk of non-fibrosis–related LSM variability. Moreover, LSM changes should be carefully interpreted on an individual case basis, taking into account changes in BMI, ALT, and the IQR/M values between exams. Despite the quantitative nature of liver stiffness, we recommend that for monitoring purposes in the clinic, only changes in fibrosis class should be given full consideration. This study shows that elastometry alone may not be sufficient for a reliable follow-up of fibrosis in patients with chronic liver diseases. Testing whether the combination of elastometry with a different noninvasive method would provide a more accurate reflection of fibrosis progression is worthy of future studies.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at http://dx.doi.org/10.1016/j.cgh.2014.07.037.

References

1. 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.
2. Marcellin P, Ziol M, Bedossa P, et al. Non-invasive assessment of liver fibrosis by stiffness measurement in patients with chronic hepatitis B. Liver Int 2009;29:242–247.
3. Wong VW, Vergniol J, Wong GL, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 2010;51:454–462.
4. Kazemi F, Kettaneh A, N'Kontchou G, et al. Liver stiffness measurement selects patients with cirrhosis at risk of bearing large oesophageal varices. J Hepatol 2006;45:230–235.
5. Bureau C, Metivier S, Peron JM, et al. Transient elastography accurately predicts presence of significant portal hypertension in patients with chronic liver disease. Aliment Pharmacol Ther 2008;27:1261–1268.
6. Castera L, Le Bail B, Roudot-Thoraval F, 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.
7. Robic MA, Procopet B, Metivier S, et al. Liver stiffness accurately predicts portal hypertension related complications in patients with chronic liver disease: a prospective study. J Hepatol 2011;55:1017–1024.
8. Poynard T, Vergniol J, Ngo Y, et al. Staging chronic hepatitis C in seven categories using fibrosis biomarker (FibroTest) and transient elastography (FibroScan). J Hepatol 2014;60:706–714.
9. Vergniol J, Foucher J, Terrebonne E, et al. Noninvasive tests for fibrosis and liver stiffness predict 5-year outcomes of patients with chronic hepatitis C. Gastroenterology 2011;140:1970–1979.
10. Hezode C, Castera L, Roudot-Thoraval F, et al. Liver stiffness diminishes with antiviral response in chronic hepatitis C. Aliment Pharmacol Ther 2011;34:656–663.
11. Enomoto M, Mori M, Ogawa T, et al. Usefulness of transient elastography for assessment of liver fibrosis in chronic hepatitis B: regression of liver stiffness during entecavir therapy. Hepatol Res 2010;40:853–861.
12. Kim BK, Fung J, Yuen MF, et al. Clinical application of liver stiffness measurement using transient elastography in chronic liver disease from longitudinal perspectives. World J Gastroenterol 2013;19:1890–1900.
13. Garcia-Tsao G, Friedman S, Iredale J, et al. Now there are many (stages) where before there was one: in search of a pathophysiological classification of cirrhosis. Hepatology 2010;51:1445–1449.
14. Boursier J, Konate A, Guilluy M, et al. Learning curve and interobserver reproducibility evaluation of liver stiffness measurement by transient elastography. Eur J Gastroenterol Hepatol 2008;20:693–701.
15. 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;56:968–973.
16. Boursier J, Konate A, Gorea G, et al. Reproducibility of liver stiffness measurement by ultrasonographic elastometry. Clin Gastroenterol Hepatol 2008;6:1263–1269.
17. 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.
18. Poynard T, Ingiliz P, Elkrief L, et al. Concordance in a world without a gold standard: a new non-invasive methodology for improving accuracy of fibrosis markers. PLoS One 2008;3:e3857.

19. Roca B, Resino E, Torres V, et al. Interobserver discrepancy in liver fibrosis using transient elastography. J Viral Hepat 2012; 19:711–715.
20. Perazzo H, Fernandes FF, Gomes A, et al. Interobserver variability in transient elastography analysis of patients with chronic hepatitis C. Liver Int 2014 Mar 28. Epub ahead of print.
21. Lucidarme D, Foucher J, Le Bail B, et al. Factors of accuracy of transient elastography (fibroscan) for the diagnosis of liver fibrosis in chronic hepatitis C. Hepatology 2009;49:1083–1089.
22. Myers RP, Crotty P, Pomier-Layrargues G, et al. Prevalence, risk factors and causes of discordance in fibrosis staging by transient elastography and liver biopsy. Liver Int 2010;30: 1471–1480.
23. Sagir A, Erhardt A, Schmitt M, et al. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. Hepatology 2008;47:592–595.
24. Arena U, Vizzutti F, Corti G, et al. Acute viral hepatitis increases liver stiffness values measured by transient elastography. Hepatology 2008;47:380–384.
25. Arena U, Lupsor Platon M, Stasi C, et al. Liver stiffness is influenced by a standardized meal in patients with chronic hepatitis C virus at different stages of fibrotic evolution. Hepatology 2013;58:65–72.

Reprint requests
Address requests for reprints to: Vlad Ratziu, MD, Hôpital Pitié Salpêtrière, Service d'Hépatogastroentérologie, 47-83 Bd de l'Hopital, Paris 75013, France. e-mail: vlad.ratziu@upmc.fr; fax: 33-1-42-16-14-27.

Acknowledgments
The following are members of the LIDO (Liver Injury in Diabetes and Obesity) Study Group: André Grimaldi, Philippe Giral, Eric Bruckert, Arnaud Basdevant, Karine Clement, Jean-Michel Oppert, Agnès Hartemann-Heurtier, Fabrizio Andreelli, Sophie Gombert, Sophie Jacqueminet, Arnaud Cocaul, Fabienne Fouffelle, Joseph Moussalli, Dominique Thabut, Pascal Lebray, Philippe Podevin, Dominique Bonnefont-Rousselot, Randa Bittar, Yves Benhamou, Carole Bernhardt, Hôpital Pitié-Salpêtrière; Christian Boitard, Etienne Larger, Agnès Sola, Martine El-Etr, Hôpital Cochin; Jean-François Gautier, Hôpital Saint-Louis; Lawrence Serfaty, Chantal Housset, Jacqueline Capeau, Hôpital Saint-Antoine, all in Paris, France.

Conflicts of interest
The authors disclose no conflicts.

Funding
Supported in part from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. HEALTH-F2-2009-241762 for the project FLIP.

Supplementary Material

Variability According to Maximum Interquartile Range/Median Between Paired Examinations

Thirty-six paired examinations, 72.2% of which (26) were performed by senior operators, had both LSMs with IQR/M $\leq$10%. In this subgroup we found a discordance rate >20% in 33.3% (12), >30% in 22.2% (8), and >50% in 5.6% (2) of paired readings, resulting in 1 fibrosis class change in 25.0% (9) and $\geq$2 fibrosis classes change in 2.8% (1) of paired examinations. Two hundred eighty-seven paired examinations, 75.3% of which (216) were performed by senior operators, had a maximum IQR/M >10% and $\leq$20%. In this subgroup we found a discordance rate of >20% in 48.1% (138), >30% in 32.8% (94), and >50% in 13.6% (39) of paired readings, resulting in 1 fibrosis class change in 28.6% (82) and $\geq$2 fibrosis classes change in 10.5% (30) of paired readings. Two hundred eight paired examinations, 66.3% of which (138) were performed by senior operators, had a maximum IQR/M >20% and $\leq$30%. In this subgroup we found a discordance rate of >20% in 54.8% (114), >30% in 38.5% (80), and >50% in 11.5% (24) of paired readings, resulting in 1 fibrosis class change in 36.5% (76) and $\geq$2 fibrosis classes change in 10.1% (21) of paired examinations.

We found a trend toward an association between maximum IQR/M between paired examinations and operator experience (paired readings performed by senior operators 18.66 $\pm$ 5.69 vs at least 1 non-senior operator in paired readings 19.69 $\pm$ 5.97, $P = .065$).

Variability According to Time Interval Between Measurements

We performed a subanalysis among 374 paired readings (70.4% of all pairs) performed at least 6 months apart. There was a statistically significant difference regarding baseline fibrosis stage among readings performed less and more than 6 months apart; patients with a shorter follow-up had higher baseline LSM values (no/mild fibrosis 38.2% vs 54.8%, respectively; $P < .001$). Two hundred seventy-six of 374 pairs (73.8%) were performed by senior operators (vs 66.2% in paired readings with less than 6-month follow-up; $P = .078$), and 26.5% were performed by the same operator (vs 44.6% in paired TE with less than 6-month follow-up; $P < .001$). In this subgroup of paired TE (TE performed $\geq$6 months apart), the mean absolute variation was 2.9 $\pm$ 5.5 kPa (vs 3.9 $\pm$ 7.2; $P = .093$), and the mean relative variation was 30.0% $\pm$ 54.8% (vs 27.9% $\pm$ 25.4%; $P = .649$). In addition, discordance >20% was found in 51.3% of paired readings (vs 45.9%; $P = .249$), >30% in 33.4% (vs 36.3%; $P = .523$), and >50% in 9.4% (vs 19.1%; $P = .002$). One fibrosis class change occurred in 115 paired readings (30.7%) and $\geq$2 fibrosis classes change in 29 paired readings (7.8% vs 14.6%; $P = .015$). Hence, at least 1 fibrosis class change occurred in 38.5% of patients (vs 47.8%; $P = .048$).

Variability According to Alanine Aminotransferase Levels

With regard to ALT levels, 88 paired readings had ALT $\geq$100 IU/L, and only 22 had ALT $\geq$200 IU/L at baseline and/or follow-up TE. Among the 278 paired examinations with baseline and follow-up ALT <100 U/L, we found a discordance >20% in 48.2% of cases, >30% in 32.4%, and >50% in 10.8%, resulting in 1 fibrosis class change in 30.6% and $\geq$2 fibrosis classes change in 8.3%.

Determinants of Variability in Paired Examinations Without Significant Modifications of Body Mass Index

To rule out the effect of weight changes on LSM variability, we assessed the factors associated with variability among the 312 paired TE readings with stable BMI (less than 5% change in BMI) (Supplementary Table 4). Despite BMI stability, 47.8% of paired examinations showed >20% discordance, 31.4% >30%, and 10.6% >50% discordance. This resulted in 1 fibrosis class change in 32.4% paired readings and in $\geq$2 fibrosis classes change in 8.3%. Having 2 different operators was significantly associated with discordance rate of >20%, but not with fibrosis class changes. Age, sex, etiology of liver disease, IQR/M, and success rate were not associated with any outcomes. With regard to fibrosis stage at LSM1, cirrhosis ($\geq$12.5 kPa) was associated with relative variability, significant fibrosis ($\geq$7 kPa) with $\geq$1 fibrosis class change, and severe fibrosis ($\geq$9.5 kPa) with $\geq$2 classes change. Doubling of ALT and 10% change in SteatoTest also had an impact on fibrosis class changes. Multivariate analyses were not performed because of small sample size.

Determinants of Variability in Paired Examinations Performed by the Same Operators

To rule out the effect of interoperator variability on paired LSMs, we assessed the determinants of variability when performed by the same operators in 169 paired TE measurements; all but 4 of them were performed by senior operators (Supplementary Table 5). Discordance rates >20%, >30%, and >50% were still found in 39.6%, 27.2%, and 9.5% pairs, respectively, which resulted in fibrosis class change of 1 (26.0%) or $\geq$2 classes (8.3%). Maximum IQR/M was significantly associated with >20% and >30% discordance and with 1 or more fibrosis classes change. The time interval between TE measurements was not predictive of any outcomes, further confirming that it was a confounding factor related to the higher number of senior operators. Even when the same (senior) operator performed both TE readings, the fibrosis stage at LSM1

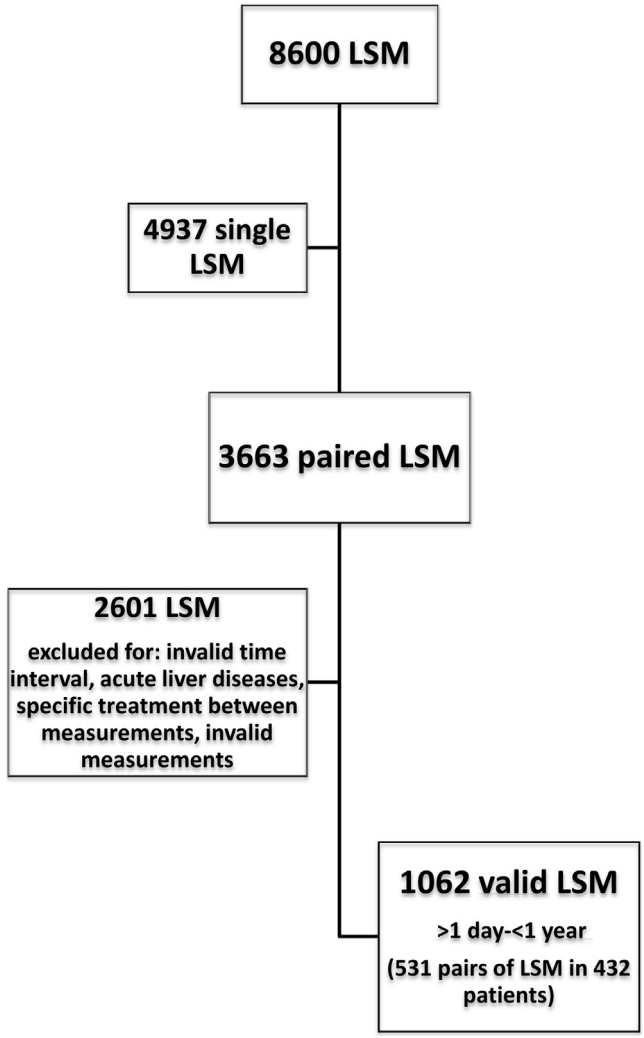
was associated with variability of TE readings; significant fibrosis (≥7 kPa) was associated with ≥1 fibrosis class change, severe fibrosis (≥9.5 kPa) with >20% discordance and with ≥2 fibrosis classes change, and cirrhosis (≥12.5 kPa) with >50% discordance. Absolute and relative BMI changes were associated with >20% discordance. Here again, doubling of ALT was associated with 20% and 30% discordance and with 1 or more fibrosis classes change. Multivariate analyses were not performed because of small sample size.

Fibrosis Class Changes According to Disease-Adjusted Liver Stiffness Measurement Cutoffs

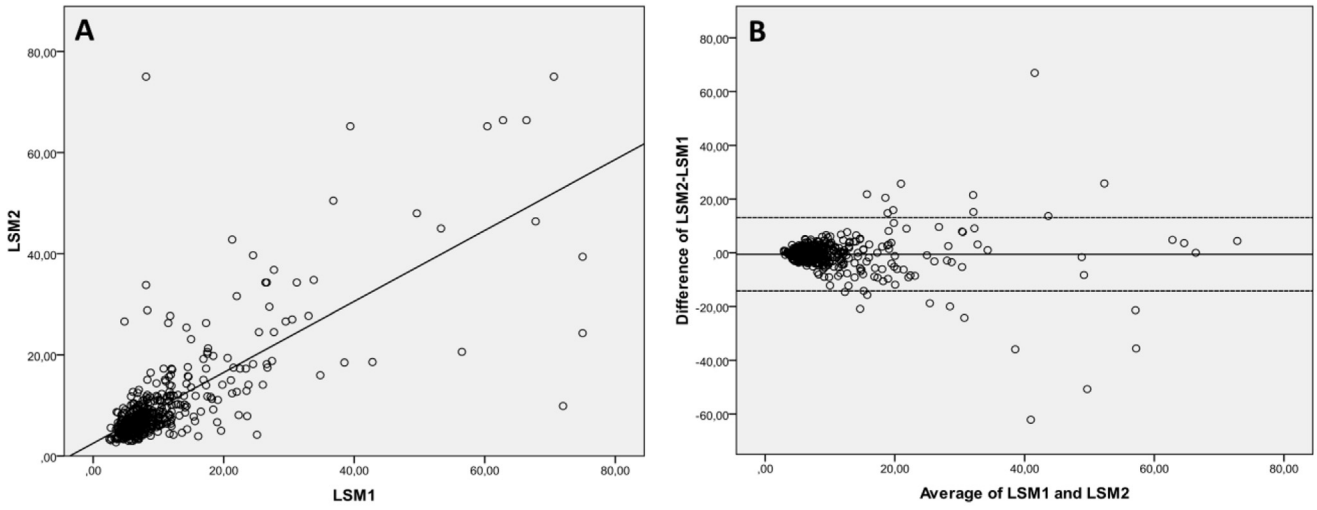
To confirm that LSM variability occurs in a substantial proportion of cases independent of the etiology of the liver disease and respective LSM cutoffs, we assessed the frequency of fibrosis class changes in paired readings from monoinfected HCV, HBV, and NAFLD patients by using etiology-specific LSM cutoffs.

Among 232 paired readings in patients with HCV, 27.2% had 1 fibrosis class change, and 10.8% had 2 or more classes change. Similarly, the 64 paired readings in HBV showed 1 fibrosis class change in 21.9% and ≥2 classes change in 14.1%. Finally, we found 1 fibrosis class change in 36.7% and ≥2 fibrosis classes change in 8.9% of 79 paired readings in NAFLD patients.

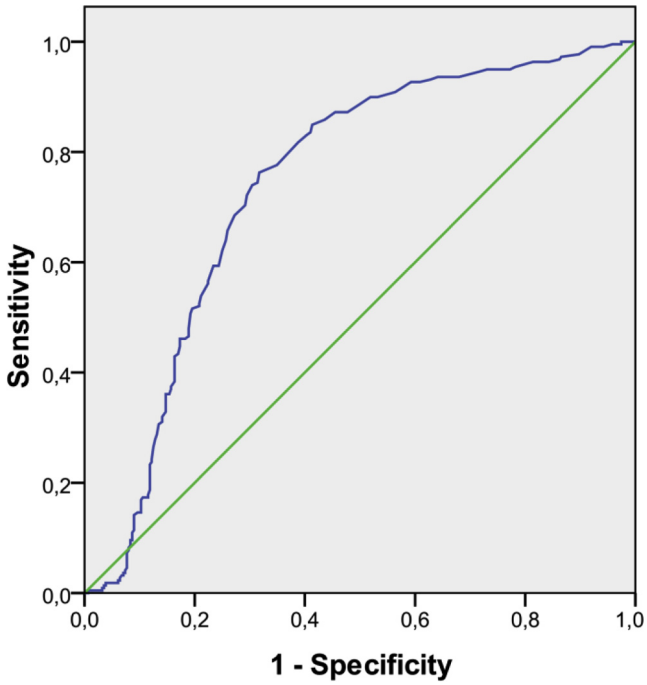
Overall, in these 375 paired readings reanalyzed according to etiology-specific cutoffs, discordance between no-mild fibrosis (F0-1 for HCV and HBV, F0-2 for NAFLD) and significant fibrosis (F2-4 for HCV and HBV, F3-4 for NAFLD) occurred in 83 cases (22.1%). Thirty-four paired readings (9.1%) showed discordance for the diagnosis of cirrhosis (F4).



Supplementary Figure 1. Flow chart representing the paired LSMs selection process.



Supplementary Figure 2. Correlation between baseline (LSM1) and follow-up (LSM2) paired LSMs. (*A*) Scatter plot representing correlation between baseline (LSM1) and follow-up (LSM2) paired LSMs. (*B*) Bland–Altman plot representing relationship between mean LSM for each pair of examinations [(LSM1+LSM2)/2] and difference between baseline and follow-up paired readings (LSM2-LSM1). *Solid line* is at mean of difference, and *dashed lines* are at ±1.96 SD.



Supplementary Figure 3. Receiver operating characteristic curve of baseline LSM (LSM1) and at least 1 fibrosis class change between paired LSMs. ROC curve representing predictive performance of baseline LSM (LSM1) for at least 1 fibrosis class change between paired LSMs. Area under the curve, 0.740; $P < .001$. Best cutoff point: LSM1 7.15 kPa, sensitivity 74%, and specificity 70%.

Supplementary Table 1. Variability in Stage of Fibrosis Between Baseline (LSM1) and Follow-up (LSM2) Paired LSMs

	LSM1					
LSM2	Absent/mild fibrosis	Significant fibrosis	Severe fibrosis	Early cirrhosis	Advanced cirrhosis	Total
Absent/mild fibrosis	213	44	16	6	1	280
Significant fibrosis	40	45	20	5	2	112
Severe fibrosis	9	17	17	13	2	58
Early cirrhosis	2	3	11	11	16	43
Advanced cirrhosis	1	3	2	6	26	38
Total	265	112	66	41	47	531

Supplementary Table 2. Factors Associated With Relative Variability in Univariate Analyses

	Non-discordant	Discordant	OR	95% CI	*P* value
Discordance >20%	n = 267	n = 264			
Different operators	165 (61.8)	197 (74.6)	1.82	1.25–2.64	.002
At least 1 non-senior operator	64 (24.0)	87 (33.0)	1.56	1.07–2.28	.022
Maximum IQR/M between LSMs	18.29 ± 5.75	19.63 ± 5.75	1.04	1.01–1.07	.008
Significant fibrosis at LSM1	113 (42.3)	153 (58.0)	1.88	1.33–2.65	<.001
Severe fibrosis at LSM1	53 (19.9)	101 (38.3)	2.50	1.69–3.70	<.001
Cirrhosis at LSM1	30 (11.2)	58 (22.0)	2.22	1.38–3.59	.001
Absolute BMI changes (kg/m^2)	0.67 ± 0.85	1.01 ± 1.33	1.35	1.10–1.65	.003
Relative BMI changes ≥7%	21 (10.3)	38 (18.3)	1.94	1.09–3.43	.022
Absolute WC changes (*cm*)	4.35 ± 4.42	5.35 ± 5.37	1.04	1.00–1.09	.044
Relative WC changes ≥10%	27 (13.7)	48 (25.3)	2.13	1.26–3.59	.004
Doubling of ALT	4 (2.3)	14 (8.3)	3.88	1.25–12.05	.015
Discordance >30%	n = 349	n = 182			
Different operators	226 (64.8)	136 (74.7)	1.61	1.08–2.40	.019
At least 1 non-senior operator	89 (25.5)	62 (34.1)	1.51	1.02–2.23	.038
Maximum IQR/M between LSMs	18.55 ± 5.72	19.73 ± 5.84	1.04	1.00–1.07	.026
Significant fibrosis at LSM1	156 (44.7)	110 (60.4)	1.89	1.31–2.72	.001
Severe fibrosis at LSM1	79 (22.6)	75 (41.2)	2.40	1.63–3.53	<.001
Cirrhosis at LSM1	43 (12.3)	45 (24.7)	2.34	1.47–3.72	<.001
Relative BMI changes ≥7%	32 (11.7)	27 (19.7)	1.86	1.06–3.25	.029
Doubling of ALT	7 (3.0)	11 (9.6)	3.37	1.27–8.94	.018
Discordance >50%	n = 466	n = 65			
At least 1 non-senior operator	123 (26.4)	28 (43.1)	2.11	1.24–3.59	.005
Time interval between LSMs (*mo*)	7.98 ± 3.26	6.62 ± 3.80	0.89	0.83–0.96	.002
Severe fibrosis at LSM1	126 (27.0)	28 (43.1)	2.04	1.20–3.48	.008
Cirrhosis at LSM1	67 (14.4)	21 (32.3)	2.84	1.59–5.08	<.001

NOTE. All data are expressed as frequency and percentage or as mean ± SD. Significant fibrosis was defined as ≥7 kPa, severe fibrosis as ≥9.5 kPa, and cirrhosis as ≥12.5 kPa.
LSM1, LSM at baseline; WC, waist circumference.

Supplementary Table 3. Factors Associated With Fibrosis Class Changes in Univariate Analyses

	Non-discordant	Discordant	OR	95% CI	*P* value
≥1 Fibrosis class change	n = 312	n = 219			
Different operators	201 (64.4)	161 (73.5)	1.53	1.05–2.24	.027
At least 1 non-senior operator	77 (24.7)	74 (33.8)	1.56	1.07–2.28	.022
Maximum IQR/M between LSMs	18.39 ± 5.63	19.76 ± 5.91	1.04	1.01–1.07	.008
Time interval between LSMs (*mo*)	8.11 ± 3.24	7.39 ± 3.48	0.94	0.89–0.99	.015
Significant fibrosis at LSM1	99 (31.7)	167 (76.3)	6.91	4.67–10.23	<.001
Severe fibrosis at LSM1	54 (17.3)	100 (45.7)	4.02	2.70–5.97	<.001
Cirrhosis at LSM1	37 (11.9)	51 (23.3)	2.26	1.42–3.59	<.001
Baseline BMI (kg/m^2)	23.98 ± 3.52	24.96 ± 3.70	1.08	1.02–1.14	.005
Doubling of ALT	5 (2.4)	13 (9.4)	4.15	1.44–11.91	.005
Relative SteatoTest changes ≥10%	58 (65.2)	37 (82.2)	2.47	1.03–5.96	.040
≥2 Fibrosis classes change	n = 479	n = 52			
At least 1 non-senior operator	124 (25.9)	27 (51.9)	3.09	1.73–5.53	<.001
Time interval between LSMs (*mo*)	7.94 ± 3.31	6.63 ± 3.57	0.90	0.83–0.97	.008
Significant fibrosis at LSM1	226 (47.2)	40 (76.9)	3.73	1.91–7.29	<.001
Severe fibrosis at LSM1	120 (25.1)	34 (65.4)	5.65	3.08–10.38	<.001
Cirrhosis at LSM1	72 (15.0)	16 (30.8)	2.51	1.33–4.77	.004

NOTE. All data are expressed as frequency and percentage or as mean ± SD. Significant fibrosis was defined as ≥7 kPa, severe fibrosis as ≥9.5 kPa, and cirrhosis as ≥12.5 kPa.
LSM1, LSM at baseline.

Supplementary Table 4. Factors Associated With Variability and Fibrosis Class Changes in Paired Examinations Without Significant Modifications of BMI

	Non-discordant	Discordant	OR	95% CI	*P* value
Discordance >20%	n = 163	n = 149			
Different operators	100 (61.3)	113 (75.8)	1.98	1.21–3.23	.006
At least 1 non-senior operator	31 (19.0)	41 (27.5)	1.62	0.95–2.75	.075
Significant fibrosis at LSM1	70 (42.9)	85 (57.0)	1.77	1.13–2.76	.013
Severe fibrosis at LSM1	29 (17.8)	58 (38.9)	2.95	1.75–4.95	<.001
Cirrhosis at LSM1	13 (8.0)	31 (20.8)	3.03	1.52–6.05	.001
Relative WC changes ≥10%	15 (10.1)	26 (19.7)	2.18	1.10–4.31	.024
Doubling of ALT	4 (3.4)	11 (11.2)	3.64	1.12–11.80	.031
Discordance >30%	n = 214	n = 98			
Significant fibrosis at LSM1	98 (45.8)	57 (58.2)	1.65	1.02–2.67	.043
Severe fibrosis at LSM1	48 (22.4)	39 (39.8)	2.29	1.36–3.83	.001
Cirrhosis at LSM1	22 (10.3)	22 (22.4)	2.53	1.32–4.83	.004
Doubling of ALT	6 (4.0)	9 (13.2)	3.64	1.24–10.67	.020
Discordance >50%	n = 279	n = 33			
Cirrhosis at LSM1	35 (12.5)	9 (27.3)	2.61	1.12–6.08	.032
≥1 Fibrosis classes change	n = 185	n = 127			
Time interval between LSMs (*mo*)	8.33 ± 2.92	7.17 ± 3.36	0.89	0.83–0.96	.002
Significant fibrosis at LSM1	59 (31.9)	96 (75.6)	6.61	3.97–11.01	<.001
Severe fibrosis at LSM1	28 (15.1)	59 (46.5)	4.87	2.86–8.28	<.001
Cirrhosis at LSM1	16 (8.6)	28 (22.0)	2.99	1.54–5.79	.001
Doubling of ALT	3 (2.3)	12 (14.3)	7.22	1.97–26.43	.001
Relative SteatoTest changes ≥10%	41 (61.2)	28 (82.4)	2.96	1.08–8.12	.041
≥2 Fibrosis classes change	n = 286	n = 26			
Significant fibrosis at LSM1	136 (47.6)	19 (73.1)	2.99	1.22–7.34	.013
Severe fibrosis at LSM1	71 (24.8)	16 (61.5)	4.85	2.10–11.16	<.001

NOTE. All data are expressed as frequency and percentage or as mean ± SD. Significant fibrosis was defined as ≥7 kPa, severe fibrosis as ≥9.5 kPa, and cirrhosis as ≥12.5 kPa.
LSM1, LSM at baseline; WC, waist circumference.

Supplementary Table 5. Factors Associated With Variability and Fibrosis Class Changes in Paired Examinations Performed by the Same Operator

	Non-discordant	Discordant	OR	95% CI	*P* value
Discordance >20%	n = 102	n = 67			
Maximum IQR/M between LSMs	17.20 ± 5.17	20.10 ± 6.12	1.10	1.04–1.16	.002
Severe fibrosis at LSM1	20 (19.6)	23 (34.3)	2.14	1.06–4.33	.032
Absolute BMI changes (kg/m^2)	0.55 ± 0.68	1.06 ± 1.65	1.53	1.04–2.25	.032
Relative BMI changes ≥7%	5 (6.6)	9 (17.6)	3.04	0.96–9.69	.051
Doubling of ALT	1 (1.5)	4 (11.1)	8.25	0.89–76.85	.049
Discordance >30%	n = 123	n = 46			
Maximum IQR/M between LSMs	17.61 ± 5.37	20.33 ± 6.23	1.09	1.02–1.16	.007
Doubling of ALT	1 (1.3)	4 (16.7)	15.60	1.65–147.37	.010
Discordance >50%	n = 153	n = 16			
Cirrhosis at LSM1	23 (15.0)	6 (37.5)	3.39	1.12–10.24	.035
≥1 Fibrosis classes change	n = 111	n = 58			
Maximum IQR/M between LSMs	17.32 ± 5.30	20.31 ± 6.05	1.10	1.04–1.17	.002
Significant fibrosis at LSM1	36 (32.4)	45 (77.6)	7.21	3.46–15.03	<.001
Severe fibrosis at LSM1	18 (16.2)	25 (43.1)	3.91	1.90–8.08	<.001
Cirrhosis at LSM1	14 (12.6)	15 (25.9)	2.42	1.07–5.44	.030
Doubling of ALT	0 (0.0)	5 (15.2)	—	—	.003
≥2 Fibrosis classes change	n = 155	n = 14			
Severe fibrosis at LSM1	35 (22.6)	8 (57.1)	4.57	1.49–14.06	.009

NOTE. All data are expressed as frequency and percentage or as mean ± SD. Significant fibrosis was defined as ≥7 kPa, severe fibrosis as ≥9.5 kPa, and cirrhosis as ≥12.5 kPa.
LSM1, liver stiffness measurement at baseline.