Diagnosis of Fibrosis and Cirrhosis Using Liver Stiffness Measurement in Nonalcoholic Fatty Liver Disease

Vincent Wai-Sun Wong,[1,2] Julien Vergniol,[3] Grace Lai-Hung Wong,[1,2] Juliette Foucher,[3] Henry Lik-Yuen Chan,[1,2] Brigitte Le Bail,[4,5] Paul Cheung-Lung Choi,[6] Mathurin Kowo,[3] Anthony Wing-Hung Chan,[6] Wassil Merrouche,[3] Joseph Jao-Yiu Sung,[1,2] and Victor de Lédinghen[3,4]

Nonalcoholic fatty liver disease (NAFLD) is one of the most common liver diseases in affluent countries. Accurate noninvasive tests for liver injury are urgently needed. The aim of this study was to evaluate the accuracy of transient elastography for the diagnosis of fibrosis and cirrhosis in patients with NAFLD and to study factors associated with discordance between transient elastography and histology. Two hundred forty-six consecutive patients from two ethnic groups had successful liver stiffness measurement and satisfactory liver biopsy specimens. The area under the receiver-operating characteristics curve (AUROC) of transient elastography for F3 or higher and F4 disease was 0.93 and 0.95, respectively, and was significantly higher than that of the aspartate aminotransferase–to–alanine aminotransferase ratio, aspartate aminotransferase–to–platelet ratio index, FIB-4, BARD, and NAFLD fibrosis scores (AUROC ranged from 0.62 to 0.81, $P < 0.05$ for all comparisons). At a cutoff value of 7.9 kPa, the sensitivity, specificity, and positive and negative predictive values for F3 or greater disease were 91%, 75%, 52%, and 97%, respectively. Liver stiffness was not affected by hepatic steatosis, necroinflammation, or body mass index. Discordance of at least two stages between transient elastography and histology was observed in 33 (13.4%) patients. By multivariate analysis, liver biopsy length less than 20 mm and F0-2 disease were associated with discordance. *Conclusion:* Transient elastography is accurate in most NAFLD patients. Unsatisfactory liver biopsy specimens rather than transient elastography technique account for most cases of discordance. With high negative predictive value and modest positive predictive value, transient elastography is useful as a screening test to exclude advanced fibrosis. Liver biopsy may be considered in NAFLD patients with liver stiffness of at least 7.9 kPa. (HEPATOLOGY 2010;51:454-462.)

Nonalcoholic fatty liver disease (NAFLD) is one of the most common chronic liver diseases worldwide.[1] It is strongly associated with metabolic syndrome and obesity,[2,3] and may progress to cirrhosis and hepatocellular carcinoma.[4,5] The prog-nosis depends heavily on histological severity. Although patients with simple steatosis have excellent prognosis, those with nonalcoholic steatohepatitis tend to progress and have hepatic complications.[6]

Traditionally, liver biopsy is the gold standard for the assessment of hepatic necroinflammation and fibrosis. However, the procedure carries a small risk of complica-

tions and may not be acceptable to some patients. Because a standard liver biopsy sample only represents approximately 1/50,000 of the whole liver mass, sampling bias may occur. When both lobes of the livers underwent biopsy during bariatric surgery, fibrosis stage was discordant between the two samples in half of the cases.[7] Noninvasive tests for NAFLD are urgently needed.[8,9]

Transient elastography by Fibroscan is a noninvasive method for the diagnosis of liver fibrosis. It has high degree of accuracy and reproducibility in predicting bridging fibrosis and cirrhosis in patients with viral hepatitis.[10-13] Nevertheless, NAFLD patients are underrepresented in previous validation studies. Whether factors other than fibrosis, such as hepatic steatosis and prehepatic fat, may affect liver stiffness is uncertain. Factors associated with inaccurate measurements have not been evaluated.

In this study, we aimed to evaluate the accuracy of transient elastography and biochemical tests for the diagnosis of fibrosis and cirrhosis in a large cohort of NAFLD patients, and to test whether liver stiffness is altered by hepatic steatosis, inflammation, and obesity. We also aimed to identify factors associated with discordance between liver stiffness measurements (LSM) and histology.

Patients and Methods

Consecutive patients with NAFLD undergoing liver biopsies at the University Hospital of Pessac, France, and Prince of Wales Hospital, Hong Kong, were prospectively recruited. We included patients age 18 years or older. Men who consumed more than 30 g alcohol per day and women who consumed more than 20 g alcohol per day were excluded. Patients with secondary causes of hepatic steatosis (such as chronic use of systemic corticosteroids), positive hepatitis B surface antigen, anti-hepatitis C virus antibody, or histological evidence of other concomitant chronic liver diseases were also excluded. Because the aim of transient elastography was to diagnose significant fibrosis and early cirrhosis, patients with clinical and radiological evidence of cirrhosis were excluded (for example, bilirubin $\geq$30 μmol/L, albumin <35 g/L, international normalized ratio >1.3, platelet count <150 $\times$ 10^9/L, ascites, varices, splenomegaly). All patients gave informed written consent.

Clinical Assessment. Comprehensive clinical assessment was performed. Co-morbid illness and drug/herb intake was recorded with a standard questionnaire. Anthropometric tests included body weight, body height, and waist circumference measurements. Body mass index (BMI) was calculated as weight (kg) divided by height (m) squared. Waist circumference was measured at a level midway between the lower rib margin and iliac crest with

the tape all around the body in the horizontal position. On the day of liver biopsy, a fasting venous blood sample was taken for albumin, bilirubin, alanine aminotransferase (ALT), glucose, total cholesterol, and triglycerides.

In patients with complete biochemical data, the performance of transient elastography was compared with that of other prediction scores. The aspartate aminotransferase (AST)-to-platelet ratio index was calculated as AST (/upper limit of normal)/platelet count ($\times$10^9/L) $\times$ 100.[14] FIB-4 was calculated as age $\times$ AST (U/L)/platelet count ($\times$10^9/L) $\times$ $\sqrt{}$ $\sqrt{\text{ALT}}$ (U/L).[15] Cutoff values for NAFLD patients were adopted.[16] The NAFLD fibrosis score was calculated according to the following formula: $-1.675 + 0.037 \times$ age (years) $+ 0.094 \times$ BMI (kg/m^2) $+ 1.13 \times$ impaired fasting glycemia (IFG)/diabetes (yes = 1, no = 0) $+ 0.99 \times$ AST/ALT ratio $- 0.013 \times$ platelet ($\times$10^9/L) $- 0.66 \times$ albumin (g/dL).[17] The BARD score was the weighted sum of three variables (BMI $\geq$ 28 = 1 point, AST/ALT ratio $\geq$ 0.8 = 2 points, diabetes = 1 point).[18]

Histological Assessment. In this study, liver histology serves as the gold standard for evaluating the diagnostic accuracy of transient elastography. Percutaneous liver biopsy was performed using the 16G Temno or Menghini needle. Liver biopsy specimens were fixed in formalin and embedded in paraffin. Liver histology was assessed by experienced histopathologists (B.L.B., P.C.C.) who were blinded to the clinical data. Liver specimens shorter than 15 mm were excluded. Histological scoring was performed according to the system reported by Kleiner et al.[19] Grade of steatosis was defined according to Kleiner et al.: 0 = steatosis < 5%, 1 = steatosis 5% to 33%, 2 = steatosis > 33% − 66%, 3 = steatosis > 66%. Fibrosis was staged from 0 to 4: stage 0 = absence of fibrosis; stage 1 = perisinusoidal or portal; stage 2 = perisinusoidal and portal/periportal; stage 3 = septal or bridging fibrosis; and stage 4 = cirrhosis.

Transient Elastography. LSM was performed within 1 week before liver biopsy by using transient elastography according to the instructions and training provided by the manufacturer. Measurements were performed on the right lobe of the liver through intercostal spaces with the patient lying in dorsal decubitus with the right arm in maximal abduction. Ten successful acquisitions were performed on each patient. The median value represented the liver elastic modulus. Only cases with 10 successful acquisitions were evaluated. The liver stiffness was expressed in kiloPascal (kPa). The success rate was calculated as the number of successful measurements divided by the total number of measurements. The operators were blinded to all clinical data and the diagnoses of the patients.

Statistical Analysis. Statistical tests were performed using the Statistical Package for Social Sciences version 16.0. Continuous variables were expressed as mean ± standard deviation or median (interquartile range [IQR]) as appropriate. Receiver-operating characteristics curves were constructed to assess the overall accuracy of LSM and to identify optimal cutoffs. The optimal cutoffs of LSM for F2, F3, and F4 disease were chosen at points with the highest Youden's index. The relationship between steatosis, NAFLD activity score, BMI, and LSMs was adjusted by fibrosis stage in a multiple linear regression model. Significant discordance between transient elastography and histology was defined as a difference in fibrosis stage by 2 points or more. In the assessment of discordance, both cutoff values identified in this study and those reported by Yoneda et al.[20] were used. Quantitative variables between groups were compared by unpaired t test, Mann-Whitney U test, and one-way analysis of variance followed by Bonferroni test. Categorical variables were compared by chi-squared test or Fisher's exact test. The area under the receiver operating characteristics curves of different noninvasive tests was compared by the Delong test. All statistical tests were two-sided. Significance was taken as $P < 0.05$.

Results

From May 2003 to April 2009, 309 consecutive patients with NAFLD underwent transient elastography and liver biopsies. A total of 35 patients were excluded because of liver biopsy length less than 15 mm. Twenty-eight (10.2%) patients were excluded because of failure to obtain 10 valid LSM acquisitions. Two hundred forty-six patients with valid LSM acquisitions and satisfactory liver biopsy specimens were included in the analysis. Patients who failed LSM acquisitions had higher BMI (35.6 ± 6.3 versus 28.0 ± 4.5 kg/m², $P < 0.001$) and waist circumference (114 ± 14 versus 94 ± 12 cm, $P < 0.001$). Valid LSM acquisitions were obtained in 62 of 63 (98.4%) patients with BMI less than 25 kg/m², 114 of 117 (97.4%) patients with BMI 25 to 30 kg/m², and 70 of 94 (74.5%) patients with BMI of 30 kg/m² or higher. The rate of successful acquisitions at the same BMI was similar in whites and Chinese. Thirty-one (12.6%) and 25 (10.2%) patients had advanced fibrosis and cirrhosis, respectively (Table 1).

Accuracy of Transient Elastography. The LSMs of patients with F0, F1, F2, F3, and F4 disease were 5.7 ± 1.8, 6.8 ± 2.4, 7.8 ± 2.4, 11.8 ± 5.2, and 25.1 ± 17.1 kPa, respectively ($P < 0.0001$ by analysis of variance). Patients with F3 and F4 disease had significantly higher LSM than those with less fibrosis (Fig. 1). Overall, the accuracy of transient elastography to detect F2 or higher,

F3 or higher, and F4 disease was good, with areas under the receiver operating curve (AUROCs) of 0.84, 0.93, and 0.95, respectively (Table 2). The corresponding AUROCs were 0.87, 0.94, and 0.94, respectively, in the French cohort, and 0.84, 0.92, and 0.97, respectively, in the Chinese cohort.

The best LSM cutoff for F2 or greater disease was 7.0 kPa (Table 2). The negative predictive value to exclude F2 or greater disease was 84% (95% confidence interval [CI], 78%-90%). Cutoff values of 5.8 kPa and 9.0 kPa had greater than 90% sensitivity and specificity to rule out and rule in F2 disease, respectively.

The best cutoff for F3 or greater disease was 8.7 kPa (Table 2). The negative predictive value to exclude F3 or greater disease was 95% (95% CI, 91%-98%). Cutoff values of 7.9 and 9.6 kPa had greater than 90% sensitivity and specificity to rule out and rule in F3 disease, respectively.

The best cutoff for F4 disease was 10.3 kPa (Table 2). The negative predictive value to exclude cirrhosis was 99% (95% CI, 98%-100%). The same cutoff value also had greater than 90% sensitivity to rule out cirrhosis. A cutoff value of 11.5 kPa had greater than 90% specificity to detect cirrhosis.

Factors Affecting LSM. Steatosis grade ($P = 0.31$), NAFLD activity score ($P = 0.31$), serum ALT ($P = 0.39$), and BMI ($P = 0.29$) did not influence LSM after adjusting for fibrosis stage (Fig. 2). Similarly, whites and Chinese had similar LSMs at the same fibrosis stage ($P = 0.22$).

Prevalence of Discordance and Risk Factors. Discordance of at least two stages between transient elastography and histology was observed in 33 (13.4%) patients according to the cutoffs derived in this study. Transient elastography predicted a higher fibrosis stage in 30 cases and a lower fibrosis stage in three cases. Using cutoffs reported by Yoneda et al.,[20] discordance was also observed in 33 (13.4%) patients. Transient elastography predicted a higher fibrosis stage in 23 cases and a lower fibrosis stage in 10 cases. Discordances occurred in seven (10.0%) patients with BMI of at least 30 kg/m² and 26 (14.8%) patients with BMI less than 30 kg/m². Thus, among 94 obese patients, 24 failed to have valid LSM acquisitions, and seven had discordant results. The overall success rate of transient elastography (valid measurements plus correct classification) in patients with BMI of at least 30 kg/m² was 67.0%. Among 37 patients with BMI of 35 kg/m² or higher, 15 (41%) failed LSM acquisition, and two (5%) had discordance between LSM and histology. To avoid overfitting, Yoneda's cutoffs were used to identify risk factors for discordance.

Table 1. Clinical Characteristics of NAFLD Patients With and Without Discordant Results

Variable	All (N = 246)	French Cohort (N = 128)	Chinese Cohort (N = 118)	No Discordance (N = 213)	Discordance (N = 33)	P*
Age (years)	51 ± 11	53 ± 13	49 ± 9	52 ± 11	47 ± 13	0.028
Male sex	135 (54.9%)	74 (57.8%)	61 (51.7%)	117 (54.9%)	18 (54.5%)	0.97
Ethnicity						0.007
Caucasian	128 (52.0%)	128 (100%)	–	118 (55.4%)	10 (30.3%)	
Chinese	118 (48.0%)	–	118 (100%)	95 (44.6%)	23 (69.7%)	
Body mass index (kg/m²)	28.0 ± 4.5	29.1 ± 5.1	26.9 ± 3.4	28.1 ± 4.7	27.7 ± 3.5	0.64
<25 kg/m²	63 (25.6%)	29 (22.7%)	34 (28.8%)	58 (27.2%)	5 (15.2%)	0.085
25-30 kg/m²	113 (45.9%)	52 (40.6%)	61 (51.7%)	92 (43.2%)	21 (63.6%)	
≥30 kg/m²	70 (28.5%)	47 (36.7%)	23 (19.5%)	63 (29.6%)	7 (21.2%)	
Waist circumference (cm)	94 ± 12	99 ± 13	90 ± 10	94 ± 13	93 ± 9	0.55
Fasting glucose (mmol/L)	6.4 ± 2.6	6.5 ± 3.1	6.3 ± 2.3	6.5 ± 2.7	6.1 ± 2.2	0.52
Total cholesterol (mmol/L)	5.4 ± 1.3	5.5 ± 1.4	5.3 ± 1.1	5.3 ± 1.3	5.6 ± 0.9	0.18
Triglycerides (mmol/L)	2.2 ± 3.2	1.9 ± 1.3	2.4 ± 3.8	2.3 ± 3.5	2.1 ± 1.7	0.77
ALT (IU/L)	75 ± 54	82 ± 64	67 ± 39	75 ± 55	75 ± 47	0.96
Diabetes	89 (36.2%)	37 (28.9%)	52 (44.1%)	77 (37.4%)	12 (36.4%)	0.91
Hypertension	99 (40.2%)	50 (39.1%)	49 (41.5%)	87 (45.1%)	12 (36.4%)	0.35
Steatosis grade (1/2/3)	78/87/81	36/42/50	42/45/31	65/80/68	13/7/13	0.19
Fibrosis stage (0/1/2/3/4)	70/75/45/31/25	14/42/36/21/15	56/33/9/10/10	55/67/36/31/24	15/8/9/0/1	0.013
NAFLD activity score	4 (3-5)	4 (3-5)	3 (2-4)	4 (3-5)	3 (2-4)	0.073
Liver biopsy length (mm)	21 ± 7	24 ± 8	18 ± 3	22 ± 7	19 ± 4	0.002
Success rate (%)†	87 ± 16	88 ± 17	87 ± 14	87 ± 16	88 ± 15	0.62
Median LSM (kPa)	9.2 ± 8.2	9.7 ± 9.7	8.6 ± 6.4	9.3 ± 8.8	8.2 ± 2.8	0.48
IQR (kPa)	1.7 ± 1.9	1.6 ± 1.6	1.7 ± 2.1	1.7 ± 2.0	1.6 ± 1.1	0.72
IQR/median ratio	0.17 ± 0.09	0.17 ± 0.09	0.18 ± 0.09	0.17 ± 0.09	0.18 ± 0.09	0.73

*Comparison between patients with and without discordance.

†Success rate = Number of valid LSM acquisitions divided by the total number of measurements.

ALT, alanine aminotransferase; HDL, high-density lipoprotein; IQR, interquartile range; LDL, low-density lipoprotein; LSM, liver stiffness measurement.

By univariate analysis, younger age, Chinese ethnicity, lower fibrosis stage, and shorter liver biopsy lengths were associated with discordance (Table 1). Discordance occurred in 25 of 144 (17.4%) patients with liver biopsy lengths smaller than 20 mm, compared with 8 of 102 (7.8%) patients with liver biopsy lengths 20 mm or greater ($P = 0.031$). Similarly, discordance occurred in 32 of 190 (16.8%) patients with F0F1F2 disease, but only 1 of 56 (1.8%) patient with F3 or higher disease ($P = 0.002$). Conversely, performance indices of transient elastography were not associated with discordance. Discordance occurred in 30 of 231 (13.0%) patients with at least 60% valid LSM acquisitions out of all measurements and 3 of 15 (20.0%) patients with valid LSM acquisitions below 60% of all measurements ($P = 0.43$). Discordance occurred in 3 of 25 (12.0%) patients with IQR/LSM ratio above 0.3 and 30 of 221 (13.6%) patients with ratio below 0.3 ($P = 1.0$). By multivariate analysis, only liver biopsy length less than 20 mm (odds ratio, 2.7; 95% CI, 1.1-6.3; $P = 0.024$) and F3 or greater disease (odds ratio, 0.084; 95% CI, 0.011-0.63; $P = 0.016$) remained independent factors associated with discordance.

Comparison Between Transient Elastography and Other Biochemical Tests. As shown in Table 3, AU-ROC of transient elastography was significantly higher than that of AST/ALT ratio, AST-to-platelet ratio index, FIB-4, NAFLD fibrosis score, and BARD score in the diagnosis of both advanced fibrosis and cirrhosis. Among the biochemical tests, the FIB-4 index was superior to AST/ALT ratio ($P = 0.0008$), AST-to-platelet ratio index ($P = 0.017$), and BARD score ($P = 0.021$) for the detection of F3 or greater disease, and superior to AST/ALT ratio ($P = 0.0061$) and BARD score ($P = 0.0031$) for the detection of cirrhosis.

In an 'intention-to-treat' analysis, all 274 patients who underwent transient elastography and liver biopsy were analyzed, and the 28 patients in whom LSM could not be obtained were considered as not correctly classified. At the cutoff of 8.7 kPa, the negative predictive value of transient elastography in excluding F3 or greater disease remained high at 89.3% (Table 4). However, the positive predictive value was modest at 48.5%. Although the combined sensitivities and specificities of the biochemical tests were generally lower than those of transient elastography, they
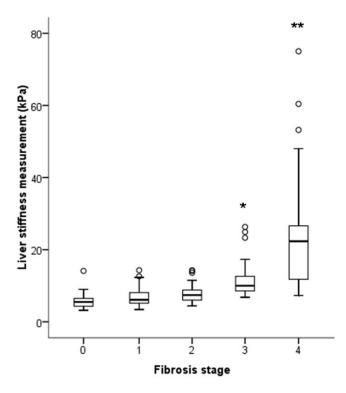
Fig. 1. Boxplot of liver stiffness measurements according to fibrosis stage. The bottom and top of each box represent the 25th and 75th percentiles, giving the interquartile range. The line through the box indicates the median, and the error bars indicate the 10th and 90th percentiles. Note: *$P < 0.01$ compared with F0, F1, and F4, $P < 0.05$ compared with F2; **$P < 0.001$ compared with all other fibrosis stages (comparison with Bonferroni correction).

also had negative predictive values over 80% at the low cutoff values (Table 4).

Clinical Use of Transient Elastography. Two strategies using transient elastography to diagnose and exclude advanced fibrosis in NAFLD patients were evaluated: (1) use a single cutoff with the best overall sensitivity and specificity; (2) use two cutoffs with high sensitivity and specificity.

If liver biopsies were reserved for patients with LSM of at least 8.7 kPa, only 79 (32.1%) would require the pro-

cedure. Eight (4.8%) patients with F3 disease and one (0.6%) patient with cirrhosis would be missed (Table 2).

When two cutoffs were used, 148 (60.2%) patients had LSM below the low cutoff of 7.9 kPa, and the negative predictive value was 96.6% (95% CI, 93.7%-99.5%) (Table 2, Fig. 3). Fifty-eight (23.6%) patients had LSM above the high cutoff of 9.6 kPa, and the positive predictive value was 72.4% (95% CI, 60.9%-83.9%). If liver biopsies were performed only in patients with LSMs between 7.9 and 9.6 kPa, 40 (16.3%) patients required the procedure. Five (3.4%) patients with F3 disease would be missed, and 16 (27.6%) patients without advanced fibrosis would be misclassified. Alternatively, if liver biopsies were performed in all patients with LSM above 7.9 kPa, 98 (39.8%) would require the procedure.

Discussion

In this large prospective cohort of NAFLD patients, transient elastography had high accuracy in detecting advanced fibrosis and cirrhosis. Successful measurement could be obtained in more than 97% of patients with BMI below 30 kg/m^2 and 75% of obese patients. LSM was not affected by hepatic steatosis, necroinflammation, or obesity. Most discordance between transient elastography and histology occurred in patients with short liver biopsy lengths and mild or no fibrosis. In addition, transient elastography had superior performance to other noninvasive biochemical tests in diagnosing advanced fibrosis and cirrhosis.

Transient elastography has been validated in chronic viral hepatitis, cholestatic liver disease, and alcoholic liver disease.[11,21-24] According to a meta-analysis of nine studies, the pooled estimates of sensitivity and specificity are 87% and 91%, respectively, for cirrhosis and 70% and 84%, respectively, for F2 or higher disease.[10] In a study of 97 Japanese patients with NAFLD, transient elastography had good overall accuracy, with AUROCs of 0.87, 0.90, and 0.99 for F2 or greater, F3 or greater, and F4 diseases,

Table 2. Accuracy of Transient Elastography

Stage	AUROC	Cutoff (kPa)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR−
≥F2	0.84 (0.79–0.90)	5.8	91.1	50.3	56.1	89.0	1.8	0.18
		7.0	79.2	75.9	69.6	84.0	3.3	0.27
		9.0	52.5	91.7	81.5	73.5	6.3	0.52
≥F3	0.93 (0.89–0.96)	7.9	91.1	75.3	52.0	96.6	3.7	0.12
		8.7	83.9	83.2	59.5	94.6	5.0	0.19
		9.6	75.0	91.6	72.4	92.6	8.9	0.27
F4	0.95 (0.91–0.99)	10.3	92.0	87.8	46.0	99.0	7.5	0.091
		10.3	92.0	87.8	46.0	99.0	7.5	0.091
		11.5	76.0	91.0	48.7	97.1	8.4	0.26

For each fibrosis stage, cutoffs with sensitivity >90%, highest overall accuracy and specificity >90% were presented.

AUROC, area under the receiver-operating characteristics curve; LR, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.
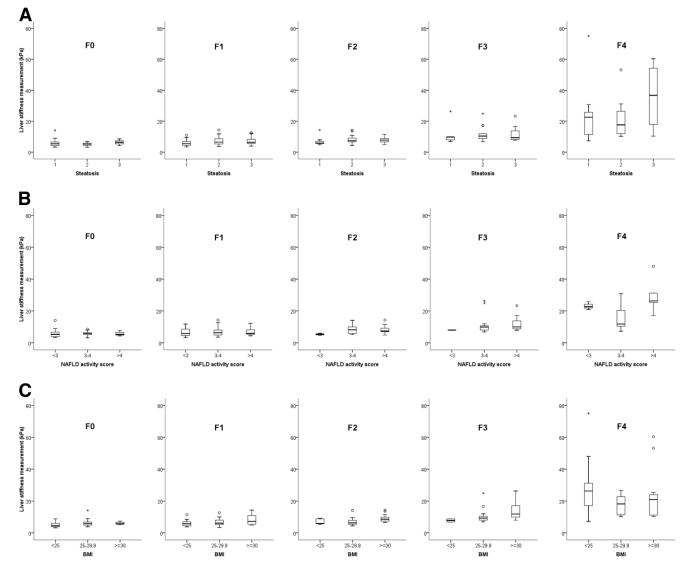
Fig. 2. Boxplot of liver stiffness measurements in relation to the (A) degree of steatosis, (B) NAFLD activity score, and (C) BMI. One graph for each fibrosis stage is provided.

respectively.[20] Similar levels of accuracy and reproducibility were observed in 50 pediatric NAFLD patients in Italy.[25] However, in this study, only 10 patients had F3 or higher fibrosis. Our study confirms that transient elastog-

raphy works well in both white and Asian NAFLD patients. At cutoff values of 8.7 and 10.3 kPa, the negative predictive values to exclude F3 or greater disease and cirrhosis were 95% and 99%, respectively. The adoption of

Table 3. Area Under the Receiver Operating Characteristics Curves for Transient Elastography and Noninvasive Markers for the Diagnosis of Advanced Fibrosis and Cirrhosis

		≥F3			F4		
	N	AUROC (95% CI)	AUROC for Fibroscan (95% CI)	P	AUROC (95% CI)	AUROC for Fibroscan (95% CI)	P
AST/ALT	245	0.66 (0.58–0.74)	0.93 (0.89–0.96)	<0.0001	0.66 (0.55–0.77)	0.95 (0.91–0.99)	<0.0001
APRI	245	0.74 (0.67–0.82)	0.93 (0.89–0.96)	<0.0001	0.75 (0.64–0.85)	0.95 (0.91–0.99)	<0.0001
FIB-4	245	0.80 (0.74–0.87)	0.93 (0.89–0.96)	0.0004	0.81 (0.73–0.89)	0.95 (0.91–0.99)	0.001
NAFLD fibrosis score	228	0.75 (0.67–0.83)	0.92 (0.89–0.96)	<0.0001	0.80 (0.69–0.92)	0.95 (0.91–0.99)	0.022
BARD score	244	0.69 (0.61–0.77)	0.93 (0.89–0.96)	<0.0001	0.62 (0.50–0.75)	0.95 (0.91–0.99)	<0.0001

P values refer to the comparison between transient elastography and other noninvasive tests.

ALT, alanine aminotransferase; APRI, AST-to-platelet ratio index; AST, aspartate aminotransferase; AUROC, area under the receiver-operating characteristics curve; CI, confidence interval.

Table 4. Comparative Performance of Noninvasive Tests for the Diagnosis of Advanced Fibrosis in 274 NAFLD Patients

Noninvasive tests	Cutoffs	Se (%)	Sp (%)	PPV (%)	NPV (%)	+LR	−LR
Transient elastography							
(kPa)*	7.9	77.3	68.8	44.0	90.5	2.5	0.33
(failure n = 28)	8.7	71.2	76.0	48.5	89.3	3.0	0.38
	9.6	63.6	83.7	55.3	87.9	3.9	0.43
AST/ALT	0.8	39.7	79.7	37.9	80.9	2.0	0.76
	1.0	20.6	90.1	39.4	78.4	2.1	0.88
APRI	0.5	65.1	72.3	42.3	86.9	2.4	0.48
	1.5	6.3	97.0	40.0	76.9	2.1	0.97
FIB-4	1.30	65.1	80.2	50.6	88.0	3.3	0.44
	2.67	20.6	95.5	59.1	79.4	4.6	0.83
NAFLD fibrosis score	−1.455	73.3	69.5	43.6	89.0	2.4	0.38
	0.676	18.3	96.3	61.1	78.6	4.9	0.85
BARD score	2	61.9	65.8	36.1	84.7	1.8	0.58

*In this "intention to treat" analysis, patients in whom ten valid liver stiffness measurements could not be acquired were included and counted as incorrect classification by transient elastography.

ALT, alanine aminotransferase; APRI, AST-to-platelet ratio index; AST, aspartate aminotransferase; LR, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; Se, sensitivity; Sp, specificity.

transient elastography could potentially spare two thirds of NAFLD patients from liver biopsies. Because the prevalence of NAFLD is high in many affluent countries, this approach would be cost saving.

Conversely, the positive predictive value of transient elastography and other noninvasive tests to diagnose advanced fibrosis in NAFLD patients remains modest. Therefore, the main value of these tests is to exclude advanced fibrosis as screening tests. Based on our data, it is reasonable to consider liver biopsy in patients whose LSM is 7.9 kPa or above. When transient elastography is not available, the biochemical tests reported in this study are reasonable screening tests despite a lower overall accuracy.

LSM has been shown to be spuriously increased in patients with acute hepatitis and extrahepatic cholestasis, indicating that the stiffness of the liver is not attributable to fibrosis alone.[26-29] One unique feature of NAFLD patients is the accumulation of subcutaneous, prehepatic, and hepatic fat. Whether this would affect LSM has major clinical implications. In patients with chronic hepatitis C, hepatic steatosis does not appear to influence LSM, although patients with severe steatosis were underrepre-

sented.[26] In this study, we clearly showed that hepatic steatosis did not increase LSM in NAFLD subjects. Although subcutaneous and prehepatic fat thickness was not measured, patients with high BMI also did not have increased LSM after adjusting for fibrosis stage. Moreover, ALT level and the NAFLD activity score did not influence LSM. This is likely because severe necroinflammation is rare in NAFLD subjects, and a milder degree of necroinflammation has no major impact on LSM. Besides, ALT level in NAFLD subjects mainly reflects the degree of hepatic steatosis and correlates poorly with necroinflammation.[30]

In patients with chronic hepatitis C, discordance between transient elastography and histology occurs more commonly if the IQR/LSM ratio is high.[31] In our study, discordance occurred mainly in patients with shorter liver biopsy lengths and lower fibrosis stages. Both factors indicate that the discordance was attributable to understaging by histology as a result of sampling bias. One possible explanation of the phenomenon is that the distribution of fibrous tissue may be less even in NAFLD patients. In a study of 41 subjects undergoing right-lobe and left-lobe
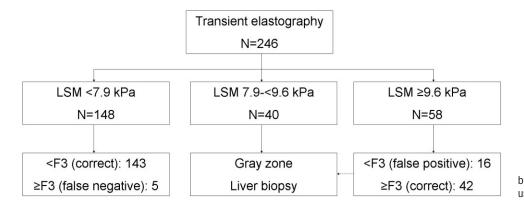
Fig. 3. Prediction of advanced fibrosis in patients with NAFLD, by using transient elastography.

liver biopsies during bariatric surgery, the kappa coefficient for fibrosis staging was only 0.53.[7] Although we cannot recommend relying on transient elastography regardless of the IQR/LSM ratio because most of our patients had IQR/LSM ratio less than 0.3 at inclusion, our study serves as a reminder that when a noninvasive test disagrees with histological results, the latter may be inaccurate. By mathematical modeling, the AUROC of a noninvasive test is limited by the biopsy sensitivity and specificity even if the test has perfect accuracy.[32]

Our study has several limitations. First, liver biopsy was used as the gold standard, and liver biopsy specimens were assessed by two pathologists. Sampling bias could not be excluded. However, liver biopsy is currently the only reference standard, and biopsy specimens were assessed by experienced pathologists. Second, patients recruited at referral centers likely had more advanced disease. However, the negative predictive values to exclude advanced fibrosis and cirrhosis would be even higher in the primary care setting. The inclusion of both whites and Chinese further increases the external validity of this study. Third, a significant proportion of obese subjects were not analyzed because of failed LSM. The problem may be solved in the future with the development of probes for obese subjects. In a study of 84 obese subjects, at least five measurements could be acquired in over 90% by using the new obese probe, compared with less than 80% by using the standard probe.[33]

In conclusion, transient elastography can be performed in most NAFLD patients and is accurate. The measurement and accuracy are not affected by hepatic steatosis, necroinflammation, and obesity. Unsatisfactory liver biopsy specimens rather than transient elastography technique account for most cases of discordance. With high negative predictive value and modest positive predictive value, transient elastography is useful as a screening test to exclude advanced fibrosis.

References

1. Torres DM, Harrison SA. Diagnosis and therapy of nonalcoholic steatohepatitis. Gastroenterology 2008;134:1682-1698.
2. Wong VW, Hui AY, Tsang SW, Chan JL, Tse AM, Chan KF, et al. Metabolic and adipokine profile of Chinese patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2006;4:1154-1161.
3. Wong VW, Hui AY, Tsang SW, Chan JL, Wong GL, Chan AW, et al. Prevalence of undiagnosed diabetes and postchallenge hyperglycaemia in Chinese patients with non-alcoholic fatty liver disease. Aliment Pharmacol Ther 2006;24:1215-1222.
4. Bugianesi E, Leone N, Vanni E, Marchesini G, Brunello F, Carucci P, et al. Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. Gastroenterology 2002;123:134-140.
5. Hui AY, Wong VW, Chan HL, Liew CT, Chan JL, Chan FK, et al. Histological progression of non-alcoholic fatty liver disease in Chinese patients. Aliment Pharmacol Ther 2005;21:407-413.
6. Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. HEPATOLOGY 2006;44:865-873.
7. Merriman RB, Ferrell LD, Patti MG, Weston SR, Pabst MS, Aouizerat BE, et al. Correlation of paired liver biopsies in morbidly obese patients with suspected nonalcoholic fatty liver disease. HEPATOLOGY 2006;44:874-880.
8. Guha IN, Parkes J, Roderick PR, Harris S, Rosenberg WM. Non-invasive markers associated with liver fibrosis in non-alcoholic fatty liver disease. Gut 2006;55:1650-1660.
9. Wong VW, Wong GL, Chim AM, Tse AM, Tsang SW, Hui AY, et al. Validation of the NAFLD fibrosis score in a Chinese population with low prevalence of advanced fibrosis. Am J Gastroenterol 2008;103:1682-1688.
10. Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM. Ultrasound-based transient elastography for the detection of hepatic fibrosis: systematic review and meta-analysis. Clin Gastroenterol Hepatol 2007;5:1214-1220.
11. 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.
12. 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.
13. Wong GL, Wong VW, Choi PC, Chan AW, Chum RH, Chan HK, et al. Assessment of fibrosis by transient elastography compared with liver biopsy and morphometry in chronic liver diseases. Clin Gastroenterol Hepatol 2008;6:1027-1035.
14. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. HEPATOLOGY 2003;38:518-526.
15. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. HEPATOLOGY 2006;43:1317-1325.
16. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ; Nash Clinical Research Network. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2009;7:1104-1112.
17. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. HEPATOLOGY 2007;45:846-854.
18. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. Gut 2008;57:1441-1447.
19. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. HEPATOLOGY 2005;41:1313-1321.
20. Yoneda M, Mawatari H, Fujita K, Endo H, Iida H, Nozaki Y, 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.
21. 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.
22. Gomez-Dominguez E, Mendoza J, Rubio S, Moreno-Monteagudo JA, Garcia-Buey L, Moreno-Otero R. Transient elastography: a valid alternative to biopsy in patients with chronic liver disease. Aliment Pharmacol Ther 2006;24:513-518.
23. Chan HL, Wong GL, Choi PC, Chan AW, Chim AM, Yiu KK, et al. Alanine aminotransferase-based algorithms of liver stiffness measurement by transient elastography (Fibroscan) for liver fibrosis in chronic hepatitis B. J Viral Hepatol 2009;16:36-44.

24. Nahon P, Kettaneh A, Tengher-Barna I, Ziol M, de Ledinghen V, Douvin C, et al. Assessment of liver fibrosis using transient elastography in patients with alcoholic liver disease. J Hepatol 2008;49:1062-1068.

25. 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.

26. 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.

27. 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.

28. Wong GL, Wong VW, Choi PC, Chan AW, Chim AM, Yiu KK, et al. Increased liver stiffness measurement by transient elastography in severe acute exacerbation of chronic hepatitis B. J Gastroenterol Hepatol 2009; 24:1002-1007.

29. 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.

30. Wong VW, Wong GL, Tsang SW, Hui AY, Chan AW, Choi PC, et al. Metabolic and histological features of non-alcoholic fatty liver disease patients with different serum alanine aminotransferase levels. Aliment Pharmacol Ther 2009;29:387-396.

31. 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.

32. Mehta SH, Lau B, Afdhal NH, Thomas DL. Exceeding the limits of liver histology markers. J Hepatol 2009;50:36-41.

33. De Ledinghen V, Fournier C, Foucher J, Miette V, Vergniol J, Rigalleau V, et al. New Fibroscan probe for obese patients: a pilot study of feasibility and performances in patients with BMI >30 kg/m^2. J Hepatol 2009;50(Suppl 1):S359.