HEPATOLOGY
Official Journal of the American Association for the Study of Liver Diseases

Liver Stiffness in Nonalcoholic Fatty Liver Disease: A Comparison of Supersonic Shear Imaging, FibroScan, and ARFI With Liver Biopsy

Christophe Cassinotto,[1,2] Jérome Boursier,[3,4] Victor de Lédinghen,[2,5] Jérome Lebigot,[4,6] Bruno Lapuyade,[1] Paul Cales,[3,4] Jean-Baptiste Hiriart,[5] Sophie Michalak,[4,7] Brigitte Le Bail,[2,8] Victoire Cartier,[4,6] Amaury Mouries,[1] Frédéric Oberti,[3,4] Isabelle Fouchard-Hubert,[3] Julien Vergniol,[5] and Christophe Aubé[4,6]

Nonalcoholic fatty liver disease (NAFLD) has become a major public health issue. The goal of this study was to assess the clinical use of liver stiffness measurement (LSM) evaluated by supersonic shear imaging (SSI), FibroScan, and acoustic radiation force impulse (ARFI) in a cohort of NAFLD patients who underwent liver biopsy. A total of 291 NAFLD patients were prospectively enrolled from November 2011 to February 2015 at 2 French university hospitals. LSM was assessed by SSI, FibroScan (M probe), and ARFI within two weeks prior to liver biopsy. Calculations of the area under the receiver operating curve (AUROC) were performed and compared for the staging of liver fibrosis. AUROC for SSI, FibroScan, and ARFI were 0.86, 0.82, and 0.77 for diagnoses of ≥F2; 0.89, 0.86, and 0.84 for ≥F3; and 0.88, 0.87, and 0.84 for F4, respectively. SSI had a higher accuracy than ARFI for diagnoses of significant fibrosis (≥F2) (P = 0.004). Clinical factors related to obesity such as body mass index ≥ 30 kg/m^2, waist circumference ≥102 cm or increased parietal wall thickness were associated with LSM failures when using SSI or FibroScan and with unreliable results when using ARFI. In univariate analysis, FibroScan values were slightly correlated with NAFLD activity score and steatosis (R = 0.28 and 0.22, respectively), whereas SSI and ARFI were not; however, these components of NAFLD did not affect LSM results in multivariate analysis. The cutoff values for SSI and FibroScan for staging fibrosis with a sensitivity ≥90% were very close: 6.3/6.2 kPa for ≥F2, 8.3/8.2 kPa for ≥F3, and 10.5/9.5 kPa for F4. *Conclusion:* Although obesity is associated with an increase in LSM failure, the studied techniques and especially SSI provide high value for the diagnosis of liver fibrosis in NAFLD patients. (HEPATOLOGY 2015; 00:000–000)

As a result of the growing obesity epidemic, nonalcoholic fatty liver disease (NAFLD) has become a major public health issue. NAFLD currently represents the leading cause of chronic liver disease in Western countries and the second leading etiology among adults awaiting liver transplantation in the United States.[1-3] Liver-related morbidity and mortality in NAFLD patients are linked to the development of nonalcoholic steatohepatitis (NASH), which can progress to fibrosis and cirrhosis lesions.[4] NASH is also associated with an increased risk of hepatocellular carcinoma development and an increased risk of

Abbreviations: ARFI, acoustic radiation force impulse; AUROC, area under the receiver operating characteristics curve; BMI, body mass index; IQR, interquartile range; LSM, liver stiffness measurement; NAFLD, nonalcoholic fatty liver disease; NAS, NAFLD activity score; NASH, nonalcoholic steatohepatitis; SSI, supersonic shear imaging.

From the [1]Department of Diagnostic and Interventional Imaging, Hôpital Haut-Lévêque, University Hospital of Bordeaux, Pessac, France; [2]INSERM U1053, University of Bordeaux, Bordeaux, France; the [3]Department of Hepatology, University Hospital of Angers, France; [4]Laboratory HIFIH, UPRES 3859, LUNAM University, Angers, France; [5]Centre d'Investigation de la Fibrose Hépatique, Hôpital Haut-Lévêque, University Hospital of Bordeaux, Pessac, France; the [6]Department of Radiology, University Hospital of Angers, France; the [7]Department of Pathology, University Hospital of Angers, France; the [8]Department of Pathology, Hôpital Pellegrin, University Hospital of Bordeaux, Bordeaux, France.

Received July 22, 2015; accepted December 10, 2015.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.28394/suppinfo.

cardiovascular mortality and type 2 diabetes mellitus.[5] Liver biopsy is considered the gold standard for the evaluation and quantification of liver fibrosis in NAFLD patients. However, this diagnostic method is invasive and often painful and can be associated with rare but severe complications.[6]

The development of noninvasive diagnostic tools to assess liver fibrosis has been an increasingly studied area of research. In recent years, several noninvasive diagnostic tools, based on liver stiffness measurement (LSM) methods such as FibroScan (Echosens, Paris, France), acoustic radiation force impulse (ARFI, Siemens Medical Solutions, Mountain View, CA) elastography, and supersonic shear imaging (SSI, Supersonic Shear Imagine, Aix-en-Provence, France) have been developed to quantify liver fibrosis through the measurement of mechanical or ultrasound shear wave propagation through the hepatic parenchyma. Numerous studies have confirmed the efficiency and reproducibility of these techniques in the diagnosis of severe fibrosis and cirrhosis in patients affected by a chronic viral hepatitis.[7-12] NAFLD is a challenge for these noninvasive techniques based on a physical approach, as the increased soft tissue/fat in obese patients leads to a poorer transmission of the ultrasound or mechanical beam, leading to failures in measurements or unreliable results. For FibroScan, failures of measurement or noninterpretable results occur in 5% and 15% of cases, respectively, with the main limiting factor being obesity.[13] ARFI elastography demonstrates a decrease in its diagnostic performances that increases with the extent of overweight.[14] Regarding the SSI technique, no study has focused on the effects of overweight on its diagnostic performance for assessing liver fibrosis. Because overweight and NAFLD are intimately linked, it is important to question how this new technique behaves in NAFLD/NASH patients. Moreover, the impact on liver stiffness of major components of NAFLD such as steatosis, inflammation, and hepatocyte ballooning remains unclear.

The aim of this study was to prospectively compare the diagnostic accuracy of SSI, FibroScan, and ARFI for the staging of fibrosis in NAFLD-related chronic liver disease using liver histology as the reference standard.

Patients and Methods

Patients. We prospectively enrolled consecutive adult patients with NAFLD who had undergone liver biopsy from November 2011 to February 2015. Patients were recruited from two French academic centers (the University Hospital of Angers and University Hospital of Bordeaux) that act as tertiary centers for the management of liver diseases (108 patients recruited from the University Hospital of Bordeaux had been included in a previous study comparing several noninvasive tests in patients with chronic liver diseases of mixed etiologies).[12] Included were patients who had LSM performed using SSI, ARFI, and FibroScan within the 2 weeks before liver biopsy. Exclusion criteria were high alcohol consumption (i.e., >21 drinks, on average, per week in men and >14 drinks, on average, per week in women),[4,15] associated causes of liver disease (alcoholic, viral, or other causes of liver injury), and a liver biopsy length of <10 mm and/or fewer than six portal spaces and/or more than two fragments, except in cases of cirrhosis. All physicians who performed LSM experiments were blinded to the results of other noninvasive methods and liver biopsies. An ethics committee approved the design of this cross-sectional study, and written informed consent was obtained from all patients.

Liver Histological Examination. Histopathological staging for liver fibrosis served as the reference standard. Liver biopsies were fixed in formalin and embedded in paraffin. All biopsy specimens were analyzed by one of three pathologists who specialized in liver diseases and were blinded to the noninvasive test results. Fibrosis was staged according to the NASH Clinical Research Network scoring system[16] as follows: stage 0 (F0), absence of fibrosis; stage 1 (F1), perisinusoidal or portal; stage 2 (F2), perisinusoidal and portal/periportal; stage 3 (F3), septal or bridging fibrosis; stage 4 (F4), cirrhosis. The grade of steatosis was also defined according to Kleiner et al.[16]: S0, <5%; S1, 5%-33% (mild); S2, 34%-66% (moderate); and S3, >67% (marked). Disease activity was quantified using the NAFLD activity score (NAS), which is based on the severity of steatosis (0-3), inflammation (0-3), and hepatocellular ballooning (0-2).

Address reprint requests to: Christophe Cassinotto, M.D., Unité d'Imagerie Abdominale, Digestive et Interventionnelle; Service de Radiologie Diagnostique et Interventionnelle, Hôpital Haut-Lévêque, Centre Hospitalier Universitaire de Bordeaux, 1, Avenue de Magellan, 33604 Pessac, France. E-mail: christophe.cassinotto@chu-bordeaux.fr; fax: +33-557656472.

Copyright © 2015 by the American Association for the Study of Liver Diseases.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep.28394

Potential conflict of interest: Nothing to report.

Morphological and Biological Parameters. For all patients, the following parameters were determined at the time of liver stiffness measurement. Clinical parameters included age, sex, body mass index (BMI), waist circumference, history of diabetes, and hypertension. A blood sample was obtained to quantify the platelet count, prothrombin time, total bilirubin levels, and gamma-glutamyl transpeptidase, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, albumin, cholesterol, triglyceride, and HDL cholesterol levels. Intercostal wall thickness was also measured in millimeters during liver ultrasound examination.

SSI. SSI is integrated into a conventional ultrasound device (Aixplorer, Supersonic Imagine, Aix-en-Provence, France). Using radiation force and ultrafast ultrasound imaging, the supersonic shear imaging technique allows one to remotely generate and follow a transient plane shear wave that is propagating in vivo in real time. The radial forces that generate the shear wave are focused at increasing depths, causing a shear wave front that propagates in the area scanned. The same emitting probe receives, at high frequency, the propagation speed data from the shear waves, allowing for the updating, in real time, of the received data. Each sequence of three shear wave fronts lasts approximately one second. The tissue shear modulus, representing its stiffness, can be estimated from the shear wave local velocity. The ultrasound system then generates a real-time color map of the elasticity encoded pixel by pixel in an image superimposed on the standard B-mode.[17,18]

One of six experienced abdominal imaging radiologists performed SSI examinations in accordance with the manufacturer's instructions. LSM was performed through intercostal spaces while the patient was lying in dorsal decubitus with the right arm in maximal abduction. The operator, who was assisted by a real-time B-mode ultrasound image, targeted under apnea, a region of the right lobe of the liver with good spatial resolution for B-mode ultrasound imaging. Once a color map with complete and homogeneous filling was obtained in the zone, a region of interest 15 mm in diameter was positioned in the center of the color map in a zone free of large vascular structures and at least 15 mm below the capsule. As described in the literature and in accordance with the manufacturer's recommendations, fewer than 10 values were used for SSI technique.[11,19] Five measurements were performed on each patient. The median value of the five SSI measurements (with their interquartile range) expressed in kilopascals was used as the representative measurement. Measurements were classified as failed when no or little signal was obtained in the SSI box for all acquisitions. Because no reliability criteria have been validated for SSI or ARFI techniques, we applied the reliability criteria of FibroScan for the three techniques used in this study.[20] SSI results were then considered as unreliable when the interquartile range (IQR)/LSM was >0.30 in patients with LSM $\geq$ 7.1 kPa.

FibroScan. Transient elastography was performed with FibroScan (Echosens, Paris, France) using the standard probe (also named the M probe). Examinations were performed by one of three trained nurses (with experience performing more than 5000 LSM) blinded to SSI, ARFI, and biological results. The objective was to obtain a total of 10 valid measurements (defined as a successful liver stiffness measurement examination), with the maximum number of attempts set at 20. Liver stiffness evaluation was considered as unreliable when IQR/LSM was >0.30 in patients with LSM $\geq$ 7.1 kPa.[20]

ARFI. ARFI imaging was performed by one of the six radiologists noted above who did not perform the SSI examination. Each investigator had more than 2 years of experience in ARFI elastography. Details of the technique and the examination procedure have been described in previous reports.[21] ARFI elastography is integrated into a conventional ultrasound device (Acuson S2000, Siemens Medical Solutions, Mountain View, CA). ARFI was performed under the same conditions as SSI but in patients with apnea or with quiet breathing. Ten valid measurements were performed on each patient. The median value was determined as the representative measurement. In extension of the reliability rules applied for FibroScan, unreliable results were defined as an IQR/LSM $>$ 30% in patients with LSM $\geq$ 1.54 m/s.

Statistical Analysis. Statistical analysis was performed using NCSS 9 (NCSS software, Kaysville, UT). Data were expressed as the mean $\pm$ standard deviation or the median (IQR) and compared using either the two-sample t test or the Wilcoxon signed-rank test, according to data distribution. Percentages were compared using the χ^2 test or Fisher's exact test as appropriate. The influence of different factors on LSM failures or unreliable results was analyzed using logistic regression. The influence of clinical, biological and histological factors on LSM was analyzed using stepwise forward multiple regression. For independent variables, we selected those that influenced each dependent variable that was studied according to the univariate analysis with $P < 0.1$. The diagnostic performances of noninvasive tests were assessed in a per-protocol strategy (i.e., with exclusion of failed measurements and unreliable results) by correlating tests results and histological

Table 1. Clinical Characteristics of NAFLD Patients

Variable	All (N = 291)	Bordeaux Cohort (n = 174)	Angers Cohort (n = 117)	P
Male sex	172 (59.1)	93 (53.4)	79 (67.5)	0.02*
Age, years	56.7 ± 12 (18-80)	57 ± 12 (18-80)	56.3 ± 12 (21-79)	0.5
Body mass index, kg/m^2	32.1 ± 6 (21.3-57.4)	31.8 ± 6 (21.3-52)	32.5 ± 6 (22-57.4)	0.4
<25	28 (9.6)	20 (11.5)	8 (6.8)	0.2
25-29.9	88 (30.2)	55 (31.6)	33 (28.2)	0.5
≥30	175 (60.1)	99 (56.9)	76 (65)	0.2
Waist circumference, cm	110 ± 16 (71-193)	107 ± 16 (71-193)	113.5 ± 16 (84-168)	0.002*
Intercostal wall thickness, mm	24.8 ± 7.4 (11-64)	24.9 ± 6.7 (11-64)	24.8 ± 8 (11-61)	1
Type 2 diabetes	153 (52.6)	93 (53.4)	60 (51.3)	0.7
Hypertension	167 (57.4)	101 (58)	66 (56.4)	0.8
Metabolic syndrome	181 (62.2)	104 (59.8)	77 (65.8)	0.3
ALT, IU/L	71.2 ± 50.7 (12-332)	74.3 ± 57 (12-332)	66.7 ± 38.7 (15-229)	0.2
Alkaline phosphatase	79.2 ± 42.3 (26-578)	83.5 ± 41 (26-578)	74.5 ± 32 (32-221)	0.04*
Cholesterol, mmol/L	5 ± 1.4 (1.3-12.7)	5 ± 1.4 (1.3-8.9)	5 ± 1.4 (2.6-12.7)	1
Length of liver biopsy, mm	26.8 ± 11 (10-70)	25.7 ± 11 (10-70)	28.4 ± 9 (10-60)	0.03*
Histologic fibrosis stage				
F0	25 (8.6)	11 (6)	14 (12)	0.09
F1	60 (20.6)	31 (18)	29 (25)	0.2
F2	80 (27.5)	46 (26)	34 (29)	0.6
F3	77 (26.5)	44 (25)	33 (28)	0.6
F4	49 (16.8)	42 (24)	7 (6)	<0.0001*
NAFLD activity score				
0	6 (2)	1 (1)	5 (4)	
1	23 (8)	12 (7)	11 (9)	
2	23 (8)	13 (7)	10 (9)	
3	44 (15)	18 (10)	26 (22)	
4	56 (19)	30 (17)	26 (22)	
5	85 (29)	59 (34)	26 (22)	
6	41 (14)	29 (17)	12 (10)	
7	11 (4)	10 (6)	1 (1)	
8	2 (1)	2 (1)	0 (0)	
Steatosis grade				
S0 (<5%)	10 (3)	5 (3)	5 (4)	0.5
S1 (5%-33%)	109 (38)	44 (25)	65 (56)	<0.0001*
S2 (34%-66%)	93 (32)	62 (36)	31 (27)	0.1
S3 (>66%)	79 (27)	63 (36)	16 (14)	<0.0001*

Unless indicated otherwise, results are expressed as mean ± standard deviation (range) or number (percentage).
*$P < 0.05$.
Abbreviations: ALT, alanine aminotransferase; NAFLD, nonalcoholic fatty liver disease.

fibrosis stage using the nonparametric Spearman correlation coefficient, the results of which were compared with each other according to Guilford et al.,[22] and by building receiver operating characteristics curves. Areas under the receiver operating characteristics curves (AUROCs) and the 95% confidence intervals of the AUROC values were calculated for the detection of histological fibrosis stage 2 or higher ($F \geq 2$), fibrosis stage 3 or higher ($F \geq 3$), and cirrhosis (F4). Optimal cutoff values were determined for SSI, FibroScan, and ARFI to predict significant fibrosis, severe fibrosis, and cirrhosis according to at least 90% sensitivity, at least 90% specificity, the Youden index (maximum sum of sensitivity and specificity), and the best diagnostic accuracy (i.e., the maximum sum of true positives and true negatives over the total number of patients). Finally, AUROC values for noninvasive tests were compared using the paired method described by Zhou et al.[23] Because multiple pairwise tests were performed on a single set of data, Bonferroni correction was used to reduce the risks of false-positive results. A Bonferroni-adjusted significance level of 0.0167 was calculated for the comparison of AUROC values for SSI, FibroScan, and ARFI.

Results

From November 2011 to February 2015, 328 NAFLD patients underwent liver biopsy and LSM. Fourteen patients were excluded due to a liver biopsy length < 10 mm, and 23 patients were excluded because one of the three elastographic methods was not available at the time of the measurements due to the unavailability of ultrasound devices (defective device or maintenance). Finally, 291 patients having had LSM with SSI,

Table 2. Success and Reliability of SSI, M Probe of FibroScan, and ARFI

Success and Reliability Measurements*	SSI	FibroScan	ARFI	P		
				SSI Versus FibroScan	SSI Versus ARFI	FibroScan Versus ARFI
Total (N = 291)						
Failure	38 (13)	42 (14.4)	2 (0.7)	0.6	<0.00001*	<0.00001*
Unreliable results	21 (7.2)	26 (8.9)	53 (18.2)	0.5	0.0001*	0.001*
Reliable results	232 (79.7)	223 (76.6)	236 (81)	0.4	0.7	0.2
Waist circumference <102 cm (n = 86)†						
Failure	3 (3.5)	3 (3.5)	0 (0)	1	0.08	0.08
Unreliable results	4 (4.7)	7 (8.1)	3 (3.5)	0.4	0.7	0.2
Reliable results†	79 (91.9)	76 (88.4)	83 (96.5)	0.4	0.2	0.04*
Waist circumference ≥ 102 cm (n = 205)†						
Failure	35 (17.1)	39 (19)	2 (1)	0.6	<0.00001*	<0.00001*
Unreliable results	17 (8.3)	19 (9.3)	50 (24.4)	0.7	<0.00001*	<0.00001*
Reliable results†	153 (74.6)	147 (71.7)	153 (74.6)	0.5	1	0.5
BMI <30 kg/m^2 (n = 116)†						
Failure	6 (5.2)	4 (3.4)	0 (0)	0.5	0.01*	0.04*
Unreliable results	5 (4.3)	13 (11.2)	6 (5.2)	0.05*	0.8	0.09
Reliable results†	105 (90.5)	99 (85.3)	110 (94.8)	0.2	0.2	0.02*
BMI ≥ 30 kg/m^2 (n = 175)†						
Failure	32 (18.3)	38 (21.7)	2 (1.1)	0.4	<0.00001*	<0.00001*
Unreliable results	16 (9.1)	13 (7.4)	47 (26.9)	0.6	<0.00001*	<0.00001*
Reliable results†	127 (72.6)	124 (70.9)	126 (72)	0.7	0.9	0.8

Data are expressed as number (percentage). Comparison of proportions was performed using the χ^2 test or Fisher's exact test as appropriate.
*Failure: 0 valid measurements. Unreliable results: IQR/M > 0.3 with LSM ≥ 7.1. Reliable results: IQR/M ≤ 0.30 or > 0.3 with LSM < 7.1.
†Very reliable results and reliable results together (i.e., all patients without failure or unreliable results according to Boursier et al. definition).[20]
Abbreviations: ARFI, acoustic radiation force impulse imaging; BMI, body mass index; IQR, interquartile range; SSI, supersonic shear imaging.

M probe of FibroScan, and ARFI were included for analysis. Patient characteristics are summarized in Table 1. The number of patients with cirrhosis was lower in the Angers cohort than in the Bordeaux cohort (6% versus 24%, respectively; $P < 0.0001$). Otherwise, both groups had similar rates of obesity, type 2 diabetes, hypertension, and metabolic syndrome. With the exception of the SSI median stiffness value for diagnoses of F3 or greater disease, there were no significant differences between both centers in median levels of LSM using SSI, FibroScan, or ARFI, adjusted for fibrosis stage, or for the correlation between LSM and fibrosis stage (Supporting Table 1), allowing results from both centers to be pooled together.

Failures of Measurements or Unreliable Results. LSM failures occurred in 13% of cases (38 of 291) with SSI, 14.4% (42 of 291) with FibroScan and 0.7% with ARFI (2 of 291) (Table 2 and Figure 1). Unreliable results were observed in 7.2% of cases with SSI (21 of 291 patients), 8.9% of cases with FibroScan (26 of 291), and 18.2% of cases with ARFI (53 of 291 patients). Failure rates were lower for ARFI than for SSI or FibroScan (both $P < 0.0001$), whereas unreliable results were higher for ARFI than for SSI or FibroScan ($P = 0.0001$ and $P = 0.001$, respectively). Very reliable results (as defined for FibroScan [i.e., IQR/LSM < 0.10]) were more frequent with SSI than for FibroScan or ARFI (both $P < 0.0001$). Overall, there were no differences in the reliability of results, with 79.7% reliable results produced by SSI (232 of 291), 76.6% by FibroScan (223 of 291), and 81% by ARFI (236 of 291) ($P = 0.4$ for SSI versus FibroScan, $P = 0.7$ for SSI versus ARFI, and $P = 0.2$ for FibroScan versus ARFI).

Factors Associated with LSM Failures or Unreliable Results. Using logistic regression, we determined that the variables associated with failure of measurement or unreliable results with at least two of the three techniques studied included: waist circumference ($R^2 = 0.19$ for SSI, 0.16 for FibroScan, and 0.18 for ARFI), BMI

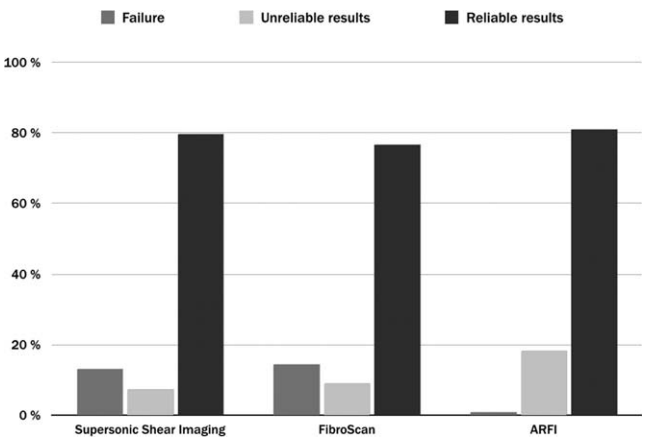


Fig. 1. Success and reliability of liver stiffness measurements.

Table 3. SSI, FibroSscan, and ARFI Values According to Fibrosis Stage

Variable	Fibrosis Stage					Correlation Coefficient
	F0	F1	F2	F3	F4	
SSI, kPa	6 (± 1.9)	6.3 (± 2.7)	8 (± 3.4)	12 (± 5)	17 (± 13.6)	0.72 ($P < 0.00001$)
FibroScan, kPa	6 (± 2)	6.7 (± 3.4)	8.4 (± 4.9)	12 (± 5.7)	23.6 (± 23.6)	0.67 ($P < 0.00001$)
ARFI, m/s	1.02 (± 0.33)	1.02 (± 0.27)	1.13 (± 0.45)	1.45 (± 0.7)	1.88 (± 1.37)	0.59 ($P < 0.00001$)

Data are expressed as the median ± IQR. Correlation between SSI, FibroScan, ARFI, and fibrosis stage was tested using the nonparametric Spearman's correlation coefficient.

Abbreviations: ARFI, acoustic radiation force impulse; IQR, interquartile range; SSI, supersonic shear imaging.

($R^2 = 0.10$ for SSI, 0.08 for FibroScan, and 0.16 for ARFI), intercostal wall thickness ($R^2 = 0.08$ for SSI, 0.05 for FibroScan, and 0.08 for ARFI), diabetes ($R^2 = 0.04$ for SSI and 0.05 for FibroScan and ARFI), presence of metabolic syndrome ($R^2 = 0.07$ for FibroScan and 0.10 for ARFI), and ALT ($R^2 = 0.03$ for FibroScan and 0.02 for ARFI). For each of the three elastography techniques, proportions of reliable results were lower in patients with a waist circumference ≥ 102 cm (Table 2): 74.6% versus 91.9% for SSI ($P < 0.00001$), 71.7% versus 88.4% for FibroScan ($P < 0.00001$), and 74.6% versus 96.5% for ARFI ($P < 0.00001$). Similarly, proportions of reliable results were lower in patients with BMI ≥30 kg/m^2: 72.6% versus 90.5% for SSI ($P < 0.00001$), 70.9% versus 85.3% for FibroScan ($P < 0.00001$), and 72% versus 94.8% for ARFI ($P < 0.00001$). However, no significant differences appeared among the three elastography techniques regarding the proportions of reliable results among patients with a waist circumference ≥ 102 cm or BMI ≥ 30 kg/m^2.

Confounding Factors of Liver Stiffness in NAFLD Patients. By univariate analysis, among many factors associated with liver stiffness, those associated with all 3 techniques were: fibrosis stage (Spearman's correlation coefficient from 0.55 to 0.68), aspartate aminotransferase (0.18 to 0.34), gamma-glutamyl transpeptidase (0.17 to 0.30), albumin (−0.22 to −0.31), age (0.16 to 0.25), platelet count (−0.20 to −0.29), diabetes (0.23 to 0.29), metabolic score (0.23 to 0.28), waist circumference (0.18 to 0.26), hypertension (0.18 to 0.26), alkaline phosphatase (0.13 to 0.26), prothrombin time (−0.12 to −0.25), and cholesterol (−0.14 to −0.28) (Supporting Table 2). In stepwise multiple regression, four independent variables were found to be associated with LSM results when using any of the three techniques: fibrosis stage, alkaline phosphatase, albumin, and waist circumference. It is notable that the presence of inflammation and hepatocyte ballooning (assessed via the NAS score), and steatosis were slightly correlated in the univariate analysis with FibroScan results

Table 4. Diagnostic Performances of SSI, M Probe of FibroScan, and ARFI

Fibrosis Stage	AUROC (95% CI)	Best Accuracy	Aim	Cutoff	Sensitivity	Specificity
SSI, kPa						
≥F2	0.86 (0.79-0.90)	80% (185/232)	Sensitivity ≥ 90%	6.3	90% (148/164)	50% (34/68)
			Specificity ≥ 90%	8.7	71% (116/164)	90% (61/68)
≥F3	0.89 (0.83-0.92)	85% (196/232)	Sensitivity ≥ 90%	8.3	91% (91/100)	71% (93/132)
			Specificity ≥ 90%	10.7	71% (71/100)	90% (119/132)
F4	0.88 (0.82-0.92)	87% (202/232)	Sensitivity ≥ 90%	10.5	90% (34/38)	72% (139/194)
			Specificity ≥ 90%	14.4	58% (22/38)	90% (174/194)
FibroScan, kPa						
≥F2	0.82 (0.76-0.87)	77% (172/223)	Sensitivity ≥ 90%	6.2	90% (141/156)	45% (30/67)
			Specificity ≥ 90%	9.8	60% (93/156)	90% (60/67)
≥F3	0.86 (0.80-0.90)	79% (175/223)	Sensitivity ≥ 90%	8.2	90% (83/92)	61% (80/131)
			Specificity ≥ 90%	12.5	57% (52/92)	90% (118/131)
F4	0.87 (0.79-0.92)	89% (198/223)	Sensitivity ≥ 90%	9.5	92% (34/37)	62% (116/186)
			Specificity ≥ 90%	16.1	65% (24/37)	90% (168/186)
ARFI, m/s						
≥F2	0.77 (0.70-0.83)	74% (175/236)	Sensitivity ≥ 90%	0.95	90% (144/160)	36% (31/76)
			Specificity ≥ 90%	1.32	56% (90/160)	91% (69/76)
≥F3	0.84 (0.78-0.89)	79% (186/236)	Sensitivity ≥ 90%	1.15	90% (84/93)	63% (90/143)
			Specificity ≥ 90%	1.53	59% (55/93)	90% (128/143)
F4	0.84 (0.78-0.89)	84% (199/236)	Sensitivity ≥ 90%	1.3	90% (35/39)	67% (132/197)
			Specificity ≥ 90%	2.04	44% (17/39)	90% (177/197)

Abbreviations: ARFI, acoustic radiation force impulse; AUROC, area under the receiver operating characteristics curve; CI, confidence interval; SSI, supersonic shear imaging.

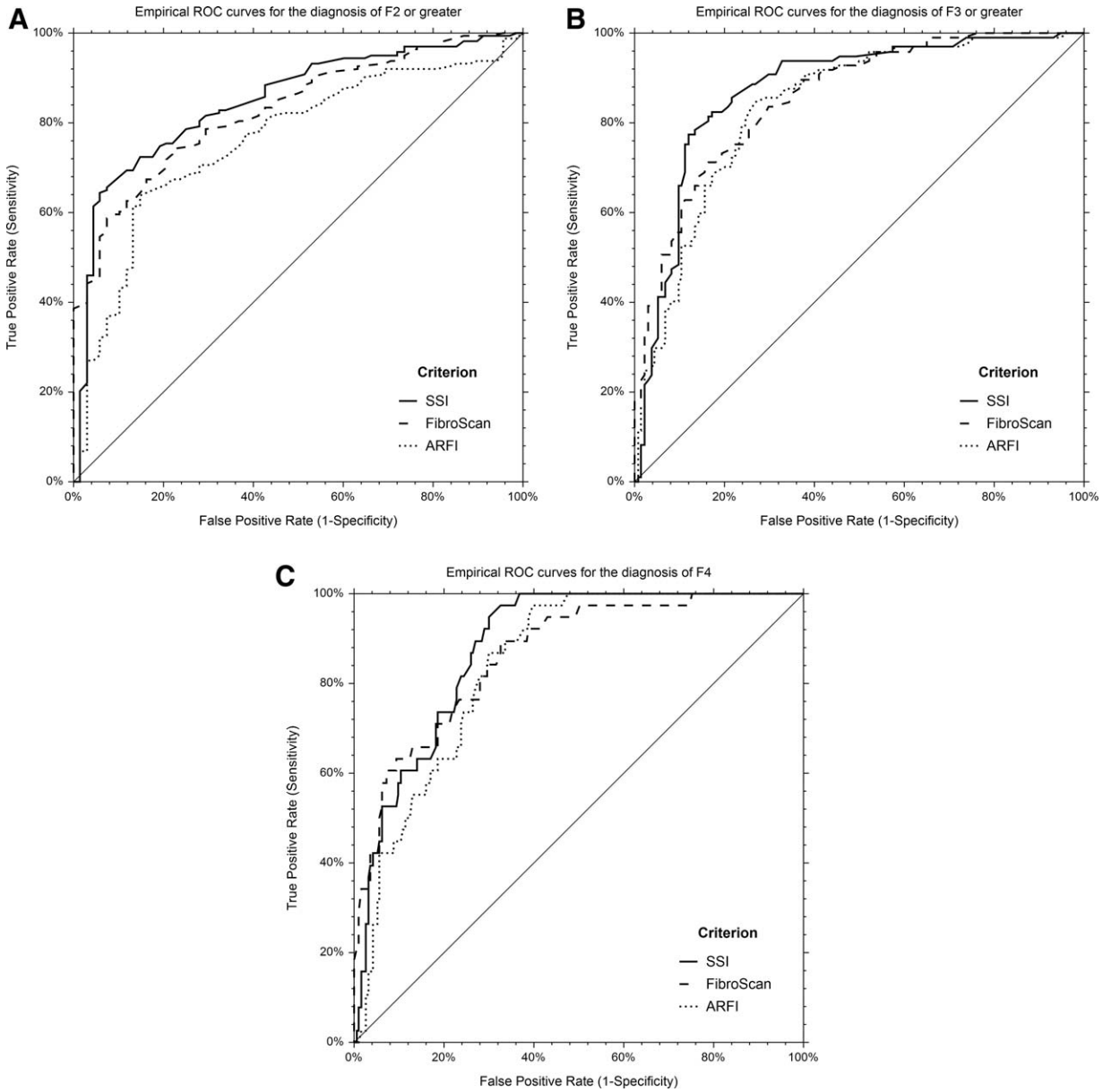


Fig. 2. ROC curves for SSI, FibroScan, and ARFI for the diagnosis of (A) significant fibrosis (F2 or greater), (B) severe fibrosis (F3 or greater), (C) and cirrhosis (F4). SSI had a higher accuracy than ARFI for the diagnosis of significant fibrosis (≥F2) ($P = 0.004$). No significant differences were observed among the three elastography techniques for the diagnosis of severe fibrosis and cirrhosis.

(Spearman's correlation coefficient of 0.28 and 0.22, respectively), but not with SSI or ARFI results. However, neither NAS nor steatosis affected the LSM results of FibroScan in the multivariate analysis.

Diagnostic Performances of Elastography Techniques. The median values of LSM for each stage of fibrosis are presented in Table 3. SSI, FibroScan, and ARFI results demonstrated a strong correlation with histological fibrosis stage (Supporting Fig. 1). However, the correlation between SSI results and fibrosis stage was significantly superior to the correlation between ARFI results and fibrosis stage ($r^2 = 0.72$ versus 0.59, $P = 0.0003$; respectively), whereas the difference between SSI and FibroScan correlation coefficients was not significant ($r^2 = 0.72$ versus 0.67; $P = 0.1$) as well as the difference between FibroScan and ARFI correlation coefficients ($r^2 = 0.67$ versus 0.59; $P = 0.08$). Measurements by SSI, FibroScan, and ARFI correlated well with each other ($r^2 = 0.70$ for SSI and FibroScan, 0.70 for SSI and ARFI, and 0.56 for M and ARFI). Diagnostic performance according to the AUROC values for the diagnosis of significant fibrosis, severe fibrosis, and cirrhosis was good for SSI, which ranged from 0.86 to 0.89; good for FibroScan, which ranged from

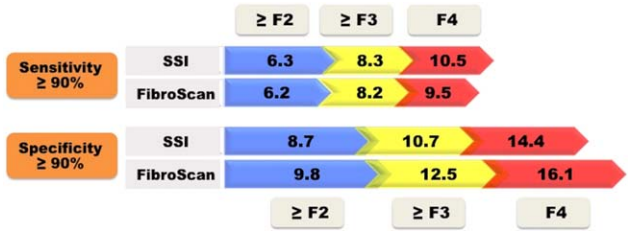


Fig. 3. Optimal liver stiffness cutoff values for SSI and FibroScan.

0.82 to 0.87; and fair or good for ARFI, which ranged from 0.77 to 0.84 (Table 4). The AUROC values for diagnosing severe fibrosis or cirrhosis were particularly good, between 0.86 and 0.89, for SSI or FibroScan.

Pairwise Comparison of Diagnostic Performances. Pairwise comparisons of AUROC values between SSI, FibroScan, and ARFI (Fig. 2) were performed on 231 patients (60 patients were excluded from the analysis due to failure of measurement either with SSI, FibroScan, or ARFI). For the diagnosis of significant fibrosis (≥F2), SSI had a significantly better diagnostic performance than ARFI (AUROC of 0.85, 0.83, and 0.76 for SSI, FibroScan and ARFI, respectively; $P = 0.5$ for SSI versus FibroScan; $P = 0.004$ for SSI versus ARFI; and $P = 0.07$ for FibroScan versus ARFI). For the diagnosis of severe fibrosis (≥F3), no significant differences were observed between SSI, FibroScan, and ARFI (AUROC of 0.87, 0.86, and 0.84, respectively; $P = 0.5$ for SSI versus FibroScan; $P = 0.2$ for SSI versus ARFI; and $P = 0.5$ for FibroScan versus ARFI). Similarly, no significant differences were observed for the diagnosis of cirrhosis (F4) between SSI, FibroScan, and ARFI (AUROC of 0.88, 0.86, and 0.84, respectively; $P = 0.5$ for SSI versus FibroScan; $P = 0.07$ for SSI versus ARFI; and $P = 0.5$ for FibroScan versus ARFI).

Optimal Liver Stiffness Cutoff Values. The optimal LSM cutoff values, according to a sensitivity of at least 90% and a specificity of at least 90% for the diagnosis of F2 or greater, F3 or greater, and F4 diseases, are described in Table 4. It is notable that the cutoff values for SSI and FibroScan for ruling out diseases with at least 90% sensitivity were very close: 6.3 kPa and 6.2 kPa for the diagnosis of F2 or greater, 8.3 kPa and 8.2 kPa for the diagnosis of severe fibrosis, and 10.5 kPa and 9.5 kPa for the diagnosis of cirrhosis, respectively (Fig. 3). The LSM cutoff values for SSI and FibroScan, according to the Youden index and the best diagnostic accuracy, were also very close (Supporting Table 3). However, SSI generated lower stiffness values than FibroScan for ruling in diseases with a specificity of at least 90%: 8.7 kPa and 9.8 kPa for the diagnosis of F2 or greater, 10.7 kPa and 12.5 kPa for the diagnosis of severe fibrosis, and 14.4 kPa and 16.1 kPa for the diagnosis of cirrhosis, respectively.

Discussion

In this prospective cohort of NAFLD patients who received liver biopsies, SSI, FibroScan, and ARFI performed well with regard to diagnosing different stages of liver fibrosis. These techniques achieved good diagnostic performance, especially SSI with AUROC values ranging from 0.86 to 0.89 and FibroScan with AUROC values from 0.82 to 0.87, whereas ARFI AUROC values ranged from 0.77 to 0.84. The diagnostic accuracy of SSI was greater than that of ARFI in the diagnosis of significant fibrosis or greater disease (AUROC of 0.85 versus 0.76; $P = 0.004$). Another important finding of our study is that several cutoff values for SSI and FibroScan were very close, which can be of great use in routine daily practice.

Although several validation studies have previously confirmed the accuracy of testing LSM using FibroScan or ARFI in NAFLD, our study is the first to assess the accuracy of SSI and to compare these three techniques in this field. The diagnostic performance of the M probe for FibroScan in our patient population was consistent with previous prospective studies.[24,25] Conversely, our results concerning ARFI accuracy were less striking than those in previous studies. In a preliminary study by Yoneda et al.,[26] analysis was performed on 54 NAFLD patients, including few patients with severe fibrosis or cirrhosis (F3, n = 4; and F4, n = 6). Palmeri et al.[27] found that the best accuracy for ARFI in NAFLD patients occurred when distinguishing between patients with no or moderate fibrosis (F0 to F2) and those with severe fibrosis or cirrhosis (F3-F4). Unfortunately, in our study, the improved accuracy of ARFI for diagnosing F2 or greater disease was not confirmed.

Indeed, the major challenge when measuring LSM in NAFLD patients is the lower success rate in obese patients. Consistent with previous results,[28,29] obesity was, in our study, the main feature associated with a significant increase in the failure rate when using FibroScan. We showed that failed, unreliable, and reliable LSM rates when using SSI were very close to those of FibroScan, with a great impact of obesity on LSM failure rates. Reliable results when using SSI were obtained in approximately 90% of patients with a BMI < 30 kg/m^2 (~85% for FibroScan) but only in nearly 73% of patients with a BMI ≥ 30 kg/m^2 (71% for FibroScan). The interference by obesity on the ARFI technique was less conspicuous, with a failure rate that was not significantly increased in obese patients. However, the impact

of obesity on ARFI diagnostic accuracy and reliability should not be underappreciated, as a much higher number of patients had unreliable LSM in obese patients.

In a multivariate analysis, LSM was mainly associated with fibrotic stage. Multivariate analysis also revealed that steatosis, inflammation, or hepatocyte ballooning (indirectly assessed via the NAS score) did not have a significant effect on LSM results. This result can be considered as positive because the impact of different confounding factors is frequently a limitation for most diagnostic techniques. However, the lack of a link between these factors and LSM can also be considered to be negative when diagnosing NAFLD patients, because some NASH patients may have no fibrosis but do present with liver injury lesions with inflammation or hepatocyte ballooning. Therefore, LSM may not aid in the diagnosis of early stage NASH without fibrosis. Moreover, it is interesting to note that in univariate analysis, FibroScan results had a slight correlation with steatosis and NAFLD activity score (R = 0.22 and 0.28), whereas SSI and ARFI did not (with close correlation coefficients: R = 0.10 and 0.07 for NAS, and R = 0.03 and 0.0002 for steatosis, respectively). We can suppose that these differences are due to technical differences between the technique of FibroScan and those of SSI and ARFI, as the velocity of the mechanical wave generated by the probe of FibroScan is recorded on about 4 cm into the liver, whereas the shear wave propagation issued from the focused ultrasonic beams of ARFI and SSI are recorded at a very low amplitude.

Our study has several limitations. First, liver biopsy was used as the reference standard to assess and quantify liver fibrosis. The diagnostic efficiency of biopsy is normally limited by the variability in the representativeness of the samples and by inter- and intraobserver variability. To reduce these limitations, we excluded all patients with biopsies that did not meet specific quality criteria. Moreover, the main results of our study were sustained when biopsies between 10 and 15 mm were excluded (data not shown). A second limitation of our study is the lack of use of the XL probe used for FibroScan. The XL probe, which is used for overweight patients, has been shown to increase the rate of reliable results significantly, especially in obese patients, with an accuracy similar to that of the M probe.[30] It would have been interesting to analyze the added value of the XL probe when SSI and the M probe were also available; unfortunately, the XL probe was not available at the time of LSM for a part of our study population. Third, our study focused on LSM techniques and did not analyze the impact or clinical value of the biological fibrosis tests that are usually performed in the field of NAFLD. Further studies are needed to assess the ways in which LSM and biological tests could be used together to provide the best diagnostic performance for NAFLD patients.

In conclusion, this prospective study confirmed that SSI, FibroScan, and ARFI are valuable diagnostic tools for the staging of liver fibrosis in NAFLD patients. However, the diagnostic accuracy of SSI was superior to that of ARFI for the diagnosis of F2 or greater disease states. Most of the cutoff values for SSI for the diagnosis of different stages of liver disease were quite close to those for FibroScan, which could be of great importance for the applicability of this technique and its wide dissemination among radiologists and hepatologists in their daily practice. However, as for the M probe of FibroScan, the SSI technique remains limited by a significant increase in its failure rate in cases of obesity, whereas ARFI presents with a great increase of unreliable results.

References

1. Williams CD, Stengel J, Asike MI, Torres DM, Shaw J, Contreras M, et al. Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. Gastroenterology 2011; 140:124-131.
2. Wong VW, Wong GL, Yeung DK, Lau TK, Chan CK, Chim AM, et al. Incidence of non-alcoholic fatty liver disease in Hong Kong: a population study with paired proton-magnetic resonance spectroscopy. J Hepatol 2015;62:182-189.
3. Wong RJ, Aguilar M, Cheung R, Perumpail RB, Harrison SA, Younossi ZM, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. Gastroenterology 2015;148:547-555.
4. Sanyal AJ, Brunt EM, Kleiner DE, Kowdley KV, Chalasani N, Lavine JE, et al. Endpoints and clinical trial design for nonalcoholic steatohepatitis. HEPATOLOGY 2011;54:344-353.
5. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther 2011;34: 274-285.
6. Cadranel JF, Rufat P, Degos F. Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AFEF). HEPATOLOGY 2000;32:477-481.
7. Castéra L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, et al. Prospective comparison of transient elastography, FibroTest™, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. Gastroenterology 2005;128:342-350.
8. 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.
9. Nierhoff J, Chávez Ortiz AA, Herrmann E, Zeuzem S, Friedrich-Rust M. The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: a meta-analysis. Eur Radiol 2013;23:3040-3053.
10. Bota S, Herkner H, Sporea I, Salzl P, Sirli R, Neghina AM, et al. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver Int 2013;33:1138-1147.
11. Ferraioli G, Tinelli C, Dal Bello B, Zicchetti M, Filice G, Filice C, et al. Accuracy of real-time shear wave elastography for assessing liver

fibrosis in chronic hepatitis C: a pilot study. Hepatology 2012;56:2125-2133.
12. Cassinotto C, Lapuyade B, Mouries A, Hiriart JB, Vergniol J, Gaye D, et al. Non-invasive assessment of liver fibrosis with impulse elastography: comparison of Supersonic Shear Imaging with ARFI and FibroScan®. J Hepatol 2014;61:550-557.
13. Castéra L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology 2010;51:828-835.
14. Cassinotto C, Lapuyade B, Aït-Ali A, Vergniol J, Gaye D, Foucher J, et al. Liver fibrosis: non-invasive assessment with acoustic radiation force impulse elastography-comparison with FibroScan M and XL probes and FibroTest™ in patients with chronic liver disease. Radiology 2013;269:283-292.
15. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. Gastroenterology 2012;142:1592-1609.
16. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for non-alcoholic fatty liver disease. Hepatology 2005;41:1313-1321.
17. Bercoff J, Tanter M, Fink M. Supersonic shear imaging: a new technique for soft tissue elasticity mapping. IEEE Trans Ultrason Ferroelectr Freq Control 2004;51:396-409.
18. Muller M, Gennisson JL, Deffieux T, Tanter M, Fink M. Quantitative viscoelasticity mapping of human liver using supersonic shear imaging: preliminary in vivo feasability study. Ultrasound Med Biol 2009;35:219-229.
19. Sporea I, Gradinaru-Tascau O, Bota S, Popescu A, Sirli R, Jurchis A, et al. How many measurements are needed for liver stiffness assessment by 2D-shear wave elastography (2D-SWE) and which value should be used: the mean or median? Med Ultrason 2013;15:268-272.
20. Boursier J, Zarski JP, de Lédinghen V, Rousselet MC, Sturm N, Lebail B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology 2013;57:1182-1191.
21. Friedrich-Rust M, Wunder K, Kriener S, Sotoudeh F, Richter S, Bojunga J, et al. Liver fibrosis in viral hepatitis: noninvasive assessment with acoustic radiation force impulse imaging versus transient elastography. Radiology 2009;252:595-604.
22. Guilford J.P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill; 1965.
23. Zhou X, Obuchowski N, McClish D. Statistical Methods in Diagnostic Medicine. 1st ed. New York: John Wiley & Sons; 2002.
24. Wong VW, Vergniol J, Wong GL, Foucher J, Chan HL, Le Bail B, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 2010;51:454-462.
25. Wong VW, Vergniol J, Wong GL, Foucher J, Chan AW, Chermak F, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol 2012;107:1862-1871.
26. Yoneda M, Suzuki K, Kato S, Fujita K, Nozaki Y, Hosono K, et al. Nonalcoholic fatty liver disease: US-based acoustic radiation force impulse elastography. Radiology 2010;256:640-647.
27. Palmeri ML, Wang MH, Rouze NC, Abdelmalek MF, Guy CD, Moser B, et al. Noninvasive evaluation of hepatic fibrosis using acoustic radiation force-based shear stiffness in patients with nonalcoholic fatty liver disease. J Hepatol 2011;55:666-672.
28. Petta S, Di Marco V, Cammà C, Butera G, Cabibi D, Craxì A. Reliability of liver stiffness measurement in non-alcoholic fatty liver disease: the effects of body mass index. Aliment Pharmacol Ther 2011;33:1350-1360.
29. Myers RP, Pomier-Layrargues G, Kirsch R, Pollett A, Duarte-Rojo A, Wong D, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology 2012;55:199-208.
30. de Lédinghen V, Wong VW, Vergniol J, Wong GL, Foucher J, Chu SH, et al. Diagnosis of liver fibrosis and cirrhosis using liver stiffness measurement: comparison between M and XL probe of FibroScan®. J Hepatol 2012;56:833-839.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.28394/suppinfo.

Supplementary Table S1. Supersonic shear imaging, Fibroscan and ARFI median values according to the fibrosis stage in Bordeaux and Angers cohort.

	Study site	F0	F1	F2	F3	F4	**Correlation coefficient**
SSI (kPa)	Bordeaux	6 (±2)	6.1 (±2.2)	8 (±3.8)	10.5 (±5.7)	17 (±13.6)	0.74
	Angers	6.6 (±1.9)	6.8 (±3)	8.7 (±4.3)	13.3 (±5.8)	23.5 (±9.7)	0.63
	p-value	0.1	0.2	0.3	0.0008*	0.3	0.1
Fibroscan (kPa)	Bordeaux	6.4 (±3)	7.5 (± 3.2)	8.8 (± 5.2)	12 (± 5.5)	22.7 (±23.3)	0.65
	Angers	6 (±1.9)	6.2 (±3.2)	8.1 (±4.7)	13.2 (±7.1)	38.2 (±54.8)	0.66
	p-value	0.9	0.2	0.6	0.3	0.3	0.9
ARFI (m/s)	Bordeaux	1.03 (±0.36)	1.01 (±0.24)	1.24 (±0.46)	1.54 (±0.7)	1.92 (±1.32)	0.58
	Angers	1.15 (±0.25)	1.06 (±0.31)	1.13 (±0.44)	1.85 (±1)	2.91 (±1.74)	0.50
	p-value	0.5	0.4	0.3	0.09	0.06	0.3

Except for the correlation coefficient (Spearman's correlation coefficient), data are expressed as medians (± interquartile range). Comparison of values between the different subgroups was performed using non-parametric Wilcoxon rank-sum test, and Spearman's correlation coefficients were compared each other according to Guilford et al.[22]

Abbreviations: SSI, Supersonic Shear Imaging; ARFI, Acoustic Radiation Force Impulse; kPa, KiloPascals.

Supplementary Table S2. Factors influencing liver stiffness measurements.

Univariate analysis was performed using Spearman's correlation coefficient (ρ) and multivariate analysis using stepwise forward regression (R^2).

	Supersonic Shear Imaging			FibroScan			ARFI		
	ρ	**p-value**	**R^2**	**ρ**	**p-value**	**R^2**	**ρ**	**p-value**	**R^2**
Gender	-0.03	0.7		-0.03	0.6		-0.15	0.01	
Age	0.25	0.00008		0.20	0.002		0.16	0.006	
Body mass index	0.09	0.2		0.15	0.02		0.21	0.0003	
Parietal wall thickness	-0.04	0.5		0.09	0.2		0.04	0.5	
Metabolic score (0-5)	0.28	0.000009		0.23	0.0002		0.23	0.0009	
Metabolic syndrome	0.24	0.0001		0.16	0.01		0.17	0.004	
Waist circumference	0.18	0.004	**0.03**	0.19	0.002	**0.02**	0.26	0.000006	**0.07**
Diabetes	0.29	0.000003		0.29	0.000005		0.23	0.0001	
Hypertension	0.26	0.00003		0.25	0.0001		0.18	0.002	
AST	0.29	<0.000001	**0.03**	0.34	<0.000001		0.18	0.003	
ALT	-0.01	0.9		0.05	0.4		-0.05	0.4	
GGT	0.26	0.00004		0.30	0.000002	**0.01**	0.17	0.004	
Total bilirubin	-0.02	0.7		-0.07	0.3		-0.01	0.9	
Platelet count	-0.29	0.000003		-0.20	0.002	**0.02**	-0.20	0.0006	
Prothrombin time	-0.25	0.00008	**0.01**	-0.17	0.006		-0.12	0.04	
Alkaline phosphatase	0.13	0.03	**0.04**	0.14	0.03	**0.09**	0.26	0.000009	**0.04**
Albumin	-0.31	<0.000001	**0.09**	-0.22	0.0007	**0.01**	-0.29	0.000001	**0.02**
Ferritin	-0.05	0.4		-0.04	0.6		-0.14	0.02	
Cholesterol	-0.18	0.006		-0.17	0.009		-0.14	0.02	

Histologic fibrosis stage	0.68	<0.000001	**0.19**	0.67	<0.000001	**0.33**	0.55	<0.000001	**0.22**
NAFLD activity score	0.10	0.1		0.28	0.000007		0.07	0.3	
Steatosis	0.03	0.7		0.22	0.0006		0.002	1	
Total R^2			**0.40**			**0.48**			**0.36**

AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma-glutamyl transpeptidase; NAFLD, Non-alcoholic fatty liver disease.

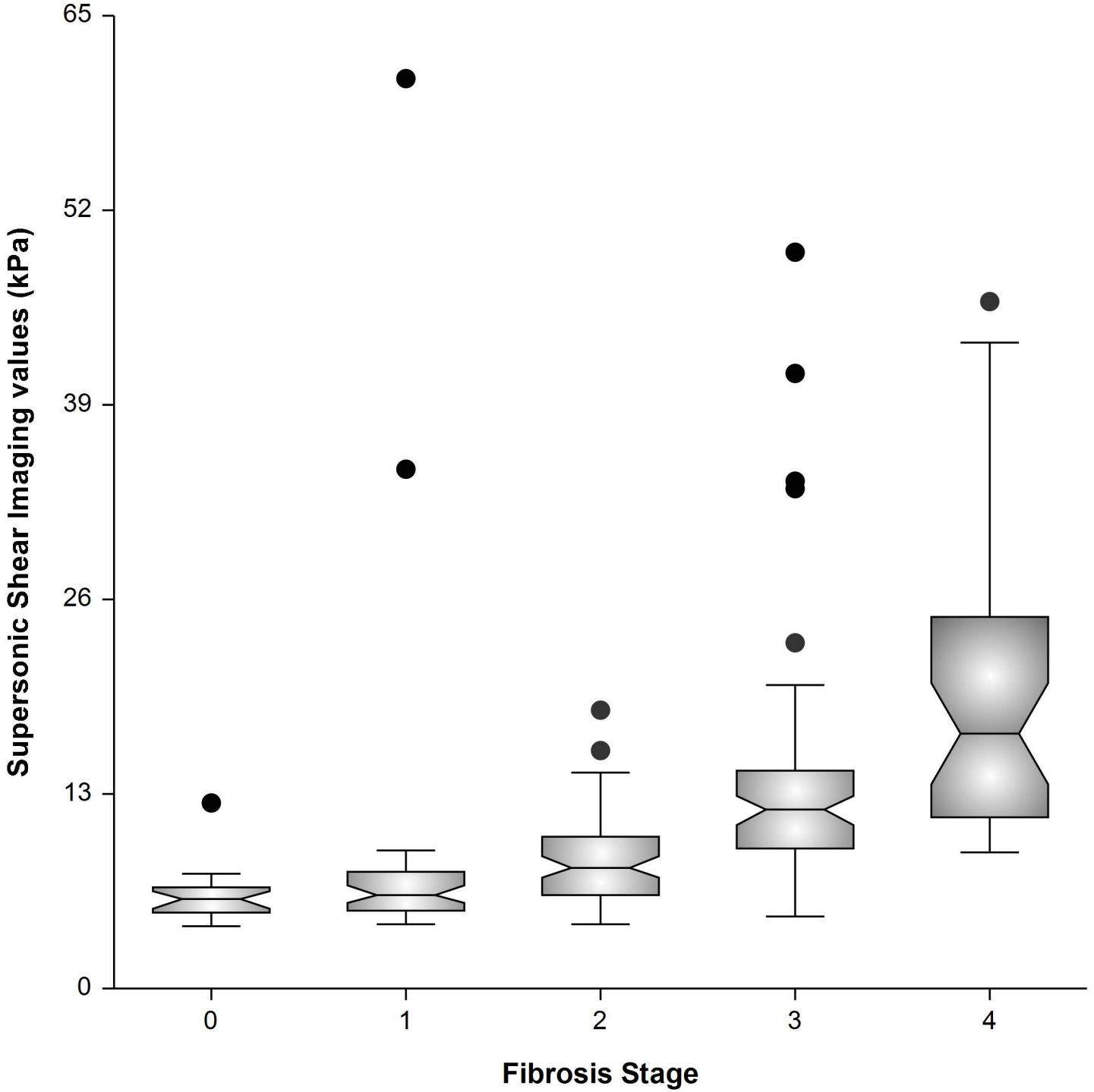
Supplementary Table S3. Cut-off values of Supersonic shear imaging, Fibroscan, and ARFI according to the Younden index, and the best diagnostic accuracy.

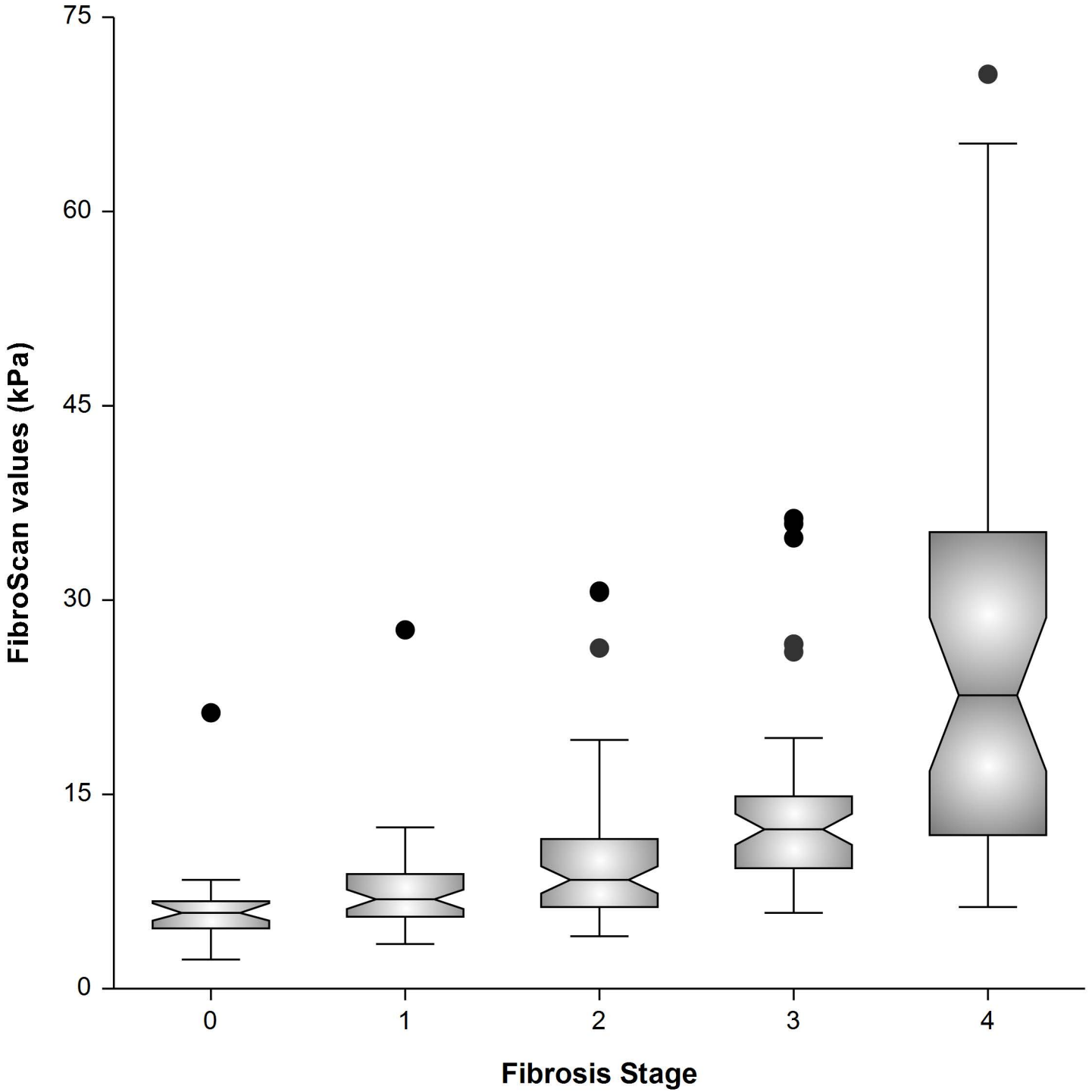
Variable	*Fibrosis stage*	*AUROC (95%CI)*	*Aim*	*Cut-off*	*Sensitivity*	*Specificity*	*Accuracy*
SSI	**≥F2**	*0.86*	*Younden*	***8.9***	68% (111/164)	94% (64/68)	75% (175/232)
(kPa)		*(0.79-0.90)*	*Best accuracy*	***6.2***	93% (152/164)	49% (33/68)	80% (185/232)
	≥F3	*0.89*	*Younden*	***9.3***	84% (84/100)	83% (110/132)	84% (194/232)
		(0.83-0.92)	*Best accuracy*	***10.2***	79% (79/100)	89% (117/132)	85% (196/232)
	F4	*0.88*	*Younden*	***10***	95% (36/38)	69% (133/194)	73% (169/232)
		(0.82-0.92)	*Best accuracy*	***20.3***	42% (16/38)	96% (186/194)	87% (202/232)
Fibroscan	**≥F2**	*0.82*	*Younden*	***8.5***	72% (113/156)	79% (53/67)	74% (166/223)
(kPa)		*(0.76-0.87)*	*Best accuracy*	***6.3***	89% (139/172)	49% (33/67)	77% (172/223)
	≥F3	*0.86*	*Younden*	***9.3***	82% (75/92)	75% (98/131)	78% (173/223)
		(0.80-0.90)	*Best accuracy*	***10.9***	72% (66/92)	83% (109/131)	79% (175/223)
	F4	*0.87*	*Younden*	***10.2***	89% (33/37)	68% (126/186)	71% (159/223)
		(0.79-0.92)	*Best accuracy*	***20***	60% (22/37)	95% (176/186)	89% (198/223)
ARFI	**≥F2**	*0.77*	*Younden*	***1.28***	59% (95/160)	90% (68/76)	69% (163/236)
(m/s)		*(0.70-0.83)*	*Best accuracy*	***1.07***	80% (128/160)	62% (47/76)	74% (175/236)
	≥F3	*0.84*	*Younden*	***1.26***	81% (75/93)	78% (111/143)	79% (186/236)
		(0.78-0.89)	*Best accuracy*	***1.26***	81% (75/93)	78% (111/143)	79% (186/236)

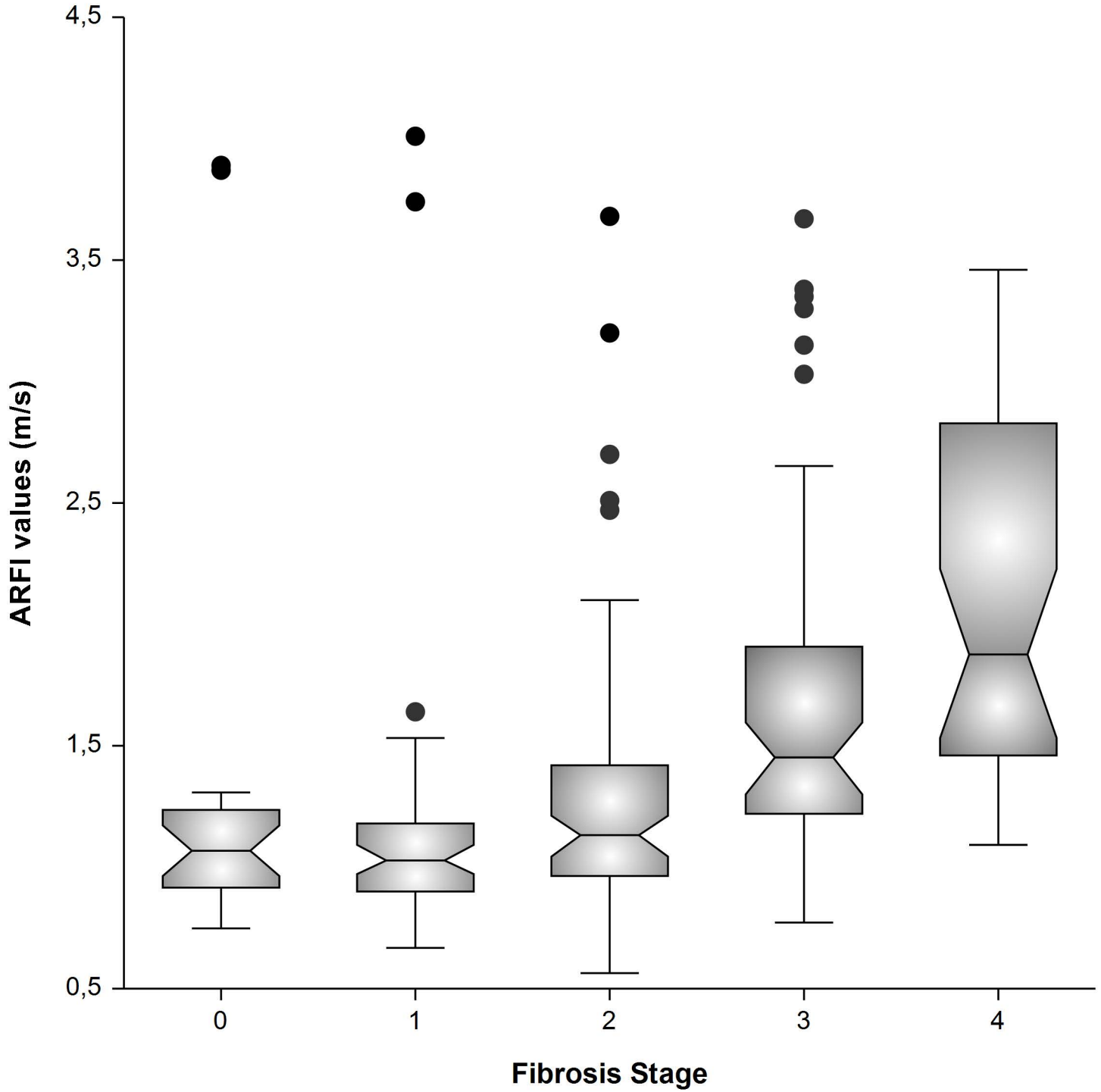
F4	*0.84*	*Younden*	***1.39***	85% (33/39)	74% (145/197)	75% (178/236)
	(0.78-0.89)	*Best accuracy*	***2.51***	41% (16/39)	93% (183/197)	84% (199/236)

Abbreviations: AUROC, areas under the receiver operating characteristic curves; CI, confidence intervals; SSI, Supersonic shear imaging.

Supplementary Figure S1. Box plots showing the distribution of Supersonic Shear Imaging (A), FibroScan (B) and ARFI (C) values in the study population according to histological fibrosis stage. The top and bottom of the box are the 25th and 75th percentiles, respectively. The length of the box is the interquartile range, and the median (50th percentile) is the line drawn through the box.


65
52
39
26
13
0
Supersonic Shear Imaging values (kPa)
0
1
2
3
4
Fibrosis Stage


75
60
45
30
15
0
FibroScan values (kPa)
0
1
2
3
4
Fibrosis Stage


4,5
3,5
2,5
1,5
0,5
ARFI values (m/s)
0
1
2
3
4
Fibrosis Stage