Liver INTERNATIONAL
OFFICIAL JOURNAL OF THE INTERNATIONAL ASSOCIATION FOR THE STUDY OF THE LIVER

Liver International ISSN 1478-3223

CLINICAL STUDIES

The combination of a blood test and Fibroscan improves the non-invasive diagnosis of liver fibrosis

Jérôme Boursier[1,2], Julien Vergniol[3,4], Apollinaire Sawadogo[1,2], Taoufiq Dakka[5], Sophie Michalak[1,6], Yves Gallois[1,7], Véronique Le Tallec[2], Frédéric Oberti[1,2], Isabelle Fouchard-Hubert[1,2], Nina Dib[1,2], Marie Christine Rousselet[1,6], Anselme Konaté[1,2], Naïma Amrani[5], Victor de Ledinghen[3,4] and Paul Calès[1,2]

1 HIFIH Laboratory, IFR 132, University, PRES UNAM, Angers, France
2 Hepato-Gastroenterology Department, CHU, Angers, France
3 Hepato-Gastroenterology Department, Haut Leveque Hospital, Pessac, CHU, Bordeaux, France
4 INSERM U889, Victor Segalen University, Bordeaux, France
5 Hepato-Gastroenterology Exploration Unit, Ibn Sina Hospital, CHU, Rabat, Morocco
6 Pathology Department, CHU, Angers, France
7 Biochemistry Department, CHU, Angers, France

Keywords
blood fibrosis test – cirrhosis – FibroMeter – liver fibrosis – liver stiffness – non-invasive diagnosis – sequential diagnostic algorithm – ultrasonographic elastometry

Correspondence
Paul Calès, Service d'Hépato-Gastroentérologie, CHU, 49933 Angers Cedex 09, France
Tel: +33 2 41 35 34 10
Fax: +33 2 41 35 41 19
e-mail: paul.cales@univ-angers.fr

Received 3 June 2009
Accepted 10 July 2009

DOI:10.1111/j.1478-3231.2009.02101.x

Abstract

Background and aims: Blood tests and liver stiffness evaluation (LSE) by ultrasonographic elastometry are accurate tools for diagnosing liver fibrosis. We evaluated whether their synchronous combination in new scores could improve the diagnostic accuracy and reduce liver biopsy requirement in algorithm. *Methods:* Three hundred and ninety patients with chronic liver disease of miscellaneous causes were included. Five blood fibrosis tests were evaluated: APRI, FIB-4, Hepascore, Fibrotest and FibroMeter. The reference was fibrosis Metavir staging. *Results: Diagnosis of significant fibrosis (Metavir $F \geq 2$).* The most accurate synchronous combination was FibroMeter+LSE, which provided a significantly higher area under the receiver operating characteristic curve (0.892) than LSE alone (0.867, $P = 0.011$) or Fibrometer (0.834, $P < 10^{-3}$). An algorithm using the FibroMeter+LSE combination and then a liver biopsy in indeterminate cases had 91.9% diagnostic accuracy and required significantly fewer biopsies (20.2%) than previously published Bordeaux algorithm (28.6%, $P = 0.02$) or sequential algorithm for fibrosis evaluation (SAFE) (55.7%, $P < 10^{-3}$). The Angers algorithm performance was not significantly different between viral hepatitis and other causes. *Diagnosis of cirrhosis*. The most accurate synchronous combination was LSE+FibroMeter, which provided $\geq 90\%$ predictive values for cirrhosis in 90.6% of patients vs 87.4% for LSE ($P = 0.02$) and 57.9% for FibroMeter ($P < 10^{-3}$). An algorithm including the LSE+FibroMeter combination, and then a liver biopsy in indeterminate cases, had a significantly higher diagnostic accuracy than the SAFE algorithm (91.0 vs 79.8%, $P < 10^{-3}$), and required significantly fewer biopsies than the Bordeaux algorithm (9.3 vs 25.3%, $P < 10^{-3}$). *Conclusion:* The synchronous combination of a blood test plus LSE improves the accuracy of the non-invasive diagnosis of liver fibrosis and, consequently, markedly decreases the biopsy requirement in the diagnostic algorithm, notably to $< 10\%$ in cirrhosis diagnosis.

The non-invasive diagnosis of liver fibrosis has gained considerable attention over the last 10 years. Initially, several studies demonstrated that some blood markers correlated well with the fibrosis stages measured on liver specimens (1–3). This was followed by the development of blood tests combining several fibrosis markers to improve the diagnostic accuracy. The first generation of simple blood tests combined common indirect fibrosis markers into a simple ratio, like APRI (4) and FIB-4 (5). The second generation of calculated blood tests combined fibrosis markers by logistic regression, either indirect and direct markers like Fibrotest (6), FibroMeter (7) and Hepascore (8), or direct markers like the ELF score (9) and Fibrospect (10). The third step was liver stiffness evaluation (LSE) by ultrasonographic elastometry (11).

Finally, blood fibrosis tests were combined into *sequential algorithms* in order to increase the diagnostic accuracy and limit the rate of liver biopsy (12–16). These sequential algorithms are usually based on a stepwise diagnosis including intervals of reliable diagnosis by blood tests in a first step, followed by a biopsy for the indeterminate cases. However, clinical applicability is somewhat difficult and a biopsy is still required in 20–50% of patients.

The association of non-invasive fibrosis tests with a *synchronous combination* circumvents these limitations while improving diagnostic accuracy and resolving the discordances between tests (17). In this setting, Castera *et al.* (18) have suggested that an association of LSE and blood fibrosis tests could improve the accuracy in patients with chronic hepatitis C. However, the statistical differences between the area under the receiver operating characteristic curves (AUROCs) – the only diagnostic index used – of the fibrosis tests and their synchronous combination were not evaluated. In another recent study, Castera *et al.* (19) have suggested that an association of LSE and blood test does not improve the cirrhosis diagnosis but only a few non-recent blood tests were evaluated.

Our primary aim was to improve the diagnostic accuracy of liver fibrosis by developing synchronous combinations of blood fibrosis tests and LSE. The ensuing secondary aim was to reduce liver biopsy requirement in diagnostic algorithms in comparison with previously published algorithms.

Patients and methods

Patients

We prospectively enrolled 390 patients with chronic liver disease (CLD) hospitalized for a percutaneous liver biopsy in the University Hospitals of Angers, Bordeaux (France) and Rabat (Morocco): 194 patients from April 2004 to June 2007 at Angers (group A, derivation set), and 196 from September 2003 to April 2007 at Bordeaux and Rabat (group B, validation set). Patients with cirrhosis complications (ascites, variceal bleeding, systemic infection and hepatocellular carcinoma) were not included. Blood fibrosis tests and LSE were performed within 1 week before biopsy. All patients gave their informed consent. The study protocol conformed to the ethical guidelines of the current Declaration of Helsinki and was approved by the local Ethics committee.

Methods

Histological liver fibrosis assessment

A percutaneous liver biopsy was performed using Menghini's technique with a 1.4–1.6-mm-diameter needle. Liver fibrosis was evaluated according to Metavir staging and significant fibrosis was defined by Metavir stages $F \geq 2$. Although initially designed for chronic viral hepatitis, Metavir staging was also validated in alcoholic liver diseases (20) and used in non-alcoholic fatty liver disease (21). Liver fibrosis evaluation was performed by two senior experts with a consensus reading in group A and by a senior expert pathologist in group B. Fibrosis staging was considered as reliable when the liver specimen length was ≥ 15 mm and/or the portal tract number ≥ 8 (22).

Fibrosis blood tests

The following blood tests were performed according to published formulas or patents: APRI, FIB-4, Fibrotest, Hepascore and FibroMeter, where cause-specific scores were used (7, 23, 24). All blood assays were performed in the same laboratories of each centre. The interlaboratory reproducibility was excellent for these tests (25).

Liver stiffness evaluation

Liver stiffness evaluation (FibroScan®, EchoSens™, Paris, France) was performed by an experienced observer (> 50 LSE before this study), blinded to patient data. LSE conditions were those recommended by the manufacturer (26). LSE was stopped when 10 valid measurements were recorded. The LSE result was expressed as the median and the interquartile range (kPa) of all valid measurements performed (26).

Statistical analysis

Quantitative variables were expressed as mean $\pm$ standard deviation, unless otherwise specified. The diagnostic cut-off of fibrosis tests was calculated according to the highest Youden index (sensitivity+specificity $-$ 1).

Accuracy of fibrosis tests

The performance of fibrosis tests was mainly expressed as AUROC which were compared with the Delong test (27). The reliable individual diagnosis was determined by the classical intervals of $\geq 90\%$ negative (NPV) and positive (PPV) predictive values (precise definitions in supporting information) as well as an intermediate diagnosis, with respect to the diagnostic target, in the indeterminate zone between these classical intervals (23).

Synchronous combination of fibrosis tests

To identify the best combination of fibrosis tests for the diagnosis of significant fibrosis, we performed a forward binary logistic regression in the derivation set (group A) using blood fibrosis tests and LSE results as independent variables. Using the regression score combining the independent tests, we constructed a new diagnostic test for significant fibrosis called the *significant fibrosis index* (SF-index). This test was validated in groups B and A+B. We constructed the *cirrhosis index* (C-index) for the diagnosis of cirrhosis in the same way.

New algorithms

We determined the thresholds of 90% NPV and 90% PPV of the SF-index (precise definitions in supporting

information). A patient was considered as having no or mild fibrosis (Metavir F0/F1) when the SF-index values corresponded to NPV $\geq$ 90%, and as having significant fibrosis (Metavir $F \geq 2$) when the SF-index values corresponded to PPV $\geq$ 90% (precise definitions in supporting information). A liver biopsy was required when the SF-index value was in the indeterminate zone between the two previous thresholds. This Angers algorithm was compared with the Bordeaux algorithm (18) and with the sequential algorithm for fibrosis evaluation (SAFE) (16). We constructed another Angers algorithm designed for the diagnosis of cirrhosis by using the C-index in the same way.

Sample size

It was found to show a significant difference for the diagnosis of significant fibrosis between FibroMeter and the synchronous combination of fibrosis tests in the validation population. With α risk: 0.05, β risk: 0.20, significant fibrosis prevalence: 0.70, AUROC correlation: 0.70, and in a bilateral test, the sample size was 159 patients for the following hypothesis of AUROC: FibroMeter: 0.84, synchronous combination: 0.90. The softwares were SPSS, version 11.5.1 (SPSS Inc., Chicago, IL, USA) and SAS 9.1 (SAS Institute Inc., Cary, NC, USA).

Results

Patients

The mean age of the patients was 52.4 years; 67.9% were males, and 74.4% had significant fibrosis (Table 1). 89.5% of the patients had a liver biopsy considered as reliable. LSE failure occurred in 12 patients (3.1%). All five blood tests and LSE were available in 332 patients.

Diagnosis of significant fibrosis

Synchronous combination of fibrosis tests

Combination of non-invasive tests. The LSE AUROC for the diagnosis of significant fibrosis was significantly higher than that of Hepascore, FIB-4 and APRI, and was not significantly different from FibroMeter and Fibrotest AUROCs (Table S1). Significant fibrosis was independently diagnosed by FibroMeter at the first step and by LSE at the second step in the derivation set as well as in the validation set (Table S2). The regression score of the SF-index was: 3.6224FibroMeter+0.4408LSE result − 3.9850 (group A).

Performance of the significant fibrosis index. The SF-index AUROCs were not significantly different between groups A and B. SF-index AUROC was significantly higher than that of FibroMeter or LSE in the whole population (Table 2). The SF-index improved the classification rate among fibrosis stages (Fig. 1): its misclassification rate was significantly lower than LSE in Metavir $F \leq 1$ stages and significantly lower than FibroMeter in Metavir $F \geq 2$ stages (details not shown).

Discordances between liver stiffness evaluation and FibroMeter. Discordances were calculated in the whole population. FibroMeter and LSE were discordant in 103 (27.0%) patients, of whom 68 (66.0%) were correctly classified by the SF-index (Table S3). Finally, the SF-index correctly classified 316 (82.7%) patients and improved the correct classification (i.e. discordances between FibroMeter and LSE resolved by the SF-index) in 33 (8.6%) patients.

Table 1. Patient characteristics at inclusion

	Group			
	All ($n = 390$)	A ($n = 194$)	B ($n = 196$)	P*
Age (years)	52.4 ± 13.4	50.8 ± 12.7	53.9 ± 14.0	0.03
Male sex (%)	67.9	68.0	67.9	0.97
Cause of liver disease (%)				$< 10^{-3}$
Virus	48.7	54.1	43.4	
Alcohol	27.2	26.3	28.1	
NAFLD	4.9	9.8	0.0	
Other	19.2	9.8	28.6	
Metavir fibrosis stage (%)				$< 10^{-3}$
F0	7.2	4.1	10.2	
F1	18.5	19.6	17.3	
F2	23.1	26.3	19.9	
F3	20.3	27.3	13.3	
F4	31.0	22.7	39.3	$< 10^{-3}$
Significant fibrosis (%)	74.4	76.3	72.4	0.39
Reliable biopsy (%)	89.5	95.3	82.6	$< 10^{-3}$
IQR/LSE result < 0.21 (%)	59.4	58.5	60.3	0.73

*By t-test or χ^2 between the groups A and B.

IQR, interquartile range (kPa); LSE, liver stiffness evaluation; NAFLD, non-alcoholic fatty liver disease.

Table 2. The area under the receiver operating characteristic curve comparisons between synchronous combinations (FM+LSE indexes), FibroMeter and liver stiffness evaluation, as a function of diagnostic target and patient group

	Significant fibrosis			Cirrhosis		
Patient group	All	A	B	All	A	B
AUROC						
FM	0.834	0.839	0.843	0.835	0.822	0.839
LSE	0.867	0.889	0.850	0.923	0.931	0.922
FM+LSE index*	0.892	0.917	0.874	0.917	0.923	0.913
Comparison (*P*)†						
FM vs LSE	0.162	0.150	0.839	$< 10^{-3}$	10^{-3}	0.004
FM vs FM+LSE index*	$< 10^{-3}$	$< 10^{-3}$	0.210	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$
LSE vs FM+LSE index*	0.011	0.081	0.042	0.458	0.463	0.445

*SF-index for significant fibrosis, C-index for cirrhosis.

†By Delong test.

A, Angers; AUROC, area under the receiver operating characteristic curve; B, Bordeaux and Rabat; FM, FibroMeter; LSE, liver stiffness evaluation.

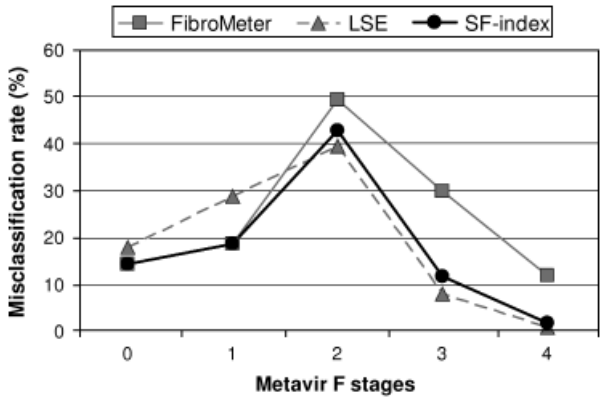

Fig. 1. Misclassification rate (%) for significant fibrosis of FibroMeter, liver stiffness evaluation (LSE) and their synchronous combination (SF-index) as a function of Metavir fibrosis stages. According to the highest Youden index, the diagnostic cut-offs used for significant fibrosis were: FibroMeter, 0.538; LSE, 6.9 kPa; SF-index, 0.753.

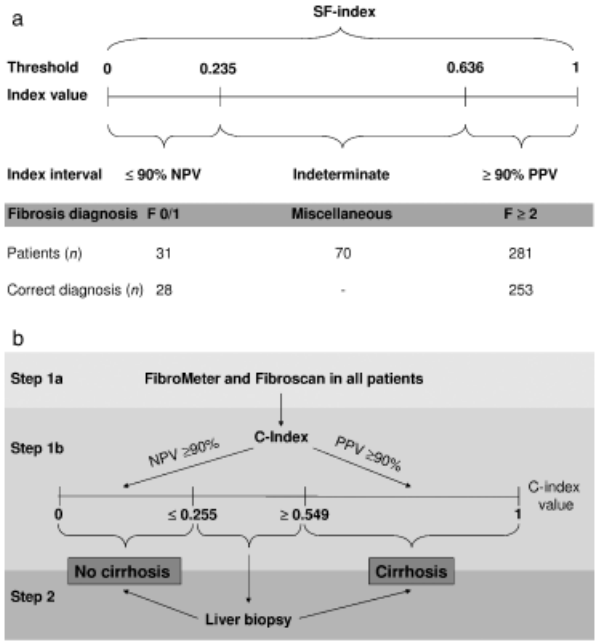

Fig. 2. The Angers algorithms designed for the diagnosis of significant fibrosis (a) or cirrhosis (b). A specific score combining FibroMeter and liver stiffness evaluation (LSE) is initially used (the SF-index for significant fibrosis or the C-index for cirrhosis), and a liver biopsy is subsequently required in the case of an indeterminate diagnosis. C-index, cirrhosis index; NPV, negative predictive value; PPV, positive predictive value; SF, significant fibrosis.

Methods reliably classifying 100% of patients

The Angers algorithm. This algorithm based on predictive values of the SF-index is described in Figure 2a. The SF-index included significantly more patients than FibroMeter or LSE in the intervals of $\geq$90% predictive values, especially in the NPV interval (Table 3). By using the SF-index with $\geq$90% predictive values and liver biopsy required in the remaining 18.3% of the patients, the Angers algorithm provided a correct diagnosis of significant fibrosis in 91.9% of the patients.

Algorithm comparison. Sequential algorithm for fibrosis evaluation had a significantly higher accuracy (97.0%) at the cost of a significantly higher biopsy rate (Table 4). The Angers algorithm required a significantly lower biopsy rate. Thus, the Angers algorithm had the best compromise with a high correct classification and a low biopsy requirement, reflected by a much lower biopsy/accuracy ratio. The accuracies of SAFE, Bordeaux and the Angers algorithms were not significantly different between patients with chronic viral hepatitis and those with other causes of CLD (detailed data not shown).

Diagnosis without liver biopsy (23). In the indeterminate interval between the thresholds of 90% predictive values of the SF-index (Fig. 2a), 92.9% of the patients had no severe fibrosis (F0: 20.0%, F1: 40.0%, and F2: 32.9%) (Fig. S1). Thus, it was possible to obtain three intervals of reliable diagnosis: F0/F1 in the $\leq$ 90% NPV interval, F1 $\pm$ 1 in the intermediate interval

Table 3. The Angers algorithms

		Rate (%) of patients included in the intervals defined by 90% predictive values			Accuracy (%)	
Diagnostic target	Fibrosis test	≥90% NPV	Indeterminate*	≥90% PPV	Fibrosis test(s)	Algorithm†
Significant fibrosis (F≥2)	FibroMeter	0.3	36.4	63.4	57.3	93.7
	LSE	0.5	28.8	70.7	64.1	92.9
	SF-index	8.1	18.3	73.6	73.6	91.9
Cirrhosis (F4)	FibroMeter	44.2	42.1	13.6	52.1	94.2
	LSE	68.3	12.6	19.1	78.8	91.4
	C-index	70.4	9.4	20.2	81.7	91.1

Rates of patients included and correctly classified by fibrosis tests in the intervals of ≥90% predictive values for the diagnosis of significant fibrosis or cirrhosis in the whole population, as a function of fibrosis test(s).

*Proportion of patients for whom the diagnosis remains uncertain (NPV and PPV $< 90\%$), thus requiring a liver biopsy. Comparison of patient rates by the McNemar test: significant fibrosis: LSE vs FibroMeter: $P = 0.006$, SF-index vs FibroMeter or LSE: $P < 10^{-3}$; cirrhosis: FibroMeter vs C-index or LSE: $P < 10^{-3}$, C-index vs LSE: $P = 0.02$.

†Based on the intervals of ≥90% predictive values of the fibrosis test(s) and liver biopsy in the subsequent interval of indeterminate diagnosis. Thus, algorithm accuracy is calculated as the sum of patients correctly classified by the fibrosis test(s) in the whole population (4th result column) and biopsy requirement (2nd result column) where accuracy is 100% by definition. Comparison of rates by the McNemar test between FibroMeter and the C-index for cirrhosis: $P = 0.04$, others: *P*: NS.

LSE, liver stiffness evaluation; NPV, negative predictive value; PPV, positive predictive value.

Table 4. Diagnostic indices of sequential algorithm for fibrosis evaluation (16), Bordeaux (18) and the Angers algorithms for the diagnosis of significant fibrosis or cirrhosis

	Significant fibrosis (F≥2)			Cirrhosis (F4)		
	SAFE	Bordeaux	Angers	SAFE	Bordeaux	Angers
DA (%)*	97.0	90.4	91.9	79.8	94.3	91.0
Se (%)	100.0	90.6	98.8	44.0	89.0	74.7
Spe (%)	88.6	89.8	72.7	93.4	96.3	97.1
NPV (%)	100.0	77.5	95.5	81.5	95.9	91.1
PPV (%)	96.1	96.1	90.9	71.4	90.0	90.7
+LR	8.80	8.86	3.62	6.62	23.84	25.73
−LR	0.0	0.11	0.02	0.60	0.11	0.26
DOR	NA	84.3	214.3	11.0	208.8	98.8
Biopsy (%)†	55.7	28.6	20.2	6.6	25.3	9.3
Biopsy/DA	0.57	0.32	0.22	0.08	0.27	0.10

Three hundred and thirty-two patients having Fibrotest, FibroMeter, APRI and LSE were available.

The prevalence of significant fibrosis was 73.5% and that of cirrhosis was: 27.4%.

*Comparison of algorithm diagnostic accuracies in the whole population, by the McNemar test: significant fibrosis: SAFE vs Bordeaux or Angers, $P < 10^{-3}$; Bordeaux vs Angers, $P = 0.50$; cirrhosis: SAFE vs Bordeaux or Angers, $P < 10^{-3}$; Bordeaux vs Angers: $P = 0.04$.

†Comparison of rates of biopsy required by the algorithms in the whole population, by the McNemar test: significant fibrosis: SAFE vs Bordeaux or Angers, $P < 10^{-3}$; Bordeaux vs Angers, $P = 0.02$; cirrhosis: Bordeaux vs SAFE or Angers, $P < 10^{-3}$; SAFE vs Angers, $P = 0.26$.

Biopsy, rate of liver biopsy required in the algorithm; biopsy/DA, ratio (rate of liver biopsy required)/(diagnostic accuracy); DA, diagnostic accuracy; DOR, diagnostic odds ratio; −LR: negative likelihood ratio; +LR: positive likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; SAFE, sequential algorithm for fibrosis evaluation; Se, sensitivity; Spe, specificity.

(correct classification: 92.9%) and $F \geq 2$ (F3 ± 1) in the ≥90% PPV interval. This classification correctly classified 90.6% of the patients without any biopsy performed.

Sensitivity analysis

By multivariate analysis, the accuracy of the SF-index was not associated with the biopsy length, the ratio interquartile range/LSE result or body mass index. In biopsy ≥25 mm, the SF-index still had a significantly higher AUROC (0.895) than FibroMeter (0.816, $P < 10^{-3}$) or LSE (0.860, $P = 0.03$). The SF-index was not dependent on the putative reliability of LSE: in interquartile range/LSE result ratio < 0.21 and ≥ 0.21 (28), AUROC were 0.881 and 0.906 ($P = 0.51$) respectively.

Diagnosis of cirrhosis

Synchronous combination of fibrosis tests

Combination of non-invasive tests. Liver stiffness evaluation had a significantly higher AUROC than the blood tests for the diagnosis of cirrhosis (Table S1). The most accurate combination of fibrosis tests was LSE+FibroMeter (Table S2). The regression score of the C-index was: 0.1162LSE result+1.9714FibroMeter − 4.6616 (group A). In each group tested, the C-index had a significantly higher AUROC than FibroMeter, but the difference with LSE was not significant (Table 2).

Discordances between liver stiffness evaluation and FibroMeter. FibroMeter and LSE were discordant in

79 patients, of whom 54 (68.4%) were correctly classified by the C-index (Table S3). Finally, the C-index correctly classified 329 (86.1%) patients and improved correct classification in 29 (7.6%) patients.

Methods reliably classifying 100% of the patients

The Angers algorithm. This algorithm, based on predictive values of the C-index, is described in Figure 2b. The C-index included significantly more patients (90.6%) than LSE or FibroMeter in the intervals of $\geq$90% predictive values (Table 3). By using a liver biopsy for the remaining 9.4% of the patients, the Angers algorithm provided a correct diagnosis of cirrhosis in 91.1% of the patients.

Algorithm comparison. The Bordeaux algorithm had a significantly higher accuracy (Table 4). However, the Angers algorithm had a significantly lower biopsy rate. Thus, again, the Angers algorithm had the best compromise with a high correct classification and a low biopsy requirement.

Diagnosis without liver biopsy (23). In the indeterminate interval between the thresholds of 90% predictive values of the C-index (Fig. 2b), 91.6% had significant fibrosis ($F \geq 2$) (Fig. S1). Thus, it was possible to obtain three intervals of a reliable diagnosis: no cirrhosis ($F \leq 3$), significant fibrosis (F3 $\pm$ 1) and cirrhosis (F4) that correctly classified 90.3% of patients without any biopsy performed.

Discussion

Study design

This study compared for the first time the diagnostic accuracy of LSE and five blood tests, and their synchronous combination, in a large population of patients with various causes of CLD. The liver fibrosis reference in our study was the Metavir staging on liver biopsy. This reference is hampered by inter-observer variability (29) and sampling variability (30). Nevertheless, histological reading was performed by expert pathologists who have been shown to provide the best inter-observer reproducibility (29). Moreover, SF-index accuracy was not independently influenced by liver specimen length. In addition, in liver specimens $\geq$25 mm displaying the highest reliability for liver fibrosis evaluation (30), the SF-index still had a significantly higher accuracy than FibroMeter and LSE.

A recent study has suggested that LSE is accurate when the ratio interquartile range/LSE result is < 0.21 (28). Although 59.4% of our patients had this ratio < 0.21, all LSE were included in the statistical analysis according to the 'intention-to-diagnose' principle (31). In addition, discarding the so-called inaccurate LSE would not have been within the scope of a combination study. Finally, SF-index accuracy was not independently influenced by this ratio.

Accuracy of blood tests and liver stiffness evaluation

The AUROC of LSE, FibroMeter and Fibrotest were not statistically different for the diagnosis of significant fibrosis. However, Figure 1 shows that the best-performing blood test had the lowest misclassification rate in F0–F1 stages, whereas LSE had the lowest misclassification rate in $F \geq 2$ stages, especially in F3 and F4. This suggests that these tests are complementary. We thus hypothesized that their synchronous combination could cumulate their advantages, erase their drawbacks and thus significantly increase their single diagnostic accuracy for liver fibrosis.

Accuracy gain by the synchronous combination

Our results show that a synchronous combination blood test+LSE increases the AUROC for the diagnosis of significant fibrosis (Table 2). However, AUROC is a global index of performance that is not meaningful in clinical practice. Indeed, physicians want to know whether a fibrosis test result is reliable in a patient. Therefore, we have developed the comparison of patient rates with high predictive values (23). The present study clearly shows that the synchronous combination blood test+LSE (SF-index) significantly improves the patient rate with $\geq$90% predictive value for significant fibrosis (Table 3) in a higher proportion than could have suggested the AUROC comparison. The indeterminate zone of fibrosis tests was thus reduced: the SF-index provided a highly significantly lower rate of biopsy than the blood test or LSE. Finally, the synchronous combination provided two other advantages: the SF-index offered the lowest misclassification rate provided by each single test in each fibrosis stage (the blood test in F0/F1 and LSE in $F \geq 2$, see Fig. 1), and the SF-index resolved 66% of discordant cases between the blood test and LSE. Alternatively, by identifying an intermediate diagnosis (F1 $\pm$ 1) in the indeterminate zone of the SF-index, this fully non-invasive diagnosis classified 100% of the population with 90.6% accuracy.

The C-index produced no significant increase in AUROC for the diagnosis of cirrhosis compared with LSE alone. However, the combination LSE+blood test also provided two advantages for the diagnosis of cirrhosis: the patient rate with $\geq$90% predictive values was significantly higher than with LSE alone (Table 3), thus resulting in a very low rate of biopsy required (9%), and it resolved 68.4% of discordant cases between LSE and blood test. Alternatively, by identifying an intermediate diagnosis ($F \geq 2$) in the indeterminate zone of the C-index, this fully non-invasive diagnosis classified 100% of patients with 90.3% accuracy.

Algorithms for liver fibrosis diagnosis

Significant fibrosis

In our study, the accuracies of Bordeaux and SAFE algorithms were similar to those published previously (16, 18, 32), thus providing an independent external validation of these algorithms. Moreover, the accuracies of the SAFE, Bordeaux and Angers algorithms were not significantly different between patients with chronic viral hepatitis and those with other CLD. Because SAFE and Bordeaux algorithms were elaborated in chronic viral C hepatitis, this suggests that these algorithms could be extended to other causes of CLD.

Our results show that SAFE had a significantly higher diagnostic accuracy for significant fibrosis than the Bordeaux and Angers algorithms. However, a sequential combination of fibrosis tests has the major disadvantage of leaving a large final indeterminate zone responsible for a high rate of biopsy required, which is unsuitable in a screening setting (13–16). In our study, SAFE selected two and three times more patients for a biopsy than the Bordeaux and Angers algorithms respectively. In this respect, the Angers algorithm offered the best compromise between a high diagnostic accuracy (91.9%) and the lowest biopsy rate (20.2%). Finally, it should be noticed that a part of the apparent misclassifications of patients provided by an algorithm are in fact attributable to the misclassification of liver biopsy used as the reference. The fully non-invasive diagnosis based on three reliable intervals (23) might compensate for this previous uncertainty.

At first glance, it seems that SAFE would be the less expensive strategy: APRI is used as the first-line test with virtually no cost, and Fibrotest is only performed at the second step in a subgroup of patients. However, this apparently cost-effective strategy is dramatically counteracted by the high biopsy rate required by SAFE, liver biopsy being an expensive procedure in some countries (approximately 1000 euros for direct costs in France).

Cirrhosis

The Angers algorithm also provided the best compromise between a high diagnostic accuracy for cirrhosis (91.0%) and a very low rate of biopsy (9.3%). SAFE also had a low biopsy rate but an unacceptable significantly lower diagnostic accuracy (79.8%) than the other two algorithms, which was because of causes other than viral hepatitis (data not shown).

How is this integrated into clinical practice?

In a first step, the biological and LSE values are recorded and computerized (on a website for example) to obtain the combined index. A second step, based on a liver biopsy, may be required in a minority of patients. In cirrhosis, this simple two-step diagnosis had an accuracy of 91% and a biopsy was required in less than 10% of the patients (Fig. 2b). Finally, the practitioner has to evaluate the clinical reliability of the results, especially those of fibrosis test results.

Limitations of this study

The present study has two limitations. First, the high prevalence of significant fibrosis (72%) as it influences diagnostic indices like accuracy (33) and predictive values. LSE was thus favoured in this study because it had a higher accuracy in $F \geq 2$ stages. However, this limitation can be easily circumvented with the misclassification profile, which shows the rate of misclassified patients in each fibrosis stage (Fig. 1).

The second limitation is the inclusion of patients with various CLD causes. This was mainly because of the sample size required. However, there was no significant difference in the diagnostic accuracy of the Angers algorithm between chronic viral hepatitis and other causes. Nevertheless, further studies should explore the possibility of increased accuracy via CLD cause-specific combination indexes.

In conclusion, the synchronous combination blood test+LSE significantly increases the diagnostic accuracy for significant fibrosis compared with a single test, resolves a majority of discordant results between non-invasive tests and improves the reliable individual diagnosis for significant fibrosis and cirrhosis while reducing liver biopsy requirement. Alternatively, one might use a fully non-invasive diagnosis, where the liver biopsy is replaced by an intermediate diagnosis with respect to the diagnostic target, at the cost of only a 1% decrease in accuracy.

Acknowledgements

French and Morocco research departments for grant (PAI Volubilis).

Potential conflict of interest: Paul Calès, Frederic Oberti and Isabelle Fouchard-Hubert have stock ownership in BioLiveScale Inc. that has a license for FibroMeters from Angers University.

References

1. Afdhal NH, Nunes D. Evaluation of liver fibrosis: a concise review. *Am J Gastroenterol* 2004; **99**: 1160–74.
2. Croquet V, Vuillemin E, Ternisien C, *et al*. Prothrombin index is an indirect marker of severe liver fibrosis. *Eur J Gastroenterol Hepatol* 2002; **14**: 1133–41.
3. Oberti F, Valsesia E, Pilette C, *et al*. Noninvasive diagnosis of hepatic fibrosis or cirrhosis. *Gastroenterology* 1997; **113**: 1609–16.
4. Wai CT, Greenson JK, Fontana RJ, *et al*. A simple non-invasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003; **38**: 518–26.
5. Sterling RK, Lissen E, Clumeck N, *et al*. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006; **43**: 1317–25.

6. Imbert-Bismut F, Ratziu V, Pieroni L, *et al.* Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. *Lancet* 2001; **357**: 1069–75.
7. Cales P, Oberti F, Michalak S, *et al.* A novel panel of blood markers to assess the degree of liver fibrosis. *Hepatology* 2005; **42**: 1373–81.
8. Adams LA, Bulsara M, Rossi E, *et al.* Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. *Clin Chem* 2005; **51**: 1867–73.
9. Rosenberg WM, Voelker M, Thiel R, *et al.* Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology* 2004; **127**: 1704–13.
10. Patel K, Gordon SC, Jacobson I, *et al.* Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. *J Hepatol* 2004; **41**: 935–42.
11. Sandrin L, Fourquet B, Hasquenoph JM, *et al.* Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Biol* 2003; **29**: 1705–13.
12. Sebastiani G, Vario A, Guido M, *et al.* Stepwise combination algorithms of non-invasive markers to diagnose significant fibrosis in chronic hepatitis C. *J Hepatol* 2006; **44**: 686–93.
13. Bourliere M, Penaranda G, Renou C, *et al.* Validation and comparison of indexes for fibrosis and cirrhosis prediction in chronic hepatitis C patients: proposal for a pragmatic approach classification without liver biopsies. *J Viral Hepat* 2006; **13**: 659–70.
14. Sebastiani G, Vario A, Guido M, *et al.* Sequential algorithms combining non-invasive markers and biopsy for the assessment of liver fibrosis in chronic hepatitis B. *World J Gastroenterol* 2007; **13**: 525–31.
15. Bourliere M, Penaranda G, Ouzan D, *et al.* Optimized stepwise combination algorithms of non-invasive liver fibrosis scores including Hepascore in hepatitis C virus patients. *Aliment Pharmacol Ther* 2008; **28**: 458–67.
16. Sebastiani G, Halfon P, Castera L, *et al.* SAFE biopsy: a validated method for large-scale staging of liver fibrosis in chronic hepatitis C. *Hepatology* 2009; **49**: 1821–7.
17. Boursier J, Cales P. Combination of fibrosis tests: sequential or synchronous? *Hepatology* 2009; **50**: 656–7.
18. 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**: 343–50.
19. 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.
20. Michalak S, Rousselet M C, Bedossa P, *et al.* Respective roles of porto-septal fibrosis and centrilobular fibrosis in alcoholic liver disease. *J Pathol* 2003; **201**: 55–62.
21. Laine F, Bendavid C, Moirand R, *et al.* Prediction of liver fibrosis in patients with features of the metabolic syndrome regardless of alcohol consumption. *Hepatology* 2004; **39**: 1639–46.
22. Nousbaum JB, Cadranel JF, Bonnemaison G, *et al.* Clinical practice guidelines on the use of liver biopsy. *Gastroenterol Clin Biol* 2002; **26**: 848–78.
23. Cales P, De Ledinghen V, Halfon P, *et al.* Evaluating accuracy and increasing the reliable diagnosis rate of blood tests for liver fibrosis in chronic hepatitis C. *Liver Int* 2008; **28**: 1352–62.
24. Cales P, Laine F, Boursier J, *et al.* Comparison of blood tests for liver fibrosis specific or not to NAFLD. *J Hepatol* 2009; **50**: 165–73.
25. Cales P, Veillon P, Konate A, *et al.* Reproducibility of blood tests of liver fibrosis in clinical practice. *Clin Biochem* 2008; **41**: 10–8.
26. Boursier J, Konate A, Gorea G, *et al.* Reproducibility of liver stiffness measurement by ultrasonographic elastometry. *Clin Gastroenterol Hepatol* 2008; **6**: 1263–9.
27. Delong ER, Delong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44**: 837–45.
28. 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**: 1083–9.
29. Rousselet MC, Michalak S, Dupre F, *et al.* Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology* 2005; **41**: 257–64.
30. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003; **38**: 1449–57.
31. Bossuyt PM, Reitsma JB, Bruns DE, *et al.* The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 2003; **49**: 7–18.
32. Castera L, Sebastiani G, Le Bail B, *et al.* Prospective comparison of two algorithms combining non invasive tests for staging of liver fibrosis in chronic hepatitis C. *Hepatology* 2007; **46**: 320A.
33. Poynard T, Halfon P, Castera L, *et al.* Standardization of ROC curve areas for diagnostic evaluation of liver fibrosis markers based on prevalences of fibrosis stages. *Clin Chem* 2007; **53**: 1615–22.

Supporting information

Additional supporting information may be found in the online version of this article:

Figure S1. Sequential algorithm.

Figure S2. Intervals of $\geq 90\%$ negative (NPV) and positive (PPV) predictive values of SF-score, for the diagnosis of significant fibrosis (Metavir F$\geq$2). NPV is $\geq 90\%$ in the 8.1% of patient with a score ≤ 0.235; PPV is $\geq 90\%$ in the 73.6% of patients with a score ≥ 0.636.

Figure S3. Diagnosis based on high predictive values for significant fibrosis (3a) or cirrhosis (3b) and intermediate fibrosis stages in the remaining patients.

Table S1. AUROCs of blood tests and liver stiffness evaluation (LSE) as a function of diagnostic target.

Table S2. Combination by logistic regression. Fibrosis tests independently associated with significant fibrosis or cirrhosis, as a function of patient group.

Table S3. Discordances. Impact of FM+LSE index on discordances between FibroMeter (FM) and liver stiffness evaluation (LSE) for the diagnosis of significant fibrosis ($F \geq 2$) or cirrhosis (F4) in the whole population.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.