

Comparison of Eight Diagnostic Algorithms for Liver Fibrosis in Hepatitis C: New Algorithms Are More Precise and Entirely Noninvasive

Jérôme Boursier,^{1,2} Victor de Ledinghen,^{3,4} Jean-Pierre Zarski,^{5,6} Isabelle Fouchard-Hubert,^{1,2} Yves Gallois,^{2,7} Frédéric Oberti,^{1,2} Paul Calès^{1,2} and multicentric groups from SNIFF 32, VINDIAG 7, and ANRS/HC/EP23 FIBROSTAR studies

The sequential algorithm for fibrosis evaluation (SAFE) and the Bordeaux algorithm (BA), which cross-check FibroTest with the aspartate aminotransferase-to-platelet ratio index (APRI) or FibroScan, are very accurate but provide only a binary diagnosis of significant fibrosis (SAFE or BA for Metavir F \geq 2) or cirrhosis (SAFE or BA for F4). Therefore, in clinical practice, physicians have to apply the algorithm for F \geq 2, and then, when needed, the algorithm for F4 ("successive algorithms"). We aimed to evaluate successive SAFE, successive BA, and a new, noninvasive, detailed classification of fibrosis. The study included 1785 patients with chronic hepatitis C, liver biopsy, blood fibrosis tests, and FibroScan (the latter in 729 patients). The most accurate synchronous combination of FibroScan with a blood test (FibroMeter) provided a new detailed (six classes) classification (FM+FS). Successive SAFE had a significantly ($P < 10^{-3}$) lower diagnostic accuracy (87.3%) than individual SAFE for F \geq 2 (94.6%) or SAFE for F4 (89.5%), and required significantly more biopsies (70.8% versus 64.0% or 6.4%, respectively, $P < 10^{-3}$). Similarly, successive BA had significantly ($P \leq 10^{-3}$) lower diagnostic accuracy (84.7%) than individual BA for F \geq 2 (88.3%) or BA for F4 (94.2%), and required significantly more biopsies (49.8% versus 34.6% or 24.6%, respectively, $P < 10^{-3}$). The diagnostic accuracy of the FM+FS classification (86.7%) was not significantly different from those of successive SAFE or BA. However, this new classification required no biopsy. **Conclusion:** SAFE and BA for significant fibrosis or cirrhosis are very accurate. However, their successive use induces a significant decrease in diagnostic accuracy and a significant increase in required liver biopsy. A new fibrosis classification that synchronously combines two fibrosis tests was as accurate as successive SAFE or BA, while providing an entirely noninvasive (0% liver biopsy) and more precise (six versus two or three fibrosis classes) fibrosis diagnosis. (HEPATOLOGY 2012;55:58-67)

Several fibrosis algorithms combining different fibrosis tests have been proposed to improve the accuracy of the noninvasive diagnosis of liver fibrosis in chronic hepatitis C.¹⁻⁵ These decision-making algorithms were developed to provide an accurate diagnosis of liver fibrosis and limit liver biopsy to indeterminate cases. They use either two blood tests in a sequential procedure, as in the sequential algorithm for

fibrosis evaluation (SAFE),⁶ or are based on agreement between a blood test and FibroScan (Echosens, Paris, France) results, as in the Bordeaux algorithm (BA).¹

Although the accuracy of SAFE and BA has been shown to be excellent for the diagnosis of significant fibrosis or cirrhosis,^{1,2,6-8} they have some limitations in clinical practice. First, SAFE uses the aspartate aminotransferase-to-platelet ratio index (APRI) as a first-line

Abbreviations: BA, Bordeaux algorithm; SAFE, sequential algorithm for fibrosis evaluation.

From the ¹Department of Hepatogastroenterology, University Hospital, Angers, France; ²HIFIH Laboratory, UPRES 3859, Institut Fédératif de Recherche (IFR) 132, University of Angers, Pôle de Recherche et d'Enseignement Supérieur Université Nantes Angers Le Mans (PRES UNAM), France; ³Department of Hepatogastroenterology, Haut-Lévêque University Hospital, Pessac, France; ⁴Institut National de la Santé et de la Recherche Médicale (INSERM) U889, Victor Segalen University, Bordeaux, France; ⁵Department of Liver-Gastroenterology, University Hospital, Grenoble, France; ⁶Institut National de la Santé et de la Recherche Médicale (INSERM)/UJF U823, IAPC, IAB, Grenoble, France; and ⁷Department of Biochemistry, University Hospital, Angers, France.

Received February 21, 2011; accepted August 23, 2011.

Supported by the Program Hospitalier de Recherche Clinique (PHRC) of the French Department of Health in 1994 and 2002 for SNIFF 32 and the Agence Nationale de Recherche sur le Sida et les Hépatites (ANRS) for FIBROSTAR HC/EP23.

fibrosis test, then FibroTest as a second-line test, and, if necessary, liver biopsy when the diagnosis remains undetermined. This implies several diagnostic steps.⁹ Second, the rate of required liver biopsy in SAFE and BA ranges from 30% to 50% for the diagnosis of significant fibrosis and from 20%-30% for cirrhosis.^{6,7} These rates seem inconsistent with a “noninvasive” diagnostic procedure for liver fibrosis screening. Third, SAFE was developed for a binary diagnosis of significant fibrosis or cirrhosis, which is insufficient for patient management in clinical practice. Indeed, a noninvasive diagnosis of significant fibrosis could indicate either moderate/severe fibrosis or cirrhosis. Thus, to achieve an accurate diagnosis, physicians have to use the SAFE for significant fibrosis first, and then, if significant fibrosis is diagnosed, the SAFE for cirrhosis. This adds a diagnostic step and increases the rate of misclassified patients and the rate of required liver biopsy. Finally, the BA was presented in the pivotal study as a three-diagnostic-class algorithm,¹ but further evaluation focused only on the binary diagnosis of significant fibrosis or cirrhosis.⁷

We have developed several statistical techniques to improve the noninvasive diagnosis of liver fibrosis. These include blood tests adapted to a diagnostic target,¹⁰ synchronous combinations of fibrosis tests^{8,11} to improve diagnostic accuracy, and reliable diagnosis intervals for fibrosis tests to improve diagnostic precision.^{12,13} Finally, a synchronous combination of FibroScan and FibroMeter using these methods in a one-step procedure resulted in an accurate noninvasive classification of fibrosis.¹⁴ This classification provided a precise diagnosis (six diagnostic classes), with robust and high diagnostic accuracy, and eliminated the need for liver biopsy.

The aim of the present study was to evaluate the accuracy of SAFE and BA for the noninvasive diagnosis of liver fibrosis in clinical practice and compare them with our new noninvasive classification of fibrosis, which synchronously combines fibrosis tests.

Patients and Methods

Patients. We pooled the populations of three published studies, SNIFF 32,¹³ VINDIAG 7,¹⁴ and

FIBROSTAR ANRS/HC/EP23,¹⁵ all of which had very similar inclusion and exclusion criteria. Patients were included if they had chronic hepatitis C, defined as both positive anti-hepatitis C virus antibodies and hepatitis C virus RNA in serum. Exclusion criteria were other causes of chronic hepatitis (hepatitis B or HIV coinfection, alcohol consumption > 30 g/day in men or >20 g/day in women in the 5 years before inclusion, hemochromatosis, or autoimmune hepatitis), cirrhosis complications (ascites, variceal bleeding, systemic infection, hepatocellular carcinoma), and antifibrotic treatment in the preceding 6 months. Patients were included from nine centers for SNIFF 32, three centers for VINDIAG 7, and 19 centers for FIBROSTAR, all located in France. Patients included in both VINDIAG 7 and FIBROSTAR were excluded from the FIBROSTAR population for the statistical analysis of the present study. All patients gave informed consent. Study protocols conformed to the ethical guidelines of the current Declaration of Helsinki and received approval from local ethics committees.

Liver Biopsy. Liver fibrosis was evaluated according to Metavir fibrosis (F) staging. Significant fibrosis was defined as Metavir F \geq 2, severe fibrosis as Metavir F \geq 3, and cirrhosis as Metavir F4. Histological liver fibrosis evaluation was performed by blinded senior pathologists in each center. In the FIBROSTAR study, liver fibrosis was centrally evaluated by two senior experts with a consensus reading in cases of discordance. All pathologists involved in the three studies were hepatology specialists. Histological results were used as reference for the evaluation of noninvasive tests.

Blood Fibrosis Tests. Fasting blood samples were collected immediately before or no more than 3 months after liver biopsy. Blood samples were processed independently in each center, except for hyaluronic acid, α 2-macroglobulin, haptoglobin, and apolipoprotein A1, which were tested centrally in the FIBROSTAR study. Fibrotest,¹ FibroMeter^{2G},¹⁶ and APRI¹⁷ were calculated according to published or patented formulas. We have demonstrated the excellent interlaboratory reproducibility of these tests.¹⁸

Liver Stiffness Evaluation. FibroScan was available in the VINDIAG 7 and FIBROSTAR studies.

Address reprint requests to: Jérôme Boursier, M.D., Service d'Hépatogastroentérologie, CHU, 49933 Angers Cedex 09, France. E-mail: JeBoursier@chu-angers.fr; fax: (33) 2 41 35 41 19.

Copyright © 2011 by the American Association for the Study of Liver Diseases.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep.24654

Potential conflict of interest: Dr. Calès owns stock in and consults for Biolivescale.

Dr. Zarski consults for and is on the speakers' bureau of Roche, Merck, Bristol-Myers Squibb, Gilead, Janssen, and Novartis.

Drs. Oberti and Fouchard-Hubert own stock in Biolivescale.

Additional Supporting Information may be found in the online version of this article.

Table 1. Main Characteristics of Study Populations

	Populations*			P Value†
	All	#1	#2	
Patients (n)	1785	1056	729	—
Age, years	48.0 ± 12.4	45.6 ± 12.5	51.7 ± 11.2	<10 ⁻³
Male sex (%)	60.2	59.5	61.3	0.426
Biopsy length (mm)	23 ± 9	21 ± 8	25 ± 9	<10 ⁻³
Biopsy length ≥ 15 mm (%)	79.0	75.3	84.3	<10 ⁻³
Metavir F (%)				0.016
0	4.2	4.4	4.0	
1	41.1	43.5	37.7	
2	26.5	27.0	25.8	
3	15.5	14.0	17.6	
4	12.7	11.2	15.0	
Metavir A (%)				0.436
0	2.0	1.6	2.3	
1	41.3	42.8	39.7	
2	45.7	44.3	47.1	
3	11.1	11.3	10.9	
FibroScan result (kPa)	—	—	10.0 ± 7.9	—
IQR/median ratio	—	—	0.21 ± 0.19	—
IQR median ratio < 0.21 (%)	—	—	67.8	—
Prothrombin index (%)	93 ± 10	93 ± 10	94 ± 9	0.005
Platelets (G/L)	212 ± 68	212 ± 70	211 ± 66	0.943
AST (IU/L)	62 ± 50	60 ± 50	66 ± 51	0.008
ALT (IU/L)	94 ± 87	94 ± 90	94 ± 82	0.928
Bilirubin (μmol/L)	12 ± 14	11 ± 12	13 ± 16	0.028

*Population #1 corresponds to SNIFF 32 cohort (n = 1056) with liver biopsy and blood fibrosis tests available. Population #2 corresponds to pooled VINDIAG 7 (n = 349) and FIBROSTAR (n = 380) cohorts with liver biopsy, blood fibrosis tests, and FibroScan available.

†Between populations #1 and #2.

FibroScan examinations were performed under fasting conditions by an experienced observer (>50 examinations before the study), blinded for patient data. Examination conditions were those recommended by the manufacturer.¹⁹ FibroScan examinations were stopped when 10 valid measurements were recorded. Results (in kilopascals) were expressed as the median of all valid measurements. A FibroScan result was considered reliable when the interquartile range (IQR)/median ratio (IQR/M) was <0.21.²⁰

Fibrosis Algorithms

Characteristics of the eight fibrosis algorithms evaluated in the present study are detailed in the glossary and summarized in Supporting Table 1 in the Supporting Material.

Decision-Making Algorithms. *SAFE.* SAFE for the diagnosis of significant fibrosis (SAFE for F ≥ 2; Fig. 1A), SAFE for the diagnosis of cirrhosis (SAFE for F4; Fig. 1B), and SAFE for the simultaneous diagnosis of significant fibrosis and cirrhosis (SAFE for F ≥ 2 and F4. Supporting Fig. 3) were determined according to data published by Sebastiani et al.⁶

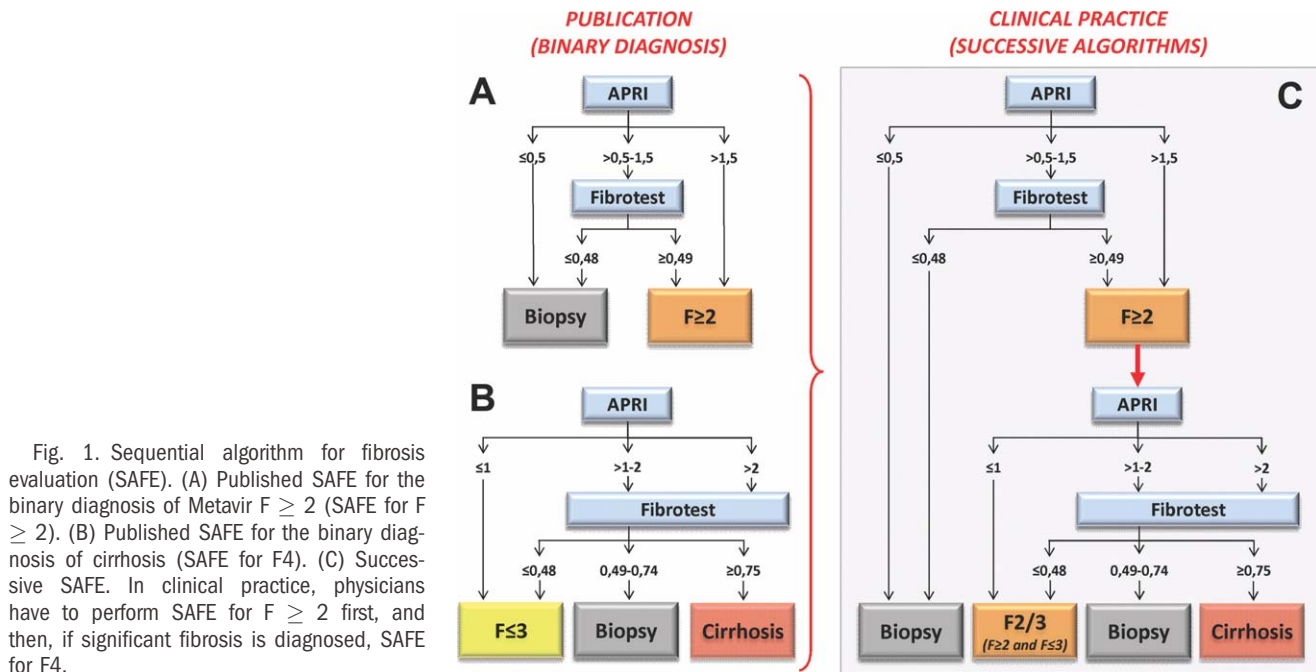
Bordeaux Algorithm. BA for the diagnosis of significant fibrosis (BA for F ≥ 2; Fig. 2A) and BA for the diagnosis of cirrhosis (BA for F4; Fig. 2B) were

determined according to data published by Castera et al.⁷

Successive Algorithms An algorithm constructed for a binary diagnosis of liver fibrosis, such as SAFE or BA, provides only limited data for the management of patients in clinical practice. Indeed, when the noninvasive diagnosis provided by the algorithm specific to significant fibrosis is F ≥ 2, the physician has to apply the cirrhosis-specific algorithm in a second step to determine whether the patient has cirrhosis. We used the term “successive algorithms” to describe this consecutive use of algorithms in clinical practice. In the present study, we evaluated: Successive SAFE, which corresponds to the use of SAFE for F ≥ 2 followed by SAFE for F4 when necessary (Fig. 1C); and Successive Bordeaux algorithms (Successive BA), which corresponds to the use of BA for F ≥ 2 followed by BA for F4 when necessary. Successive BA presents as a three-diagnostic-class algorithm (Fig. 2C).

New Noninvasive Classification of Fibrosis

The new noninvasive classification of fibrosis was derived from the synchronous combination of FibroMeter and FibroScan results as described.¹⁴ The method is detailed in the glossary in the Supporting Information and summarized in Supporting Fig. 4. Briefly,



two fibrosis indexes combining FibroMeter and FibroScan are derived by binary logistic regression: the clinically significant fibrosis (CSF) index (diagnostic target: Metavir F ≥ 2) and the severe fibrosis (SF) index (diagnostic target: Metavir F ≥ 3). The reliable diagnosis intervals (see the glossary in the Supporting Information for a precise definition) of these two indexes are then determined according to a method that has been described.¹³ Finally, the association of the reliable diagnoses from the CSF index and SF index determines the FM+FS classification, which includes six diagnostic classes of fibrosis stages (F0/1, F1/2, F2 \pm 1,

F2/3, F3 \pm 1, and F4) and eliminates the need for required liver biopsy.

Statistical Analysis

All FibroScan examinations, reliable or not, were included in the initial statistical analysis. Then sensitivity analysis was performed in patients with reliable FibroScan results.

SAFE was evaluated in the SNIFF32 (called population #1 in the present study), VINDIAG 7 (population #2a), and FIBROSTAR (population #2b) cohorts. Because FibroScan was not available in the SNIFF32

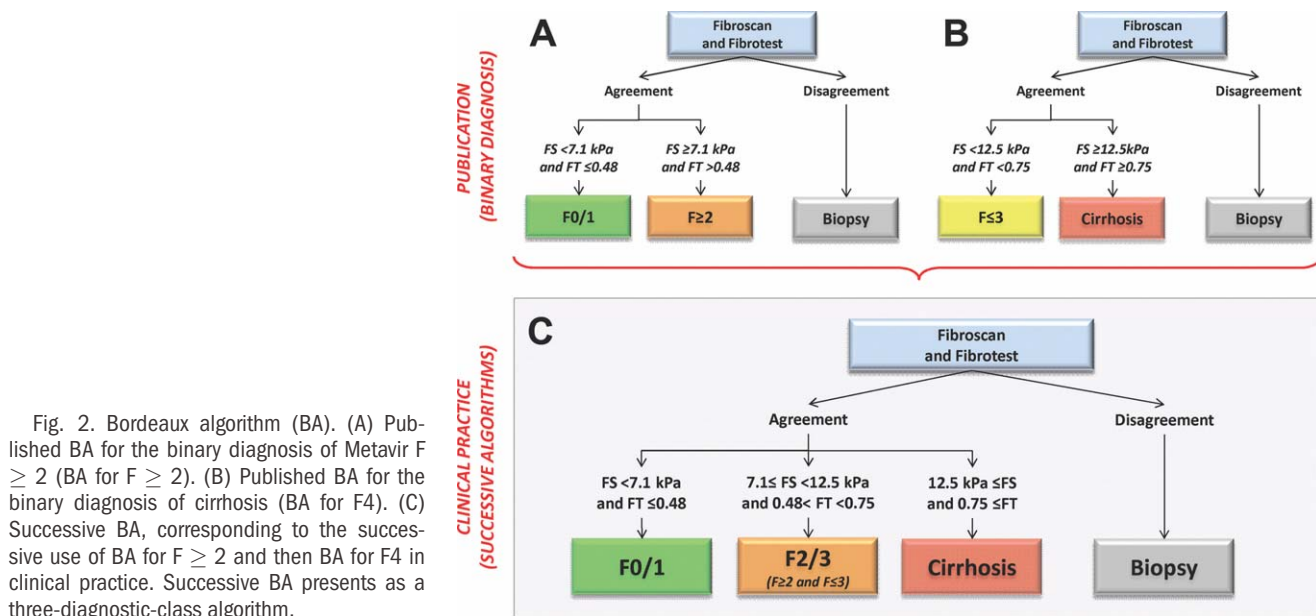


Table 2. Accuracy of SAFE and Bordeaux Algorithm (BA) Constructed for the Binary Diagnosis of Significant Fibrosis (Metavir F \geq 2) or Cirrhosis (Metavir F4) as a Function of Study Population

Diagnostic Target	F \geq 2				F4			
	SAFE			BA #2	SAFE			BA #2
	All	#1	#2		All	#1	#2	
DA (%)	94.6	96.0	92.5	88.3*	89.5	90.7	87.6	94.2†
Sensitivity (%)	100.0	100.0	100.0	88.0	61.5	60.7	62.3	86.5
Specificity (%)	88.2	91.8	82.1	88.7	93.4	94.2	92.1	95.4
NPV (%)	100.0	100.0	100.0	84.5	94.6	95.4	93.3	97.7
PPV (%)	91.0	92.9	88.7	91.4	56.2	54.6	57.9	76.1
+LR	8.46	12.17	5.58	7.78	9.25	10.44	7.86	18.99
−LR	0.0	0.0	0.0	0.14	0.41	0.42	0.41	0.14
OR	NA	NA	NA	57.6	22.4	25.0	19.2	133.8
LB (%)	64.0	68.8	57.0	34.6†	6.4	6.2	6.7	24.6†
DA without LB (%)	85.1	87.3	82.7	82.1	88.7	90.1	86.7	92.2

Abbreviations: DA, diagnostic accuracy (rate of correctly classified patients); NPV, negative predictive value; PPV, positive predictive value; +LR, positive likelihood ratio; −LR, negative likelihood ratio; OR, odds ratio; NA, not available (null −LR); LB, rate of liver biopsy required by the algorithm; DA without LB, rate of correctly classified patients by noninvasive tests in the subgroup of patients for whom liver biopsy was not required.

* $P = 0.010$ versus SAFE in population #2.

† $P < 10^{-3}$ versus SAFE in population #2.

study, BA was evaluated only in the VINDIAG 7 and FIBROSTAR populations (i.e., population #2). The VINDIAG 7 study provided the exploratory population of the FM+FS classification,¹⁴ which was then validated in the FIBROSTAR population.

Performance for the evaluated fibrosis algorithms was expressed as the rate of correctly classified patients according to liver biopsy results, the rate of required liver biopsy, and the performance profile as described²¹; comparisons were done with the paired McNemar test. Statistical software was SPSS, version 11.5 (SPSS Inc., Chicago, IL).

Results

Patients

The characteristics of the patients included in the three populations have been described^{13–15} and are summarized in Table 1. More detailed characteristics are presented in Supporting Table 2. The prevalence of significant fibrosis (Metavir F \geq 2) was 52.2% in population #1, 67.8% in population #2a, and 49.4% in population #2b ($P < 10^{-3}$). The prevalence of Metavir F stages over the entire population of 1785 patients was: F0: 4.2%, F1: 41.1%, F2: 26.5%, F3: 15.5%, and F4: 12.7%.

Decision-Making Algorithms

Binary Diagnosis of Significant Fibrosis. In the entire population, SAFE for F \geq 2 provided 94.6% diagnostic accuracy but required liver biopsy in 64.0% of patients (Table 2). Because FibroScan was not available in population #1, SAFE and BA were com-

pared in population #2: BA for F \geq 2 provided significantly lower diagnostic accuracy than SAFE for F \geq 2 (88.3% versus 92.5%, $P = 0.010$) but required a significantly lower rate of liver biopsy (34.6% versus 57.0%, $P < 10^{-3}$).

Binary Diagnosis of Cirrhosis. In the entire population, SAFE for F4 provided 89.5% diagnostic accuracy and required liver biopsy in 6.4% of patients (Table 2). In population #2, BA for F4 provided significantly higher diagnostic accuracy than SAFE for F4 (94.2% versus 87.6%, $P < 10^{-3}$) but required a significantly higher rate of liver biopsy (24.6% versus 6.7%, $P < 10^{-3}$).

SAFE for F \geq 2 and F4. The SAFE for F \geq 2 and F4 published by Sebastiani et al.⁶ provided excellent diagnostic accuracy (97.0%) but required a very high rate of liver biopsy (85.2%) (Table 3).

Successive Algorithms

Successive SAFE. In the entire population, Successive SAFE provided significantly lower diagnostic accuracy (87.3%) than individual SAFE for F \geq 2 (94.6%, $P < 10^{-3}$) or SAFE for F4 (89.5%, $P < 10^{-3}$) (Table 3). Moreover, Successive SAFE required a significantly higher rate of liver biopsy (70.8%) than SAFE for F \geq 2 (64.0%, $P < 10^{-3}$) or SAFE for F4 (6.4%, $P < 10^{-3}$). The use of FibroTest was required in 49.2% of patients with Successive SAFE versus 35.8% with SAFE for F \geq 2 ($P < 10^{-3}$) or 22.2% with SAFE for F4 ($P < 10^{-3}$). Finally, the accuracy of the noninvasive diagnosis (i.e., the rate of correctly classified patients by noninvasive tests in the subgroup of

Table 3. Comparison of Diagnostic Accuracies (%) and Rates of Required Liver Biopsy (LB, %) Between Decision-Making Algorithms Constructed for a Binary Diagnosis of Liver Fibrosis (Bold Values) and Either Successive Algorithms or the New FM+FS Classification, as a Function of Study Population

Fibrosis Algorithm Type	Name	Population					
		All		#1		#2	
		Accuracy	LB	Accuracy	LB	Accuracy	LB
Decision-making algorithm	SAFE for F \geq 2	94.6	64.0	96.0	68.8	92.5	57.0
	SAFE for F4	89.5	6.4	90.7	6.2	87.6	6.7
	SAFE for F \geq 2 and F4	97.0	85.2	97.8	87.6	95.8	81.7
	BA for F \geq 2	—	—	—	—	88.3	34.6
Successive algorithm	BA for F4	—	—	—	—	94.2	24.6
	Successive SAFE	87.3*	70.8*	89.6†	75.7*	84.1*	63.8*
	Successive BA	—	—	—	—	84.7‡	49.8‡
Noninvasive classification of fibrosis	FM+FS classification	—	—	—	—	86.7§	0.0

* $P \leq 10^{-3}$ versus SAFE for F \geq 2 or SAFE for F4.

† $P < 10^{-3}$ versus SAFE for F \geq 2 and $P = 0.059$ versus SAFE for F4.

‡ $P \leq 10^{-3}$ versus BA for F \geq 2 or BA for F4.

§ $P > 0.118$ versus Successive SAFE or Successive BA.

patients without liver biopsy) was 56.5% with Successive SAFE whereas it was 85.1% and 88.7% with SAFE for F \geq 2 and SAFE for F4, respectively.

Successive BA. In population #2, Successive BA had significantly lower diagnostic accuracy (84.7%) than individual BA for F \geq 2 (88.3%, $P = 10^{-3}$) or BA for F4 (94.2%, $P < 10^{-3}$) (Table 3). Also, Successive BA required a significantly higher rate of liver biopsy (49.8%) than BA for F \geq 2 (34.6%, $P < 10^{-3}$) or BA for F4 (24.6%, $P < 10^{-3}$). Finally, the accuracy of the noninvasive diagnosis was 69.6% with Successive BA compared with 82.1% and 92.2% with, respectively, BA for F \geq 2 and BA for F4.

New Noninvasive Classifications of Fibrosis

There was no discrepancy between the reliable diagnoses of the CSF index and SF index (Supporting Fig. 4), with thus a required liver biopsy rate of 0%. The diagnostic accuracy of the FM+FS classification was not significantly different between populations #2a

(derivation) and #2b (validation): 87.7% and 85.8%, respectively ($P = 0.461$). Despite the absence of required liver biopsy, the FM+FS classification provided high diagnostic accuracy (86.7%), with no significant difference from that of Successive SAFE or Successive BA (Table 3). The FM+FS classification provided a lower rate of large discrepancies (≥ 2 F stages: 1.2%; Table 4) compared with Successive SAFE (3.1%, $P = 0.015$) or Successive BA (2.4%, $P = 0.115$). The rate of correctly classified patients was $>85\%$ in all diagnostic classes of the FM+FS classification (except for the F2/3 class: 74.2%), whereas it was $<73\%$ in all diagnostic classes of Successive SAFE or Successive BA (Supporting Fig. 5). The FM+FS classification provided the best performance profile,²¹ especially in F ≥ 2 stages: the rate of correctly classified patients was the highest ($>80\%$) and the most homogeneous over the fibrosis stages, compared with the other algorithms (Fig. 3).

Sensitivity Analysis

We evaluated the influence of age, sex, biopsy length, Metavir F, and alanine aminotransferase (ALT) level on the diagnostic accuracy of successive algorithms and FM+FS classification. The influence of FibroScan examination characteristics (success rate, IQR/M) was also evaluated for Successive BA and FM+FS classification.

Successive SAFE. By stepwise forward binary logistic regression, the rate of well-classified patients by Successive SAFE was independently associated with ALT (first step), age (second step), Metavir F (third step), and sex (fourth step; Supporting Table 3). The diagnostic accuracy of Successive SAFE as a function

Table 4. Rates (%) of Fibrosis Stage Discrepancy*

Fibrosis Algorithm	Population	Stage Discrepancy				
		0	1	2	3	≥ 2
Successive SAFE	All	87.3	10.2	1.8	0.6	2.5
	#1	89.6	8.4	1.5	0.6	2.1
	#2	84.1	12.8	2.4	0.7	3.1
Successive BA	#2	84.7	12.9	2.4	0.0	2.4
FM+FS classification	#2	86.7	12.1	1.2	0.0	1.2†

*The difference in fibrosis stages between liver biopsy and fibrosis algorithm results, as a function of study population.

† $P = 0.015$ versus Successive SAFE and $P = 0.115$ versus Successive BA in population #2.

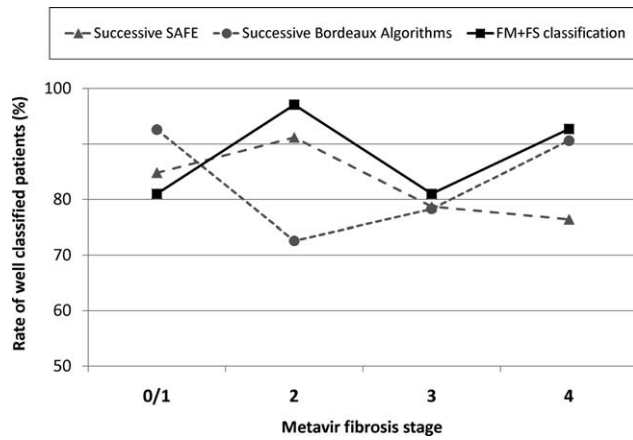


Fig. 3. Performance profile. Rate of correctly classified patients by successive algorithms and the new FM+FS classification as a function of Metavir fibrosis stage in population #2. Because of the low prevalence of F0 patients, F0 and F1 stages were pooled.

of each of these influencing factors is detailed in Supporting Table 4. The combination of age, sex, and ALT level showed that diagnostic accuracy of Successive SAFE decreased in patients with high ALT level especially in the subgroup of men ≥ 50 years old, in which only 64.6% were well classified (Fig. 4A).

Successive BA. The rate of well-classified patients by Successive BA was only independently associated with Metavir F (Supporting Table 3). The diagnostic accuracy of Successive BA was significantly lower in Metavir F2 or F3 stages compared with F0/1 or F4 (Supporting Table 4).

FM+FS Classification. The rate of well-classified patients by FM+FS classification was independently associated with Metavir F (first step), IQR/M (second step), ALT (third step), and age (fourth step, Supporting Table 3). Diagnostic accuracy of FM+FS

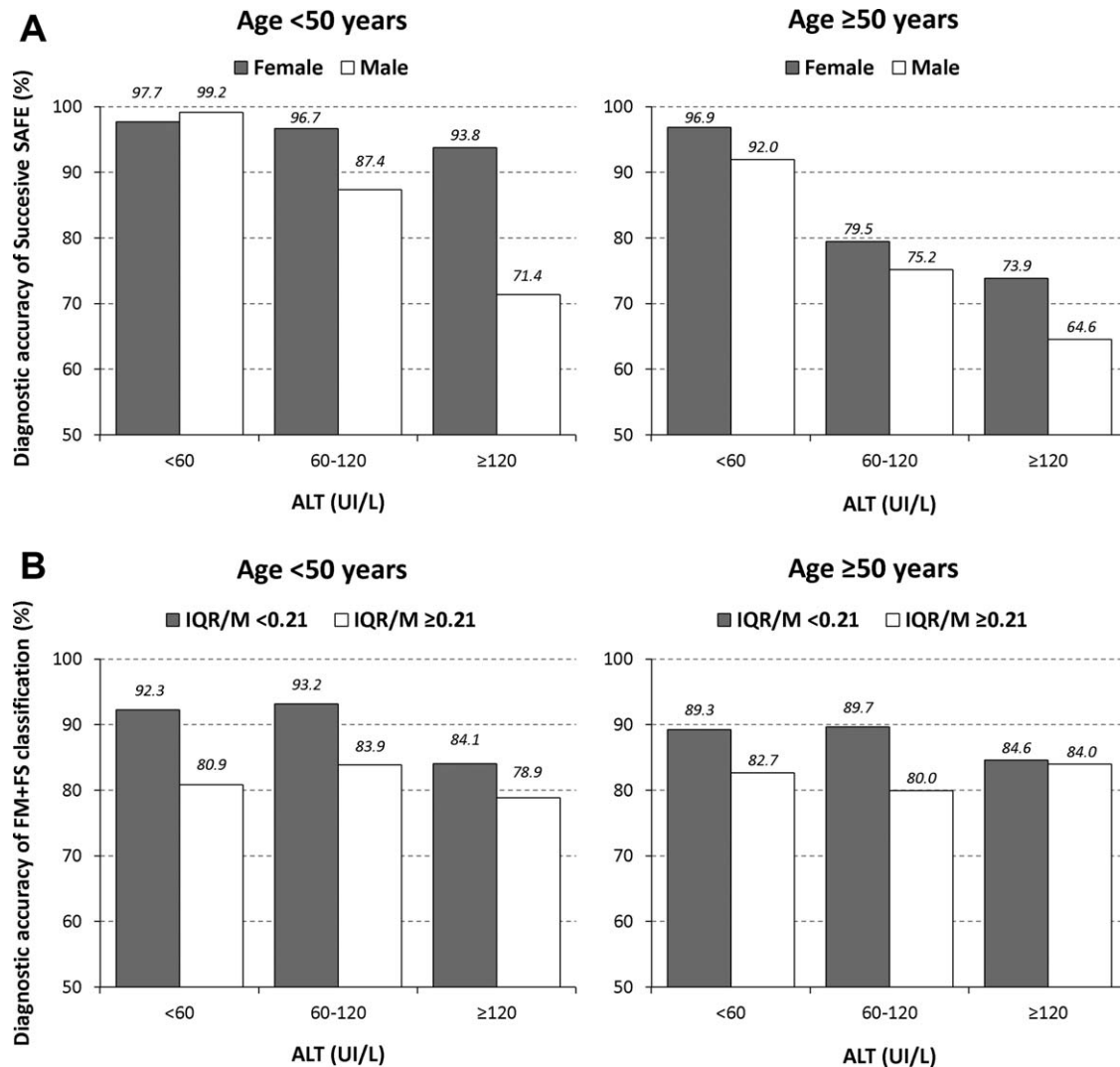


Fig. 4. Accuracy (% well-classified patients) of fibrosis algorithms as a function of combined influencing factors. (A) Accuracy of Successive SAFE as a function of combined age, sex, and ALT level. (B) Accuracy of FM+FS classification as a function of combined age, IQR/median ratio (IQR/M), and ALT level.

classification as a function of each of these influencing factors is detailed in Supporting Table 4. Diagnostic accuracy of FM+FS classification was always higher than 80%, whatever the combination of age, ALT, and IQR/M (Fig. 4B).

Discussion

This study evaluated the accuracy of two published fibrosis algorithms, the SAFE and the BA, in a large cohort of chronic hepatitis C patients. Castera et al. recently compared SAFE and BA constructed for the binary diagnoses of significant fibrosis or cirrhosis.⁷ However, their work had two limitations: first, it included a relatively small subset of 302 patients from two centers, and second, the prevalence of significant fibrosis or cirrhosis in the population studied (76% and 24%, respectively) was higher than observed in a reference population (48% and 12%, respectively) including more than 33,000 patients with chronic hepatitis C.²² This epidemiological limit induced a misvaluation of the overall accuracy, the predictive values, and probably the sensitivity and specificity of the algorithms studied.

Our study has several noteworthy points. First, we performed a direct and independent comparison of SAFE and BA in a large cohort of 729 patients with chronic hepatitis C. Second, our study had a multicenter design.¹³⁻¹⁵ Third, the prevalence of fibrosis stages in our population was very close to that of the reference population²² described above. In this setting, compared with the study by Castera et al.,⁷ we found a higher negative predictive value for BA for $F \geq 2$ and lower positive predictive values for SAFE for F4 and BA for F4. Our results are probably more representative of the real accuracy of SAFE and BA in clinical practice.

All the studies that have evaluated SAFE and BA have demonstrated their excellent accuracy for the diagnosis of significant fibrosis or cirrhosis.^{2,6-8} Two caveats should, however, be kept in mind. First, SAFE for $F \geq 2$ was impaired in all these studies by a very high rate of required liver biopsy (>50%). Indeed, because it considers negative predictive values of APRI and FibroTest as insufficient, SAFE for $F \geq 2$ recommends the use of liver biopsy when the blood fibrosis tests suggest no/mild fibrosis.² It should be noted that this implies to perform of liver biopsy in the subgroup of patients with a good prognosis.

Second, fibrosis algorithms intended for a simple binary diagnosis of fibrosis do not provide sufficient information for the management of patients in clinical

practice. Indeed, physicians have to answer two questions: (1) whether the patient needs antiviral therapy (i.e., is there any significant fibrosis in genotype 1 chronic hepatitis C); and (2) whether the patient needs screening for hepatocellular carcinoma and esophageal varices (i.e., is there any cirrhosis). Thus, physicians first have to apply the algorithm for the diagnosis of significant fibrosis and then, if the noninvasive diagnosis is $F \geq 2$, apply the algorithm for the diagnosis of cirrhosis. This successive use of algorithms for binary diagnosis leads to greater rates of misclassified patients and of liver biopsy. In this setting, our results clearly demonstrated that Successive SAFE (Fig. 1C) and Successive BA (Fig. 2C) had significantly lower diagnostic accuracies and required significantly higher rates of liver biopsy than single algorithms (respectively SAFE for $F \geq 2$ or SAFE for F4, and BA for $F \geq 2$ or BA for F4) (Table 3). Moreover, Successive SAFE required a significantly higher rate of FibroTest use, compared with SAFE for $F \geq 2$ or SAFE for F4. It is also of note that the results for Successive SAFE or Successive BA were the same when the algorithm for the binary diagnosis of cirrhosis was performed first, and then followed, if necessary, by the algorithm for the binary diagnosis of significant fibrosis (data not shown). Taken together, these results show that the accuracy of SAFE and BA for the diagnosis of fibrosis in clinical practice has been overestimated in published studies.

Also, Sebastiani et al. proposed an algorithm for the simultaneous detection of significant fibrosis and cirrhosis (Supporting Fig. 3).⁶ Despite very high diagnostic accuracy (97.0%), this algorithm required liver biopsy in almost all patients (85.2%), thus greatly limiting its interest for the noninvasive diagnosis of fibrosis.

The association of FibroMeter and FibroScan, which was shown to be the best combination among six noninvasive fibrosis tests,¹⁴ serves as the foundation of the FM+FS classification (Supporting Fig. 4). This new noninvasive classification of fibrosis had several advantages compared with SAFE and BA. First, the FM+FS classification required no liver biopsy. Second, the FM+FS classification provided a more precise diagnosis (six diagnostic classes) than Successive SAFE (two classes) or Successive BA (three classes) (Supporting Table 1). Third, despite the absence of liver biopsy requirement, the diagnostic accuracy of the FM+FS classification was not significantly different from those of Successive SAFE or Successive BA. Finally, the FM+FS classification provided the best performance profile compared with Successive SAFE or Successive

BA (Fig. 3). It should be noted that the reference for liver fibrosis in our study was liver biopsy, which should be considered as a “best standard” but not a “gold standard.”²³ Thus, the diagnostic accuracy of the FM+FS classification was probably underestimated in our study.²⁴

Finally, the FM+FS classification significantly improved the noninvasive diagnosis of liver fibrosis by avoiding liver biopsy and refined the precision of fibrosis diagnosis while maintaining very high accuracy. Thus, between published decision-making algorithms and our new noninvasive classification of fibrosis, the FM+FS classification appears to be the most appropriate for clinical use. Because it requires several steps and calculations (Supporting Fig. 4), the use of the FM+FS classification in clinical practice may, at first glance, seem complex. However, once all these steps are computerized, physicians need only provide the results of FibroScan and FibroMeter.

The length of liver biopsy had no influence on the diagnostic accuracy of Successive SAFE, Successive BA, or the FM+FS classification. The accuracy of Successive SAFE was independently influenced by age, sex, and ALT level. Indeed, Successive SAFE was quite inaccurate at high ALT levels, especially in older men (Fig. 4A). The success rate of FibroScan had no influence on the diagnostic accuracy of Successive BA or the FM+FS classification. In this setting, it has been already been shown that IQR/M was the only FibroScan characteristic that had a significant impact on its accuracy.²⁰ In our study, IQR/M influenced the accuracy of the FM+FS classification but not that of Successive BA. This was due to the high rate of liver biopsy (49.8%) required by Successive BA. In fact, the FM+FS classification is more sensitive to the influence of IQR/M than Successive BA because its diagnosis depends on the FibroScan results in all patients (0% liver biopsy required). In addition to IQR/M, the accuracy of the FM+FS classification was independently influenced by age and ALT level. However, the accuracy of the FM+FS classification remained higher than 80% in the various subgroups resulting from the combination of these three parameters (Fig. 4B).

In conclusion, SAFE and BA for binary diagnoses of significant fibrosis or cirrhosis have excellent diagnostic accuracy in chronic hepatitis C. However, in clinical practice, the significant fibrosis algorithm and the cirrhosis algorithm have to be used successively, which induces a significant decrease in diagnostic accuracy and a significant increase in the rate of required liver biopsy. A new noninvasive classification of fibrosis synchronously combining FibroScan and FibroMeter

results allows for an entirely noninvasive (0% liver biopsy required) and precise (six fibrosis classes) diagnosis of liver fibrosis with a diagnostic accuracy (87%) that is not significantly different from those of SAFE and BA.

Acknowledgment: The **SNIFF 32 study** was conducted by the following: **Angers:** Sophie Michalak, Anselme Konaté, Catherine Ternisien, Alain Chevailler, Françoise Lunel, Marie Christine Rousselet, Wael Mansour; **Provence-Alpes-Côte-d'Azur (PACA):** Philippe Halfon, Marc Bourlière, Denis Ouzan, Albert Tran, Danielle Botta, Christophe Renou, Christophe Sattonnet, Marie-Christine Saint-Paul, Thierry Benderitter, Stéphane Garcia, Henri Pierre Bonneau, Guillaume Penaranda; **Tours:** Yannick Bacq, Anne de Muret, Marie-Claude Bréchet; **Grenoble:** Vincent Leroy, Nathalie Sturm, Marie-Noëlle Hilleret, Patrice Faure, Jean-Charles Renversez, Françoise Morel, Candice Trocme; **Bordeaux:** Victor de Ledinghen, Juliette Foucher, Laurent Castéra, Patrice Couzigou, Pierre-Henri Bernard, Wassil Merrouche, Paulette Bioulac-Sage, Brigitte Le Bail; and **Clichy:** Claude Degott, Valérie Paradis.

The VINDIAG 7 study was conducted by the following: **Grenoble:** Vincent Leroy, Nathalie Sturm, Marie-Noëlle Hilleret, Patrice Faure, Jean-Charles Renversez, Françoise Morel, Nathalie Sturm, Candice Trocme; **Bordeaux:** Victor de Ledinghen, Juliette Foucher, Laurent Castéra, Patrice Couzigou, Pierre-Henri Bernard, Wassil Merrouche, Paulette Bioulac-Sage, Brigitte Le Bail.

The FIBROSTAR study was conducted by the following: **Hepatologists:** R. Poupon, A. Poujol, Saint-Antoine, Paris; A. Abergel, Clermont-Ferrand; J.P. Bronowicki, Nancy; J.P. Vinel, S. Metivier, Toulouse; V. De Ledinghen, J. Foucher, Bordeaux; O. Gorla, Rouen; M. Maynard-Muet, C. Treppe, Lyon; Ph. Mathurin, Lille; D. Guyader, H. Danielou, Rennes; O. Rogeaux, Chambéry; S. Pol, Ph. Sogni, Cochin, Paris; A. Tran, Nice; P. Calès, Angers; P. Marcellin, T. Asselah, Clichy; M. Bourlière, V. Oulès, Saint Joseph, Marseille; D. Larrey, Montpellier; F. Habersetzer, Strasbourg; M. Beaugrand, Bondy; V. Leroy, MN Hilleret, Grenoble.

Biologists: R-C. Boisson, Lyon Sud; M-C. Gelineau, B. Poggi, Hôtel Dieu, Lyon; J-C. Renversez, Candice Trocme, Grenoble; J. Guéchet, R. Lasnier, M. Vaubourdolle, Paris; H. Voitot, Beaujon, Paris; A. Vassault, Necker, Paris; A. Rosenthal-Allieri, Nice; A. Lavoine, F. Ziegler, Rouen; M. Bartoli, C. Lebrun, Chambéry; A. Myara, Paris Saint-Joseph; F. Guerber, A. Pottier, Eliblio, Vizille.

Pathologists: E-S. Zafrani, Créteil; N. Sturm, Grenoble.

Methodologists: A. Bechet, J-L Bosson, A. Paris, S. Royannais, CIC, Grenoble; A. Plages, Grenoble

We also thank the following contributors: Gilles Hunault, Pascal Veillon, Gwenaëlle Soulard; and Kevin L. Erwin (for English proofreading).

References

1. Castera L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology* 2005;128:343-350.
2. Sebastiani G, Vario A, Guido M, Noventa F, Plebani M, Pistis R, et al. Stepwise combination algorithms of non-invasive markers to diagnose significant fibrosis in chronic hepatitis C. *J Hepatol* 2006;44:686-693.
3. Bourliere M, Penaranda G, Renou C, Botta-Fridlund D, Tran A, Portal I, 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;13:659-670.
4. Bourliere M, Penaranda G, Ouzan D, Renou C, Botta-Fridlund D, Tran A, et al. Optimized stepwise combination algorithms of non-invasive liver fibrosis scores including Hepascore in hepatitis C virus patients. *Aliment Pharmacol Ther* 2008;28:458-467.
5. Becker L, Salameh W, Sferruzza A, Zhang K, Chen R, Malik R, et al. Validation of Hepascore, compared to simple indices of fibrosis, in US patients with chronic hepatitis C virus infection. *Clin Gastroenterol Hepatol* 2009;7:696-701.
6. Sebastiani G, Halfon P, Castera L, Pol S, Thomas DL, Mangia A, et al. SAFE biopsy: a validated method for large-scale staging of liver fibrosis in chronic hepatitis C. *HEPATOLOGY* 2009;49:1821-1827.
7. Castera L, Sebastiani G, Le Bail B, de Ledinghen V, Couzigou P, Alberti A. Prospective comparison of two algorithms combining non-invasive methods for staging liver fibrosis in chronic hepatitis C. *J Hepatol* 2010;52:191-198.
8. Boursier J, Vergniol J, Sawadogo A, Dakka T, Michalak S, Gallois Y, et al. The combination of a blood test and Fibroscan improves the non-invasive diagnosis of liver fibrosis. *Liver Int* 2009;29:1507-1515.
9. Rashid M, Mitchell JD, Cramp ME, Cross TJ. Limitations of the algorithm for the SAFE biopsy: a noninvasive fibrosis measure in chronic hepatitis C. *HEPATOLOGY* 2010;51:354-355.
10. Boursier J, Bacq Y, Halfon P, Leroy V, De Ledinghen V, De Muret A, et al. Improved diagnostic accuracy of blood tests for severe fibrosis and cirrhosis in chronic hepatitis C. *Eur J Gastroenterol Hepatol* 2009; 21:28-38.
11. Boursier J, Cales P. Combination of fibrosis tests: sequential or synchronous? *HEPATOLOGY* 2009;50:656-657.
12. Cales P, Boursier J, de Ledinghen V, Halfon P, Bacq Y, Leroy V, et al. Evaluation and improvement of a reliable diagnosis of cirrhosis by blood tests. *Gastroenterol Clin Biol* 2008;32:1050-1060.
13. Cales P, De Ledinghen V, Halfon P, Bacq Y, Leroy V, Boursier J, et al. Evaluating accuracy and increasing the reliable diagnosis rate of blood tests for liver fibrosis in chronic hepatitis C. *Liver Int* 2008;28: 1352-1362.
14. Boursier J, De Ledinghen V, Zarski J, Rousselet MC, Sturm N, Foucher J, et al. A new combination of blood test and Fibroscan for accurate non-invasive diagnosis of liver fibrosis stages in chronic hepatitis C. *Am J Gastroenterol* 2011;106:1255-1263.
15. Zarski H, Sturm N, Guechot J, Paris A. Independent and prospective comparison of 9 surrogate markers for the diagnosis of liver fibrosis in chronic hepatitis C [Abstract]. *HEPATOLOGY* 2009;50(Suppl 4): 1061A.
16. Leroy V, Halfon P, Bacq Y, Boursier J, Rousselet MC, Bourliere M, et al. Diagnostic accuracy, reproducibility and robustness of fibrosis blood tests in chronic hepatitis C: a meta-analysis with individual data. *Clin Biochem* 2008;41:1368-1376.
17. Wai CT, Greenon 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.
18. Cales P, Veillon P, Konate A, Mathieu E, Ternisien C, Chevailler A, et al. Reproducibility of blood tests of liver fibrosis in clinical practice. *Clin Biochem* 2008;41:10-18.
19. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. *J Hepatol* 2008;48:835-847.
20. 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.
21. Halfon P, Bacq Y, De Muret A, Penaranda G, Bourliere M, Ouzan D, et al. Comparison of test performance profile for blood tests of liver fibrosis in chronic hepatitis C. *J Hepatol* 2007;46:395-402.
22. Thein HH, Yi Q, Dore GJ, Krahn MD. Estimation of stage-specific fibrosis progression rates in chronic hepatitis C virus infection: a meta-analysis and meta-regression. *HEPATOLOGY* 2008;48:418-431.
23. Bedossa P, Carrat F. Liver biopsy: the best, not the gold standard. *J Hepatol* 2009;50:1-3.
24. Mehta SH, Lau B, Afdhal NH, Thomas DL. Exceeding the limits of liver histology markers. *J Hepatol* 2009;50:36-41.