Criteria Grid
Best Practices and Interventions for the Prevention and Awareness of Hepatitis C

Best Practice/Intervention:	Calès P. et al. (2015) Cirrhosis Diagnosis and Liver Fibrosis Staging: Transient Elastometry Versus Cirrhosis Blood Test. J Clin Gastroenterol, 49(6): 512-9.
Date of Review:	February 11, 2016
Reviewer(s):	Christine Hu

Part A

Category:	**Basic Science** ☐ **Clinical Science** ☒ **Public Health/Epidemiology** ☐ **Social Science** ☐ **Programmatic Review** ☐
Best Practice/Intervention:	**Focus:** **Hepatitis C** ☒ **Hepatitis C/HIV** ☐ **Other:** _____ **Level:** **Group** ☒ **Individual** ☐ **Other:** _____ **Target Population:** **patients with chronic HCV infection** **Setting:** **Health care setting/Clinic** ☒ **Home** ☐ **Other:** _____ **Country of Origin:** **France** **Language:** **English** ☒ **French** ☐ **Other:** _____

Part B

	YES	NO	N/A	COMMENTS
Is the best practice/intervention a meta-analysis or primary research? ***Please go to Comments section****.*	☒	☐	☐	Primary research: to compare the accuracies of CirrhoMeter and Fibroscan for the diagnosis of cirrhosis and to develop a precise fibrosis classification to increase diagnosis precision.
The best practice/intervention shows evidence of "scale up" ability	☒	☐	☐	A total of 679 patients were included in the prognostic study along with 1110 patients with prognostic study.
The best practice/intervention shows evidence of transferability	☐	☒	☐	

The best practice/intervention shows evidence of adaptation	☐	☒	☐	
Do the methodology/results described allow the reviewer(s) to assess the generalizability of the results?	☒	☐	☐	Results are generalizable.
	YES	**NO**	**N/A**	**COMMENTS**
Are the best practices/methodology/results described applicable in developed countries?	☒	☐	☐	
Are the best practices/methodology/results described applicable in developing countries?	☒	☐	☐	Results are applicable to developing countries as long as clinical settings utilize CirrhoMeter or Fibroscan for diagnosis and prognosis of cirrhosis.
The best practice/intervention has utilized a program evaluation process	☐	☒	☐	
Consultation and feedback with community has taken place	☐	☒	☐	
The best practice/intervention is sensitive to gender issues	☐	☒	☐	
The best practice/intervention is sensitive to multicultural and marginalized populations	☐	☒	☐	
The best practice/intervention is easily accessed/available electronically	☒	☐	☐	Journal subscription required for access at http://journals.lww.com/
Is there evidence of a cost effective analysis? If so, what does the evidence say? ***Please go to Comments section***	☐	☒	☐	
How is the best practice/intervention funded? ***Please go to Comments section***	☐	☒	☐	No funding stated
Is the best practice/intervention dependent on external funds?	☐	☐	☒	

Other relevant criteria: _____	☐	☐	☐	- Results find that noninvasive tests are good tools for diagnosing cirrhosis - Fibroscan is more accurate than CirrhoMeter for cirrhosis diagnosis, but the latter is better for prognostication.

Cirrhosis Diagnosis and Liver Fibrosis Staging
Transient Elastometry Versus Cirrhosis Blood Test

Paul Calès, MD,† Jérôme Boursier, MD, PhD,*† Frédéric Oberti, MD,*†*
Derek Bardou, Jean-Pierre Zarski, MD,‡§*
Victor de Lédinghen, MD, PhD,‖¶ and Multicentric Group

Introduction: Elastometry is more accurate than blood tests for cirrhosis diagnosis. However, blood tests were developed for significant fibrosis, with the exception of CirrhoMeter^{2G} developed for cirrhosis. We compared the performance of Fibroscan and CirrhoMeter^{2G}, and classic binary cirrhosis diagnosis versus new fibrosis staging for cirrhosis diagnosis.

Methods: The diagnostic population included 679 patients with hepatitis C and liver biopsy (Metavir staging and morphometry), Fibroscan, and CirrhoMeter^{2G}. The prognostic population included 1110 patients with chronic liver disease and both tests.

Results: *Binary diagnosis:* AUROCs for cirrhosis were: Fibroscan: 0.905; CirrhoMeter^{2G}: 0.857; and $P = 0.041$. Accuracy (Youden cutoff) was: Fibroscan: 85.4%; CirrhoMeter^{2G}: 79.2%; and $P < 0.001$. *Fibrosis classification* provided 6 classes (F0/1, F1/2, F2 ± 1, F3 ± 1, F3/4, and F4). Accuracy was: Fibroscan: 88.2%; CirrhoMeter^{2G}: 88.8%; and $P = 0.77$. A simplified fibrosis classification comprised 3 categories: discrete (F1 ± 1), moderate (F2 ± 1), and severe (F3/4) fibrosis. Using this simplified classification, CirrhoMeter^{2G} predicted survival better than Fibroscan (respectively, $\chi^2 = 37.9$ and 19.7 by log-rank test), but both predicted it well ($P < 0.001$ by log-rank test). *Comparison:* binary diagnosis versus fibrosis classification, respectively, overall accuracy: CirrhoMeter^{2G}: 79.2% versus 88.8% ($P < 0.001$); Fibroscan: 85.4% versus 88.2% ($P = 0.127$); positive predictive value for cirrhosis by Fibroscan: Youden cutoff (11.1 kPa): 49.1% versus cutoffs of F3/4 (17.6 kPa): 67.6% and F4 classes (25.7 kPa): 82.4%.

Conclusions: Fibroscan's usual binary cutoffs for cirrhosis diagnosis are not sufficiently accurate. Fibrosis classification should be preferred over binary diagnosis. A cirrhosis-specific blood test markedly attenuates the accuracy deficit for cirrhosis diagnosis of usual blood tests versus transient elastometry, and may offer better prognostication.

Key Words: noninvasive diagnosis, diagnostic accuracy, liver fibrosis, transient elastometry, blood tests, diagnostic tests

(*J Clin Gastroenterol* 2015;49:512–519)

Noninvasive tests for liver lesions were initially blood tests targeted mainly for cirrhosis, but accuracy was fair to poor.[1] Some years later, teams started focusing their attention on constructing blood tests with significant fibrosis as the diagnostic target.[2] However, cirrhosis is again becoming the main diagnostic target as it modifies therapeutic schemas—even the most recent—in chronic hepatitis C (CHC). Cirrhosis diagnosis also significantly influences care management, especially complication screening. Even after a successful treatment, the diagnosis of cirrhosis remains a clinical challenge, for example, sustained virological response in cirrhosis does not suppress the risk of hepatocellular carcinoma.[3] In addition, in many countries, a significant proportion of patients cannot be treated because current treatments are too expensive, or at best, treatments are restricted to the most severe cases. Moreover, transient elastometry is becoming a popular diagnostic tool as it provides excellent diagnostic performance for cirrhosis,[4,5] outperforming blood tests in particular.[6] However, almost all blood tests were constructed for significant fibrosis. It might therefore be expected that their accuracy for cirrhosis be lesser than that of transient elastometry. Of note, we have recently constructed a blood test specifically designed for cirrhosis called CirrhoMeter^{2G}.[7] Care management for a patient without fibrosis will differ greatly from that of a patient with severe fibrosis. This raises the issue of fibrosis staging. Most noninvasive tests are constructed and/or evaluated according to a single diagnostic target, thus providing a binary (yes or no) diagnosis. Fibrotest and FibroMeter are among the few tests offering a fibrosis classification method, that is, reflecting histologic staging for these 2, but they differ in methods and accuracies.[8] For Fibroscan, the most popular classification method cumulates the cutoffs determined for significant fibrosis and cirrhosis. In fact, only 2 classification methods with valid statistical bases have been described.[9,10] Finally, the accuracies of CirrhoMeter^{2G} and

Received for publication January 22, 2014; accepted March 16, 2014.
From the *Liver-Gastroenterology Department, CHU Angers; †HIFIH Laboratory, LUNAM University, Angers; ‡Liver-Gastroenterology Department, University Hospital; §INSERM/UJF U823, IAPC, IAB, University, Grenoble; ‖Liver-Gastroenterology Unit, University Hospital, Pessac; and ¶INSERM U1053, Bordeaux University, Bordeaux, France.
Multicentric Group: Centers of study ANRS HC EP 23 Fibrostar.
P.C.: planning and conducting the study, collecting and interpreting data, statistical analysis, drafting the manuscript, approving the final draft submitted. J.B.: planning and conducting the study, collecting and interpreting data, drafting the manuscript, approving the final draft submitted. F.O.: collecting and interpreting data, approving the final draft submitted. D.B.: collecting and interpreting data, approving the final draft submitted. J.-P.Z.: planning and conducting the Fibrostar study, collecting and interpreting data, approving the final draft submitted. V.d.L.: collecting and interpreting data, approving the final draft submitted.
Supported by ANRS (French national agency for AIDS and viral hepatitis) for the HC EP 23 Fibrostar study.
P.C. is a consultant for BioLiveScale Inc. that has a license for FibroMeter and CirrhoMeter from Angers University. BioLiveScale had no role in the present study. The other authors declare that they have nothing to disclose.
Reprints: Paul Calès, MD, Service d'Hépato-Gastroentérologie, CHU, 49933 Angers Cedex 09, France (e-mail: paul.cales@univ-angers.fr).
Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website, www.jcge.com.

Fibroscan have never been thoroughly compared and a fibrosis classification for them has never been developed according to validated statistical methods.

Therefore, the main aim of this study was to compare the accuracies of CirrhoMeter^{2G} and Fibroscan for the diagnosis of cirrhosis. The secondary aim was to develop a precise fibrosis classification, with the goal of increasing diagnosis precision over that offered by available staging. We conducted this study on a large diagnostic population of patients with liver fibrosis caused by CHC, and then on a prognostic population with miscellaneous causes.

PATIENTS AND METHODS

Population

The diagnostic population, composed of 2 subgroups, was provided by the VINDIAG 7 study.[11] The first subgroup included patients with CHC hospitalized for a percutaneous liver biopsy. They were prospectively enrolled from March 2004 to September 2008 in 3 tertiary centers in France (Angers, Bordeaux, and Grenoble) if they had anti-HCV antibodies, HCV RNA in serum, and available liver biopsy, blood markers, and valid Fibroscan results. Exclusion criteria were additional causes of liver disease, particularly HIV or HBV coinfection, putative antifibrotic treatment in the preceding 6 months, and alcohol consumption of >30 g/d in the past 5 years. The second subgroup comprised other patients recruited in the multicenter Fibrostar study promoted by the French National Agency for Research in AIDS and Hepatitis.[12] This study prospectively included 512 patients with CHC. Patients with cirrhosis complications (ascites, variceal bleeding, systemic infection, and hepatocellular carcinoma) were not included. Blood tests and Fibroscan were conducted in the week preceding biopsy. Patient duplication in the 2 subgroups was corrected for and patients with incomplete data were not included (n = 50). Finally, 679 patients were included in the present diagnostic study.

The prognostic cohort (SNIFF 64) included 1110 patients with miscellaneous causes of chronic liver disease.[13] The study protocol conformed to the ethical guidelines of the current Declaration of Helsinki and received approval from the local Ethics Committee. Patients gave written consent.

Methods

Histologic Assessment

Liver biopsies were performed using Menghini's technique with a 1.4 to 1.6 mm diameter needle. Biopsy specimens were fixed in a formalin-alcohol-acetic solution and embedded in paraffin; 5-μm-thick sections were then cut and stained with hematoxylin-eosin-saffron. Liver fibrosis was evaluated according to fibrosis Metavir (F_M) staging. Significant fibrosis was defined as Metavir stages $F_M \geq 2$, and cirrhosis as F_M4. Liver fibrosis was evaluated by 2 senior experts with a consensus reading in case of discordance in Angers and in the Fibrostar study, and by a senior expert in Bordeaux and Grenoble. Fibrosis staging was considered as reliable when liver specimen length was ≥ 15 mm and/or portal tract number was ≥ 8.[14] These liver specimen findings served as a reference for the liver fibrosis evaluations by noninvasive tests. The area of liver fibrosis was centrally measured by automated morphometry as previously described,[15] with the area of portal fibrosis[16] as the main independent histologic judgment criterion.

Fibrosis Blood Tests

The following blood tests were calculated according to scores detailed in the supplementary material (Supplemental Digital Content 1, http://links.lww.com/JCG/A137): FibroMeter^{2G} [10] and CirrhoMeter^{2G}.[7] CirrhoMeter^{2G} includes the same markers as FibroMeter^{2G} but with different coefficients. Biomarkers are: platelet count, protrombin index, AST, α2-macroglobulin, hyaluronate, urea, age, and sex. All blood assays were performed in the same laboratories of each center, or centralized in the Fibrostar study. We have previously demonstrated the excellent interlaboratory reproducibility of these blood fibrosis tests.[17]

Liver Stiffness Evaluation

Fibroscan (Echosens, Paris, France) examination was conducted by experienced observers (> 50 examinations before the study), blinded for patient data. Examination conditions were those recommended by the manufacturer.[18] Results (kPa) were expressed as the median and the interquartile range of all valid measurements (maximum: 10). Fibroscan reliability was evaluated according to the most recent criteria based on interquartile range and median.[19] Patients were included regardless of the reliability level.

Statistics

Quantitative variables were expressed as mean ± SD. The diagnostic cutoffs of fibrosis tests were calculated (a posteriori cutoff) according to the highest Youden index (sensitivity + specificity − 1). The diagnostic accuracy of each test was expressed as the area under the receiver operating characteristic (AUROC) and the overall accuracy (rate of well-classified patients according to F_M). AUROCs were compared using the nonparametric Delong test.[20] Data were reported according to STARD statements[21] and analyzed on an intention to diagnose basis.

Fibrosis Classification

The estimated fibrosis staging by noninvasive tests was called fibrosis classification. This was developed using our previously described percentile method[10] evaluated elsewhere.[8] Briefly, the test values were first segmented according to patient percentiles (usually 2.5%). These 20 percentiles were plotted against F_M stages in a table. They were then grouped in different classes of estimated fibrosis (F) stages to obtain a probability of $\geq 75\%$ for a maximum of $3F_M$ stages per class. For example, with Fibroscan, from the second to the eighth percentile, the probability was $\geq 78\%$ (mean, 88%) for F_M 1 or 2, thus determining the F1/2 class. Each fibrosis class was determined by 2 cutoffs: lower and upper. The lower cutoffs were easily deduced from the first percentile of each F class, for example, 3.9 kPa for the F1/2 class (Fig. 1). We thus developed a new (a posteriori thresholds) *fibrosis classification* for Fibroscan and CirrhoMeter^{2G}.

Sample Size Calculation

The size of the population was necessary to detect a significant difference between the 2 tests for the diagnosis of cirrhosis. With an α risk of 0.05, β risk of 0.05, cirrhosis prevalence of 0.12, AUROC correlation of 0.82, and bilateral testing, the required sample size was 659 patients

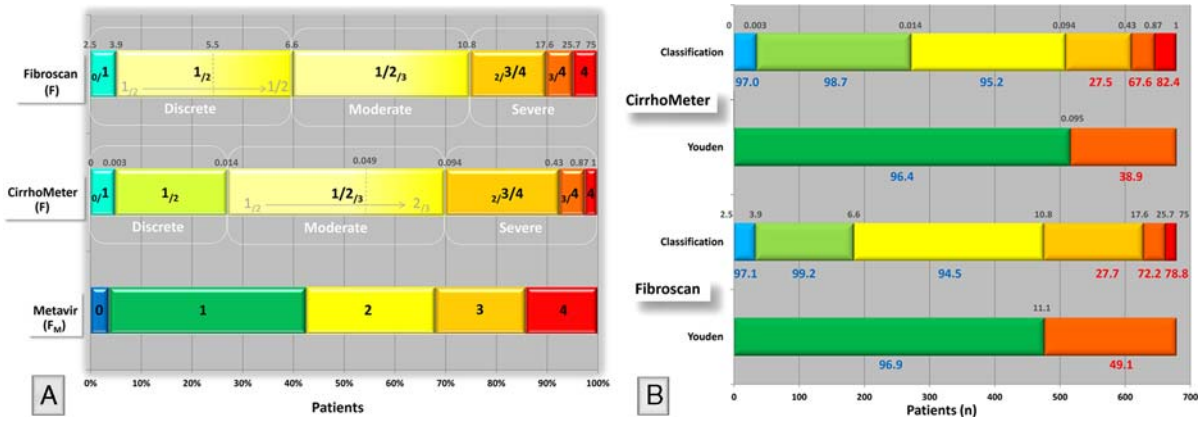

FIGURE 1. A, Classification in fibrosis stages by Metavir staging and noninvasive tests. The colored bars depict fibrosis stages as a function of patient prevalence on the *x*-axis. Within the bars, the numbers depict the fibrosis stages of fibrosis classes (black numbers) or subclasses (gray numbers); the most probable fibrosis stage(s) is (are) indicated by a larger figure. The gray numbers above the bars indicate the cutoffs between the fibrosis classes and subclasses of the diagnostic tests (values from 0 to 1 for CirrhoMeter^{2G} and from 2.5 to 75 kPa for Fibroscan). The white categories depict simplified fibrosis classification. B, Comparison of predictive values for cirrhosis between fibrosis classification and binary diagnosis. Binary diagnosis is shown as a function of the Youden cutoff. Blue and red figures indicate, respectively, the negative predictive value and positive predictive value for cirrhosis of each bar. The gray numbers above the bars indicate the cutoffs between the fibrosis classes.

for the following expected AUROC values: Fibroscan: 0.92; CirrhoMeter^{2G}: 0.90.[7]

RESULTS

Populations

The main characteristics of the diagnostic and prognostic populations are described in Table 1. Briefly, both populations had close characteristics except for etiology; however, CHC was the predominant cause in the prognostic population (76%). Follow-up duration was 4.0 ± 1.7 years in the prognostic population.

Binary Diagnosis

Binary diagnosis calculations were based on raw results (expressed as a score from 0 to 1 for CirrhoMeter^{2G} and in kPa for Fibroscan). The AUROC for cirrhosis of Fibroscan, 0.905 (95% CI, 0.871-0.938), was significantly higher than that of CirrhoMeter^{2G}, 0.857 (95% CI, 0.813-0.900), $P = 0.041$. Table 2 shows that several diagnostic indices [overall accuracy, positive predictive value (PPV), specificity] were better with Fibroscan than with CirrhoMeter^{2G}.

Fibrosis Classification

Construction

We obtained the same 6 fibrosis classes with both tests: F0/1, F1/2, F2 ± 1, F3 ± 1, F3/4, and F4 (Fig. 1A). It was possible to further refine staging within some fibrosis classes (Fig. 1A). (Detailed data on classification refinement, presented in Fig. 1A, are provided in Tables S1 and S2, Supplemental Digital Content 2 and 3, http://links.lww.com/JCG/A138; http://links.lww.com/JCG/A139.)

Performance

Accuracies are detailed in Table 3.

Overall accuracy: The accuracy of CirrhoMeter^{2G} fibrosis classification, 88.8% (86.4% to 91.1%), was not significantly different from that of Fibroscan, 88.2% (85.8% to 90.6%) ($P = 0.773$).

Accuracy by fibrosis stages: In both the tests, accuracy was optimal (≥ 82%) in every F_M stage, except F_M0, where it was ≤28% (Table 3). This limit was circumvented by fibrosis classification, as significant discordance (>1 F_M stage) between F_M0 (by liver biopsy) and fibrosis classification was observed in only 4% of patients with both

TABLE 1. Population Characteristics

	Population	
	Diagnostic	**Prognostic**
Patients (n)	679	1110
Male (%)	61.3	58.5
Age (y)	51.4 ± 11.2	48.8 ± 14.8
Body mass index (kg/m^2)	24.8 ± 4.0	25.3 ± 4.7
Cause (%)		
Hepatitis C virus	100	76.0
Nonalcoholic fatty liver disease	0	7.7
Alcoholic liver disease	0	3.4
Others	0	13.0
Metavir (%)		NA
F_M0	3.7	—
F_M1	38.9	—
F_M2	25.5	—
F_M3	17.8	—
F_M4	15.0	—
Significant fibrosis (%)	58.3	58.4/50.1*
Reliable biopsy (%)	93.5	NA
Liver stiffness (Fibroscan)		
Median (kPa)	10.0 ± 7.9	10.0 ± 10.3
IQR/median <0.21 (%)	67.8	61.4
Very reliable (%)†	15.7	17.2
Reliable (%)†	76.9	72.6
Poorly reliable (%)†	7.3	10.2
CirrhoMeter^{2G}	0.11 ± 0.20	0.11 ± 0.22

*According to CirrhoMeter^{2G} and Fibroscan classifications, respectively.

†According to the most recent criteria.[19]

F_M indicates fibrosis Metavir stage, IQR, interquartile range, NA, not available

TABLE 2. Cirrhosis Diagnosis

	Cutoff *	Sensitivity (%)	Specificity (%)	Predictive Value (%)		Overall Accuracy (%)	Likelihood Ratio	
				Positive	Negative		Positive	Negative
Binary diagnosis								
By Youden index								
CirrhoMeter^{2G}	0.095	*82.3*	78.7	38.9	*96.4*	79.2	3.87	0.22
Fibroscan	11.1	***83.3***	85.8	49.1	***96.9***	85.4	5.85	**0.19**
P†	—	1	< 0.001	0.051	0.680	< 0.001	—	—
Classification								
CirrhoMeter^{2G}	F4	14.6	99.5	**82.4**	87.6	87.5	**28.34**	0.86
	F3 ± 1	*82.3*	78.6	38.7	*96.4*	79.1	3.84	0.23
Fibroscan	F4	27.1	**98.8**	78.8	89.2	88.7	22.56	0.74
	F3 ± 1	***83.3***	84.6	47.1	***96.9***	84.4	5.40	0.20

Comparison of diagnostic indices between tests as a function of staging: binary diagnosis or classification (with 2 diagnostic cutoffs: < F4 and < F3 ± 1 classes).

Best diagnostic indices are indicated in bold (plus italics when observed in 2 stagings).

*Cutoffs by maximum Youden indices for raw tests (binary diagnosis) or lower cutoffs of class(es) for classification. Unit: kPa for Fibroscan.

†*P* by paired McNemar test for overall accuracy, sensitivity, and specificity. Predictive value comparisons are less powerful (unpaired χ^2 test).

tests. The accuracy was thus not homogenous as a function of Metavir F_M, but it was with fibrosis classes for both tests.

Diagnostic indices for cirrhosis: Cirrhosis sensitivity could be calculated in 2 ways. First, in the F4 class sensitivity was 14.6% for CirrhoMeter^{2G} and 27.1% for Fibroscan ($P = 0.033$); this is the proportion of all cirrhosis in this class. Second, the sensitivity of the classes including F4 (ie, pooled F3 ± 1, F3/4, and F4 classes) was 82.3% and 83.3% ($P = 1$), respectively (Table 3: F_M4 row). The PPV for cirrhosis significantly increased as a function of fibrosis class ($P < 0.001$ by ANOVA) with both tests: from ≤5% in classes without F4 to around 28% in F3 ± 1, around 70% in F3/4, and around 80% in F4 class (Table 3: bottom rows).

TABLE 3. Fibrosis Classifications

	%		
	CirrhoMeter^{2G}	Fibroscan	*P**
Overall accuracy	88.8	88.2	0.773
Accuracy by Metavir F_M stage†			
F_M0	28.0	16.0	0.549
F_M1	90.2	93.6	0.211
F_M2	95.4	95.4	1
F_M3	94.2	85.1	**0.027**
F_M4	82.3	83.3	1
Accuracy by fibrosis class‡			
F0/1	78.8	85.7	0.705
F1/2	90.7	88.2	0.432
F2 ± 1	92.4	91.1	0.587
F3 ± 1	83.0	84.2	0.809
F3/4	88.2	91.7	0.937
F4	82.4	78.8	0.940
Cirrhosis rate in F4 class (%)§	14.6	27.1	**0.033**
Cirrhosis prevalence by fibrosis class‖			
F0/1	3.0	2.9	0.499
F1/2	1.3	0.8	0.953
F2 ± 1	4.8	5.5	0.726
F3 ± 1	27.5	27.7	0.962
F3/4	67.6	72.2	0.676
F4	82.4	78.8	0.940

Comparison of performance characteristics between CirrhoMeter^{2G} and Fibroscan.

*P by paired McNemar test for overall accuracy and accuracy by F_M stage. Other comparisons are less powerful (unpaired χ^2 test). Significant differences are indicated in bold.

†Sensitivity for the F_M stage.

‡Positive predictive value for F stages.

§Sensitivity for cirrhosis.

‖Positive predictive value for cirrhosis.

Validation

The fibrosis classifications of both tests were independently validated by a progressive increase in the area of portal fibrosis as a function of their 6 classes (Fig. 2), $P < 0.001$ by ANOVA. We then compared between fibrosis classes the mean values of the following fibrosis descriptors: liver morphometry, Metavir staging, and noninvasive tests. Most fibrosis descriptors were significantly different between the adjacent classes F1/2, F2 ± 1, F3 ± 1, and F3/4 with both tests (details in Tables S3 and S4, Supplemental Digital Content 4 and 5, http://links.lww.com/JCG/A140; http://links.lww.com/JCG/A141 in supplement). In addition, Fibroscan had more significant differences in fibrosis descriptors between F3/4 and F4 classes, especially with the area of portal or whole fibrosis, whereas CirrhoMeter^{2G} had more significant differences between F0/1 and F1/2 classes.

Precision

Classification precision was evaluated with 4 criteria comparing the F stages, provided by test classification, and histologic reference, with stages expressed as a numerical score (Table 3). First, F correlations between test classification and histologic reference were evaluated by nonparametric correlations. Second, F agreement was evaluated by concordance index between test classification and Metavir reference. Third, F exactness was tested by the F difference between test classification and Metavir staging. Finally, classification dispersion was evaluated by the mean F stage number per class. Briefly, Fibroscan had a better index than CirrhoMeter^{2G} in each comparison (Table 4).

Simplification

Construction: We were able to merge the 6 fibrosis classes into 3 fibrosis categories. The first category (patients, CirrhoMeter^{2G}: 27.1%, Fibroscan: 40.0%) was characterized by *discrete* baseline fibrosis ($F_{0/}1$ and $F1_{/2}$, ie, F1 ± 1) and by a small progression in the area of portal

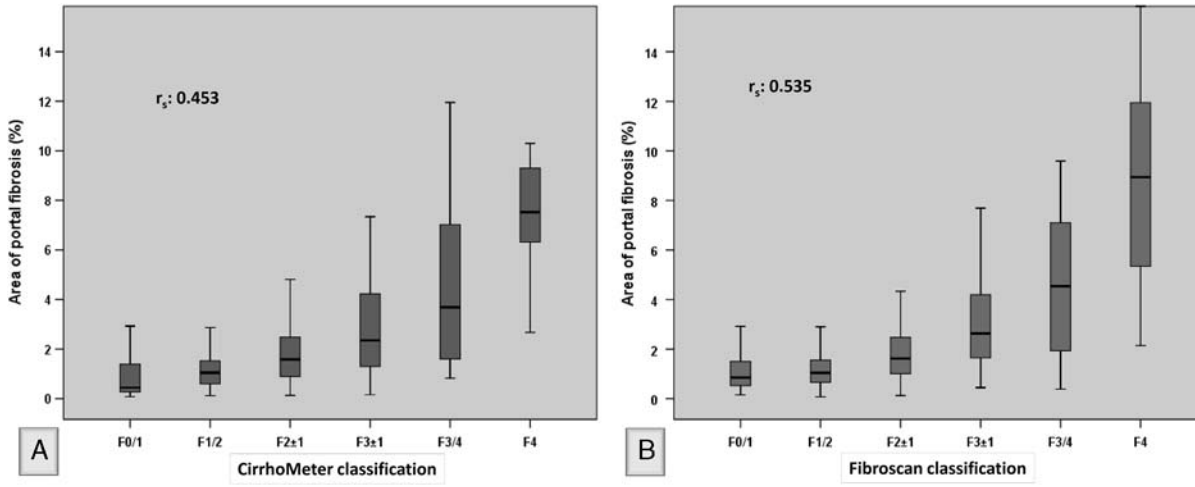

FIGURE 2. Comparison of area of portal fibrosis (*y*-axis) as a function of fibrosis classification (*x*-axis). A, CirrhoMeter^{2G}; B, Fibroscan.

fibrosis (≈0.6% until the next fibrosis class, Table S5, Supplemental Digital Content 6, http://links.lww.com/JCG/A142). The second category (42.9%, 34.9%) included patients with *moderate* baseline fibrosis (F1/$2_{/3}$ or F2 ± 1) with a risk of accelerated progression in the area of portal fibrosis (≥ 1.5%). The third category (30.0%, 25.0%) corresponded to *severe* baseline fibrosis ($F_{2/}3/4$, $F_{3/}4$, and F4, ie, F3/4) and accelerated progression in the area of portal fibrosis (≥ 1.8%).

Validation: This simplified classification was validated as follows. First, all fibrosis descriptors were significantly different between the adjacent classes of categories (Tables S3 and S4, Supplemental Digital Content 4 and 5, http://links.lww.com/JCG/A140; http://links.lww.com/JCG/A141 with the exception of sinusoidal fibrosis, which was roughly stable across Metavir fibrosis stages). Second, significant differences in overall survival as a function of these 3 categories ($P < 0.001$ by log-rank test) were observed in the cohort of 1110 patients (Fig. 3). In addition, CirrhoMeter^{2G} categories were more discriminant ($\chi^2 = 37.9$ by log-rank test) than Fibroscan ($\chi^2 = 19.7$) in predicting differences in survival.

Comparison of Staging Systems

Overall accuracy of CirrhoMeter^{2G} was significantly higher with fibrosis classification (88.8%) compared with binary diagnosis (79.2%, $P < 0.001$). The overall accuracy of Fibroscan was not significantly different between fibrosis classification (88.2%) and binary diagnosis (85.4%, $P = 0.127$).

Other diagnostic indices are depicted in Table 2. Taken together, these results showed that fibrosis classification provided globally similar or better diagnostic indices for cirrhosis when compared with binary diagnosis, regardless of the class cutoff. Thus, sensitivity and negative predictive value (NPV) for cirrhosis were maximized by classes including F4 (F3 ± 1, F3/4, and F4), whereas specificity and PPV for cirrhosis were optimized by the F4 class. A comparison of PPV and NPV for cirrhosis between binary diagnosis and fibrosis classification is shown in Figure 1B. Briefly, the PPV of binary diagnosis by classic cutoff was insufficient: 49.1% for Fibroscan and 38.9% for CirrhoMeter^{2G}, whereas F4 classes of classifications provided much higher PPV (around 70% to 80%).

TABLE 4. Fibrosis Classification Precision

	Metavir	CirrhoMeter^{2G}	*P**	Fibroscan
Correlation†				
F Metavir	—	0.578	0.420	**0.602**
Area of portal fibrosis	0.697‡	0.453	0.034	**0.535**
Agreement				
With F Metavir§	—	0.705	—	**0.746**
Exactness				
F difference with Metavir‖	—	0.76 ± 0.60	< 0.001	**0.68 ± 0.57**
Dispersion				
Mean n F stages/class	1	2.63 ± 0.533	< 0.001	**2.45 ± 0.587**

Comparison between CirrhoMeter^{2G} and Fibroscan.

Classification was used here as a numerical score: F0/1: 0.5, F1/2: 1.5, F2 ± 1: 2, F3 ± 1: 3, F3/4: 3.5, F4: 4. Best results between noninvasive tests are depicted in bold.

*Between CirrhoMeter^{2G} and Fibroscan by Steiger test for correlation and paired *t* test for difference.

†Spearman ρ, all with $P < 0.001$.

‡$P < 0.001$ versus CirrhoMeter^{2G} or Fibroscan.

§Intraclass correlation coefficient.

‖Absolute difference in F stage (mean ± SD).

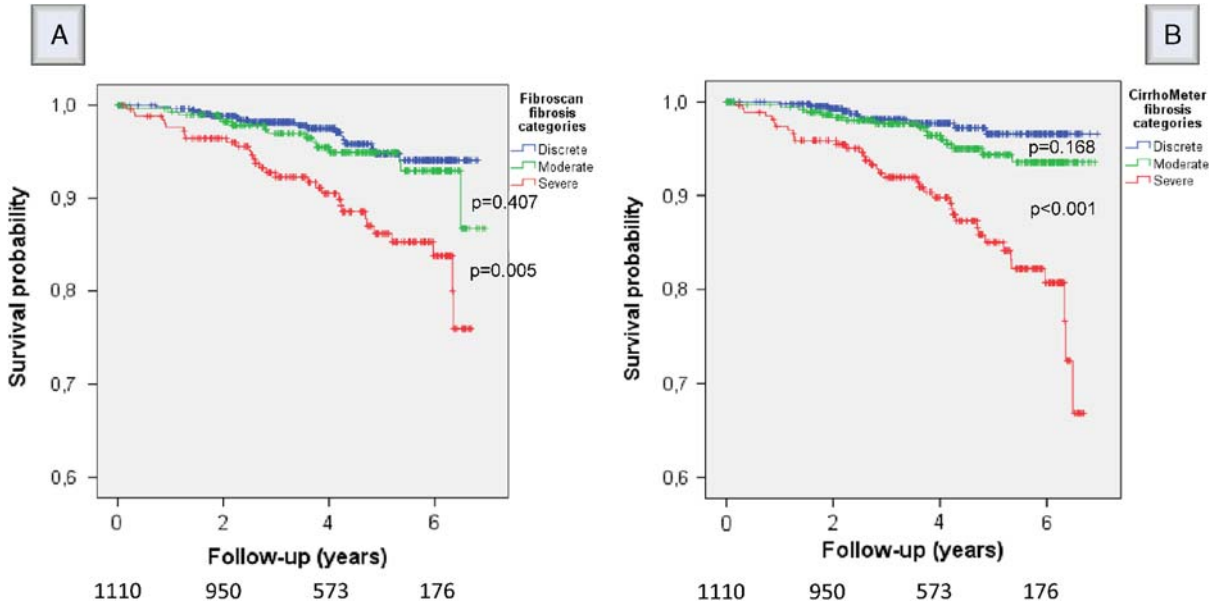

FIGURE 3. Kaplan-Meir plots of overall survival as a function of the 3 fibrosis categories for Fibroscan (A) and CirrhoMeter^{2G} (B). The figures in the bottom line indicate the number of patients remaining in follow-up at years 0, 2, 4, and 6.

DISCUSSION

The present study has 2 original aspects: we developed fibrosis classifications for Fibroscan and CirrhoMeter^{2G}; and we provided a detailed comparison between both tests.

Binary Diagnosis

Binary diagnosis has several flaws. There is no consensus on the cutoff for cirrhosis diagnosis with Fibroscan. Indeed, in the literature, this diagnostic cutoff varies from 11.9 to 16.2 kPa,[4,5,22,23] that is, in the F3 ± 1 class in the present study. It should be noted that cutoffs are sometimes determined by classic rules (eg, the Youden index) or according to expert preferences, such as maximized sensitivity. These diverging rules naturally contribute to the variability of reported cutoff values. It should be noted that the cutoff for 90% PPV was 33.5 kPa, whereas the usual cutoff of 14 kPa used by clinicians had only a PPV of 57% in our series (details not shown).

CirrhoMeter^{2G} is a recently developed test that has been validated in 2 large multicentric populations: in 382 HCV patients, AUROC for cirrhosis were 0.90 and 0.93 for CirrhoMeter^{2G} and Fibroscan, respectively,[12] and in 1204 patients with chronic viral hepatitis (HBV, HCV, HIV coinfection) they were 0.86 and 0.90, respectively,[6] figures similar to those in the present population. Finally, in a recent systematic review comparing 15 blood tests in HCV patients, FibroMeter or CirrhoMeter^{2G} had the highest AUROC for cirrhosis: 0.91.[24]

Fibrosis Classification

Our new fibrosis classification was validated by several characteristics: good correlation with the area of portal fibrosis, high accuracy, and significant differences in fibrosis descriptors.

This classification had 2 main advantages. First, accuracy was increased compared with binary diagnosis. Thus, the fibrosis classification erased the accuracy differences observed between blood test and Fibroscan for binary cirrhosis diagnosis. Second, precision was increased from 2 to 6 fibrosis classes. It might be argued that 2 classes of the fibrosis classification included 3 fibrosis stages, but that is still better than the 4 fibrosis stages comprised in the F0-3 class of binary diagnosis. Moreover, these 2 classes included a much lesser proportion of patients than the F0-3 class. In addition, in the 2 fibrosis classes including the threshold of significant fibrosis (F1/2 and F2 ± 1), it was possible to further refine the fibrosis staging (Fig. 1A).

Considering Fibroscan, the classification used by clinicians combines cutoffs for significant fibrosis and cirrhosis, thus providing a 3-class classification: F0/1, F2/3, and F4. We have shown that its accuracy was only 51% to 65%.[8] This low accuracy, despite a modest precision, is because of the addition of the errors of each cutoff. In fact, this empirical method has no statistical basis. Finally, we propose a simplified fibrosis classification with only 3 fibrosis categories: discrete (F1 ± 1), moderate (F2 ± 1), and severe (F3/4). This simplified classification differs from that classically used for Fibroscan (and blood tests) and provides higher accuracy. It can be used together with detailed classifications such as Child-Pugh classes and score.

Our new detailed fibrosis classification has other interesting aspects as well. In addition to accuracy, the other diagnostic indices (sensitivity, specificity, NPV, PPV) were superior or equal to those of binary diagnosis. The F4 class had a high PPV (around 80%) and non-F4 classes had high NPV (≥ 95%). This was obtained at the expense of an intermediate zone (F3 ± 1, F3/4). However, it should be noted that the classification method stated the cirrhosis PPV: around 28% in the F3 ± 1 class and around 70% in the F3/4 class.

One can assume that some variability might also occur during the determination of cutoffs for the fibrosis classes. However, this putative flaw would have less impact in a fibrosis classification than in a binary diagnosis. Indeed, in the latter, an error means the opposite diagnosis (F0-3 vs. F4), whereas with classification the error would be limited

to a maximum of a single fibrosis stage (eg, F1/2 vs. F2 ± 1). This explains the robust accuracy of fibrosis classification observed in large independent series.[8]

Limits

The present study does have some limits. First, although liver biopsy is the best standard for fibrosis tests, it cannot be considered as a "gold" standard.[25] However, the comparison between diagnostic tests should not be affected by this putative chance of misdiagnosis. In addition, classification was validated by prognostication. Second is the rate of patients misclassified by fibrosis classification in F_M0. However, the rate of significant discordance (>1 F_M stages) was only 4% in the F0/1 class.

Fibrosis classification might be considered as unnecessary regarding the 2 usual diagnostic targets used in chronic viral hepatitis: significant fibrosis and cirrhosis. However, fibrosis classification has excellent prognostication and diagnostic precision. Patient follow-up may differ according to the precise stage. For example, a patient with F2 has no immediate risk of complication and can be cared for without immediate intervention, whereas an F3 patient should be closely monitored for active intervention (drug, complication screening), knowing that care management will depend on the local resources.

Finally, as diagnostic cutoff of tests can differ according to etiology, prognostic value of diagnostic tests might also differ according to etiology. However, we have observed that the present diagnostic tests had the highest prognostic value in chronic viral hepatitis as well as in nonalcoholic fatty liver disease or alcoholic liver disease.[13]

Clinical Application

Staging Interpretation

The classifications include several fibrosis classes that overlap Metavir fibrosis targets, implying different clinical care provision paths. There are 2 ways to handle this issue.

First, classification can be classically interpreted as a function of Metavir staging. Considering significant fibrosis, in the $F1_{/2}$ class of Fibroscan, there is a significant gradient of fibrosis degree from the $F1_{/2}$ to the F1/2 subclass (Fig. 1A). In a class or subclass including F1 and F2 stages, the clinician might additionally monitor the patient and consider active intervention if significant progression (eg, ≥ 0.10 in CirrhoMeter^{2G} or ≥ 5 kPa with Fibroscan) is observed; an alternative option would be to consider another test such as liver biopsy. Considering cirrhosis, the classification should be considered according to the application. In clinical trials on cirrhosis, patients can be selected by the F4 class; in trials excluding cirrhosis, patients classified as ≥ F3 ± 1 should be excluded. In clinical practice, patients with $F_{3/}4$ class should be considered as putative cirrhotic patients. Indeed, cirrhosis is more prevalent in these cases and furthermore several complications classically attributed to cirrhosis have been described in F3 patients with CHC.[26] Finally, in the $F_{2/}3/4$ class, the clinician might additionally monitor the patient or conduct another examination such as liver biopsy, that is, in only 15% to 22% of patients.

Second, classification can be interpreted in different ways. Indeed, some patients are in putative broad classes, such as $F1/2_{/3}$ and $F_{2/}3/4$. However, $F1/2_{/3}$ can be considered clinically as F2 (especially in the $F2_{/3}$ subclass of Fibroscan) due to the very poor observer variability in Metavir F2 and the significant differences compared with adjacent classes observed with other fibrosis descriptors. For this same (latter) reason, $F_{2/}3/4$ can be considered as F3. By extension, the fibrosis classification might be considered per se with its own fibrosis test (F_T) stages: F_T1 to F_T6. Finally, according to prognostication, classification may be resumed as discrete (F_T1-2), moderate (F_T3), and severe (F_T4-6) fibrosis.

Elastometry or Blood Test?

Elastometry can be considered as the reference noninvasive test and Fibroscan might be superior to ARFI.[22] However, the usual cutoffs for establishing a binary diagnosis using Fibroscan are inappropriate and should be replaced by other cutoff values such as those determined in the present study (Fig. 1). A cirrhosis-specific blood test compares more favorably with transient elastometry than do the usual blood tests. It could be an alternative option when transient elastometry is unavailable, unsuccessful, or unreliable (Fig. 4). Thus, when Fibroscan fails, physicians can call upon a blood test (FibroMeter^{2G} or CirrhoMeter^{2G}) and have reasonable faith in its results as the main causes of failure for Fibroscan (metabolic syndrome, cholestasis, right cardiac failure) are different from those of FibroMeter^{2G} or CirrhoMeter^{2G} (renal dysfunction, vitamin K deficiency, marked inflammatory syndrome), although both test results are altered by acute hepatitis. However, CirrhoMeter^{2G} had 2 advantages in our study: it offered more discriminant and precise diagnosis in low F_M stages, and higher prognostic value. This latter important putative advantage should be confirmed in further studies.

Which Blood Test?

CirrhoMeter^{2G} competes with FibroMeter^{2G}. In fact, each test has an area of better accuracy: CirrhoMeter^{2G} results should be preferred when CirrhoMeter^{2G} class is ≥ F2 ± 1 (personal data), with the knowledge that both tests can be simultaneously calculated as they are based on the same markers.

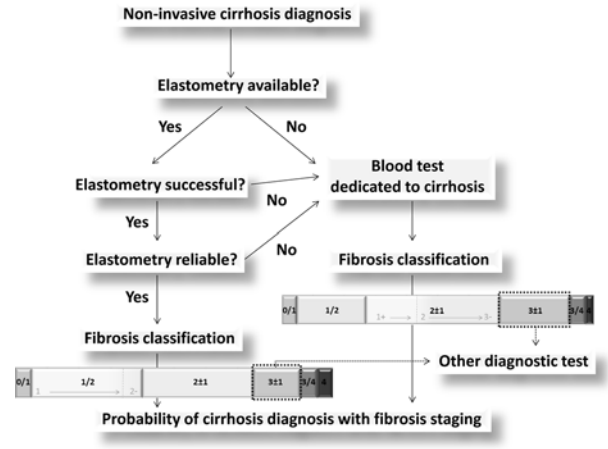

FIGURE 4. Cirrhosis diagnosis in clinical practice using noninvasive tests. Dashed rectangle indicates intermediate zone. Another possibility is to combine transient elastometry and a blood test, which decreases the intermediate zone. Optionally, a blood test dedicated to significant fibrosis, such as FibroMeter, may be preferred when CirrhoMeter^{2G} indicates F2 ± 1.

CONCLUSIONS

Currently available noninvasive tests offer good performance for cirrhosis diagnosis. Fibroscan outperforms CirrhoMeter^{2G} for cirrhosis diagnosis, but the latter may offer better prognostication. A precise fibrosis classification offers several advantages over classic binary diagnosis. However, there is still an intermediate zone that should be interpreted within the clinical context and further refined by complementary means, such as imagery studies.[27]

ACKNOWLEDGMENTS

The authors thank other investigators from Fibrostar study.

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. Vergniol, Bordeaux; O. Goria, Rouen; M. Maynard-Muet, C. Trepo, 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. Bourliere, 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 Trocmé, Grenoble; J. Guéchot, R. Lasnier, M. Vaubourdolle, Paris; H. Voitot, Beaujon, Paris; A. Vassault, Necker, Paris; A. Rosenthal-Allieri, Nice; A. Lavoinne, F. Ziegler, Rouen; M. Bartoli, C. Lebrun, Chambéry; A. Myara, Paris Saint-Joseph; F. Guerber, A. Pottier, Elibio, Vizille; M.C. Beauvieux, Bordeaux.

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

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

The authors also thank the following contributors: Sandrine Bertrais, Pascal Veillon, Gwénaëlle Soulard, and Kevin L. Erwin (for English proofreading; support: Angers University).

REFERENCES

1. Poynard T, Aubert A, Bedossa P, et al. A simple biological index for detection of alcoholic liver disease in drinkers. *Gastroenterology*. 1991;100:1397–1402.
2. Oberti F, Valsesia E, Pilette C, et al. Noninvasive diagnosis of hepatic fibrosis or cirrhosis. *Gastroenterology*. 1997;113: 1609–1616.
3. Mallet V, Gilgenkrantz H, Serpaggi J, et al. Brief communication: the relationship of regression of cirrhosis to outcome in chronic hepatitis C. *Ann Intern Med*. 2008;149:399–403.
4. Ganne-Carrie N, Ziol M, de Ledinghen V, et al. Accuracy of liver stiffness measurement for the diagnosis of cirrhosis in patients with chronic liver diseases. *Hepatology*. 2006;44: 1511–1517.
5. 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.
6. Degos F, Perez P, Roche B, et al. Diagnostic accuracy of FibroScan and comparison to liver fibrosis biomarkers in chronic viral hepatitis: a multicenter prospective study (the FIBROSTIC study). *J Hepatol*. 2010;53:1013–1021.
7. Boursier J, Bacq Y, Halfon P, 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.
8. Boursier J, Bertrais S, Oberti F, et al. Comparison of accuracy of fibrosis degree classifications by liver biopsy and non-invasive tests in chronic hepatitis C. *BMC Gastroenterol*. 2011;11:132.
9. Cales P, de Ledinghen V, Halfon P, et al. Evaluating the accuracy and increasing the reliable diagnosis rate of blood tests for liver fibrosis in chronic hepatitis C. *Liver Int*. 2008;28:1352–1362.
10. Leroy V, Halfon P, Bacq Y, 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.
11. Boursier J, de Ledinghen V, Zarski JP, 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.
12. Zarski JP, Sturm N, Guechot J, et al. Comparison of nine blood tests and transient elastography for liver fibrosis in chronic hepatitis C: The ANRS HCEP-23 study. *J Hepatol*. 2012;56:55–62.
13. Bertrais S, Boursier J, Oberti F, et al. Prognostic models for mortality using the successive values of non-invasive fibrosis tests in patients with chronic liver disease. *J Hepatol*. 2014;60:S32.
14. Nousbaum JB, Cadranel JF, Bonnemaison G, et al. Clinical practice guidelines on the use of liver biopsy. *Gastroenterol Clin Biol*. 2002;26:848–878.
15. Cales P, Boursier J, Chaigneau J, et al. Diagnosis of different liver fibrosis characteristics by blood tests in non-alcoholic fatty liver disease. *Liver Int*. 2010;30:1346–1354.
16. Sandrini J, Boursier J, Chaigneau J, et al. Quantification of portal-bridging fibrosis area more accurately reflects fibrosis stage and liver stiffness than whole fibrosis or perisinusoidal fibrosis areas in chronic hepatitis C. *Mod Pathol*. 2014. [Epub ahead of print].
17. Cales P, Veillon P, Konate A, et al. Reproducibility of blood tests of liver fibrosis in clinical practice. *Clin Biochem*. 2008;41:10–18.
18. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. *J Hepatol*. 2008; 48:835–847.
19. Boursier J, Zarski JP, de Ledinghen V, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology*. 2013;57:1182–1191.
20. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837–845.
21. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic acuracy: explanation and elaboration. *Clin Chem*. 2003;49:7–18.
22. Sporea I, Bota S, Peck-Radosavljevic M, et al. Acoustic Radiation Force Impulse elastography for fibrosis evaluation in patients with chronic hepatitis C: an international multicenter study. *Eur J Radiol*. 2012;81:4112–4118.
23. Stebbing J, Farouk L, Panos G, et al. A meta-analysis of transient elastography for the detection of hepatic fibrosis. *J Clin Gastroenterol*. 2010;44:214–219.
24. Chou R, Wasson N. Blood tests to diagnose fibrosis or cirrhosis in patients with chronic hepatitis C virus infection: a systematic review. *Ann Intern Med*. 2013;158:807–820.
25. Bedossa P, Carrat F. Liver biopsy: the best, not the gold standard. *J Hepatol*. 2009;50:1–3.
26. Sanyal AJ, Fontana RJ, Di Bisceglie AM, et al. The prevalence and risk factors associated with esophageal varices in subjects with hepatitis C and advanced fibrosis. *Gastrointest Endosc*. 2006;64:855–864.
27. Berzigotti A, Abraldes JG, Tandon P, et al. Ultrasonographic evaluation of liver surface and transient elastography in clinically doubtful cirrhosis. *J Hepatol*. 2010;52:846–853.