Journal of Viral Hepatitis, 2006, **13,** 659–670 doi:10.1111/j.1365-2893.2006.00736.x

Validation and comparison of indexes for fibrosis and cirrhosis prediction in chronic hepatitis C patients: proposal for a pragmatic approach classification without liver biopsies

M. Bourliere,[1] G. Penaranda,[2] C. Renou,[3] D. Botta-Fridlund,[4] A. Tran,[5] I. Portal,[4] L. Lecomte,[1] P. Castellani,[1] M. A. Rosenthal-Allieri,[6] R. Gerolami,[4] D. Ouzan,[7] R. Deydier,[2] C. Degott[8] and P. Halfon[9] [1]*Department of Hepato-Gastroenterology, Saint-Joseph Hospital, Marseille, France;* [2]*Department of Biostatistics and Epidemiology, CDLPharma, Marseille, France;* [3]*Department of Hepato-Gastroenterology, Hyères Hospital, Hyères, France;* [4]*Department of Hepato-Gastroenterology, La Conception Hospital, Marseille, France;* [5]*Department of Hepato-Gastroenterology, Archet Hospital, Nice, France;* [6]*Immunology Central Laboratory, Archet Hospital, Nice, France;* [7]*Department of Hepato-Gastroenterology, Arnault Tzanck Institute, Saint-Laurent du Var, France;*[8]*Anatomopathology Laboratory, Beaujon Hospital, Clichy, France; and* [9]*Virology Department, Alphabio Laboratory, Marseille, France*

Received June 2005; accepted for publication September 2005

SUMMARY. Noninvasive indexes have been developed to predict fibrosis staging. The aim of this study was to assess the diagnostic accuracy of these indexes in comparison with liver histology in hepatitis C virus (HCV)-infected patients. A total of 235 consecutive patients with HCV infection from the Fibropaca multicentre independent study were included in this paper. FibroTest (FT), aspartate aminotransferase to platelet ratio index (APRI) and Forns score were assessed in the cohort and compared with liver histology performed on the same day. The main end point was the area under characteristic curves (AUCs) for the diagnosis of significant fibrosis (F2–F4) and cirrhosis (F4) by the METAVIR classification. Mean age was 46 (±11) years, 55% were males, 42% ($n = 99$) had significant fibrosis (F2–F4) and 7% ($n = 16$) had cirrhosis (F4). For the diagnosis of significant fibrosis, respective AUCs of FT, APRI and Forns score were 0.81 (95% confidence interval: 0.76–0.86), 0.71 (0.67–0.79) and 0.76 (0.70–0.82); for cirrhosis prognosis, AUCs of FT and APRI were 0.82 (0.77–0.87) and 0.81 (0.76–0.86) (AUCs not significantly different). Using each index independently, all patients were classified by FT, 214 (91%) patients were classified by APRI and 129 (55%) by Forns score. There were significantly more cases of discordances between APRI and liver biopsy than between FT or Forns score and liver biopsy ($P < 0.05$). Performing all scores (FT, Forns and APRI) without liver biopsy allowed fibrosis to be well evaluated in 191 patients (81.3%), including patients with FT failure. Liver biopsy remained mandatory to evaluate fibrosis in 44 patients (18.7%). Our study shows that performing all the tests and liver biopsy improves the diagnostic accuracy for liver fibrosis in chronic hepatitis C patients without patent comorbidities. The combination of all tests with liver biopsy allowed 225/235 (96%) patients to be correctly classified. The combination of all tests without liver biopsy allowed 191/235 (81.3%) patients to be correctly classified; liver biopsy remained mandatory in some patients (18.7%).

Keywords: aspartate aminotransferase to platelet ratio index, chronic hepatitis C, cirrhosis, fibrosis, fibrosis biochemical markers, FibroTest, Forns, hepatitis C virus, liver biopsy.

INTRODUCTION

Liver biopsy is recommended for the management of patients infected by hepatitis C virus (HCV) and is currently the 'gold standard' in assessing liver histology, but it is occasionally prone to difficulties. The difficulties associated with liver biopsy include its highly invasive nature and a risk of complications with morbidity between 0.3 and 0.6% and mortality of 0.05% [1].

Many studies have been performed to evaluate the use of readily available laboratory test to predict significant fibrosis or cirrhosis in patients with HCV and substantially reduce the number of biopsies performed for the management of HCV infection [2–4].

The FibroTest (FT) score is computed with the patient's age, sex and results of analyses of serum haptoglobin, α2-macroglobulin, apolipoprotein A1, γ-glutamyl transpeptidase (GGT)

Abbreviations: HCV, hepatitis C virus; FT, FibroTest; AUCs, area under curves; GGT, γ-glutamyl transpeptidase; NPV, negative predictive value; PPV, positive predictive value; APRI, aspartate aminotransferase to platelet ratio index; US, ultrasound.

Correspondence: Dr Philippe Halfon, Laboratoire Alphabio, 23 rue de Friedland, 13006 Marseille, France. E-mail: philippe.halfon@alphabio.fr

and bilirubin levels. French investigators reported that with appropriate cut-off values, the FT gave a 100% negative predictive value (NPV) for the absence of significant fibrosis and a 91% positive predictive value (PPV) for its presence [2].

Wai *et al.* [3] developed the aspartate aminotransferase (AST) to platelet ratio index (APRI) which is the ratio between AST and platelet count. They reported that, when appropriate cut-off values were applied, APRI index gave a 90% NPV for the absence of significant fibrosis and a 91% PPV for its presence. For cirrhosis prediction, the APRI index gave a 100% NPV for the absence of cirrhosis and a 95% PPV for its presence.

Forns *et al.* [4] developed the Forns score which is an algorithm including platelet count, GGT, age and cholesterol level. They reported that with appropriate cut-off values, the Forns score gave a 96% NPV for the absence of significant fibrosis and a 66% PPV for its presence.

We sought to validate and compare each of these three tests in their accuracy in predicting either the presence or the absence of significant fibrosis and the presence or absence of cirrhosis among a cohort of patients with chronic hepatitis C prospectively recruited. Moreover, we tried to find out the interest of combining scores with or without liver biopsy in order to improve the accuracy in predicting fibrosis or cirrhosis and to reduce the needs of liver biopsy.

PATIENTS AND METHODS

Patients

The cohort study included 235 patients with chronic hepatitis C from the Fibropaca study. Fibropaca was a French national, multicentre, prospective, cross-sectional study on 520 patients, which was performed in five centres in the southeast region, known for their specific expertise in hepatitis C: Saint-Joseph Hospital and La Conception Hospital (Marseille), Archet Hospital (Nice), Hyères Hospital (Hyères) and Arnault Tzanck Institute (St. Laurent du Var). All these patients were consecutively recruited from November 2002 to December 2003. All patients had chronic HCV infection documented by a positive of HCV RNA in serum. Signed informed consent was obtained from all of the patients before their inclusion. Liver biopsy and biochemical markers were performed the same day. Liver biopsy was performed in each centre and analysed by the resident pathologist. For all patients, ultrasound (US) examination was performed before liver biopsy. Information relating to the patient demographics data, risk factors, virological status, clinical examinations, biological data [platelets, prothrombin time (PT) ratio, serum albumin level] was prospectively recorded in each centre on the day of biopsy. However, serum cholesterol was recorded in only 235 patients. This explains that the actual cohort included only 235 patients. All data were anonymously recorded in the database.

Methods

Liver biopsies

Liver biopsy examinations were performed in each centre and analysed by the local pathologist, with evaluation of fibrosis stage and activity grade according to the METAVIR scoring system [5]. Fibrosis was staged on a scale of 0–4: F0 = no fibrosis, F1 = portal fibrosis without septa, F2 = few septa, F3 = numerous septa without cirrhosis, F4 = cirrhosis. The grading of activity assessed by the METAVIR system (based on the intensity of necroinflammatory activity, mainly on necrosis) was scored as follows: A0 = no histological activity, A1 = mild activity, A2 = moderate activity, A3 = severe activity [5].

To assess liver biopsy quality, Regev quality criteria were used (15 mm or more in length, five or more portal tracts and one fragment) [6]. A biopsy between 10 and 15 mm of length, with less than five portal tracts, or fragmented is considered as 'fair quality biopsy'; a 'poor quality biopsy' is under 10 mm of length.

A second liver examination was performed blindly by a reference pathologist (CD) for 148 patients. These include all patients who had a discordance of two points or more on fibrosis stage between FT and liver biopsy.

Biochemical markers

To perform FT, serum samples were taken, on the day of biopsy, for the determination of five serum biochemical markers: α2-macroglobulin, haptoglobin, GGT, total bilirubin and apolipoprotein A1. Biochemical marker analysis was performed in accredited laboratories following the guidelines recommended for FT assessment by the authors of the initial publication [2]. In the hospital-based cohort, GGT and total bilirubin levels were measured by Hitachi 917 Analyzer (Roche Diagnostics Corp., Mannheim, Germany) and Roche Diagnostics reagents (Roche Diagnostics Corp., Mannheim, Germany). α2-Macroglobulin, apolipoprotein A1 and haptoglobin were measured using a Modular Analyzer (BNII, Dade Behring, Marburg, Germany). Platelets were measured by Beckman Coulter LH 750 (Beckman Coulter France S.A., Roissy, France). Biochemical assays were performed on fresh serum, decanted and stored for 72 h maximum at +2 °C/+8 °C, protected from light. The assays of the specific proteins (α2-macroglobulin, haptoglobin and apolipoprotein A1) were carried out on serum stored at +2 °C/+8 °C for 5 days. All biochemical parameters and FT determinations were performed without the knowledge of liver biopsy results. FT formula is available on the USPTO website (http://www.uspto.gov; Patent No. 6 631 330) [7]. Moreover, stored blood samples were collected for further biochemical determination and further evaluations of noninvasive fibrosis markers. The sera were collected and stored after the patients gave their signed informed consent according to French regulations.

Predictive models.

FT, APRI and Forns score were assessed and compared in this study for 235 patients with complete serum biochemical markers. FT calculations were assessed through the Internet link of the BioPredictive group using the FT formula described above. APRI score formula and Forns score formula were taken from the respective publications [3,4]. To compare fibrosis score with liver biopsy, we used the cut-off values described in the initial publication to discriminate the absence of fibrosis and the presence of significant fibrosis or cirrhosis. We also compared fibrosis determined by the FT on a scale of 0–4 with respect to METAVIR fibrosis staging as described by the authors [8]. For the FT score from 0 to 0.21, fibrosis was staged F0, from 0.22 to 0.27 F0F1, from 0.28 to 0.31 F1, from 0.32 to 0.48 F1F2, from 0.49 to 0.58 F2, from 0.59 to 0.72 F3, from 0.73 to 0.74 F3F4 and from 0.75 to 1 F4.

Discordance determination

Significant discordance between FT and biopsy was defined as a discordance of at least two stages of fibrosis in the METAVIR scoring system. Significant discordance between APRI or Forns score and biopsy was defined as discordance between the result of the test according to the respective cut-off value of each test and the result of fibrosis staging in the METAVIR scoring system.

In patients with an F4 stage on FT, APRI or liver biopsy, all radiological (US examination, CT, MRI) or endoscopic examinations and biochemical parameter analyses were performed to demonstrate cirrhosis or portal hypertension. We considered that the patients had cirrhosis if they had at least two criteria of cirrhosis among radiological (portal hypertension or liver morphology abnormalities at US, CT or MRI) or biological criteria [low blood platelets (<150 G/L) or low PT ratio (<80%)].

We considered FT failure to be an isolated abnormal value of one of the five FT components attributable to a clinically identified condition such as haemolysis with haptoglobin <0.30 g/L, inflammation or sepsis with haptoglobin >2 g/L and/or α2-macroglobulin >3 g/L, Gilbert's disease or an extra-hepatic cholestasis with elevated bilirubin and/or GGT. We considered APRI failure to be an abnormal value of one of the two components, especially normal AST value, that can lead to the underestimation of fibrosis. We considered Forns score failure to be an isolated abnormal value of one of the three biological Forns score components attributable to a clinically identified condition such as hypercholesterolaemia, drug-induced GGT elevation or inflammation leading to high platelet value. We considered biopsy failure to be a short biopsy with few portal tracts.

To summarize, we considered that

- Discordance was highly attributable to biopsy failure, if the biopsy was of poor quality (size <10 mm) with no associated abnormal marker value.
- Discordance was considered moderately attributable to biopsy failure, if the biopsy was not of good quality (size between 10 and 15 mm) with fewer than five portal tracts or fragmented but with no FT, APRI or Forns score failure.
- Discordance was considered highly attributable to FT, APRI or Forns score failure, if one of the components of each test had an abnormal value attributable to a clinically identified condition.
- Discordance was considered moderately attributable to FT, APRI or Forns score failure, if one of the components had an abnormal value in the absence of a clinically identified condition and was paired with a biopsy of good quality.
- Discordance was considered undetermined, if there was no risk of failure or there was a simultaneous risk of failure between two modes of evaluation (FT, APRI, Forns score and liver biopsy).

Moreover, in case of undetermined discordance between two tests or between one test and liver biopsy, we looked at the results of the other tests in order to determine the potential interest of those tests in resolving the discordance.

Statistical analysis

The area under the curve (AUC), sensitivity, specificity and PPV and NPV were calculated using SAS V8.0 software (SAS Institute Inc., Cary, NC, USA). AUCs were compared using Hanley and McNeil method [9]. Distributions were compared using Kruskal–Wallis test, and Spearman rank correlation coefficient was assessed. The main end point was the AUC for the diagnosis of significant fibrosis (F2F3F4 *vs* F0F1) and cirrhosis (F4 *vs* F0F1F2F3) by the METAVIR classification.

RESULTS

Patients characteristics

A total of 235 patients with HCV infection from the Fibropaca cohort were included in this study. Patients' characteristics at the time of liver biopsy are shown in Table 1. The mean age was 46 ± 11 years and 129 (55%) were males.

Liver biopsy characteristics

We had no major clinical complications of liver biopsy in our 235 patients. The mean biopsy size was 16 ± 7.5 mm; the mean number of portal tracts was 9.4 ± 5. According to Regev size quality criteria, 117/235 (50%) patients had a biopsy ≥15 mm, 84/235 (36%) from 10 to 15 mm and 34/235 (15%) less than 10 mm; 216/235 (92%) had ≥5 portal tracts in liver biopsy. Fifteen per cent (35/235) had a biopsy ≥25 mm. According to METAVIR liver fibrosis staging, 99 (42%) patients had significant fibrosis (F2–F4) and 16 (7%) had cirrhosis (F4).

Table 1 Characteristics of the 235 patients with chronic hepatitis C

Variable	Mean (±SD)
Age	46 (11)
Male, *n* (%)	129 (55)
AST (IU/L)	52.8 (37.5)
ALT (IU/L)	75.2 (60.3)
γ-GT (IU/L)	81.9 (115.9)
Bilirubin (mg/dL)	10.9 (5.1)
α2-Macroglobulin (g/L)	2.8 (1.1)
Platelets (G/L)	210.3 (62.2)
Prothrombin time ratio	93.2%
Hyaluronate (μg/L)	50.8 (90.8)
Stage of fibrosis, *n* (%)	
0	34 (14.4)
1	102 (43.4)
2	42 (17.9)
3	41 (17.5)
4	16 (6.8)

Fibrosis and cirrhosis diagnosis

Fibrosis stages on liver biopsy and FT, APRI and Forns score were fairly correlated: respective correlation coefficients were 0.56, 0.43 and 0.49 ($P < 0.0001$; Fig. 1). Respective AUCs of FT, APRI and Forns score for the diagnosis of a significant fibrosis (≥F2) were 0.81 (95% confidence interval: 0.76–0.86), 0.71 (0.67–0.79) and 0.76 (0.70–0.82; NS) (Table 2, Fig. 2). Using the suggested cut-off values of FT, APRI and Forns score, we observed the following distribution of patients: FT scores were ≤0.1 in 30 (13%) patients and >0.6 in 68 (29%), APRI scores were ≤0.5 in 105 (45%) and >1.5 in 29 (12%) and Forns scores were <4.21 in 93 (40%) and >6.9 in 37 (15%).

According to the respective cut-off values for the absence of fibrosis, FT (≤0.1) had a high NPV of 90%, while APRI (<0.5) and Forns score (<4.21) had 71 and 79%, respectively; the respective specificities were 20, 55 and 54% (Table 2). For the presence of significant fibrosis, respective PPV of FT (>0.6), APRI (>1.5) and Forns score (>6.9) were 79, 76 and 83% with sensitivity of 55, 22 and 30%, respectively. Together, using respective lower and upper cut-off values, fibrosis scoring by liver biopsy was identical in 39% of patients to scoring by FT, in 41% by APRI, and in 44% by Forns score (not significant).

According to the respective sensitive cut-off values recommended for significant fibrosis detection, 3 of 99 patients with fibrosis on liver biopsy (3% false negative) had an FT score ≤0.1, 30 (30% false negative) had an APRI score ≤0.5 and 20 (20% false negative) had a Forns score <4.21.

Respective AUCs of the FT and APRI for the diagnosis of cirrhosis (F4) were 0.82 (0.77–0.87) and 0.81 (0.76–0.86; NS) (Fig. 3). Wai *et al.* suggested two cut-off values, 1.00 and

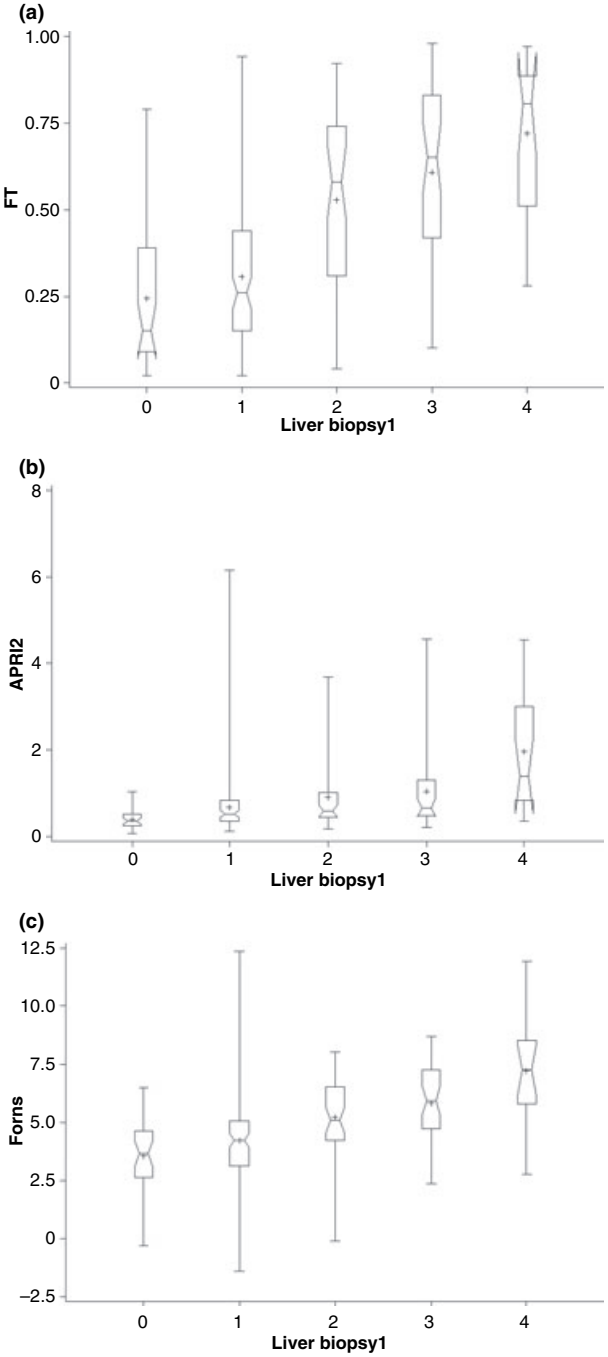

Fig. 1 Notched box plots showing the relations between the stages of fibrosis and FibroTest (a), APRI (b) and Forns score (c). The horizontal line inside each box represents the median, and the width of each box is the 95% confidence interval. Correlation between fibrosis stages and FT, APRI and Forns score was 0.56, 0.43 and 0.49, respectively ($P < 0.0001$), with significant relation between METAVIR fibrosis stages and noninvasive fibrosis indexes (Kruskal–Wallis $P < 0.0001$ for each case).

Table 2 Predictive values of FT, APRI and Forns score using original cut-off values for the diagnosis of absence and presence of fibrosis

	FT		APRI		Forns score	
	No fibrosis: ≤0.1 ($n = 30$)	Fibrosis: ≥0.6 ($n = 68$)	No fibrosis: ≤0.5 ($n = 105$)	Fibrosis: ≥1.5 ($n = 29$)	No fibrosis: <4.21 ($n = 93$)	Fibrosis: >6.9 ($n = 37$)
Sensitivity (%)	97	55	70	22	80	30
Specificity (%)	20	90	55	95	54	96
NPV (%)	90	73	71	63	79	65
PPV (%)	47	79	53	76	56	83
Diagnostic accuracy (%)	52	75	61	62	65	68
AUC	0.81 [0.76; 0.86]		0.71 [0.67; 0.79]		0.76 [0.70; 0.82]	

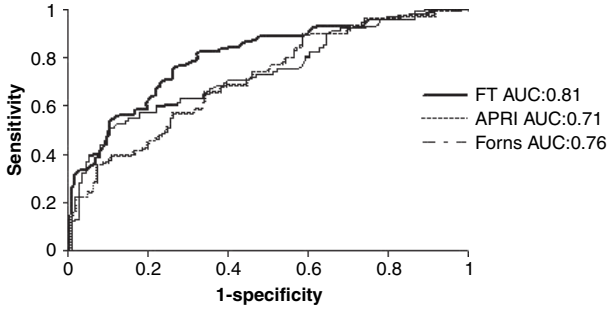

Fig. 2. ROC curves of FT, APRI and Forns score indexes in discrimination of patients with and without significant fibrosis. Respective AUCs were 0.81, 0.71 and 0.76.

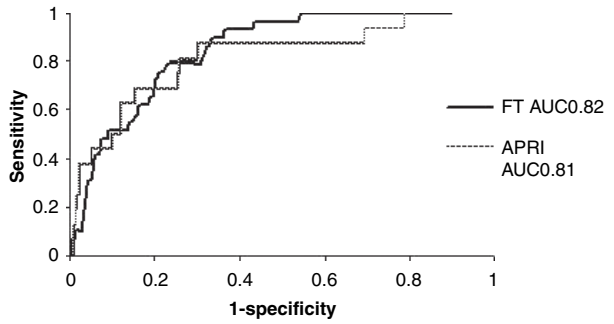

Fig. 3. ROC curves of FT and APRI indexes in discrimination of patients with and without cirrhosis. Respective AUCs were 0.82 and 0.81.

2.00, to predict the presence and absence of cirrhosis. Using these cut-off values, scores were <1.00 in 185 (79%) patients and >2.00 in 15 (6%)(Table 3). For patients with APRI of 1.00 or less, 180 of 185 (NPV of 97%) did not have cirrhosis on liver biopsy and 5 patients had cirrhosis on liver biopsy and are classified falsely by APRI. For the 15 patients with APRI >2.00, six (PPV 40%) had cirrhosis on liver biopsy, four were F3 on liver biopsy and five were below or equal to F2. Together, using APRI below the lower cut-off value (1.00) and above the upper cut-off value (2.00), 79% ($n = 186$) of

Table 3 Predictive values of APRI in cirrhosis prognosis

	APRI	
	No cirrhosis: ≤1.0 ($n = 185$)	Cirrhosis: >2 ($n = 15$)
Sensitivity (%)	69	38
Specificity (%)	82	96
NPV (%)	97	96
PPV (%)	22	40
Diagnostic accuracy (%)	81	92
AUC	0.81 [0.76; 0.86]	

the patients had identical fibrosis scoring with liver biopsy. Among the 16 patients with cirrhosis on liver biopsy, five had an APRI score ≤1.00 (31% false-negative rate).

With FT, but not with APRI and Forns score, one can assess fibrosis stage by stage. AUC analysis between FT and liver biopsy showed a poor diagnostic value between each fibrosis stage (AUCs ranged from 0.59 to 0.74). Nevertheless, the AUCs were good for differentiating F1 from F2 (AUC 0.74) and gave better diagnostic values for differences of two or more fibrosis stages, as reflected by the AUCs (from 0.73 to 0.92).

Using each test specifically, all patients were classified by FT using the continuous scale provided by Poynard *et al.* [8], 214 patients (91%) were classified by APRI (21 patients had an APRI score between 1 and 1.5) and only 130 patients (55%) were classified by Forns score (105 patients had a Forns score between 4.21 and 6.90).

Among our cohort, only 69 patients (29%) were in total agreement regarding liver biopsy, FT, APRI and Forns score. This low number is mainly due to the limitations of some tests, for instance Forns, in classifying patients.

Discordance analysis

A total of 46/235 patients (19%) were discordant for fibrosis staging between liver biopsy and FT. Surprisingly, there was

no difference regarding discordance between liver biopsy and FT according to the length of biopsy: 6/35 discordance (17%) in patients with liver biopsy ≥25 mm and 40/200 (20%) in patients with liver biopsy ≤25 mm (P = NS). This was confirmed by the second liver examination performed blindly by a reference pathologist. Of 131 patients, 37 (28%) were discordant for the presence (APRI < 0.5) or absence (APRI > 1.5) of significant fibrosis between liver biopsy and APRI. Of 199 patients, 11 (5.5%) were discordant for the presence (APRI < 1) or absence (APRI > 2) of cirrhosis between liver biopsy and APRI. Of 130 patients, 26 (20%) were discordant for the presence or absence of significant fibrosis between liver biopsy and Forns score. APRI had significantly more discordant cases than FT ($P = 0.03$) or Forns score ($P = 0.05$) for the detection of significant fibrosis as compared with liver biopsy.

Discordance analysis with cirrhosis

FT vs liver biopsy discordances

There were 21 patients with cirrhosis on either test (liver biopsy or FT). Cirrhosis was diagnosed in 4 patients by liver biopsy and in 17 by FT (Table 4). The biological, radiological and endoscopic evaluations suggested cirrhosis in four patients. Among these patients, one was determined F4 by FT and APRI, one was determined F4 by FT alone and two by liver biopsy alone. Seventeen other patients had no biological, radiological and endoscopic signs of cirrhosis. Among these 17 discordant patients, the combination of all tests allowed 13 patients to be well classified.

APRI vs liver biopsy discordances

There were 19 patients with cirrhosis on either test (liver biopsy or APRI). Cirrhosis was diagnosed by liver biopsy in 10 patients and by APRI in 9 patients (Table 5). The biological, radiological and endoscopic evaluation suggested cirrhosis in three patients. Among the 16 discordant patients without biological, radiological and endoscopic signs of cirrhosis, the combination of all tests allowed 13 patients to be well reclassified.

Discordance analysis without cirrhosis

FT vs liver biopsy discordances

Twenty-five patients without cirrhosis had discordance between FT and liver biopsy (Table 6). Among the 142 patients with FT < F2 (absence of fibrosis), 13 were discordant with liver biopsy. Among these 13 discordant patients, the combination of all tests allowed 12 patients to be well reclassified.

Among the 50 patients with FT = F2 or F3, 12 were discordant with liver biopsy. Among these 12 discordant cases, the combination of all tests allowed all patients to be well reclassified.

APRI vs liver biopsy discordances

Thirty-five patients without cirrhosis had discordance between APRI and liver biopsy (Table 7). Among the 105 patients with APRI score ≤0.5, 28 were discordant with liver biopsy. Among these 28 discordant patients, the combination of all tests allowed 24 patients to be well reclassified.

Table 4 FT *vs* liver biopsy discordances with cirrhosis and combination with APRI diagnosis

FT *vs* liver biopsy discordances with cirrhosis			APRI diagnosis
4 patients with biological, radiological and endoscopic signs of cirrhosis	2 patients F4 by FT	Agree with FT in 1 case Agree with liver biopsy in 1 case	
	2 patients F4 by liver biopsy		Agree with FT in 2 cases
17 patients without biological, radiological and endoscopic signs of cirrhosis	15 patients F4 by FT	6 FT failure	Agree with liver biopsy
		3 liver biopsy failure	Agree with FT in 1 case Undetermined in 2 cases
		6 undetermined	Agree with FT in 1 case Agree with liver biopsy in 3 cases Undetermined in 2 cases
	2 patients F4 by liver biopsy	2 FT: F2	Agree with FT

Biological, radiological and endoscopic evaluations suggested cirrhosis in 4 patients. Among these patients, 1 was determined F4 by FT and APRI, 1 was determined F4 by FT alone and 2 by liver biopsy alone. The 17 other patients had no biological, radiological and endoscopic signs of cirrhosis. Two patients were classified F4 by biopsy; they were F2 by FT and APRI failed to detect cirrhosis in these patients. Fifteen were classified F4 by FT and below F2 by biopsy. Cirrhosis was detected by APRI in only 2 cases. We therefore classified these patients according to the size of biopsy and evaluation of FT components. Six were FT failure, 3 were a liver biopsy failure and 6 were unexplained. In case of FT failure, APRI always agreed with liver biopsy. In case of liver biopsy failure, APRI agreed with FT in one case (1/3). In case of unexplained discordance between FT and liver biopsy, APRI agreed in one patient with FT, agreed with liver biopsy in 3 patients and was undetermined in 2 other patients.

Table 5 APRI *vs* liver biopsy discordances with cirrhosis and combination with FT diagnosis

APRI *vs* liver biopsy discordances with cirrhosis		FT diagnosis
3 patients with biological, radiological and endoscopic signs of cirrhosis	1 patient F4 by APRI	Agree with APRI
	2 patients F4 by liver biopsy	Agree with APRI (FN)
16 patients without biological, radiological and endoscopic signs of cirrhosis	8 patients F4 by APRI	Agree with APRI in 5 cases F3 by FT in 3 cases
	8 patients F4 by liver biopsy	F4 by FT in 3 cases F3 by FT in 2 cases F2 by FT in 3 cases

FN, false negative. There were 19 patients with cirrhosis on either test (liver biopsy or APRI). Cirrhosis was diagnosed by liver biopsy in 10 patients and by APRI in 9 patients. The biological, radiological and endoscopic evaluation suggested cirrhosis in 3 patients. Among these patients, 1 was determined F4 by APRI and 2 by liver biopsy. The 16 remaining patients had no biological, radiological and endoscopic signs of cirrhosis. Interestingly, among the 8 patients classified F4 by APRI, 5 were classified F4 by FT and 3 were classified F3 by FT. In these patients, even if liver biopsies were over 15 mm length in 6 patients and all had more than 5 portal tracts, the concordance between APRI and FT suggested that these patients had significant fibrosis, severe fibrosis or cirrhosis. Conversely, among the 8 patients classified F4 by liver biopsy, 4 were classified F4 by FT, 2 F3 by FT and 2 F2 by FT. Again in this situation, the concordance between liver biopsy and FT in 5 patients suggested that they had severe liver fibrosis or cirrhosis. Among these 5 patients, 2 had normal or very low AST value that can explain APRI failure. Among these 16 discordant patients with cirrhosis, the combination of all tests allowed 13 patients to be well reclassified.

Table 6 Analysis of the 25 discordant cases without cirrhosis between FT *vs* liver biopsy and combination with APRI and Forns diagnosis

FT *vs* liver biopsy discordances		APRI and Forns diagnosis
FT < F2 (no fibrosis) ($N = 13$)	4 liver biopsy failure	APRI and Forns scores confirmed failures
	1 FT failure	APRI and Forns scores confirmed failure
	8 undetermined	APRI and/or Forns agreed with FT in 7 cases 1 undetermined for APRI and Forns
FT = F2 or FT = F3 ($N = 12$)	1 liver biopsy failure	Confirmed by Forns score
	9 FT failure	Confirmed by APRI and Forns scores in 5 cases Undetermined by APRI and Forns scores in 4 cases
	1 both liver biopsy and FT failure	APRI and Forns showed no fibrosis
	1 case with no failure of liver biopsy nor FT	APRI and Forns scores confirmed FT

Twenty-five patients without cirrhosis had discordance between FT and liver biopsy. Among the 142 patients with FT <F2 (absence of fibrosis), 13 were discordant with liver biopsy: 7 were F2 and 6 were F3 by liver biopsy. Among these 13 patients, 4 were biopsy failure and 1 was FT failure. In these 5 cases, APRI and Forns score confirmed the failure. Among the remaining 8 patients with undetermined failure, in 1 case APRI and Forns score were undetermined and in the remaining 7 patients APRI and/or Forns score agreed with FT. Among these 13 discordant patients, the combination of all tests allowed 12 patients to well reclassified.Among the 50 patients with FT = F2 or F3, 12 were discordant with liver biopsy: 5 were F0 and 7 were F1 for liver biopsy. Among these 12 patients, 9 were FT failure. Of these 9 patients, 5 were confirmed by APRI and Forns score, and for the other 4 patients, APRI and Forns score were undetermined. Among the 3/12 remaining patients, 1 was a liver biopsy failure confirmed by Forns score, 1 was both liver biopsy and FT failure (Forns score and APRI showed no fibrosis), and in the last discordant case with no failure of liver biopsy and FT, Forns score and APRI confirmed FT result. Among these 12 discordant cases, the combination of all tests allowed all patients to be well reclassified.

Among the 14 patients with APRI score ≥1.5 and ≤2.0, 7 were discordant with liver biopsy. Among these seven patients, the combination of all tests allowed six patients to be well reclassified.

Forns vs liver biopsy discordances

Twenty-five patients without cirrhosis had discordance between Forns score and liver biopsy (Table 8). Among the 93 patients with Forns score <4.21 (absence of fibrosis), 19

Table 7 Analysis of the 35 discordant cases without cirrhosis between APRI *vs* liver biopsy and combination with FT and Forns diagnosis

APRI *vs* liver biopsy discordances		FT and Forns diagnosis
APRI <0.5 (no fibrosis) ($N = 28$)	8 liver biopsy failure	FT and Forns scores confirmed failures
	20 undetermined	FT and/or Forns agreed with APRI in 12 cases
		FT and/or Forns agreed with liver biopsy in 4 cases
		Undetermined in 4 cases
APRI >1.5 (fibrosis) ($N = 7$)	4 APRI failure	FT and Forns scores confirmed failures
	3 undetermined	FT and Forns agreed with APRI in 2 cases
		Undetermined in 1 case

Thirty-five patients without cirrhosis had discordance between APRI and liver biopsy. Among the 105 patients with APRI score ≤0.5, 28 were discordant with liver biopsy: 14 were F2 by liver biopsy and 14 were F3 by liver biopsy. Among these 28 cases, 8 were biopsy failure confirmed by FT and Forns score. Among the remaining 20 patients with undetermined failure, FT and/or Forns score agreed with APRI in 12 cases, with liver biopsy in 4 cases, and in the remaining 4 cases discordance could not be explained. Among these 28 discordant patients, the combination of all tests allowed 24 patients to be well reclassified. Among the 14 patients with APRI score ≥1.5 and ≤2.0, 7 were discordant with liver biopsy: all 7 patients were diagnosed F1 by liver biopsy. Among these 7 patients, 4 were APRI failure confirmed by FT and/or Forns score. Among the remaining 3 patients with undetermined failure, FT and Forns score agreed with APRI in 2 cases, and in the last case discordance failure could not be explained. Among these 7 patients, the combination of all tests allowed 6 patients to be well reclassified.

Table 8 Analysis of the 25 discordant cases without cirrhosis between Forns *vs* liver biopsy and combination with FT and APRI diagnosis

Forns *vs* liver biopsy discordances		FT and APRI diagnosis
Forns <4.21 (no fibrosis) ($N = 19$)	4 liver biopsy failure	FT and APRI scores confirmed failures
	15 undetermined	FT and/or APRI agreed with Forns in 10 cases
		FT and/or APRI agreed with liver biopsy in 4 cases
		Undetermined in 1 case
Forns >6.9 (fibrosis) ($N = 6$)	2 liver biopsy failure	FT and APRI scores confirmed failures
	4 undetermined	FT and/or APRI agreed with Forns in 2 cases
		FT and/or APRI agreed with liver biopsy in 2 cases

Twenty-five patients without cirrhosis had discordance between Forns score and liver biopsy. Among the 93 patients with Forns score <4.21 (absence of fibrosis), 19 were discordant with liver biopsy: 10 were F2 by liver biopsy and 9 were F3 by liver biopsy. Among these 19 cases, 4 were biopsy failure confirmed by FT and APRI. Among the remaining 15 patients with undetermined failure, FT and/or APRI agreed with Forns score in 10 cases, with liver biopsy in 4 cases, and in 1 case discordance could not be explained. Among these 19 discordant patients, the combination of all tests allowed 18 patients to be well reclassified. Among the 36 patients with Forns score >6.9 (presence of fibrosis), 6 were discordant with liver biopsy: all 6 patients were diagnosed F1 by liver biopsy. Among these 6 patients, 2 were liver biopsy failure confirmed by FT and APRI. Among the remaining 4 patients with undetermined failure, FT and/or APRI score agreed with Forns score in 2 cases and with liver biopsy in 2 cases. The combination of all tests allowed all 6 patients to be well reclassified.

were discordant with liver biopsy. Among these 19 discordant patients, the combination of all tests allowed 18 patients to be well reclassified.

Among the 37 patients with Forns score >6.9 (presence of fibrosis), six were discordant with liver biopsy. The combination of all tests allowed all six patients to be well reclassified.

FT vs APRI discordances

There were 26 cases of discordance between FT and APRI (Table 9). Among these 26 discordant patients, the combination of all tests allowed 24 patients to be well reclassified.

FT vs Forns discordances

There were 16 cases of discordance between FT and Forns score (Table 10). Among these 16 discordant patients, the combination of all tests allowed 14 patients to be well reclassified.

APRI vs Forns discordances

There were two cases of discordance between APRI and Forns score: both cases were diagnosed F0F1 by Forns score and F2F4 by APRI. The two cases were APRI failure confirmed by liver biopsy and FT.

Table 9 Analysis of the 26 discordant cases without cirrhosis between FT *vs* APRI and combination with liver biopsy and Forns diagnosis

FT *vs* APRI discordances ($N = 26$)	Liver biopsy and Forns diagnosis
6 FT failure	Liver biopsy and Forns scores confirmed failures
7 APRI failure	Liver biopsy and Forns scores confirmed failures
13 undetermined	Liver biopsy and Forns score agreed with APRI in 9 cases Liver biopsy and Forns score agreed with FT in 2 cases Undetermined in 2 cases

There were 26 cases of discordance between FT and APRI: 6 were diagnosed F0F1 by FT and ≥F2 by APRI, and 20 were diagnosed F0F1 by APRI and ≥F2 by FT. Among these 26 cases, 7 were APRI failure and 6 were FT failure confirmed by liver biopsy and Forns score. Among the remaining 13 patients with undetermined failure, liver biopsy and/or Forns score agreed with APRI in 9 cases, with FT in 2 cases, and in the remaining 2 cases discordance could not be explained. Among these 26 discordant patients, the combination of all tests allowed 24 patients to be well reclassified.

Table 10 Analysis of the 16 discordant cases without cirrhosis between FT *vs* Forns and combination with liver biopsy and APRI diagnosis

FT *vs* Forns discordances ($N = 16$)	Liver biopsy and APRI diagnosis
1 FT failure	Liver biopsy and/or APRI confirmed failures
9 Forns failure	Liver biopsy and/or APRI confirmed failures
6 undetermined	Liver biopsy and/or APRI agreed with FT in 2 cases Liver biopsy and/or APRI agreed with Forns in 3 cases Undetermined in 1 case

There were 16 cases of discordance between FT and Forns score: 4 were diagnosed F0F1 by FT and ≥F2 by Forns score, and 12 were diagnosed F0F1 by Forns score and ≥F2 by FT. Among these 16 cases, 9 were Forns score failure and 1 was FT failure. Among the remaining 6 patients with undetermined failure, liver biopsy and/or APRI agreed with FT in 2 cases, with Forns score in 3 cases, and in 1 case discordance could not be explained. Among these 16 discordant patients, the combination of all tests allowed 14 patients to be well reclassified.

Overall analysis

When liver biopsy, FT, APRI and Forns were analysed for each patient, 225/235 (96%) were well classified.

If only liver biopsy had been performed, we would have missed cirrhosis in four patients (two with biological, endoscopic or radiological signs of cirrhosis and two with cirrhosis on FT and APRI) and would have misdiagnosed cirrhosis in two patients. Moreover, in 12 patients, liver biopsy would have over- or underestimated fibrosis. Therefore, 18 patients (7.6%) would have been misdiagnosed by liver biopsy, whatever the length of biopsy.

If only FT had been performed, we would have missed two patients with biological, endoscopic or radiological signs of cirrhosis. Seventeen patients (7.2%) had FT Failure based on an isolated abnormal value of one of the five FT components attributable to a clinically identified condition, and 10 patients (4.2%) had an over- or underestimation of fibrosis without any FT component abnormalities.

If both FT and liver biopsy had been performed, 189 patients (81%) would have agreed in both tests and 46 would have been discordant [17 (7%) as a result of FT failure; 29 (12%) would have been undetermined]. Among them, four patients had biological, endoscopic or radiological cirrhosis diagnosed by only one test.

In a pragmatic approach, if FT had been performed with APRI and Forns in each patient, 82 patients (35%) would have had a complete concordance between the three scores, 92 patients (39%) would have had a concordance between FT and APRI (71 patients) or between FT and Forns (21 patients), 17 patients would have had FT failure diagnosed both by APRI and Forns and 44 patients (18.7%) would have had a discordance between FT and both APRI and Forns. In this situation, liver biopsy would have been necessary, and in 33 cases it would have reclassified the patients (Fig. 4). However, even with this pragmatic approach, cirrhosis would have been missed in two patients. We suggest an algorithm combining the three tests (Fig. 5), which can be used to evaluate the necessity of liver biopsy on our data set.

DISCUSSION

As liver biopsy is an invasive procedure with associated morbidity and mortality [1], several attempts have been made to find accurate noninvasive markers of fibrosis and/or necroinflammatory activity [2–4]. Our study was aimed at comparing these different attempts in their efficacy in predicting fibrosis and cirrhosis in patients with chronic HCV. The three scores studied were only moderately associated with METAVIR fibrosis stages. There was considerable overlap of FT, APRI and Forns score interquartile ranges for F2–F4 fibrosis and those for F0 or F1 fibrosis (Fig. 1).

The diagnostic accuracy of FT for significant fibrosis was comparable to that of the original study (AUC: 0.81 [0.76;

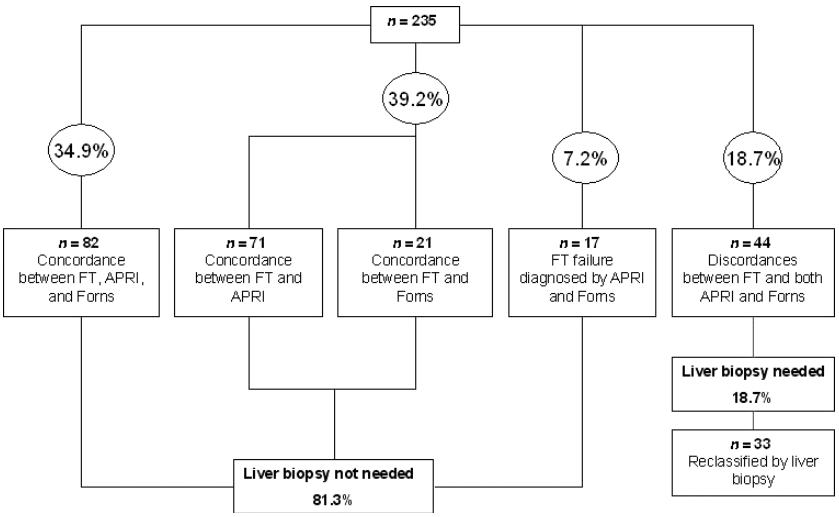

Fig. 4. Proposed algorithm for chronic hepatitis C patients without comorbidities according to FT, Forns and APRI scores results.

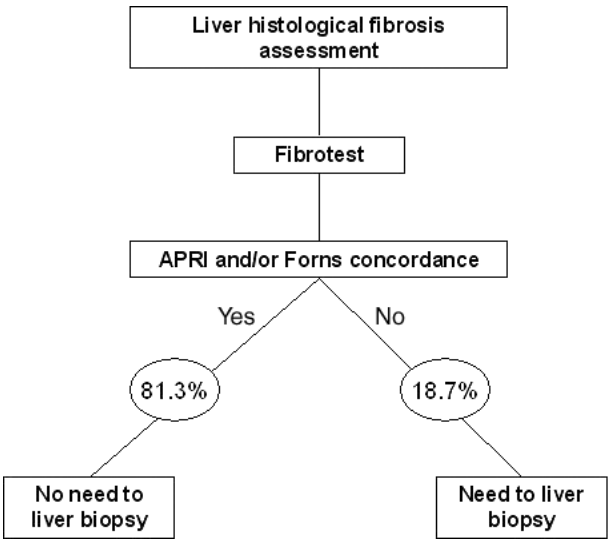

Fig. 5. Simplified algorithm for chronic hepatitis C.

0.86] *vs* 0.84, respectively) [2] and to those of others [10–13].

The diagnostic accuracy of APRI score for significant fibrosis and cirrhosis was lower than those reported in the original study [0.71 (0.65–0.77) *vs* 0.88 (0.80–0.96) for the diagnosis of significant fibrosis ($P < 0.05$) and 0.81 (0.76–0.86) *vs* 0.94 (0.89–1.00) for the diagnosis of cirrhosis ($P < 0.05$)] [3]. Our findings are similar to those in other studies [10,11,14,15]. The discrepancy with the original study may be due to the 38% of patients who had normal AST value in our cohort. Moreover, the main problem with the APRI remains uncertain regarding the appropriate definition of the upper normal limit for AST.

In our study, the diagnostic accuracy of Forns score for significant fibrosis was comparable to that of the original study [AUC: 0.76 (0.70–0.82) *vs* 0.81] [4]. Our findings are similar to those in another study [10].

Using the respective cut-off values of FT, APRI and Forns score for the absence of fibrosis or the presence of significant fibrosis or cirrhosis given in the original studies, each test classified, correctly according to liver biopsy, the same percentage of patients (92 of 235 (39%) for FT, 97 of 235 (41%) for APRI and 103 of 235 (44%) for Forns test). However, the advantage of FT over APRI and Forns score is the almost linear relationship between FT and fibrosis stage, which allows all patients to be classified; in our study, APRI was unable to classify 21 patients (9%) and Forns score was unable to classify 106 patients (45%).

For cirrhosis prognosis, the respective AUCs for FT, APRI and Forns score were 0.82, 0.83 and 0.86, with no significant differences between AUCs.

For the diagnosis of significant fibrosis, FT, APRI and Forns score had good specificity, but APRI and Forns score had very low sensitivity. However, for the diagnosis of the absence of fibrosis, FT had a better sensitivity but a lower specificity than APRI and Forns score.

Regarding all these results, the three biochemical markers studied show overall good diagnostic accuracy for significant fibrosis compared with liver biopsy.

However, in our large prospective hepatitis C cohort, only 35% of patients showed agreement between all biochemical markers. This figure may be due in part to the variability in measurements of the biochemical components for each test. For FT, all components can be measured using a single autoanalyser with minimal variability after a required standardization [16]. For APRI, AST has a poor sensitivity for fibrosis detection, and the uncertainty regarding the appropriate definition of upper normal limit remains a problem. This problem may explain the limitation of the test and the lower diagnostic power of this test as compared with FT in the literature [14]. For Forns score, the standardization of platelet count may be a problem [17].

Moreover, pathological situations such as haemolysis, inflammation, sepsis or Gilbert's disease can lead to FT fail-

ure, hypercholesterolaemia can lead to Forns score failure and normal AST value can lead to APRI failure.

In our large prospective hepatitis C cohort, liver biopsy missed cirrhosis diagnosis in 4 patients, misdiagnosed cirrhosis in 2 patients and incorrectly estimated fibrosis in 12 patients. This finding emphasizes that liver biopsy is not the gold standard it was thought to be for fibrosis assessment. Indeed, Regev *et al.* [6] demonstrated that variability in the distribution of fibrosis within the liver is a potential limitation. In a recent paper, Bedossa *et al.* [18] demonstrated that the biopsy should be at least 25 mm long to evaluate fibrosis accurately with a semiquantitative score (METAVIR scoring system). An ideal but impossible study for the validation of biochemical markers would be to perform laparoscopy with two biopsies of 20 mm to reach a total length of 40 mm. Note that inter- and intrapathologist variability is reported to be 20% (6). In our study, however, the discordance between FT and liver biopsy was surprisingly not influenced by the length of biopsy, and this was confirmed by the second liver examination performed blindly by a reference pathologist in the discordance cases. Nevertheless, even if liver biopsy is not a 'gold standard', it remains mandatory in HCV patients with comorbidities in order to determine specific abnormalities such as alcoholic liver damage, steatosis or iron overload.

One of the weaknesses of our study is the distribution of liver fibrosis in the cohort, with 42% of patients with significant fibrosis, reflecting the fact that most centres were referral centres for liver disease. This may be a limitation of our study, as noninvasive markers of fibrosis may have different diagnostic accuracy depending on the prevalence of significant fibrosis in the studied population. The design of a similar study in community-based cohorts with milder disease may have distinct results.

Our study shows that performing all the tests and liver biopsy allows for better diagnostic accuracy of liver fibrosis. In our study, cirrhosis would have been missed by liver biopsy in four patients, by FT in two patients, and by APRI in five patients. This point underlines the fact that any single test could not be a gold standard. Looking at discordances between two tests or between one test and liver biopsy, the combination of all tests allows 225/235 (96%) patients to be well classified.

However, in a pragmatic approach, performing all tests (FT, Forns and APRI) without liver biopsy allowed fibrosis to be well evaluated in 191 patients (81.3%) including patients with FT failure. Liver biopsy remained mandatory to evaluate fibrosis in 44 patients (18.7%) without considering comorbidities that may need more liver biopsies.

New noninvasive procedures such as FibroScan are in development [12, 19]. This technique has a good diagnostic accuracy for fibrosis detection and may be interesting for studying portal hypertension. The combination of FibroScan with biochemical markers may be another way to improve diagnostic accuracy for fibrosis detection.

REFERENCES

1 Cadranel JF, Rufat P, Degos F. Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AFEF). *Hepatology* 2000; 32: 477–481.

2 Imbert-Bismut F, Ratziu V, Pieroni L, Charlotte F, Benhamou Y, Poynard Tfor the MULTIVIRC group, Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. *Lancet* 2001; 357: 1069–1075.

3 Wai CT, Greenson JK, Fontana RJ *et al.* A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003; 38: 518–526.

4 Forns X, Ampurdanes S, Llovet JM *et al.* Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. *Hepatology* 2002; 36: 986–992.

5 Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The Metavir Cooperative Study Group. *Hepatology* 1996; 24: 289–293.

6 Regev A, Berho M, Jeffers LJ *et al.* Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol* 2002; 97: 2614–2618.

7 Poynard T. Diagnosis method of inflammatory or cancerous disease using the biochemical markers. United States Patent No US 6 631 330 B1, 2003.

8 Poynard T, Imbert-Bismut F, Munteanu M *et al.* Overview of the diagnostic value of biochemical markers of the liver fibrosis (Fibrotest, HCV Fibrosure) and necrosis (Actitest) in patients with chronic hepatitis C. *Comp Hepatol* 2004; 3: 8.

9 Hanley JT, McNeil BJ. A method of comparing the areas under a receiver operating characteristic (ROC) curve. *Radiology* 1983; 143: 29–36.

10 Foucher J, Vergniol J, Castéra L *et al.* Fibrosis evaluation in chronic liver diseases: comparison of Fibroscan with liver biopsy, Fibrotest, Forns score, APRI, Hyaluronan, Prothrombin time, and AST/ALT Ratio. *J Hepatol* 2005; 42 (Suppl. 2): 78 (Abstract).

11 Calès P, Oberti F, Fouchard-Hubert I *et al.* Evaluation of different morphological characteristics of liver fibrosis by blood markers. *J Hepatol* 2005; 42 (Suppl. 2): 201 (Abstract).

12 Castéra 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: 343–350.

13 Rossi E, Adams L, Prins A *et al.* Validation of the Fibrotest biochemical markers score in assessing liver fibrosis in hepatitis C patients. *Clin Chem* 2003; 49 (3): 450–454.

14 Le Calvez S, Thabut D, Messous D *et al.* The predictive value of fibrotest vs. APRI for the diagnosis of fibrosis in chronic hepatitis C. *Hepatology* 2004; 39: 862–863.

15 Lackner C, Struber G, Liegl B *et al.* Comparison and validation of simple noninvasive tests for prediction of fibrosis in chronic hepatitis C. *Hepatology* 2005; 41 (6): 1376–1382.

16 Halfon P, Imbert-Bismut F, Messous D *et al.* A prospective assessment of the inter-laboratory variability of biochemical markers of fibrosis (Fibrotest) and activity (Actitest) in patients with chronic liver disease. *Comp Hepatol* 2002; 2: 3–7.
17 Moroff G. The influence of various hematology analyzers on component platelets count. *Transfus Med Rev* 2005; 19: 155–166.
18 Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003; 38: 1449–1457.
19 Ziol M, Handra-Luca A, Kettaneh A *et al.* Non-invasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. *Hepatology* 2005; 41: 48–54.