Prospective Evaluation of FIBROSpect II for Fibrosis Detection in Hepatitis C and B Patients Undergoing Laparoscopic Biopsy

Lennox J. Jeffers, MD[1]; Rafael A. Cortes, MD[1];
Pablo A. Bejarano, MD[2]; Esther Oh, MS[3];
Arie Regev, MD[1]; Katie M. Smith, PhD[3];
Maria De Medina[1]; Margaret Smith-Riggs, PhD[3];
Marlene Colon[1]; Keith Hettinger, MS[3];
Sandra Jara, MD[1]; Tulia Patricia
Mendez[1]; Eugene R. Schiff, MD[1]

[1]*Center for Liver Diseases, Division of Hepatology,*
[2]*Department of Pathology, University of Miami
Leonard Miller School of Medicine, Miami, Fla.;*
[3]*Prometheus Laboratories, San Diego, Calif.*

Address correspondence to:
Dr. Eugene R. Schiff, Center for Liver
Diseases, Division of Hepatology,
University of Miami Leonard Miller School
of Medicine, 1500 N. W. 12th Avenue,
Suite 1101 ET, Miami, FL 33136-1090;
Fax: 305-243-6681;
E-mail: eschiff@med.miami.edu.

Abstract: Serum markers of liver fibrosis are difficult to validate, due to the sampling error and observer variability associated with percutaneous liver biopsies. Laparoscopic biopsy decreases sampling error and increases the reliability of histopathologic assessment. We prospectively evaluated the FIBROSpect[SM] II serum marker test for viral liver fibrosis against laparoscopic biopsies by studying 145 patients with chronic hepatitis B or C who underwent laparoscopy in a tertiary care setting. Serum samples obtained at biopsy were tested with FIBROSpect II to assess the degree of fibrosis. Multiple biopsies were obtained from each patient and scored blindly using the Batts-Ludwig system. An average biopsy stage was calculated and the performance of the test panel assessed. FIBROSpect II was able to rule in significant fibrosis (stages 2–4), with a likelihood ratio of 2.6. It correctly indicated absence of disease in 74% of stages 0–1 patients and correctly predicted significant disease in 67% of stages 2–4 patients. Test correlation was highest with Batts-Ludwig stages 3 (77%) and 4 (96%) and lowest with stage 2 (43%). Multiple biopsies from 52% of patients differed by at least 1 stage. In 13 patients (9%), cirrhosis was detected by laparoscopy but not histologically; in 4 (3%), a stage of 4 was obtained, but cirrhosis was not evident by laparoscopy. FIBROSpect II provided valuable additional information for assessing fibrosis. The discordance in fibrosis stage seen in multiple biopsies from the same patient underscores the need to consider all available information when assessing fibrosis. This study confirms and extends results of previous studies evaluating FIBROSpect II using percutaneous liver biopsy.

There is a growing need for a simple, safe, effective, and noninvasive method of assessing hepatic fibrosis in patients with chronic liver disease. Worldwide, over 200 million people are infected with hepatitis C virus (HCV) alone, and the incidence of chronic hepatitis is expected to increase 3-fold by 2020.[1] Estimates suggest that 20% of HCV-infected patients will develop cirrhosis, with 1–4% of those patients developing hepatocellular carcinoma.[2] Although great progress has been made in antiviral therapy, treatment is costly, there are significant side effects, and up to 50% of treated HCV patients may not achieve long-term sustained viral clearance.[3] Patients infected with genotype 1 HCV require more aggressive treatment, which creates a need to assess fibrosis to guide their treatment.[4] In addition, hepatitis B virus (HBV) is only inhibited, and not often cleared, by antiviral treatments, requiring continued assessment of fibrosis.[5] Studies are in progress to determine if maintenance therapy can significantly reduce the progression of fibrosis, even in patients who do not completely clear the viruses.[4] Noninvasive serum marker tests are critically needed to manage the increasing number of patients needing repeated assessment of liver fibrosis, both to evaluate the effectiveness of therapy and to indicate treatment options for particular types of patients.

Histopathologic examination of a liver biopsy remains the standard method by which patients are evaluated for liver disease and the extent of fibrosis. Several semiquantitative scoring systems that use integers to represent the severity of fibrosis based on the architectural disturbance of the liver and the location of fibrosis have been introduced. The degree of fibrosis does not increase linearly throughout these stages: Comparing Metavir fibrosis stages with image analysis,[6] the amount of fibrosis in stage F2 was found to increase 3-fold above that of stage F0, to increase 7-fold in stage F3, and 12-fold in stage F4. Although the degree of fibrosis as measured by these systems is an important guide for prognosis and treatment,[4] the rate of fibrosis progression varies greatly between patients, and many patients may never develop advanced disease or cirrhosis.[7] This potentially leads to serial biopsies in relatively asymptomatic patients solely to monitor the progression of disease.

Moreover, repeated biopsy carries an inherent risk for the patient, and it is widely recognized that liver biopsy often yields an inaccurate assessment of hepatic inflammation and fibrosis due to sampling errors and interobserver variability and experience.[6,8-11] Clinical considerations limit the number and size of the biopsy samples obtained, even though the accuracy of histologic assessment is known to be critically dependent on the biopsy length, the degree of fragmentation of tissue core, and the number of obtained samples.[6,11] In addition, liver biopsy is invasive, costly, frequently associated with pain, and causes serious complications in 0.1% of patients.[5] Not surprisingly, many patients are reluctant to submit to repeated biopsies.

The error rate and risks associated with biopsy have intensified the search for noninvasive serum marker tests that can identify patients with more severe fibrosis, who require more stringent management and treatment, and assess the progression of disease. Several different biomarkers or combinations of markers have been proposed.[12-15] However, validation of these tests has been complicated by the lack of a gold standard for comparison. Most studies have, out of necessity, relied on fibrosis stages derived from a single percutaneous biopsy and have attributed any discordance to the failure of the fibrosis markers when, in fact, it is widely accepted that there is considerable inaccuracy in biopsy results.

The goal of our study was to use the more accurate technique of laparoscopic examination and biopsy[8,10] to prospectively evaluate the clinical utility of FIBROSpect[SM] II, a panel of 3 serum markers directly related to extracellular matrix turnover. To minimize sampling error and provide a more complete liver assessment, multiple biopsies were taken from both the right and left lobes, when clinically feasible. All livers were also visually evaluated for cirrhosis by the laparoscopist during the procedure.

Serum marker tests potentially reflect not only a more global assessment of liver disease than a single biopsy, but, more importantly, they are not subject to interobserver variation or sampling errors. Measurement of fibrosis matrix components theoretically might detect smaller or more rapid changes in the extent of fibrosis than those reflected by the histologic fibrosis stage. Therefore, these tests may provide an important adjunct to the use of biopsy in the management of liver disease.

Methods

Patients

Consecutive patients over the age of 18 years who had chronic HBV or HCV infection and who were undergoing laparoscopy-guided liver biopsy were recruited. All

patients accepted into the study had active infections. For HBV, this was demonstrated by a positive hepatitis B surface antigen or HBV DNA level of greater than 100 IU/mL. For HCV, this meant a positive HCV antibody or HCV RNA level of greater than 60 IU/mL. Each patient was infected with only HCV or HBV, and no patient was coinfected with HIV. No patient had received treatment for hepatitis for 6 months prior to the biopsy. Patients who were taking medications that affect liver function or could interfere with the FIBRO*Spect* II test were excluded. In addition, patients with a history of excessive alcohol intake, decompensated cirrhosis, nonalcoholic steatohepatitis, drug-induced hepatitis, primary biliary cirrhosis, autoimmune hepatitis, alpha-1-antitrypsin deficiency, hemochromatosis, or Wilson disease were excluded. The study was performed in accordance with US Food and Drug Administration regulations for Good Clinical Practice and approved by the Institutional Review Boards (IRB) of the University of Miami School of Medicine and the Western IRB. All patients signed informed consent for the study and procedures.

Study Design

Demographic information, medical history, and viral etiology were collected for the cohort of 145 patients. Concomitant medications were recorded, as were standard biochemical lab results that were obtained within 6 months of biopsy. Serum samples from each patient were collected either before the administration of preoperative drugs on the day of biopsy or within 5 days after biopsy. Samples were given a coded study identification number and were shipped frozen for FIBRO*Spect* II analysis. Because liver fibrosis can occur in patches, biopsies were obtained from multiple sites during the laparoscopic examination to evaluate the status of the entire liver, not just the most severely affected areas. A single institutional pathologist, who was selected for skill and experience in liver fibrosis assessment, reviewed all biopsies. The pathologist had no knowledge of patient identity or clinical information. For comparison to the FIBRO*Spect* II results, the average biopsy stage across all samples was used for each patient. Additionally, the biopsy scores and FIBRO*Spect* II results were compared to the liver cirrhosis assessment made by the laparoscopist during the procedure.

Testing

All testing was conducted at Prometheus Laboratories in San Diego, California. The test panel was composed of 3 serum markers that directly reflect extracellular matrix turnover: hyaluronic acid (HA), tissue inhibitor of metalloproteinase-1 (TIMP-1), and alpha2-macroglobulin (A2M). Serum HA was measured in an enzyme-linked immunosorbent assay (ELISA) using HA-binding protein (Corgenix), TIMP-1 was measured by a sandwich ELISA (Amersham Biosciences), and A2M was measured by nephelometry (Beckman Coulter). The analytical performance of these 3 assays has been validated in the clinical laboratory, with intra-assay and interassay variability of 2–11% and 1–11%, respectively. Test results were generated by technologists blinded to clinical, histologic, and laboratory findings. For each patient, the levels of the 3 markers were combined using a logistic regression algorithm to generate a FIBRO*Spect* II index score ranging from 1 to 100. Three results were reported: the individual index score, a dichotomized interpretation of no/mild fibrosis versus significant fibrosis, and the associated probabilities for the Metavir fibrosis stages. The dichotomized result was generated using a cutoff index score of 42, which was previously determined in a large cohort of chronic HCV patients. A score of less than 42 was interpreted as consistent with no/mild disease, defined as fibrosis stages 0 or 1. An index score of 42 or greater was interpreted as consistent with significant disease, defined as fibrosis stages 2, 3, or 4. We used the dichotomized result to calculate the test performance measures of accuracy, sensitivity, and specificity.

Liver Biopsy and Scoring

Macroscopic evaluation of cirrhosis was made by the laparoscopist based on observation of the following criteria: diffuse nodules on the liver surface, shallow nodules slightly protruding from the liver, and liver hardness by palpitation or rigidity when lifted with a blunt probe. A verbal description was recorded at the time of laparoscopy, and a picture of each liver was obtained as documentation. Whenever possible, biopsy samples were taken from both the right and left lobes, with sites chosen to obtain the best overall evaluation of liver condition. Biopsies were obtained using an automatic 16-gauge tru-cut needle (biopsy gun). All biopsies included in the analysis contained at least 6 portal tracts. Formalin-fixed, paraffin-embedded sections were stained with hematoxylin and eosin and with Masson's Trichrome. Slides were labeled with patient identification numbers and then reviewed blindly by the pathologist. Histologic findings were assessed according to the standard grading and staging method based on the Scheuer system,[16] as modified by Batts and Ludwig.[17,18] This is similar to the Metavir system commonly used in Europe, with stages 0–1 comparable to Metavir stages F0–F1, and stages 2–4 comparable to Metavir stages F2–F4.

Statistical Analysis

Patient baseline characteristics were descriptively summarized and reported as mean and range. Continuous

variables between groups were compared by analysis of variance. A biopsy was considered adequate for analysis if it contained at least 6 portal tracts. Clinical performance for FIBRO*Spect* II was calculated versus the fibrosis stage determined from biopsies, as follows. When multiple biopsy samples for a patient were available for analysis, the scores were averaged and rounded to the nearest whole number. For the 4 patients in which only 1 biopsy with at least 6 portal tracts was obtained, the biopsy score from this single biopsy was used in the analysis. Only the 2 dichotomized interpretations of "consistent with stages 0–1" and "consistent with stages 2–4" were used in calculating the test performance measures. For comparison of grade and stage variability between multiple biopsy samples from one individual, the biopsy scores of all individual biopsies were used. For comparison between the right and left lobes of the liver, the biopsy scores for each lobe were averaged and the comparison was made between these averages.

Results

Demographic and Clinical Data

The majority (n=138, 95%) of the 145 patients enrolled in the study had chronic HCV infection, but a small percentage (n=7, 5%) was infected with HBV. There did not appear to be a difference between these patient groups in terms of demographics, number of biopsy samples obtained, or histopathologic scoring (data not shown). Only gamma glutamyl transferase, aspartate aminotransferase, and alanine aminotransferase levels were slightly lower in the HBV patients, although these levels were within the range seen for the HCV patients. Because these laboratory values were not used directly in any calculation, both groups of patients were analyzed together (Table 1).

Patient characteristics were similar to that of other clinical viral hepatitis populations,[9,15] except that a very high percentage of our patients had significant disease. Only 23 patients (16%) had mild disease, defined as stage 0 or 1, whereas 53 patients (36%) were stage 2 and 69 patients (48%) were stages 3–4. This resulted in an overall prevalence of 84% for fibrosis stages 2–4. HCV genotype was available for all 138 HCV patients; of these, 120 patients (87%) were genotype 1, and 9 patients (6.5%) were genotype 2, with 2 of these 9 infected with 2A/2C and 1 with 2A/2C/4E. There were also 5 patients (3.6%) with genotype 3, and 4 patients (2.9%) with genotype 4. Of the HCV genotype 1 patients, 50.8% had previously received treatment, whereas 25% of genotype 4 patients had been previously treated. A higher percentage of genotype 2 (77.8%) and genotype 3 (80%) patients had previously received treatment. Previous treatment was reported

Table 1. Patient Characteristics

Characteristic	All subjects
Age, years (range)	52 (30–73)
Gender, male (%)	81 (55.9%)
BMI, kg/m²	27.0 (16.8–51.3)
Prothrombin INR	1.0 (0.5–1.8)
Platelet count (10^9 cells/L)	197 (43–404)
Albumin, g/dL	4.0 (2.2–5.2)
Total bilirubin, mg/dL	0.67 (1.2–4.2)
GGT (U/L)	84 (10–480)
AST (U/L)	62 (13–510)
ALT (U/L)	79 (16–659)
ALP (U/L)	82 (29–260)
Batts-Ludwig fibrosis (average)	
0	4 (3%)
1	19 (13%)
2	53 (37%)
3	39 (27%)
4	30 (21%)

The mean and range are given for all results except fibrosis stage, for which the average and percent of the total subjects are indicated. N=145 for all characteristics except BMI, where complete data were available for only 141 subjects.

BMI=body mass index; GGT=gamma glutamyl transferase; AST= aspartate aminotransferase; ALT=alanine amniotransfrase; ALP=alkaline phosphatase; INR=international normalized ratio.

for 3 of the 7 patients (42.8%) infected with HBV. All the patients had active viral infections at the time of the study and had been off treatment for at least 6 months.

A total of 432 biopsies were taken from the 145 patients; however, 12 biopsies had fewer than 6 portal tracts and were not included in our analysis. There were 118 biopsies (28%) with a length of less than 1.5 cm and 302 (72%) with a length of at least 1.5 cm. There were 42 biopsies (10%) with a length of at least 2.0 cm and only 2 biopsies (0.4%) with a length of at least 2.5 cm. Of the patients with multiple biopsies taken, 92% had at least 1 biopsy of 1.5 cm or longer in length, and 68% had 2 or more biopsies of 1.5 cm or longer, with additional smaller biopsies taken from other areas of the liver. Two or more analyzable biopsy samples were provided by 98% of the patients, with 84% providing a biopsy from both the right and left lobes (Table 2).

Table 2. Distribution of Biopsy Samples

Biopsy samples	No. of subjects
1	4 (3%)
2	43 (30%)
3	62 (43%)
4	36 (25%)
Right lobe only	2 (1%)
Left lobe only	21 (14%)
Both lobes	122 (84%)

The number (percent) of biopsies obtained from different lobes of the liver. A total of 145 patients were biopsied; 141 had more than 1 analyzable biopsy.

Figure 1. ROC curves for FIBRO*Spect* II in the discrimination of no/mild from significant fibrosis: F0–F1 vs F2–F4 (black triangles, bold black line), AUROC 0.77; F0–F2 vs F3–F4 (open triangles, gray line), AUROC 0.83. The diagonal where no discrimination occurs is also shown for reference.

ROC=receiver operating characteristic; AUROC=area under the ROC curve.

FIBROSpect II Performance

Serum markers reflect the fibrotic process throughout the entire liver; therefore, to compare the FIBRO*Spect* II results to biopsy stage, we averaged the stages from all the biopsies of a patient's liver to form the patient's global score. Because biopsy stages in histologic scoring systems are represented only by integers, the averaged biopsy stage had to be rounded to the nearest whole number. Thus, a biopsy average of 1.0–1.4 was rounded down to stage 1, and a score of 1.5–1.9 was rounded up to stage 2. To minimize sampling errors, the laparoscopic findings for all patients were reviewed without reference to the histopathologic stage, and a determination of cirrhosis was made. If the visual evaluation at laparoscopy indicated cirrhosis, the stages of all the biopsies taken from the patient were set to stage 4, even if histopathologic evidence had indicated a different stage. The overall ability of FIBRO*Spect* II to discriminate between no/mild fibrosis and more significant fibrosis is demonstrated by the receiver operating characteristic (ROC) curves in Figure 1. A test that cannot differentiate between disease and no disease would generate a 45-degree line. The area under the ROC curve (AUROC) is a measure of diagnostic accuracy, so a test that does not discriminate between disease and no disease would have an AUROC of 0.50. For discriminating between stages 0–1 versus stages 2–4, the AUROC of FIBRO*Spect* II was 0.77 (95% confidence interval [CI], 0.672–0.867) and for discriminating between stages 0–2 versus stages 3–4, the AUROC was 0.83 (95% CI, 0.764–0.898).

The dichotomized interpretation of no/mild fibrosis (stages 0–1) or significant fibrosis (stages 2–4) was used to calculate the overall test performance measures shown in Table 3. The sensitivity, specificity, and accuracy of FIBRO*Spect* II were all near 70%, and the likelihood ratio for a positive test result in our population was 2.6.
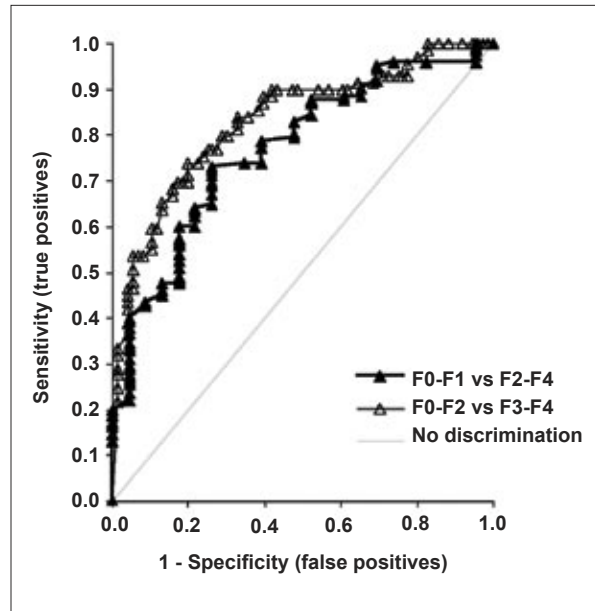
The ability of the FIBRO*Spect* II test to predict significant disease is excellent (93%) in this high prevalence population (84% stages 2–4). The negative predictive value is low (30%); however, the low proportion of patients with stages 0–1 biopsies (n=23, 16%) limits the reliability of the negative predictive value estimate from this study population.

The FIBRO*Spect* II test had the best concordance with biopsy stages 3 and 4 and was least concordant when the biopsy was stage 2 (Table 4). Because of the small number of stages 0 and 1 patients, these patients have been grouped together. When comparing serum test and biopsy stage, 6 patients were interpreted as having significant disease by serum test but had a histologic stage of 0–1 (false-positive). In addition, 40 patients were interpreted as having no/mild disease by serum but had a histologic stage of 2–4 (false-negative). However, 30 (75%) of these false-negative results were classified as stage 2 by histopathologic evidence. Only 1 of the 30 stage 2 patients was deemed cirrhotic by laparoscopic inspection.

In addition to the dichotomized interpretation, each patient has an individual index score that can range from 1–100. The relationship between the index score and
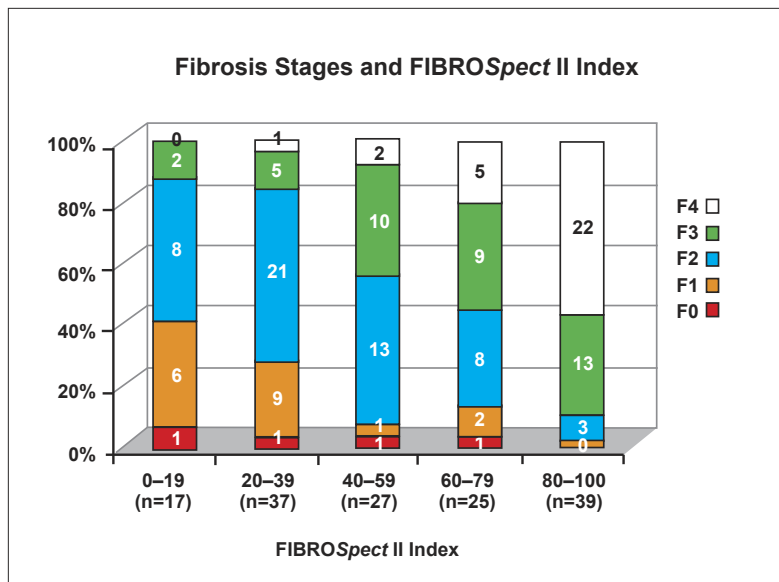
Figure 2. Distribution of FIBRO*Spect* II index scores across fibrosis stages. Each bar has been normalized to 100%, and the number of subjects (n) in each category and the number of patients in each stage are indicated.

Table 3. FIBRO*Spect* II Overall Performance

Parameter	Value	95% confidence interval
Prevalence (F2–F4)	84.1%	
Sensitivity	67.2%	58.1–75.4%
Specificity	73.9%	51.6–89.9%
Accuracy	68.3%	60.0–75.7%
PPV	93.2%	85.7–97.5%
NPV	29.8%	18.5–43.4%

Dichotomized test scores from all 145 subjects were used. Scores below 42 were interpreted as consistent with stages 0–1 fibrosis; those equal to or greater than 42 were interpreted as stages 2–4 fibrosis.

PPV=positive predictive value; NPV=negative predictive value.

average fibrosis stage is shown in Figure 2. Each bar is normalized to 100%, even though there were many more severely affected patients; this results in fewer patients in the bars representing lower index scores. There was good correlation between index scores and average biopsy stage (Spearman r=0.657; *P*>.001).

Sampling Error

Each patient's liver was also evaluated visually for cirrhosis at the time of laparoscopy. For 17 subjects (12%), the histologic evidence and visual results were inconsistent. Thirteen of the patients (9%) had features that indicates

cirrhosis such as a nodular surface; however, the histologic stage was not scored as stage 4. Ten of these 13 patients had an average stage of 3, and 3 patients had an average stage of 2. An example of such a patient is shown in Figure 3; this patient is clearly cirrhotic, although the single biopsy obtained was scored as stage 2. The FIBRO*Spect* II test was consistent with significant fibrosis, with an index value of 99 (range 0–100). In 4 patients (3%), there was histologic evidence of cirrhosis (stage 4) but no macroscopic evidence at laparoscopy. The FIBRO*Spect* II test indicated significant fibrosis for these 4 nonnodular cirrhotic patients and for 11 of the 13 cirrhotic patients identified by laparoscopy.

To further explore the discordance between FIBRO*Spect* II results and biopsy stage, we analyzed the variation in the fibrosis stage of multiple biopsies from the same patient. In 48% of the patients with multiple biopsies, there was no difference in the stage of fibrosis. In 47% of patients, there was a difference of 1 stage; in 5% of patients, there was a difference of 2 stages. A similar variation was seen when the degree of inflammation was compared across biopsies from the same patient. In 44% of patients with multiple biopsies, there was no difference in grade between biopsies, in 52% of patients, there was a difference of 1 grade, and in 4% of patients, there was a difference of 2 grades.

In the 122 patients with a biopsy from both the right and left lobes, there was no consistent unidirectional difference in the stage of biopsies from the right versus the left lobe. In 23 patients, the right lobe biopsy had a lower stage than that of the left lobe, whereas in 20 patients, it was the reverse, with a higher stage in the right lobe than

Table 4. Concordance with Batts-Ludwig Staging

Average Stage	0	1	2	3	4	Total
FS II positive	2	4	23	30	29	88
FS II negative	2	15	30	9	1	57
Total	4	19	53	39	30	145
Number concordant Percent	17 74%		23 43%	30 77%	29 97%	105 72%

FS II=FIBRO*Spect* II.

in the left. For 79 patients (65%), both the right and left lobe biopsies had the same stage.

The grade of inflammation also showed variation between the right and left lobes. In 38 patients, the right lobe had a lower grade of inflammation, whereas in 16 patients, the left lobe had a lower grade. For 68 of the patients (56%), both the right and left lobe biopsies had the same grade. Similar differences in the histologic evidence of biopsies from the right and left lobes of the liver were found in a prior study[9] from our group.

Discussion

This study shows that the combination of 3 serum markers as measured by FIBRO*Spect* II can distinguish no/mild fibrosis (stages 0–1) from more significant fibrosis (stages 2–4) in the majority of patients.

FIBRO*Spect* II test results were reported both as an index score and as a dichotomized interpretation indicating either no/mild or significant disease. We found a good correlation between the FIBRO*Spect* II index scores and average biopsy stage (Spearman r=0.657; *P*>.001). The index score allowed the reliability of the test result to be assessed for each patient. The cutoff index value used for the dichotomized interpretation was determined in a large cohort of chronic HCV patients.[15,19] Index scores of less than 42 were dichotomized as consistent with Metavir F0–F1, and scores of 42 and greater as consistent with Metavir F2–F4. The probability of an accurate result naturally increases as the index score diverges from 42. For example, in this study, an index score above 60 was associated with only a 6% chance of being stages 0–1, a 17% chance of being stage 2, and a 77% chance of being stages 3–4, for a total of a 94% chance of being stages 2–4. An index value of 80 was associated with only a 2.5% chance of being a false-negative (stages 0–1), an 8% chance of being stage 2, and a 90% chance of being stages 3–4, for a total of a 98% chance of being stages 2–4.
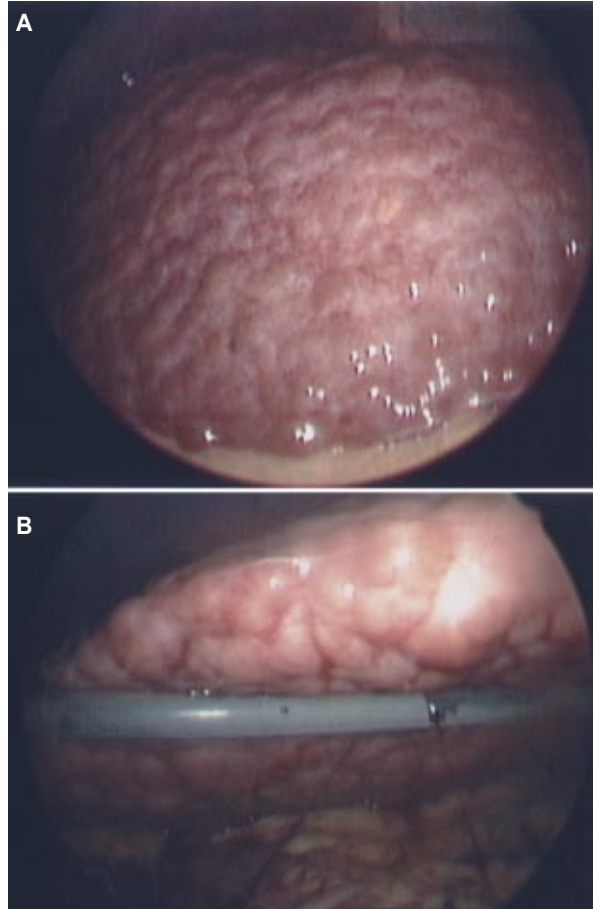
Figure 3. View of a cirrhotic liver at laparoscopy, with the top of the right lobe (A) and the underside of the left lobe, which reveals patchy granularity and severe nodularity (B). This patient was staged as F2 by histopathology. The FIBRO*Spect* II score was positive, with an index of 99 out of 100 possible.

We used the dichotomized test interpretations of no/mild (stages 0–1) and significant fibrosis (stages 2–4) for the overall performance analysis. Using average biopsy stage for comparison with test results, the likelihood ratio for a positive test result in our population was 2.6, and the ROC curves demonstrated good discrimination with AUROC of 0.77 (stages 2–4) and 0.83 (stages 3–4). The sensitivity, specificity, and accuracy were all close to 70%. The positive predictive value was 93%, whereas the negative predictive value was 30%. However these predictive values are strongly influenced by the prevalence of significant disease in the population; the high disease prevalence in patients scheduled for laparoscopy did not allow for a good estimation of these parameters.

Although patients with decompensated cirrhosis were excluded based on clinical assessment, 24 asymptomatic

patients who had one or more biochemical indications of cirrhosis (platelet count <100 million/mL, bilirubin >2 mg/dL, albumin <3.5 g/dL) were included. These patients serve as a validation of the test's ability to identify significant disease. FIBRO*Spect* II did correctly identify 21 of these patients as having significant disease, and 1 patient with biopsy stage 1 and a decreased albumin (3.3 g/dL) was correctly identified as having no/mild disease. However, 2 patients were incorrectly classified.

Overall, FIBRO*Spect* II correctly identified 67% of the 122 patients with significant fibrosis (stages 2–4). The greatest discordance between serum marker test interpretation and average biopsy stage was observed at stage 2, at the point of dichotomization. This is also the point at which disease is just beginning. Of the 40 false-negative FIBRO*Spect* II results, 30 patients (75%) were stage 2 by biopsy. Therefore, the test is best able to identify either patients with early disease or those with advanced serious disease. Other serum marker tests or panels are also least accurate in the middle of the range, and several of them include an indeterminate range that may encompass up to 50% of the patients tested.[20] As an example, results allowing the calculation of the aspartate transaminase-to-platelet ratio index (APRI)[10] were available for 143 subjects of this study. APRI values for one third (33%) were between 0.5 and 1.5, the range where APRI cannot determine the degree of fibrosis. For the 52 stage 2 patients, 16 (30%) had APRI values in this indeterminate range and only 3 (8%) would have been correctly predicted as F2 or greater by APRI. For the 39 stage 3 patients, 21 (54%) were in the indeterminate range, with only 5 (13%) correctly predicted. FIBRO*Spect* II was able to predict significant fibrosis in 43% of the stage 2 patients and in 77% of the stage 3 patients, with no indeterminate range.

In assessing the clinical value of serum marker tests like FIBRO*Spect* II, it is important to consider the role they would play in disease management and the setting in which they would be used. As more treatment options become available, the role of serum marker tests is changing. FIBRO*Spect* II was originally proposed as a way of identifying patients with mild disease in the general population where antiviral treatment and biopsy might be deferred. An example is the study by Christensen and associates[21] of 142 Alaska natives/American Indians chronically infected with HCV. They found that the use of FIBRO*Spect* II would have avoided biopsies in 44% of their stages 0–1 cohort, with only 4 Ishak stage 3 patients (3%) incorrectly categorized as having mild disease. All of the more severely affected patients were correctly identified. In their study population, there was a prevalence of significant disease of 38%, sensitivity of 93%, specificity of 66%, accuracy of 76%, and NPV of 94%. A similar study by Zaman and colleagues[22] studied Oregon patients

with chronic HCV, comparing FIBRO*Spect* II results with Metavir biopsy scores. The overall prevalence of significant fibrosis was 36.1%, with sensitivity of 71.8%, specificity of 73.9%, and accuracy of 73.1%. The AUROC was excellent at 0.826.

In contrast, the prevalence of significant disease in our study population of patients scheduled for laparoscopy was very high (84%). Many of the patients had already been treated unsuccessfully; 53% of our HCV patients had already undergone some type of antiviral treatment, and 43% of our HBV patients had already undergone previous treatment. Instead of using FIBRO*Spect* II to select those with mild disease, a better use of the test in this clinical context would be to identify patients with late-stage disease, who would be targeted for a more intensive regimen of treatment or for cirrhosis and hepatocellular cancer follow-up. By using an index value of 60 or greater as a cutoff, we would have correctly identified 60 of the 64 patients as having significant disease, and incorrectly identified only 4 patients (stages 0–1). A cutoff index value of 80 would have correctly identified all but 1 of the 40 patients in this group, for a probability of 97% of stages 2–4 disease.

It is also important to evaluate the performance of FIBRO*Spect* II in the context of the known errors inherent in the assessment of fibrosis by liver biopsy histopathology. Although liver biopsy provides important information about the location and extent of fibrosis and the degree of hepatic inflammation, it is not always a good comparison standard. Previous studies have used the fibrosis stage of a single liver biopsy specimen as the gold standard for evaluating noninvasive fibrosis tests and have attributed discordance to the failure of serum marker tests. However, discordance may also come from errors in the biopsy stage.

Fibrosis due to chronic viral hepatitis is part of the wound-healing response of the liver. Inflammation and deposition of extracellular matrix proteins are initially localized to the area around the portal tracts. Repeated injury due to chronic disease eventually leads to the substitution of hepatocytes with extracellular matrix, then to bridging fibrosis, and then to frank cirrhosis.[23] The standard procedure of percutaneous liver biopsy ends up sampling only 0.002% of the entire liver mass, usually from the right lobe; therefore, a single biopsy may miss localized or patchy fibrosis. Bedossa and colleagues[6] studied the correlation of Metavir stage and fibrosis percent, as measured by image analysis. The amount of fibrosis they observed appeared to increase in a nonlinear fashion, with little change in early stages and a more rapid increase beginning at stage F2. A similar relationship was also found in an earlier study by Pilette and associates.[24] These investigators compared three measurements of fibrosis: area-by-image analysis, fibrosis stage, and serum markers.

Although all were well correlated, serum markers were better correlated with the area of fibrosis than with the semiquantitative fibrosis stages.

Differences of one stage have previously been reported between biopsy samples taken from within the same liver. When two separate biopsies were taken from the right lobe of the liver through a single skin puncture in 29 chronic HCV patients,[8] 44.8% differed by at least 1 stage. Regev and colleagues[9] sampled both the right and left lobes of chronic HCV patients at laparoscopy and found that 33.1% of patients had a difference of at least 1 stage between lobes: 14.5% of these patients had cirrhosis in one lobe, whereas stage 3 fibrosis was detected in the other. In this study, 52% of patients had a stage difference of at least 1 in biopsies taken from different sites within their livers, and 8 (6.5%) had cirrhosis in 1 lobe and stage 3 in the other. All our biopsies were read by a single, experienced pathologist; therefore, these differences likely reflect heterogeneity in the distribution of fibrosis. In 5 patients, 2 biopsies were obtained in which one scored as stage 1 and the other as stage 2; this is the point where the averaging and rounding of the averaged biopsy stage would affect the interpretation of the results. Which stage is the correct one to use in comparison to the serum test results for these cases? Because the fibrosis stage is a whole number, we had to round the average value of 1.5 to stage 2 for the calculation of test performance. In fact, 3 of these patients were reported as consistent with F0–F1 by FIBRO*Spect* II and therefore as false-negatives, whereas the other 2 patients were reported as consistent with F2–F4. Although validation studies confirm that the intersample variability of the FIBRO*Spect* II assays is no more than 11%, information about the variability of repeated samples from the same subject obtained at different times is not available. This issue merits further study.

Small and therefore unrepresentative samples make histologic diagnosis more inaccurate. A biopsy size of 1.5 cm with 5 portal tracts is generally considered adequate for analysis;[9] however, even then, appreciable sampling error can occur. The study by Bedossa and associates[6] investigated how biopsy length influenced the accuracy of fibrosis stage. A large area of each biopsy was reviewed to establish a reference Metavir stage against which the stages obtained from shorter "virtual" sections were compared. Only 65% of 1.5-cm biopsies and 75% of 2.5-cm biopsies were scored as having the same stage as the reference. Therefore, this group recommended that biopsies used for fibrosis staging be at least 2.5 cm in length. A sample length of 2.0 cm with 11 portal tracts was recommended in a recent study by Colloredo and associates.[25] However, Rousselet and colleagues[11] found that it was the experience of the pathologist, rather than

specimen length, that was the critical factor in obtaining a reproducible fibrosis stage.

In actual practice, even at specialized centers, many patients do not have biopsies meeting these standards. Clinical considerations limit biopsy size; unfortunately, there is no straightforward way in which the treating physician can use length, number of portal tracts, and biopsy fragmentation to determine the expected accuracy of the reported stage and grade. At most hospitals, liver biopsies are not always read by the same pathologist, and the level of experience varies. Poynard and colleagues[14] recently used predetermined risk factors to assign discordance to either liver biopsy failure or to failure of a serum marker test in a large cohort of patients. Discordance was observed in 29% of the patients in their series but was attributable to the failure of serum markers in only 2.4% of patients. In Poynard's study, for the group of biopsies scored at an experienced academic center, mean biopsy length was 1.6 cm, with 31% being at least 1.5 cm in length and containing at least 5 portal tracts. Only 14% of the biopsies were 2.5 cm or longer. In our study, 432 biopsies were taken, with 420 containing 6 or more portal tracts. Of the patients, 92% (134) had at least 1 biopsy that was at least 1.5 cm in length, with additional smaller biopsies taken from other areas of the liver. Because FIBRO*Spect* II provided a more global indication of liver fibrosis, an average of all the biopsies from each subject was used for our comparison. The results were not changed when biopsies smaller than 1.5 cm were removed from the analysis. Only 12 of the 145 subjects had more than 1 biopsy of at least 2.0 cm, and these longer biopsies comprised only 10% of the total. Thus, we could not make a further analysis. However the biopsy series of Christensen and associates[21] did show increased concordance of biopsy stage and FIBRO*Spect* II score for biopsies of greater than 2 cm.

To reduce biopsy sampling error, we corrected biopsy scores using the assessment made visually by the laparoscopist. All biopsy samples were read by the same experienced pathologist. In earlier comparisons of biopsy histopathologic evidence to laparoscopic examination,[10,26] 32–61% of cases with visual evidence of cirrhosis were missed by histopathologic evidence. In our study, 17 patients (12%) had histopathologic scores that were not consistent with laparoscopic assessment. However, this did not affect the overall performance calculated for FIBRO*Spect* II because it combined stages 0–1 (no/mild disease) and 2–4 (significant disease). In all of these 17 patients, the fibrosis stage was read as stage 2 or 3.

We have demonstrated that FIBRO*Spect* II can successfully distinguish severe fibrosis (stages 3–4) from no/mild fibrosis (stages 0–1) in a predominantly HCV-infected population. This further supports the ability of

the serum marker test to correctly predict severe disease, even in the absence of other, more obvious biochemical evidence. FIBRO*Spect* II score was a less reliable predictor for patients with stage 2 disease. However, the index score could be used to assess the reliability of the test result for any given patient. Because this test is not subject to interobserver variability or to sampling error, it may offer distinct advantages in assessing fibrosis, particularly in situations where repeated testing is indicated. Unlike percutaneous liver biopsy, there is minimal risk to the patient. Thus, the FIBRO*Spect* II test panel may be a useful adjunct to biopsy in the management of HCV and HBV liver disease. Further study of FIBRO*Spect* II is warranted to assess its ability to accurately measure fibrotic changes over time and to detect both the accumulation of fibrosis and its decrease in patients undergoing antiviral or antifibrotic therapy.

References

1. Parkes J, Guha IN, Roderick PJ, et al. Performance of serum marker panels for liver fibrosis in chronic hepatitis C. *J Hepatol*. 2006;44:462-474.

2. Pawlotsky J-M. Therapy of hepatitis C: from empiricism to eradication. *Hepatology*. 2006;43:S207-S220.

3. Feld JJ, Liang TJ. Identifying patients with progressive liver injury. *Hepatology*. 2006;43:S194-S206.

4. Strader DB, Wright T, Thomas DL, et al. AASLD practice guideline: diagnosis, management, and treatment of hepatitis C. *Hepatology*. 2004;39:1147-1171.

5. Poordad FF. FIBROSpect II: a potential noninvasive test to assess hepatic fibrosis. *Expert Rev Mol Diagn*. 2004;4:593-597.

6. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology*. 2003;38:1449-1457.

7. Pounard T, Ratziu V, Charlotte F, et al. Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis C. *J Hepatol*. 2001;34:730-739.

8. Siddique I, El-Naga HA, Madda JP, et al. Sampling variability on percutaneous liver biopsy in patients with chronic hepatitis C virus infection. *Scan J Gastroenterol*. 2003;38:427-432.

9. 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.

10. Wietzke-Braun P, Braun F, Schott P, et al. Is laparoscopy an advantage in the diagnosis of cirrhosis in chronic hepatitis C virus infection? *World J Gastroenterol*. 2003;9:745-750.

11. Rousselet MC, Michalak S, Dupre F, et al. Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology*. 2005;41:257-264.

12. Forns X, Ampurdanes S, Llovet JM, et al. Identification of chronic hepatitis C patients by a simple predictive model. *Hepatology*. 2002;36:986-992.

13. 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.

14. Poynard T, Munteanu M, Imbert-Bismut F, et al. Prospective analysis of discordant results between biochemical markers and biopsy in patients with chronic hepatitis C. *Clin Chem*. 2004;50:1344-1355.

15. Patel K, Gordon SC, Jacobson I, et al. Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. *J Hepatol*. 2004;41:935-942.

16. Scheuer PJ. Classification of chronic viral hepatitis: a need for reassessment. *J Hepatol*. 1991;13:372-374.

17. Batts KP, Ludwig J. Chronic hepatitis: an update on terminology and Reporting. *Am J Surg Path*. 1995;19:1409-1417.

18. Ludwig J. The nomenclature of chronic active hepatitis: An obituary. Gastroenterology. 1993;105:274-278.

20. 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;41:1376-1382.

19. Oh E, Smith KM, Patel K, et al. An improved biomarker panel (FIBROSpect II) for differentiating mild from more server liver fibrosis in chronic HCV patients [abstract S1639]. *Gastroenterology*. 2004;4:(suppl 2):A707.

21. Christensen C, Bruden D, Livingston S, et al. Diagnostic accuracy of a fibrosis serum panel FIBROSpect II compared to Knodell and Ishak liver biopsy scores in hepatitis C patients. *J Viral Hepat*. 2006,13:652-658.

22. Zaman A, Rosen HR, Ingram K, et al. Assessment of FIBROSpect II to detect hepatic fibrosis in chronic hepatitis C patients. *Am J Med*. 2007;120:280.e9-14.

23. Bataller R, Brenner DA. Liver fibroisis. *J Clin Invest*. 2005;115:209-218.

24. Pilette C, Rousselet MC, Bedossa P, et al. Histopathological evaluation of liver fibrosis: quantitative image analysis vs semi-quantitative scores. *J Hepatol*. 1998;28:439-446.

25. Colloredo G, Guido M, Sonzogni A, et al. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample the milder the disease. *J Hepatol*. 2003;39:239-244.

26. Poniachik J, Bernstein DE, Reddy KR, et al. The role of laparoscopy in the diagnosis of cirrhosis. *Gastrointest Endosc*. 1996;43:586-571.