Validation of the FibroTest Biochemical Markers Score in Assessing Liver Fibrosis in Hepatitis C Patients

Enrico Rossi,1* Leon Adams,2 Alexander Prins,1 Max Bulsara,3 Bastiaan de Boer,4 George Garas,2 Gerry MacQuillan,5 David Speers,6 and Gary Jeffrey2,5

Background: Determining the stage of fibrosis by liver biopsy is important in managing patients with hepatitis C virus infection. We investigated the predictive value of the proprietary FibroTest score to accurately identify significant fibrosis in Australian hepatitis C patients.

Methods: Serum obtained from 125 confirmed hepatitis C patients before antiviral therapy was analyzed for haptoglobin, α₂-macroglobulin, apolipoprotein A1, bilirubin, and γ-glutamyltransferase activity, and the FibroTest score was computed. Liver fibrosis pathology was staged according to a defined system on a scale of F0 to F4. We used predictive values and a ROC curve to assess the accuracy of FibroTest scores.

Results: The prevalence of significant fibrosis defined by liver biopsy was 0.38. The most useful single test for predicting significant fibrosis was serum α₂-macroglobulin (cutoff value, 2.52 g/L; sensitivity, 75%; specificity, 67%). The negative predictive value of a FibroTest score <0.1 was 85%, and the positive predictive value of a score >0.6 was 78%. Although 33 of the 125 patients had FibroTest scores <0.1 and were therefore deemed unlikely to have fibrosis, 6 (18%) had significant fibrosis. Conversely, of the 24 patients with scores >0.6 who were likely to have significant fibrosis, 5 (21%) had mild fibrosis. Of the 125 patients in the cohort, 57 (46%) could have avoided liver biopsy, but discrepant results were recorded in 11 of those 57 (19%).

Conclusion: The FibroTest score could not accurately predict the presence or absence of significant liver fibrosis.

© 2003 American Association for Clinical Chemistry

Liver biopsy is recommended for the management of patients infected by hepatitis C virus (HCV). In Australia, demonstration of a minimum stage of fibrosis on liver biopsy is required to qualify for standard combination antiviral therapy. The difficulties associated with liver biopsy include its highly invasive nature and the development of substantial complications in 0.3% of patients (1). Although liver biopsy is acknowledged as the gold standard for evaluating fibrosis, it is occasionally prone to sampling error as a result of the heterogeneous distribution of pathologic changes.

Previous studies have described markers with high predictive values for the detection of cirrhosis, whereas detection of early-stage fibrosis has been neglected (2). A proprietary fibrosis score (FibroTest) based on biochemical serum markers has been described that could substantially reduce the number of biopsies performed for the management of HCV infection (3). The FibroTest score is computed by accessing a proprietary website and entering the patient’s age, sex, and results for serum haptoglobin, α₂-macroglobulin, apolipoprotein A1, γ-glutamyltransferase, and bilirubin analyses. Investigators from France reported that with appropriate cutoff values, the FibroTest score in HCV patients gave either a 100% negative predictive value for the absence of clinically significant fibrosis or a 91% positive predictive value for its presence (3).

We independently assessed the predictive value of the FibroTest score to distinguish patients whose liver biopsy revealed insignificant fibrosis from those with clinically
significant fibrosis who would qualify for antiviral therapy.

**Materials and Methods**

**STUDY PARTICIPANTS**

Patients with HCV RNA detectable by PCR who had undergone liver biopsy between January 1998 and November 2001 were eligible. We consecutively recruited a total of 125 patients: 82 males and 43 females (mean age, 40 years). The study was approved by the ethics committee of the Sir Charles Gairdner Hospital.

**HISTOLOGIC STAGING**

Biopsies were staged blindly according to the METAVIR group scoring system (4) by one pathologist (B.D.B.) with a special interest in liver pathology. To maintain consistency, METAVIR scores were reported in batches of 20–25 patients.

Every biopsy specimen was staged on a scale of F0 to F4: F0, no fibrosis; F1, portal fibrosis without septa; F2, few septa; F3, many septa without cirrhosis; and F4, cirrhosis. Patients with scores of F0 or F1 were considered to have insignificant fibrosis, and those with scores of F2, F3, or F4 were considered to have clinically significant fibrosis that qualified for combination antiviral therapy.

**FibroTest score**

Serum taken at the time of biopsy and before any antiviral therapy was analyzed for haptoglobin, α2-macroglobulin, and apolipoprotein A1 by nephelometry (Beckman Immage) and γ-glutamyltransferase and total bilirubin on an automated analyzer (Hitachi 917). The FibroTest score was computed on the Experts-MD website (www.experts-md.com), accessed from July to November 2001, by entering the patient’s age, sex, and results for the five biochemical analytes listed. More recently, this service was located at www.biopredictive.com (accessed September 2002).

**STATISTICS**

The sensitivity, specificity, positive and negative predictive values, likelihood ratio, and ROC curves were calculated using Stata software (5). A maximum likelihood estimate of the parameters for a smooth fitting ROC curve was based on the binormal distribution.

**Results**

**PREVALENCE OF FIBROSIS**

The following distribution of METAVIR fibrosis stages was observed on liver biopsy: no fibrosis in 25 of 125 patients (F0 = 20%); portal fibrosis in 52 of 125 (F1 = 42%); few septa in 26 of 125 (F2 = 21%); numerous septa in 13 of 125 (F3 = 10%); and cirrhosis in 9 of 125 (F4 = 7%). The main endpoint was the identification of patients with significant fibrosis (F2, F3, or F4), who were eligible for government-subsidized antiviral therapy, and those with insignificant fibrosis (F0 or F1). The prevalence of significant fibrosis was 0.38 (48 of 125 patients).

**SERUM MARKERS**

Of the five serum markers studied, two showed significant positive correlation with increasing fibrosis stage: α2-macroglobulin and bilirubin (P for trend <0.0001 and 0.04, respectively). Significant negative trends with fibrosis stage were observed for haptoglobin and apolipoprotein A1 (P = 0.03 and 0.02, respectively). γ-Glutamyltransferase activities showed a positive trend with fibrosis stage that did not achieve significance (P = 0.06).

The most useful test for predicting F2 to F4 fibrosis was α2-macroglobulin, which gave a sensitivity of 75% and a specificity of 67% at a 2.52 g/L cutoff. There was a rapid decrease in sensitivity and specificity for the other markers, shown in decreasing order: bilirubin, sensitivity 61% and specificity 53% at a 10 μmol/L cutoff; γ-glutamyltransferase, 57% and 55% at a 45 U/L cutoff; apolipoprotein A1, 26% and 50% at a 1.41 g/L cutoff; and haptoglobin, 20% and 79% at a 0.56 g/L cutoff. Box plots of medians and interquartile ranges for serum α2-macroglobulin, haptoglobin, apolipoprotein A1 and bilirubin for each fibrosis grade are shown in Fig. 1.

**FibroTest score and fibrosis stage**

Using the suggested FibroTest score cutoff values, we observed the following distribution of patients: scores were <0.1 in 33 of 125 (26%); >0.6 in 24 of 125 (19%); and indeterminate (0.1–0.6) in the remaining 68 of 125 (54%) patients. There was moderate correlation between METAVIR fibrosis stage and FibroTest score (r = 0.59; P <0.00001). Fig. 2 shows box plots for FibroTest scores in patients with insignificant fibrosis (F0 and F1) and those who had significant (F2 to F4) fibrosis.

The sensitivity, specificity, and predictive values of FibroTest scores at various cutoff values are shown in Table 1. The sensitivity and specificity of a FibroTest score >0.1 for the presence of fibrosis stages F2 to F4 were 92% and 29%, respectively, and for FibroTest scores >0.6 the sensitivity and specificity were 42% and 94%, respectively. For our cohort, where the prevalence of F2 to F4 fibrosis was 0.38, the negative predictive value of a score <0.1 was 85%. The positive predictive value of a score >0.6 for the presence of F2 to F4 fibrosis was 78%, and a score of 0.6 gave a likelihood ratio of 6.4.

The diagnostic value of the FibroTest score was also assessed by ROC curve analysis, shown in Fig. 3, which gave an area (SE) under the curve of 0.739 (0.050).

**Discussion**

FibroTest is a composite algorithm-based score computed from age, sex, and serum α2-macroglobulin, bilirubin, γ-glutamyltransferase activity, apolipoprotein A1, and haptoglobin. We have examined the accuracy of FibroTest in identifying liver fibrosis in a cross-sectional analysis of HCV patients in whom the prevalence of biopsy-confirmed significant liver fibrosis was 0.38. In the original report, FibroTest could have achieved a 46% reduction in the number of liver biopsies required to manage...
chronic HCV infection, and we sought to confirm these encouraging initial findings (3). Subsequent reports from the same group have described longitudinal validation of FibroTest by demonstrating a significant decrease of the score in HCV patients who had sustained virologic response to treatment (6). In a further study, FibroTest scores were shown to be more accurate at predicting liver fibrosis than an index comprising the clinical historical features of HCV infection (7).

We found the individual serum markers used to calculate FibroTest scores ranked as follows from the most to the least informative in terms of sensitivity and specificity: α2-macroglobulin, bilirubin, γ-glutamyltransferase activity, apolipoprotein A1, and haptoglobin. The diagnostic value of an increased serum α2-macroglobulin concentration in staging fibrosis in alcoholic liver disease has been reported previously (8). This acute-phase protein is produced at sites of inflammation and liver fibrosis by a variety of cells, including hepatocytes and stellate cells. Bilirubin concentrations significantly increased with fibrosis severity, possibly reflecting early cholestasis. Serum γ-glutamyltransferase activity combined with prothrombin time and apolipoprotein A1 in the form of an index have previously been used to determine fibrosis severity, with high sensitivity and specificity recorded for

Fig. 1. Relationships between serum markers and fibrosis stage.
(A), α2-macroglobulin; (B), haptoglobin; (C), apolipoprotein A1; (D), bilirubin. The top and bottom of each box are the 25th and 75th centiles, giving the interquartile range. The line through the box is the median, and the error bars are the 5th and 95th centiles.

Fig. 2. FibroTest scores for patients with fibrosis stages F0 and F1 and F2 to F4.
The top and bottom of each box are 25th and 75th centiles, giving the interquartile range. The line through the box is the median, and the error bars are the 5th and 95th centiles.
the detection of cirrhosis in alcoholic liver disease (9).
Haptoglobin showed a negative correlation with fibrosis stage. This has been reported previously in HCV patients and may be a response to the increased hepatocyte growth factor seen in liver damage (10).

The FibroTest score was only moderately associated with fibrosis stage. As shown in Fig. 2, there was considerable overlap of FibroTest interquartile ranges for F2 to F4 fibrosis and those for F0 or F1 fibrosis. This is reflected in the sensitivities, specificities, and predictive values shown in Table 1. At the recommended cutoff values, the negative predictive value of a FibroTest score <0.1 for the presence of fibrosis stages F2 to F4 was 85%, the positive predictive value of a score >0.6 was 78%, and a score of 0.6 gave a likelihood ratio of 6.4. Superior data were reported in the original study (3), in which a FibroTest score <0.1 gave a 100% negative predictive value for the presence of fibrosis and a score >0.6 gave a 91% positive predictive value, with a score of 0.6 giving a likelihood ratio of 12.9. The area under the ROC curve in our study was 0.739, compared with 0.837 in the original study.

If the published FibroTest score cutoff values were used, liver biopsy could have been avoided in 33 of 125 (26%) HCV patients with scores <0.1 (deemed unlikely to have significant fibrosis) and 24 of 125 (19%) patients with scores >0.6 (likely to have significant fibrosis). The remaining 68 of 125 (54%) with indeterminate scores would have required biopsy for fibrosis staging (3). Of the 33 HCV patients with scores <0.1, a false-negative result, F2 to F4 fibrosis, was observed in 6 (18%) patients, compared with none in the original report. Although 33 fewer liver biopsies would have been performed, treatable fibrosis would have been missed in 6 patients.

Conversely, of the 24 patients with scores >0.6, a false-positive result, F0 to F1 fibrosis, was observed in 5 patients (21%), compared with a 91% certainty of F2 to F4 fibrosis in the original report. We would have performed 24 fewer liver biopsies, but treated 5 patients with insignificant fibrosis. Applying cutoff limits of <0.1 and >0.6 to our cohort would have avoided liver biopsy in 57 patients at the cost of discrepant results in 11 (19%) patients, false negatives for 6, and false positives for 5 patients. The 11 cases with discrepant results were submitted to repeat pathology staging, but the results were unchanged.

Compared with the original report, we obtained lower positive and negative predictive values for FibroTest, and a possible reason may be a difference in instrumentation. Although both studies used automated nephelometry to analyze serum haptoglobin, α2-macroglobulin, and apolipoprotein A1, different instrument platforms were used. We used the Beckman Immage, whereas the original work was performed on the Dade-Behring BNII, and the two systems may not give comparable results, leading to different FibroTest scores.

Several previous studies described the use of biochemical markers to assess various stages of fibrosis. For example, a panel of six biochemical markers used to detect cirrhosis (F4 fibrosis) in HCV patients correctly classified 85% of patients, but there was no attempt to apply the panel to stage fibrosis (11). Our work is more comparable to a study that used patient age and a panel of four markers (γ-glutamyltransferase activity, cholesterol, prothrombin time, and platelet count) to distinguish fibrosis stages F2 to F4 from F0 to F1 in HCV patients (12). When this predictive model was validated on 125 HCV patients, a negative predictive value of 96% and a relatively low positive predictive value of 66% were obtained, although the model could be applied to only 39% of patients. In a review of the application of serum markers for hepatic fibrosis, the authors forecast a probable future role while acknowledging that no individual marker was presently clinically applicable (13).

Our study design was cross-sectional, and we have not tested whether repeated determinations of FibroTest

Table 1. Sensitivity, specificity, and predictive values of FibroTest scores.

<table>
<thead>
<tr>
<th>Fibrosis score cut off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Likelihood ratio</th>
<th>Positive Predictive value</th>
<th>Negative Predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.92</td>
<td>0.29</td>
<td>1.3</td>
<td>0.45</td>
<td>0.85</td>
</tr>
<tr>
<td>0.2</td>
<td>0.83</td>
<td>0.52</td>
<td>1.7</td>
<td>0.52</td>
<td>0.83</td>
</tr>
<tr>
<td>0.3</td>
<td>0.75</td>
<td>0.61</td>
<td>1.9</td>
<td>0.54</td>
<td>0.80</td>
</tr>
<tr>
<td>0.4</td>
<td>0.67</td>
<td>0.78</td>
<td>3.0</td>
<td>0.65</td>
<td>0.79</td>
</tr>
<tr>
<td>0.5</td>
<td>0.56</td>
<td>0.85</td>
<td>3.9</td>
<td>0.70</td>
<td>0.76</td>
</tr>
<tr>
<td>0.6</td>
<td>0.42</td>
<td>0.94</td>
<td>6.4</td>
<td>0.78</td>
<td>0.72</td>
</tr>
<tr>
<td>0.7</td>
<td>0.35</td>
<td>0.94</td>
<td>5.4</td>
<td>0.77</td>
<td>0.70</td>
</tr>
<tr>
<td>0.8</td>
<td>0.22</td>
<td>0.96</td>
<td>5.8</td>
<td>0.79</td>
<td>0.66</td>
</tr>
<tr>
<td>0.9</td>
<td>0.08</td>
<td>0.97</td>
<td>3.2</td>
<td>0.57</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Fig. 3. ROC curve of FibroTest score, computed from five biochemical analytes and age and sex.
Area (SE) under the curve, 0.739 (0.050).
FibroTest scores might be useful to track the progression of fibrosis in individual patients. However, there are data describing the application of pre- and posttreatment FibroTest scores to HCV patients who had pre- and posttreatment liver biopsies available, showing that scores significantly decreased in those patients who responded to treatment (6).

Several other biochemical indicators have been tested specifically in HCV patients as potential serum markers of fibrosis, e.g., hyaluronic acid and matrix metalloproteinases with their tissue inhibitors. Hyaluronic acid was reported (14) to have a sensitivity and specificity of 85% and 88%, respectively, for detecting severe fibrosis (grades 4 and 5 of a 5-stage system). The ratio of matrix metalloproteinase-2 to tissue inhibitor metalloproteinase-1 was found to correlate with severity of fibrosis in HCV patients (15). It is possible that adding one or more of these tests to the algorithm would produce a final score with improved predictive values for grading fibrosis.

In conclusion, in clinical practice it has been suggested that the FibroTest score might be applied to patients who either have contraindications or refuse liver biopsy for the management of their HCV. We found that FibroTest scores could not accurately predict either the presence or absence of significant fibrosis and could not reliably be used to reduce the need for liver biopsy.

This study was supported by a grant from the Sir Charles Gairdner Hospital Research Foundation.

References