Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study

François Imbert-Bismut, Vlad Ratziu, Laurence Pieroni, Frederic Charlotte, Yves Benhamou, Thierry Poynard, for the MULTIVIRC group

Summary

Background Liver biopsy is thought mandatory for management of patients with hepatitis C virus (HCV) infection, especially for staging fibrosis. We aimed, in our prospective study, to assess the predictive value of a combination of basic serum biochemical markers for diagnosis of clinically significant fibrosis (including early stages).

Methods We assessed liver-biopsy patients with detectable HCV by PCR, for eligibility, and took a blood sample on the day of the procedure. The analysis was done in a first-year period for 205 patients and then tested in a second period on 134 patients. We devised a fibrosis index that included the most informative markers (combined with age and sex) for the first-year group. 11 serum markers were assessed as well as fibrosis stage: F0=no fibrosis and F1=portal fibrosis; and for clinically significant fibrosis, F2=few septa, F3=many septa, and F4=cirrhosis. Statistical analysis was by logistic regression, neural connection, and receiver-operating characteristic (ROC) curves.

Findings First-year and second-year patient-group characteristics and biochemical markers did not differ. The overall frequency of clinically significant fibrosis was 40% (138 patients). The most informative markers were: \( \alpha_1 \) macroglobulin, \( \alpha_2 \) globulin (or haptoglobin), \( \gamma \) globulin, apolipoprotein A\(_1\), \( \gamma \) glutamyltranspeptidase, and total bilirubin. The areas (SD) under the ROC curves for the first-year (0.836 [0.430]) and second-year groups (0.870 [0.340]) did not differ (p=0.44). With the best index, a high negative predictive value (100% certainty of absence of F2, F3, or F4) was obtained for scores ranging from zero to 0–10 (12% [41] of all patients), and high positive predictive value (>90% certainty of presence of F2, F3, or F4) for scores ranging from 0–60 to 1–00 (34% [115] of all patients).

Interpretation A combination of basic serum markers could be used to substantially reduce the number of liver biopsies done in patients with chronic HCV infection.

Lancet 2001; 357: 1069–75

Introduction

Liver biopsy is thought mandatory for management of patients infected by hepatitis C virus (HCV), particularly to stage fibrosis.\(^3,4\) However, after biopsy 30% of patients feel pain, 0–3% have severe complications, and 0–03% die.\(^5,6\) Several markers have substantial predictive values for diagnosis of cirrhosis,\(^4,10\) but none are available for diagnosis of earlier stages—eg, with few septa (the start of bridging fibrosis). No prospective studies have been done in a large population infected only by HCV. We aimed to prospectively assess the predictive value of a combination of basic serum biochemical markers for the diagnosis of clinically significant fibrosis (ranging from a few septa to cirrhosis) and necroinflammatory activity (necrosis and inflammation). If markers with high positive or negative predictive values of important fibrosis can be obtained, fewer liver biopsies would need to be done and thus the cost and risk of liver biopsy would be lessened.\(^3,4\)

Methods

Patients

From August, 1997, to March, 2000, all liver-biopsy patients who gave informed consent and with detectable HCV by PCR were assessed for eligibility and had a blood sample taken on the day of biopsy. Patients belonged to a single centre cohort, Cohorte Hépatite C Pitié-Salpêtrière (DOSVIRC). This cohort included all patients with HCV infection (defined by a positive serological test by at least a second-generation ELISA) attending the liver and gastrointestinal unit of Pitié-Salpêtrière Hospital, Paris, France.\(^11\) A questionnaire of 129 items was completed for every patient and included: sociodemographic data; risk factors; clinical, biological, virological, and treatment information from each visit; and histological data obtained at liver biopsy. The duration of HCV infection was estimated from transfusion date or first exposure to other parenteral sources, but it could not be calculated for patients with sporadic infection or those in whom the source of infection was unknown. Exclusion criteria were coinfection with HIV, hepatitis B virus, or other liver disease, and non-interpretable liver biopsy. The analysis was done in a first-year period on 205 patients and then tested in a second period on 134 patients.

Serum markers

11 markers were assessed: \( \alpha_1 \) macroglobulin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), \( \gamma \) glutamyltranspeptidase (GGT), total bilirubin, albumin, \( \alpha_1 \) globulin, \( \alpha_2 \) globulin, \( \beta \) globulin, \( \gamma \) globulin, and apolipoprotein A\(_1\), \( \alpha_1 \) globulins mainly consisted of \( \alpha_1 \) macroglobulin and haptoglobin; a retrospective assessment of haptoglobin was done for the two periods. Interleukin 10 (IL-10), tumour growth factor \( \beta_1 \) (TGF \( \beta_1 \), hepatocyte growth factor, apolipoprotein A\(_1\), and apolipoprotein B were also assessed in the second period.
AST, ALT, GGT, and total bilirubin were measured by autoanlyser (Hitachi 917 Automate; Mannheim, Germany) and Roche Diagnostics reagents (Mannheim, Germany). Albumin was assessed by bromocresol green method.\textsuperscript{10} Serum protein electrophoresis for $\alpha_1$, globulin, $\alpha_2$, globulin, $\beta$, globulin, and $\gamma$ globulin was done in an automatic system (Hydrasys and Hyrys, Sebia; Issy-Les-Moulineaux, France). Apolipoprotein $A_1$, apolipoprotein $B$, $\alpha_1$, macroglobulin, and haptoglobin were measured in serum samples with an automatic nephelometer (BNII, Dade Behring; Marburg, Germany). Plasma TGF $\beta_1$ and hepatocyte growth factor concentrations were measured with Quantikine human TGF $\beta_1$ immunoassay and Quantikine human hepatocyte growth factor immunoassay, respectively (R and D Systems; Minneapolis, MN, USA). Latent TGF $\beta_1$, was activated to the immunoreactive form by acid and then neutralised. Plasma IL-10 was measured with an immunoassay kit (Beckman Coulter Immunootech; Marseille, France). To prevent contamination with platelet-derived TGF $\beta_1$, blood samples were centrifuged 1 h after they were obtained, and serum samples were stored at $-80^\circ C$ until assays were done (<1 year for cytokines).

**Histological staging and grading**

Histological features of liver specimens were analysed with the METAVIR group scoring system.\textsuperscript{13,14} Liver biopsy specimens of more than 10 mm in length were fixed, paraffin-embedded, and stained with at least haematoxylin and eosin safran—and Masson's trichrome or picrosirius red for collagen. Every biopsy specimen was staged on a scale of F0 to F4: F0=no fibrosis, F1=portal fibrosis without septa, F2=few septa, F3=numerous septa without cirrhosis, and F4=cirrhosis. Histological activity, a measure of intensity of necroinflammatory lesions, was graded as follows: A0=no histological activity, A1=mild activity, A2=moderate activity, and A3=severe activity. The METAVIR scoring system was assessed by one pathologist (FC), who was unaware of patient characteristics.

**Statistical analysis**

Statistical analysis was by logistic regression, neural connection, and receiver-operating characteristic (ROC) curves.\textsuperscript{10,11} Analysis was done in the first year and then tested in the second period. In accordance with the METAVIR scoring system, patients were divided into several groups. The main endpoint was the identification of patients with substantial fibrosis (F2, F3, or F4) versus those without (F0 or F1). F2, F3, and F4 categories were grouped together because F2 is generally chosen as a threshold for treatment of chronic HCV infection.\textsuperscript{1} In secondary analyses, patients were also grouped by activity grades: those without much activity (A0 or A1), and patients with substantial activity (A2 or A3). Patients were also classified into three overall severity groups: without important histological features (A<2 and F<2); important lesions (A>2 and F>2, or both); and widespread fibrosis or cirrhosis (F3 or F4).

First, factors that differed significantly between these groups were identified by univariate analyses: $x^2$, student $t$ test, Mann-Whitney, and variance analysis with the Bonferroni all-pair-wise multiple comparison. The independent discriminative value of markers for the diagnosis of fibrosis was then assessed by logistic regression analysis. The third step was to construct an index that combined the independent factors. The best index for discrimination was the logistic regression function that combined the most discriminatory

422 patients with chronic HCV assessed

45 patients excluded: 37 HIV coinfection 7 hepatitis B infection 1 liver transplantation

38 patients excluded: 30 fibrosis not stageable 8 one or more markers missing

339 patients included

205 first-year group 134 second-year group

Figure 1: Study design

Independent markers were identified during the first year and their predictive values validated during the second period. HCV=hepatitis C virus.
IL-10, TGF β1, hepatocyte growth factor, apolipoprotein A1, and apolipoprotein B were assessed in the second period only, to explain the value of markers identified in the first period. However, their diagnostic values were assessed by the same methods as the other markers.

**Results**

**Participants**

We assessed 442 patients with chronic HCV infection for eligibility (figure 1). We excluded 45 because of: HIV coinfection (37), hepatitis B virus coinfection (seven), and transplantation (one). We could not stage fibrosis in 30 of the remaining 377 patients, and at least one of the 11 markers was missing in eight, which left 339 included patients. Patient characteristics and biochemical markers did not differ between first-year and second-year samples (table 1). The overall frequency of clinically significant fibrosis was 40% (138 patients).

**Diagnosis of fibrosis**

Table 2 shows diagnostic values (area under ROC curves) of biochemical markers and their independent association with fibrosis (logistic regression). Because ALT and AST were highly correlated (r=0.88), only ALT was used. The fibrosis indices combined ten or the six most informative markers (α1 macroglobulin, α2 globulin, total bilirubin, γ globulin, apolipoprotein A1, and GGT), or five markers (α1 globulin and γ globulin excluded; haptoglobin included). All indices had high diagnostic values in both year groups, whether groups were analysed separately or together (table 2). The ROC curves for the three indices were very similar (figure 2). The areas (SD) under the curves did not differ: 0.851 (0.370), 0.847 (0.370), and 0.837 (0.370), respectively. Figure 3 shows the box plots of the five-marker and six-marker fibrosis indices (in which scores range from zero to 1.00), for each fibrosis stage. With the six-marker index, a high negative predictive value (>90% certainty of absence of F2, F3, or F4) was obtained for scores from zero to 0.20. Of these 119 patients with low scores (35% of total) there were 13 false negatives: four F2, A0; six F2, A1; and three F2, A2. A high positive predictive value (>90% certainty of presence of F2, F3, or F4) was obtained for scores from 0.80 to 1.00. Of these 50 patients with high scores (15% of total), there were five false positives: two F1, A1, and three F1, A2.

When analysis was done on the second-year group only (table 3), and with predictive values greater than or equal to 90% deemed acceptable, biopsies could have been avoided in 62 patients of 134 (46%) with scores below 0.10 (16 patients, all F0 or F1) or above 0.60 (46 patients, only four F1). Neural connection methods gave similar results to logistic regression. The number of patients correctly classified (whether with important fibrosis or not) by Multilayer Perceptron with α1 macroglobulin, haptoglobin, GGT, total bilirubin, apolipoprotein A1, age, and sex was 130 of 163 (80%) for the training group, 61 of 82 (74%) for the validation group, and 59 of 81 (73%) for the test group.

In the second-year group, the addition of IL-10, TGF β1, apolipoprotein A1, and apolipoprotein B increased the area (SD) under the curve slightly to 0.889 (0.340), but did not differ (p=0.60) from the six-marker fibrosis index. None of the extra markers added significant diagnostic value to the indices by logistic regression or neural connection (data not shown).

43 patients had ALT lower than 35 IU/L, and ten of those had important fibrosis. The diagnostic value of the six-marker fibrosis index was still high, with an area under the ROC curve of 0.758 (0.590). Two patients with a score greater than 0.80 had cirrhosis. Among the 29 patients with scores lower than 0.20, 25 had no notable fibrosis.

**Tables 1 and 2**

**Table 1: Characteristics of included patients**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>First year</th>
<th>Second year</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>205</td>
<td>134</td>
<td>339</td>
</tr>
<tr>
<td>Age at biopsy (years)</td>
<td>47 (14)</td>
<td>48 (13)</td>
<td>47 (13)</td>
</tr>
<tr>
<td>Male</td>
<td>108 (53%)</td>
<td>88 (66)</td>
<td>196 (58%)</td>
</tr>
<tr>
<td>Female</td>
<td>97 (47%)</td>
<td>46 (34)</td>
<td>143 (42%)</td>
</tr>
</tbody>
</table>

**Table 2: Diagnostic value of biochemical markers for clinically significant fibrosis**

<table>
<thead>
<tr>
<th>Markers</th>
<th>First year</th>
<th>p</th>
<th>Second year</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>AST</td>
<td>0.773 (0.570)</td>
<td>0.13</td>
<td>0.679 (0.580)</td>
<td>0.35</td>
</tr>
<tr>
<td>α1 macroglobulin</td>
<td>0.749 (0.570)</td>
<td>0.0001</td>
<td>0.740 (0.460)</td>
<td>0.0001</td>
</tr>
<tr>
<td>ALT</td>
<td>0.725 (0.570)</td>
<td>0.09</td>
<td>0.564 (0.580)</td>
<td>0.11</td>
</tr>
<tr>
<td>Haptoglobin (decrease)*</td>
<td>0.704 (0.570)</td>
<td>0.02</td>
<td>0.654 (0.580)</td>
<td>0.006</td>
</tr>
<tr>
<td>γ globulin</td>
<td>0.680 (0.570)</td>
<td>0.16</td>
<td>0.670 (0.460)</td>
<td>0.59</td>
</tr>
<tr>
<td>GGT</td>
<td>0.672 (0.570)</td>
<td>0.03</td>
<td>0.705 (0.460)</td>
<td>0.01</td>
</tr>
<tr>
<td>Total bilirubin</td>
<td>0.611 (0.570)</td>
<td>0.69</td>
<td>0.726 (0.460)</td>
<td>0.008</td>
</tr>
<tr>
<td>Apo A1 (decrease)*</td>
<td>0.554 (0.570)</td>
<td>0.14</td>
<td>0.647 (0.580)</td>
<td>0.12</td>
</tr>
<tr>
<td>Albumin (decrease)*</td>
<td>0.514 (0.570)</td>
<td>0.12</td>
<td>0.662 (0.460)</td>
<td>0.53</td>
</tr>
<tr>
<td>α1 globulin</td>
<td>0.518 (0.570)</td>
<td>0.30</td>
<td>0.577 (0.580)</td>
<td>0.80</td>
</tr>
<tr>
<td>α2 globulin (decrease)*</td>
<td>0.508 (0.570)</td>
<td>0.007</td>
<td>0.518 (0.580)</td>
<td>0.03</td>
</tr>
<tr>
<td>β globulin</td>
<td>0.475 (0.570)</td>
<td>0.75</td>
<td>0.601 (0.580)</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Further reading:

- **For personal use. Only reproduce with permission from The Lancet Publishing Group.**
such fibrosis alone plus ALT (R²=0.297, p<0.0001). A high negative predictive value (>90% certainty of absence of F3 or F4) was obtained for scores from zero to 0.80. For scores greater than 0.80, there was a high positive predictive value (>85% certainty of presence of F3 or F4).

For diagnosis of cirrhosis or widespread fibrosis, the six-marker fibrosis index had a very large area (SD) under the ROC curve, 0.923 (0.370). A high negative predictive value (>90% certainty of absence of F3 or F4) was obtained for scores from zero to 0.80. For scores greater than 0.80, there was a high positive predictive value (>85% certainty of presence of F3 or F4).

For personal use. Only reproduce with permission from The Lancet Publishing Group.
Discussion

Our results show that a combination of five or six basic biochemical markers can have high positive or negative predictive value for diagnosis of clinically significant fibrosis, even at the early stage of a few septa. The most informative markers were, in decreasing rank: α₂ macroglobulin, haptoglobin, GGT, γ globulin, total bilirubin, and apolipoprotein A1.

α₂ globulin is mainly composed of α₂ macroglobulin and haptoglobin. Fibrosis was associated with an increase of α₂ macroglobulin and a decrease of haptoglobin, which masked the diagnostic value of α₂ globulin in univariate analysis. We have previously noted a significant diagnostic value of increased α₂ macroglobulin for fibrosis staging in patients with alcoholic liver disease, which has been confirmed. α₂ macroglobulin is an acute-phase protein and is produced at sites of inflammation and liver fibrosis by hepatocytes, stellate cells, and granuloma cells. Moreover, α₂ macroglobulin is related to fibrosis since it is a feature of stellate cell activation. It is also a proteinase inhibitor, and increased synthesis can inhibit catabolism of matrix proteins and enhance fibrotic processes in the liver. Haptoglobin was strongly and negatively associated with fibrosis, as already noted; this association was not related to haemolysis, hypersplenism, or hepatic insufficiency. Furthermore, we did not find an association with unconjugated bilirubin (data not shown).

These opposing correlations with fibrosis of α₂ macroglobulin (positive) and haptoglobin (negative) could be explained by the different roles of hepatocyte growth factor and TGF β1 in fibrogenesis and acute phase response. As seen in experimental fibrosis, increase in hepatocyte growth factor could account for the unexpected fall in reduction in TGF β1, rise of α₂ macroglobulin, and decrease of haptoglobin. Transduction with hepatocyte growth factor gene suppresses increase of TGF β1, and the factor stimulates synthesis of α₂ macroglobulin and reduces synthesis of haptoglobin.

GGT is associated with fibrosis and has been used with prothrombin and apolipoprotein A1 as a serum marker. Its diagnostic value in our study was independent of other factors, especially transaminases and bilirubin. GGT and bilirubin were both associated with hepatocyte growth factor. Early cholestasis or an increase of epidermal growth factor could be one explanation for the GGT increase with severity of fibrosis. γ Globulin serum concentration is associated with cirrhosis and portosystemic shunts. In our study, although lower than in patients with cirrhosis, it was already higher in patients with non-cirrhotic fibrosis than in those with scores of F1 or F0.
Apolipoprotein A1, serum concentration is associated with fibrosis and used with prothrombin and GGT as a serum marker.23 It is trapped on extracellular matrices. Apolipoproteins, especially ApoA1, also interact with HCV capsid proteins. Neither apolipoprotein A1 nor apolipoprotein B significantly added to the diagnostic value of apolipoprotein A1 alone. Albumin had no independent diagnostic value, probably because we had excluded patients with severe cirrhosis.

Finally, serum cytokines did not add much diagnostic value to the biochemical markers, which are easier and cheaper to measure than cytokines. Measurements of electrophoresis compounds (α-globulins and γ-globulins) are outdated semi-quantitative assessments. Their replacement by haptoglobin in a five-marker index did not significantly alter predictive value.

Extra-cellular matrix component markers (procollagen III peptide, hyaluronate, collagens, or collagenases) have diagnostic value, but are not significantly greater than other biochemical markers or prothrombin time18–19 when measured in serum. We did not include prothrombin time or platelets because we wanted a combination of inexpensive biochemical markers, which could be easily and automatically analysed.

We are confident that our study samples are representative of most patients. Clinical, histological, and biochemical characteristics of our prospective population were stable during the 33 months of the study and similar to those for populations in recent large randomised trials.20 We did not include patients with obvious decompensated cirrhosis. Inclusion of patients with severe liver disease would have artificially improved the predictive values of the logistic function. On the other hand, we included patients with very slight histological features (17% without fibrosis), and 43 (13%) with ALT lower than 35 IU/mL, who would usually not be included in randomised trials.

Results were similar with logistic regression or neural network. We preferred the logistic regression method, because with the neural connection the weight accorded to each factor is not certain. Whatever the statistical method, sampling variation poses potential difficulties, especially in early stages of disease when fibrosis might be unevenly distributed. Therefore, our biochemical markers might provide a more accurate description of fibrogenic events that occur across the whole liver. Our study was cross-sectional, and such markers should be longitudinally assessed. Use of our fibrosis score is inexpensive biochemical markers, which could be easily and automatically analysed.

Although our results need to be repeated by another centre, we think that the number of biopsies in the management of chronic HCV infection could be reduced (by up to 46% according to our result in the second-year period). Diagnostic value of our fibrosis index was highly reproducible between the two year groups. Patients were treated according to fibrosis stage and grade (1–4). If a decision not to treat were made without biopsy, with a fibrosis score of less than 0·20, only 13 of 119 patients would have been false negatives. Of these 13, none had cirrhosis or extensive fibrosis (F3 or F4) and only three had moderate activity. If a treatment decision without biopsy would have been made with a fibrosis score of greater than 0·80, only five patients of 50 would have been false positives. Of these, three had moderate activity, which justified treatment despite having only portal fibrosis. Two of these three patients underwent a transvenous liver biopsy, which showed raised portocaval gradients of 19 and 13 mm Hg. Therefore, these patients probably had substantial fibrosis. In this instance, two patients (4%) would have been unnecessarily treated. Therefore, our index can detect most patients who have moderate and severe histological activity but who do not have clinically significant fibrosis.

Other algorithms for fibrosis diagnosis should be compared with our fibrosis index.21 In clinical practice we plan to use this index to make decisions about treatment in patients who have contraindications to, or who refuse, liver biopsy. We also plan to prospectively test the hypothesis that use of this index, and a resulting reduction of liver biopsies, could be cost effective in management of HCV infection.

Contributors

P Imbert-Bismut and Laurence Pieroni did the biochemical assays and assessed quality control. V Ratziu and Y Benhamou recorded data and were responsible for execution of the study. P Charlot read liver biopsies. T Poynard initiated and designed the study, analysed data, and devised the fibrosis scores. All investigators helped to write the manuscript.

MULTIVIRC group (Paris): Christine Schall (administration); Frédéric Charlotte (anatomopathologie); Bruno Roux, Pierre Coriat (anesthesie-réanimation); Bernhard Hamique, Françoise Imbert-Bismuth, Laurence Pieroni, Annie Piron, Lina Khalil (biochimie); Françoise Mercadent (Centre de Transfusion Sanguine); Alain Pavie, Richard Dorent, Iradj Gandjbakhch (chirurgie cardiaque); Olivier Deckmyn, Jean René Vidaud (computer unit); Cécile Blot, Muriel Lechevalier, Hélène Villette (data management); Olivier Chosidow, Camille Francés (dermatologie); André Grimaldi (diabétologie); Philippe Giral, Eric Bruckett (endocrinologie); Pierre Opolon, Thierry Poynard, Yves Benhamou, Vlad Ratziu, Joseph Moussali, Corinne Regimbhand, Vincent Di Martino, Marie Bochet, Brigitte Bernard, Julien Taïeb, Dominique Thubat, Mercedes Torres, Luminum Bonyháyi, Marie Hélène Sayeg, Hélène Scapa, Marie Hélène Desailles (haut-gastroentérologie); Pascale Ghillani, Lucille Musset (inmunochimie); Brigitte Autran, Patrice Débré, Ionnis Theodorou, Guy Gorochov (immunologie cellulaire); Françoise Briaire, Martin Dusis, Christine Katlama (maladies infectieuses); Patrice Cacoub, Olivier Bousquet, Serge Hersov, Sonie Pelletier, Jean-Charles Piette (médecine interne); André Augereau, Laurence Leenhardt (ménécine nucleaire); Gilbert Deray, Sylvie Mouquet (nephrologue); Jean-Marc Leger, Olivier Lyon-Caen (neurologie); Eric Antoine, David Khayat (oncologie médicale); Marie-Hélène Fievret, Alain Thuiller (pharmacie); Olivier Ludicarme, Philippe Grenier (radiologie); Pierre Bourgeois, Nathalie Wrona (rhumatologie); Laurent Hannoun, Dominique Borie, Dominique Heraut (transplantation hépatique et chirurgie); Olivier Bitker (transplantation rénale); Jean-Marie Hurazus, Vincent Thibaut (virologie); Michel Vidaud, Martine Olivi, Franck Bonvouloir (EIA CNRS 8067 et Faculté de Pharmacie).

Acknowledgments

This study was supported by grants (ARC5117) from Association pour la Recherche sur le Cancer.

References


16 SPSS. Neural connection 1.0 user’s guide. Chicago, IL, 1995.


