Accuracy of a predictive model for severe hepatic fibrosis or cirrhosis in chronic hepatitis C

Agostino Colli, Alice Colucci, Silvia Paggi, Mirella Fraquelli, Sara Massironi, Marco Andreoletti, Vittorio Michela, Dario Conte

INTRODUCTION

Hepatitis C virus (HCV) infection is a major cause of chronic liver disease worldwide: nearly 80% of HCV-infected patients develop chronic infection, and about 20% progress to cirrhosis[1]. Liver histology is currently considered as the “reference standard” for evaluating hepatic damage on the basis of the degree of necroinflammatory activity and fibrosis, with the latter having prognostic significance and playing a major role in therapeutic decision making[2-3].

However, liver biopsy is an invasive procedure with mild and severe complications of 20% and 0.5%, respectively[4-5]. Furthermore, its sensitivity in diagnosing liver cirrhosis is not absolute and the rate of false negative results is approximately 30%-50%. In addition, as it has been well evidenced by the recent studies, an adequate liver specimen should be at least 2.5 cm long[6] and 1.4 mm wide[7] including at least 6-8 portal tracts[8]. This has recently led various groups to investigate the non-invasive methods of detecting severe fibrosis/cirrhosis, including only biochemical (ALT, AST, GGT) tests[9-10] or in combination with hematological (platelet count) tests[11], test panels[12-15], serum “markers” of fibrosis (such as hyaluronic acid or procollagen peptides)[16-17] and ultrasonographic parameters (e.g., liver surface nodularity (LSN), portal blood flow)[18-19], but none of which have proved to be capable of avoiding liver biopsy.

The aim of this prospective study was to evaluate the accuracy of a model based on the sequential combination of a set of simple biochemical tests (Bonacini score, BS)[20-21] and liver surface ultrasound (US) examination[22] in diagnosing severe fibrosis/cirrhosis in a large series of consecutive patients with chronic HCV infection undergoing liver biopsy.

Abstract

AIM: To assess the accuracy of a model in diagnosing severe fibrosis/cirrhosis in chronic hepatitis C virus (HCV) infection.

METHODS: The model, based on the sequential combination of the Bonacini score (BS: ALT/AST ratio, platelet count and INR) and ultrasonography liver surface characteristics, was applied to 176 patients with chronic HCV infection. Assuming a pre-test probability of 35%, the model defined four levels of post-test probability of severe fibrosis/cirrhosis: <10% (low), 10-74% (not diagnostic), 75-90% (high) and >90% (almost absolute). The predicted probabilities were compared with the observed patients’ distribution according to the histology (METAVIR).

RESULTS: Severe fibrosis/cirrhosis was found in 67 patients (38%). The model discriminated patients in three comparable groups: 34% with a very high (>90%) or low (<10%) probability of severe fibrosis, 33% with a probability ranging from 75% to 90%, and 33% with an uncertain diagnosis (i.e., a probability ranging from 10% to 74%). The observed frequency of severe fibrosis/cirrhosis was within the predefined ranges.

CONCLUSION: The model can correctly identify 67% of patients with a high (>75%) or low (<10%) probability of cirrhosis, leaving only 33% of the patients still requiring liver biopsy.

Key words: Liver fibrosis; Ultrasonography; Bonacini score; Liver biopsy; Hepatitis C

Colli A, Colucci A, Paggi S, Fraquelli M, Massironi S, Andreoletti M, Michela V, Conte D. Accuracy of a predictive model for severe hepatic fibrosis or cirrhosis in chronic hepatitis C. World J Gastroenterol 2005; 11(46): 7318-7322
http://www.wjgnet.com/1007-9327/11/7318.asp
MATERIALS AND METHODS

All anti-HCV positive patients (EIA III, Abbott Laboratories, Chicago, IL, USA) with detectable HCV-RNA serum levels (AmpliCor HCV kit, Roche, Molecular Systems, Basel, Switzerland) were evaluated for enrollment between September 2001 and June 2004.

The presence of decompensated liver disease (i.e., portosystemic encephalopathy, jaundice or ascites detected by means of US) and or an absolute contraindication to liver biopsy (PLT < 60 000/mm$^3$, PT < 60%) excluded patients from the study, whereas patients with an increase of ≥ 1.5 UNL in the serum alanine aminotransferase (ALT) levels recorded twice in the previous 6 months, were included in the study, whose protocol was approved by the pertinent ethics committee after having obtained their written informed consent. Socio-demographic and clinical data were recorded as those on past and/or current alcohol intake for which a semi-quantitative questionnaire with three pre-defined levels of daily consumption (<30, 30-80 and >80 g) was used. At the time of liver biopsy, the laboratory data included AST, ALT (IU/L r.v. < 40), platelet count (10$^7$/µL r.v. 130 000-400 000), the international normalized ratio (INR 0.8-1.2), level of albumin (g/dL r.v. 3.5-5.0), total bilirubin (mg/dL r.v. 0.3-1.0) and hemoglobin (g/dL r.v. 14-18 for men and 12-16 for women), as well as HCV genotyping by means of restriction fragment length polymorphism after amplification of the 5’ non-coding region of the HCV genome.$^{[35]}$ As detailed in Table 1, the “cirrhosis discriminant score” (range 0-11) of each patient was calculated according to the method of Bonacini et al.$^{[19]}$, based on ALT/AST ratio, platelet count and INR. After an overnight fast, a US liver scan was performed using an ATL HDI 5000 equipment (Advanced Technology Laboratories, Bothell, WA, USA) and both 3.5 and a 5-12 MHz transducer by one of three gastroenterologists (AC, SM or MA) blinded to the clinical, biochemical and histologic data. A 5-12 MHz transducer was used to obtain multiple scans of the outer 2-3 cm of the liver parenchyma and of both lobes. LSN was considered positive when the liver surface appeared as a dotted or an irregular line and/or the liver parenchyma was not homogeneous, with areas of different echogenicity, reflecting an underlying nodularity. The interobserver agreement for this sign was calculated according to K statistics.

A US-guided transcostal or subcostal liver biopsy was performed (by AC or MF) using an 18-gauge needle (Biomol Hospital Service, Pomezia, Rome, Italy). We considered acceptable only specimens ≥ 2.0 cm long including ≥ 12 portal tracts, that were fixed in formalin and stained with hematoxylin-eosin, silver impregnation for reticulin, and Masson’s trichrome or picrosirius red for collagen. In case of inadequacy of the sample, a second biopsy was obtained.

The specimens evaluated by the same pathologist (VM) unaware of the patients’ characteristics were staged according to METAVIR scoring system.$^{[30]}$, where F0 indicates the lack of fibrosis, F1 corresponds to portal fibrosis without septa, F2 to the presence of few septa, F3 and F4 to the finding of numerous septa without or with cirrhosis, respectively. Histologic findings were considered as the reference standard for the presence and degree of fibrosis.

**Table 1 Determinants of Bonacini score**

| Laboratory Parameters | Score |
|-----------------------|-------|
| INR                   |       |
| ≤ 1.1                 | 0     |
| 1.1-1.4               | 1     |
| >1.4                  | 2     |
| ALT/AST ratio         |       |
| ≤ 1.7                 | 0     |
| 1.7-1.2               | 1     |
| ≥ 1.9-0.6             | 2     |
| PLT ×1000/mm$^3$      |       |
| ≤ 340                 | 0     |
| 340-280               | 1     |
| 279-220               | 2     |
| 219-160               | 3     |
| 159-100               | 4     |
| 99-40                 | 5     |
| <40                   | 6     |

and F4 to the finding of numerous septa without or with cirrhosis, respectively. Histologic findings were considered as the reference standard for the presence and degree of fibrosis.

**Statistical analysis**

The pre-test probability of severe fibrosis/cirrhosis (i.e., a staging score 3-4) was estimated to be about 35% on the basis of recent data from comparable series.$^{[2,32,37-39]}$ A predictive model was obtained by the sequential application of the BS and US liver surface examination, which could obtain the post-test probability of severe fibrosis/cirrhosis. As the BS and US signs of LSN could be considered conditionally independent, the probability of severe fibrosis/cirrhosis obtained by calculating the BS has been considered as the pre-test probability before the US examination. The post-test probabilities were calculated on the basis of post-test odds as previously described.$^{[40,41]}$ The values of sensitivity, specificity, positive and negative likelihood ratios (LR+ and LR-) of both BS and US in diagnosing severe fibrosis have been previously defined.$^{[36]}$

In detail, three possibilities had to be taken into account. Firstly, for a staging score 0-2, a BS ≤ 3 had a sensitivity of 58% and a specificity of 85%, accounting for a LR+ of 4. Accordingly, in our model in the presence of a BS ≤3, the probability of stage 0-2 increased from a pre-test one of 65% to 88%. Secondly, the sensitivity and specificity for severe fibrosis in the presence of a BS of 4-7 were 50%, making it impossible to modify the 35% pre-test probability of severe fibrosis/cirrhosis on the basis of a LR+ and a LR- of about one. Thirdly, the specificity of 100% for a BS > 7 led to an almost absolute post-test probability of staging score 3-4, with a LR+ higher than 100. As far as the US finding of LSN was concerned, an inter-observer variability of 0.80 was reported. Our own previous data$^{[32]}$ demonstrated a sensitivity of 54% and a specificity of 95% for a staging score 3-4, with a LR+ and a LR- of 11.6 and 0.5, respectively. By combining the BS and the US signs of LSN, five groups of patients (I-V) were defined (Figure 1) with different post-test probability of severe fibrosis/cirrhosis. In the last group, with BS ≥ 7 the post-test probability of a staging score of 3-4 was close to 100% regardless of the US results. Finally, as detailed in Figure 1, to test the predictive accuracy and clinical usefulness of the model, four clinically acceptable ranges of post-test probability of severe fibrosis/cirrhosis were predefined based on BS and US findings: low probability (<10%), not diagnostic (10-74%), high probability (75-90%) and almost absolute probability (>90%).

**RESULTS**
During the recruitment period, 1,089 patients were investigated for liver disease, and 363 were chronically infected with HCV. Of them, 184 were excluded for the refusal of liver biopsy (#11) because of already established cirrhosis with Child-Pugh score $\geq 5$ (#67), or because of normal or slightly elevated ALT levels (<1.5 times the upper normal limit) or uncompleted ALT data (less than six distinct determinations) (#106). The remaining 179 fulfilled the inclusion criteria. Based on the inadequate US findings, three were excluded accounting for an overall fulfillment of the inclusion criteria. Based on the inadequate US examination was technically inadequate in three cases), to the vast majority of enrolled patients (98%, because the US examination was technically inadequate in three cases), has been proved to have an adequate discriminating power and an accurate calibration. Its discriminating power is demonstrated by the fact that we were able to divide the patients into three comparable groups: 34% with a very high (>90%) or low (<10%) probability of severe fibrosis/cirrhosis.

**DISCUSSION**

In this prospective study, we have evaluated the accuracy of a non-invasive predictive model, consisting of the sequential combination of laboratory data score (BS) and US finding of a nodular liver surface, in detecting severe liver fibrosis or cirrhosis in patients with chronic hepatitis C. In clinical practice, it is easy to use the two elements with their operative characteristics individually validated in previous studies.[19,20,32] The BS is based on simple and widely available laboratory tests (ALT, AST, INR and platelet count) that assure its repeatability, and LSN can be interpreted by the model, are shown in Table 3. Overall, the diagnostic accuracy of the model was 86.5% in groups I, IV, and V, accounting for 67% of the patients. Sixteen of one hundred and nineteen patients were incorrectly defined (1 in group I with fibrosis stage $\geq 3$, 15 in group IV with fibrosis stage $<3$). In groups II and III (33% of the total), the model did not show any predictive role and did not significantly modify the pre-test probability of severe fibrosis/cirrhosis.

**Table 2** Clinical and biochemical characteristics of 176 consecutive patients with chronic hepatitis C included in

| Age (yr) | % | Mean±SD |
|----------|---|---------|
| Male     | 96 | 55      |
| Female   | 80 | 45      |
| BMI (kg/m²) | | |
| $<25$    | 90 | 51      |
| $\geq 25$| 86 | 49      |
| Alcohol intake (g/d) | | |
| $\leq 30$ | 142 | 80    |
| 31-80    | 24 | 14      |
| $>80$    | 10 | 6       |
| HCV genotype | | |
| 1        | 92 | 52      |
| 2        | 60 | 34      |
| 3        | 19 | 11      |
| 4        | 5  | 3       |
| Platelet count | | |
| 181±128  | | |
| AST (IU/L) | 78.5±101 |
| ALT (IU/L) | 125±178  |
| AST/ALT   | 1.6±1.1  |
| INR       | 1.05±0.1 |
| Hemoglobin (g/dL) | 14.5±2.1 |
| Albumin (g/dL) | 4.3±0.6 |
| Bilirubin (mg/dL) | 0.9±0.2 |

During the recruitment period, 1,089 patients were investigated for liver disease, and 363 were chronically infected with HCV. Of them, 184 were excluded for the refusal of liver biopsy (#11) because of already established cirrhosis with Child-Pugh score $>5$ (#67), or because of normal or slightly elevated ALT levels (<1.5 times the upper normal limit) or uncompleted ALT data (less than six distinct determinations) (#106). The remaining 179 fulfilled the inclusion criteria. Based on the inadequate US findings, three were excluded accounting for an overall fulfillment of the inclusion criteria. Based on the inadequate US examination was technically inadequate in three cases), has been proved to have an adequate discriminating power; thus, few patients have a very low or high probability of fibrosis, and most of them are in the gray zone of uncertainty.[19,20,32] On the contrary, the sequential model, which is applicable to the vast majority of enrolled patients (98%, because the US examination was technically inadequate in three cases), has been proved to have an adequate discriminating power and an accurate calibration. Its discriminating power is demonstrated in the fact that we were able to divide the patients into three comparable groups: 34% with a very high (>90%) or low (<10%) probability of severe fibrosis,
33% with a probability ranging between 75% and 90%,
and 33% with an uncertain diagnosis (i.e., with a fibrosis
score 3-4 probability between 10% and 74%). The precise
calibration of the model is demonstrated by the fact that
the observed frequency of severe fibrosis/cirrhosis was
within the predefined ranges for each group (Table 3).

The validity of the model is further supported by the
finding that the 38% frequency of severe fibrosis/cirrhosis
in the study population was similar to the theoretical value
of 35% assumed by considering series of patients with
comparable characteristics, represented by a $\geq 1.5$ times
increase in ALT levels recorded twice during a period of 6
months or more.

The operative characteristics of US are slightly different
from those indicated by our own recent data[32,38], which
were used for the model. However, the variations in LR+
and LR- remained within the broad confidence intervals
of the original estimates and did not critically challenge
the calibration of the model. Further studies are
needed to assess the generalizability of the model, i.e., its
reproducibility and transportability, for which independent
validations are advisable.

Although our simple model could detect the presence
or absence of severe fibrosis in nearly two-thirds of
the patients, thus avoiding the need for liver biopsy,
some limitations of the study require further discussion.
Because of its intrinsic characteristics, the model was
unable to detect moderate (stage 2) fibrosis, which
together with grading is relevant in therapeutic decision
making. Furthermore, the study seemed to underestimate
the diagnostic role of liver biopsy, which was not used
only to define staging and grading, but also to detect
concomitant pathologies. However, the latter role which
is already questionable in patients with liver disease of
uncertain etiology[15-16] was irrelevant in our series of patients
with chronic HCV infection. Conversely, liver biopsy
still has some limitations in clinical practice, such as the
difficulty in obtaining a liver sample of adequate size, as
recently reported by some studies[7,8], and which can be
considered as representative of the whole liver, given the
patchy distribution of liver fibrosis in chronic hepatitis
C. The above limitations in our opinion are a major issue
in evaluating the diagnostic accuracy of non-invasive
diagnostic tests compared to liver biopsy. In fact in these
types of studies, the major drawback is the imperfection
of the reference standard itself, which can significantly
affect the estimates of the operative characteristics of the
test under investigation. To reduce the possible sample
size effect, in the present study we obtained all the liver
specimens $\geq 3$ cm with $\geq 12$ portal tracts from patients.

Previous studies assessing the individual role of BS
and US findings of LSN in detecting severe fibrosis have
found that they are inaccurate because their low power of
discrimination means that a large proportion of patients
still require liver biopsy[8,9,12,13]. It has recently been shown
that the use of laboratory tests[10,11] and specific fibrosis
markers, such as hyaluronic acid[8,17], is highly accurate
in diagnosing cirrhosis. However, the considerable
differences in the study design and the particularly high
prevalence of symptomatic cirrhosis sometimes make
it difficult to compare these data to our own. It has also
recently been shown that relatively complex laboratory
scores are accurate in discriminating stage 0-1 from stage
2-4, but fail to confirm severe fibrosis[12-15] although the
aim of these studies was to detect patients with minimal or
absent fibrosis with stage 3-4 found in only 15% (or 37% in
our series), possibly because they included patients with
normal ALT levels and a younger age.

Recently, preliminary data from patients with chronic
HCV liver disease[5,12] suggest a promising role of fibroscan
(one-dimensional (1-D) transient elastography) in
predicting the degree of liver fibrosis with a good accuracy
in discriminating stage 0-2 from 3-4, and a LR+ and LR-
of 5.7 and 0.16, respectively, similar to those observed in
the present series for the US sign of LSN (LR+ and LR-
of 7.5 and 0.5, respectively).

In conclusion, our data suggest that the use of liver
biopsy to detect severe fibrosis can be avoided in about
two-thirds of patients. Furthermore, the model can also be
used in patients aged more than 65 years for whom liver
biopsy could be questioned.

REFERENCES

1. Seeff LB, Hoofnagle JH. National Institutes of Health
Consensus Development Conference: management of hepatitis
C. 2002. Hepatology 2002; 36: S1-S2
2. Khan MH, Farrell GC, Byth K, Lin R, Weltman M, George
J, Samarasinghe D, Kench J, Kaba S, Crewe E, Liddle C
Which patients with hepatitis C develop liver complications?
Hepatology 2000; 31: 513-520
3. Poynard T, Ratziu V, Bemmanov Y, Di Martino V, Bedossa
P, Opolon P. Fibrosis in patients with chronic hepatitis C:
detection and significance. Semin Liver Dis 2000; 20: 47-55
4. Cadanel 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 (AFEP). Hepatology 2000; 32: 477-481
5. Piccione F, Sagnelli E, Pasquale G, Giusti G. Complications
following percutaneous liver biopsy. A multicentre
retrospective study on 68,276 biopsies. J Hepatol 1986; 2: 165-173
6. Poniachik J, Bernstein DE, Reddy KR, Jeffers Lj, Coelho-
Little ME, Civantos F, Schiff ER. The role of laparoscopy in
the diagnosis of cirrhosis. Gastrointest Endosc 1996; 43: 568-571
7. Bedossa P, Dargère D, Paradis V. Sampling variability of liver
fibrosis in chronic hepatitis C. Hepatology 2003; 38:1449-1457
8. Colloredo G, Guido M, Sonzogni A, Leandro G. Impact of liver
biopsy size on the histological evaluation of chronic viral
hepatitis: the smaller the sample the milder the disease. J
Hepatol 2003; 39: 239-244
9. Scheuer PJ. Liver biopsy size matters in chronic hepatitis: bigger
is better. Hepatology 2003; 38: 1356-1358
10. Giannini E, Botta F, Fasoli A, Ceppa P, Rioso D, Lantieri PB,
Celle G, Testa R. Progressive liver functional impairment is
associated with an increase in AST/ALT ratio. Dig Dis Sci
1999; 44: 1249-1253
11. Sheth SG, Flamm SL, Gordon FD, Chopra S. AST/ALT ratio
predicts cirrhosis in patients with chronic hepatitis C virus
infection. Am J Gastroenterol 1998; 93: 44-48
12. Haber MM, West AB, Haber AD, Reuben A. Relationship of
aminotransferases to liver histological status in chronic
hepatitis C. Am J Gastroenterol 1995; 90: 1250-1257
13. Assy N, Minuk GY. Serum aspartate but not alanine
aminotransferase levels help to predict the histological
features of chronic hepatitis C viral infections in adults. Am J
Diagnostic value of serum type IV collagen test in comparison with platelet count for predicting the fibrotic stage in patients with chronic hepatitis C virus infection: a prospective study. Lancet 2001; 357: 1069-1075

Forns X, Ampurdanés S, Llovet JM, Aponte J, Quintó L, Martínez-Bauer E, Bruguera M, Sánchez-Tapias JM, Rodes J. Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. Hepatology 2002; 36: 986-992

Bonacini M, Hadi G, Govindarajan S, Lindsay KL. Utility of a discriminant score for diagnosing advanced fibrosis or cirrhosis in patients with chronic hepatitis C virus infection. Am J Gastroenterol 1997; 92: 1302-1304

Saadeh S, Carmell G, Carey WD, Younossi Z, Barnes D, Di Bona D, Schepis F, Heathcote EJ, Zeuzem S, Gramantieri L, Venturoli N, Piscaglia F, Siringo Ziol M, Kazemi F, de Lédinghen V, Marcellin P, Dhumeaux D, Lindor KD. An assessment of the role of liver biopsies in chronic hepatitis C. Hepatology 2001; 33: 196-200

Peynard T, Munteanu M, Imbert-Bismut F, Charlotte F, Thabut D, Le Calvez S, Messous D, Thibault V, Benhamou Y, Moussalli J, Ratziu V. Prospective analysis of discordant results between biochemical markers and biopsy in patients with chronic hepatitis C. Clin Chem 2004; 50: 1344-1355

Le Calvez S, Thabut D, Messous D, Munteanu M, Ratziu V, Imbert-Bismut F, Peynard T. The predictive value of Fibrotest vs. APRI for the diagnosis of fibrosis in chronic hepatitis C. Hepatology 2004; 39: 862-863; author reply 863

Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, Lok AS. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 2003; 38: 518-526

Sud A, Hui JM, Farrell GC, Bandara P, Kench JG, Fung C, Lin R, Samarasinghe D, Liddle C, McCaughan GW, George J. Improved prediction of fibrosis in chronic hepatitis C using measures of insulin resistance in a probability index. Hepatology 2004; 39: 1239-1247

Rossi E, Adams L, Prins A, Bulsara M, de Boer B, Garas G, MacQuillan G, Spears D, Jeffrey G. Validation of the FibroTest biochemical markers score in assessing liver fibrosis in hepatitis C patients. Clin Chem 2003; 49: 450-454

Oberti F, Valsesia E, Pilette C, Rousselet MC, Bedossa P, Aubé C, Gallois Y, Riflett H, Mgaya MA, Pennent-Fontbonne D, Cales P. Noninvasive diagnosis of hepatic fibrosis or cirrhosis. Gastroenterology 1997; 113: 1609-1616

Muraiwaki Y, Ikuta Y, Okamoto K, Koda M, Kawasaki H. Diagnostic value of serum markers of connective tissue turnover for predicting histological staging and grading in patients with chronic hepatitis C. J Gastroenterol 2001; 36: 399-406

Muraiwaki Y, Koda M, Okamoto K, Mimura K, Kawasaki H. Diagnostic value of serum type IV collagen test in comparison with platelet count for predicting the fibrotic stage in patients with chronic hepatitis C. J Gastroenterol Hepatol 2001; 16: 777-781

McHutchison JG, Blatt LM, de Medina M, Craig JR, Conrad A, Schiff ER, Tong MJ. Measurement of serum hyaluronic acid in patients with chronic hepatitis C and its relationship to liver histology. Consensus Interferon Study Group. J Gastroenterol Hepatol 2000; 15: 945-951