"Liverscore" is predictive of both liver fibrosis and activity in chronic hepatitis C

Shoukat Ali Arain, Qamar Jamal, Amir Omair

Abstract

AIM: To formulate a noninvasive index predictive of severity of liver fibrosis and activity in chronic hepatitis C.

METHODS: This cross sectional study was conducted on polymerase chain reaction positive, treatment naïve patients. Fibrosis was staged on a five point scale from F0-F4 and activity was graded on a four point scale from A0-A3, according to the METAVIR system. Patients were divided into two overall severity groups, minimal disease (< F2 and < A2) and significant disease (≥ F2 or ≥ A2). Eleven markers were measured in blood. Statistically, the primary outcome variable was identification of minimal and significant overall disease. Indices were formulated using β regression values of different combinations of nine statistically significant factors.

Diagnostic performance of these indices was assessed through receiver-operating characteristic curve analysis.

RESULTS: A total of 98 patients were included and of these 46 had an overall clinically significant disease. Our final six marker index, Liverscore for Hepatitis C, consisted of age, alanine transaminase, gamma-glutamyl transpeptidase, apolipoprotein A-1, alpha-2 macroglobulin and hyaluronic acid. The area under the curve was found to be 0.813. On a 0-1 scale, negative predictive value at a cutoff level of ≤ 0.40 was 83%, while positive predictive value at ≥ 0.80 remained 89%. Altogether, 61% of the patients had these discriminative scores.

CONCLUSION: This index is discriminative of minimal and significant overall liver disease in a majority of chronic hepatitis C patients and can help in clinical decision making.

Key words: Chronic hepatitis C; Staging and grading; Liver fibrosis and activity; Noninvasive assessment; Liverscore

Peer reviewer: Hussein M Atta, MD, PhD, Department of Surgery, Faculty of Medicine, Minia University, Mir-Aswan Road, El-Minia 61519, Egypt

Arain SA, Jamal Q, Omair A. "Liverscore" is predictive of both liver fibrosis and activity in chronic hepatitis C. World J Gastroenterol 2011 November 7; 17(41): 4607-4613 Available from: URL: http://www.wjgnet.com/1007-9327/full/v17/i41/4607.htm DOI: http://dx.doi.org/10.3748/wjg.v17.i41.4607

INTRODUCTION

Chronic hepatitis C (CHC) induces injury and inflammation of the liver, which appears to be responsible for
the associated fibrogenesis[1]. Morbidity and mortality in CHC is associated with the development of cirrhosis and its complications. The rate of fibrosis progression varies markedly from person to person and over time. The risk of developing cirrhosis varies from 10% to 20% over a period of 20 years[2]. Treatment of CHC is complex, costly, and associated with side effects that are difficult to accept in a population that is predominantly asymptomatic. Furthermore, about half of the patients with genotype 1, and slightly lesser than that in other genotypes, fail to respond to anti-viral therapy[3,4].

Treatment decisions are recommended to be individualized on the basis of severity of the liver disease, treatment response rates, co-morbid conditions and the readiness of the patient for treatment[5]. Therefore, assessment of the fibrosis stage and rapidity of progression of fibrosis (necro-inflammation) may help in determining the prognosis and the need of therapy in an individual patient. A prevalence peak of advanced fibrosis and cirrhosis in CHC patients is expected during this decade. Thus, increasing numbers of patients will require assessment. Furthermore, with the development of anti-fibrotic therapies, there will be a need for regular and more frequent monitoring[6].

In underdeveloped countries like Pakistan, with a high prevalence of disease and resource constraints[7,8], it seems unrealistic to offer treatment to all patients. There is a need to identify patient categories to rationalize the need for therapy. For patients showing minimal disease, treatment may be deferred with follow up for disease progression. Treatment may be offered to patients with progressive disease or else, safer, better tolerated, and more cost effective therapy will become available.

Liver biopsy is the current tool for the assessment of liver disease; nonetheless it is an invasive procedure and may be associated with complications. Moreover, the biopsy facility is not available in remote areas and is not always possible[9]. Alternative strategies are being actively evaluated, such as imaging and non-invasive biochemical monitoring of liver disease. Some of the biochemical markers, especially panels of multiple markers in the form of indices are promising, and may reduce the number of liver biopsies for assessment of liver disease[10]. The purpose of this study was to evaluate the predictive value of noninvasive biomarkers for the diagnosis of overall liver disease categories in CHC.

**MATERIALS AND METHODS**

This was a cross sectional study to determine the diagnostic accuracy of noninvasive biomarkers, conducted at Ziauddin University Hospital, Karachi, Pakistan from June 2006 to July 2010. The study was approved by Ethics Review Committee of the university. Treatment naïve polymerase chain reaction proven CHC patients, of 20 to 60 years of age, were included in the study. Patients with HBV co-infection and diabetes mellitus were excluded. Patients with history of other chronic inflammatory conditions, alcohol intake, and blood disorders requiring frequent blood transfusions were also excluded. A questionnaire was completed for every patient to document possible variables, such as demographic factors, and to rule out causes of exclusion. Written informed consent was obtained from each patient.

**Liver biopsy**

A percutaneous liver biopsy was performed with a 16-18 gauge modified Menghini aspiration needle (Surecu® TSK, Japan). Tissue was formalin fixed, paraffin-embedded, and processed for light microscopic examination. Along with standard hematoxylin and eosin staining, slides were also stained with connective tissue stains. Only liver biopsy specimens of more than 10 mm length and having not less than five portal tracts were included in the study[11].

Biopsy specimens with evident pathology, but without identification of the correct number of portal tracts, were also included. Histological features of the liver biopsy specimens were analyzed according to the METAVIR group scoring system[11,12], by one pathologist (Jamal Q), without any knowledge of the clinical or biochemical data. Every specimen was staged for fibrosis on a five-point scale; F0 = no fibrosis; F1 = portal fibrosis without septa; F2 = portal fibrosis with rare septa; F3 = numerous septa without cirrhosis; and F4 = cirrhosis. There is an element of subjectivity in classifying patients in different stages, especially in biopsy specimens, because of the majority of hepatic lobules being partial. However, we considered any fibrosis beyond portal tracts as significant (F2). Fibrosis was considered F3 when 50% or more portal tracts showed fibrous septa extending beyond the portal tracts. Necroinflammatory lesions were graded on a four point scale on the basis of an algorithm based on the severity of focal lobular necro-inflammation and piecemeal necrosis; A0 = no histological activity, A1 = mild activity, A2 = moderate activity, and A3 = severe activity[12].

On the basis of fibrosis stage and necro-inflammatory grade patients were divided into two overall severity groups; minimal disease \( \leq \) F2 and \(< \ A2 and significant disease \( \geq \) F2 or \( \geq \) A2[13].

**Biochemical markers**

Chemical analysis was carried out for alanine transaminase (ALT), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase (ALP), total and direct bilirubin, apolipoprotein A-1 (Apo A-1), haptoglobin, alpha-2 macroglobulin (A2M), hyaluronic acid (HA), hydroxyproline (HYP) and proline.

The ALT assay was performed by an International Federation of Clinical Chemistry standardized ultraviolet enzymatic method, GGT was assayed by an enzymatic colorimetric method, the ALP assay was performed by a colorimetric method and total and direct bilirubin assays were performed according to the method described by Jendrassik and Grof. All enzymatic activities were measured at 37°C. All the above assays were performed using reagents from Roche Diagnostics®. A2M and haptoglobin assays were performed by an immuno-turbidimetric method using polyclonal rabbit anti-human antibodies (DakoCytomation Denmark Code. Nos. Q0102 &
### Table 1 Baseline characteristics of all patients (n = 98)

| Characteristic          | Value            | Significance |
|-------------------------|------------------|--------------|
| Age (yr)                | 36.0 ± 10.6      |              |
| Male                    | 52 (53)          |              |
| Activity grade          |                  |              |
| Absent (A0)             | 32 (33)          |              |
| Mild (A1)               | 40 (41)          |              |
| Moderate (A2)           | 20 (20)          |              |
| Severe (A3)             | 6 (6)            |              |
| Fibrosis stage          |                  |              |
| No fibrosis (F0)        |                  |              |
| Portal fibrosis without septa (F1) | 35 (36) |              |
| Portal fibrosis with rare septa (F2) | 29 (30) |              |
| Numerous septa without cirrhosis (F3) | 10 (10) |              |
| Cirrhosis (F4)          |                  |              |

Results are shown as mean ± SD or n (%).

### Table 2 Univariate analysis for significant and minimal disease

| Characteristic          | Significant (n = 46) | Minimal (n = 52) | P value |
|-------------------------|----------------------|------------------|---------|
| Age (yr)                | 39.6 ± 11.4          | 32.9 ± 8.7       | 0.002   |
| Gender                  |                      |                  |         |
| Male                    | 28 (61)              | 24 (46)          | 0.15    |
| Female                  | 18 (39)              | 28 (56)          |         |
| Bilirubin total (mg/dL) | 0.62 ± 0.34          | 0.49 ± 0.30      | 0.04    |
| Bilirubin direct (mg/dL)| 0.21 ± 0.17          | 0.15 ± 0.08      | 0.04    |
| ALT (U/L)               | 53.2 ± 39.2          | 32.60 ± 15.90    | 0.002   |
| ALP (U/L)               | 80.4 ± 26.9          | 81.70 ± 32.30    | 0.78    |
| GGT (U/L)               | 51.8 ± 47.2          | 23.80 ± 16.90    | < 0.001 |
| Haptoglobin (g/L)       | 1.05 ± 0.45          | 1.05 ± 0.49      | 0.99    |
| A2M (g/L)               | 2.56 ± 0.64          | 2.24 ± 0.55      | 0.009   |
| Apo-A1 (mg/dL)          | 96.85 ± 23.70        | 114.2 ± 29.80    | 0.002   |
| Hydroxyproline (μmol/L)| 9.7 ± 7.0            | 13.5 ± 9.30      | 0.03    |
| Proline (μmol/L)        | 123.0 ± 78.10        | 144.2 ± 106.4    | 0.27    |
| Hyaluronic acid (mg/mL) | 100 ± 150            | 23.90 ± 16.40    | 0.001   |

Statistical analysis was performed using SPSS version 17.0 (Chicago, IL). A χ² test was used to compare the qualitative and independent t test was used to compare the quantitative variables. P value of less than 0.05 was considered statistically significant.

### RESULTS

Initially 104 CHC patients were enrolled. Six patients were excluded; the biopsy specimens of four patients had less than five portal tracts, one biopsy showed a granuloma, and assays for three biochemical markers could not be performed in one patient.

#### Baseline characteristics

Table 1 shows the baseline characteristics of 98 patients, 52 (53%) patients were male (Male: Female = 1.15:1). The mean age was 36.0 ± 10.6 years. On liver biopsy, 26 (27%) patients had an activity grade of A2-3 and 42 (43%) had F2-4 fibrosis stages. Except for 4 patients, all the patients with A2 or A3 activity on liver biopsy also had F2-4 fibrosis. The major determinant of overall significant disease category was thus found to be the stage of the disease. In aggregate, 46 (47%) patients were classified as having clinically significant overall disease (A2-3 or F2-4).

#### Univariate analysis

The demographic and biochemical variables were compared for their association with different overall disease categories (Table 2). The mean age of patients with significant disease was significantly higher (39.6 ± 11.4 year vs 32.9 ± 8.7 year, P = 0.002), while gender distribution was not different between significant and minimal disease groups (P = 0.15). The mean total and direct bilirubin levels were significantly higher in the significant disease group (P = 0.04 for both). The mean ALT level was significantly higher in significant disease group than the minimal disease group (P = 0.002), but the mean difference of ALP level was statistically insignificant between the two groups (P = 0.78). Mean GGT was significantly higher in the significant disease group (P < 0.001). Mean difference of haptoglobin was not statistically significant according to overall disease categories (P = 0.99). The mean of A2M was significantly higher as in the significant disease group (P = 0.002), but mean Apo-A1 was significantly lower in the significant disease group (P = 0.002). The mean hydroxyproline was also significantly lower in the significant disease group (P = 0.03), while proline was not statistically different in minimal and significant
In the univariate analysis, nine variables were identified as significantly associated with two groups of overall significance (p = 0.001).

### Formulation and assessment of indices

In the univariate analysis, nine variables were identified as significantly associated with two groups of overall significant disease. These included age of the patient, bilirubin total and direct, ALT, GGT, A2M, Apo-A1, HYP and HA. As bilirubin total and direct were highly correlated (r = 0.85), we included bilirubin total only. Although gender was not significantly associated with histological categories, we included it in some indices to see if it improved their performance. Various combinations of the factors identified in the univariate analysis, were assessed by logistic regression. By keeping the overall disease categories as a binary variable, regression coefficients of the factors were obtained to generate indices. The diagnostic value of these indices was assessed by an area under ROC (AUROC) curve. In the nine marker index, age, gender, bilirubin total, ALT, GGT, A2M, Apo-A1, HYP and HA were included. Bilirubin total was excluded in the eight marker index. Gender and HYP were excluded in the seven marker index. Gender, total bilirubin, and HYP were excluded in the six marker index. AUROC (±SE) was found to be 0.831 (0.05) for both the nine and eight marker indices and 0.813 (0.05) for both the seven and six marker indices. Although indices having HYP as a component i.e., the nine and eight marker indices performed slightly better, we excluded these indices because the difference was marginal and measurement of HYP is expensive and is not available in routine clinical laboratories. The AUROC was similar for seven and six marker indices.

Finally, we selected a six marker index formulated from beta regression values keeping overall disease category as a binary variable. The formula used to generate this index is given below (with ALT and GGT expressed in U/L, A2M as G/L, Apo-A1 as mg/dL and HA as ng/mL).

\[
\text{Six marker index} = \beta_0 + \beta_1 \times \text{Age} + \beta_2 \times \text{ALT} + \beta_3 \times \text{GGT} + \beta_4 \times \text{A2M} + \beta_5 \times \text{Apo-A1} + \beta_6 \times \text{HYP}
\]

The obtained scores were converted to a standardized index ranging from 0-1 through percent ranking as Liver- scoring for Hepatitis C (Table 3).

The ROC curve showing the sensitivity and specificity for overall disease categories is shown in Figure 1, with an AUROC (±SE) of 0.813 (0.05). The specificity, sensitivity, positive predictive value (PPV) and negative predictive value (NPV) were calculated at different levels. Important cutoff points are shown in Table 4. Negative predictive value of Liverscore for Hepatitis C at a cutoff level of 0.30 was 84%. Six out of 31 patients who were below this cutoff were wrongly diagnosed (false nega-
tives). All patients had F2 fibrosis and minimal activity. At a cutoff point of 0.40, NPV remained at 83%. Seven out of 40 patients below this cutoff were false negatives, again all had F2 fibrosis. For the positive predictive value, to identify the presence of significant fibrosis, cutoff levels of $\geq 0.70$ and $\geq 0.80$ were compared, and had 82% and 89% positive predictive values, respectively. Out of 29 patients with a score of 0.70 and above, five were false positives, all with F1 fibrosis. At a cutoff point of 0.80, two patients out of 20 in this study were false positives. Again both had F1 fibrosis.

Thus, a liverscore for hepatitis C of 0.40 or below reliably excludes the presence of significant disease and a score of 0.80 and above confirms the presence of significant disease.

**DISCUSSION**

The rate of fibrosis progression in CHC patients varies markedly from person to person, and only a minority suffers from long term complications\(^{[6]}\). The current use of liver biopsy for the assessment of liver histology has many drawbacks. Noninvasive assessment of liver histology has been the focus of research for many years. Isolated markers of liver cell injury and fibrosis have not proved to be sufficiently reliable for clinical use\(^{[7]}\). Development of indices consisting of multiple markers is now being focused to distinguish between minimal and clinically significant fibrosis categories.

Our final index, liverscore for hepatitis C, consisted of six markers, ALT, GGT, A2M, Apo-A1, hyaluronic acid and age of the patient. AUROC with this index was found to be 0.813 for overall disease. This index was also evaluated for fibrosis stage and activity grade separately, with clinically acceptable diagnostic performance (data not shown). The negative predictive value of Liverscore for Hepatitis C at a cutoff level of $\leq 0.40$ was 83%. All the patients (7/40) diagnosed as false negatives had F2 fibrosis and minimal activity. For the positive predictive value, a cutoff level $\geq 0.80$ was found suitable, with a PPV of 89%. At this cutoff, two patients out of 20 were false positives, both had F1 fibrosis. The diagnostic performance of this index in terms of AUROC is comparable to other similar indices reported in the literature.

Forns’ index\(^{[10]}\) consists of age of the patient, GGT, cholesterol, and platelets. AUROC was 0.86 in the formulation and 0.81 in the validation group. Using the cut off score of $< 4.2$, presence of significant fibrosis could be excluded in 36% (125/351) patients, with a NPV of 96% in the formulation group. The majority of the patients in this cohort had genotype 1. This index includes cholesterol, which is metabolized differently in genotype 3; it has been suggested\(^{[7]}\) that this index might not perform well in patients having genotype 3, the most common genotype affecting the Pakistani population\(^{[10]}\).

AST to platelet ratio index (APRI) is a very simple and widely validated index that amplifies the opposing effects of liver fibrosis on AST and platelet counts\(^{[10]}\). The AUROC curve of APRI for prediction of significant fibrosis remained 0.80 in training and 0.88 in the validation set. In one study from Pakistan\(^{[16]}\), it showed an AUROC of 0.82 for significant fibrosis. At a cutoff point of $< 0.5$, the authors could exclude the presence of significant fibrosis in 36% (43/120) patients, with an NPV of 78%. Our index performed slightly better than this at a cutoff point of $\leq 0.40$, with the exclusion of 41% patients and an NPV of 83%. Our index has a better NPV and these scores were present in 41% (40/98) patients. This has a clinical advantage of identifying patients that can safely be deferred for urgent treatment.

Patented Fibrotest\(^{®}\) for fibrosis consists of age and gender of the patient, GGT, total bilirubin, haptoglobin, A2M and Apo A-1. In addition, the same authors have also reported Actitest\(^{®}\) for necroinflammation, which includes ALT in addition to Fibrotest\(^{®}\) biomarkers\(^{[17]}\). This is the most widely validated noninvasive marker and is in clinical use. The AUROC for the identification of liver fibrosis was 0.84 and 0.87 for the formulation and validation group, respectively. The PPV of this index was excellent ($> 90\%$ certainty of presence of F2, F3 or F4) for scores ranging from 0.60 to 1.00 (34% of all patients). This index could exclude the presence of significant fibrosis in 12% of patients, with a high negative predictive value ($100\%$ certainty of absence of F2, F3 or F4) for scores ranging from zero to 0.10. This high accuracy of Fibrotest was not uniform in different populations. A study carried out by Rossi et al., in the Australian population demonstrated, a PPV of 78% at a Fibrotest\(^{®}\) score of $> 0.6$ and an NPV of 85% at $< 0.1\%$. We included all the components of Fibrotest\(^{®}\) in our initial analysis; however, haptoglobin and gender of the patients were not discriminative of minimal and clinically significant disease in our cohort of patients in univariate analysis; Thus, these were not included in the final index.

Hepascore consists of age and gender of the patient, bilirubin, A2M and HA. For significant fibrosis, it showed an AUROC of 0.85 and 0.82 in the training and validation groups, respectively. A score of $\geq 0.5$ was 92% specific and 67% sensitive in the training set and 89% specific and 63% sensitive in the validation group. At this cutoff it provided high PPVs of 87% and 88% in the training and validation set, respectively\(^{[20]}\). Authors have not reported negative predictive values for significant fibrosis. We have all the data required for the calculation of Hepascore and will evaluate the performance of this score in our patients.

All the above mentioned scores either predict fibrosis or activity. One advantage of our index is its prediction for overall disease, which includes both fibrosis stage and activity grade. Both these histopathological categories are important for prognosis and making treatment decisions\(^{[23]}\). Furthermore, the majority of the previous indices have been reported in populations infected predominantly with HCV genotype 1. Evidence points towards the possibility that HCV genotype 3 associated CHC is a metabolically different disease\(^{[24]}\). Our index
might perform better in genotype 3 patients, because it is formulated in a population predominantly infected with this genotype\[18\].

All the factors included in our index are available and easily programmable on automated instruments in routine clinical laboratories. Furthermore, factors included in the Liverscore for Hepatitis C have physiological rationale.

Gamma-glutamyl transpeptidase is synthesized by the liver cells, its synthesis increases with fibrosis. The mechanisms for this increase could be the stimulation of GGT synthesis by epidermal growth factor during fibrogenesis\[29\]. ALT is synthesized by hepatocytes and its release into serum is related to liver cell injury\[28\]. The synthesis of A2M increases during stellate cell activation in the course of fibrogenesis, and its serum concentration increases with fibrosis\[27\]. In liver fibrosis, Apo A-1 release from the hepatocytes is hampered by the collagen fibers decreasing its serum levels\[28\]. Hyaluronic acid is a nonsulfated glycosaminoglycan and is major component of extracellular matrix. Among the direct markers of liver fibrosis, HA has been most extensively studied in CHC. It increases in the liver during fibrogenesis and is released into the systemic circulation during remodeling. Recent indices consisting of HA in combination with indirect markers have shown promising results\[23\].

Some of the markers we evaluated were not helpful in differentiating minimal from significant disease. Total and direct bilirubin were significantly associated with different histological categories, but were highly correlated ($r = 0.85$). Direct bilirubin was therefore excluded. Total bilirubin was included in the seven marker index, but its exclusion did not affect the diagnostic value and was thus excluded from the final index. Serum levels of ALP are known to be raised in both alcoholic and non-alcoholic liver disease with advanced histological changes. In addition, ALP has shown a discriminative value for advanced fibrosis and cirrhosis previously\[29\]. It was also associated with mild and advanced fibrosis categories in our study, but not for overall disease, the primary outcome in our study, thus, we did not include ALP in our index. HYP and proline are amino acids present in collagen in large quantities. The HYP content of liver biopsies is found to increase with advancing stage of fibrosis\[30\]. The evidence that fibrosis is a dynamic two way process with fibrosis and its degradation occurring simultaneously, prompted us to include these amino acids as products of collagen degradation in our panel of biomarkers. We expected their high levels in serum because of the greater amount of collagen undergoing remodeling in advanced stages. We found no association of these amino acids of collagen degradation with severity of liver fibrosis or necroinflammatory activity. HYP, however, was statistically associated with overall disease category, but actually decreased. In the only study we could find, predictive value of these amino acids in sera of CHC patients was evaluated for advanced (F3 and 4) and mild (F0, 1 and 2) fibrosis\[33\]. Proline was not significantly different between the two groups, while HYP was found to be increased with advanced fibrosis, but showed a low (0.525) area under ROC curve. We found no study comparing these amino acids in minimal and significant fibrosis or overall disease. The evidence that their serum levels do not increase with increasing fibrosis might be explained by slower fibrosis degradation in advanced fibrosis. It has been shown that accumulation of fibrosis is the net effect of increased fibrogenesis and its decreased degradation\[33\].

One limitation of our study was that we could not validate our results in a different cohort of patients. This was not possible because of the smaller number of patients recruited in our study. We recommend an independent study for the validation of our index.

In conclusion, a liverscore for hepatitis C of 0.40 or below excludes the presence of significant disease. Thus, it can reliably exclude around 41% of CHC patients that do not require an urgent treatment. A score of 0.80 and above confirms the presence of significant disease. Using these cutoff values, the severity of the liver disease can reliably be predicted in around 61% of the CHC patients.
