Assessing Liver Fibrosis with Serum Marker Models

*Enrico Rossi,1 Leon A Adams,2,3 Max Bulsara,4 Gary P Jeffrey2,3
1Department of Clinical Biochemistry, PathWest Laboratory Medicine, Queen Elizabeth II Medical Centre, 2Department of Gastroenterology and Hepatology, Sir Charles Gairdner Hospital, 3School of Medicine and Pharmacology, The University of Western Australia, 4School of Population Health, The University of Western Australia, Nedlands, WA 6009 Australia.
*For correspondence: Dr Enrico Rossi e-mail: ric.rossi@health.wa.gov.au

Abstract
Chronic liver disease is characterised by liver fibrosis, which may lead to cirrhosis. Conventional serum-based liver function tests do not give information on either the presence or the rate of progress of liver fibrosis. The reference diagnostic test to detect fibrosis is liver biopsy, a procedure subject to various limitations, including risk of patient injury and sampling error.

Serum markers have been evaluated for the determination of fibrosis either singly or combined as a panel of markers, however diagnostic accuracy is greatest in studies using a panel together with an algorithm, which generates a predictive score. Serum marker models, especially those targeted at hepatitis C, have multiplied in spectacular fashion over the last five years, with most models regularly achieving a median area under the receiver operating characteristic curve (ROCC) of 0.80 versus liver biopsy. Five years after publication of the first major serum marker model, the first study to document clinical outcomes reported that applying the model to hepatitis C patients improved prediction of decompensated cirrhosis and survival compared to liver biopsy.

An obstacle to widespread adoption of serum marker models has been the lack of uniform performance indicators, such as diagnostic odds ratios and likelihood ratios. At present, serum marker models are not considered sufficiently reliable to replace liver biopsy in patients with chronic liver disease. However with continued evaluation in parallel with liver biopsy rapid advances are being made.

Liver Fibrosis
Liver fibrosis can accompany almost any chronic liver disease characterised by the presence of inflammation or hepatobiliary distortion. Fibrosis or scarring arises as a result of wound repair and is the net result of the balance between fibrinogenesis (production of extracellular matrix) and fibrolysis (degradation of extracellular matrix). Scar formation alters liver structure and the liver responds with regeneration. A review of liver fibrosis was recently published by Friedman.¹

Progressive fibrosis of the hepatic parenchyma leads to cirrhosis, nodule formation, altered hepatic function and risk of liver-related morbidity and mortality. The commonest liver diseases causing fibrosis and possible cirrhosis are chronic viral hepatitis or steatohepatitis associated with either alcohol or obesity. Other aetiologies include autoimmune attack on hepatocytes (autoimmune hepatitis) or biliary epithelium (primary biliary cirrhosis, primary sclerosing cholangitis), inherited metabolic conditions such as haemochromatosis, neonatal liver disease, parasitic liver disease such as schistosomiasis, chronic inflammatory conditions such as sarcoidosis, drug toxicity and vascular derangements. Cirrhosis typically develops over many years or decades, although occasionally it occurs rapidly, for example in neonatal liver disease. Once considered irreversible, ample evidence now exists that reversal of cirrhosis is possible when the underlying pathogenic insult is eliminated, for example the causative virus.³

Liver Fibrosis in Hepatitis C
Serum models were initially developed to predict fibrosis in patients chronically infected with the hepatitis C virus and most published data on serum marker model systems has been obtained in these patients. In addition, the natural history of liver fibrosis is best understood for this condition, where the course of liver fibrosis is very variable, ranging from decades of viraemia with little fibrosis to rapid onset of cirrhosis in 10-15 years. The available evidence shows that host factors rather than viral factors correlate with fibrosis progression. The main risk factors for more rapid progression include:
Assessing Liver Fibrosis with Serum Marker Models

*Enrico Rossi,1 Leon A Adams,2,3 Max Bulsara,4 Gary P Jeffrey2,3
1Department of Clinical Biochemistry, PathWest Laboratory Medicine, Queen Elizabeth II Medical Centre, 2Department of Gastroenterology and Hepatology, Sir Charles Gairdner Hospital, 3School of Medicine and Pharmacology, The University of Western Australia, 4School of Population Health, The University of Western Australia, Nedlands, WA 6009 Australia.
*For correspondence: Dr Enrico Rossi e-mail: ric.rossi@health.wa.gov.au

Abstract
Chronic liver disease is characterised by liver fibrosis, which may lead to cirrhosis. Conventional serum-based liver function tests do not give information on either the presence or the rate of progress of liver fibrosis. The reference diagnostic test to detect fibrosis is liver biopsy, a procedure subject to various limitations, including risk of patient injury and sampling error.

Serum markers have been evaluated for the determination of fibrosis either singly or combined as a panel of markers, however diagnostic accuracy is greatest in studies using a panel together with an algorithm, which generates a predictive score. Serum marker models, especially those targeted at hepatitis C, have multiplied in spectacular fashion over the last five years, with most models regularly achieving a median area under the receiver operating characteristic curve (ROCC) of 0.80 versus liver biopsy. Five years after publication of the first major serum marker model, the first study to document clinical outcomes reported that applying the model to hepatitis C patients improved prediction of decompensated cirrhosis and survival compared to liver biopsy.

An obstacle to widespread adoption of serum marker models has been the lack of uniform performance indicators, such as diagnostic odds ratios and likelihood ratios. At present, serum marker models are not considered sufficiently reliable to replace liver biopsy in patients with chronic liver disease. However with continued evaluation in parallel with liver biopsy rapid advances are being made.

Liver Fibrosis
Liver fibrosis can accompany almost any chronic liver disease characterised by the presence of inflammation or hepatobiliary distortion. Fibrosis or scarring arises as a result of wound repair and is the net result of the balance between fibrinogenesis (production of extracellular matrix) and fibrolysis (degradation of extracellular matrix). Scar formation alters liver structure and the liver responds with regeneration. A review of liver fibrosis was recently published by Friedman.1

Progressive fibrosis of the hepatic parenchyma leads to cirrhosis, nodule formation, altered hepatic function and risk of liver-related morbidity and mortality. The commonest liver diseases causing fibrosis and possible cirrhosis are chronic viral hepatitis or steatohepatitis associated with either alcohol or obesity. Other aetiologies include autoimmune attack on hepatocytes (autoimmune hepatitis) or biliary epithelium (primary biliary cirrhosis, primary sclerosing cholangitis), inherited metabolic conditions such as haemochromatosis, neonatal liver disease, parasitic liver disease such as schistosomiasis, chronic inflammatory conditions such as sarcoidosis, drug toxicity and vascular derangements. Cirrhosis typically develops over many years or decades, although occasionally it occurs rapidly, for example in neonatal liver disease. Once considered irreversible, ample evidence now exists that reversal of cirrhosis is possible when the underlying pathogenic insult is eliminated, for example the causative virus.3

Liver Fibrosis in Hepatitis C
Serum models were initially developed to predict fibrosis in patients chronically infected with the hepatitis C virus and most published data on serum marker model systems has been obtained in these patients. In addition, the natural history of liver fibrosis is best understood for this condition, where the course of liver fibrosis is very variable, ranging from decades of viraemia with little fibrosis to rapid onset of cirrhosis in 10-15 years. The available evidence shows that host factors rather than viral factors correlate with fibrosis progression. The main risk factors for more rapid progression include:
Rossi E

has been developed to describe the morphological changes of a few septae (F2); septal fibrosis without cirrhosis (F3) and portal fibrosis without septae (F1); portal fibrosis with fibrosis are described as follows: chronic hepatitis without fibrosis with the five point scale META VIR system, where fibrosis in hepatitis C patients are compared to biopsy results obtained disease throughout the liver, multiple passes of a trucut needle technique for assessing fibrosis, the presence of uniform accuracy of liver biopsy include the use of a semi-quantitative status of the liver. Factors which improve the diagnostic source of additional information such as steatosis and iron liver biopsy assessed histopathologically has long been the 'gold standard' for describing liver histology, disease activity and liver fibrosis. Biopsy also provides a unique additional 'gold standard' for describing liver histology, disease activity and liver fibrosis. Biopsy also provides a unique additional source of additional information such as steatosis and iron status of the liver. Factors which improve the diagnostic accuracy of liver biopsy include the use of a semi-quantitative technique for assessing fibrosis, the presence of uniform disease throughout the liver, multiple passes of a trucut needle and a biopsy of 2 cm or greater in length.

The most widely used systems for grading activity and staging fibrosis are the semi-quantitative Ishak and META VIR systems. Most serum marker models used to predict fibrosis in hepatitis C patients are compared to biopsy results obtained with the five point scale META VIR system, where fibrosis is described as follows: chronic hepatitis without fibrosis (F0); portal fibrosis without septae (F1); portal fibrosis with a few septae (F2); septal fibrosis without cirrhosis (F3) and complete cirrhosis (F4). A separate scale, the Brunt system, has been developed to describe the morphological changes of non-alcoholic fatty liver disease.

Limitations of Liver Biopsy
There are several issues impacting on the use of liver biopsy that prevent its routine use as a clinical tool. Some are beyond the scope of this review, for example lack of manpower to undertake biopsies on all patients who require it, associated cost and risk of patient injury. However, three limitations especially relevant to the application of serum marker models, should be briefly discussed.

Fibrosis Staging Systems
Although histologic staging of fibrosis is widely used, it is based on two flawed assumptions: firstly, it is not appropriate to describe a continuous variable such as the amount of fibrosis with categorical values such as fibrosis stages. Secondly, the staging systems assume a linear increase in the severity of fibrosis between stages, although it is recognised that this is not true. A serum-based model giving an algorithm-based score is a continuous variable and may be a more valid parameter.

Sampling Error
Sampling error is an intrinsic problem of biopsy. A 10-15 mg sample of tissue represents a tiny fraction of an organ weighing 1500 g. Even a disease like hepatitis C that affects the liver relatively uniformly will vary from lobule to lobule, although the error is typically not greater than one fibrosis stage. In one study, simultaneous biopsies were taken laparoscopically from right and left hepatic lobes from 124 patients with hepatitis C. A difference of at least one stage between right and left lobes was documented in 33% of patients, which could not be attributed to intra-observer variation, which was low. Only two patients had a difference of two fibrosis stages. Sampling error is especially evident in small biopsies.

Inter-observer Variation
The third limitation is inter-observer variability amongst pathologists in categorising the degree of fibrosis, which is considered to be up to 20%. Assessment of fibrosis remains subjective and it is difficult to compare results of different studies using different scoring systems, for example Ishak and META VIR.

The Case for Serum Marker Models
Aside from the limitations of liver biopsy, there is an urgent need to develop non-invasive serum markers for the following reasons:

1. There is increasing evidence that even advanced fibrosis is reversible. Having shown that severe disease is amenable to therapy, a requirement arises for more frequent testing than allowed by liver biopsy.

2. It is expected that antifibrotic therapies will be developed which will require early and regular monitoring of response to establish effectiveness and optimise dosing. As noted above, the need for regular monitoring will greatly exceed what is appropriate for liver biopsy.

Several non-invasive diagnostic imaging tests for fibrosis and cirrhosis, which do not involve testing serum, have been
older age at infection; concurrent liver disease due to hepatitis B virus or alcohol (>50g per day); male gender; hepatic steatosis; infection with human immunodeficiency virus; immunosuppression and iron overload.

Standard clinical indices cannot distinguish between degrees of fibrosis and clinical management requires identification of risk factors for progressive fibrosis and determining the duration of infection, even if the latter is an estimate. Although information on progression of fibrosis is extremely valuable, estimates of progression are tempered by the observation that fibrosis progression is not entirely linear and more advanced stages are probably associated with accelerating progression. However, an estimate of the current degree of fibrosis is valuable for the following reasons:

1. The actual stage of fibrosis will indicate the likelihood of response to treatment, with advanced stages generally having an inferior response rate;
2. If progression is slow, treatment with antiviral therapy may be less urgent;
3. The approximate time to the development of cirrhosis can be estimated.

Liver Biopsy
Liver biopsy assessed histopathologically has long been the ‘gold standard’ for describing liver histology, disease activity and liver fibrosis. Biopsy also provides a unique additional source of additional information such as steatosis and iron status of the liver. Factors which improve the diagnostic accuracy of liver biopsy include the use of a semi-quantitative technique for assessing fibrosis, the presence of uniform disease throughout the liver, multiple passes of a trucut needle and a biopsy of 2 cm or greater in length.

The most widely used systems for grading activity and staging fibrosis are the semi-quantitative Ishak and META VIR systems. Most serum marker models used to predict fibrosis in hepatitis C patients are compared to biopsy results obtained with the five point scale META VIR system, where fibrosis is described as follows: chronic hepatitis without fibrosis (F0); portal fibrosis without septae (F1); portal fibrosis with a few septae (F2); septal fibrosis without cirrhosis (F3) and complete cirrhosis (F4). A separate scale, the Brunt system, has been developed to describe the morphological changes of non-alcoholic fatty liver disease.

Limitations of Liver Biopsy
There are several issues impacting on the use of liver biopsy that prevent its routine use as a clinical tool. Some are beyond the scope of this review, for example lack of manpower to undertake biopsies on all patients who require it, associated cost and risk of patient injury. However, three limitations especially relevant to the application of serum marker models, should be briefly discussed.

Fibrosis Staging Systems
Although histologic staging of fibrosis is widely used, it is based on two flawed assumptions: firstly, it is not appropriate to describe a continuous variable such as the amount of fibrosis with categorical values such as fibrosis stages. Secondly, the staging systems assume a linear increase in the severity of fibrosis between stages, although it is recognised that this is not true. A serum-based model giving an algorithm-based score is a continuous variable and may be a more valid parameter.

Sampling Error
Sampling error is an intrinsic problem of biopsy. A 10-15 mg sample of tissue represents a tiny fraction of an organ weighing 1500 g. Even a disease like hepatitis C that affects the liver relatively uniformly will vary from lobule to lobule, although the error is typically not greater than one fibrosis stage. In one study, simultaneous biopsies were taken laparoscopically from right and left hepatic lobes from 124 patients with hepatitis C. A difference of at least one stage between right and left lobes was documented in 33% of patients, which could not be attributed to intra-observer variation, which was low. Only two patients had a difference of two fibrosis stages. Sampling error is especially evident in small biopsies.

Inter-observer Variation
The third limitation is inter-observer variability amongst pathologists in categorising the degree of fibrosis, which is considered to be up to 20%. Assessment of fibrosis remains subjective and it is difficult to compare results of different studies using different scoring systems, for example Ishak and META VIR.

The Case for Serum Marker Models
Aside from the limitations of liver biopsy, there is an urgent need to develop non-invasive serum markers for the following reasons:

1. There is increasing evidence that even advanced fibrosis is reversible. Having shown that severe disease is amenable to therapy, a requirement arises for more frequent testing than allowed by liver biopsy.
2. It is expected that antifibrotic therapies will be developed which will require early and regular monitoring of response to establish effectiveness and optimise dosing. As noted above, the need for regular monitoring will greatly exceed what is appropriate for liver biopsy.

Several non-invasive diagnostic imaging tests for fibrosis and cirrhosis, which do not involve testing serum, have been
to be identified, it is useful to consider the properties of an ideal marker:
1. Specific for liver;
2. Readily available and standardised between all laboratories performing diagnostic biochemistry/haematology;
3. Not subject to false positive results, for example due to inflammation;
4. Identifies the stage of fibrosis.

Most current serum markers are not liver specific, or may represent impaired hepatic clearance or are affected by inflammation. In addition, coexisting pathologies such as haemolysis (causing a decrease in haptoglobin (Hap) levels and/or increase in Bil levels) or rheumatoid arthritis (increase in hyaluronic acid (HA) levels) are associated with changes in levels of serum markers.

Although no ideal marker exists, several have been identified as possible useful indicators of fibrosis. Single markers often correlate with fibrosis in large groups of patients but are not sufficiently predictive in the individual patient, especially when used longitudinally over time. A systematic review compared single and multiple markers versus liver biopsy up to 2002 and noted diagnostic accuracy was greatest in studies using multiple markers.  

In practice, serum markers are therefore used in combination where they have achieved a greater likelihood for success in discriminating minimal from significant fibrosis. Usually, three or more markers are used in combination in an algorithm to generate a score, which is then used to give a fibrosis prediction. The serum markers listed below have been chosen because they are common components of published serum models used to make fibrosis predictions. The list is not intended to be complete. A brief rationale is given for the use of each marker followed by comments on the available methods for analysis. Newer approaches such as proteomics, metabolomics and clinical glycomics are expected to yield more novel biomarkers.

Major Serum Fibrosis Markers

**HA**

This mucopolysaccharide is a glycosaminoglycan, a high molecular weight polymer present in joints and in some tissues such as liver. It is found in synovial fluid and serum levels are elevated in various chronic liver diseases due to HA production by hepatic stellate cells and decreased clearance by sinusoidal endothelial cells. Serum levels are normally <50 µg/L and elevated levels correlate reasonably well with the degree of liver fibrosis in alcoholic liver disease and hepatitis C.  

**α-2-Macroglobulin (α-2-M)**

This is a high molecular weight protein synthesised in hepatocytes and stellate cells which is reasonably abundant in human serum, where normal levels are typically from 0.66 to 2.65 g/L. The functions of α-2-M are not well understood but it does inhibit the catabolism of matrix proteins by acting as a broad-spectrum inhibitor of nearly all enzymes that split proteins internally (endoproteases). Serum levels increase with the degree of liver fibrosis.

The preferred methods for analysis are immunonephelometry and immunoturbidimetry and reagents are available as commercial kits from manufacturers of immunonephelometry platforms such as the Beckman Coulter IMMAGE and Dade-Behring BNII.

**Collagen Markers**

This diverse group of markers includes pro-collagen peptides, proteins such as type I, type III and type IV collagen and collagen metabolites such as laminin. For example, the N-terminal propeptide of type III collagen (PIIINP) is a valuable fibrosis marker that has been validated in alcoholic liver disease, hepatitis C, and non-alcoholic fatty liver disease. Serum levels increase with the degree of liver fibrosis. A typical analysis method for PIIINP is by heterogeneous immunoassay using magnetic particle separation techniques on an automated analyser (Bayer Healthcare AG, Leverkusen, Germany).

**Apolipoprotein A1 (Apo A1)**

This is the major protein component found in high-density lipoprotein. Serum concentrations are negatively associated with liver fibrosis, i.e. levels decrease as the extent of fibrosis increases. Decreased levels are also seen in uncontrolled diabetes, nephrotic syndrome, some diets and smoking. As for α-2-M, the preferred method for analysis of Apo A1 is an immunonephelometry platform such as the Beckman Coulter IMMAGE or Dade-Behring BNII.

**Haptoglobin**

This serum protein binds any free haemoglobin present in the circulation. Hap is an acute phase protein whose concentrations
to be identified, it is useful to consider the properties of an ideal marker:
1. Specific for liver;
2. Readily available and standardised between all laboratories performing diagnostic biochemistry/haematology;
3. Not subject to false positive results, for example due to inflammation;
4. Identifies the stage of fibrosis.

Most current serum markers are not liver specific, or may represent impaired hepatic clearance or are affected by inflammation. In addition, coexisting pathologies such as haemolysis (causing a decrease in haptoglobin (Hap) levels and/or increase in Bil levels) or rheumatoid arthritis (increase in hyaluronic acid (HA) levels) are associated with changes in levels of serum markers.

Although no ideal marker exists, several have been identified as possible useful indicators of fibrosis. Single markers often correlate with fibrosis in large groups of patients but are not sufficiently predictive in the individual patient, especially when used longitudinally over time. A systematic review compared single and multiple markers versus liver biopsy up to 2002 and noted diagnostic accuracy was greatest in studies using multiple markers. In practice serum markers are therefore used in combination where they have achieved a greater likelihood for success in discriminating minimal from significant fibrosis. Usually, three or more markers are used in combination in an algorithm to generate a score, which is then used to give a fibrosis prediction. The serum markers listed below have been chosen because they are common components of published serum models used to make fibrosis predictions. The list is not intended to be complete. A brief rationale is given for the use of each marker followed by comments on the available methods for analysis. Newer approaches such as proteomics, metabolomics and clinical glycomics are expected to yield more novel biomarkers.

Major Serum Fibrosis Markers

**HA**
This mucopolysaccharide is a glycosaminoglycan, a high molecular weight polymer present in joints and in some tissues such as liver. It is found in synovial fluid and serum levels are elevated in various chronic liver diseases due to HA production by hepatic stellate cells and decreased clearance by sinusoidal endothelial cells. Serum levels are normally <50 µg/L and elevated levels correlate reasonably well with the degree of liver fibrosis in alcoholic liver disease and hepatitis C. HA has been used on its own to exclude significant fibrosis or more recently in combination with other markers.

Testing for serum HA is currently not widely available. It is available commercially as a self contained kit in the 96-well ELISA format (Corgenix, Colorado, US) the most efficient use requires a plate washer to perform the multiple rinses and washes as well as a plate reader to read the final absorbance. If demand increases, it is anticipated that serum HA will become available on commercial automated platforms.

**α-2-Macroglobulin (α-2-M)**
This is a high molecular weight protein synthesised in hepatocytes and stellate cells which is reasonably abundant in human serum, where normal levels are typically from 0.66 to 2.65 g/L. The functions of α-2-M are not well understood but it does inhibit the catabolism of matrix proteins by acting as a broad-spectrum inhibitor of nearly all enzymes that split proteins internally (endoproteases). Serum levels increase with the degree of liver fibrosis. The preferred methods for analysis are immunonephelometry and immunoturbidimetry and reagents are available as commercial kits from manufacturers of immunonephelometry platforms such as the Beckman Coulter IMMAGE and Dade-Behring BNII.

**Collagen Markers**
This diverse group of markers includes pro-collagen peptides, proteins such as type I, type III and type IV collagen and collagen metabolites such as laminin. For example, the N-terminal propeptide of type III collagen (PIIINP) is a valuable fibrosis marker that has been validated in alcoholic liver disease, hepatitis C, and non-alcoholic fatty liver disease. Serum levels increase with the degree of liver fibrosis. A typical analysis method for PIIINP is by heterogeneous immunooassay using magnetic particle separation techniques on an automated analyser (Bayer Healthcare AG, Leverkusen, Germany).

**Apolipoprotein A1 (Apo A1)**
This is the major protein component found in high-density lipoprotein. Serum concentrations are negatively associated with liver fibrosis, i.e. levels decrease as the extent of fibrosis increases. Decreased levels are also seen in uncontrolled diabetes, nephrotic syndrome, some diets and smoking. As for α-2-M, the preferred method for analysis of Apo A1 is an immunonephelometry platform such as the Beckman Coulter IMMAGE or Dade-Behring BNII.

**Haptoglobin**
This serum protein binds any free haemoglobin present in the circulation. Hap is an acute phase protein whose concentrations
to be identified, it is useful to consider the properties of an ideal marker:

1. Specific for liver;
2. Readily available and standardised between all laboratories performing diagnostic biochemistry/haematology;
3. Not subject to false positive results, for example due to inflammation;
4. Identifies the stage of fibrosis.

Most current serum markers are not liver specific, or may represent impaired hepatic clearance or are affected by inflammation. In addition, coexisting pathologies such as haemolysis (causing a decrease in haptoglobin (Hap) levels and/or increase in Bil levels) or rheumatoid arthritis (increase in hyaluronic acid (HA) levels) are associated with changes in levels of serum markers.

Although no ideal marker exists, several have been identified as possible useful indicators of fibrosis. Single markers often correlate with fibrosis in large groups of patients but are not sufficiently predictive in the individual patient, especially when used longitudinally over time. A systematic review compared single and multiple markers versus liver biopsy up to 2002 and noted diagnostic accuracy was greatest in studies using multiple markers.13

In practice, serum markers are therefore used in combination where they have achieved a greater likelihood for success in discriminating minimal from significant fibrosis. Usually, three or more markers are used in combination in an algorithm to generate a score, which is then used to give a fibrosis prediction. The serum markers listed below have been chosen because they are common components of published serum models used to make fibrosis predictions. The list is not intended to be complete. A brief rationale is given for the use of each marker followed by comments on the available methods for analysis. Newer approaches such as proteomics, metabolomics and clinical glycomics are expected to yield more novel biomarkers.14

Major Serum Fibrosis Markers

HA
This mucopolysaccharide is a glycosaminoglycan, a high molecular weight polymer present in joints and in some tissues such as liver. It is found in synovial fluid and serum levels are elevated in various chronic liver diseases due to HA production by hepatic stellate cells and decreased clearance by sinusoidal endothelial cells. Serum levels are normally <50 µg/L and elevated levels correlate reasonably well with the degree of liver fibrosis in alcoholic liver disease15 and hepatitis C.16 HA has been used on its own to exclude significant fibrosis15,16 or more recently in combination with other markers.

Testing for serum HA is currently not widely available. It is available commercially as a self contained kit in the 96-well ELISA format (Corgenix, Colorado, US) the most efficient use requires a plate washer to perform the multiple rinses and washes as well as a plate reader to read the final absorbance. If demand increases, it is anticipated that serum HA will become available on commercial automated platforms.

α-2-Macroglobulin (α-2-M)
This is a high molecular weight protein synthesised in hepatocytes and stellate cells which is reasonably abundant in human serum, where normal levels are typically from 0.66 to 2.65 g/L. The functions of α-2-M are not well understood but it does inhibit the catabolism of matrix proteins by acting as a broad-spectrum inhibitor of nearly all enzymes that split proteins internally (endoproteases). Serum levels increase with the degree of liver fibrosis.17

The preferred methods for analysis are immunonephelometry and immunoturbidimetry and reagents are available as commercial kits from manufacturers of immunonephelometry platforms such as the Beckman Coulter IMMAGE and Dade-Behring BNII.

Collagen Markers
This diverse group of markers includes pro-collagen peptides, proteins such as type I, type III and type IV collagen and collagen metabolites such as laminin. For example, the N-terminal propeptide of type III collagen (PIIINP) is a valuable fibrosis marker that has been validated in alcoholic liver disease,18 hepatitis C,7 and non-alcoholic fatty liver disease.7 Serum levels increase with the degree of liver fibrosis. A typical analysis method for PIIINP is by heterogeneous immunooassay using magnetic particle separation techniques on an automated analyser (Bayer Healthcare AG, Leverkusen, Germany).

Apolipoprotein A1 (Apo A1)
This is the major protein component found in high-density lipoprotein. Serum concentrations are negatively associated with liver fibrosis, i.e. levels decrease as the extent of fibrosis increases.19 Decreased levels are also seen in uncontrolled diabetes, nephrotic syndrome, some diets and smoking. As for α-2-M, the preferred method for analysis of Apo A1 is an immunonephelometry platform such as the Beckman Coulter IMMAGE or Dade-Behring BNII.

Haptoglobin
This serum protein binds any free haemoglobin present in the circulation. Hap is an acute phase protein whose concentrations

---

increase in a wide variety of inflammatory conditions and in nephrotic syndrome. Concentrations decrease in in vivo haemolysis whether caused by autoimmune, iso-immune or mechanical reasons. Haptoglobin levels also decrease with increasing stages of fibrosis.\textsuperscript{20} The preferred method for analysis is immunonephelometry.

**Matrix Metalloproteinases (MMPs)**

MMPs and their tissue inhibitors (TIMPs) have been shown to correlate with the development of liver fibrosis, for example circulating MMP 1 concentrations are significantly reduced as fibrosis grades increase in hepatitis C,\textsuperscript{21} whereas TIMP 1 levels increase.

The excess collagen deposition in liver, which is characteristic of fibrosis, is the result of both decreased collagen degradation mediated by increased TIMPs and increased collagen synthesis.

Testing for MMPs and their tissue inhibitors is not currently widely available in laboratories performing diagnostic biochemistry. Commercial kits are in the 96-well ELISA format, however if demand increases it is anticipated that these analytes will become available on commercial automated platforms.

**Constructing an Algorithm-Based Serum Model**

**Prerequisites**

A set of minimum prerequisites for constructing a serum model to predict liver fibrosis can be identified. The first requirement is for a relatively homogeneous set of patients, usually with a single liver disease who are usually treatment naive with respect to antiviral therapy. For example in chronic hepatitis C, those patients with regular high alcohol intakes or co-infection with hepatitis B or human immunodeficiency viruses would typically be excluded.

A second requirement is for a pre-treatment liver biopsy and histologic staging which achieves certain minimum standards, usually for length and number of portal tracts. Further conditions may include biopsy staging conducted by the same pathologist who is blinded to the clinical data.

The third requirement concerns the serum samples, usually obtained at the same time as the liver biopsy. Constructing a serum model implies prior identification of a candidate group of potential serum markers of fibrosis from which the final panel of markers will be selected. For example in the first report describing serum markers used in combination to generate a score which could predict liver fibrosis in hepatitis C patients, a total of 11 candidate markers were assessed, whereas only five markers were used in the final model.\textsuperscript{20}

**Statistical Analyses**

Most models have been developed by following these general rules. Two groups of patients are required, a training set in which all candidate serum markers are measured and a validation set in which the performance of the final model is assessed. Sometimes the two sets are created by random selection from one pool of patients or alternatively the validation set can be entirely separate, for example from another centre.

An essential requirement is to establish the desired fibrosis stage endpoints, normally there is no attempt made to predict individual fibrosis stages, instead a binary ‘presence’ or ‘absence’ is used. For example the simplest variant would be a single endpoint of significant fibrosis defined using the META VIR system as a grade of F2, F3 or F4. Examples of other common endpoints are the presence of META VIR grade of F3 or F4 for defining advanced fibrosis and META VIR grade F4 for cirrhosis.

The predictive model itself is commonly formulated by performing univariate analysis on the candidate serum markers in patients with and without the desired endpoints in the training set. Those markers from the univariate analysis, together with other desired variables such as age at biopsy or gender found to be significant predictors (p<0.05), are then subjected to multivariate analysis by forward logistic regression to identify independent factors associated with either the presence or absence of the desired endpoint. Equations giving a score that could best predict the desired endpoints are constructed by entering different sets of independent variables into the regression model. The diagnostic value of each equation can be assessed by comparing the areas under the ROC. An ideal equation would have an area under the curve of 1.0, whereas 0.5 indicates an equation of no diagnostic value. The equations are typically simplified by constructing a score system, for example from 0.0 to 1.0, and the best cut-off points within that range are selected from the ROC by calculating sensitivity, specificity and positive and negative predictive values.

**Performance Depends on Disease Prevalence**

The utility of serum models for detecting fibrosis is critically dependent on the prevalence of liver fibrosis in the population being investigated. Thus if positive and negative predictive values are quoted for the detection of significant fibrosis using a serum model, these are only applicable at the quoted prevalence of significant fibrosis in that particular population. In addition the particular characteristics of a serum model could make it suitable for a population. Thus if a serum model is being applied to patients where the prevalence of significant fibrosis is expected to be low, for example non-alcoholic fatty liver disease, it is preferable for the model to deliver a high
increase in a wide variety of inflammatory conditions and in nephrotic syndrome. Concentrations decrease in in vivo haemolysis whether caused by autoimmune, iso-immune or mechanical reasons. Haem levels also decrease with increasing stages of fibrosis. The preferred method for analysis is immunonephelometry.

Matrix Metalloproteinases (MMPs)
MMPs and their tissue inhibitors (TIMPs) have been shown to correlate with the development of liver fibrosis, for example circulating MMP 1 concentrations are significantly reduced as fibrosis grades increase in hepatitis C, whereas TIMP 1 levels increase.

The excess collagen deposition in liver, which is characteristic of fibrosis, is the result of both decreased collagen degradation mediated by increased TIMPs and increased collagen synthesis.

Testing for MMPs and their tissue inhibitors is not currently widely available in laboratories performing diagnostic biochemistry. Commercial kits are in the 96-well ELISA format, however if demand increases it is anticipated that these analytes will become available on commercial automated platforms.

Constructing an Algorithm-Based Serum Model
Prerequisites
A set of minimum prerequisites for constructing a serum model to predict liver fibrosis can be identified. The first requirement is for a relatively homogeneous set of patients, usually with a single liver disease who are usually treatment naive with respect to antiviral therapy. For example in chronic hepatitis C, those patients with regular high alcohol intakes or co-infection with hepatitis B or human immunodeficiency viruses would typically be excluded.

A second requirement is for a pre-treatment liver biopsy and histologic staging which achieves certain minimum standards, usually for length and number of portal tracts. Further conditions may include biopsy staging conducted by the same pathologist who is blinded to the clinical data.

The third requirement concerns the serum samples, usually obtained at the same time as the liver biopsy. Constructing a serum model implies prior identification of a candidate group of potential serum markers of fibrosis from which the final panel of markers will be selected. For example in the first report describing serum markers used in combination to generate a score which could predict liver fibrosis in hepatitis C patients, a total of 11 candidate markers were assessed, whereas only five markers were used in the final model.

Statistical Analyses
Most models have been developed by following these general rules. Two groups of patients are required, a training set in which all candidate serum markers are measured and a validation set in which the performance of the final model is assessed. Sometimes the two sets are created by random selection from one pool of patients or alternatively the validation set can be entirely separate, for example from another centre.

An essential requirement is to establish the desired fibrosis stage endpoints, normally there is no attempt made to predict individual fibrosis stages, instead a binary ‘presence’ or ‘absence’ is used. For example the simplest variant would be a single endpoint of significant fibrosis defined using the META VIR system as a grade of F2, F3 or F4. Examples of other common endpoints are the presence of META VIR grade of F3 or F4 for defining advanced fibrosis and META VIR grade F4 for cirrhosis.

The predictive model itself is commonly formulated by performing univariate analysis on the candidate serum markers in patients with and without the desired endpoints in the training set. Those markers from the univariate analysis, together with other desired variables such as age at biopsy or gender found to be significant predictors (p<0.05), are then subjected to multivariate analysis by forward logistic regression to identify independent factors associated with either the presence or absence of the desired endpoint. Equations giving a score that could best predict the desired endpoints are constructed by entering different sets of independent variables into the regression model. The diagnostic value of each equation can be assessed by comparing the areas under the ROC. An ideal equation would have an area under the curve of 1.0, whereas 0.5 indicates an equation of no diagnostic value. The equations are typically simplified by constructing a score system, for example from 0.0 to 1.0, and the best cut-off points within that range are selected from the ROC by calculating sensitivity, specificity and positive and negative predictive values.

Performance Depends on Disease Prevalence
The utility of serum models for detecting fibrosis is critically dependent on the prevalence of liver fibrosis in the population being investigated. Thus if positive and negative predictive values are quoted for the detection of significant fibrosis using a serum model, these are only applicable at the quoted prevalence of significant fibrosis in that particular population. In addition the particular characteristics of a serum model could make it suitable for a population. Thus if a serum model is being applied to patients where the prevalence of significant fibrosis is expected to be low, for example non-alcoholic fatty liver disease, it is preferable for the model to deliver a high...
negative predictive value to allow the maximum number of patients to avoid liver biopsy. An advantage of the European Liver Fibrosis Group (ELFG) model is that the adoption of different score thresholds delivered changes favouring either negative or positive predictive values, depending on the population being studied.7

Rapid Proliferation of Serum Models
The first report describing serum markers used in combination to generate a score, which could predict liver fibrosis in hepatitis C patients, appeared in Lancet in 2001 and has been widely quoted.20 In the intervening time, there has been an explosion of interest in the area such that it is unusual to read an issue of any specialist hepatology journal, which does not describe a new serum model. However what is lacking is good data comparing serum models with each other. The problem is compounded by the commercialisation of some of the models, with the result that if the all-important algorithm used to calculate the score is not published, comparative studies are not possible.

FibroTest
FibroTest was first described for hepatitis C patients in 2001,20 and is licensed to BioPredictive (www.biopredictive.com). FibroTest uses five serum markers, Apo A1, Hap, α-2-M, gamma glutamyl transpeptidase (γGT) activity and Bil, together with the age and gender of the patient to calculate a score. In the original report, FibroTest scores from 0 to 0.10 provided 100% negative predictive value for the absence of significant fibrosis (defined as F2, F3 or F4 by Metavir) while scores from 0.60 to 1.00 had a >90% positive predictive value for significant fibrosis for hepatitis C patients. Scores from 0.11 to 0.59 were indeterminate and liver biopsy was recommended. In an independent validation of FibroTest, the negative predictive value of a score <0.10 was 85% and the positive predictive value of a score >0.60 was 78%.22

FibroTest has also been applied to detect liver fibrosis in patients with chronic hepatitis B infection.23 For application in non-alcoholic fatty liver disease FibroTest has been modified and presented as NashTest by including the following additional parameters: height, weight, serum triglycerides, cholesterol, and both AST and ALT.24

Fibrospect
FIBROspect II was first described for hepatitis C patients in 2004 and is licensed by Prometheus Laboratories (California, US).25 Fibrospect uses three serum markers, α-2-M, HA and TIMP, to calculate a score. When applied to 696 patients with hepatitis C, a score ≤0.36 excluded significant fibrosis with a negative predictive value of 76% and a score >0.36 detected significant fibrosis with a positive predictive value of 74%.

ELFG
In a thorough international multicentre study, the ELFG developed an algorithm combining age and three serum markers: HA, PIIINP and TIMP.1 In the same paper, the algorithm was applied to three chronic liver diseases; hepatitis C, alcoholic liver disease and non-alcoholic fatty liver disease where it achieved areas under ROC of 0.77, 0.94 and 0.87, respectively. When histologic grading obtained by three pathologists was compared, the agreement between pathologists ranged from very good to moderate (kappa scores 0.97-0.46).7

Hepascore
Hepascore requires the measurement of serum Bil, γGT activity, α-2-M and HA levels.26 Hepascore is a score from 0.00 to 1.00 calculated from the results of these four analyses and the age and sex of the patient. Hepascore has been validated in hepatitis C patients, where a score ≥0.50 provided a positive predictive value of 88% for significant fibrosis (META VIR score of F2 or above) and a score <0.5 had a negative predictive value of 95% for the absence of advanced fibrosis (META VIR score of F3 or above).26

Fibrometer
In a thorough study Cales et al. measured a total of 51 serum markers and were able to calculate and compare five previously described serum models, including FibroTest, Fibrospect II and the European Liver Fibrosis models.27 In addition they proposed Fibrometer, a new serum model that is claimed to outperform previous models. The six tests required to calculate Fibrometer are platelets, PT index, AST, α-2-M, HA and urea.

Assessing Serum Model Performance
Serum models are assessed against the prevailing liver biopsy gold standard, although it is a flawed standard. In a new approach, Poynard et al. assessed the risk factors for discordances between the FibroTest serum marker model and biopsy, and then classified them as attributable to either biopsy or marker failure.28 Discordance was attributable to failure of serum markers in 2.4%, to biopsy failure in 18% and was not attributable in a further 8% of patients. The most frequent reason for marker failure was a false negative result due to inflammation affecting the serum results, whereas biopsy failures were usually due to false negative staging associated with smaller biopsy size, fragmented biopsy and steatosis. In a similar study, discordance was attributable to failure of FibroTest serum marker model in 5%, to biopsy failure in 4% and was not attributable in a further 9% of patients.29

There is five years experience with the first reported serum marker model and the five year prognostic value of the serum
negative predictive value to allow the maximum number of patients to avoid liver biopsy. An advantage of the European Liver Fibrosis Group (ELFG) model is that the adoption of different score thresholds delivered changes favouring either negative or positive predictive values, depending on the population being studied.7

Rapid Proliferation of Serum Models
The first report describing serum markers used in combination to generate a score, which could predict liver fibrosis in hepatitis C patients, appeared in Lancet in 2001 and has been widely quoted.20 In the intervening time, there has been an explosion of interest in the area such that it is unusual to read an issue of any specialist hepatology journal, which does not describe a new serum model. However what is lacking is good data comparing serum models with each other. The problem is compounded by the commercialisation of some of the models, with the result that if the all-important algorithm used to calculate the score is not published, comparative studies are not possible.

FibroTest
FibroTest was first described for hepatitis C patients in 2001,20 and is licensed to BioPredictive (www.biopredictive.com). FibroTest uses five serum markers, Apo A1, Hap, α-2-M, gamma glutamyl transpeptidase (γGT) activity and Bil, together with the age and gender of the patient to calculate a score. In the original report, FibroTest scores from 0 to 0.10 provided 100% negative predictive value for the absence of significant fibrosis (defined as F2, F3 or F4 by Metavir) while scores from 0.60 to 1.00 had a >90% positive predictive value for significant fibrosis for hepatitis C patients. Scores from 0.11 to 0.59 were indeterminate and liver biopsy was recommended. In an independent validation of FibroTest, the negative predictive value of a score <0.10 was 85% and the positive predictive value of a score >0.60 was 78%.22

FibroTest has also been applied to detect liver fibrosis in patients with chronic hepatitis B infection.23 For application in non-alcoholic fatty liver disease FibroTest has been modified and presented as NashTest by including the following additional parameters: height, weight, serum triglycerides, cholesterol, and both AST and ALT.24

Fibrospect
FIBROSpect II was first described for hepatitis C patients in 2004 and is licensed by Prometheus Laboratories (California, US).25 Fibrospect uses three serum markers, α-2-M, HA and TIMP, to calculate a score. When applied to 696 patients with hepatitis C, a score <0.36 excluded significant fibrosis with a negative predictive value of 76% and a score >0.36 detected significant fibrosis with a positive predictive value of 74%.

ELFG
In a thorough international multicentre study, the ELFG developed an algorithm combining age and three serum markers: HA, PIIINP and TIMP 1.7 In the same paper, the algorithm was applied to three chronic liver diseases; hepatitis C, alcoholic liver disease and non-alcoholic fatty liver disease where it achieved areas under ROCC of 0.77, 0.94 and 0.87, respectively. When histologic grading obtained by three pathologists was compared, the agreement between pathologists ranged from very good to moderate (kappa scores 0.97-0.46).7

Hepascore
Hepascore requires the measurement of serum Bil, γGT activity, α-2-M and HA levels.26 Hepascore is a score from 0.00 to 1.00 calculated from the results of these four analyses and the age and sex of the patient. Hepascore has been validated in hepatitis C patients, where a score ≥0.50 provided a positive predictive value of 88% for significant fibrosis (META VIR score of F2 or above) and a score <0.5 had a negative predictive value of 95% for the absence of advanced fibrosis (META VIR score of F3 or above).26

Fibrometer
In a thorough study Cales et al. measured a total of 51 serum markers and were able to calculate and compare five previously described serum models, including FibroTest, Fibrospect II and the European Liver Fibrosis models.27 In addition they proposed Fibrometer, a new serum model that is claimed to outperform previous models. The six tests required to calculate Fibrometer are platelets, PT index, AST, α-2-M, HA and urea.

Assessing Serum Model Performance
Serum models are assessed against the prevailing liver biopsy gold standard, although it is a flawed standard. In a new approach, Poynard et al. assessed the risk factors for discordances between the FibroTest serum marker model and biopsy, and then classified them as attributable to either biopsy or marker failure.28 Discordance was attributable to failure of serum markers in 2.4%, to biopsy failure in 18% and was not attributable in a further 8% of patients. The most frequent reason for marker failure was a false negative result due to inflammation affecting the serum results, whereas biopsy failures were usually due to false negative staging associated with smaller biopsy size, fragmented biopsy and steatosis. In a similar study, discordance was attributable to failure of FibroTest serum marker model in 5%, to biopsy failure in 4% and was not attributable in a further 9% of patients.29

There is five years experience with the first reported serum marker model and the five year prognostic value of the serum
Combining non-invasive methods for assessment of liver fibrosis

Combinaison des méthodes non-invasives pour l’évaluation de la fibrose hépatique

M. Bourlière, G. Pénaranda, X. Adhouté, V. Oules, P. Castellani

Department of Hepato-gastroenterology Hôpital Saint Joseph, 26 boulevard de Louvain, 13008 Marseille, France
Department of Biostatistics, CDL Pharma, Marseille, France

Summary

Non-invasive markers of liver fibrosis, including biochemical scores using simple biochemical parameters and transient elastography (TE), have been developed over the past decade to either replace or reduce the need for liver biopsy (LB) in the assessment of liver fibrosis. Although their diagnostic accuracy in liver fibrosis is promising, around 20% of patients are likely to be misclassified if these tests or LB are used alone. However, using a combination of several biochemical scores (Fibropaca algorithm, Leroy algorithm) or one biochemical score with TE (Bordeaux algorithm) will increase diagnostic accuracy for fibrosis and reduce the need for LB. Stepwise combination algorithms of non-invasive scores (SAFE biopsy) also improve the diagnostic performance in chronic hepatitis C (CHC) compared with the use of a single non-invasive score. Other sequential stepwise algorithms have been developed in CHC with similar performance results. Comparisons of different combinations of non-invasive methods indicate that the SAFE biopsy, Fibropaca algorithm and Bordeaux algorithm are excellent and comparable in the non-invasive diagnosis of liver fibrosis in HCV patients, and will markedly reduce the need for LB. They may also be useful in clinical practice and particularly for large-scale screening of HCV patients.

Résumé

Depuis une dizaine d’année plusieurs marqueurs non invasifs de fibrose utilisant des paramètres biochimiques simples ou l’élastométrie impulsionelle (EI) ont été développés pour la détermination de la fibrose hépatique afin soit de remplacer ou de diminuer le besoin d’une biopsie hépatique. Bien que les performances diagnostique de ces méthodes soient bonnes, près de 20% des patients seront mal classés si ces tests ou la biopsie...
Two studies have addressed the issue of discordance between TE and FT compared with LB. More false-negative results were observed with TE than with FT, and it appeared that TE more often underestimated, whereas FT more frequently overestimated, the LB results. In addition, the combination of TE and FT may be of value in HCV-infected patients with normal transaminases (ALT) and in chronic hepatitis B inactive carriers. In the latter population, the combination of TE and FT allowed the exclusion of a ≥ F2 diagnosis in nearly 80%.

Combination of non-invasive biochemical scores

In a prospective study of 235 HCV patients, without co-morbidities, who had undergone LB, FT, APRI and the Forns index on the same day, the diagnostic performance for ≥F2 and F4 did not differ among these tests. However, all patients were classified using FT, 214 (91%) patients were classified using APRI and 129 (55%) using the Forns index. There were significantly more cases of discordance between APRI and LB than between FT or the Forns index and LB (p<0.05). Moreover, the combination of all scores with LB allowed 225/235 (96%) patients to be correctly classified. Interestingly, the combination of FT, APRI and the Forns index (Fibropanca algorithm) allowed 191/235 (81.3%) patients to be correctly classified due to a concordance between FT, APRI and/or the Forns index (Fig. 2). However, in this case, 18 patients (8%) were discordant between the serum tests and LB. This means that, with this combination, LB remains mandatory for only a few patients (18.7%) [29].

Leroy et al. [19], in a prospective study of 180 CHC patients, compared the diagnostic performance of six non-invasive scores (MP3, FT, FM, HS, the Forns index and APRI) vs LB in the diagnosis of liver fibrosis. The overall diagnostic performance of the tests, as determined by AUROCs for the diagnosis of ≥F2, were comparable among themselves and ranged from 0.86 for FM to 0.78 for the Forns index. For a diagnosis of ≥F3, FM had the best performance, but its superiority over the other scores was statistically significant only compared with the Forns index.

Using logistic regression, the statistical independence of scores was assessed to construct a logical algorithm for combinations. Statistical independence was demonstrated for MP3, FT and APRI, and was explained by the fact that these scores do not share any biological parameters. When the diagnostic performance of paired-combination scores was evaluated, it was interesting to note that combinations such as MP3 + APRI, FT + APRI and MP3 + FT produced better results than when used as single tests. The best combination (FT + APRI) selected one-third of patients for whom either the absence of ≥F2 or presence of extensive fibrosis could be predicted with more than 90% certainty without LB. A clinical-management algorithm using the combination of FT and APRI was derived from these results (Fig. 3).

Stepwise combination algorithms of non-invasive scores

Sebastiani et al. [38] attempted to increase the diagnostic performance of non-invasive tests of liver fibrosis by combining them in sequential algorithms. They evaluated the diagnostic performance of FT, APRI and the Forns index in 190 CHC patients at the same time as LB. The accuracy of the three scores ranged from 57.1% to 81.2% for the diagnosis of ≥F2 in patients with elevated ALT, from 54.9% to 86.3% for the diagnosis of ≥F2 in patients with persistent normal alanine aminotransferase (PNALT) and from 78.1% to 85.9% for the diagnosis of F4. Overall, FT had the best performance scores for all parameters considered, and also when combined in stepwise algorithms to achieve optimized accuracy (> 90%) while minimizing the number of necessary LB.

These algorithms were developed in the first 190 patients and validated prospectively in 100 further patients. Three different algorithms were developed (Fig. 4) for the diagnosis of ≥F2 in patients with elevated or normal ALT, and for the diagnosis of F4. Significant fibrosis in patients with elevated ALT was identified with an accuracy range of 94–94.2%, using APRI as the screening test, followed by FT in APRI non-classified cases and restricting LB to patients classified as F0–F1 by non-invasive tests. This algorithm avoided biopsy in around 50% of cases with no instances of underestimations and a 0% rate of overestimations.

Using a slightly modified algorithm in patients with normal ALT, ≥F2 was identified with an accuracy ranging from 94–95% using APRI as the screening test, followed by
Two studies have addressed the issue of discordance between TE and FT compared with LB. More false-negative results were observed with TE than with FT, and it appeared that TE more often underestimated, whereas FT more frequently overestimated, the LB results [22]. In addition, the combination of TE and FT [34, 35], or other serum markers [36], may be of value in HCV-infected patients with normal transaminases (ALT) and in chronic hepatitis B inactive carriers [37]. In the latter population, the combination of TE and FT allowed the exclusion of a ≥ F2 diagnosis in nearly 80%.

Combination of non-invasive biochemical scores

In a prospective study of 235 HCV patients, without co-morbidities, who had undergone LB, FT, APRI and the Forns index on the same day, the diagnostic performance for ≥F2 and F4 did not differ among these tests [29]. However, all patients were classified using FT, 214 (91%) patients were classified using APRI and 129 (55%) using the Forns index. There were significantly more cases of discordance between APRI and LB than between FT or the Forns index and LB (p<0.05). Moreover, the combination of all scores with LB allowed 225/235 (96%) patients to be correctly classified. Interestingly, the combination of FT, APRI and the Forns index (Fibropanca algorithm) allowed 191/235 (81.3%) patients to be correctly classified due to a concordance between FT, APRI and/or the Forns index (Fig. 2). However, in this case, 18 patients (8%) were discordant between the serum tests and LB. This means that, with this combination, LB remains mandatory for only a few patients (18.7%) [29].

Leroy et al. [19], in a prospective study of 180 CHC patients, compared the diagnostic performance of six non-invasive scores (MP3, FT, FM, HS, the Forns index and APRI) vs LB in the diagnosis of liver fibrosis. The overall diagnostic performance of the tests, as determined by AUROCs for the diagnosis of ≥F2, were comparable among themselves and ranged from 0.86 for FM to 0.78 for the Forns index. For a diagnosis of ≥ F3, FM had the best performance, but its superiority over the other scores was statistically significant only compared with the Forns index.

Using logistic regression, the statistical independence of scores was assessed to construct a logical algorithm for combinations. Statistical independence was demonstrated for MP3, FT and APRI, and was explained by the fact that these scores do not share any biological parameters. When the diagnostic performance of paired-combination scores was evaluated, it was interesting to note that combinations such as MP3 + APRI, FT + APRI and MP3 + FT produced better results than when used as single tests. The best combination (FT + APRI) selected one-third of patients for whom either the absence of F2 or presence of extensive fibrosis could be predicted with more than 90% certainty without LB. A clinical-management algorithm using the combination of FT and APRI was derived from these results (Fig. 3).

Stepwise combination algorithms of non-invasive scores

Sebastiani et al. [38] attempted to increase the diagnostic performance of non-invasive tests of liver fibrosis by combining them in sequential algorithms. They evaluated the diagnostic performance of FT, APRI and the Forns index in 190 CHC patients at the same time as LB. The accuracy of the three scores ranged from 57.1% to 81.2% for the diagnosis of ≥F2 in patients with elevated ALT, from 54.9% to 86.3% for the diagnosis of ≥ F2 in patients with persistent normal alanine aminotransferase (PNALT) and from 78.1% to 85.9% for the diagnosis of F4. Overall, FT had the best performance scores for all parameters considered, and also when combined in stepwise algorithms to achieve optimized accuracy (> 90%) while minimizing the number of necessary LB.

These algorithms were developed in the first 190 patients and validated prospectively in 100 further patients. Three different algorithms were developed (Fig. 4) for the diagnosis of ≥ F2 in patients with elevated or normal ALT, and for the diagnosis of F4. Significant fibrosis in patients with elevated ALT was identified with an accuracy range of 94–94.2%, using APRI as the screening test, followed by FT in APRI non-classified cases and restricting LB to patients classified as F0–F1 by non-invasive tests. This algorithm avoided biopsy in around 50% of cases with no instances of underestimations and a 5% rate of overestimations.

Using a slightly modified algorithm in patients with normal ALT, ≥F2 was identified with an accuracy ranging from 94–95% using APRI as the screening test, followed by
Two studies have addressed the issue of discordance between TE and FT compared with LB. More false-negative results were observed with TE than with FT, and it appeared that TE more often underestimated, whereas FT more frequently overestimated, the LB results [22]. In addition, the combination of TE and FT [34, 35], or other serum markers [36], may be of value in HCV-infected patients with normal transaminases (ALT) and in chronic hepatitis B inactive carriers [37]. In the latter population, the combination of TE and FT allowed the exclusion of a ≥ F2 diagnosis in nearly 80%.

**Combination of non-invasive biochemical scores**

In a prospective study of 235 HCV patients, without co-morbidities, who had undergone LB, FT, APRI and the Forns index on the same day, the diagnostic performance for ≥F2 and F4 did not differ among these tests [29]. However, all patients were classified using FT, 214 (91%) patients were classified using APRI and 129 (55%) using the Forns index. There were significantly more cases of discordance between APRI and LB than between FT or the Forns index and LB (p<0.05). Moreover, the combination of all scores with LB allowed 225/235 (96%) patients to be correctly classified. Interestingly, the combination of FT, APRI and the Forns index (Fibropaca algorithm) allowed 191/235 (81.3%) patients to be correctly classified due to a concordance between FT, APRI and/or the Forns index (Fig. 2). However, in this case, 18 patients (8%) were discordant between the serum tests and LB. This means that, with this combination, LB remains mandatory for only a few patients (18.7%) [29].

Leroy et al. [19], in a prospective study of 180 CHC patients, compared the diagnostic performance of six non-invasive scores (MP3, FT, FM, HS, the Forns index and APRI) vs LB in the diagnosis of liver fibrosis. The overall diagnostic performance of the tests, as determined by AUROCs for the diagnosis of ≥F2, were comparable among themselves and ranged from 0.86 for FM to 0.78 for the Forns index. For a diagnosis of ≥ F3, FM had the best performance, but its superiority over the other scores was statistically significant only compared with the Forns index.

Using logistic regression, the statistical independence of scores was assessed to construct a logical algorithm for combinations. Statistical independence was demonstrated for MP3, FT and APRI, and was explained by the fact that these scores do not share any biological parameters. When the diagnostic performance of paired-combination scores was evaluated, it was interesting to note that combinations such as MP3 + APRI, FT + APRI and MP3 + FT produced better results than when used as single tests. The best combination (FT + APRI) selected one-third of patients for whom either the absence of ≥ F2 or presence of extensive fibrosis could be predicted with more than 90% certainty without LB. A clinical-management algorithm using the combination of FT and APRI was derived from these results (Fig. 3).

**Stepwise combination algorithms of non-invasive scores**

Sebastiani et al. [38] attempted to increase the diagnostic performance of non-invasive tests of liver fibrosis by combining them in sequential algorithms. They evaluated the diagnostic performance of FT, APRI and the Forns index in 190 CHC patients at the same time as LB. The accuracy of the three scores ranged from 57.1% to 81.2% for the diagnosis of ≥F2 in patients with elevated ALT, from 54.9% to 86.3% for the diagnosis of ≥ F2 in patients with persistent normal alanine aminotransferase (PNALT) and from 78.1% to 85.9% for the diagnosis of F4. Overall, FT had the best performance scores for all parameters considered, and also when combined in stepwise algorithms to achieve optimized accuracy (> 90%) while minimizing the number of necessary LB.

These algorithms were developed in the first 190 patients and validated prospectively in 100 further patients. Three different algorithms were developed (Fig. 4) for the diagnosis of ≥ F2 in patients with elevated or normal ALT, and for the diagnosis of F4. Significant fibrosis in patients with elevated ALT was identified with an accuracy range of 94–94.2%, using APRI as the screening test, followed by FT in APRI non-classified cases and restricting LB to patients classified as F0-F1 by non-invasive tests. This algorithm avoided biopsy in around 50% of cases with no instances of underestimations and a 5% rate of overestimations.

Using a slightly modified algorithm in patients with normal ALT, ≥ F2 was identified with an accuracy ranging from 94–95% using APRI as the screening test, followed by
Combining non-invasive methods for assessment of liver fibrosis

**Figure 4** Sequential algorithms for detection of fibrosis (SAFE biopsy), from Sebastiani et al. [58].

(A) Algorithm for detection of ≥F2 in CHC patients with elevated ALT: F0/F1 corresponds to an APRI score < 0.5; ≥F2 corresponds to an APRI score ≥ 1.5; and N.C corresponds to an APRI score 0.5 and < 1.5. (B) Algorithm for detection of ≥F2 in CHC patients with PNALT. (C) Algorithm for detection of F4 in CHC patients. FT: FibroTest; LB: Liver Biopsy; CHC: Chronic Hepatitis C; PNALT: Persistent Normal Alanine Aminotransferase.

**Algorithm séquentiels pour le diagnostic de la fibrose hépatique (Biopsie SAFE) selon Sebastiani et al [58].** (A) Algorithme pour la détection de F ≥ 2 au cours de l’hépatite C avec transaminases élevées. F0/F1 correspond à un score APRI < 0.5; F ≥ 2 correspond à un score APRI ≥ 1.5; et N.C correspond à un score APRI compris entre 0.5 et 1.49. (B) Algorithme pour la détection de F ≥ 2 au cours de l’hépatite C avec transaminases normales. (C) Algorithme pour la détection de la cirrhose au cours de l’hépatite C. FT : FibroTest ; LB : Biopsie Hépatique ; CHC : Hépatite chronique C ; PNALT : Taux sérique de l’alanine aminotransférase normal.

**Figure 3** Algorithm for detection of fibrosis in CHC patients with elevated transaminases, from Leroy et al. [19].

Algorithm pour le diagnostic de la fibrose hépatique au cours de l’hépatite C avec élévation des transaminases selon Leroy et al. [19].

Identification of F4 was achieved with an accuracy range of 93-95% using APRI as the screening test, followed by FT in APRI non-classified and ≥F2 cases, and restricting LB to patients classified as F2-F3 by FT. This algorithm avoided biopsies in 60-70% of the cases with a 1% rate of underdiagnosed cases and 3-6% overdiagnosed cases.

Sequential algorithms for fibrosis evaluation (SAFE biopsy) improved the diagnostic performance in CHC compared with a single non-invasive score. The need for LB was reduced by 50-70%, but could not be completely avoided. Recently, a large-scale validation of the SAFE biopsy was carried out in 2035 CHC patients [39]. Subgroup analyses were conducted to determine whether age, gender, HCV genotype and BMI could modify its performance. Overall, SAFE biopsy for ≥F2 avoided 47.1% of LB with >95% accuracy. SAFE biopsy for F4 also performed excellently and waived 82% of LB. The performance of SAFE biopsy was excellent across all major HCV genotypes, although it was reduced for ≥F2 in patients aged >50 years and, for F4, in patients with BMI >30kg/m². For these reasons, specific cutoffs for FT were identified that allowed optimization of SAFE biopsies. This large-scale, multicenter, validation study confirmed that SAFE biopsy is an effective way of combining non-invasive markers with a markedly reduced need for LB to correctly classify liver fibrosis in CHC.

Recently, we assessed the diagnostic performances of FT, HS, APRI and the Forns index in 467 CHC patients [20]. We confirmed that both FT and HS have similar profiles for the diagnosis of ≥F2 or F4 with misclassification rates of around 20%. In addition, to increase diagnostic accuracy for ≥F2, to decrease underestimations of fibrosis and to reduce the need for LB, we developed five sequential stepwise algorithms in a population of 293 patients and validated them in 179 patients (Fig. 5). These algorithms were more accurate (> 90%) than either FT, HS, the Forns index or APRI alone (p<0.001). Algorithm 1 is similar to those reported in the original study, with an accuracy rate of 90% and no biopsy needed in 44% of cases. The algorithm with the highest rate of avoided biopsies (45%) was algorithm 2, using APRI as the screening test, followed by HS in FT in APRI F0- F1 and non-classified cases, and restricting LB to patients classified as F0-F1 by FT. Again, this algorithm avoided biopsy in around 50% of cases, with no underestimations and a 5% rate of overestimations. APRI non-classified cases and restricting LB to patients classified as F0- F1 in non-invasive tests. Algorithms 4 and 5 had the highest overall accuracy rates (94% in the validation set), but only avoided biopsy in around 30% of cases.
1. What are the Public Health Implications of Hepatitis C?
Hepatitis C is a major health problem. The global prevalence of chronic hepatitis C is estimated to average 3% (ranging from 0.1 to 5% in different countries): there are some 150 million chronic HCV carriers throughout the world, of whom an estimated 4 million are in the USA and 5 million in Western Europe. The prevalence seems to be higher in Eastern Europe than in Western Europe. In industrialized countries, HCV accounts for 20% of cases of acute hepatitis, 70% of cases of chronic hepatitis, 40% of cases of end-stage cirrhosis, 60% of cases of hepatocellular carcinoma and 30% of liver transplants.

The incidence of new symptomatic infections has been estimated to be 1–3 cases/100,000 persons annually. The actual incidence of new infections is obviously much higher (the majority of cases being asymptomatic). The incidence is declining for two reasons: (a) transmission by blood products has been reduced to near zero; (b) universal precautions have markedly reduced transmission in medical settings. Intravenous drug use remains the main mode of transmission; but, even here, the rate of transmission is diminishing due to a heightened awareness of the risk of needle sharing and, in some countries, the availability of needle-exchange programs.

2. What is the Natural History of Hepatitis C?
What are the Factors Influencing the Disease?
Hepatitis C is a disease with various rates of progression. In general, its course is slowly progressive. About 15% of HCV-infected individuals recover spontaneously; an additional 25% have an asymptomatic illness with persistently normal aminotransferases and generally benign histological lesions; hence, about 40% of patients recover or have a benign outcome. In those with biochemical evidence of chronic hepatitis, the majority have only mild to moderate necro-inflammatory lesions and minimal fibrosis: their long-term outcome is unknown and, probably, most of them will not succumb to the liver disease. About 20% of patients with chronic hepatitis C develop cirrhosis in 10–20 years, and may die of complications of cirrhosis in the absence of liver transplantation. Thus, hepatitis C is a dichotomous disease in which a subset of patients will die from liver-related causes, but in which the majority will probably live out their normal life span.

Several cofactors play an important role in the development of cirrhosis: (a) the age at the time of infection (on average, patients who acquire the disease at an older age have a more rapidly progressing disease, while progression is slower in younger patients; (b) alcoholism (all studies show that alcohol is a very important co-factor in the progression of chronic hepatitis to cirrhosis); (c) co-infection with HIV; (d) co-infection with hepatitis B virus.

The incidence of hepatocellular carcinoma is 1–4% per year in patients with cirrhosis. This risk supports the necessity of regular monitoring by ultrasonography and measurement of alphafetoprotein in patients with established or suspected cirrhosis. Development of hepatocellular carcinoma is rare in patients with chronic hepatitis C who do not have cirrhosis.

3. Diagnostic Tests
ELISA tests are easy to use and inexpensive, and are the best tests for initial screening. These tests are reliable in most immunocompetent patients who replicate HCV. They are less sensitive in hemodialyzed and in immunocompromised patients.

In low-risk settings, such as blood banks and other...
1. What are the Public Health Implications of Hepatitis C?

Hepatitis C is a major health problem. The global prevalence of chronic hepatitis C is estimated to average 3% (ranging from 0.1 to 5% in different countries): there are some 150 million chronic HCV carriers throughout the world, of whom an estimated 4 million are in the USA and 5 million in Western Europe. The prevalence seems to be higher in Eastern Europe than in Western Europe. In industrialized countries, HCV accounts for 20% of cases of acute hepatitis, 70% of cases of chronic hepatitis, 40% of cases of end-stage cirrhosis, 60% of cases of hepatocellular carcinoma and 30% of liver transplants.

The incidence of new symptomatic infections has been estimated to be 1–3 cases/100,000 persons annually. The actual incidence of new infections is obviously much higher (the majority of cases being asymptomatic). The incidence is declining for two reasons: (a) transmission by blood products has been reduced to near zero; (b) universal precautions have markedly reduced transmission in medical settings. Intravenous drug use remains the main mode of transmission; but, even here, the rate of transmission is diminishing due to a heightened awareness of the risk of needle sharing and, in some countries, the availability of needle-exchange programs.

2. What is the Natural History of Hepatitis C? What are the Factors Influencing the Disease?

Hepatitis C is a disease with various rates of progression. In general, its course is slowly progressive. About 15% of HCV-infected individuals recover spontaneously; an additional 25% have an asymptomatic illness with persistently normal aminotransferases and generally benign histological lesions; hence, about 40% of patients recover or have a benign outcome. In those with biochemical evidence of chronic hepatitis, the majority have only mild to moderate necro-inflammatory lesions and minimal fibrosis: their long-term outcome is unknown and, probably, most of them will not succumb to the liver disease. About 20% of patients with chronic hepatitis C develop cirrhosis in 10–20 years and may die of complications of cirrhosis in the absence of liver transplantation. Thus, hepatitis C is a dichotomous disease in which a subset of patients will die from liver-related causes, but in which the majority will probably live out their normal life span.

Several cofactors play an important role in the development of cirrhosis: (a) the age at the time of infection (on average, patients who acquire the disease at an older age have a more rapidly progressing disease, while progression is slower in younger patients); (b) alcoholism (all studies show that alcohol is a very important co-factor in the progression of chronic hepatitis to cirrhosis); (c) co-infection with HIV; (d) co-infection with hepatitis B virus.

The incidence of hepatocellular carcinoma is 1–4% per year in patients with cirrhosis. This risk supports the necessity of regular monitoring by ultrasonography and measurement of alphafetoprotein in patients with established or suspected cirrhosis. Development of hepatocellular carcinoma is rare in patients with chronic hepatitis C who do not have cirrhosis.

3. Diagnostic Tests

ELISA tests are easy to use and inexpensive, and are the best tests for initial screening. These tests are reliable in most immunocompetent patients who replicate HCV. They are less sensitive in hemodialyzed and in immunocompromised patients.

In low-risk settings, such as blood banks and other
1. What are the Public Health Implications of Hepatitis C?
Hepatitis C is a major health problem. The global prevalence of chronic hepatitis C is estimated to average 3% (ranging from 0.1 to 5% in different countries): there are some 150 million chronic HCV carriers throughout the world, of whom an estimated 4 million are in the USA and 5 million in Western Europe. The prevalence seems to be higher in Eastern Europe than in Western Europe. In industrialized countries, HCV accounts for 20% of cases of acute hepatitis, 70% of cases of chronic hepatitis, 40% of cases of end-stage cirrhosis, 60% of cases of hepatocellular carcinoma and 30% of liver transplants.

The incidence of new symptomatic infections has been estimated to be 1–3 cases/100 000 persons annually. The actual incidence of new infections is obviously much higher (the majority of cases being asymptomatic). The incidence is declining for two reasons: (a) transmission by blood products has been reduced to near zero; (b) universal precautions have markedly reduced transmission in medical settings. Intravenous drug use remains the main mode of transmission; but, even here, the rate of transmission is diminishing due to a heightened awareness of the risk of needle-sharing and, in some countries, the availability of needle-exchange programs.

2. What is the Natural History of Hepatitis C? What are the Factors Influencing the Disease?
Hepatitis C is a disease with various rates of progression. In general, its course is slowly progressive. About 15% of HCV-infected individuals recover spontaneously; an additional 25% have an asymptomatic illness with persistently normal aminotransferases and generally benign histological lesions; hence, about 40% of patients recover or have a benign outcome. In those with biochemical evidence of chronic hepatitis, the majority have only mild to moderate necro-inflammatory lesions and minimal fibrosis: their long-term outcome is unknown and, probably, most of them will not succumb to the liver disease. About 20% of patients with chronic hepatitis C develop cirrhosis in 10–20 years, and may die of complications of cirrhosis in the absence of liver transplantation. Thus, hepatitis C is a dichotomous disease in which a subset of patients will die from liver-related causes, but in which the majority will probably live out their normal life span.

Several cofactors play an important role in the development of cirrhosis: (a) the age at the time of infection (on average, patients who acquire the disease at an older age have a more rapidly progressing disease, while progression is slower in younger patients; (b) alcoholism (all studies show that alcohol is a very important co-factor in the progression of chronic hepatitis to cirrhosis); (c) co-infection with HIV; (d) co-infection with hepatitis B virus.

The incidence of hepatocellular carcinoma is 1–4% per year in patients with cirrhosis. This risk supports the necessity of regular monitoring by ultrasonography and measurement of alphafetoprotein in patients with established or suspected cirrhosis. Development of hepatocellular carcinoma is rare in patients with chronic hepatitis C who do not have cirrhosis.

3. Diagnostic Tests
ELISA tests are easy to use and inexpensive, and are the best tests for initial screening. These tests are reliable in most immunocompetent patients who replicate HCV. They are less sensitive in hemodialyzed and in immunocompromised patients.

In low-risk settings, such as blood banks and other
1. What are the Public Health Implications of Hepatitis C?
Hepatitis C is a major health problem. The global prevalence of chronic hepatitis C is estimated to average 3% (ranging from 0.1 to 5% in different countries): there are some 150 million chronic HCV carriers throughout the world, of whom an estimated 4 million are in the USA and 5 million in Western Europe. The prevalence seems to be higher in Eastern Europe than in Western Europe. In industrialized countries, HCV accounts for 20% of cases of acute hepatitis, 70% of cases of chronic hepatitis, 40% of cases of end-stage cirrhosis, 60% of cases of hepatocellular carcinoma and 30% of liver transplants.

The incidence of new symptomatic infections has been estimated to be 1–3 cases/100000 persons annually. The actual incidence of new infections is obviously much higher (the majority of cases being asymptomatic). The incidence is declining for two reasons: (a) transmission by blood products has been reduced to near zero; (b) universal precautions have markedly reduced transmission in medical settings. Intravenous drug use remains the main mode of transmission; but, even here, the rate of transmission is diminishing due to a heightened awareness of the risk of needle sharing and, in some countries, the availability of needle-exchange programs.

2. What is the Natural History of Hepatitis C? What are the Factors Influencing the Disease?
Hepatitis C is a disease with various rates of progression. In general, its course is slowly progressive. About 15% of HCV-infected individuals recover spontaneously; an additional 25% have an asymptomatic illness with persistently normal aminotransferases and generally benign histological lesions; hence, about 40% of patients recover or have a benign outcome. In those with biochemical evidence of chronic hepatitis, the majority have mild to moderate necro-inflammatory lesions and minimal fibrosis: their long-term outcome is unknown and, probably, most of them will not succumb to the liver disease. About 20% of patients with chronic hepatitis C develop cirrhosis in 10–20 years and may die of complications of cirrhosis in the absence of liver transplantation. Thus, hepatitis C is a dichotomous disease in which a subset of patients will die from liver-related causes, but in which the majority will probably live out their normal life span.

Several cofactors play an important role in the development of cirrhosis: (a) the age at the time of infection (on average, patients who acquire the disease at an older age have a more rapidly progressing disease, while progression is slower in younger patients; (b) alcoholism (all studies show that alcohol is a very important co-factor in the progression of chronic hepatitis to cirrhosis); (c) co-infection with HIV; (d) co-infection with hepatitis B virus.

The incidence of hepatocellular carcinoma is 1–4% per year in patients with cirrhosis. This risk supports the necessity of regular monitoring by ultrasonography and measurement of alphafetoprotein in patients with established or suspected cirrhosis. Development of hepatocellular carcinoma is rare in patients with chronic hepatitis C who do not have cirrhosis.

3. Diagnostic Tests
ELISA tests are easy to use and inexpensive, and are the best tests for initial screening. These tests are reliable in most immunocompetent patients who replicate HCV. They are less sensitive in hemodialyzed and in immunocompromised patients.

In low-risk settings, such as blood banks and other
general screening situations where approximately 25% of ELISA positive results may be false, a supplemental specificity test, such as a strip immunoblot assay, is recommended to avoid unwarranted notification of false positives. Then, a qualitative HCV RNA test should be performed if anti-HCV positivity is confirmed.

In high-risk populations and in clinical settings where hepatitis C is suspected, a positive ELISA should be confirmed by a qualitative HCV RNA test.

In patients with acute hepatitis of unknown cause, an ELISA test should be performed first. If hepatitis A and B tests are negative, then a qualitative HCV RNA test must be performed.

In ELISA-negative patients with chronic hepatitis of unknown cause, particularly in hemodialyzed and immunocompromised patients, a qualitative HCV RNA test should be performed.

Genotyping and quantitative HCV RNA tests are only recommended prior to the treatment of patients.

4. Who Should be Screened for Hepatitis C?
General screening is not advisable. Screening should be limited to risk groups: (a) persons who have (or might have) received blood products prior to initiation (1991) of second-generation ELISA test; (b) hemophiliacs; (c) hemodialyzed patients; (d) children born to mothers who have hepatitis C; (e) current or previous users of intravenous drugs; (f) donors for organ or tissue transplantation.

5. How Can the Transmission of Hepatitis C be Prevented?
The two main sources of infection are intravenous drug use and administration of blood products. The latter source has almost completely disappeared since 1991.

Sexual transmission is very uncommon: the prevalence of HCV infection in stable partners of homosexual or heterosexual individuals infected with HCV is very low, but is higher in persons with multiple partners. The use of condoms in stable monogamous relationships is not justified; the use of condoms is strongly encouraged in patients with multiple partners.

Pregnancy is not contraindicated in HCV-infected women. Routine HCV screening is not recommended in pregnant women.

HCV vertical transmission is uncommon: the prevalence of transmission from mother to child is less than 6%. The risk of transmission appears to be greater in women with high levels of viremia or HIV co-infection. The mode of delivery (cesarean section/vaginal) does not appear to influence the rate of HCV transmission from mother to child.

There is no association between breast feeding and transmission of HCV infection from mother to child.

There are insufficient data concerning the risk of vertical transmission of in vitro fertilization in patients with hepatitis C to make recommendations at this time.

Nosocomial HCV infection is efficiently prevented by the observance of universal precautions.

6. Which Patients Should be Treated?
The decision to treat is a complex issue which must take into consideration numerous variables: age of the patients, general state of health, risk of cirrhosis, likelihood of response, and other medical conditions that may decrease life expectancy or contraindicate the use of interferon or ribavirin.

Does the decision to treat depend on the histologic lesions?
It is appropriate and important to obtain a percutaneous liver biopsy before beginning therapy. The liver biopsy provides an opportunity to grade the severity of necro-inflammation and to stage the progression of fibrosis, which may then be considered in relation to the supposed duration of the disease, clinical status and biochemical abnormalities to make therapeutic decisions. The biopsy also provides a baseline in individual patients. There is agreement that patients with moderate/severe necro-inflammation and/or fibrosis should be treated.

Does the decision to treat depend on the age of the patient?
The physiological age of the patient is more important than the chronological age of the patient. Factors to be considered in older patients include overall health status with a special assessment of the cardiovascular system to determine the potential risk of a decrease in hemoglobin level if treatment with ribavirin is being considered.

Does the decision to treat depend on the clinical manifestations?
In the early stages, in the absence of advanced cirrhosis, there is a poor correlation between the clinical manifestations and the histological lesions of the disease. Overall, clinical status may affect the decision to treat with regard to quality of life. Studies have shown the abatement of symptoms in patients in whom treatment has induced sustained loss of HCV RNA.
The hepatitis C virus (HCV) is one of the leading known causes of liver disease in the United States. It is a common cause of cirrhosis and hepatocellular carcinoma (HCC) as well as the most common reason for liver transplantation. At least 4 million people in this country are believed to have been infected with HCV. Following the identification of hepatitis A and hepatitis B, this disorder was categorized in 1974 as “non-A, non-B hepatitis.” In 1989, the hepatitis C virus was identified and found to account for the majority of those patients with non-A, non-B hepatitis. In March 1997, the National Institutes of Health (NIH) held a Consensus Development Conference regarding management and treatment of HCV infection. This led to an important, widely distributed NIH Consensus Statement that, for several years, defined the standard of care. Now 5 years later, knowledge of hepatitis C has increased dramatically, leading to the need to reexamine the approaches to management and treatment. This conference was convened with the aim of reviewing the most recent developments regarding management, treatment options, and the widening spectrum of potential candidates for treatment and of updating the 1997 Consensus Statement.

This NIH Consensus Development Conference on Management of Hepatitis C: 2002 was held June 10-12, 2002. The primary sponsors of this meeting were the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the Office of Medical Applications of Research (OMAR) of the NIH. The cosponsors were the National Institute of Child Health and Human Development (NICHD); the National Cancer Institute (NCI); the National Center for Complementary and Alternative Medicine (NCCAM); the National Institute on Drug Abuse (NIDA); the National Institute of Allergy and Infectious Diseases (NIAID); the National Heart, Lung, and Blood Institute (NHLBI); the Centers for Medicare & Medicaid Services (CMS); the Centers for Disease Control and Prevention (CDC); the U.S. Food and Drug Administration (FDA); and the U.S. Department of Veterans Affairs (VA).

The Agency for Healthcare Research and Quality (AHRQ) provided support to the NIH Consensus Development Conference on Management of Hepatitis C: 2002 through its Evidence-based Practice Center program. Under contract to the AHRQ, the Johns Hopkins University Evidence-based Practice Center developed the systematic review and analysis that served as a reference for discussion at the Conference.

This two-and-a-half-day conference examined the current state of knowledge regarding the management of hepatitis C and identified directions for future research. During the first day-and-a-half of the conference, experts presented the latest hepatitis C research findings to an independent non-Federal Consensus Development Panel. After weighing this scientific evidence, the panel drafted a statement, addressing the following key questions:

- What is the natural history of hepatitis C?
- What is the most appropriate approach to diagnose and monitor patients?
- What is the most effective therapy for hepatitis C?
- Which patients with hepatitis C should be treated?
- What recommendations can be made to patients to prevent transmission of hepatitis C?
- What are the most important areas for future research?

On the final day of the conference, the panel chairperson read the draft statement to the conference audience.
dle-aged transfused subjects. The actual risk is likely intermediate between these two ranges, on the order of 10 to 15 percent. There is little evidence that virologic factors, including viral load, viral genotype, and quasispecies diversity significantly affect the risk of progression of liver disease. However, many host factors increase this risk, including older age at time of infection, male gender, and an immunosuppressed state such as that associated with human immunodeficiency virus (HIV) infection. Concurrent chronic hepatitis B also appears to increase the risk of progressive liver disease. In addition, higher levels of alcohol use play an important role in promoting the development of progressive liver disease, with strong evidence for the detrimental effects of 30 g/day in men (~equivalent to 2 beers, 2 glasses of wine, or 2 mixed drinks) and 20 g/day in women. Lower amounts of alcohol also may increase the risk of liver damage associated with HCV. Other factors, including iron overload, nonalcoholic fatty liver disease, schistosomal co-infection, potentially hepatotoxic medications, and environmental contaminants, also may have important effects.

In the United States, deaths associated with chronic hepatitis C are currently more likely to be due to decompensated cirrhosis than to HCC. Data from death certificates suggest that there are 10,000 to 12,000 deaths yearly in the United States due to hepatitis C, but these may be underestimates. The only treatment option for persons who have developed decompensated cirrhosis is liver transplantation. Currently, HCV is the primary reason for liver transplantation in the United States. Little is known about the clinical course and risks of HCV-related complications in persons who have been infected for longer than two decades.

HCV accounts for an estimated one-third of HCC cases in the United States. HCC rarely occurs in the absence of cirrhosis or advanced fibrosis. The incidence of HCV-related HCC continues to rise in United States and worldwide, in part because of the increasing numbers of persons who have been chronically infected for decades, the presence of comorbid factors, and the longer survival of persons with advanced liver disease due to improved management of complications. Risk factors for HCC in persons with chronic HCV infection are largely the same as those for the development of decompensated cirrhosis. Some but not all studies suggest that treatment with interferon and ribavirin may reduce the risk of developing HCC in HCV patients with cirrhosis, but more data are needed.

**Extrahepatic Manifestations of HCV Infection**

Patients with chronic hepatitis C can present with extrahepatic manifestations or syndromes considered to be of immunologic origin, such as rheumatoid symptoms, keratoconjunctivitis sicca, lichen planus, glomerulonephritis, lymphoma, and essential mixed cryoglobulinemia. Cryoglobulins have been detected in the serum of up to one-half of patients with chronic hepatitis C, but the clinical features of mixed cryoglobulinemia are uncommon. Chronic hepatitis C is also related to porphyria cutanea tarda. Psychological disorders including depression have been associated with HCV infection in up to 20 to 30 percent of cases.

2. **What is the most appropriate approach to diagnose HCV infection and monitor infected patients?**

Various tests are available for the diagnosis and monitoring of HCV infection. Tests that detect antibodies against the virus include the enzyme immunoassay (EIA), which contains HCV antigens from the core and non-structural genes, and the recombinant immunoblot assay. The same HCV antigens are used in both EIAs and the immunoblot assays. Target amplification techniques using either polymerase chain reaction (PCR) or transcription-mediated amplification (TMA) have been developed as qualitative tests for HCV RNA, whereas both target amplification (PCR) and signal amplification techniques (branched DNA) may be used to measure HCV RNA levels. Liver biopsy can provide direct histologic assessment of liver injury due to HCV but cannot be used to diagnose HCV infection.

**HCV Serologic Assays**

EIA tests are reproducible, inexpensive, and FDA-approved for use in the diagnosis of HCV infection. They are suitable for screening at-risk populations and are recommended as the initial test for patients with clinical liver disease. The very high sensitivity and specificity of the version 3 (third-generation) EIAs (sensitivity of greater than 99 percent, specificity of 99 percent in immunocompetent patients) obviate the need for a confirmatory immunoblot assay in the diagnosis of individual patients with clinical liver disease, particularly those with risk factors for HCV infection. A negative EIA test is sufficient to exclude a diagnosis of chronic HCV infection in immunocompetent patients. Rarely, patients on hemodialysis and patients with immune deficiencies may have false-negative EIAs. Conversely, false-positive EIAs may occur in patients with autoimmune disorders. In these patients, an assay for HCV RNA is necessary for diagnosis of chronic infection. The immunoblot assay is still useful as a supplemental assay for persons screened in nonclinical settings and in persons with a positive EIA who test negative for HCV RNA.
dilated blood vessels, with a prevalence of approximately 20 percent in the United States. The risk of liver damage associated with HCV is influenced by factors such as the duration of infection, the presence of comorbid factors, and the longer survival of persons who have been infected for decades.

In the United States, deaths associated with chronic hepatitis C are currently more likely to be due to decompensated cirrhosis than to HCC. Data from death certificates suggest that there are 10,000 to 12,000 deaths yearly in the United States due to hepatitis C, but these may be underestimates. The only treatment option for persons who have developed decompensated cirrhosis is liver transplantation. Currently, HCV is the primary reason for liver transplantation in the United States. Little is known about the clinical course and risks of HCV-related complications in persons who have been infected for longer than two decades.

HCV accounts for an estimated one-third of HCC cases in the United States. HCC rarely occurs in the absence of cirrhosis or advanced fibrosis. The incidence of HCV-related HCC continues to rise in United States and worldwide, in part because of the increasing numbers of persons who have been chronically infected for decades, the presence of comorbid factors, and the longer survival of persons with advanced liver disease due to improved management of complications. Risk factors for HCC in persons with chronic HCV infection are largely the same as those for the development of decompensated cirrhosis. Some but not all studies suggest that treatment with interferon and ribavirin may reduce the risk of developing HCC in HCV patients with cirrhosis, but more data are needed.

Extrahepatic Manifestations of HCV Infection

Patients with chronic hepatitis C can present with extrahepatic manifestations or syndromes considered to be potential hepatotoxic medications, and environmental contaminants, also may have important effects.

In the United States, deaths associated with chronic hepatitis C are currently more likely to be due to decompensated cirrhosis than to HCC. Data from death certificates suggest that there are 10,000 to 12,000 deaths yearly in the United States due to hepatitis C, but these may be underestimates. The only treatment option for persons who have developed decompensated cirrhosis is liver transplantation. Currently, HCV is the primary reason for liver transplantation in the United States. Little is known about the clinical course and risks of HCV-related complications in persons who have been infected for longer than two decades.

HCV accounts for an estimated one-third of HCC cases in the United States. HCC rarely occurs in the absence of cirrhosis or advanced fibrosis. The incidence of HCV-related HCC continues to rise in United States and worldwide, in part because of the increasing numbers of persons who have been chronically infected for decades, the presence of comorbid factors, and the longer survival of persons with advanced liver disease due to improved management of complications. Risk factors for HCC in persons with chronic HCV infection are largely the same as those for the development of decompensated cirrhosis. Some but not all studies suggest that treatment with interferon and ribavirin may reduce the risk of developing HCC in HCV patients with cirrhosis, but more data are needed.

Extrahepatic Manifestations of HCV Infection

Patients with chronic hepatitis C can present with extrahepatic manifestations or syndromes considered to be of immunologic origin, such as rheumatoid symptoms, keratoconjunctivitis sicca, lichen planus, glomerulonephritis, lymphoma, and essential mixed cryoglobulinemia. Cryoglobulins have been detected in the serum of up to one-half of patients with chronic hepatitis C, but the clinical features of mixed cryoglobulinemia are uncommon. Chronic hepatitis C is also related to porphyria cutanea tarda. Psychological disorders including depression have been associated with HCV infection in up to 20 to 30 percent of cases.

2. What is the most appropriate approach to diagnose HCV infection and monitor infected patients?

Various tests are available for the diagnosis and monitoring of HCV infection. Tests that detect antibodies against the virus include the enzyme immunoassay (EIA), which contains HCV antigens from the core and non-structural genes, and the recombinant immunoblot assay. The same HCV antigens are used in both EIAs and the immunoblot assays. Target amplification techniques using either polymerase chain reaction (PCR) or transcription-mediated amplification (TMA) have been developed as qualitative tests for HCV RNA, whereas both target amplification (PCR) and signal amplification techniques (branched DNA) may be used to measure HCV RNA levels. Liver biopsy can provide direct histologic assessment of liver injury due to HCV but cannot be used to diagnose HCV infection.

HCV Serologic Assays

EIA tests are reproducible, inexpensive, and FDA-approved for use in the diagnosis of HCV infection. They are suitable for screening at-risk populations and are recommended as the initial test for patients with clinical liver disease. The very high sensitivity and specificity of the version 3 (third-generation) EIAs (sensitivity of greater than 99 percent, specificity of 99 percent in immunocompetent patients) obviate the need for a confirmatory immunoblot assay in the diagnosis of individual patients with clinical liver disease, particularly those with risk factors for HCV infection. A negative EIA test is sufficient to exclude a diagnosis of chronic HCV infection in immunocompetent patients. Rarely, patients on hemodialysis and patients with immune deficiencies may have false-negative EIAs. Conversely, false-positive EIAs may occur in patients with autoimmune disorders. In these patients, an assay for HCV RNA is necessary for diagnosis of chronic infection. The immunoblot assay is still useful as a supplemental assay for persons screened in nonclinical settings and in persons with a positive EIA who test negative for HCV RNA.
Qualitative HCV RNA Assays

Acute or chronic HCV infection in a patient with a positive EIA test should be confirmed by a qualitative HCV RNA assay with a lower limit of detection of 50 IU/mL or less (approximately 100 viral genes/mL). However, confirmation may be unnecessary in a patient who has evidence of liver disease and obvious risk factors for HCV infection. The FDA-approved manual and semiautomated, qualitative, HCV PCR assays have a lower limit of detection of 50 to 100 IU/mL. More recently, a transcription-mediated amplification assay has been developed with a lower limit of detection on the order of 5 to 10 IU/mL, but it has yet to be approved for use by the FDA. The specificity of these assays for detecting HCV RNA exceeds 98 percent. A single positive qualitative assay for HCV RNA confirms active HCV replication, but a single negative does not exclude viremia and may reflect only a transient decline in viral level below the level of detection of the assay. A followup qualitative HCV RNA should be performed to confirm the absence of active HCV replication. Once HCV infection is confirmed, repeat testing using a qualitative assay with a limit of detection of 50 IU/mL or less is not helpful in the management of untreated patients, except for determining whether an acute infection has resolved spontaneously. Until future studies determine whether the sustained virological response (SVR) will be sustained over the long term following successful antiviral treatment, periodic measurements of HCV RNA may need to be performed.

Quantitative HCV RNA Assays

Testing for HCV RNA level (or viral load) with a quantitative assay such as quantitative PCR (qPCR) or branched DNA (bDNA) signal amplification assay provides accurate information on HCV viral levels. An HCV RNA standard has been introduced to permit normalization of reported viral titers in IU/mL. The reported IU does not represent the actual number of viral particles in a preparation. Significant variability exists between available assays. The reportable range, accuracy, and precision of each assay needs to be monitored, and appropriate dilutions of sample material should be performed to obtain accurate quantitative results. The clinical utility of serial HCV viral levels in a patient is predicated on continued use of the same specific quantitative assay that was used in the initial determination of the viral level. While there is little correlation between disease severity or disease progression with the absolute level of HCV RNA, quantitative determination of the HCV level provides important information on the likelihood of response to treatment in patients undergoing antiviral therapy. In clinical trials of combination interferon and ribavirin reported to date, a positive response to antiviral therapy in patients infected with all common genotypes (genotypes 1, 2, and 3) has been correlated with low viral levels.

ALT

Testing for serum ALT levels is the most inexpensive and noninvasive, but relatively insensitive, means of assessing disease activity. A single determination of ALT level gives limited information about the severity of the underlying liver disease. In most studies, a weak association exists between the degree of ALT elevation and severity of the histopathological findings on liver biopsy. Serial determinations of ALT levels over time may provide a better means of assessing liver injury, but the accuracy of this approach has not been well documented. Patients who initially have a normal ALT level should undergo serial measurements over several months to confirm the persistence of normal ALT levels. Although loss or reduction in HCV RNA is the primary indicator of response to antiviral therapy, the resolution of elevated ALT levels with antiviral therapy appears to be an important indicator of disease response. Nevertheless, pegylated interferon can cause mild elevations of ALT during therapy, and ALT levels are insensitive in detecting disease progression to cirrhosis.

Noninvasive Tests of Fibrosis

Various noninvasive tests of hepatic fibrosis have been examined for monitoring patients with chronic HCV infection. These include routinely available laboratory tests, such as liver-associated chemistries, platelet count, and prothrombin time, as well as specific serum markers of fibrosis and inflammation not currently widely available or well validated. No single test or panel of serologic markers can provide an accurate assessment of intermediate stages of hepatic fibrosis. Similarly, quantitative tests of liver function and radiologic imaging of the liver are sensitive for diagnosing advanced cirrhosis but are not useful in assessing hepatic fibrosis and early cirrhosis.

Liver Biopsy

Liver biopsy provides a unique source of information on fibrosis and assessment of histology. Liver enzymes have shown little value in predicting fibrosis. Extracellular matrix tests can predict severe stages of fibrosis but cannot consistently classify intermediate stages of fibrosis. Moreover, only liver biopsy provides information on possible contributions of iron, steatosis, and concurrent alcoholic liver disease to the progression of chronic hepatitis C toward cirrhosis. Although unexpected etiologies of liver disease are rarely discovered on liver biopsies from patients undergoing evaluation of chronic hepatitis C, the
Prospective comparison of six non-invasive scores for the diagnosis of liver fibrosis in chronic hepatitis C

Vincent Leroy1,2,*, Marie-Noëlle Hilleret1, Nathalie Sturm3, Candice Trocmé4, Jean-Charles Renversez5, Patrice Faure5, Françoise Morel4, Jean-Pierre Zarski1,2

1Département d’Hépato-gastroentérologie, CHU de Grenoble, BP 217, 38043 Grenoble cedex 9, France
2Laboratoire d’Immunochimie, CEA, GIRDRC, INSERM U 548, Grenoble, France
3Département d’Anatomie Pathologique et Cytologique, CHU de Grenoble, France
4Laboratoire d’Enzymologie, DBPC, GREPI EA 2938, CHU de Grenoble, France
5Laboratoire de Biochimie, DBI, CHU de Grenoble, France

See Editorial, pages 751–755

Background/Aims: Non-invasive markers of liver fibrosis have recently been developed as an alternative to liver biopsy. The aim of this study was to compare the diagnostic performance of 6 scores (MP3, Fibrotest, Fibrometer, Hepascore, Forns’ score and APRI).

Methods: We studied 180 chronic hepatitis C patients. Liver fibrosis was staged according to the METAVIR scoring system.

Results: Overall diagnostic performance of scores determined by AUROCs ranged from 0.86 for Fibrometer to 0.78 for Forns’ score (NS) for discriminating F0F1 versus F2F3F4. For discriminating F0F1F2 versus F3F4, AUROCs ranged from 0.91 for Fibrometer to 0.78 for Forns’ score (p < 0.02). Significant or extensive fibrosis was predicted in 10–86% of patients with positive predictive value (PPV) ranging from 55% to 94%. Using logistic regression, statistical independence was demonstrated for MP3, Fibrotest and APRI. Diagnostic performance of paired-combination scores was then evaluated. The best combinations could select one-third of patients for whom either absence of significant fibrosis or presence of extensive fibrosis could be predicted with more than 90% of certainty.

Conclusions: Current non-invasive scores give reliable information on liver fibrosis in one-third of chronic hepatitis C patients, especially when used in combination.

Keywords: PIIINP; Metalloproteinase; Fibrotest; Fibrometer; APRI; METAIVIR; Fibrosis

1. Introduction

Clinical management of chronic hepatitis C is dependent on the extent of liver fibrosis. Liver biopsy, the gold standard, is still recommended in the majority of patients [1]. However, it is an invasive procedure responsible for severe complications in about 0.5% of cases [2]. Sample variability is another limitation. The biopsy specimen appears to be poorly reliable when its length is inferior to 15 mm [3]. Moreover, liver fibrosis is evaluated by histological scores, which have inter-observer variability especially among non-expert pathologists...
(r = 0.70) and MP3 (r = 0.69) (p < 0.001). Fig. 1 shows the box-plots of fibrosis scores according to METAVIR fibrosis stages. Weaker correlations were also found between scores and histological activity, especially for APRI (r = 0.56), MP3 (r = 0.49) and Fibrometer (r = 0.47) (p < 0.001). In each case, correlation with fibrosis persisted after statistical adjustment of activity. There was no statistical interaction between biopsy length and correlation coefficients.

3.3. Overall diagnostic performance of serum markers

Areas under ROC curves were used for evaluating the overall diagnostic performance of scores. For discriminating F0F1 versus F2F3F4, AUROCs ranged from 0.86 for Fibrometer to 0.78 for Forns’ score (Fig. 2a). Fibrometer had a better AUROC than Forns’ score (p < 0.02) but the significance disappeared after the Dunn-Sidak correction due to multiple comparisons. For discriminating F0F1F2 versus F3F4, AUROCs were better and ranged from 0.91 for Fibrometer to 0.78 for Forns’ score. The only significant difference was observed between Fibrometer and Forns’ score (p < 0.02). The difference was nearly significant between Fibrometer and APRI (p = 0.07). Details about AUROCs and 95% confidence intervals are given in Table 2. AUROCs were not altered by separating patients according to genotype 1 (versus non 1), length of biopsies <25 mm versus ≥25 mm, and presence or absence of sinusoidal fibrosis, steatosis or intra-hepatic iron load (data not shown). A sensitivity analysis excluding the 19 patients with biopsies smaller than 15 mm or with less than 7 portal tracts was also performed, showing very similar results (Table 2).

![Fig. 1. Score values according to METAVIR fibrosis stages. The top and bottom of each box are the 25th and the 75th centiles. The line through the box is the median, and the error bars are the 5th and 95th centiles. Spearman correlation coefficients were as follows: FT, r = 0.70; MP3, r = 0.69; FM, r = 0.71; HS, r = 0.60; FS, r = 0.55; APRI, r = 0.59, p < 0.001 for each correlation test.](image-url)
in 31/93 (33.3%) of cases compared to 11/87 (12.6%) of patients who had FT values predicting either F0 or F4 (p < 0.01).

3.6. Combination of scores

Multiple stepwise logistic regression analysis was performed to test the statistical independence of scores. In logistic regression including the six scores, MP3 (p < 0.001) and APRI (p < 0.05) were the only variables independently associated to significant and extensive fibrosis. When MP3 was removed from the analysis, Fibrotest (p < 0.001) and APRI (p < 0.02) were the remaining variables associated to fibrosis. In a model including MP3 and Fibrotest, both variables were independently associated to fibrosis. Thus, we analysed the diagnostic performance of paired combination of independent scores. As expected, simultaneous use of two scores improved NPV and PPV while slightly decreasing the number of selected patients. Diagnostic values of the best combinations associating MP3 and APRI and Fibrotest and APRI are shown in Table 4. For example, the concomitant presence of Fibrotest >0.59 and APRI >2 improved the PPV for significant fibrosis to 96.7% and for extensive fibrosis to 92.2%. By contrast, the concomitant presence of Fibrotest <0.22 and APRI <0.5, observed in 13.2% of patients, could rule out significant fibrosis with NPV of 94.1%.

Fig. 3. Relationships between the percentage of patients selected according to different cut-offs and the positive predictive value for the diagnosis of extensive (F3F4) fibrosis.

Table 4

<table>
<thead>
<tr>
<th>Score</th>
<th>Cut-off</th>
<th>%</th>
<th>Significant fibrosis (F2F3F4)</th>
<th>Extensive fibrosis (F3F4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sen Spe PPV NPV</td>
<td>Sen Spe PPV NPV</td>
</tr>
<tr>
<td>MP3</td>
<td>&gt;0.20</td>
<td>86.0</td>
<td>95.6 23.9 56.5 84.0</td>
<td>100 19.5 33.1 100</td>
</tr>
<tr>
<td></td>
<td>&gt;0.30</td>
<td>55.3</td>
<td>82.4 72.7 75.8 80.0</td>
<td>92.2 59.4 47.5 95</td>
</tr>
<tr>
<td></td>
<td>&gt;0.40</td>
<td>24.6</td>
<td>44.0 95.5 90.9 62.2</td>
<td>60.8 89.8 70.5 85.2</td>
</tr>
<tr>
<td></td>
<td>&gt;0.50</td>
<td>10.1</td>
<td>18.7 98.9 94.4 54.0</td>
<td>31.4 98.4 88.9 78.3</td>
</tr>
<tr>
<td>APRI</td>
<td>&gt;0.50</td>
<td>82.4</td>
<td>91.6 26.8 55.9 75.9</td>
<td>93.5 21.8 31.6 89.7</td>
</tr>
<tr>
<td></td>
<td>&gt;1.00</td>
<td>58.2</td>
<td>79.5 63.4 68.8 75.4</td>
<td>89.1 53.8 42.7 92.8</td>
</tr>
<tr>
<td></td>
<td>&gt;1.50</td>
<td>42.4</td>
<td>72.3 87.8 85.7 75.8</td>
<td>87.0 74.8 57.1 93.7</td>
</tr>
<tr>
<td></td>
<td>&gt;2.00</td>
<td>32.1</td>
<td>57.8 93.9 90.6 68.8</td>
<td>73.9 84.0 64.2 91.6</td>
</tr>
<tr>
<td>Forns</td>
<td>&gt;4.20</td>
<td>73.1</td>
<td>88.4 42.4 60.8 78.3</td>
<td>91.7 34.1 35.2 91.3</td>
</tr>
<tr>
<td></td>
<td>&gt;6.90</td>
<td>24.6</td>
<td>41.9 92.9 85.7 61.2</td>
<td>54.2 87.0 61.9 82.9</td>
</tr>
<tr>
<td>Fibrotest</td>
<td>&gt;0.22</td>
<td>67.3</td>
<td>89.0 52.8 65.9 82.5</td>
<td>94.1 41.9 39.0 94.7</td>
</tr>
<tr>
<td></td>
<td>&gt;0.32</td>
<td>51.1</td>
<td>75.8 74.2 75.0 75.0</td>
<td>90.2 64.3 50.0 94.3</td>
</tr>
<tr>
<td></td>
<td>&gt;0.59</td>
<td>27.8</td>
<td>45.1 89.9 82.0 61.5</td>
<td>66.7 87.6 68.0 86.9</td>
</tr>
<tr>
<td>Hepascore</td>
<td>&gt;0.50</td>
<td>35.4</td>
<td>53.8 83.9 77.8 63.5</td>
<td>76.5 81.1 61.9 89.6</td>
</tr>
<tr>
<td></td>
<td>&gt;0.84</td>
<td>20.8</td>
<td>33.0 92.0 81.1 56.7</td>
<td>47.1 89.8 64.9 80.9</td>
</tr>
</tbody>
</table>

4. Discussion

Several non-invasive tests combining biological parameters have recently been developed with the objective of replacing liver biopsy [7–16]. We assessed here the diagnostic performance of Fibrotest, Fibrometer, Hepascore, MP3, Forns’ score and APRI. Other scores using routine laboratory parameters such as AST/ALT ratio were not studied for they were recently found to be of less value than APRI [18]. The first finding of our study is an independent validation of scores, with overall diagnostic performances very similar to those originally reported. To our knowledge, it is the first independent validation for Fibrometer and Hepascore. The Fibrometer had the best diagnostic performance but its superiority over other scores was significant only for Forns’ score. We thus did not confirm the recently suggested superiority of Fibrometer over Fibrotest [15]. It should be pointed out that in this study the superiority of Fibrometer was observed only in the training group of
patients. In two studies, the Fibrotest was found to have a greater diagnostic performance than APRI and Forns’ score, but study groups came again from the original population that served for Fibrotest calculation [19,20]. Altogether, these results show that Fibrometer, Fibrotest, MP3, APRI, Forns’ score and Hepascore have overall diagnostic performances close to each other.

The most important issue is to know whether such accuracy is sufficient to waive liver biopsy. For each score, cut-offs have been defined in order to discriminate patients according to two relevant threshold of fibrosis. The first is METAVIR F0F1 versus F2F3F4 because patients with no or mild fibrosis usually do not need to receive antiviral treatment. The second is F0F1F2 versus F3F4 because patients with severe fibrosis or cirrhosis need to be treated and to be screened for portal hypertension, liver failure and hepato-cellular carcinoma. Applying the different cut-offs originally described, NPV and PPV for the diagnosis of fibrosis were variable amongst the scores. This can be explained by the fact that some cut-offs were not designed to discriminate either significant or extensive fibrosis. For example, the cut-offs proposed in Forns’ study were designed to detect significant and not extensive fibrosis [12]. Interestingly, there was a clear correlation between the number of selected patients for a given cut-off and predictive values. These data support the concept that scores have close diagnostic performance and that stringency of cut-offs mainly indicate sensitivity, specificity and predictive values. The use of higher cut-offs can select one-third of patients for whom the probability of significant fibrosis is 80–90% and the probability of extensive fibrosis is 60–70%. To the opposite, the use of lower cut-offs can select 20% of patients for whom the probability of no/mild fibrosis is 80%. Although these predictive values are promising, we should keep in mind that 1 in 5 patients would be misclassified if liver biopsy was not performed.

Another limitation of this strategy is that nearly 50% of patients have intermediate values and cannot be classified. Fibrotest overrides this limitation and uses several cut-offs that allow a conversion to the METAVIR scoring system [21]. However, we observed discordances of at least two fibrosis stages between Fibrotest and biopsy in 23% of cases, a result consistent with that reported in two recent studies [22,23]. The only parameter associated to discordance was an intermediate Fibrotest value, especially indicating F2. Similar data were recently reported for the Fibrometer, which has less diagnostic accuracy when its value suggests METAVIR F2 [15]. This could be explained by the fact that fibrosis area in F2 stage is close to that observed in F1 in morphometric studies [3]. Moreover, it is obvious that F2 determination is the only stage allowing both underestimation and overestimation of two stages. Contrary to Poynard et al. [22], we did not find steatosis associated to discordance. We cannot exclude that in some cases discordance between serum markers and histology could be attributable to biopsy examination failure due to the heterogeneity of fibrosis in the liver. In a laparoscopic study, discordancess in fibrosis stage were reported in 33% of patients when left and right liver lobes were compared [24]. However, 98% of discordances were of only one stage according to the Sheuer classification. The main factor of biopsy failure appears to be an inadequate size of biopsy samples. Colloredo et al. [25] suggested that biopsies smaller than 20 mm long and 1.4 mm wide and containing less than 11 portal tracts led to an underestimation of fibrosis. In an elegant study comparing virtual biopsies to whole liver resections, Bedossa et al. [3] showed that the rate of correct METAVIR fibrosis staging ranged from 65% when biopsy measured 15 mm long to 90% when biopsy length was of 40 mm. The authors concluded that a biopsy length of at least 25 mm would be necessary to evaluate fibrosis accurately. A major strength of our study is that biopsies were of greater quality than that published in previous studies. In the study by Halfon et al. [23], discordancess were attributed to liver biopsy in 22% of cases but reasons were not fully given. In our study, the size of biopsies was not associated to discordance, which appears to be attributable to scores’ failure rather than biopsy failure. This is strengthened by the fact that diagnostic performance is improved when using scores in combination.

An original aspect of our study was to test the statistical independence of scores, in order to propose a logical algorithm. Interestingly, some combinations (including MP3 + APRI, Fibrotest + APRI and MP3 + Fibrotest) gave greater results than single scores. The statistical independence can be explained by the fact that these scores do not share the same biological parameters. Different results were reported by Calès...
et al. [15] who also performed a multivariate analysis and found the Fibrometer to be the only independent score associated to fibrosis. However, this analysis was again done on the training group of patients and results were not given after removing Fibrometer from the analysis. The same concept of combining two non-invasive methods of liver fibrosis was proposed by Castera et al. [26]. Concordance between Fibrotest and transient elastography (Fibroscan) led to a more precise evaluation of liver fibrosis compared to each method used alone. In this study, combination of Fibrotest and APRI did not perform better than Fibrotest used alone but authors did not give any detail about the way they combined both scores. Interestingly, Sebastiani et al. [27] recently proposed an algorithm based on sequential utilization of APRI followed by Fibrotest. Although this algorithm appears to be complex, it highlights the concept of combining non-invasive score of fibrosis to increase their diagnostic accuracy. Our results strongly suggest that the diagnosis of extensive fibrosis (F3F4) can be made with more than 90% certainty when Fibrotest is greater than 0.59 and APRI is greater than 2, such scenario being met in 20% of cases. By contrast, significant fibrosis (F2F3F4) can be ruled out with the same certainty when Fibrotest is lower than 0.22 and APRI is lower than 0.5, this situation being observed in 13.2% of patients. In conclusion, we believe that in these two situations a clinical decision can be made without the need of a liver biopsy and we propose here a simple algorithm (Fig. 4). For patients with intermediate values, important but only partial information on liver fibrosis can be obtained by scores. Clinical consequences of misclassification should always be taken into consideration before deciding to perform a liver biopsy, greater than 25 mm, which still remains the gold standard of liver fibrosis evaluation.

References

Other non-invasive markers of liver fibrosis

Autres marqueurs non-invasifs de fibrose hépatique

V. Leroy

1 Clinique d’Hépato-gastroentérologie, pôle DIGI-DUNE, CHU de Grenoble
2 INSERM/UJF U823, IAPC, IAB, Grenoble

Summary

An intensive research effort in the field of non-invasive evaluation of liver fibrosis has recently permitted the description and validation of several serum markers of fibrosis, mainly in chronic hepatitis C patients. In addition to the commonly used tests such as FibroTest or FibroMeters, other either indirect (aspartate aminotransferase, prothrombin time, platelets) or direct (PIIINP, hyaluronic acid, metalloproteinases) markers, usually used in combination, have been evaluated. Simple scores such as APRI or FIB-4 have also been widely studied and have revealed interesting, albeit non-comprehensive, data on liver fibrosis, especially in terms of significant, extensive fibrosis or cirrhosis. These simple scores may be proposed as a first-line investigation, bearing in mind their limitations and comparing them with more accurate methods for evaluating liver fibrosis if necessary. Other scores, including direct serum markers, which can be difficult to assess, have given disappointing results that, in general, were either similar to, or only slightly better than, the results of the simpler tests. Further studies are needed to identify new markers that are more accurate and, above all, able to predict the outcome of liver fibrosis.

Résumé

Un effort de recherche majeur dans le domaine de l’évaluation non-invasive de la fibrose hépatique a permis ces dernières années la description et pour certains la validation de nombreux marqueurs sériques de fibrose essentiellement dans l’hépatite chronique C. A côté des scores déjà utilisés en clinique tels que le FibroTest ou le FibroMètre d’autres marqueurs de fibrose indirects (aspartate aminotransférase, taux de prothrombine, plaquettes) ou directs (PIIINP, acide hyaluronique, métalloprotéinases), le plus souvent utilisés en combinaison, ont été évalués. Certains scores simples tels que l’APRI ou le FIB-4 ont été largement étudiés et apportent des informations intéressantes bien qu’incomplètes en particulier en terme d’exclusion de fibrose significative, extensive et de cirrhose. Il peut donc être proposé de les utiliser en première intention en gardant en mémoire leurs limites et en les confrontant...
Other non-invasive markers of liver fibrosis

Autres marqueurs non-invasifs de fibrose hépatique

V. Leroy

Corresponding author: Clinique d’Hépato-gastroentérologie, pôle DIGI-DUNE, CHU de Grenoble, BP 217 38043 Grenoble cedex 9
E-mail address : Vleroy@chu-grenoble.fr

Summary
An intensive research effort in the field of non-invasive evaluation of liver fibrosis has recently permitted the description and validation of several serum markers of fibrosis, mainly in chronic hepatitis C patients. In addition to the commonly used tests such as FibroTest or FibroMeters, other either indirect (aspartate aminotransferase, prothrombin time, platelets) or direct (PIIINP, hyaluronic acid, metalloproteinases) markers, usually used in combination, have been evaluated. Simple scores such as APRI or FIB-4 have also been widely studied and have revealed interesting, albeit non-comprehensive, data on liver fibrosis, especially in terms of significant, extensive fibrosis or cirrhosis. These simple scores may be proposed as a first-line investigation, bearing in mind their limitations and comparing them with more accurate methods for evaluating liver fibrosis if necessary. Other scores, including direct serum markers, which can be difficult to assess, have given disappointing results that, in general, were either similar to, or only slightly better than, the results of the simpler tests. Further studies are needed to identify new markers that are more accurate and, above all, able to predict the outcome of liver fibrosis.

© 2008 Elsevier Masson SAS. All rights reserved.

Résumé
Un effort de recherche majeur dans le domaine de l’évaluation non-invasive de la fibrose hépatique a permis ces dernières années la description et pour certains la validation de nombreux marqueurs sériques de fibrose essentiellement dans l’hépatite chronique C. A côté des scores déjà utilisés en clinique tels que le FibroTest ou le FibroMètre d’autres marqueurs de fibrose indirects (Aspartate aminotranferase, taux de prothrombine, plaquettes) ou directs (PIIINP, acide hyaluronique, métalloprotéinases), le plus souvent utilisés en combinaison, ont été évalués. Certains scores simples tels que l’APRI ou le FIB-4 ont été largement étudiés et apportent des informations intéressantes bien qu’incomplètes en particulier en terme d’exclusion de fibrose significative, extensive et de cirrhose. Il peut donc être proposé de les utiliser en première intention en gardant en mémoire leurs limites et en les confrontant.

© 2008 Elsevier Masson SAS. Tous droits réservés.
In addition to the development of FibroTest and FibroMeters, a large number of other putative serum markers, used alone or in combinations, have been evaluated for the assessment of liver fibrosis. These are usually classified as indirect or direct markers, according to their relationship to fibrosis remodeling. Direct markers are either present in the extracellular matrix or involved in the regulation of fibrogenesis or fibrilysis. In contrast, although indirect markers are not directly related to fibrosis physiology, their concentrations are altered for various reasons, as liver fibrosis develops. The goal of this article is to review the diagnostic performance of the main new validated serum tests and to elucidate their utility in clinical practice.

Indirect markers

A large number of indirect serum markers have been tested over the past few years, especially for the diagnosis of cirrhosis. They mainly include aspartate aminotransferase (AST) and alanine aminotransferase (ALT) serum levels, platelet count and prothrombin time, either tested alone or in combination. Prothrombin time, a reproducible test, is a well-known marker of hepatic dysfunction. It also appears to be altered even in the earlier stages of fibrosis; speculative reasons for why this should be include the possible trapping of thrombin by the hepatic sinusoidal capillary network, and thrombin binding to the receptors expressed on the surface of activated hepatic stellate cells. Prothrombin time was negatively correlated to fibrosis in some studies that generally reported good positive predictive values (PPV) for the diagnosis of cirrhosis [1, 2]. For example, PPVs were 74% and 95% with the use of 85% and 70% prothrombin time thresholds, respectively.

AST levels usually increase in cases of liver fibrosis due to reduced clearance by the smaller sinusoidal capillary network. Some studies have reported that the AST/ALT ratio might be associated with the presence of cirrhosis, with a PPV approaching 90% for a ratio > 1 in viral--but not alcoholic--chronic liver disease [3, 4]. However, such a diagnostic utility has not been confirmed by other studies [5]. In the study by Pohl et al. [3], association of the AST/ALT ratio with thrombocytopenia < 150 g/L improved its diagnostic performance. Platelets decrease as liver fibrosis progresses not only because they are sequestered in the spleen in cases of portal hypertension, but also as a consequence of reduced thrombopoietin production. Taken individually, these cost-free blood tests may help in the diagnosis of cirrhosis, but they have not been tested for the diagnosis of less-advanced stages of fibrosis. However, this issue has been subsequently evaluated by more recent studies, which combined these markers with each other to produce optimized, statistically significant scores.

APRI

This test was originally proposed by Wai et al. [6] for a retrospective cross-sectional study of 270 chronic hepatitis C patients, of whom 47% had significant fibrosis and 15% had cirrhosis. By multivariate analyses, both AST and platelet count were significantly associated with fibrosis, leading the authors to devise a simplified score that could be calculated as: APRI = (AST/upper limit of normal) x 100/platelet count. The area under the receiver operating characteristic curve (AUROC) was 0.88 for significant fibrosis and 0.94 for cirrhosis. A score < 0.50, which was observed in 29% of patients, ruled out significant fibrosis and a negative predictive value (NPV) of 86%. A score < 1.0 excluded cirrhosis with an NPV of 98%, whereas a score > 2.0, observed in 15% of the patients, suggested cirrhosis with a PPV of 93%. However, on applying these thresholds, only 51% of patients were diagnosed with significant fibrosis and 81% with cirrhosis.

Subsequent studies have mainly looked at chronic hepatitis C patients with or without HIV co-infection and reported poorer diagnostic performance. A recently published meta-analysis of 22 studies, including 4266 patients mainly with chronic hepatitis C, reported interesting findings [7].
The prevalence of significant fibrosis and cirrhosis were 40% and 15%, respectively. Summary AUROCs were 0.76 [95% confidence interval (CI), 0.74 - 0.79] for significant fibrosis and 0.82 (95% CI, 0.79 - 0.86) for cirrhosis, with significant variability across the studies. With a threshold of 0.50, the sensitivity was 81% and the NPV was 80%. With a threshold of 1.0, the ability to exclude cirrhosis was excellent, with a sensitivity of 76% and an NPV of 91%. In contrast, even when using higher thresholds, APRI had suboptimal specificity and PPV for the diagnosis of significant fibrosis and cirrhosis. For example, with a threshold of 2.0, the PPV was 50% for the diagnosis of cirrhosis. The main conclusion of this study was that the major strength of APRI is its ability to exclude significant fibrosis and especially cirrhosis.

Another study that aimed to optimize APRI use suggested using a threshold of 0.6 instead of 0.5 for the exclusion of significant fibrosis, and obtained an NPV of 87% [8]. Recently, the diagnostic performance of APRI was validated in a cohort of haemodialysis patients with chronic hepatitis C, again with an excellent performance in excluding significant fibrosis [9]. However, considerably poorer results (AUROC = 0.66) were reported in a population of alcoholic patients with or without hepatitis C [10]. This poor performance in cases involving alcohol intake is understandable, as alcohol may have direct effects on AST levels and platelets, and APRI is a major limitation on the use of APRI in daily clinical practice.

The diagnostic performance of APRI has been compared with that of other tests, especially FibroTest and FibroMeters [11 - 13]. In these studies, the performance of APRI was no different from that of the other tests; however, these studies were also underpowered and only involved straightforward statistical tools such as the AUROC. In a recent meta-analysis of the individual data from 925 patients, it was found that the AUROC of APRI was significantly lower than that of FibroMeters (0.79 vs 0.84), but similar to that of FibroTest (0.79 vs 0.80) [14]. However, greater variability due to a lack of reproducibility was observed for APRI compared with FibroTest and FibroMeters. In a recent study involving 272 HIV/HCV co-infected patients, lower accuracy was seen with APRI compared with either FibroMeters or FibroTest [15].

In conclusion, APRI is a simple and cost-free test that can easily be used as a first-line investigation by clinicians and perhaps even included in more sophisticated algorithms. Its major strength is its ability to exclude, with around 90% certainty, the diagnosis of significant fibrosis or cirrhosis in a substantial number of patients. It has, however, important drawbacks, including high variability and poor diagnostic performance in cases of alcohol abuse; in addition, it cannot be used on its own in the majority of patients.

The Forns score

Forns et al. studied 476 chronic hepatitis C patients, of whom 75% had no or mild fibrosis, and divided them into a test cohort of 351 patients and a validation cohort of 125 patients. On the basis of multivariate analyses, the researchers derived a score that combined age, platelet count, cholesterol levels and gamma-glutamyl transferase (γ-GT), according to the formula: (Forns) score = 7.811 - 3.131 x ln platelets (g/L) + 0.781 x ln γ-GT (IU/L) + 3.647 x ln age (years) - 0.014 x cholesterol. Its overall diagnostic performance for significant fibrosis estimated by AUROCs was 0.86 in the test cohort and 0.81 in the validation cohort. A score > 4.2, which was observed in 39% of patients, had a sensitivity of 94% and an NPV of 96% for the diagnosis of significant fibrosis, whereas a score > 6.9, observed in 15% of patients, had a PPV of 66%. Half of the patients fell into the indeterminate group. No thresholds were defined for the diagnosis of extensive fibrosis or cirrhosis.

A limited number of studies have attempted to validate these results by comparing the Forns score with other serum markers [13,16 - 18]. Although the diagnostic performance of the Forns score reported in those studies was closer to that originally described, it was still significantly poorer than that of either FibroTest or FibroMeters in two of the studies [13, 16]. Also, several criticisms directed at these studies include concerns over the lipid abnormalities seen in chronic hepatitis C, an aspect that has not been extensively studied so far.

In conclusion, this test is of potential interest, but it does not appear to have been sufficiently validated for use in routine clinical practice.

FibrolIndex

Koda et al. [19] recently proposed a score that combined platelets, AST levels and gamma globulin, according to the formula: FibrolIndex = 1.738 - 0.064 x platelets (10⁴/mm³) + 0.005 x AST (IU/L) + 0.463 x gamma globulin (g/dL). In that study, 360 chronic hepatitis C patients were divided into a test cohort (n = 240) and a validation cohort (n = 120). AUROCs for predicting significant fibrosis were 0.83 in the test cohort and 0.82 in the validation cohort. The authors claimed that these results were superior to those observed with APRI (0.78) or the Forns score (0.78), but they do not appear to have performed statistical comparisons of the AUROCs. Two thresholds were determined that had low sensitivity and good specificity for the diagnosis of significant fibrosis. With the 1.25 threshold, specificity was 94% and the PPV was 87% whereas, with the 2.25 threshold, specificity was 97% and the PPV was 94%. Nevertheless, in a recent study comparing a panel of non-invasive scores, the diagnostic performance of FibrolIndex was significantly lower than that of FibroTest, and similar to that of APRI and the Forns score [18].

FIB-4

The FIB-4 is another test that combines age, AST, ALT and platelets: FIB-4 = age (year) x AST (IU/L) / [platelets (10⁴/L) x ALT (IU/L)]¹/² [20]. It was originally constructed from a series of 832 HIV/HCV co-infected patients included in the APRICOT study. Patients were assigned to either a training (n = 555) or a validation (n = 277) group. Liver fibrosis, evaluated by the Ishak score, was classified as mild (36%), moderate (44%) or advanced (20%). The AUROC of this index to differentiate
The prevalence of significant fibrosis and cirrhosis were 40% and 15%, respectively. Summary AUROCs were 0.76 [95% confidence interval (CI), 0.74 - 0.79] for significant fibrosis and 0.82 (95% CI, 0.79 - 0.86) for cirrhosis, with significant variability across the studies. With a threshold of 0.50, the sensitivity was 81% and the NPV was 80%. With a threshold of 1.0, the ability to exclude cirrhosis was excellent, with a sensitivity of 76% and an NPV of 91%. In contrast, even when using higher thresholds, APRI had suboptimal specificity and PPV for the diagnosis of significant fibrosis and cirrhosis. For example, with a threshold of 2.0, the PPV was 50% for the diagnosis of cirrhosis. The main conclusion of this study was that the major strength of APRI is its ability to exclude significant fibrosis and especially cirrhosis.

Another study that aimed to optimize APRI use suggested using a threshold of 0.6 instead of 0.5 for the exclusion of significant fibrosis, and obtained an NPV of 87% [8]. Recently, the diagnostic performance of APRI was validated in a cohort of haemodialysis patients with chronic hepatitis C, again with an excellent performance in excluding significant fibrosis [9]. However, considerably poorer results (AUROC = 0.66) were reported in a population of alcoholic patients with or without hepatitis C [10]. This poor performance in cases involving alcohol intake is understandable, as alcohol may have direct effects on AST levels and platelets; nevertheless, it is a major limitation on the use of APRI in daily clinical practice.

The diagnostic performance of APRI has been compared with that of other tests, especially FibroTest and FibroMeters [11 - 13]. In these studies, the performance of APRI was no different from that of the other tests; however, these studies were also underpowered and only involved straightforward statistical tools such as the AUROC. In a recent meta-analysis of the individual data from 925 patients, it was found that the AUROC of APRI was significantly lower than that of FibroMeters (0.79 vs 0.84), but similar to that of FibroTest (0.79 vs 0.80) [14]. However, greater variability due to a lack of reproducibility was observed for APRI compared with FibroTest and FibroMeters. In a recent study involving 272 HIV/HCV co-infected patients, lower accuracy was seen with APRI compared with either FibroMeters or FibroTest [15].

In conclusion, APRI is a simple and cost-free test that can easily be used as a first-line investigation by clinicians and perhaps even included in more sophisticated algorithms. Its major strength is its ability to exclude, with around 90% certainty, the diagnosis of significant fibrosis or cirrhosis in a substantial number of patients. It has, however, important drawbacks, including high variability and poor diagnostic performance in cases of alcohol abuse; in addition, it cannot be used on its own in the majority of patients.

### The Forns score

Forns et al. studied 476 chronic hepatitis C patients, of whom 75% had no or mild fibrosis, and divided them into a test cohort of 351 patients and a validation cohort of 125 patients. In the test cohort of 351 patients, they derived a score that combined age, platelet count, cholesterol levels and gamma-glutamyl transferase (γ-GT), according to the formula: (Forns) score = 7.811 - 3.131 x ln platelets (g/L) + 0.781 x ln γ-GT (IU/L) + 3.647 x ln age (years) - 0.014 x cholesterol. Its overall diagnostic performance for significant fibrosis estimated by AUROCs was 0.86 in the test cohort and 0.81 in the validation cohort. A score < 4.2, which was observed in 39% of patients, had a sensitivity of 94% and an NPV of 96% for the diagnosis of significant fibrosis, whereas a score > 6.9, observed in 15% of patients, had a PPV of 66%. Half of the patients fell into the indeterminate group. No thresholds were defined for the diagnosis of extensive fibrosis or cirrhosis.

A limited number of studies have attempted to validate these results by comparing the Forns score with other serum markers [13,16 - 18]. Although the diagnostic performance of the Forns score reported in those studies was closer to that originally described, it was still significantly poorer than that of either FibroTest or FibroMeters in two of the studies [13, 16]. Also, several criticisms directed at these studies include concerns over the lipid abnormalities seen in chronic hepatitis C, an aspect that has not been extensively studied so far.

In conclusion, this test is of potential interest, but it does not appear to have been sufficiently validated for use in routine clinical practice.

### FibrolIndex

Koda et al. [19] recently proposed a score that combined platelets, AST levels and gamma globulin, according to the formula: FibrolIndex = 1.738 – 0.064 x platelets (10^4/mm^3) + 0.005 x AST (IU/L) + 0.463 x gamma globulin (g/dL). In that study, 360 chronic hepatitis C patients were divided into a test cohort (n = 240) and a validation cohort (n = 120). AUROCs for predicting significant fibrosis were 0.83 in the test cohort and 0.82 in the validation cohort. The authors claimed that these results were superior to those observed with APRI (0.78) or the Forns score (0.78), but they do not appear to have performed statistical comparisons of the AUROCs. Two thresholds were determined that had low sensitivity and good specificity for the diagnosis of significant fibrosis. With the 1.25 threshold, specificity was 94% and the PPV was 87% whereas, with the 2.25 threshold, specificity was 97% and the PPV was 94%. Nevertheless, in a recent study comparing a panel of non-invasive scores, the diagnostic performance of FibrolIndex was significantly lower than that of FibroTest, and similar to that of APRI and the Forns score [18].

### FIB-4

The FIB-4 is another test that combines age, AST, ALT and platelets: FIB-4 = age (year) x AST (IU/L)/[platelets (10^9/L) x ALT (IU/L)]^1/2 [20]. It was originally constructed from a series of 832 HIV/HCV co-infected patients included in the APRICOT study. Patients were assigned to either a training (n = 555) or a validation (n = 277) group. Liver fibrosis, evaluated by the Ishak score, was classified as mild (36%), moderate (44%) or advanced (20%). The AUROC of this index to differentiate
between mild/moderate vs advanced fibrosis was 0.77. A score < 1.45 had a sensitivity of 70% and an NPV of 90% for excluding advanced fibrosis, while a threshold of 3.25 had a specificity of 97% and a PPV of 65% for the diagnosis of advanced fibrosis.

Vallet-Pichard et al. [21] recently validated these results in a large cohort of 847 HCV mono-infected patients. In this study, the AUROCs were 0.85 for the diagnosis of extensive fibrosis and 0.91 for cirrhosis. A score < 1.45 ruled out extensive fibrosis with 95% NPV while a score > 3.25 had 82% PPV for this diagnosis; 27% of patients fell into the indeterminate range. Due to wide overlapping of the scores for F0, F1 and F2, its capacity to diagnose no or mild fibrosis (F0/F1) was not tested in this study. The choice of threshold, rather than the intrinsic performance of FIB-4, is the likely explanation of its major discrepancy compared with APRI.

Interestingly, FIB-4 and FibroTest both correlated with each other and were concordant in the vast majority of cases. Also, FIB-4 was recently evaluated by other teams and compared with other markers in HCV mono-infected and HIV/HCV co-infected patients [18, 22, 23]. In these cases, its diagnostic performance was found to be better than those of the two studies of APRI, with no significant difference compared with FibroTest.

### Direct markers

The development of liver fibrosis is a complex wound-healing process, involving a range of cell types and mediators that eventually lead to the deposition and accumulation of numerous components in the extracellular matrix. Some of these components, including collagen, glycoproteins, proteoglycans and glycosaminoglycans, can be tested for in serum by various techniques, and their serum concentration reflects more or less accurately the levels of deposition in situ. Moreover, wound-healing is a dynamic process and certain markers, especially those belonging to the complex network of cytokines, growth factors, metalloproteinases and their inhibitors, may have the potential to reveal information concerning fibrogenesis and fibrolysis and, thus, the outcome of the disease.

Such an hypothesis was recently proposed by Fontana et al. [24], who found that a panel of direct serum markers was more closely correlated to fibrosis stage than to hepatic collagen content assessed by morphometry. However, none of the markers that have been tested so far is ideal, as they can have complex metabolic processes in terms of excretion and clearance, they are not specific to the liver and testing methods can be difficult to implement or insufficiently reliable.

### Collagen

Changes in extracellular matrix composition during fibrosis include accumulation of fibril-forming collagen, particularly types I and III. Type I collagen at protein and mRNA levels was found to correlate with fibrosis [25]. The most extensively studied collagen is the PIIINP, an amino-terminal peptide of procollagen III (a42 kDa) cleaved from procollagen III during its excretion by fibroblasts. For this reason, it is thought to be mainly a marker of fibrogenesis rather than fibrosis. Some clinical studies have shown a significant correlation between PIIINP levels and fibrosis stages in chronic hepatitis C patients, but its diagnostic value appears to be insufficient to support its use as a fibrosis tool on its own [26 – 28].

### Hyaluronic acid

Hyaluronic acid is a glycosaminoglycan, produced by myofibroblasts, with a structural role in the extracellular matrix. It is degraded by liver sinusoidal cells, and its increased levels in serum may be linked to the endothelial dysfunction that occurs as fibrosis progresses. Several studies have shown a significant correlation between hyaluronic acid and fibrosis, especially in chronic hepatitis C patients [1, 28, 29]. In the largest study, Mchutchison et al. [29] evaluated hyaluronic acid serum concentrations in 486 HCV-infected patients, of whom 78 (16%) had cirrhosis.

Hyaluronic acid levels < 60 μg/L excluded cirrhosis and extensive fibrosis, with NPVs of 99% and 93%, respectively, while values > 110 μg/L had a PPV of 44% for cirrhosis.

### MMPs and TIMPs

Metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) are proteins that are involved in the regulation of fibrogenesis and fibrolysis. MMPs are secreted by cell types such as fibroblasts and Kupffer cells, usually in a proenzyme form that requires further cleavage for functional capacity. They are mainly involved in the degradation of normal and pathological extracellular matrices and are, in turn, regulated by the specific inhibitors called TIMPs. Clinical studies have shown a positive correlation between TIMP-1, TIMP-2 and MMP-2 and fibrosis [28, 30]. MMP-2 is upregulated by type I collagen and may serve as a degradation agent of normal extracellular matrix to allow the development of pathological matrix.

Boeker et al. [31] showed, in 78 patients, that TIMP-1 had a sensitivity of 100% for cirrhosis and a specificity of 75%. In addition, it has been shown in HCV-infected patients that PIIINP, hyaluronic acid and TIMP-1 were strongly intercorrelated and had similar accuracy for the diagnosis of extensive fibrosis and cirrhosis [28]. Other MMPs have been less extensively studied. However, in one study, MMP-1 and MMP-9 concentrations were negatively correlated with liver fibrosis [28]. On the basis of this finding, a score called MP3 was derived, combining PIIINP and MMP-1, with an AUROC of 0.82 for the diagnosis of significant fibrosis and 0.88 for extensive fibrosis. Subsequently, the diagnostic accuracy of MP3 has proved to be similar—but not superior—to those of FibroTest and FibroMeters, and might be useful for the longitudinal follow-up of HCV-infected patients treated with interferon-α and ribavirin [13, 32]. However, as a test, MP3 has not been validated by other researchers, and its limitations include high cost, complexity and a reproduciability of serum dosages that requires the use of expert laboratories.

Overall and so far, the results obtained with direct markers have been disappointing, as they do not appear to perform any better than the more straightforward, indirect markers. This has led to recent tests that have combined both types of markers in an attempt to improve their diagnostic performance.
between mild/moderate vs advanced fibrosis was 0.77. A score < 1.45 had a sensitivity of 70% and an NPV of 90% for excluding advanced fibrosis, while a threshold of 3.25 had a specificity of 97% and a PPV of 65% for the diagnosis of advanced fibrosis.

Vallet-Pichard et al. [21] recently validated these results in a large cohort of 847 HCV mono-infected patients. In this study, the AUROCs were 0.85 for the diagnosis of extensive fibrosis and 0.91 for cirrhosis. A score < 1.45 ruled out extensive fibrosis with 95% NPV while a score > 3.25 had 82% PPV for this diagnosis; 27% of patients fell into the indeterminate range. Due to wide overlapping of the scores for F0, F1 and F2, its capacity to diagnose no or mild fibrosis (F0/F1) was not tested in this study. The choice of threshold, rather than the intrinsic performance of FIB-4, is the likely explanation of its major discrepancy compared with APRI.

Interestingly, FIB-4 and FibroTest both correlated with each other and were concordant in the vast majority of cases. Also, FIB-4 was recently evaluated by other teams and compared with other markers in HCV mono-infected and HIV/HCV co-infected patients [18, 22, 23]. In these cases, its diagnostic performance was found to be better than those of the two studies of APRI, with no significant difference compared with FibroTest.

### Direct markers

The development of liver fibrosis is a complex wound-healing process, involving a range of cell types and mediators that eventually lead to the deposition and accumulation of numerous components in the extracellular matrix. Some of these components, including collagen, glycoproteins, proteoglycans and glycosaminoglycans, can be tested for in serum by various techniques, and their serum concentration reflects more or less accurately the levels of deposition in situ. Moreover, wound-healing is a dynamic process and certain markers, especially those belonging to the complex network of cytokines, growth factors, metalloproteinases and their inhibitors, may have the potential to reveal information concerning fibrogenesis and fibrolysis and, thus, the outcome of the disease.

Such an hypothesis was recently proposed by Fontana et al. [24], who found that a panel of direct serum markers was more closely correlated to fibrosis stage than to hepatic collagen content assessed by morphometry. However, none of the markers that have been tested so far is ideal, as they can have complex metabolic processes in terms of excretion and clearance, they are not specific to the liver and testing methods can be difficult to implement or insufficiently reliable.

### Collagen

Changes in extracellular matrix composition during fibrosis include accumulation of fibril-forming collagen, particularly types I and III. Type I collagen at protein and mRNA levels was found to correlate with fibrosis [25]. The most extensively studied collagen is the PIIINP, an amino-terminal peptide of procollagen III (142 kDa) cleaved from procollagen III during its excretion by fibroblasts. For this reason, it is thought to be mainly a marker of fibrogenesis rather than fibrosis. Some clinical studies have shown a significant correlation between PIIINP levels and fibrosis stages in chronic hepatitis C patients, but its diagnostic value appears to be insufficient to support its use as a fibrosis tool on its own [26 - 28].

### Hyaluronic acid

Hyaluronic acid is a glycosaminoglycan, produced by myofibroblasts, with a structural role in the extracellular matrix. It is degraded by liver sinusoidal cells, and its increased levels in serum may be linked to the endothelial dysfunction that occurs as fibrosis progresses. Several studies have shown a significant correlation between hyaluronic acid and fibrosis, especially in chronic hepatitis C patients [1,28,29]. In the largest study, McHutchison et al. [29] evaluated hyaluronic acid serum concentrations in 486 HCV-infected patients, of whom 78 (16%) had cirrhosis.

Hyaluronic acid levels < 60 μg/L excluded cirrhosis and extensive fibrosis, with NPVs of 99% and 93%, respectively, while values > 110 μg/L had a PPV of 44% for cirrhosis.

### MMPs and TIMPs

Metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) are proteins that are involved in the regulation of fibrogenesis and fibrolysis. MMPs are secreted by cell types such as fibroblasts and Kupffer cells, usually in a proenzyme form that requires further cleavage for functional capacity. They are mainly involved in the degradation of normal and pathological extracellular matrices and are, in turn, regulated by the specific inhibitors called TIMPs. Clinical studies have shown a positive correlation between TIMP-1, TIMP-2 and MMP-2 and fibrosis [28,30]. MMP-2 is upregulated by type I collagen and may serve as a degradation agent of normal extracellular matrix to allow the development of pathological matrix.

Boeker et al. [31] showed, in 78 patients, that TIMP-1 had a sensitivity of 100% for cirrhosis and a specificity of 75%. In addition, it has been shown in HCV-infected patients that PIIINP, hyaluronic acid and TIMP-1 were strongly intercorrelated and had similar accuracy for the diagnosis of extensive fibrosis and cirrhosis [28]. Other MMPs have been less extensively studied. However, in one study, MMP-1 and MMP-9 concentrations were negatively correlated with liver fibrosis [28]. On the basis of this finding, a score called MP3 was derived, combining PIIINP and MMP-1, with an AUROC of 0.82 for the diagnosis of significant fibrosis and 0.88 for extensive fibrosis. Subsequently, the diagnostic accuracy of MP3 has proved to be similar—but not superior—to those of FibroTest and FibroMeters, and might be useful for the longitudinal follow-up of HCV-infected patients treated by interferon-α and ribavirin [13,32]. However, as a test, MP3 has not been validated by other researchers, and its limitations include high cost, complexity and a reproducibility of serum dosages that requires the use of expert laboratories.

Overall and so far, the results obtained with direct markers have been disappointing, as they do not appear to perform any better than the more straightforward, indirect markers. This has led to recent tests that have combined both types of markers in an attempt to improve their diagnostic performance.
between mild/moderate vs advanced fibrosis was 0.77. A score < 1.45 had a sensitivity of 70% and an NPV of 90% for excluding advanced fibrosis, while a threshold of 3.25 had a specificity of 97% and a PPV of 65% for the diagnosis of advanced fibrosis.

Vallet-Pichard et al. [21] recently validated these results in a large cohort of 847 HCV mono-infected patients. In this study, the AUROCs were 0.85 for the diagnosis of extensive fibrosis and 0.91 for cirrhosis. A score < 1.45 ruled out extensive fibrosis with 95% NPV while a score > 3.25 had 82% PPV for this diagnosis; 27% of patients fell into the indeterminate range. Due to wide overlapping of the scores for F0, F1 and F2, its diagnostic performance was found to be better than that of the two studies of APRI, with no significant difference compared with FibroTest.

Interestingly, FIB-4 and FibroTest both correlated with each other and were concordant in the vast majority of cases. Also, FIB-4 was recently evaluated by other teams and compared with other markers in HCV mono-infected and HIV/HCV co-infected patients [18, 22, 23]. In these cases, its diagnostic performance was found to be better than that of the two studies of APRI.

### Direct markers

The development of liver fibrosis is a complex wound-healing process, involving a range of cell types and mediators that eventually lead to the deposition and accumulation of numerous components in the extracellular matrix. Some of these components, including collagen, proteoglycans, proteoglycans and glycosaminoglycans, can be tested for in serum by various techniques, and their serum concentration reflects more or less accurately the levels of deposition in situ. Moreover, wound-healing is a dynamic process and certain markers, especially those belonging to the complex network of cytokines, growth factors, metalloproteinases and their inhibitors, may have the potential to reveal information concerning fibrogenesis and fibrolysis and, thus, the outcome of the disease.

Such an hypothesis was recently proposed by Fontana et al. [24], who found that a panel of direct serum markers was more closely correlated to fibrosis stage than to hepatic collagen content assessed by morphometry. However, none of the markers that have been tested so far is ideal, as they can have complex metabolic processes in terms of excretion and clearance, they are not specific to the liver and testing methods can be difficult to implement or insufficiently reliable.

### Collagen

Changes in extracellular matrix composition during fibrosis include accumulation of fibril-forming collagen, particularly types I and III. Type I collagen at protein and mRNA levels was found to correlate with fibrosis [25]. The most extensively studied collagen is the PIIINP, an amino-terminal peptide of procollagen III (42 kDa), cleaved from procollagen III during its excretion by fibroblasts. For this reason, it is thought to be mainly a marker of fibrogenesis rather than fibrosis. Some clinical studies have shown a significant correlation between PIIINP levels and fibrosis stages in chronic hepatitis C patients, but its diagnostic value appears to be insufficient to support its use as a fibrosis tool on its own [26 – 28].

### Hyaluronic acid

Hyaluronic acid is a glycosaminoglycan, produced by myofibroblasts, with a structural role in the extracellular matrix. It is degraded by liver sinusoidal cells, and its increased levels in serum may be linked to the endothelial dysfunction that occurs as fibrosis progresses. Several studies have shown a significant correlation between hyaluronic acid and fibrosis, especially in chronic hepatitis C patients [1, 28, 29]. In the largest study, McHutchison et al. [29] evaluated hyaluronic acid serum concentrations in 486 HCV-infected patients, of whom 78 (16%) had cirrhosis.

Hyaluronic acid levels < 60 μg/L excluded cirrhosis and extensive fibrosis, with NPVs of 99% and 93%, respectively, while values > 110 μg/L had a PPV of 44% for cirrhosis.

### MMPs and TIMPs

Metalloproteases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) are proteins that are involved in the regulation of fibrogenesis and fibrolysis. MMPs are secreted by cell types such as fibroblasts and Kupffer cells, usually in a proenzyme form that requires further cleavage for functional capacity. They are mainly involved in the degradation of normal and pathological extracellular matrices and are, in turn, regulated by the specific inhibitors called TIMPs. Clinical studies have shown a positive correlation between TIMP-1, TIMP-2 and MMP-2 and fibrosis [28, 30]. MMP-2 is upregulated by type I collagen and may serve as a degradation agent of normal extracellular matrix to allow the development of pathological matrix.

Boeker et al. [31] showed, in 78 patients, that TIMP-1 had a sensitivity of 100% for cirrhosis and a specificity of 75%. In addition, it has been shown in HCV-infected patients that PIIINP, hyaluronic acid and TIMP-1 were strongly intercorrelated and had similar accuracy for the diagnosis of extensive fibrosis and cirrhosis [28]. Other MMPs have been less extensively studied. However, in one study, MMP-1 and MMP-9 concentrations were negatively correlated with liver fibrosis [28]. On the basis of this finding, a score called MP3 was derived, combining PIIINP and MMP-1, with an AUROC of 0.82 for the diagnosis of significant fibrosis and 0.88 for extensive fibrosis. Subsequently, the diagnostic accuracy of MP3 has proved to be similar—but not superior—to those of FibroTest and FibroMeters, and might be useful for the longitudinal follow-up of HCV-infected patients treated by interferon-α and ribavirin [13, 32]. However, as a test, MP3 has not been validated by other researchers, and its limitations include high cost, complexity and a reproducibility of serum dosages that requires the use of expert laboratories.

Overall and so far, the results obtained with direct markers have been disappointing, as they do not appear to perform any better than the more straightforward, indirect markers. This has led to recent tests that have combined both types of markers in an attempt to improve their diagnostic performance.
Combination of indirect and direct markers

Hepascore

The Hepascore combines age, gender, bilirubin, gamma GT, alpha2-macroglobulin and hyaluronic acid according to the formula: Hepascore = y/(1 + y), where y = \exp(-4.185818 - 0.0249 \times \text{age} + 0.7464 \times \text{gender (male = 1, female = 0)} + 1.0039 \times \alpha_2\text{-macroglobulin} + 0.0302 \times \text{hyaluronic acid} + 0.0691 \times \text{bilirubin} - 0.0012 \times \gamma\text{-GT}). In the original study published by Adams et al. [33], including 221 HCV-infected patients, the AUROCs were 0.82, 0.90 and 0.89 in the validation groups for the diagnosis of significant fibrosis, extensive fibrosis and cirrhosis, respectively. A score < 0.5 had 74% specificity and 88% sensitivity for extensive fibrosis (NPV = 95%), while a score > 0.5 had a specificity and sensitivity of 89% and 63%, respectively, for significant fibrosis (PPV = 87%). Interestingly, a higher threshold of 0.84 had good sensitivity (71%) and an excellent NPV (94%) for excluding cirrhosis.

Hepascore diagnostic accuracy has been compared with those of other serum markers and found to be similar [15,17] or significantly lower than those of FibroTest and FibroMeters [14,22]. In the study by Bourlière et al. [17], the excellent accuracy of Hepascore to exclude cirrhosis was confirmed using the 0.84 threshold, which had a 71% sensitivity and 97% NPV, while the specificity was 88% and the PPV, only 33%.

FIBROSpect

FIBROSpect is a test initially described by Patel et al. [34] that combines hyaluronic acid, TIMP-1 and alpha2-macroglobulin. It is licensed and commercially available in the USA. In the original study of 696 HCV-infected patients, the AUROC was 0.83 for the diagnosis of significant fibrosis. This test has recently been validated in several studies [35–38]. In general, the test proved to be excellent for excluding either significant fibrosis or cirrhosis.

In the study by Zaman et al. [36], using appropriate thresholds, F2/F3/F4 was excluded with 72% sensitivity (NPV = 82%), while F3/F4 were excluded with 82% sensitivity and 97% NPV. In contrast, the PPVs were low at 61% and 21%, respectively, but it should be pointed out that the prevalences of significant (36%) and extensive (13%) fibrosis were low in this study. To our knowledge, although this score was not compared with either FibroTest or FibroMeters, its diagnostic performance appears to be similar.

Other tests

Other tests combining direct and indirect markers have been described, but only limited information on their performance is available, and they have rarely been externally validated by other teams. The SHASTA index combines hyaluronic acid, AST and albumin, and was developed in 95 HIV/HCV co-infected patients [39]. In that study, its accuracy was similar to that of FibroTest. Rosenberg et al. [40] proposed a score combining age, hyaluronic acid A, PIIINP and TIMP-1 that had good accuracy in both alcoholic and non-alcoholic chronic liver disease. Recently, Fontana et al. [24] constructed, from the HALT-C cohort, a score combining hyaluronic acid, TIMP-1 and platelet count that, again, was highly accurate for the diagnosis of cirrhosis.

Use of tests in clinical practice

Although an impressive number of biomarker tests are now becoming available, determining a rationale for their clinical use is difficult. In countries where FibroMeters, FibroTest and FibroScan are available, the use of the more straightforward non-invasive tests, such as APRI and FIB-4, could be considered pointless. However, the latter are cost-free, virtually always available and easy to calculate—although clinicians generally do not keep calculators in their pockets and do not like formulas. Most important of all, they provide, if not a comprehensive evaluation, at least some interesting clues for the status of liver fibrosis. This means that either APRI or FIB-4 may have a role in the first-line assessment of liver disease patients, although a more reliable evaluation of liver fibrosis is often necessary in the majority of patients for making therapeutic and management decisions. The information obtained from these simple tests can help in the interpretation of the findings of other diagnostic tools—all of which, including biopsy, may give erroneous results—and in the verification of findings consistent with the clinical setting. Such simple tests, as well as tests such as Hepascore, could also be incorporated into more rational algorithms, a specific and important aspect that is discussed in another article in this issue.

Conflicts of interest:

Vincent Leroy participates to clinical trials, has occasional involvements (for expert reports and advisory services), and attends conferences (either as contributor or audience member) on behalf of Roche, Schering Plough, Bayer, BMS, Gilead and Novartis.

References

Combination of indirect and direct markers

Hepascore

The Hepascore combines age, gender, bilirubin, γ-GT, α2-macroglobulin and hyaluronic acid according to the formula: Hepascore = y/(1 + y), where y = exp (-4.185818 - 0.0249 x age + 0.7464 x gender (male = 1, female = 0) - 1.0039 x α2-macroglobulin + 0.0302 x hyaluronic acid + 0.0691 x bilirubin - 0.0012 x γ-GT. In the original study published by Adams et al. [33], including 221 HCV-infected patients, the AUROCs were 0.82, 0.90 and 0.89 in the validation groups for the diagnosis of significant fibrosis, extensive fibrosis and cirrhosis, respectively. A score < 0.5 had 74% specificity and 88% sensitivity for extensive fibrosis (NPV = 95%), while a score > 0.5 had a specificity and sensitivity of 89% and 63%, respectively, for significant fibrosis (PPV = 87%). Interestingly, a higher threshold of 0.84 had good sensitivity (71%) and an excellent NPV (94%) for excluding cirrhosis.

Hepascore diagnostic accuracy has been compared with those of other serum markers and found to be similar [15,17] or significantly lower than those of FibroTest and FibroMeters [14,22]. In the study by Bourière et al. [17], the excellent accuracy of Hepascore to exclude cirrhosis was confirmed using the 0.84 threshold, which had a 71% sensitivity and 97% NPV, while the specificity was 88% and the PPV, only 33%.

FIBROSpect

FIBROSpect is a test initially described by Patel et al. [34] that combines hyaluronic acid, TIMP-1 and α2-macroglobulin. It is licensed and commercially available in the USA. In the original study of 696 HCV-infected patients, the AUROC was 0.83 for the diagnosis of significant fibrosis. This test has recently been validated in several studies [35–38]. In general, the test proved to be excellent for excluding either significant fibrosis or cirrhosis.

In the study by Zaman et al. [36], using appropriate thresholds, F2/F3/F4 was excluded with 72% sensitivity (NPV = 82%), while F3/F4 were excluded with 82% sensitivity and 97% NPV. In contrast, the PPVs were low at 61% and 21%, respectively, but it should be pointed out that the prevalences of significant (36%) and extensive (13%) fibrosis were low in this study. To our knowledge, although this score was not compared with either FibroTest or FibroMeters, its diagnostic performance appears to be similar.

Other tests

Other tests combining direct and indirect markers have been described, but only limited information on their performance is available, and they have rarely been externally validated by other teams. The SHASTA index combines hyaluronic acid, AST and albumin, and was developed in 95 HIV/HCV co-infected patients [39]. In that study, its accuracy was similar to that of FibroTest. Rosenberg et al. [40] proposed a score combining age, hyaluronic acid A, PIIINP and TIMP-1 that had good accuracy in both alcoholic and non-alcoholic chronic liver disease. Recently, Fontana et al. [24] constructed, from the HALT-C cohort, a score combining hyaluronic acid, TIMP-1 and platelet count that, again, was highly accurate for the diagnosis of cirrhosis.

Use of tests in clinical practice

Although an impressive number of biomarker tests are now becoming available, determining a rationale for their clinical use is difficult. In countries where FibroMeters, FibroTest and FibroScan are available, the use of the more straightforward non-invasive tests, such as APRI and FIB-4, could be considered pointless. However, the latter are cost-free, virtually always available and easy to calculate—although clinicians generally do not keep calculators in their pockets and do not like formulas. Most important of all, they provide, if not a comprehensive evaluation, at least some interesting clues for the status of liver fibrosis. This means that either APRI or FIB-4 may have a role in the first-line assessment of liver-disease patients, although a more reliable evaluation of liver fibrosis is often necessary in the majority of patients for making therapeutic and management decisions. The information obtained from these simple tests can help in the interpretation of the findings of other diagnostic tools—all of which, including biopsy, may give erroneous results—and in the verification of findings consistent with the clinical setting. Such simple tests, as well as tests such as Hepascore, could also be incorporated into more rational algorithms, a specific and important aspect that is discussed in another article in this issue.

Conflicts of interest:

Vincent Leroy participates to clinical trials, has occasional involvements (for expert reports and advisory services), and attends conferences (either as contributor or audience member) on behalf of Roche, Schering Plough, Bayer, BMS, Gilead and Novartis.

References

Combination of indirect and direct markers

Hepascore

The Hepascore combines age, gender, bilirubin, γ-GT, α2-macroglobulin and hyaluronic acid according to the formula: Hepascore = y/(1 + y), where y = exp (-4.185818 - 0.0249 x age + 0.7464 x gender (male = 1, female = 0) + 1.0039 x α2-macroglobulin + 0.0302 x hyaluronic acid + 0.0691 x bilirubin - 0.0012 x γ-GT. In the original study published by Adams et al. [33], including 221 HCV-infected patients, the AUROCs were 0.82, 0.90 and 0.89 in the validation groups for the diagnosis of significant fibrosis, extensive fibrosis and cirrhosis, respectively. A score < 0.5 had 74% specificity and 88% sensitivity for extensive fibrosis (NPV = 95%), while a score > 0.5 had a specificity and sensitivity of 89% and 63%, respectively, for significant fibrosis (PPV = 87%). Interestingly, a higher threshold of 0.84 had good sensitivity (71%) and an excellent NPV (94%) for excluding cirrhosis.

Hepascore diagnostic accuracy has been compared with those of other serum markers and found to be similar [15,17] or significantly lower than those of FibroTest and FibroMeters [14,22]. In the study by Bourlière et al. [17], the excellent accuracy of Hepascore to exclude cirrhosis was confirmed using the 0.84 threshold, which had a 71% sensitivity and 97% NPV, while the specificity was 88% and the PPV, only 33%.

FIBROSpect

FIBROSpect is a test initially described by Patel et al. [34] that combines hyaluronic acid, TIMP-1 and α2-macroglobulin. It is licensed and commercially available in the USA. In the original study of 696 HCV-infected patients, the AUROC was 0.83 for the diagnosis of significant fibrosis. This test has recently been validated in several studies [35 - 38]. In general, the test proved to be excellent for excluding either significant fibrosis or cirrhosis.

In the study by Zaman et al. [36], using appropriate thresholds, F2/F3/F4 was excluded with 72% sensitivity (NPV = 82%), while F3/F4 were excluded with 82% sensitivity and 97% NPV. In contrast, the PPVs were low at 61% and 21%, respectively, but it should be pointed out that the prevalences of significant (36%) and extensive (13%) fibrosis were low in this study. To our knowledge, although this score was not compared with either FibroTest or FibroMeters, its diagnostic performance appears to be similar.

Other tests

Other tests combining direct and indirect markers have been described, but only limited information on their performance is available, and they have rarely been externally validated by other teams. The SHASTA index combines hyaluronic acid, AST and albumin, and was developed in 95 HIV/HCV co-infected patients [39]. In that study, its accuracy was similar to that of FibroTest. Rosenberg et al. [40] proposed a score combining age, hyaluronic acid A, PIIINP and TIMP-1 that had good accuracy in both alcoholic and non-alcoholic chronic liver disease. Recently, Fontana et al. [24] constructed, from the HALT-C cohort, a score combining hyaluronic acid, TIMP-1 and platelet count that, again, was highly accurate for the diagnosis of cirrhosis.

Use of tests in clinical practice

Although an impressive number of biomarker tests are now becoming available, determining a rationale for their clinical use is difficult. In countries where FibroMeters, FibroTest and FibroScan are available, the use of the more straightforward non-invasive tests, such as APRI and FIB-4, could be considered pointless. However, the latter are cost-free, virtually always available and easy to calculate—although clinicians generally do not keep calculators in their pockets and do not like formulas. Most important of all, they provide, if not a comprehensive evaluation, at least some interesting clues for the status of liver fibrosis. This means that either APRI or FIB-4 may have a role in the first-line assessment of liver disease patients, although a more reliable evaluation of liver fibrosis is often necessary in the majority of patients for making therapeutic and management decisions. The information obtained from these simple tests can help in the interpretation of the findings of other diagnostic tools—all of which, including biopsy, may give erroneous results—and in the verification of findings consistent with the clinical setting. Such simple tests, as well as tests such as Hepascore, could also be incorporated into more rational algorithms, a specific and important aspect that is discussed in another article in this issue.

Conflicts of interest:

Vincent Leroy participates to clinical trials, has occasional involvements (for expert reports and advisory services), and attends conferences (either as contributor or audience member) on behalf of Roche, Schering Plough, Bayer, BMS, Gilead and Novartis.

References


V. Leroy
Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection.


Abstract

OBJECTIVES:

Needle liver biopsy has been shown to have a high rate of sampling error in patients with diffuse parenchymal liver diseases. In these cases, the sample of liver tissue does not reflect the true degree of inflammation, fibrosis, or cirrhosis, despite an adequate sample size. The aim of this study was to determine the rate and extent of sampling error in patients with chronic hepatitis C virus infection, and to assess the intraobserver variation with the commonly used scoring system proposed by Scheuer and modified by Batts and Ludwig.

METHODS:

A total of 124 patients with chronic hepatitis C virus infection underwent simultaneous laparoscopy-guided biopsies of the right and left hepatic lobes. Formalin-fixed paraffin-embedded sections were stained with hematoxylin and eosin and with trichrome. The slides were blindly coded and randomly divided among two hepatopathologists. Inflammation and fibrosis were scored according to the standard grading (inflammation) and staging (fibrosis) method based on the modified Scheuer system. Following the interpretation, the slides were uncoded to compare the results of the right and left lobes. Fifty of the samples were blindly resubmitted to each of the pathologists to determine the intraobserver variation.

RESULTS:

Thirty of 124 patients (24.2%) had a difference of at least one grade, and 41 of 124 patients (33.1%) had a difference of at least one stage between the right and left lobes. In 18 patients (14.5%), interpretation of cirrhosis was given in one lobe, whereas stage 3 fibrosis was given in the other. A difference of two stages or two grades was found in only three (2.4%) and two (1.6%) patients, respectively. Of the 50 samples that were examined twice, the grading by each pathologist on the second examination differed from the first examination in 0% and 4%, and the staging differed in 6% and 10%, respectively. All observed variations were of one grade or one stage.

CONCLUSIONS:

Liver biopsy samples taken from the right and left hepatic lobes differed in histological grading and staging in a large proportion of chronic hepatitis C virus patients; however, differences of more than one stage or grade were uncommon. A sampling error may have led to underdiagnosis of cirrhosis in 14.5% of the patients. These differences could not be attributed to intraobserver variation, which appeared to be low.
Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection.


Abstract

OBJECTIVES:

Needle liver biopsy has been shown to have a high rate of sampling error in patients with diffuse parenchymal liver diseases. In these cases, the sample of liver tissue does not reflect the true degree of inflammation, fibrosis, or cirrhosis, despite an adequate sample size. The aim of this study was to determine the rate and extent of sampling error in patients with chronic hepatitis C virus infection, and to assess the intraobserver variation with the commonly used scoring system proposed by Scheuer and modified by Batts and Ludwig.

METHODS:

A total of 124 patients with chronic hepatitis C virus infection underwent simultaneous laparoscopy-guided biopsies of the right and left hepatic lobes. Formalin-fixed paraffin-embedded sections were stained with hematoxylin and eosin and with trichrome. The slides were blindly coded and randomly divided among two hepatopathologists. Inflammation and fibrosis were scored according to the standard grading (inflammation) and staging (fibrosis) method based on the modified Scheuer system. Following the interpretation, the slides were uncoded to compare the results of the right and left lobes. Fifty of the samples were blindly resubmitted to each of the pathologists to determine the intraobserver variation.

RESULTS:

Thirty of 124 patients (24.2%) had a difference of at least one grade, and 41 of 124 patients (33.1%) had a difference of at least one stage between the right and left lobes. In 18 patients (14.5%), interpretation of cirrhosis was given in one lobe, whereas stage 3 fibrosis was given in the other. A difference of two stages or two grades was found in only three (2.4%) and two (1.6%) patients, respectively. Of the 50 samples that were examined twice, the grading by each pathologist on the second examination differed from the first examination in 0% and 4%, and the staging differed in 6% and 10%, respectively. All observed variations were of one grade or one stage.

CONCLUSIONS:

Liver biopsy samples taken from the right and left hepatic lobes differed in histological grading and staging in a large proportion of chronic hepatitis C virus patients; however, differences of more than one stage or grade were uncommon. A sampling error may have led to underdiagnosis of cirrhosis in 14.5% of the patients. These differences could not be attributed to intraobserver variation, which appeared to be low.
OBJECTIVES:

Needle liver biopsy has been shown to have a high rate of sampling error in patients with diffuse parenchymal liver diseases. In these cases, the sample of liver tissue does not reflect the true degree of inflammation, fibrosis, or cirrhosis, despite an adequate sample size. The aim of this study was to determine the rate and extent of sampling error in patients with chronic hepatitis C virus infection, and to assess the intraobserver variation with the commonly used scoring system proposed by Scheuer and modified by Batts and Ludwig.

METHODS:

A total of 124 patients with chronic hepatitis C virus infection underwent simultaneous laparoscopy-guided biopsies of the right and left hepatic lobes. Formalin-fixed paraffin-embedded sections were stained with hematoxylin and eosin and with trichrome. The slides were blindly coded and randomly divided among two hepatopathologists. Inflammation and fibrosis were scored according to the standard grading (inflammation) and staging (fibrosis) method based on the modified Scheuer system. Following the interpretation, the slides were uncoded to compare the results of the right and left lobes. Fifty of the samples were blindly resubmitted to each of the pathologists to determine the intraobserver variation.

RESULTS:

Thirty of 124 patients (24.2%) had a difference of at least one grade, and 41 of 124 patients (33.1%) had a difference of at least one stage between the right and left lobes. In 18 patients (14.5%), interpretation of cirrhosis was given in one lobe, whereas stage 3 fibrosis was given in the other. A difference of two stages or two grades was found in only three (2.4%) and two (1.6%) patients, respectively. Of the 50 samples that were examined twice, the grading by each pathologist on the second examination differed from the first examination in 0% and 4%, and the staging differed in 6% and 10%, respectively. All observed variations were of one grade or one stage.

CONCLUSIONS:

Liver biopsy samples taken from the right and left hepatic lobes differed in histological grading and staging in a large proportion of chronic hepatitis C virus patients; however, differences of more than one stage or grade were uncommon. A sampling error may have led to underdiagnosis of cirrhosis in 14.5% of the patients. These differences could not be attributed to intraobserver variation, which appeared to be low.
A Simple Noninvasive Index Can Predict Both Significant Fibrosis and Cirrhosis in Patients With Chronic Hepatitis C

Chun-Tao Wai, Joel K. Greenson, Robert J. Fontana, John D. Kalbfleisch, Jorge A. Marrero, Hari S. Conjeevaram, and Anna S.-F. Lok

Information on the stage of liver fibrosis is essential in managing chronic hepatitis C (CHC) patients. However, most models for predicting liver fibrosis are complicated and separate formulas are needed to predict significant fibrosis and cirrhosis. The aim of our study was to construct one simple model consisting of routine laboratory data to predict both significant fibrosis and cirrhosis among patients with CHC. Consecutive treatment-naive CHC patients who underwent liver biopsy over a 25-month period were divided into 2 sequential cohorts: training set (n = 192) and validation set (n = 78). The best model for predicting both significant fibrosis (Ishak score ≥ 3) and cirrhosis in the training set included platelets, aspartate aminotransferase (AST), and alkaline phosphatase with an area under ROC curves (AUC) of 0.82 and 0.92, respectively. A novel index, AST to platelet ratio index (APRI), was developed to amplify the opposing effects of liver fibrosis on AST and platelet count. The AUC of APRI for predicting significant fibrosis and cirrhosis were 0.80 and 0.89, respectively, in the training set. Using optimized cut-off values, significant fibrosis could be predicted accurately in 51% and cirrhosis in 81% of patients. The AUC of APRI for predicting significant fibrosis and cirrhosis in the validation set were 0.88 and 0.94, respectively.

In conclusion, our study showed that a simple index using readily available laboratory results can identify CHC patients with significant fibrosis and cirrhosis with a high degree of accuracy. Application of this index may decrease the need for staging liver biopsy specimens among CHC patients. (HEPATOLOGY 2003;38:518-526.)

Histologic examination of the liver is an integral part of the evaluation of patients with chronic hepatitis C (CHC). Knowledge of the stage of liver fibrosis is essential for prognostication and decisions on antiviral treatment. CHC patients with no or minimal fibrosis at presentation appear to progress slowly and treatment possibly could be delayed or withheld. On the other hand, patients with significant fibrosis (i.e., septal or bridging fibrosis) progress almost invariably to cirrhosis over a 10- to 20-year period so antiviral treatment should be strongly considered. For patients with cirrhosis, surveillance for hepatocellular carcinoma and gastroesophageal varices should be considered also.

Liver biopsy is currently the gold standard in assessing liver histology. Although percutaneous liver biopsy is in general a safe procedure, it is costly and does carry a small risk for complication. In addition, there could be sampling error because only 1/50,000 of the organ is sampled. Furthermore, inter- and intraobserver discrepancies of 10% to 20% in assessing hepatic fibrosis have been reported, which may lead to understaging of cirrhosis. Hence, there is a need to develop accurate and reliable noninvasive means to assess the severity of hepatic fibrosis.

Noninvasive approaches to assess histology in CHC patients include clinical symptoms and signs, routine laboratory tests, serum markers of fibrosis and inflammation, quantitative assays of liver function, and radiologic imag-
A Simple Noninvasive Index Can Predict Both Significant Fibrosis and Cirrhosis in Patients With Chronic Hepatitis C

Chun-Tao Wai, Joel K. Greenson, Robert J. Fontana, John D. Kalbfleisch, Jorge A. Marrero, Hari S. Conjeevaram, and Anna S.-F. Lok

Information on the stage of liver fibrosis is essential in managing chronic hepatitis C (CHC) patients. However, most models for predicting liver fibrosis are complicated and separate formulas are needed to predict significant fibrosis and cirrhosis. The aim of our study was to construct one simple model consisting of routine laboratory data to predict both significant fibrosis and cirrhosis among patients with CHC. Consecutive treatment-naive CHC patients who underwent liver biopsy over a 25-month period were divided into 2 sequential cohorts: training set (n = 192) and validation set (n = 78). The best model for predicting both significant fibrosis (Ishak score ≥ 3) and cirrhosis in the training set included platelets, aspartate aminotransferase (AST), and alkaline phosphatase with an area under ROC curves (AUC) of 0.82 and 0.92, respectively. A novel index, AST to platelet ratio index (APRI), was developed to amplify the opposing effects of liver fibrosis on AST and platelet count. The AUC of APRI for predicting significant fibrosis and cirrhosis were 0.80 and 0.89, respectively, in the training set. Using optimized cut-off values, significant fibrosis could be predicted accurately in 51% and cirrhosis in 81% of patients. The AUC of APRI for predicting significant fibrosis and cirrhosis in the validation set were 0.88 and 0.94, respectively. In conclusion, our study showed that a simple index using readily available laboratory results can identify CHC patients with significant fibrosis and cirrhosis with a high degree of accuracy. Application of this index may decrease the need for staging liver biopsy specimens among CHC patients. (HEPATOLOGY 2003;38:518-526.)

Histologic examination of the liver is an integral part of the evaluation of patients with chronic hepatitis C (CHC). Knowledge of the stage of liver fibrosis is essential for prognostication and decisions on antiviral treatment. CHC patients with no or minimal fibrosis at presentation appear to progress slowly and treatment possibly could be delayed or withheld. On the other hand, patients with significant fibrosis (i.e., septal or bridging fibrosis) progress almost invariably to cirrhosis over a 10- to 20-year period so antiviral treatment should be strongly considered. For patients with cirrhosis, surveillance for hepatocellular carcinoma and gastroesophageal varices should be considered also.

Liver biopsy is currently the gold standard in assessing liver histology. Although percutaneous liver biopsy is in general a safe procedure, it is costly and does carry a small risk for complication. In addition, there could be sampling error because only 1/50,000 of the organ is sampled. Furthermore, inter- and intraobserver discrepancies of 10% to 20% in assessing hepatic fibrosis have been reported, which may lead to understaging of cirrhosis. Hence, there is a need to develop accurate and reliable noninvasive means to assess the severity of hepatic fibrosis.

Noninvasive approaches to assess histology in CHC patients include clinical symptoms and signs, routine laboratory tests, serum markers of fibrosis and inflammation, quantitative assays of liver function, and radiologic imag-
A Simple Noninvasive Index Can Predict Both Significant Fibrosis and Cirrhosis in Patients With Chronic Hepatitis C

Chun-Tao Wai, Joel K. Greenson, Robert J. Fontana, John D. Kalbleisch, Jorge A. Marrero, Hari S. Conjeevaram, and Anna S.-F. Lok

Information on the stage of liver fibrosis is essential in managing chronic hepatitis C (CHC) patients. However, most models for predicting liver fibrosis are complicated and separate formulas are needed to predict significant fibrosis and cirrhosis. The aim of our study was to construct one simple model consisting of routine laboratory data to predict both significant fibrosis and cirrhosis among patients with CHC. Consecutive treatment-naive CHC patients who underwent liver biopsy over a 25-month period were divided into 2 sequential cohorts: training set (n = 192) and validation set (n = 78). The best model for predicting both significant fibrosis (Ishak score ≥ 3) and cirrhosis in the training set included platelets, aspartate aminotransferase (AST), and alkaline phosphatase with an area under ROC curves (AUC) of 0.82 and 0.92, respectively. A novel index, AST to platelet ratio index (APRI), was developed to amplify the opposing effects of liver fibrosis on AST and platelet count. The AUC of APRI for predicting significant fibrosis and cirrhosis were 0.80 and 0.89, respectively, in the training set. Using optimized cut-off values, significant fibrosis could be predicted accurately in 51% and cirrhosis in 81% of patients. The AUC of APRI for predicting significant fibrosis and cirrhosis in the validation set were 0.88 and 0.94, respectively. In conclusion, our study showed that a simple index using readily available laboratory results can identify CHC patients with significant fibrosis and cirrhosis with a high degree of accuracy. Application of this index may decrease the need for staging liver biopsy specimens among CHC patients. (HEPATOLOGY 2003;38:518-526.)

Histologic examination of the liver is an integral part of the evaluation of patients with chronic hepatitis C (CHC). Knowledge of the stage of liver fibrosis is essential for prognostication and decisions on antiviral treatment. CHC patients with no or minimal fibrosis at presentation appear to progress slowly and treatment possibly could be delayed or withheld. On the other hand, patients with significant fibrosis (i.e., septal or bridging fibrosis) progress almost invariably to cirrhosis over a 10- to 20-year period so antiviral treatment should be strongly considered. For patients with cirrhosis, surveillance for hepatocellular carcinoma and gastroesophageal varices should be considered also.

Liver biopsy is currently the gold standard in assessing liver histology. Although percutaneous liver biopsy is in general a safe procedure, it is costly and does carry a small risk for complication. In addition, there could be sampling error because only 1/50,000 of the organ is sampled. Furthermore, inter- and intraobserver discrepancies of 10% to 20% in assessing hepatic fibrosis have been reported, which may lead to understaging of cirrhosis. Hence, there is a need to develop accurate and reliable noninvasive means to assess the severity of hepatic fibrosis.

Noninvasive approaches to assess histology in CHC patients include clinical symptoms and signs, routine laboratory tests, serum markers of fibrosis and inflammation, quantitative assays of liver function, and radiologic imag-
The development of noninvasive markers of liver fibrosis is a clinical and research priority. The aspartate aminotransferase-to-platelet ratio index (APRI) is a promising tool with limited expense and widespread availability. Our objective was to systematically review the performance of the APRI in hepatitis C virus (HCV)–infected patients. Random effects meta-analyses and areas under summary receiver operating characteristic curves (AUC) were examined to characterize APRI accuracy for significant fibrosis (stages 2-4) and cirrhosis. In 22 studies (n = 4,266), the summary AUCs of the APRI for significant fibrosis and cirrhosis were 0.76 [95% confidence interval (CI), 0.74-0.79] and 0.82 (95% CI, 0.79-0.86), respectively. For significant fibrosis, an APRI threshold of 0.5 was 81% sensitive and 50% specific. At a 40% prevalence of significant fibrosis, this threshold had a negative predictive value (NPV) of 80%, but could reduce the necessity of liver biopsy by only 35%. For cirrhosis, a threshold of 1.0 was 76% sensitive and 71% specific. At a 15% cirrhosis prevalence, the NPV of this threshold was 91%. Higher APRI thresholds had suboptimal positive predictive values except in settings with a high prevalence of cirrhosis. APRI accuracy was not affected by the prevalence of advanced fibrosis, or study and biopsy quality. However, the accuracy for cirrhosis was greater in studies including human immunodeficiency virus (HIV)/HCV–co-infected patients. Conclusion: The major strength of the APRI is the exclusion of significant HCV-related fibrosis. Future studies of novel markers should demonstrate improved accuracy and cost-effectiveness compared with this economical and widely available index. (HEPATOLOGY 2007;46:912-921.)

Abbreviations: APRI, aspartate aminotransferase to platelet ratio index; AST, aspartate aminotransferase; AUC, area under the summary receiver operating characteristic curve; CI, confidence interval; DOR, diagnostic odds ratio; HCV, hepatitis C virus; HIV, human immunodeficiency virus; NPV, negative predictive value; PPV, positive predictive value; SROC, summary receiver operating characteristic.

From the Liver Unit, Division of Gastroenterology, Department of Medicine; University of Calgary, Calgary, Alberta, Canada.

Received April 10, 2007; accepted May 28, 2007.

Dr. Myers is funded by a Clinical Investigator Award from the Alberta Heritage Foundation for Medical Research. Dr. Shaheen is supported by a grant from the Canadian Liver Foundation.

*Both authors contributed equally to this work.

Address reprint requests to: Dr. Robert P. Myers, Liver Unit, University of Calgary, G126, 3330 Hospital Drive NW, Calgary, AB, Canada T2N 4N1.
E-mail: rpmyers@ucalgary.ca; fax: 403-210-9368.

Copyright © 2007 by the American Association for the Study of Liver Diseases.
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/hep.21835

Potential conflict of interest: The authors have no competing interests to disclose. Although Dr. Myers has published manuscripts on the FibroTest, he has never had a relationship with Biopredictive, the company marketing this test.
The development of noninvasive markers of liver fibrosis is a clinical and research priority. The aspartate aminotransferase-to-platelet ratio index (APRI) is a promising tool with limited expense and widespread availability. Our objective was to systematically review the performance of the APRI in hepatitis C virus (HCV)–infected patients. Random effects meta-analyses and areas under summary receiver operating characteristic curves (AUC) were examined to characterize APRI accuracy for significant fibrosis (stages 2-4) and cirrhosis. In 22 studies (n = 4,266), the summary AUCs of the APRI for significant fibrosis and cirrhosis were 0.76 [95% confidence interval (CI), 0.74-0.79] and 0.82 (95% CI, 0.79-0.86), respectively. For significant fibrosis, an APRI threshold of 0.5 was 81% sensitive and 50% specific. At a 40% prevalence of significant fibrosis, this threshold had a negative predictive value (NPV) of 80%, but could reduce the necessity of liver biopsy by only 35%. For cirrhosis, a threshold of 1.0 was 76% sensitive and 71% specific. At a 15% cirrhosis prevalence, the NPV of this threshold was 91%. Higher APRI thresholds had suboptimal positive predictive values except in settings with a high prevalence of cirrhosis. APRI accuracy was not affected by the prevalence of advanced fibrosis, or study and biopsy quality. However, the accuracy for cirrhosis was greater in studies including human immunodeficiency virus (HIV)/HCV–co-infected patients. Conclusion: The major strength of the APRI is the exclusion of significant HCV-related fibrosis. Future studies of novel markers should demonstrate improved accuracy and cost-effectiveness compared with this economical and widely available index. (HEPATOLOGY 2007;46:912-921.)
At the lower recommended threshold of 1.0, the summary sensitivity and specificity were 76% (95% CI 68-82%) and 71% (69-73%), respectively (Table 2). At a more specific threshold of 2.0, these figures were 49% (95% CI 43-55%) and 91% (90-93%), respectively. At the 15% prevalence of cirrhosis observed in the included studies, the estimated PPV and NPV of the 1.0 threshold were 32% and 94%, respectively. At the 2.0 threshold, the estimated PPV and NPV were 50% and 91%, respectively. The inversely proportional relationship between PPV and NPV for these thresholds at variable cirrhosis prevalence rates is illustrated in Fig. 6.

According to the meta-regression analysis, APRI accuracy for the detection of cirrhosis was greater in studies with a higher proportion of males ($P = 0.001$), younger participants ($P = 0.04$), and HIV/HCV–co-infected patients ($P = 0.03$). The DOR for cirrhosis was 44.9 (95% CI, 13.1-153.6) in the 2 studies including HIV/HCV–co-infected patients versus 9.5 (6.9-13.1) in the studies including only HCV-monoinfected participants. The other covariates were not significant (data not shown). An analysis for funnel plot asymmetry suggested possible publication bias for the prediction of cirrhosis ($P < 0.0005$).

**Discussion**

In this systematic review, we summarize the diagnostic accuracy of the APRI for the prediction of HCV-related fibrosis. In an era in which the number of fibrosis markers is growing rapidly, many clinicians, patients, researchers, and policy makers are confused as to the optimal measure. Because the APRI is based on routinely performed, inexpensive laboratory parameters, it is potentially the ideal tool because most HCV-infected patients reside in regions with limited healthcare resources. Our systematic review suggests that the accuracy of the APRI is perhaps less than initially described. In Wai and colleagues’ original study, the AUC for significant fibrosis and cirrhosis in the training and validation cohorts were 0.80 to 0.88 and 0.89 to 0.94, respectively. In our systematic review, the APRI had modest accuracy for significant fibrosis (AUC 0.76; DOR 6.69). Because these point estimates are difficult to translate into clinical practice, we calculated summary sensitivities and specificities. Moreover, we provide predictive values at varying fibrosis prevalence rates, with the aim of providing practically useful information for clinicians in a variety of practice settings (Figs. 3 and 6). According to these analyses, the primary strength of the APRI is the exclusion of significant fibrosis. Based on our bivariate meta-analysis, the 0.5 threshold was 81% sensitive and 50% specific. Assuming a 47% prevalence of significant fibrosis (as observed in the included studies), this translates into an estimated PPV of 59% and NPV of 75%. Although these predictive values appear suboptimal, the NPV was more acceptable in lower prevalence settings, such as typically observed in community-based cohorts. For example, at a prevalence of 30% to 40%, the estimated NPV ranged from 80% to 86%; at the same time, the PPV did not exceed 52% (Fig. 3). On the contrary, a cutoff of 1.5 was more specific (91%) but less sensitive (35%). The PPV of this threshold did not reach 80% until the prevalence of significant fibrosis exceeded 50%, which is typically observed only in referral centers. Based on these analyses, we suggest that

**Fig. 3.** NPV (solid lines) and PPV (dotted lines) of the APRI for significant fibrosis according to prevalence. Curves are illustrated for the recommended APRI thresholds of 0.5 (black lines) and 1.5 (gray lines).
APRI (0.76), FibroTest (0.79), FibroMeter (0.78), and HepaScore (0.76) were not significantly different. With respect to transient elastography, Castera et al. did not find a statistically significant difference between the FibroScan and APRI for METAVIR F2-F4 fibrosis (AUC 0.83 vs. 0.78), although the FibroScan was more accurate for cirrhosis (0.95 vs. 0.83). This large insignificantly difference between the APRI, which is inexpensive and available for all HCV-infected patients, versus other more costly and specialized fibrosis measures underscores 2 important points. First, before the latter tools are widely used, their incremental cost-effectiveness, or the tradeoff between increased accuracy (if there is any) and cost, should be demonstrated. Second, further research must identify novel markers with improved accuracy over conventional measures. These studies should include the APRI as a benchmark for diagnostic performance.

A strength of our review is our analysis for heterogeneity in APRI accuracy, which was significant for the primary outcome. Despite examining 9 patient and study-specific covariates, including study quality, fibrosis stage distribution, and inclusion of HIV/HCV-co-infected patients, we could not explain this finding. Biopsy length, specifically, was not significant in the meta-regression analyses, although this has been reported to affect the accuracy of some fibrosis markers. Because most of the studies did not exclude suboptimal biopsy specimens (or did not report biopsy characteristics), we would encourage future investigators to be rigorous in their assessment of the quality of liver biopsies, ensuring adequate length and portal tract number. Although these negative findings may reflect a type II error, we hypothesize that they likely relate to “unquantifiable” differences between studies such as the quality of histopathologists, heterogeneous patient populations, and different assays and reference ranges for AST (reported in only 1 of the studies). An individual patient data meta-analysis would be useful to explore these issues. Interestingly, the APRI was more accurate for the identification of cirrhosis in HIV/HCV-co-infected patients. This finding was surprising because we hypothesized that its accuracy may be diminished in co-infected patients because of HIV-related or antiretroviral-related thrombocytopenia. Because this analysis included only 2 studies of co-infected patients, it warrants confirmation.

Our systematic review has several limitations. Although we identified 22 eligible studies including more than 4,200 patients, our funnel plot analysis for cirrhosis suggested the possibility of publication or other small sample size-related biases. This may relate to our inclusion of only published manuscripts. Potentially eligible abstracts were identified, but most did not report sufficient data, and many were subsequently published, often in altered form. Even so, because tests for publication bias have not been fully validated in meta-analyses of diagnostic test accuracy, these findings must be interpreted cautiously. A second limitation is that we have focused our analysis on HCV-infected patients only. The APRI has been examined in hepatitis B, but the few published studies suggest reduced accuracy. Therefore, to avoid introducing further heterogeneity, we restricted our analysis to HCV. Ideally, we would have also examined other test characteristics such as cost-effectiveness and impact on clinical outcomes. Because of a scarcity of publications, we could not address these important issues. However, Ngo et al. recently described an association between APRI scores and 5-year survival without HCV-related complications (AUC 0.82). In another study, the APRI 6 months after the end of antiviral therapy was highly predictive of hepatocellular carcinoma development and survival (AUCs 0.87). Finally, we have considered the accuracy of the APRI in isolation, rather than in combination with other measures. As reported by Sebastiani et al., a stepwise algorithm including the APRI and other markers may improve diagnostic performance. Because of an absence of similar publications, we could not examine this issue.

In summary, our systematic review suggests that the APRI has moderate diagnostic utility for the prediction of fibrosis in HCV-infected patients. Its major role appears to be the exclusion of significant fibrosis and cirrhosis, which can be achieved with acceptable accuracy in at least one third and three quarters of patients, respectively. Future studies of novel fibrosis markers should demonstrate improved accuracy and cost-effectiveness compared with this simple, economical, and widely available index.

References

A Prospective Analysis of the Prognostic Value of Biomarkers (FibroTest) in Patients with Chronic Hepatitis C

Yen Ngo,1 Mona Munteanu,2 Djamilia Messous,3 Frederic Charlotte,4 Françoise Imbert-Bismut,3 Dominique Thabut,1 Pascal Lebray,1 Vincent Thibault,5 Yves Benhamou,1 Joseph Moussalli,1 Vlad Ratziu,1 and Thierry Poynard1*

Background: FibroTest, a noninvasive method of measuring biomarkers of liver fibrosis, is an alternative to liver biopsy for determining the severity of chronic hepatitis C virus (HCV) infection. We compared the 5-year prognostic value of the FibroTest with biopsy staging for predicting cirrhosis decompensation and survival in patients with chronic HCV infection.

Methods: Fibrosis stage was assessed on the same day by FibroTest and biopsy in a prospective cohort of 537 patients. Disease classification at baseline was 157 patients with severe fibrosis (FibroTest > 0.58), 137 with moderate fibrosis (FibroTest 0.32–0.58), and 243 with no or minimal fibrosis (FibroTest < 0.32).

Results: In 64 untreated patients with severe fibrosis, survival without HCV complications was 73% [95% confidence interval (CI), 59%–86%; 13 complications], and survival without HCV-related death was 85% (95% CI, 73%–96%; 7 HCV deaths). Survival rates were higher in patients with moderate fibrosis, 199% (95% CI, 97%–100%; 1 complication; P < 0.001) and 100% (no HCV death; P < 0.001) for patients with and without HCV-related complications, respectively, and in patients with minimal fibrosis [100% (no complication; P < 0.001 vs severe) and 100% (no HCV death; P < 0.001 vs severe), respectively]. FibroTest was a better predictor than biopsy staging for HCV complications, with area under the ROC curves (AUROC) = 0.96 (95% CI, 0.93%–0.97%) vs 0.91 (95% CI, 0.85%–0.94%; P = 0.01), respectively; it was also a better predictor for HCV deaths: AUROC = 0.96 (95% CI, 0.93%–0.98%) vs 0.87 (95% CI, 0.70%–0.94%; P = 0.046), respectively. The prognostic value of FibroTest was still significant (P < 0.001) in multivariate analyses after taking into account histology, treatment, alcohol consumption, and HIV coinfection.

Conclusion: The FibroTest measurement of HCV biomarkers has a 5-year prognostic value similar to that of liver biopsy.

© 2006 American Association for Clinical Chemistry

Finding the best method to evaluate and manage patients infected with the hepatitis C virus (HCV)6 continues to be a challenge (1,2). Liver biopsy for determining disease grade and stage has limitations (3–5) and risks (6); noninvasive alternatives to liver biopsy in patients infected with HCV (7) include 2 combinations of simple serum biochemical markers: FibroTest (FT) (Biopredictive) for the assessment of fibrosis, and ActiTest (AT) (Biopredictive) for the assessment of necroinflammatory activity (8–16). With biopsy as the standard of reference, the diagnostic value of FT for a diagnosis of advanced fibrosis (bridging fibrosis), as estimated by the area under the ROC curve (AUROC), is 0.73–0.86 (8). AUROC values inside the range of other studies have been reported, but the usefulness of FT compared with biopsy was not

1 Service d’Hépato-Gastroentérologie, Groupe Hospitalier Pitié-Salpêtrière, CNRS, Paris, France.
2 Biopredictive, Paris, France.
3 Laboratoire de Biochimie, Groupe Hospitalier Pitié-Salpêtrière, Paris, France.
5 Laboratoire de Virologie, Groupe Hospitalier Pitié-Salpêtrière, Paris, France.
6 Address correspondence to this author at: Service d’Hépato-Gastroentérologie, Groupe Hospitalier Pitié-Salpêtrière, 47 Boulevard de l’Hôpital 75651 Paris Cedex 13, France. Fax (33-1)-42-16-14-27; e-mail tpoynard@teaser.fr.

Received March 30, 2006; accepted August 2, 2006.
Previously published online at DOI: 10.1373/clinchem.2006.070961
A Prospective Analysis of the Prognostic Value of Biomarkers (FibroTest) in Patients with Chronic Hepatitis C

Yen Ngo,1 Mona Munteanu,2 Djamila Messous,3 Frederic Charlotte,4 Françoise Imbert-Bismut,3 Dominique Thabut,1 Pascal Lebray,1 Vincent Thibault,5 Yves Benhamou,1 Joseph Moussalli,1 Vlad Ratziu,1 and Thierry Poynard1*

Background: FibroTest, a noninvasive method of measuring biomarkers of liver fibrosis, is an alternative to liver biopsy for determining the severity of chronic hepatitis C virus (HCV) infection. We compared the 5-year prognostic value of the FibroTest with biopsy staging for predicting cirrhosis decompensation and survival in patients with chronic HCV infection.

Methods: Fibrosis stage was assessed on the same day by FibroTest and biopsy in a prospective cohort of 537 patients. Disease classification at baseline was 157 patients with severe fibrosis (FibroTest > 0.58), 137 with moderate fibrosis (FibroTest 0.32–0.58), and 243 with no or minimal fibrosis (FibroTest < 0.32).

Results: In 64 untreated patients with severe fibrosis, survival without HCV complications was 73% [95% confidence interval (CI), 59%–086%; 13 complications], and survival without HCV-related death was 85% (95% CI, 73%–96%; 7 HCV deaths). Survival rates were higher in patients with moderate fibrosis, 99% (95% CI, 97%–100%; 1 complication; P < 0.001) and 100% (no HCV death; P < 0.001) for patients with and without HCV-related complications, respectively, and in patients with minimal fibrosis [100% (no complication; P < 0.001 vs severe) and 100% (no HCV death; P < 0.001 vs severe), respectively]. FibroTest was a better predictor than biopsy staging for HCV complications, with area under the ROC curves (AUROC) = 0.96 (95% CI, 0.93%–0.97%) vs 0.91 (95% CI, 0.85%–0.94%; P = 0.01), respectively; it was also a better predictor for HCV deaths: AUROC = 0.96 (95% CI, 0.93%–0.98%) vs 0.87 (95% CI, 0.70%–0.94%; P = 0.046), respectively.

Conclusion: The FibroTest measurement of HCV biomarkers has a 5-year prognostic value similar to that of liver biopsy.

© 2006 American Association for Clinical Chemistry

Finding the best method to evaluate and manage patients infected with the hepatitis C virus (HCV) continues to be a challenge (1, 2). Liver biopsy for determining disease grade and stage has limitations (3–5) and risks (6); noninvasive alternatives to liver biopsy in patients infected with HCV (7) include 2 combinations of simple serum biochemical markers: FibroTest (FT) (Biopredictive) for the assessment of fibrosis, and ActiTest (AT) (Biopredictive) for the assessment of necroinflammatory activity (8–16). With biopsy as the standard of reference, the diagnostic value of FT for a diagnosis of advanced fibrosis (bridging fibrosis), as estimated by the area under the ROC curve (AUROC), is 0.73–0.86 (8). AUROC values inside the range of other studies have been reported, but the usefulness of FT compared with biopsy was not

1 Service d’Hépato-Gastroentérologie, Groupe Hospitalier Pitié-Salpêtrière, CNRS, Paris, France.
2 Biopredictive, Paris, France.
3 Laboratoire de Biochimie, Groupe Hospitalier Pitié-Salpêtrière, Paris, France.
5 Laboratoire de Virologie, Groupe Hospitalier Pitié-Salpêtrière, Paris, France.
* Address correspondence to this author at: Service d’Hépato-Gastroentérologie, Groupe Hospitalier Pitié-Salpêtrière, 47 Boulevard de l’Hôpital 75651 Paris Cedex 13, France. Fax (33-1)-42-16-14-27; e-mail tpoynard@teaser.fr.
Received March 30, 2006; accepted August 2, 2006.
Previously published online at DOI: 10.1373/clinchem.2006.070961

Nonstandard abbreviations: HCV, hepatitis C virus; FT, FibroTest; AT, ActiTest; AUROC, area under the ROC curve; F0, no fibrosis; F1, portal fibrosis without septa; F2, few septa; F3, numerous septa without cirrhosis; F4, cirrhosis; HCC, hepatocellular carcinoma; CI, confidence interval.
Chronic infection by hepatitis C virus (HCV) affects 170 million people worldwide. In European countries the overall prevalence in adults ranges from 1% to 2%, and more than five million Europeans are thus infected. The high prevalence of HCV infection, together with the risk of severe complications which include cirrhosis and hepatocellular carcinoma, makes this a major public health problem.

HCV was discovered more than 10 years ago, and considerable progress has since been made in our knowledge of the virus, its modes of transmission, the natural history of the infection, and also patient management, including specific therapy. Several consensus conferences have been organised in the last five years, in particular the French Association for the Study of the Liver in January 1997; the National Institutes of Health in March 1997; and the European Association for the Study of the Liver (EASL) in February 1999. Major changes have since occurred in the epidemiology of the infection and in patient management.

Firstly, recent surveys show profound changes in the characteristics of newly diagnosed patients: the proportion of patients with mild chronic hepatitis at diagnosis has increased, and the absolute number of severe cases—that is, with cirrhosis and hepatocellular carcinoma—has also increased, a large proportion of patients having been infected several decades previously. The modes of HCV transmission have also evolved, with a gradual reduction in the proportion of cases related to transfusion and an increase in the proportion related to intravenous drug use. These changes largely account for the observed changes in the HCV genotype profile, characterised by an increase in the prevalence of genotype 3 infection, which is associated with better response to treatment.

Secondly, since the 1999 EASL consensus conference, new advances have been made in the treatment of chronic hepatitis C, with the introduction of pegylated interferon (PEG IFN) which, in combination with ribavirin, gives an overall rate of sustained virological response of 55% (approximately 45% for HCV genotype 1 and 80% for HCV genotypes 2 and 3). It has also been shown in long term follow up studies that patients with sustained virological responses can be considered cured of the infection.

Thirdly, the information provided to HCV infected individuals has considerably improved, making patients more aware and better placed to participate in the management of their disease. For example, knowing that treatment efficacy has improved, some patients may insist they receive curative treatment even if they have little liver damage. More generally, the treatment target is tending to shift from the clinical disease itself (that is, the hepatic lesions of chronic hepatitis C) to the underlying viral infection. This change may have a significant impact on the indications of pretreatment investigations, and especially liver biopsy.

Fourthly, some patients have extensive fibrosis or cirrhosis at diagnosis. When treatment fails to induce a virological response, the question arises as to the need for "maintenance treatment" aimed at limiting disease progression and the risk of hepatocellular carcinoma.

Fifthly, the management of patients infected both by HCV and by human immunodeficiency virus (HIV) has changed with the improved efficacy of antiretroviral therapy, which has transformed the prognosis of HIV disease and, as a consequence, revealed the impact of HCV infection on mortality in this particular setting.

Finally, the management of intravenous drug users has markedly evolved, especially with the introduction of replacement therapy and a trend towards community based "holistic" approaches.

With the aim of assessing these changes and optimising the management strategy for HCV infected patients, a new consensus conference was held in Paris on 27–28 February 2002. The recommendations elaborated in this conference are directed towards all physicians caring for patients with HCV infection, including internists, general practitioners, and specialists in hepatology, gastroenterology, and infectious diseases.

This consensus conference followed the methodological rules developed by the French Agence Nationale d’Accréditation et d’Évaluation en Santé. It consisted of an organising committee of academic scientists, a working group who prepared the information for the jury, the experts themselves, and an independent jury composed of persons not involved in the study of HCV or in patient management.

As required by the organising committee, this conference was entirely financed by funds from the public sector. It took the form of:

(a) presentations by experts working in areas relevant to the consensus questions, during a public session;
(b) questions from and discussions among conference attendees; and
(c) closed deliberations by the jury, followed by the writing of conclusions and recommendations (available from http://afef.meditis.net).

The five questions put to the jury were:

1. Which patients should be treated?
2. What are the most appropriate investigations before treatment?
3. Which is the optimal treatment?
4. How to monitor treated patients?
5. How to monitor untreated patients?

WHICH PATIENTS SHOULD BE TREATED?

The introduction of new drugs or new drug combinations with better efficacy may modify the treatment indications in

Abbreviations: HCV, hepatitis C virus; PEG IFN, pegylated interferon; HIV, human immunodeficiency virus
Treatment of hepatitis C. The 2002 French consensus

D Dhumieux, P Marcellin, E Lerebours

Chronic infection by hepatitis C virus (HCV) affects 170 million people worldwide. In European countries the overall prevalence in adults ranges from 1% to 2%, and more than five million Europeans are thus infected. The high prevalence of HCV infection, together with the risk of severe complications which include cirrhosis and hepatocellular carcinoma, makes this a major public health problem.

HCV was discovered more than 10 years ago, and considerable progress has since been made in our knowledge of the virus, its modes of transmission, the natural history of the infection, and also patient management, including specific therapy. Several consensus conferences have been organised in the last five years, in particular the French Association for the Study of the Liver in January 1997, the National Institutes of Health in March 1997, and the European Association for the Study of the Liver (EASL) in February 1999. Major changes have since occurred in the epidemiology of the infection and in patient management.

Firstly, recent surveys show profound changes in the characteristics of newly diagnosed patients: the proportion of patients with mild chronic hepatitis at diagnosis has increased, and the absolute number of severe cases—that is, with cirrhosis and hepatocellular carcinoma—has also increased, a large proportion of patients having been infected several decades previously. The modes of HCV transmission have also evolved, with a gradual reduction in the proportion of cases related to transfusion and an increase in the proportion related to intravenous drug use. These changes largely account for the observed changes in the HCV genotype profile, characterised by an increase in the prevalence of genotype 3 infection, which is associated with a better response to treatment.

Secondly, since the 1999 EASL consensus conference, new advances have been made in the treatment of chronic hepatitis C, with the introduction of pegylated interferon (PEG IFN) which, in combination with ribavirin, gives an overall rate of sustained virological response of 55% (approximately 45% for HCV genotype 1 and 80% for HCV genotypes 2 and 3). It has also been shown in long term follow up studies that patients with sustained virological responses can be considered cured of the infection.

Thirdly, the information provided to HCV infected individuals has considerably improved, making patients more aware and better placed to participate in the management of their disease. For example, knowing that treatment efficacy has improved, some patients may insist they receive curative treatment even if they have little liver damage. More generally, the treatment target is tending to shift from the clinical disease itself (that is, the hepatic lesions of chronic hepatitis C) to the underlying viral infection. This change may have a significant impact on the indications of pretreatment investigations, and especially liver biopsy.

Fourthly, some patients have extensive fibrosis or cirrhosis at diagnosis. When treatment fails to induce a virological response, the question arises as to the need for “maintenance treatment” aimed at limiting disease progression and the risk of hepatocellular carcinoma.

Lastly, the management of patients infected both by HCV and by human immunodeficiency virus (HIV) has changed with the improved efficacy of antiretroviral therapy, which has transformed the prognosis of HIV disease and, as a consequence, revealed the impact of HCV infection on mortality in this particular setting.

Finally, the management of intravenous drug users has markedly evolved, especially with the introduction of replacement therapy and a trend towards community based “holistic” approaches.

With the aim of assessing these changes and optimising the management strategy for HCV infected patients, a new consensus conference was held in Paris on 27–28 February 2002. The recommendations elaborated in this conference are directed towards all physicians caring for patients with HCV infection, including internists, general practitioners, and specialists in hepatology, gastroenterology, and infectious diseases.

This consensus conference followed the methodological rules developed by the French Agence Nationale d’Accréditation et d’Évaluation en Santé. It consisted of an organising committee of academic scientists, a working group who prepared the information for the jury, the experts themselves, and an independent jury composed of persons not involved in the study of HCV or in patient management (see appendix).

As required by the organising committee, this conference was entirely financed by funds from the public sector. It took the form of:

(a) presentations by experts working in areas relevant to the consensus questions, during a public session;
(b) questions from and discussions among conference attendees; and
(c) closed deliberations by the jury, followed by the writing of conclusions and recommendations (available from http://afef.meditis.net and14).

The five questions put to the jury were:

1. Which patients should be treated?
2. What are the most appropriate investigations before treatment?
3. Which is the optimal treatment?
4. How to monitor treated patients?
5. How to monitor untreated patients?

WHICH PATIENTS SHOULD BE TREATED?

The introduction of new drugs or new drug combinations with better efficacy may modify the treatment indications in...
HCV infection. In this respect, the advent of the PEG IFN/ribavirin combination is important to consider. Although not excluding broader indications, the jury recalled that therapy must preferentially be offered to those patients with the best predicted risk-benefit ratio. Treatment indications remain based on histological assessment of hepatic lesions. Although high grade histological activity may be associated with accelerated disease progression, the severity of chronic hepatitis C is mainly defined by the stage of fibrosis. The jury thus considered that, whatever the grade of necroinflammatory activity, patients with stage F2 or F3 fibrosis (METAVIR scoring system) have the best potential risk-benefit ratio and are thus the best candidates for treatment.

With regard to patients with mild chronic hepatitis (F0 or F1) or chronic HCV infection associated with repeatedly normal transaminase levels (a situation in which histological lesions are also generally mild), the risk of disease progression is low and the long-term benefits of treatment have not yet been established. Previous consensus conferences did not recommend treatment for these patients. However, it was recently found that combined therapy was highly effective in patients infected with HCV genotype 2 or 3. Thus the jury recommended that patients with mild hepatitis should not be excluded from treatment, provided they are infected by HCV genotype 2 or 3.

Patients with cirrhosis generally responded poorly to previous treatments. This situation could change with the advent of new agents such as PEG IFN, which is effective in a significant proportion of cirrhotic patients. The jury recommended that these patients should not be excluded from treatment, and that some of them (those with a biochemical response to treatment but without viral eradication) qualify for “maintenance therapy” in an attempt to delay disease progression pending the development of new drugs. Specific doses of PEG IFN could not be recommended in the absence of relevant data. This maintenance therapy was also recommended for patients with extensive fibrosis (stage F3 in the METAVIR scoring system). The jury stressed that this strategy remains to be validated and that, whenever possible, such patients should be enrolled in clinical trials. This example clearly illustrates the fact that there are now two distinct treatment aims:

(a) ideally, to eradicate the virus—that is, to cure the infection; and
(b) to prevent, stabilise, or even improve hepatic lesions.

All meta-analyses of studies involving patients with acute hepatitis have shown that therapy reduces the rate of progression to chronicity, and previous consensus statements recommend that these patients should be treated. Two studies have now reported a very high success rate—better than 80%—in patients treated with relatively high doses of standard interferon, given daily (at least at the beginning of treatment). The jury thus recommended this treatment protocol, but stated that the effects of PEG IFN, alone or in combination with ribavirin, must rapidly be assessed with the aim of obtaining similar or even better efficacy.

In previous consensus conferences, HCV infected intravenous drug users were not considered as a treatment priority. Recent changes in the global management of these patients, including replacement therapy, are modifying this view. Ongoing excessive alcohol consumption, concomitant HIV or HBV infection, psychiatric disorders, and social precariousness frequently worsen the prognosis of HCV infected intravenous drug users, but the jury noted that they also have a number of favourable characteristics for treatment; indeed, diagnosis is generally made at a relatively young age, the prior duration of HCV infection is relatively short, histological lesions are usually mild, and the prevalence of genotype 3 infection is high. Given these factors, therapeutic indications in drug users could be broader than previously proposed.

The increased efficacy of antiretroviral drugs in HIV infection has also challenged previous restrictions on the indications for HCV treatment in HCV/HIV coinfected patients. Indeed, some patients who respond well to treatment for HIV infection may now die from HCV disease. The jury therefore recommended that coinfected patients with moderate to severe histological fibrosis should not be excluded from HCV treatment but stressed the risk of interactions between ribavirin and other nucleoside analogues, and the possible potentiation of antiretroviral hepatotoxicity by underlying liver disease.

**WHAT ARE THE MOST APPROPRIATE INVESTIGATIONS BEFORE TREATMENT?**

In addition to standard tests, special attention should be paid to extrahepatic manifestations, psychiatric disorders, HIV coinfection, excessive alcohol consumption and, as recently emphasised, excess body weight. As recommended by a previous conference, HCV genotypes and HCV RNA levels must be determined before treatment. The HCV genotype influences both the treatment indications and the therapeutic strategy itself, as current treatments are more effective and shorter in patients with HCV genotype 2 or 3 infection. HCV RNA quantitative assay, at least in patients treated with the PEG IFN/ribavirin combination, is not used to define the duration of treatment but may be useful in patients infected by HCV genotype 1 providing a baseline value on which to appreciate the early response, which is a good indicator of the likelihood of a sustained virological response.

A major point of discussion on the pretreatment workup was the place of liver biopsy. Clearly, liver biopsy is still the best way of evaluating fibrosis (the key parameter in prognostication and therapeutic decision making) and, as in previous conferences, the jury supported its use. However, given the good efficacy of treatment on some HCV genotypes, and reports that a fear of biopsy and its refusal by some patients may hamper access to treatment, the jury recommended that liver biopsy should not be mandatory when the decision to treat has already been taken and will not be affected by the histological result. This particularly concerns patients in whom the treatment aim is viral eradication, independently of histological lesions—that is:

(a) patients infected by HCV genotype 2 or 3, in whom efficacy is approximately 80% in clinical trials;
(b) women planning to become pregnant and wishing to avoid the (low) risk of transmitting HCV to their child;
(c) patients with symptomatic cryoglobulinaemia (viral eradication being crucial for symptom control); and
(d) HCV/HIV infected patients, when antiretroviral treatment can be postponed (the advantage in such cases is that treating HCV infection before HIV infection avoids interference between the drugs used for the two infections).

Serum markers of fibrosis have recently been proposed as an alternative to liver biopsy. Some are not used routinely, and all must be validated in large studies and by different groups before being adopted.
Identification of Chronic Hepatitis C Patients Without Hepatic Fibrosis by a Simple Predictive Model

Xavier Forns,1 Sergi Ampurdanès,1 Josep M. Llovet,1 John Aponte,2 Llorenç Quintó,2 Eva Martínez-Bauer,1 Miquel Bruguera,1 Jose Maria Sánchez-Tapias,1 and Juan Rodés1

Liver biopsy is required for staging hepatic fibrosis in patients with chronic hepatitis C, but it is an expensive procedure with occasional complications and poor patient acceptance. This cohort study was designed to assess the accuracy of a noninvasive method aimed to discriminate between patients with and without significant liver fibrosis (stages 2–4 versus 0–1). Clinically relevant variables were analyzed in a cohort of 476 consecutive untreated patients (estimation group, 351 patients; validation group, 125 patients) with chronic hepatitis C who underwent a liver biopsy. Multivariate analysis identified age, gamma glutamyl transpeptidase (GGT), cholesterol, platelet count, and prothrombin time as independent predictors of fibrosis. We constructed a model and a score system combining age, GGT, cholesterol, and platelet count that proved useful to identify patients without significant hepatic fibrosis. The area under the ROC curve was 0.86 for the estimation group and 0.81 for the validation group. Using the best cutoff score (less than 4.2), presence of significant fibrosis (F2 to F4) could be excluded with high accuracy (negative predictive value of 96%) in 125 (36%) of 351 patients. Similarly, it could be excluded with the same certainty in 49 (39%) of the 125 patients of the validation group. Only 2 patients with liver fibrosis stage 2 were incorrectly classified. In conclusion, a combination of easily accessible variables accurately predicts the absence of significant fibrosis, and might render liver biopsy unnecessary in more than one third of patients with chronic hepatitis C. (HEPATOLOGY 2002;36:986-992.)

In our geographic area 2% to 3% of the general population is chronically infected with the hepatitis C virus (HCV).1 Chronic hepatitis C is the most prevalent disease in hepatology clinics, accounting for more than 30% of the visits. In general, it is accepted that the diagnostic protocol of chronic hepatitis C includes a liver biopsy, particularly in patients with elevated amino transferase levels.2–5 Liver histology and, especially, fibrosis staging provide prognostic information6–7 and may be useful in deciding on therapeutic strategies in individual cases.2–3 However, liver biopsy is an invasive procedure that may cause undesirable events, such as pain in 20% to 30% of the cases, major complications in 0.5%, and even death.8 Other than the complications derived from the procedure and the frequent poor patient acceptance, the direct cost of such procedure is high.2–3

Nowadays, chronic hepatitis C is often recognized at an early stage of the disease, and liver fibrosis is mild or absent in about 80% of the patients undergoing a liver biopsy.6 Thus, the finding of surrogate markers of liver fibrosis/absence of fibrosis would be relevant to reduce the number of liver biopsies in patients with chronic hepatitis C. Imbert-Bismuth et al.9 have recently shown that a combination of biochemical markers of liver fibrosis can be useful to predict the presence or absence of fibrosis and, therefore, to reduce the number of liver biopsies in patients with chronic hepatitis C. However, some of these markers, such as α2 macroglobulin, haptoglobin, or apolipoprotein A1 are not routinely used in clinical practice. This is also the case for other proposed predictors of fibrosis, such as hyaluronate concentration or type III procollagen.10–12

Using variables easily available to clinicians, we have constructed and validated a model and a score system aimed to discriminate patients with substantial fibrosis...
Identification of Chronic Hepatitis C Patients Without Hepatic Fibrosis by a Simple Predictive Model

Xavier Forns,1 Sergi Ampurdanes,1 Josep M. Llovet,1 John Aponte,2 Llorenç Quintó,2 Eva Martínez-Bauer,1 Miquel Bruguera,1 Jose Maria Sánchez-Tapias,1 and Juan Rodes1

Liver biopsy is required for staging hepatic fibrosis in patients with chronic hepatitis C, but it is an expensive procedure with occasional complications and poor patient acceptance. This cohort study was designed to assess the accuracy of a noninvasive method aimed to discriminate between patients with and without significant liver fibrosis (stages 2-4 versus 0-1). Clinically relevant variables were analyzed in a cohort of 476 consecutive untreated patients (estimation group, 351 patients; validation group, 125 patients) with chronic hepatitis C who underwent a liver biopsy. Multivariate analysis identified age, gamma glutamyl transpeptidase (GGT), cholesterol, platelet count, and prothrombin time as independent predictors of fibrosis. We constructed a model and a score system combining age, GGT, cholesterol, and platelet count that proved useful to identify patients without significant hepatic fibrosis. The area under the ROC curve was 0.86 for the estimation group and 0.81 for the validation group. Using the best cutoff score (less than 4.2), presence of significant fibrosis (F2 to F4) could be excluded with high accuracy (negative predictive value of 96%) in 125 (36%) of 351 patients. Similarly, it could be excluded with the same certainty in 49 (39%) of the 125 patients of the validation group. Only 2 patients with liver fibrosis stage 2 were incorrectly classified. In conclusion, a combination of easily accessible variables accurately predicts the absence of significant fibrosis, and might render liver biopsy unnecessary in more than one third of patients with chronic hepatitis C. (HEPATOLOGY 2002;36:986-992.)

In our geographic area 2% to 3% of the general population is chronically infected with the hepatitis C virus (HCV).1 Chronic hepatitis C is the most prevalent disease in hepatology clinics, accounting for more than 30% of the visits. In general, it is accepted that the diagnostic protocol of chronic hepatitis C includes a liver biopsy, particularly in patients with elevated amino transferase levels.2-5 Liver histology and, especially, fibrosis staging provide prognostic information6,7 and may be useful in deciding on therapeutic strategies in individual cases.2,3 However, liver biopsy is an invasive procedure that may cause undesirable events, such as pain in 20% to 30% of the cases, major complications in 0.5%, and even death.8 Other than the complications derived from the procedure and the frequent poor patient acceptance, the direct cost of such procedure is high.2,3 Nowadays, chronic hepatitis C is often recognized at an early stage of the disease, and liver fibrosis is mild or absent in about 80% of the patients undergoing a liver biopsy.6 Thus, the finding of surrogate markers of liver fibrosis/absence of fibrosis would be relevant to reduce the number of liver biopsies in patients with chronic hepatitis C. Imbert-Bismuth et al.9 have recently shown that a combination of biochemical markers of liver fibrosis can be useful to predict the presence or absence of fibrosis and, therefore, to reduce the number of liver biopsies in patients with chronic hepatitis C. However, some of these markers, such as α2 macroglobulin, haptoglobin, or apolipoprotein A1 are not routinely used in clinical practice. This is also the case for other proposed predictors of fibrosis, such as hyaluronate concentration or type III procollagen.10-12

Using variables easily available to clinicians, we have constructed and validated a model and a score system aimed to discriminate patients with substantial fibrosis...
Liver biopsy is required for staging hepatic fibrosis in patients with chronic hepatitis C, but it is an expensive procedure with occasional complications and poor patient acceptance. This cohort study was designed to assess the accuracy of a noninvasive method aimed to discriminate between patients with and without significant liver fibrosis (stages 2-4 versus 0-1). Clinically relevant variables were analyzed in a cohort of 476 consecutive untreated patients (estimation group, 351 patients; validation group, 125 patients) with chronic hepatitis C who underwent a liver biopsy. Multivariate analysis identified age, gamma glutamyl transpeptidase (GGT), cholesterol, platelet count, and prothrombin time as independent predictors of fibrosis. We constructed a model and a score system combining age, GGT, cholesterol, and platelet count that proved useful to identify patients without significant hepatic fibrosis. The area under the ROC curve was 0.86 for the estimation group and 0.81 for the validation group. Using the best cutoff score (less than 4.2), presence of significant fibrosis (F2 to F4) could be excluded with high accuracy (negative predictive value of 96%) in 125 (36%) of 351 patients. Similarly, it could be excluded with the same certainty in 49 (39%) of the 125 patients of the validation group. Only 2 patients with liver fibrosis stage 2 were incorrectly classified. In conclusion, a combination of easily accessible variables accurately predicts the absence of significant fibrosis and might render liver biopsy unnecessary in more than one third of patients with chronic hepatitis C. (HEPATOLOGY 2002;36:986-992.)

In our geographic area 2% to 3% of the general population is chronically infected with the hepatitis C virus (HCV). Chronic hepatitis C is the most prevalent disease in hepatology clinics, accounting for more than 30% of the visits. In general, it is accepted that the diagnostic protocol of chronic hepatitis C includes a liver biopsy, particularly in patients with elevated amino transferase levels. Liver histology and, especially, fibrosis staging provide prognostic information and may be useful in deciding on therapeutic strategies in individual cases. However, liver biopsy is an invasive procedure that may cause undesirable events, such as pain in 20% to 30% of the cases, major complications in 0.5%, and even death. Other than the complications derived from the procedure and the frequent poor patient acceptance, the direct cost of such procedure is high. Nowadays, chronic hepatitis C is often recognized at an early stage of the disease, and liver fibrosis is mild or absent in about 80% of the patients undergoing a liver biopsy. Thus, the finding of surrogate markers of liver fibrosis/absence of fibrosis would be relevant to reduce the number of liver biopsies in patients with chronic hepatitis C. Imbert-Bismuth et al. have recently shown that a combination of biochemical markers of liver fibrosis can be useful to predict the presence or absence of fibrosis and, therefore, to reduce the number of liver biopsies in patients with chronic hepatitis C. However, some of these markers, such as α2 macroglobulin, haptoglobin, or apolipoprotein A1 are not routinely used in clinical practice. This is also the case for other proposed predictors of fibrosis, such as hyaluronic concentration or type III procollagen.

Using variables easily available to clinicians, we have constructed and validated a model and a score system aimed to discriminate patients with substantial fibrosis.
Noninvasive Markers of Fibrosis in Nonalcoholic Fatty Liver Disease: Validating the European Liver Fibrosis Panel and Exploring Simple Markers

Indra Neil Guha,1 Julie Parkes,2 Paul Roderick,2 Dipanker Chattopadhyay,3 Richard Cross,4 Scott Harris,2 Philip Kaye,5 Alastair D. Burt,3 Steve D. Ryder,5 Guruprasad P. Aithal,5 Christopher P. Day,3 and William M. Rosenberg1

The detection of fibrosis within nonalcoholic fatty liver disease (NAFLD) is important for ascertaining prognosis and the stratification of patients for emerging therapeutic intervention. We validated the Original European Liver Fibrosis panel (OELF) and a simplified algorithm not containing age, the Enhanced Liver fibrosis panel (ELF), in an independent cohort of patients with NAFLD. Furthermore, we explored whether the addition of simple markers to the existing panel test could improve diagnostic performance. One hundred ninety-six consecutively recruited patients from 2 centers were included in the validation study. The diagnostic accuracy of the discriminant scores of the ELF panel, simple markers, and a combined panel were compared using receiver operator curves, predictive values, and a clinical utility model. The ELF panel had an area under the curve (AUC) of 0.90 for distinguishing severe fibrosis, 0.82 for moderate fibrosis, and 0.76 for no fibrosis. Simplification of the algorithm by removing age did not alter diagnostic performance. Addition of simple markers to the panel improved diagnostic performance with AUCs of 0.98, 0.93, and 0.84 for the detection of severe fibrosis, moderate fibrosis, and no fibrosis, respectively. The clinical utility model showed that 82% and 88% of liver biopsies could be potentially avoided for the diagnosis of severe fibrosis using ELF and the combined panel, respectively. The ELF panel has good diagnostic accuracy in an independent validation cohort of patients with NAFLD. The addition of established simple markers augments the diagnostic performance across different stages of fibrosis, which will potentially allow superior stratification of patients with NAFLD for emerging therapeutic strategies. (HEPATOLOGY 2008;47:455-460.)

N onalcoholic fatty liver disease (NAFLD) is emerging as a major global cause of liver disease on the background of an increasing prevalence of obesity and type 2 diabetes. NAFLD encompasses a spectrum of disease from simple steatosis through nonalcoholic steatohepatitis, to fibrosis and ultimately cirrhosis and hepatocellular carcinoma. The identification of the minority of patients with fibrosis amongst those with NAFLD is critically important for prognosis1,2 and therefore for the selection of patients that are candidates for existing and emerging therapeutic interventions. Furthermore, the identification of the subset of patients who have developed cirrhosis has clear importance for prophylaxis against variceal bleeding, surveillance for hepatocellular cancer, and the timing of transplantation. Currently, the differentiation of steatosis from steatohepatitis and fibrosis in NAFLD is dependent on histological examination of liver biopsies. However, liver biopsy is invasive and is limited by sampling error, diagnostic accuracy, and hazard to the patient.3,4 In addition, the numbers of patients with NAFLD means that use of liver biopsy in their investigation is both practically and financially impractical.

The Original European Liver Fibrosis (OELF) test is an example of a panel of markers (which highlight matrix turnover) and consists of age, tissue inhibitor of matrix metalloproteinase 1 (TIMP 1), hyaluronic acid (HA), and

See Editorial on Page 383

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; BMI, body mass index; DS, discriminant scores; ELF, Enhanced Liver fibrosis panel; HA, hyaluronic acid; NAFLD, nonalcoholic fatty liver disease; OELF, Original European Liver Fibrosis panel; TIMP1, tissue inhibitor of matrix metalloproteinase 1.

From the1Liver group,2Public Health Sciences and Medical Statistics, University of Southampton, Southampton, UK;3Institute of Cellular Medicine, Medical School, Newcastle University, UK;4iQur Ltd, Southampton General Hospital, Southampton, UK; and5University Hospital, Queen’s Medical Centre, Nottingham, UK.

C.P.D. and W.M.R. are joint senior authors for this manuscript.

Received May 21, 2007; accepted August 14, 2007.

Address reprint requests to: Dr. I.N. Guha, Mail point 805, Level C, South Academic Block, Southampton General Hospital, University of Southampton, S016 6YD. E-mail: guhaneil@hotmail.com; fax: (44) 02 380794772.

Copyright © 2007 by the American Association for the Study of Liver Diseases. Published online in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/hep.21984

Potential conflict of interest: W.M.R. is a shareholders in iQur Ltd. J.P., P.R., and W.M.R. have received grant income from Bayer/Siemens.
rithm is the combination of simple markers and ELF panel. Figure 2 shows the translation of the improved AUC scores into number of biopsies avoided using the clinical utility model described (optimal thresholds for simple and combined panels are shown in Appendix 2).

Discussion

The performance of the ELF panel in this cohort in distinguishing severe fibrosis is excellent, with an AUC of 0.9, and is comparable to the original cohort and other panel marker tests in this disease. Moreover, age is not required in this modified algorithm, simplifying the panel and allowing age to be used as an independent variable that could be useful in future prognosis studies.

The direct comparison of ELF and simple markers in the same cohort attempts to address the question of whether the inclusion of specific markers of matrix turnover confers any additional benefit in the diagnosis of liver fibrosis in comparison with simple clinical and biochemical parameters. There is a suggestion from the AUC values and clinical utility model that improvement of the algorithm can be made by combining the ELF and simple markers panel; larger studies are required to confirm these findings.

The long-term prognostic studies suggest that fibrosis, and in particular severe fibrosis, is the most important histological determinant for developing future disease. However, there are a number of reasons why the identification of other severities of fibrosis could be beneficial. In the community setting, it will facilitate the identification of patients with any fibrosis so that dietary and lifestyle interventions can be implemented early as well as deciding which of the many patients with abnormal liver function tests require referral to secondary care. In the secondary care setting, serum marker tests could be used to identify suitable candidates for new and emerging pharmacological treatments for NAFLD and liver fibrosis, the selection of the test threshold being determined by the risk–benefit ratio of the therapy; for example, an effective but potentially toxic therapy will require a threshold to be chosen with a high specificity. Finally, the identification of severe fibrosis and cirrhosis will aid stratification of treatment by serial measurement of serum markers would give clinicians valuable information to aid management decisions. In this regard there is emerging evidence that this detailed information is required for every patient presenting with abnormal liver function tests serves the best interests of neither the patients nor the funders of healthcare. Evidence presented in this study clearly demonstrates that noninvasive markers can be used to identify patients who have early fibrosis in NAFLD, making them suitable tests for screening the increasing number of patients presenting with abnormal liver function tests, obesity, and metabolic syndrome.

This study has a number of limitations. The study population has been selected from a tertiary care setting and represents a more severe disease spectrum. This is a criticism of the majority of noninvasive markers and is attributable to the requirement of the liver biopsy as a reference standard. Although these results cannot be extrapolated to other healthcare settings, with different prevalence of disease, modeling suggests that the diagnostic accuracy for the identification of severe fibrosis will improve in a community setting (unpublished data). Use of noninvasive tests may depend on simple practical considerations such as the necessity for a fasted sample and availability of specific panel components. Whereas a single ELF algorithm is used to evaluate all stages of fibrosis, the simple marker algorithm must be adjusted to achieve optimal performance. Moreover, the ELF algorithm does not require demographic or anthropometric data, potentially simplifying the acquisition of data. The economic cost of any commercial test will need to be balanced against any practical or diagnostic benefit gained.

One of the intriguing aspects of the ELF panel is the variation of performance in different diseases as shown in the original study. Although it is tempting to think of fibrosis as a common pathway for all liver disease, differences in the distribution of fibrosis and mechanisms of fibrosis may account for the variation in diagnostic accuracy, exemplified by the periportal distribution of fibrosis in hepatitis C compared with the perisinusoidal distribution of fibrosis in NAFLD and confounding effects attributable to inflammation, necrosis, and apoptosis. Other possibilities include the direct effects of the disease origin or the influence of extrahepatic manifestations on the serum markers. Although these uncertainties do not detract from the diagnostic use of ELF, clarifying these issues may provide further insights into fibrogenesis.

The true potential of serum markers may not be realized until longitudinal studies measuring serum markers against clinical outcomes are published. The ability to measure disease progression, regression, and response to treatment by serial measurement of serum markers would give clinicians valuable information to aid management decisions. In this regard there is emerging evidence that
Non-invasive tests for liver fibrosis: Encouraging or discouraging results?

Andrew K. Burroughs*, Evangelos Cholongitas

The Sheila Sherlock Hepatobiliarypancreatic and Liver Transplantation Unit, Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG, UK

In this issue Leroy et al. [1], for the first time, report a prospective comparison of diagnostic performance of six different non-invasive tests (NIT) for hepatic fibrosis (MP3, Fibrotest, Fibrometer, Hepascore, Forns’ index and APRI) using 180 patients/biopsies with chronic hepatitis C (CHC).

The NIT have two categories [2]. The first is based on serum markers: direct and indirect [3]. The indirect NIT comprise routinely available tests: Bonacini index [4] (platelet count, ALT/AST ratio, INR), Forns’ index [5] (age, platelet count, γGT, cholesterol), Fibroindex [6] (AST, platelet count, gamma globulin) and APRI index [7] (AST/platelet count). Recently, APRI, modified by adding ALT and INR, resulted in improved performance [8]. More complex and expensive indirect NIT are the Fibrotest [9] (a2-macroglobulin, haptoglobin, gamma globulin, apolipoprotein, bilirubin) and PGAA index [10] (a2-macroglobulin, γGT, apolipoprotein A1, prothrombin time). The Fibrotest is the most validated NIT for liver fibrosis [11]; cut-offs allow conversion to the five METAVIR stages (F0–F4) [12–14]. A recent comparison of five indirect NIT found them insufficiently accurate for the diagnosis of cirrhosis [15].

Direct NIT measure components of extracellular matrix in serum, such as glycoproteins (e.g. laminin, hyaluronic acid), collagen IV, pro-collagen III, metalloproteinases and tissue inhibitors of metalloproteinases, as well as cytokines implicated in the fibrogenic process (TGF-β1, TNF) [11]. Direct NIT have been used singly [16,17] or in combination (e.g. Fibrospect: hyaluronic acid, TIMP-1, a2-macroglobulin) increasing accuracy [18], but further studies are needed to assess performance in diagnosing cirrhosis and changes in fibrosis [2,11]. Moreover, indirect NIT, which are less expensive, have similar diagnostic accuracy to direct NIT [3,11].

Some scores combine direct and indirect NIT (e.g. Fibrometer comprising platelets, prothrombin time, AST, a2-macroglobulin, age, urea and hyaluronic acid) [19], or use stepwise algorithms to diagnose significant fibrosis or cirrhosis. Prospective evaluation [20,21] has shown improved accuracy compared to single markers, with some studies suggesting the need for liver biopsy (LB) can be removed in more than 50% of cases [11,21].

The second category of NIT are methodologies related to liver imaging techniques [2]. Ultrasound, computed tomography scan and magnetic resonance imaging are performed routinely, particularly if cirrhosis is suspected. Structural changes can be detected, but stage of fibrosis cannot diagnose accurately except cirrhosis. Ultrasound correctly diagnoses cirrhosis in up to 90% of cases (using 2 or 3 quantitative and qualitative parameters or Doppler signs) [22–24]. Fibroscan is based on ultrasound technology, which correlates liver fibrosis with hepatic elasticity [25,26]. Prospectively, it accurately predicted significant fibrosis (METAVIR ≥ F2) or cirrhosis (ROC curves: 0.74 and 0.94, respectively) [27] (although it has not been extensively assessed in patients with F0 or F1), and the presence of complications of cirrhosis [28]. It has high reproducibility (intra- and inter-observer agreement: 98%) [29];
Editorial

Non-invasive tests for liver fibrosis: Encouraging or discouraging results?

Andrew K. Burroughs*, Evangelos Cholongitas

The Sheila Sherlock Hepatobiliarypancreatic and Liver Transplantation Unit, Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG, UK

In this issue Leroy et al. [1], for the first time, report a prospective comparison of diagnostic performance of six different non-invasive tests (NIT) for hepatic fibrosis (MP3, Fibrotest, Fibrometer, Hepascore, Forns’ index and APRI) using 180 patients/biopsies with chronic hepatitis C (CHC).

The NIT have two categories [2]. The first is based on serum markers: direct and indirect [3]. The indirect NIT comprise routinely available tests: Bonacini index [4] (platelet count, ALT/AST ratio, INR), Forns’ index [5] (age, platelet count, γGT, cholesterol), Fibroindex [6] (AST, platelet count, gamma globulin) and APRI index [7] (AST/platelet count). Recently, APRI, modified by adding ALT and INR, resulted in improved performance [8]. More complex and expensive indirect NIT are the Fibrotest [9] (a2-macroglobulin, haptoglobin, gamma globulin, apolipoprotein, bilirubin) and PGAA index [10] (a2-macroglobulin, γGT, apolipoprotein A1, prothrombin time). The Fibrotest is the most validated NIT for liver fibrosis [11]; cut-offs allow conversion to the five METAVIR stages (F0–F4) [12–14]. A recent comparison of five indirect NIT found them insufficiently accurate for the diagnosis of cirrhosis [15].

Direct NIT measure components of extracellular matrix in serum, such as glycoproteins (e.g. laminin, hyaluronic acid), collagen IV, pro-collagen III, metalloproteinases and tissue inhibitors of metalloproteinases, as well as cytokines implicated in the fibrogenic process (TGF-β1, TNF) [11]. Direct NIT have been used singly [16,17] or in combination (e.g. Fibrospect: hyaluronic acid, TIMP-1, a2-macroglobulin) increasing accuracy [18], but further studies are needed to assess performance in diagnosing cirrhosis and changes in fibrosis [2,11]. Moreover, indirect NIT, which are less expensive, have similar diagnostic accuracy to direct NIT [3,11].

Some scores combine direct and indirect NIT (e.g. Fibrometer comprising platelets, prothrombin time, AST, a2-macroglobulin, age, urea and hyaluronic acid) [19], or use stepwise algorithms to diagnose significant fibrosis or cirrhosis. Prospective evaluation [20,21] has shown improved accuracy compared to single markers, with some studies suggesting the need for liver biopsy (LB) can be removed in more than 50% of cases [11,21].

The second category of NIT are methodologies related to liver imaging techniques [2]. Ultrasound, computed tomography scan and magnetic resonance imaging are performed routinely, particularly if cirrhosis is suspected. Structural changes can be detected, but stage of fibrosis cannot diagnose accurately except cirrhosis. Ultrasound correctly diagnoses cirrhosis in up to 90% of cases (using 2 or 3 quantitative and qualitative parameters or Doppler signs) [22–24]. Fibroscan is based on ultrasound technology, which correlates liver fibrosis with hepatic elasticity [25,26]. Prospectively, it accurately predicted significant fibrosis (METAVIR ≥F2) or cirrhosis (ROC curves: 0.74 and 0.94, respectively) [27] (although it has not been extensively assessed in patients with F0 or F1), and the presence of complications of cirrhosis [28]. It has high reproducibility (intra- and inter-observer agreement: 98%) [29];
combined with Fibrotest, diagnostic performance for significant and severe fibrosis/cirrhosis was improved [25,26]. However, Fibroscan requires costly hardware, and some operator training. With current equipment, the signal penetrates only 25–65 mm, a limiting factor in obese patients (particularly with increased fatty thoracic belt) [3], but new developments may overcome this. On the other hand, although 10 successful shots are recommended for optimal performance, 5 valid shots after a rapid operator training have been sufficient [30].

The existence of several and different scores for NIT for liver fibrosis shows that no single measure has achieved sufficiently good diagnostic performance. Thus, APRI and Forns’ index result in almost half of patients being unclassified [5,7,21] and they do not perform well in specific circumstances (e.g. Forns’ index in patients with genotype 3 CHC) [11,31]. In addition, NIT, even if they were diagnostically accurate, are not a surrogate for other histological features such as inflammation and steatosis, which have significant predictive value for the progression of fibrosis [32–34]. Moreover, most NIT are not able to identify individual stages of fibrosis, and, to date, the major limitation is the identification of intermediate stages of fibrosis [35,36]. METAVIR F0/F1 versus higher stages can be used as a cut-off for indicating antiviral therapy, whereas F0/F1/F2 versus higher stages can differentiate need for therapy and screening for portal hypertension. Reasonable diagnostic accuracy with less variability exists for detecting advanced fibrosis and cirrhosis [35,37].

However, one may ask, when cirrhosis is suspected how does ultrasound perform versus these tests? Although ultrasound has several limitations [38], it can diagnose cirrhosis [22] with high accuracy [24] (particularly in combination with clinical examination and with or without simple blood tests, such as platelet count). The diagnostic accuracy of direct NIT (e.g. hyaluronic acid) and PGAA is better than clinical examination and simple echography signs (spleen length, portal velocity) [39], but confidence intervals overlap widely. In another study, APRI compared to platelet count alone had comparable accuracy for prediction of significant fibrosis (ROC curves: 0.80 and 0.71, respectively) and cirrhosis (ROC curves: 0.90 and 0.89, respectively) [40]. Formal studies evaluating diagnosis of severe fibrosis/cirrhosis by ultrasound versus NIT, and the ancillary information gathered by either method should be performed. When evaluating NIT or considering their use in clinical practice, it is important to verify what particular question is being asked: is it to confirm or exclude cirrhosis, or to accurately stage fibrosis without a biopsy, or to confirm normality? These different end points will have different measures for positive and negative predictive values of the tests. Unfortunately these different questions are not answered in every study and results have been extrapolated wrongly from one clinical setting to another.

Other aspects of the evaluation of NIT are firstly, the lack of calibration, which compares the predicted stage with the actual stage across the spectrum of fibrosis. This is different from the assessment of discriminatory ability (ROC curves), which depends on the sensitivity, and specificity, i.e. a plot of true positive rate (sensitivity) versus false positive rate (1-specificity), and is related to the relative prevalence of fibrosis stages [37]. These parameters enable ranking of patients with different stages of fibrosis and can be used to compare tests, as Leroy et al. have done [1]. Thus, discrimination can be very high if the question is “cirrhosis versus normal”, but the same tests have poor calibration if the question is F0/F1 versus F2 METAVIR or higher. Indeed, for the latter, discriminatory ability is also reduced, as already mentioned. Leroy et al. [1] found that within a Fibrotest value ranging from 0.22 to 0.74 (i.e. F0/F1 to F3) there was 33% discordance, whereas Fibrotest values predicting either F0 or F4 only had 12.6% discordance, i.e. one can deduce the calibration is not good. Other statistical approaches, such as likelihood ratios [36] or the performance profile test [37], may better reflect the diagnostic performance of NIT.

Secondly, NIT, all of which have continuous scores, have been correlated with categorical variables, i.e. the stage scores, which are only descriptive categories (which are different amongst the various histological assessments) and do not have an arithmetical progression, e.g. stage 2 fibrosis (F2) is neither twice the severity of stage 1 (F1) nor half the severity of stage 4 (F4) in METAVIR. This partly explains the poor correlation of NIT with fibrosis stage found by Leroy et al. [1], in which only a third of patients could be classified with 90% certainty and thus “avoid” a LB.

A methodologically more correct comparison would be, between NIT scores and a quantitative measurement of liver fibrosis, such as histological digital analysis of collagen, which has only been performed by Cales et al. [19]. However, clinical correlations with quantification of liver collagen have not been extensively evaluated, but should be in the future [19]. Quantitative correlations with collagen content could help to validate NIT, and may help reduce errors, when evaluating LB of suboptimal quality.

Even discounting the issue of a gold standard for the quantification of collagen, the problem of assessing NIT is the quality of the current gold standard i.e. accurate histological assessment of stage, which is another reason for discordant results between LB and NIT [14]. Given that recent standards suggest optimal grading and staging in chronic viral hepatitis should be performed with LB samples of 20–25 mm length and/or containing ≥11 complete portal tracts [41], then NIT ideally should be assessed against biopsies, that nearly reach, or fulfil, these standards (or even better two specimens of 20 mm each)
Surprisingly, the performance of NIT has mostly been evaluated using suboptimal liver biopsy specimens (<20–25 mm and even ≤15 mm), common in the literature [42], or the quality of the biopsy was not mentioned, so that a priori, diagnostic accuracy was suboptimal from the start [42]. Thus, even a recent study [30], published in the Journal of Hepatology, had 34% of LB ≤15 mm in length with 6.3% not having the length recorded and neither was the type of needle. In contrast, 89% of LB evaluated by Leroy et al. [1] were ≥15 mm long, but only 45% ≥25 mm long. If the ideal criteria are adhered to, a single pass of percutaneous LB (PLB) or transjugular LB (TJLB) usually provides an inadequate biopsy specimen [42]. However, multiple cores of TJLB can be obtained, in contrast to PLB where more than one pass gives rise to increased complications [42]. In our centre, TJLB are performed routinely, with three passes as a standard, and the mean length is 22 mm with eight complete portal tracts [43]. Three non-fragmented cores of at least 7 mm give better reproducibility compared to one similar core for diagnosis as well as for inflammation and fibrosis [44]. TJLB performed correctly could be used for validation studies of NIT.

Thus, it is not surprising that the study by Leroy et al. [1], despite a far better than average quality of LB [42], confirmed poor accuracy of NIT, with once again a difficulty in identifying intermediate stages of fibrosis. However, they are to be applauded for conducting a thorough independent validation of NIT. Using ROC curves, Fibrometer had the best, and Forns’ index the worst, discrimination ability. The various NIT gave similar discrimination values, almost always greater than 0.80, similar to the original papers, with regard to discrimination between F0F1 versus F2F3F4 METAVIR and F0F1F2 versus F3F4 METAVIR, with significant differences only between Fibrometer and Forns’ index. However, about 20% of patients were misclassified and 50% could not be classified because of intermediate values, so that the precision is not sufficient for any individual patient. Moreover, despite the correlation of cut-off values of Fibrotest with METAVIR stages [12], there was a discordance of at least 2 fibrosis stages and biopsy, in 23% of cases, which, given the quality of LB, is not due to this as suggested by previous studies [14,45]. Interestingly, the performance of NIT was the same irrespective of HCV genotype (1 versus non-1) and the adequacy of LB (< or ≥25 mm).

The finding that adequacy of biopsy (providing ≥15 mm in length) does not appear to affect the performance of NIT is a very important one to confirm, because depending on the validity of these data, then NIT could compensate for the large proportion of “sub-optimal” biopsies (≤25 mm) which are usually obtained [42]. In addition, NIT may be able to assess the severity of scarring in histological cirrhosis, which at present can only be assessed by evaluating hepatic venous pressure gradient [46]. Conversely, NIT currently could be considered intrinsically unreliable because the basis of their evaluation is flawed i.e. inadequate biopsy specimens. This is an unresolved issue and requires evaluation of optimal biopsy specimens with NIT. Even if one takes an optimistic view that NIT can overcome the inadequacy of liver biopsy specimens, overall discordance between METAVIR fibrosis and Fibrotest was observed in 52% of cases in Leroy’s study, independent again of LB length, thus strongly imputing a problem with NIT scoring, and was clinically significant in 23% of patients.

Leroy et al. [1] could not confirm a single best NIT, but they found that a combination of 2 scores improved performance. Fibrotest and APRI ruled out significant fibrosis (F2F3F4) with a negative predictive value of 94.1%. An algorithm based on combined utilization of APRI and Fibrotest, simpler than that of Sebastiani et al. [21], excluded significant fibrosis or could diagnose extensive fibrosis (F3F4). If the clinical question relates to discriminating between these categories, this looks like a useful scheme, but of course it applies only to patients with CHC i.e. a known diagnosis. Moreover, as the authors themselves demonstrate, the best combinations could select only one-third of patients, for whom either absence of significant fibrosis or presence of extensive fibrosis could be predicted with more than 90% certainty.

Indeed, only some studies notably the Fibrometer study [19] and the European Liver Fibrosis study [20] included patients with a wide range of liver diseases. In the latter study, an algorithm using nine surrogate markers of fibrosis to exclude significant fibrosis and cirrhosis (Scheuer stages 3 and 4) had a sensitivity in excess of 90%. This is impressive, but it was not correlated with imaging assessment. It would be interesting to perform a similar study using Fibroscan, as combining their algorithm and that by Cales et al. [19] and Fibroscan could increase the diagnostic accuracy, particularly in patients with non-CHC liver disease.

NIT for liver fibrosis have the potential to become an important tool in clinical practice [47], but better validation is needed before starting to consider NIT as established tests, as illustrated by Leroy et al. [1] who on balance give discouraging results. Studies comparing collagen content of liver biopsies, and using only optimal biopsies to provide the best reference standard, still need to be performed. It is likely that an initial diagnostic biopsy will still be needed, but follow up for fibrosis could be based on NIT, providing that encouraging results will published in the future.

References

A Canadian Multicenter Retrospective Study Evaluating Transjugular Liver Biopsy in Patients With Congenital Bleeding Disorders and Hepatitis C: Is It Safe and Useful?

Jennifer L. Shin,1,2 Jerome Teitel,2 Mark G. Swain,3 Vincent G. Bain,4 Paul C. Adams,5 Kenneth Croitoru,6 Kevork Peltekian,7 Frank Schweiger,8 Martin E. Simons,9 and E. Jenny Heathcote1*

1 Department of Medicine at University Health Network, University of Toronto, Toronto, Ontario, Canada
2 St. Michael’s Hospital, University of Toronto, Toronto, Ontario, Canada
3 University of Calgary Foothills Hospital, Calgary, Alberta, Canada
4 University of Alberta Hospital, Edmonton, Alberta, Canada
5 London Health Sciences Center, University of Western Ontario, London, Ontario, Canada
6 Intestinal Diseases Research Program, McMaster University, Hamilton, Ontario, Canada
7 Queen Elizabeth II Hospital, Dalhousie University, Halifax, Nova Scotia, Canada
8 Southeast Health Regional Authority, Moncton, New Brunswick, Canada
9 Department of Medical Imaging, University Health Network, Toronto, Ontario, Canada

Prior to the introduction of virally inactivated clotting factor concentrates, the majority of individuals with congenital bleeding disorders became infected with the hepatitis C virus. Although liver biopsy is valuable in prognosis and guiding antiviral therapy, there is a reluctance to perform biopsies in this population because of the risk of hemorrhage. The purpose of this study was to evaluate the safety of transjugular liver biopsy, and the usefulness of evaluating liver histology in this patient population. Liver histopathology was assessed by the METAVIR index and compared with corrected sinusoidal pressures, platelet counts, and abdominal ultrasonography. Liver biopsy was performed at seven Canadian centers in 65 patients with hemophilia or von Willebrand’s disease. Biopsies were done on an outpatient basis, followed by a 4-hr observation period in hospital. Normal hemostasis was maintained during the peri biopsy period, with follow-up doses of factor concentrate self administered by the patient at home. One patient (1.4%) had significant bleeding leading to readmission and red cell transfusion. Liver histology showed 14 patients (22%) had cirrhosis. Ten patients had elevated corrected sinusoidal pressures; 7 of these (70%) had cirrhosis on biopsy, and the other 3 (30%) likely had cirrhosis although histology showed stage 3 fibrosis. This series represents the largest reported experience of transjugular biopsy in individuals with congenital bleeding disorders. We conclude that this procedure can be safely performed on an outpatient basis. The diagnosis of cirrhosis and/or portal hypertension was made in a substantial proportion of individuals (26%), all of whom had asymptomatic liver disease. Am. J. Hematol. 78:85–93, 2005.

Key words: hepatitis C; hemophilia; safety; fibrosis; liver biopsy
INTRODUCTION

The majority of individuals with congenital bleeding disorders who received factor concentrates that were not virally inactivated, or that were treated with first generation viral inactivation methodology were exposed to the hepatitis C virus. Almost 100% of these individuals test hepatitis C antibody-positive [1].

Earlier reports suggested that the natural history of hepatitis C infection in hemophiliacs [2–5] was a mild disease with no evidence of significant progression. Subsequent studies indicated that progressive disease did occur [6,7] and that some patients presented with complications of cirrhosis [8]. These discrepancies may be explained by differences in the duration of infection. Cirrhosis was the primary or an associated cause of death in 8–11% of deaths in hemophiliac patients [9,10]. Human immunodeficiency (HIV) co-infection is also an important contributor to the development of cirrhosis and liver failure [8,11].

Liver biopsy is an important tool in the evaluation of patients with chronic hepatitis C infection. Liver biopsy is currently the only accurate way to assess the extent of liver damage and thus to establish the potential to develop liver failure. Individuals with cirrhosis secondary to hepatitis C not only require consideration of urgent antiviral therapy but also need regular screening for the development of hepatocellular carcinoma and esophageal varices. Biopsy also provides a baseline for subsequent histological comparisons. The 2002 United States consensus conference on hepatitis C recommends that HCV-infected patients undergo pretreatment liver biopsy to assess the degree of hepatic fibrosis and inflammation, particularly in those patients with HCV genotypes other than 2 or 3 [12]. The information obtained on liver biopsy allows affected individuals to make more informed choices about the initiation or postponement of antiviral treatment.

The most frequent complication of liver biopsy is hemorrhage, hence the reluctance to perform liver biopsy in individuals with congenital bleeding disorders [6]. The transjugular approach to liver biopsy has been used with relative safety in patients with compensated liver disease, manifested by a coagulopathy and/or massive ascites [13].

The aim of our study was to evaluate the safety of transjugular liver biopsy in patients with congenital bleeding disorders and hepatitis C and to examine the usefulness of histological verification of liver disease severity prior to the initiation of antiviral therapy. Our multicenter Canadian experience is the largest reported study of transjugular liver biopsies in this population, and demonstrates that our treatment protocol for hemostasis prophylaxis allows the biopsy to be performed safely as an outpatient. The 4-hr inpatient monitoring done post biopsy in this study is significantly shorter than the typical 1- to 3-day period that is reported in the literature and may be a source of potential cost savings.

PATIENTS AND METHODS

Patient Characteristics

Subjects with congenital bleeding abnormalities and abnormal liver biochemistry were assessed at seven centers across Canada prior to consideration of antiviral treatment for chronic hepatitis C. Patients were under the joint care of a hepatologist and a hematologist who was the medical director of a comprehensive hemophilia program. All patients were seropositive for antibodies against hepatitis C virus and had serum aminotransferase (ALT) levels above the upper limit of normal at least once in the previous 6 months. All of these patients were asymptomatic from a liver disease standpoint.

Liver Biopsy Protocol

All vascular access and liver biopsies were accomplished under cover of clotting factor concentrate replacement. The exact regimen varied according to the center. The center in Toronto administered concentrate in doses calculated to raise the activities of factor VIII or IX to approximately 0.75 IU/mL immediately preprocedure using the usual replacement product (35–40 IU/kg for factor VIII, 70–100 IU/kg for factor IX). Three further doses of concentrate were administered postprocedure, in each case using half the number of units that were given as the preprocedure dose. These doses were given at approximately 12, 24, and 48 hr after the biopsy, usually self-administered at home. Some patients with mild hemophilia or von Willebrand’s disease received desmopressin (DDAVP) intravenously or subcutaneously 0.4 μg/kg (maximum dose 20 μg) for 2 days. Other centers aimed to replace to a level of 1.0 IU/mL immediately preprocedure and then maintain values not below 0.50 IU/mL for 2–4 days with factor concentrate, administered at home. Factor concentrate was not administered by continuous infusion. None of the patients had evidence of inhibitors.

Experienced interventional radiologists performed the liver biopsy. Following concentrate replacement, patients were brought immediately to the angiography suite for transjugular biopsy. Under aseptic technique, conscious sedation, and local anesthesia, the right hepatic vein was catheterized via the right internal jugular approach under sonographic guidance. Contrast-enhanced venography was routinely performed to determine anatomy and
To the Editor:

Liver biopsy is viewed as the gold standard for staging fibrosis in chronic hepatitis. An ongoing great search for noninvasive diagnostic tests aims at replacing this inconvenient and costly invasive procedure. Several models have been proposed\(^1-3\); however, most of them require special laboratory parameters that are not available in clinical practice. In a recent article, Wai et al. proposed a simple model on the basis of routinely available laboratory test results (aspartate aminotransferase–to-platelet ratio index, \([\text{APRI}]\)), which was shown to predict with high sensitivity and specificity liver fibrosis in patients with hepatitis C virus infection.\(^4\) Tests aims at replacing this inconvenient and costly invasive procedure.

Several models have been proposed\(^1-3\); however, most of them require special laboratory parameters that are not available in clinical practice. In a recent article, Wai et al. proposed a simple model on the basis of routinely available laboratory test results (aspartate aminotransferase–to-platelet ratio index, \([\text{APRI}]\)), which was shown to predict with high sensitivity and specificity liver fibrosis in patients with hepatitis C virus infection.\(^4\) These results were recently challenged by Giannini and Testa, who pointed out that the ratio of aspartate aminotransferase to alanine aminotransferase assessed fibrosis more accurately than the APRI.\(^5\)

We would like to comment on the applicability and validity of the APRI test, which we have applied to a large cohort of 484 treatment naïve patients with chronic hepatitis C (271 males; mean age, 46 ± 0.4; range, 18-68 years). All patients underwent liver biopsy as part of the screening evaluation within the course of national or international clinical trials. Staging of fibrosis was done according to the Scheuer\(^6\) score, which categorizes 4 different stages of fibrosis (F0–F4), as compared to the Ishak score used by Wai et al. referring to 6 stages (F0–F6).

Our results are in agreement with the data presented by Wai et al., although the overall sensitivity and specificity of the APRI as well as the positive predictive value and the negative predictive value for the certain cutoffs were found to be lower (Table 1). As mentioned, the different histological scoring system (i.e., Scheuer vs. Ishak score) could be responsible for these differences. Additionally, our results have been evaluated by several pathologists, not by just one as in the Wai et al. study. Interobserver differences certainly could have played a role in this respect, but in clinical practice one is confronted with this kind of situation.

From these studies it appears that fibrosis stage assessed by invasive and noninvasive approaches differ to some extent. In this respect, sampling variability and the size of liver biopsies have to be considered as important contributors to false fibrosis staging,\(^7\) and they raise the question of whether liver biopsy can still be regarded as the gold standard for fibrosis assessment. Thus, the interpretation of liver biopsy results on the background of a noninvasive fibrosis prediction method should be proven as a rational approach to improve accuracy of fibrosis staging. One should, however, be aware that a reliable noninvasive assessment of the different stages of fibrosis cannot be made in all patients. Data by Wai et al. as well as our data are in accordance, showing that a prediction concerning presence or absence of significant fibrosis was possible only in 57% and 51%, respectively, of the patients examined in both studies.

Finally, the need to confirm histologically the stage of fibrosis is suspect when clinical experience and laboratory data provide all necessary information to judge with great certainty the severity of patients’ chronic liver disease.

\(\text{Thomas Berg et al}\)

\(\text{Thomas Berg}\)\(^1\)
\(\text{Christoph Sarrazin}\)\(^2\)
\(\text{Holger Hinrichsen}\)\(^3\)
\(\text{Peter Buggisch}\)\(^4\)

\(\text{1Medizinische Klinik mit Schwerpunkt Hepatologie und Gastroenterologie}\)
\(\text{Charité - Universitätsmedizin Berlin, Campus Virchow-Klinikum}\)
\(\text{Berlin, Germany}\)

\(\text{2Medizinische Klinik und Poliklinik – Innere Medizin II Universitätskliniken des Saarlandes}\)
\(\text{Homburg/Saar, Germany}\)

\(\text{3Medizinische Universitätsklinik I, Klinik für Allgemeine Innere Medizin Christian-Albrechts-Universität}\)
\(\text{Kiel, Germany}\)

\(\text{4Medizinische Kernklinik/Gastroenterologie}\)
\(\text{Universitätsklinik Eppendorf}\)
\(\text{Hamburg, Germany}\)

\(\text{5Medizinische Klinik II, Klinikum Großhadern}\)
\(\text{Ludwig-Maximilians-Universität München, Germany}\)

\(\text{Acknowledgment: Supported in part by the German BMBF Network of Competence for Viral Hepatitis (Hep Net).}\)
Table 1. Accuracy of the AST-to-Platelet Ratio Index (APRI) in Predicting Significant Fibrosis and Cirrhosis

<table>
<thead>
<tr>
<th>APRI*</th>
<th>All Patients (n = 484) n (%)</th>
<th>Stage 0–1 (n = 231) n (%)</th>
<th>Stage 2–4 (n = 253) n (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤0.50</td>
<td>168 (35)</td>
<td>122 (53)</td>
<td>46 (18)</td>
<td>82 (91)‡</td>
<td>53 (47)‡</td>
<td>66 (61)‡</td>
<td>73 (86)‡</td>
</tr>
<tr>
<td>&gt;0.50</td>
<td>316 (65)</td>
<td>109 (47)</td>
<td>207 (82)</td>
<td>37 (41)‡</td>
<td>93 (95)‡</td>
<td>85 (88)‡</td>
<td>57 (64)‡</td>
</tr>
<tr>
<td>≤1.50</td>
<td>375 (77.5)</td>
<td>215 (93)</td>
<td>160 (63)</td>
<td>69 (89)§</td>
<td>80 (75)§</td>
<td>55 (38)§</td>
<td>88 (98)§</td>
</tr>
<tr>
<td>&gt;1.50</td>
<td>109 (22.5)</td>
<td>16 (7)</td>
<td>93 (37)</td>
<td>39 (57)§</td>
<td>93 (93)§</td>
<td>67 (57)§</td>
<td>81 (93)§</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ACTUAL FIBROSIS</th>
<th>Stage 0–2 (n = 357)</th>
<th>Stage 3–4 (n = 127)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤1.0</td>
<td>325 (67)</td>
<td>286 (80)</td>
</tr>
<tr>
<td>&gt;1.0</td>
<td>159 (33)</td>
<td>71 (20)</td>
</tr>
<tr>
<td>≤2.0</td>
<td>409 (84.5)</td>
<td>332 (93)</td>
</tr>
<tr>
<td>&gt;2.0</td>
<td>75 (15.5)</td>
<td>25 (7)</td>
</tr>
</tbody>
</table>

Abbreviations: AST, aspartate aminotransferase; PPV, positive predictive value; NPV, negative predictive value.

*Results are given according to the APRI cutoff points proposed by Wai et al. to predict the absence (APRI ≤0.50) or presence (APRI ≥1.50) of significant fibrosis and the absence (APRI ≤1.00) or presence (APRI ≥2.00) of cirrhosis.
†Staging of fibrosis was done according to the Scheuer score, which categorizes 4 different stages of fibrosis (F0, absent; F1, mild portal fibrosis without septa; F2, moderate portal fibrosis with few septa; F3, numerous septa (bridging fibrosis) without cirrhosis; F4, cirrhosis), as compared to the Ishak score used by Wai et al. referring to 6 stages (F0–F6). Significant fibrosis was defined as ≥F2. Data for prediction of cirrhosis were given for patients histologically classified as either F3–F4 or F4.
‡§Respective data for sensitivity and specificity as well as the positive and negative predictive value assessed in the study by Wai et al. for patients with Ishak fibrosis scores F0–2 vs. F3–6‡ and 5–6 vs. F0–4§ are shown in parentheses.

References


Copyright © 2004 by the American Association for the Study of Liver Diseases. Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/hep.20226

Reply:

We thank Berg et. al. for the interest expressed on our article. We are grateful that many investigators like Berg et. al. had tested our prognostic model in their patient populations. Using the AST-to-Platelet Ratio Index (APRI) to predict the absence (APRI ≤0.50) or presence (APRI ≥1.50) of significant fibrosis and the absence (APRI ≤1.00) or presence (APRI ≥2.00) of cirrhosis is a simple predictive model.

The sensitivities were lower. The proportion of patients who fell into the classifiable group was higher than in our study. We agree that the diminished sensitivities may be related to the use of a different fibrosis scoring system, as well as multiple versus single pathologists scoring the biopsies. It is also possible that the inclusion of stage 3 in the Scheuer fibrosis scoring system may affect histological interpretation. Because liver biopsies are not necessarily gold standards for assessing liver histol-
"Stepwise combination algorithms of non-invasive markers to diagnose significant fibrosis in chronic hepatitis C"

Journal of Hepatology April 2006

Giada Sebastiana, Alessandro Varioa, Maria Guidob, Franco Noventaa, Mario Plebanic, Roberta Pisticsa, Alessia Ferraria, Alfredo Albertia

a Department of Clinical and Experimental Medicine, University of Padova, via Giustiniani, 35100 Padova, Italy
b Department of Oncological and Surgical Sciences, University of Padova, Padova, Italy
c Department of Laboratory Medicine, University of Padova, Padova, Italy

Following the main article are 2 interesting letters to the editor with additional useful information.

"....Several non-invasive markers of liver fibrosis have been described but their use in place of liver biopsy is still controversial and not universally accepted due to still unsatisfactory diagnostic accuracy. Indeed, some of these methods, like APRI and Forns' index, leave many patients unclassified and all of them do not exceed 80-85% diagnostic accuracy.... As a consequence, many patients still need to have a liver biopsy taken and in those classified without biopsy misdiagnosis is expected to occur in at least 15-20%, a figure that is considered inadequate by many clinicians..... These markers have other limitations.... Our results indicate that the diagnostic performance of these non-invasive markers of liver fibrosis can be greatly improved by combining them in stepwise algorithms..... The algorithm we have developed for identifying patients with significant fibrosis (F≥2) among HCV carriers with elevated ALT had indeed more than 94% diagnostic accuracy with around 50% reduction in the number of liver biopsies needed.... We have also developed an algorithm for identifying patients with significant fibrosis among HCV carriers with PNALT (normal ALT).....there is abundant evidence in the literature that around 15-30% of them may have significant fibrosis and a definitive indication to antiviral therapy.... The algorithm we have developed to identify cirrhosis showed around 95% diagnostic accuracy and allowed to save 65-70% of liver biopsies..."

ABSTRACT

Background/Aims: In chronic hepatitis C, biopsy is the gold standard for assessment of liver fibrosis. Non-invasive markers have been proposed but their use is limited by diagnostic accuracy. Our aim was to increase the diagnostic performance of non-invasive markers of liver fibrosis by combining them in sequential algorithms.

Methods: One hundred and ninety patients with chronic hepatitis C were evaluated for AST to platelets ratio (APRI), Forns' index and Fibrotest at the time of liver biopsy and stepwise combination algorithms were developed and validated prospectively in 100 additional patients.

Results: Three algorithms were developed: (1) significant fibrosis (F≥2 by METAVIR) was identified with high diagnostic performance (>94% accuracy) using APRI as screening test, followed by Fibrotest in APRI non-classified cases and restricting liver biopsy to patients classified F0-F1 by non-invasive tests. (2) A slightly modified algorithm had similar performance when applied to hepatitis C carriers with normal ALT.
Identification of cirrhosis (95% accuracy) was achieved using a dedicated algorithm with different cut-off, reducing by 60-70% the liver biopsies needed.

Conclusions: Stepwise combination of non-invasive markers of liver fibrosis improves the diagnostic performance in chronic hepatitis C. Need for liver biopsy is reduced by 50-70% but cannot be completely avoided.

Introduction
Chronic infection with hepatitis C virus remains a major health problem with around 200 million individuals affected worldwide [1]. The natural course of chronic hepatitis C is characterised by progressive fibrosis in the inflamed liver with structural and hemodynamic changes leading to cirrhosis, which is followed by end-stage complications [2]. Accordingly, the prognostic evaluation and clinical management of still compensated chronic hepatitis C is largely based on assessment of the type and degree of liver fibrosis and several semiquantitative scoring systems have been proposed and validated [3], [4], [5], [6]. Liver biopsy is the gold standard for fibrosis staging in chronic hepatitis C as in many other chronic liver diseases. However, liver biopsy is invasive and complications occur in 0.6-5% of patients [7], [8], [9]. Moreover, liver biopsy is costly and requires hospitalisation of at least 6-18h [10]. Finally, recent studies performed in chronic hepatitis C have demonstrated that inadequate liver biopsy sample size frequently leads to underestimation of fibrosis stage [11], [12]. Furthermore, laparoscopic studies have shown that cirrhosis is missed by percutaneous liver biopsy in 10-30% of the cases [13], [14], [15]. For all these reasons a great interest and many studies have been recently dedicated to the development of non-invasive markers as surrogates of liver biopsy. These non-invasive markers include the AST to platelet ratio index (APRI) proposed by Wai et al., the Forns' index based on age, platelets, ΓGT and cholesterol and more sophisticated model like Fibrotest [16], [17], [18]. Fibrotest is based on five serological parameters including bilirubin, ΓGT, apolipoprotein A1 (ApoA1), alfa-2-macroglobulin (A2M) and haptoglobin [18]. These different methods have been usually applied individually in the different validation studies. All of them have limitations. APRI and Forns' index leave many patients unclassified while Fibrotest is more expensive and uses two uncommon parameters. Furthermore, the diagnostic accuracy of most methods has not exceed 80-85% [12]. There are no published studies in which these non-invasive markers of liver fibrosis have been combined in the attempt to improve diagnostic accuracy. We have here measured APRI, Forns' index and Fibrotest in a consecutive series of patients with chronic hepatitis C. Having first assessed the diagnostic accuracy of each individual method using liver biopsy as gold standard, we have then combined them with the aim of defining stepwise algorithms of higher diagnostic accuracy to be used in most common clinical scenarios, i.e. for identifying cases with significant fibrosis and those with cirrhosis among patients presenting with chronic HCV infection.

Patients and methods
Patients
This study included two cohorts of consecutive patients with a recent diagnosis of chronic hepatitis C who underwent percutaneous liver biopsy at the Department of Clinical and Experimental Medicine at the University of Padova. The first cohort (training set) included 190 patients seen between March 2003 and June 2004. The second cohort (validation set) consisted of 100 consecutive patients seen between July 2004 and April 2005. All patients were positive for serum HCV-RNA by polymerase chain reaction and had compensated chronic HCV infection. The exclusion criteria were any other cause of chronic liver disease, co-infection with HBV or HIV and co-morbidities that could confound the results of the non-invasive markers adopted. These included current alcohol intake (>20g/die), haemolysis, Gilbert's syndrome. All biopsies were obtained with 16G Menghini type needle. To limit the risk of fibrosis underestimation, patients with biopsy samples shorter than 1.5cm or containing less than seven portal tracts were excluded [19]. According to these criteria, 76 patients with chronic hepatitis C were excluded. Informed consent was obtained from all patients participating in the study, that was conducted according to the rules of the Declaration of Helsinki.

Discussion
Staging of hepatic fibrosis is fundamental for clinical management of chronic hepatitis C. Patients showing
F≥2 stage by METAVIR are at significant risk of developing cirrhosis within 5-10 years and are considered to have priority for antiviral therapy in all International and National guidelines and recommendations [1], [2], [21]. On the other hand, many experts believe that therapy is not mandatory in patients with no or minimal fibrosis (F0-F1), particularly when the chance of achieving a favourable response is not high, like in patients with contraindications or with poor motivation towards therapy, patients with high viral load or difficult-to-treat HCV genotypes [1]. Furthermore, assessment of the risk-benefit ratio of antiviral therapy has always been to consider the stage of liver disease in the presence of any contraindication. Thus, a liver biopsy is often necessary to assess prognosis and to decide management. Several non-invasive markers of liver fibrosis have been described but their use in place of liver biopsy is still controversial and not universally accepted due to still unsatisfactory diagnostic accuracy. Indeed, some of these methods, like APRI and Forns' index, leave many patients unclassified and all of them do not exceed 80-85% diagnostic accuracy [15], [16], [22], [23], [24]. As a consequence, many patients still need to have a liver biopsy taken and in those classified without biopsy misdiagnosis is expected to occur in at least 15-20%, a figure that is considered inadequate by many clinicians [1], [21], [25], [26]. These markers have other limitations. Most of them, such as APRI and Forns' index, are not able to identify individual stages of fibrosis. APRI cannot be completely standardised due to the variability of measurement and normality ranges of AST in different laboratories [24]. The performance of Forns' index might be modified in patients infected by HCV-3 because of lower cholesterol levels [22]. In our series cholesterol levels were indeed significantly lower in patients with HCV-3 but the diagnostic performance of Forns' index was similar in HCV-3 and non-3, probably due to the low prevalence of HCV-3 in our series. Our results indicate that the diagnostic performance of these non-invasive markers of liver fibrosis can be greatly improved by combining them in stepwise algorithms. The algorithm we have developed for identifying patients with significant fibrosis (F≥2) among HCV carriers with elevated ALT had indeed more than 94% diagnostic accuracy with around 50% reduction in the number of liver biopsies needed both in the training and in the validation cohort. We have also developed an algorithm for identifying patients with significant fibrosis among HCV carriers with PNALT. This category has not been considered in most previous studies of non-invasive markers of fibrosis. However, there is abundant evidence in the literature that around 15-30% of them may have significant fibrosis and a definitive indication to antiviral therapy [27], particularly when considering the favourable results recently reported with PEG-interferon alfa-2a plus ribavirin combination therapy [28]. The slightly modified algorithm developed for significant fibrosis in these patients had similar performance to that of patients with elevated ALT. Forns' index was excluded from this stepwise algorithm because of the minimal impact on further reducing liver biopsies needed. It should be here underlined that the prevalence of significant fibrosis (F≥2) was in our series of patients with PNALT somehow higher than that usually reported in the literature [28]. This most likely was due to the attitude in our centre to perform more frequently a liver biopsy in these patients when ALT levels are in the upper part of the normal range. Furthermore, we included only liver biopsy specimens of adequate size and Colloredo et al [11], have recently shown that the use of larger samples favours identification of higher stages of liver fibrosis. The rather high prevalence of patients with significant fibrosis in our series with and without elevated ALT should be taken into account when considering the performance of the combined algorithms described. Indeed our algorithms depend on the PPV of the non-invasive markers used and may therefore perform less well when applied to clinical settings with lower prevalence of significant fibrosis, where the negative rather than positive predictive value might be more important.

Another algorithm was developed to diagnose cirrhosis. Identification of cirrhosis in chronic hepatitis C is extremely important since patients who have reached this stage need specific management, including closer monitoring for complications and hepatocellular carcinoma. Furthermore, patients with cirrhosis may be expected to have reduced response and tolerability during antiviral therapy. The algorithm we have developed to identify cirrhosis showed around 95% diagnostic accuracy and allowed to save 65-70% of liver biopsies. Again, these performances appear superior to those reported previously for individual non-invasive markers of cirrhosis. The drastic reduction in liver biopsies needed and the fact that our algorithm requires this invasive procedure in patients with low chance of having cirrhosis are particularly important since the risk of liver biopsy complications is increased in cirrhotics [7], [9].

The diagnostic algorithms we have described may not only reduce the risk of liver biopsy, but may also
Having defined the individual diagnostic performance of the three non-invasive markers of fibrosis, we have then combined them in stepwise algorithms aimed to achieve optimised diagnostic performance (accuracy >90%) while minimising the number of liver biopsies needed. These algorithms were developed in the training set on the basis mainly of the PPV or NPV of each marker and then validated in the validation set. Three different pretest clinical scenarios were chosen for the development of these algorithms. The first scenario was for patients with elevated ALT with the aims of identifying all those with significant fibrosis (F≥2 by METAVIR) and a clear indication to antiviral treatment and of minimising the number of over-diagnosed cases and the number of liver biopsies needed. The same modelling was performed for patients with PNALT. In the third scenario, the algorithm was modelled to optimise identification of cirrhosis with the aim of detecting all patients with cirrhosis, and of minimising over-diagnosis and the number of liver biopsies needed. The best diagnostic algorithm for each of these three scenarios is reported in Fig. 1a-c. The results and performance of each algorithm in the training and in the validation set are described in Table 5.
Implementing non-invasive markers for liver fibrosis in clinical practice
Sebastiani Giada et al

To the Editor:

We read with interest the article by Leroy et al. regarding a prospective, independent validation of six non-invasive markers for liver fibrosis in chronic hepatitis C [1]. The overall performance of the markers tested, which included Fibrotest, Fibrometer, Hepascore, MP3, Forns’ index and APRI, was very similar to those originally reported, the area under the ROC curve (AUC) ranging between 0.78 and 0.86 for diagnosis of significant fibrosis (≥F2 by METAVIR). The authors tested the statistical independence of the non-invasive scores in order to propose a logical algorithm to be used in clinical practice and they found that some combinations of non-invasive markers gave a better performance than the single scores. Indeed, a combination of APRI and Fibrotest allows to predict presence of significant fibrosis with more than 90% accuracy. In their article, Leroy and colleagues also referred to the sequential algorithm that we had previously proposed to diagnose liver fibrosis in chronic hepatitis C [2]. In our algorithm, APRI is used 100% as first-line test, followed by Fibrotest and then by liver biopsy in misclassified cases. The application of this algorithm has resulted in a 50% reduction of liver biopsies to diagnose F2 fibrosis with a diagnostic accuracy of 94%.

We sought to compare the performance of Leroy’s algorithm and our algorithm for the diagnosis of significant fibrosis in chronic hepatitis C. We investigated a consecutive series of 188 monoinfected HCV patients (mean age 48.6 ± 12.4, 51.6% males) who underwent a percutaneous liver biopsy. For all patients APRI and Fibrotest were calculated using fasting serum samples obtained on the same day of liver biopsy. Patients with comorbidities were excluded. METAVIR staging was: F0–F1 = 30.5%, F2 = 45.5%, F3 = 12.5%, F4 = 11.5%. The performance of the non-invasive methods was measured as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and AUROC. The mean length of liver specimens was 17 ± 3 mm. The AUROCs for significant fibrosis were 0.75 and 0.79 for APRI and Fibrotest, respectively. Table 1 shows the performance and the main features of the two algorithms.

Table 1
Performance of two algorithms combining non-invasive markers for liver fibrosis to diagnose significant fibrosis (≥F2 by METAVIR) in 188 HCV patients

<table>
<thead>
<tr>
<th></th>
<th>Sebastiani’s algorithm</th>
<th>Leroy’s algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>APRI</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Patients in whom APRI was performed (%)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Fibrotest</td>
<td>60</td>
<td>100</td>
</tr>
<tr>
<td>Patients in whom Fibrotest was performed (%)</td>
<td>54</td>
<td>19</td>
</tr>
<tr>
<td>Liver biopsies was avoided (%)</td>
<td>100</td>
<td>91.2</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>100</td>
<td>91.2</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>69.3</td>
<td>97.3</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>83.1</td>
<td>98.1</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>100</td>
<td>87.9</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>87.8</td>
<td>93.6</td>
</tr>
<tr>
<td>AUROC (95% CI)</td>
<td>0.89 (0.72–0.97)</td>
<td>0.94 (0.86–0.99)</td>
</tr>
<tr>
<td>+ LR</td>
<td>3.2</td>
<td>3.0</td>
</tr>
<tr>
<td>− LR</td>
<td>0</td>
<td>0.09</td>
</tr>
</tbody>
</table>

APRI, AST-to-platelet-ratio; PPV, positive predictive value; NPV, negative predictive value; AUROC, area under the ROC curve; CI, confidence interval.

Table 1: Sebastiani’s algorithm: cutoff = 0.49 for significant fibrosis; Leroy’s algorithm: cutoff = 0.59 for significant fibrosis, 0.22 for no-mild fibrosis.
of our algorithm and those of Leroy’s algorithm. Both algorithms employed APRI 100% at baseline. The main difference between the two algorithms was that Fibrotest is again required for all cases according to Leroy’s algorithm, while our algorithm uses Fibrotest only in 60% of cases. Our algorithm presents with 100% NPV for the exclusion of significant fibrosis while Leroy’s algorithm showed 98% PPV for the prediction of significant fibrosis. Though the overall accuracy of both algorithms was excellent, with a slightly better performance of Leroy’s, the application of our algorithm resulted in a much greater reduction of liver biopsies (54% vs. 19%, \( p < 0.0001 \)).

Reducing the need for liver biopsies and Fibrotest implies a clear advantage in terms of risks, costs and better patient-compliance.

Leroy’s study confirms that most non-invasive markers do not overcome 75–85% accuracy in patients with chronic hepatitis C. Non-invasive markers, should be used sequentially while liver biopsy should be limited to the subset of patients with inaccurate response to non-invasive markers. Liver biopsy and non-invasive markers should be considered as agonists and not as antagonists towards the goal of correctly classifying the stage of liver fibrosis in patients with chronic hepatitis C.

References


Sebastiani Giada 1,2
Alberti Alfredo 2,3

1Digestive Diseases, Hepatology and Clinical Nutrition Department, Umberto I Hospital, Venice, Italy
2Venetian Institute of Molecular Medicine (VIMM), Padova, Italy
3Department of Clinical and Experimental Medicine, University of Padova, 35100 Via Giustiniani 2, Padova, Italy

E-mail address: alfredo.alberti@unipd.it

Available online 11 February 2008