Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease

Stuart McPherson,1 Stephen F Stewart,1 Elsbeth Henderson,1 Alastair D Burt,2 Christopher P Day2

ABSTRACT
Background Accurate evaluation of liver fibrosis in patients with non-alcoholic fatty liver disease (NAFLD) is important to identify patients who may develop complications. The aim of this study was to compare the diagnostic performance of simple non-invasive tests in identifying advanced fibrosis among patients with biopsy-proven NAFLD.

Methods Consecutive patients with biopsy proven NAFLD were recruited from the Newcastle Hospitals Fatty Liver Clinic from 2003 to 2009. The AST/ALT ratio, AST to platelet ratio index, BARD (weighted sum of BMI ≥ 28 = 1 point, AST/ALT ratio > 0.9 = 2 points, diabetes = 1 point), FIB-4 (age × AST (IU/l)/platelet count (×10^9/litre) × √ALT (IU/l)) and NAFLD fibrosis scores were calculated from blood tests taken at time of biopsy.

Results 145 patients (82 male (61%), mean age 51 ± 12 years) were included. The mean body mass index was 35.5 kg/m². 73 subjects (50%) had diabetes. 93 (19%) had advanced fibrosis (Kleiner stage 3—4). The FIB-4 score had the best diagnostic accuracy for advanced fibrosis (area under receiver operator characteristic curve (AUROC) 0.86), followed by AST/ALT ratio (AUROC 0.83), NAFLD fibrosis score (AUROC 0.81), BARD (AUROC 0.77) and AST to platelet ratio index (AUROC 0.67). The AST/ALT ratio, BARD score, FIB-4 and NAFLD fibrosis scores had negative predictive values greater than 90% (93%, 95%, 95% and 92% respectively). Positive predictive values were modest. In order to exclude advanced fibrosis liver biopsy could potentially be avoided in 69% with AST/ALT ratio, 62% with FIB-4, 52% with NAFLD fibrosis score and 38% with BARD.

Conclusions The ALT/AST ratio, FIB-4 and NAFLD fibrosis scores can reliably exclude advanced fibrosis in a high proportion of patients with NAFLD, allowing liver biopsy to be used in a more directed manner.

INTRODUCTION
Non-alcoholic fatty liver disease (NAFLD) is one of the most common causes of liver disease worldwide.1–3 In the USA approximately 30% of the population now has NAFLD.4 Obesity and insulin resistance are key risk factors for the development of NAFLD. The majority of patients with NAFLD have simple steatosis, which carries a good long-term prognosis.3–7 However, patients with non-alcoholic steatohepatitis (NASH) are at risk of progressive fibrosis, cirrhosis and hepatocellular carcinoma.7–10 As patients with NASH-related cirrhosis may have a poor prognosis,11 it is important to identify subjects with advanced fibrosis so they can be screened for complications of liver disease, such as varices and hepatocellular carcinoma and eventually entered into treatment trials aimed at reversing or preventing progression of fibrosis.

Currently, the ‘gold standard’ investigation for the assessment of hepatic fibrosis and inflammation is liver biopsy. However, this is invasive and may result in complications.13 In addition, liver biopsy only samples a small portion of the liver, which may result in sampling error.14 With the high prevalence of NAFLD in the population2 and the majority of patients having simple steatosis or low-grade NASH,15 liver biopsy may not be an appropriate investigation for many of these patients. Therefore, non-invasive tests that can reliably diagnose or exclude advanced fibrosis would be clinically beneficial to reduce the need for liver biopsy.
Several clinical scoring systems based on simple clinical or laboratory indices have been proposed to identify advanced fibrosis in patients with NAFLD and other liver diseases. These include the aspartate aminotransferase (AST)-to-platelet ratio index (APRI), the AST/alanine aminotransferase (ALT) ratio, the BARD score, the FIB-4 score and the NAFLD fibrosis score. However, many of these need to be validated prior to their widespread use. The aim of this study was to compare the diagnostic performance of a number of simple non-invasive tests in identifying advanced fibrosis in a cohort of subjects with biopsy-proven NAFLD from the UK.

PATIENTS AND METHODS
Consecutive patients with biopsy-proven NAFLD were recruited from the Newcastle Hospitals Fatty Liver Clinic from 2003 to 2009. Ethical approval was obtained from the Newcastle Hospitals Ethics Committee. Subjects who had been included in a previous study that derived the NAFLD fibrosis score were excluded from this study. Clinical and laboratory data was collected from the time of liver biopsy. Patients were excluded if they consumed more than 30 g of alcohol per day for males or more than 20 g per day for females. Patients who had evidence of coexistent liver disease were also excluded. In addition, patients whose liver biopsy was regarded as inadequate for staging purposes or who had incomplete data to calculate all the non-invasive scores were excluded.

Relevant clinical details such as gender, age, weight, height and average alcohol intake (g/day) from the preceding 6 months were obtained from all patients at the time of liver biopsy. Information regarding average alcohol intake (g/day) prior to the last 6 months was also obtained. The body mass index (BMI) was calculated by the formula: weight (kg)/height (m)\(^2\). Waist circumference (in centimetres) was measured at the midpoint between the lower costal edge and upper iliac crest following a normal expiration. Patients were identified as having diabetes if they were taking an oral hypoglycaemic drug or insulin, or had been diagnosed with diabetes according to the 2004 American Diabetic Association criteria. Blood test results from the time of liver biopsy or within 3 months were recorded.

The APRI was calculated as AST (IU/l)/(upper limit of normal)/platelet count (\(\times 10^9\)/litre)\(^{16}\). The FIB-4 score was calculated: age \(\times\) AST (IU/l)/platelet count (\(\times 10^9\)/litre)\(^{18}\). The NAFLD fibrosis score was calculated according to the following formula: \(-1.675+0.037\times \text{age (years)}+0.094\times \text{BMI (kg/m}^2\)+1.13\times \text{impaired fasting glycaemia or diabetes (yes=1, no=0)}+0.99\times \text{AST/ALT ratio}−0.013\times \text{platelet (\(\times 10^9\)/litre)}−0.66\times \text{albumin (g/dl)}\)\(^{20}\). The BARD score was the weighted sum of three variables (BMI > 28 = 1 point, AST/ALT ratio > 0.8 = 2 points, diabetes = 1 point)\(^{18}\).

Percutaneous liver biopsies were performed using an 18G BioPince liver biopsy system (Medical Devices Technologies, Gainville, Florida, USA) or a Menghini needle. Liver biopsy specimens were assessed by an experienced hepatopathologist (AB). Histological scoring was performed according to the NIH NAFLD Clinical Research Network criteria. A score of 5 or more by this system was required to make a diagnosis of NASH and a pre-requisite was a score of at least 1 for ballooning degeneration. Internal quality assurance systems demonstrated no significant intra-observer variation in the histological assessment.

All statistical analyses were performed using SPSS software version 14.0 (SPSS Inc.). Continuous normally distributed variables were represented as mean±SD. Categorical and non-normal variables were summarised as median and range. \(\chi^2\) tests were used to determine the distribution of categorical variables between groups. To compare the means of normally distributed variables between groups the Student t test was performed. To determine differences between groups for continuous non-normally distributed variables, medians were compared using the Mann–Whitney U test. The diagnostic performance of non-invasive tests was assessed by receiver operating characteristic (ROC) curves. The area under the ROC (AUROC) was used as an index to compare the accuracy of tests. The sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV) for relevant cut-offs were also displayed.

RESULTS
A total of 217 patients were identified, but 72 were excluded (65 incomplete data and seven had participated in a previous study\(^{26}\)). Therefore, data for 145 patients was analysed. The demographic and laboratory characteristics of all patients are shown in table 1. Eighty-eight patients (61%) were male and the mean age was 51±12 years. One hundred and twenty-seven patients (87%) were obese (BMI ≥ 30) and the mean BMI was 35±5 kg/m\(^2\). Ninety-three patients (63%) had NASH on liver biopsy and 27 patients (19%) had advanced fibrosis (Kleiner fibrosis stage 3 or 4).

As the identification of patients with advanced fibrosis is of clinical importance, the clinical and laboratory features of subjects with no/mild fibrosis (stage 0–2) were compared with...
patients with advanced fibrosis (stage 3–4) and the results are shown in table 2. As expected, in comparison with subjects with no/mild fibrosis, patients with advanced fibrosis were significantly older (p<0.001), more likely to be female (p=0.02) and had lower platelets counts (p=0.001), serum cholesterol levels (p=0.01) and triglyceride (p<0.001) levels. Interestingly, patients with advanced fibrosis had significantly lower serum ALT levels than patients with no/mild fibrosis (p=0.001), but the AST levels were similar (p=0.65), resulting in lower AST/ALT ratios in patients with advanced fibrosis. In addition, the APRI (p=0.01), BARD score (p<0.001), FIB-4 score (p<0.001) and NAFLD fibrosis score (p<0.001) were all significantly higher in patients with advanced fibrosis, compared with subjects with no/mild fibrosis.

In order to compare the diagnostic accuracy of the scoring systems, ROC curves were used (figure 1). Using this method, the FIB-4 score had the best diagnostic accuracy for advanced fibrosis (AUROC 0.86), followed by AST/ALT ratio (AUROC 0.83), NAFLD fibrosis score (AUROC 0.81), BARD score (AUROC 0.77) and APRI (AUROC 0.67). The sensitivity, specificity, positive predictive values (PPVs) and negative predictive values (NPVs) of each test using previously published cut-offs are shown in table 3.16 18 20 23 The FIB-4 and BARD score had the highest NPVs of 95% using their lower cut-offs, but the AST/ALT ratio and NAFLD score also performed well with NPVs of 92% and 93% respectively. The NAFLD fibrosis score and FIB-4 had the highest PPVs of 79% and 75% respectively.

As the NPVs for AST/ALT ratio, BARD score, FIB-4 score and NAFLD fibrosis score were all greater than 90% using their lower cut-offs, these tests may have sufficient accuracy to be used clinically to exclude advanced fibrosis. Using this approach, a significant proportion of patients could avoid liver biopsy using each of these tests (table 4). As the PPV were modest for all non-invasive tests, ranging from 27% to 79%, it was felt they were not accurate enough to be used as an alternative to liver biopsy.

### Table 2 Comparison between patients with fibrosis stages 0–2 and 3–4

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No/mild fibrosis (stage 0–2)</th>
<th>Advanced fibrosis (stage 3–4)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>49±12</td>
<td>61±8</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Gender (%male)</td>
<td>65%</td>
<td>41%</td>
<td>0.02†</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>35±5</td>
<td>36±7.6</td>
<td>0.13*</td>
</tr>
<tr>
<td>BMI ≥28</td>
<td>96%</td>
<td>100%</td>
<td>0.94†</td>
</tr>
<tr>
<td>Diabetes</td>
<td>49%</td>
<td>56%</td>
<td>0.54‡</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>111±11</td>
<td>113±12</td>
<td>0.34*</td>
</tr>
<tr>
<td>ALT (IU/l)</td>
<td>101±6</td>
<td>68±40</td>
<td>0.001*</td>
</tr>
<tr>
<td>AST (IU/l)</td>
<td>63±44</td>
<td>67±43</td>
<td>0.65*</td>
</tr>
<tr>
<td>GGT (IU/l)</td>
<td>122±107</td>
<td>262±398</td>
<td>0.08*</td>
</tr>
<tr>
<td>ALB (g/l)</td>
<td>45±3</td>
<td>44±6</td>
<td>0.6*</td>
</tr>
<tr>
<td>Platelets (×10³/litre)</td>
<td>267±90</td>
<td>204±78</td>
<td>0.001*</td>
</tr>
<tr>
<td>Cholesterol (mmol/l)</td>
<td>5.5±1.3</td>
<td>4.8±1.3</td>
<td>0.01*</td>
</tr>
<tr>
<td>Triglycerides (mmol/l)</td>
<td>2.9±2.4</td>
<td>1.8±0.9</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>APRI</td>
<td>0.65±0.52</td>
<td>0.94±0.58</td>
<td>0.01*</td>
</tr>
<tr>
<td>AST/ALT ratio</td>
<td>0.67±0.23</td>
<td>1.05±0.39</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>BARD score</td>
<td>2 (0–4)</td>
<td>3 (1–4)</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>FIB-4 score</td>
<td>1.25±0.79</td>
<td>2.83±1.71</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>NAFLD score</td>
<td>1.83±1.52</td>
<td>1.52±1.75</td>
<td>&lt;0.001*</td>
</tr>
</tbody>
</table>

*Student t test. †Mann–Whitney U test. ‡z² test.

**Table 2** Comparison between patients with fibrosis stages 0–2 and 3–4.

**Discussion**

The major finding of this study was that the AST/ALT ratio, BARD score, FIB-4 score and NAFLD fibrosis score all had high NPV (=92%) for advanced fibrosis in a cohort of patients with NAFLD. This suggests they could be used clinically to exclude advanced fibrosis in subjects with NAFLD. Using the simplest of these systems, the AST/ALT ratio (<0.8), to exclude advanced fibrosis, liver biopsy could have been avoided in 69% of patients in our cohort and 93% would have been classified correctly. Similarly, liver biopsy could have been avoided in more than half of patients using the FIB-4 and NAFLD fibrosis scores with similar accuracy (table 4). Given the large numbers of patients with NAFLD who are currently being referred to liver clinics for evaluation, use of these non-invasive tests could substantially reduce the number of liver biopsies being performed. This would result in significant benefit to patients by directing liver biopsy to those more likely to have advanced liver disease, as well as lead to cost savings.

In contrast to the NPV, the PPV for each test were modest ranging from 27% to 79%. Therefore, these tests do not have sufficient accuracy to be used to diagnose advanced fibrosis, particularly in view of the seriousness of this diagnosis. It would therefore seem appropriate to consider liver biopsy in all patients who have a value above the lower cut-off for the chosen non-invasive score. Clearly, liver biopsy may also be indicated in individuals for whom the diagnosis in uncertain or where a coexistent disease may be suspected.

The major advantage of using any of these simple scoring systems is that they are derived from readily available clinical and laboratory indices. Clearly, the AST/ALT ratio and BARD score are the simplest and can be calculated easily when a patient is reviewed in the clinic. The FIB-4 and NAFLD fibrosis scores require more complex calculation, but the relevant details can easily be entered onto a pre-designed Excel spreadsheet that can produce an instant result in front of the patient. Therefore, introduction of the use of these tests into daily practice should be relatively simple and will not result in extra costs.

A number of other panels of serum markers of fibrosis such as the Enhanced Liver Fibrosis panel and Fibrotest have been
evaluated in patients with NAFLD.24,25 These tests are relatively expensive as they involve the measurement of markers of matrix turnover. The Enhanced Liver Fibrosis panel has been assessed in a large cohort of patients with NAFLD, but performed only marginally better than the NAFLD fibrosis score in predicting advanced fibrosis (AUROC 0.95 vs 0.89).24 More recently hepatic fibrosis was assessed by liver stiffness measurement using transient elastography in a cohort of patients with NAFLD from Hong Kong and France.26 In that study, liver stiffness measurement was accurate for excluding advanced fibrosis (Kleiner stage 3–4) and performed better than simple non-invasive scores (AST/ALT ratio, APRI, BARD, FIB-4 and NAFLD fibrosis score). However, transient elastography is not widely available and not validated in this disease group. In addition, the success rate was poor in patients with a BMI>35, which represented 46% of patients in our cohort. A new probe for these tests as clinical tools to exclude advanced fibrosis may perform better than each alone, but this needs to be assessed in future studies.

Currently, there is very limited guidance for general practitioners regarding which patients with NAFLD to refer to secondary care for further evaluation. Many doctors overly rely on the ALT level as a surrogate marker of disease severity. However, interestingly, in this study patients with advanced fibrosis had significantly lower ALT levels than those with no/mild fibrosis, so clearly the ALT has no role in identifying patients with advanced disease. This finding was also observed in a previous study.18 As it is not realistic for all patients who have NAFLD to be referred to liver services for evaluation, clear guidance for general practitioners is needed. In this study, the increasing AST/ALT ratio was associated with advanced fibrosis and a cut-off value of >0.8 was associated with higher risk of advanced fibrosis. Therefore, an AST/ALT ratio of >0.8 could be used as a screening tool to determine which patients with NAFLD should be referred to secondary care for evaluation.

One limitation of this study is that it took place in a tertiary liver centre where there may be selection bias. Our cohort was very obese with a mean BMI of 35 kg/m² and 50% were diabetic. Approximately two thirds of the patients had NASH, which is a higher proportion of patients with progressive disease than would be seen in a more general setting. However, we propose these tests as clinical tools to exclude advanced fibrosis, and in populations with less severe disease these tests may have even higher NPV. Therefore, the results of this study should have wide applicability, but this needs to be validated.

The simple non-invasive scoring systems evaluated in the present study have a role in the assessment of fibrosis, but are not effective in differentiating patients with simple steatosis from those with NASH. This will be of significant clinical importance when effective therapies that prevent progression of NASH are available. Other serum tests such Nashstest and caspase-generated CK18 fragments have shown promise in the differentiation between simple steatosis and NASH, but need further evaluation before entering routine practice.20,29

In conclusion, this study showed that the AST/ALT ratio, BARD score, FIB-4 score and NAFLD fibrosis score can reliably exclude advanced fibrosis in subjects with NAFLD. Introduction of these scores in clinical practice may reduce the proportion of patients that require liver biopsy to diagnose mild disease.

### Competing interests None.

### Ethics approval
This study was conducted with the approval of the Newcastle Hospitals Ethics Committee.

### Provenance and peer review
Not commissioned; externally peer reviewed.

### References


Management of post-operative intestinal failure

This is an introduction to the Gut tutorial “Management of post-operative intestinal failure” hosted on BMJ Learning—the best available learning website for medical professionals from the BMJ Group.

The management of post-operative intestinal failure is presented around a case of elective small bowel resection. The tutorial aims to describe an evidence based approach, and to focus the reader’s consideration of the key management steps and issues in nutrition support. The objectives of this tutorial are to illustrate and inform on the assessment of a high output stoma patient, management of fluid and electrolyte balance in a patient, determination of energy and protein needs, and the rates of replacement and assessment as well as management of a parenterally fed patient with sepsis.

To access the tutorial (Interactive Case History), click on BMJ Learning: Take this module on BMJ Learning from the content box at the top right and bottom left of the online article. For more information please go to: http://gut.bmj.com/site/tutorials/

If prompted, subscribers must sign into Gut with their journal username and password. All users must also complete a one-time registration on BMJ Learning and subsequently log in (with a BMJ Learning username and password) on every visit.

Richard Johnston, Tim E Bowling

Nottingham Digestive Diseases Centre and Biomedical Research Unit, Queen’s Medical Centre, Nottingham, UK

Correspondence to Dr Richard Johnston, Clinical Research Fellow in Nutrition, Nottingham Digestive Diseases Centre and Biomedical Research Unit, Queen’s Medical Centre, Nottingham NG7 2UH, UK; rjohnston75@doctors.org.uk

Competing interests None.

Provenance and peer review Commissioned; not externally peer reviewed.

Gut 2010;59:1269. 10.1136/gut.2010.225110