The NAFLD Fibrosis Score: A Noninvasive System That Identifies Liver Fibrosis in Patients with NAFLD

Paul Angulo,1 Jason M. Hui,2 Giulio Marchesini,3 Elisabetta Bugianesi,4 Jacob George,2 Geoffrey C. Farrell,2 Felicity Enders,5 Sushma Saksena,6 Alastair D. Burt,6 John P. Bida,5 Keith Lindor,1 Schuyler O. Sanderson,7 Marco Lenzi,8 Leon A. Adams,1 James Kench,9 Terry M. Therneau,5 and Christopher P. Day6

Patients with nonalcoholic fatty liver disease (NAFLD) and advanced liver fibrosis are at the highest risk for progressing to end-stage liver disease. We constructed and validated a scoring system consisting of routinely measured and readily available clinical and laboratory data to separate NAFLD patients with and without advanced fibrosis. A total of 733 patients with NAFLD confirmed by liver biopsy were divided into 2 groups to construct (n = 480) and validate (n = 253) a scoring system. Routine demographic, clinical, and laboratory variables were analyzed by multivariate modeling to predict presence or absence of advanced fibrosis. Age, hyperglycemia, body mass index, platelet count, albumin, and AST/ALT ratio were independent indicators of advanced liver fibrosis. A scoring system with these 6 variables had an area under the receiver operating characteristic curve of 0.88 and 0.82 in the estimation and validation groups, respectively. By applying the low cutoff score (1.455), advanced fibrosis could be excluded with high accuracy (negative predictive value of 93% and 88% in the estimation and validation groups, respectively). By applying the high cutoff score (0.676), the presence of advanced fibrosis could be diagnosed with high accuracy (positive predictive value of 90% and 82% in the estimation and validation groups, respectively). By applying this model, a liver biopsy would have been avoided in 549 (75%) of the 733 patients, with correct prediction in 496 (90%).

Conclusion: a simple scoring system accurately separates patients with NAFLD with and without advanced fibrosis, rendering liver biopsy for identification of advanced fibrosis unnecessary in a substantial proportion of patients. (HEPATOLOGY 2007;45:846-854.)

Abbreviations: BMI, body mass index; HDL, high-density lipoprotein; HOMA, homeostatic model assessment; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; ROC, receiver operating characteristics curve.

From the 1Division of Gastroenterology and Hepatology, 3Department of Health Sciences Research, and 7Division of Anatomic Pathology, Mayo Clinic College of Medicine, Rochester, MN; 2Storr Liver Unit and 9Department of Anatomical Pathology, Westmead Hospital and Westmead Millennium Institute, University of Sydney, Australia; 3Department of Medicine and 8Pathology, Alma Mater Studiorum University, Bologna, Italy; 4Division of Gastroenterology, University of Turin, Ospedale San Giovanni Battista, Turin, Italy; 6School of Clinical Medical Sciences, University of Newcastle and Freeman Hospital, Newcastle upon Tyne, UK.

Received May 19, 2006; accepted November 2, 2006.

Address reprint requests to: Paul Angulo, M.D., Miles and Shirley Fiterman Center for Digestive Diseases, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905. E-mail: angulohernandez.paul@mayo.edu; fax: 507-266-4531.

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.21496

Potential conflict of interest: Nothing to report.
based on the severity of fibrosis. However, liver biopsy is an expensive and invasive procedure associated with a number of complications and prone to sampling error. Because of all this, efforts have been made to identify noninvasive indicators of liver fibrosis in patients with NAFLD. Noninvasive approaches for assessing the severity of fibrosis in NAFLD have included a combination of clinical features and routine laboratory investigations as well as some less readily available serum markers of fibrosis. These noninvasive approaches are, however, either insufficiently accurate in their prediction of liver fibrosis or their diagnostic accuracy has been evaluated in only a limited number of patients with NAFLD; further, most have not been validated in a separate population of NAFLD patients. A model using routinely measured clinical and laboratory variables that accurately predicts the severity of liver fibrosis in patients with NAFLD has therefore yet to be developed and validated. Hence, the purpose of this study was (1) to develop a simple noninvasive scoring system aimed at separating patients with NAFLD with and without advanced liver fibrosis by using routinely determined and easily available clinical and biochemical variables, and (2) to validate the results in a separate cohort of patients. The NAFLD fibrosis score we developed is a clinically applicable and useful method to separate patients with and without prognostically significant NAFLD.

**Materials and Methods**

**Patient Population.** A total of 733 patients with well-characterized and liver biopsy–confirmed NAFLD were included in this study. They were untreated, consecutively biopsied patients seen at the Mayo Clinic in Rochester, MN (n = 356); Newcastle, UK (n = 158); Sydney, Australia (n = 123); and Italy (n = 96) during 2000-2003. The study was approved by appropriate regulatory bodies at all centers.

The diagnosis of NAFLD was based on the following criteria: (1) elevated aminotransferases (AST and/or ALT); (2) liver biopsy showing steatosis in at least 10% of hepatocytes; and (3) appropriate exclusion of liver disease of other etiology including alcohol-induced or drug-induced liver disease, autoimmune or viral hepatitis, and cholestatic or metabolic/genetic liver disease. These other liver diseases were excluded using specific clinical, biochemical, radiographic, and/or histological criteria. Aminotransferase levels had been elevated for a mean of 17.5 months (median 9.0 months, interquartile range 6, 20). All patients had a negative history of ethanol abuse as indicated by a weekly ethanol consumption of <140 g in women and <210 g in men. History of alcohol consumption was specifically investigated by interviewing the patients, and in almost all cases, a close relative. In the Newcastle center, alcohol levels in urine were measured randomly to rule out patients who abused alcohol. Patients with clinical or imaging evidence of decompensated cirrhosis (i.e., portal systemic encephalopathy, variceal bleeding, ascites) were specifically excluded from this study because they most likely had cirrhotic-stage NAFLD regardless of what a model may predict.

Clinical and laboratory data were collected on the date a diagnostic liver biopsy was performed. A complete medical history and physical examination was undertaken in all patients. Body mass index (BMI) was calculated using the formula: weight (in kilograms)/height (in meters²). Waist circumference (to the nearest half centimeter) was measured at the midpoint between the lower border of the rib cage and the iliac crest. Laboratory evaluation included routine liver biochemistry (ALT and AST levels, total bilirubin, albumin, alkaline phosphatase, and gamma glutamyl transpeptidase); complete blood count; total cholesterol, HDL cholesterol, and total triglycerides; fasting glucose; fasting insulin; ferritin levels; viral serology for hepatitis B and C infection; autoantibodies; alpha 1 antitrypsin levels and phenotype; and ceruloplasmin levels.

The degree of insulin resistance was determined by the homeostatic model assessment (HOMA) using the formula: insulin resistance = (insulin × glucose)/22.5. HOMA correlates well with the “gold standard” hyperinsulinemic euglycemic clamp technique, but HOMA is more appropriate for studies with large patient populations. Components of the metabolic syndrome were recorded including central obesity (waist circumference >102 cm for men and >88 cm for women; or ≥90 cm in Asian men and ≥80 cm in Asian women), hyperglycemia (fasting blood glucose ≥110 mg/dl or previously diagnosed type 2 diabetes), hypertriglyceridemia (triglycerides ≥150 mg/dl or under treatment for this lipid abnormality), hypertension (blood pressure ≥130/≥85 or treatment of previously diagnosed hypertension), and low HDL cholesterol (<40 mg/dl in men or <50 mg/dl in women). The presence of diabetes mellitus (fasting glucose ≥126 mg/dl or treatment with antidiabetic drugs), obesity (BMI ≥30 kg/m², or ≥25 kg/m² in Asians), and overweight (BMI 25-29.9 kg/m², or 23-24.9 kg/m² in Asians) was also recorded. The AST/platelet ratio was calculated because it has been previously associated with the severity of fibrosis in patients with chronic hepatitis C infection using the formula: AST (xULN) × 100/platelet count (10⁹/l).

**Liver Histology.** Liver biopsies were routinely stained with hematoxylin and eosin, Masson’s trichrome, and
special stains for iron and copper. Liver biopsies were read by a single liver pathologist in each participating center, and the stage of fibrosis was scored based on the 5-point scale proposed by Brunt et al. and recently modified by Kleiner et al. Briefly, stage 0 = absence of fibrosis; stage 1 = perisinusoidal or portal; stage 2 = perisinusoidal and portal/periporal; stage 3 = septal or bridging fibrosis; and stage 4 = cirrhosis. To control for biopsy size, the length of the biopsy was measured with a hand ruler, and the number of portal areas on one cross-section was counted.

**Statistical Analysis.** The main endpoint of the study was to predict the presence or absence of advanced fibrosis (stages 3-4) by a combination of simple and clinically relevant variables. Data from each of the 4 countries were randomly separated into 2/3 and 1/3 of patients for model building and model validation, respectively. Hence, data on 480 patients were used to build a model, whereas data on 253 patients were used to validate the model. Univariate descriptive statistic was used to compare patients with and without significant fibrosis. All variables were included in a multivariate backward stepwise logistic regression analysis to identify variables independently associated with presence or absence of advanced fibrosis. The variables in the resulting model were assessed for all 2-way interactions, as well as interaction with other variables suggested by clinical knowledge. The results of the multivariate analysis were adjusted for site. Those variables with \( P < 0.05 \) by multivariate analysis were used to construct a scoring system to predict advanced fibrosis. The overall diagnostic accuracy of the scoring system was determined by calculating the area under the receiver operating characteristic (ROC) curve (the c-statistic) and its 95% confidence intervals. Validation was performed (1) in the validation dataset (\( n = 253 \)) and (2) in the full dataset (\( n = 733 \)). In both cases, cross-validation was used with 20 subgroups, so that at most 5% of the data under consideration was excluded at any one time. By employing cross-validation, the possibility of an unusually positive or negative validation subset could be assessed. ROC curve estimates from cross-validation analysis were found with the jackknife method, which provided a simple overall ROC curve from the 20 cross-validation groups. Using the ROC curve for the final model, 2 cutoff points were selected, so that the positive predictive value (PPV) and negative predictive value (NPV) for advanced fibrosis were at least 90%, on the assumption that false results of less than 10% are clinically acceptable. The diagnostic accuracy of the 2 cutoff points was determined by calculating sensitivity, specificity, PPV, NPV, and likelihood ratios. For each decile of probability of significant fibrosis, the full dataset was used to identify both the average value of the new scoring system and a 95% confidence interval (CI). All analyses were carried out using the statistical analysis software SAS Release 8.2 (SAS Institute Inc., Cary, NC).

**Results**

**Characteristics of the Patient Population.** Table 1 summarizes the clinical, laboratory, and liver biopsy data of the patient population. The 733 patients had an age range from 11 to 81 years, about half were female, and most were Caucasians. More than half were obese as defined by BMI, had central obesity as defined by waist circumference, or suffered from hypertriglyceridemia or low HDL cholesterol. About a third or more patients suffered from hyperglycemia/diabetes or hypertension. Of the total cohort, 244 (33%) patients did not have fibrosis on liver biopsy, 290 (40%) had stage 1-2 fibrosis, and 199 (27%) had advanced (stage 3-4) fibrosis.

**Liver Biopsy Size and Correlation with Advanced Fibrosis.** The mean (±SD) length of the liver biopsy was 18.7 ± 8.5 mm (median 17 mm, interquartile range 14, 22) in the total population, and similar between estimation (19 ± 8.4 mm, median 17 mm, interquartile range 14, 23) and validation (18.1 ± 8.8 mm; median 16 mm, interquartile range 13, 22) groups (\( P = 0.8 \)) (Table 1). The number of portal areas was 10 ± 4.5 (median 9.0, interquartile range 6, 15) in the total patient population, and not significantly different between estimation (9.9 ± 4.5; median 9, interquartile range 6, 15) and validation (10.1 ± 4.5; median 9, interquartile range 6, 15) groups (\( P = 0.8 \)). The liver biopsy length correlated significantly with the number of portal areas (\( r = 0.72, P < 0.001 \)) in the total patient population, and in the estimation (\( r = 0.70, P < 0.001 \)) and validation (\( r = 0.74, P < 0.001 \)) groups. Advanced fibrosis, however, did not correlate significantly with biopsy length (\( r = 0.04, P = 0.4 \)) or number of portal areas (\( r = 0.04, P = 0.6 \)) in the total patient population. Similarly, advanced fibrosis did not correlate significantly with biopsy length (\( r = 0.01, P = 0.9 \)) or number of portal areas (\( r = 0.1, P = 0.2 \)) in the estimation or validation (\( r = 0.04, P = 0.6 \) and \( r = 0.3, P = 0.8 \), respectively) groups. Further, the biopsy length and number of portal areas were not significantly associated with advanced fibrosis in univariate or multivariate analyses (Table 2).

**Predictors of Fibrosis.** Table 2 shows the univariate comparison and the results of the multivariate analysis performed in the 480 patients comprising the estimation group. By multivariate analysis, 5 variables remained significant including age, BMI, AST/ALT ratio, platelet
using these 6 variables, we constructed a scoring system (risk score formula) to distinguish between patients with (F3-F4) and without (F0-F2) advanced fibrosis. This scoring system had an area under the ROC curve of 0.88 ± 0.02 (95% CI = 0.85, 0.92) (Fig. 1). The regression formula (risk score) for prediction of severity of fibrosis based on these 6 variables is: NAFLD fibrosis score = −1.675 + 0.037 × age (years) + 0.094 × BMI (kg/m²) + 1.13 × IFG/diabetes (yes = 1, no = 0) + 0.99 × AST/ALT ratio − 0.013 × platelet (×10^9/l) − 0.66 × albumin (g/dl).

Using the area under the ROC curve, 2 cutoff points were selected to identify the presence (greater than 0.676) and absence (lower than 0.455) of significant fibrosis (Fig. 1).

By applying the low cutoff point (score below −1.455), 273 (77%) of the 355 patients without significant fibrosis were correctly identified, whereas 22
(7%) of 295 patients with a low cutoff point were incorrectly staged (Table 3). Thus, using this low cutoff point, the absence of advanced fibrosis could be excluded with high accuracy (negative predictive value of 93%).

By applying the high cutoff point (score above 0.676), 64 (50%) of 125 patients with advanced fibrosis were correctly identified, whereas only 7 (10%) of the 71 with a high cutoff point were incorrectly staged (Table 3). Using this high cutoff point, the presence of advanced fibrosis could be diagnosed with high accuracy (positive predictive value of 90%).

Overall, in the estimation group, the model predicted the presence or absence of advanced fibrosis in 337/480 = 71% of patients with a correct prediction in 337/366 or 92% [or 70% (337/480) of the total]. The incorrect prediction rate in the estimation group was only 7.9%. Thus, by applying the model to the estimation group, a liver biopsy would have been avoided in 366 (76%) patients and would be performed in only 114 (24%) of the 480 patients identified as “indeterminate”.

Validation of Results. The diagnostic accuracy of the scoring system in separating patients with and without advanced fibrosis was cross-validated in a separate set of 253 patients. The area under the ROC curve remained high in the validation set [0.82 ± 0.03 (95% CI = 0.76, 0.88)], and also after 20-fold cross-validation [0.84 ± 0.02 (95% CI = 0.81, 0.88)] (Figs. 1 and 2, respectively).

![Fig. 1. ROC curves of the scoring system in the estimation (n = 480) and validation (n = 253) groups combining 6 variables (age, BMI, hyperglycemia/diabetes, AST/ALT ratio, platelet count, and albumin) to distinguish between NAFLD patients with and without advanced fibrosis. The area under the ROC curve for the estimation and validation groups is 0.88 ± 0.02 (95% confidence intervals, 0.85, 0.92) and 0.82 ± 0.03 (95% confidence intervals, 0.76, 0.88) respectively.](image-url)
By applying the low cutoff point (score below \(-1.455\)), 127 (71%) of the 179 patients without advanced fibrosis were correctly identified, whereas 17 (12%) of 144 with a low cutoff point were incorrectly staged (Table 4). Thus, using this low cutoff point, the absence of advanced fibrosis could be excluded with high accuracy (negative predictive value of 88%).

By applying the high cutoff point (greater than 0.676), 32 (43%) of the 74 patients with advanced fibrosis were correctly identified, whereas only 7 (18%) of the 39 patients with a high cutoff point were incorrectly staged (Table 4). By using this high cutoff point, the presence of advanced fibrosis could be diagnosed with high accuracy (positive predictive value of 82%).

Overall, in the validation group, the model identified presence or absence of advanced fibrosis in \((144 + 39)/253 = 72\%\) of patients with a correct prediction in \(159/183 = 87\%\) [or 63\% (159/253) of the total]. The incorrect prediction rate in the validation group was only \((17 + 7)/183 = 13.1\%\). Thus, by applying the model to the validation group, a liver biopsy would have been avoided in 183 (72\%) patients and would be performed in only 70 (28\%) of the 253 patients identified as “indeterminate”.

### Predictive Values of the Model for Different Prevalence of Significant Fibrosis.

The prevalence of advanced fibrosis in the 4 centers was 12.5\% (Italy), 19.6\% (Newcastle), 30.9\% (Sydney), and 29.8\% (Mayo). Therefore, we calculated positive and negative predictive values of the 2 cutoff points using a wide range of prevalence of advanced fibrosis varying from 5\% to 50\%. The NPV of the low cutoff point to rule out advanced fibrosis remained high (\(\geq \text{87\%},\ Table 5\)). The PPV of the high cutoff point to diagnose advanced fibrosis also remained high, particularly for prevalence of advanced fibrosis of 10\% or more (\(\geq \text{87\%},\ Table 5\)). Thus, these 2 cutoff points may be useful to predict the severity of liver fibrosis in patients with NAFLD seen in medical centers with different prevalences of advanced fibrosis.

The estimated range of the NAFLD fibrosis score is shown in Table 6. Across the full range of the probability of stage 0-2 fibrosis (0\% to 100\%), the NAFLD fibrosis score is translated to a negative, positive, or indeterminate result.

### Discussion

In this study, we developed and validated a simple noninvasive scoring system composed of routinely measured and easily available clinical and laboratory variables to discriminate between the presence or absence of advanced fibrosis in patients with NAFLD. This index, which we call the “NAFLD fibrosis score”, was accurate in distinguishing the severity of fibrosis. Using values below the lower or above the higher cutoff points, a prediction of absence or presence of advanced fibrosis was made in 549
(75%) of the 733 patients of the total cohort, and this prediction was correct in 496 (90%) of these 549 individuals. Only 184 (25%) patients of the total cohort of 733 were considered “indeterminate”. This implies that by applying the NAFLD fibrosis score, liver biopsy could have been avoided in 75% (549 of 733) of patients in the total cohort.

The potential diagnostic accuracy of serum markers of fibrosis has been evaluated by others.19-21 The European liver fibrosis group assessed the combination of age and serum levels of hyaluronic acid, aminoterminal propeptide of type 3 collagen, and tissue inhibitor of matrix metalloproteinase 1 in predicting advanced fibrosis in patients with a wide range of liver disease.19 The proposed algorithm had an acceptable accuracy overall, but only 61 of the 912 patients studied had NAFLD, a number that is too small to derive meaningful conclusions for the NAFLD population. Further, the lack of availability of these serum markers of fibrosis in most centers makes it difficult to apply the proposed scoring system on a daily basis. A French group recently reported their experience with the FibroTest, a combination of 5 serum markers, in the prediction of advanced fibrosis (stages 3-4) in 267 patients with NAFLD.22 In that study,22 the FibroTest value was in between the proposed cutoffs of 0.30 and 0.70 in 88 patients, and thus, unable to predict the presence or absence of advanced fibrosis in 33% of the patients, a proportion slightly higher than the 25% of indeterminate cases in our series.

The combination of serum levels of hyaluronic acid and type 6 collagen 7S domain20 and serum YKL-4021 levels also have been proposed as a diagnostic tool on the basis of small studies that lacked a validation group. More recently, liver stiffness measured by the ultrasonographic-based FibroScan has been proposed as a useful noninvasive method for the identification of advanced liver fibrosis in patients with chronic hepatitis C infection.26,27 A recent prospective study of 2,114 FibroScan examinations, however, found presence of BMI greater than 28 kg/m² to be the only independent factor associated with failure of FibroScan examination for the identification of liver fibrosis.28 BMI of 28 kg/m² or greater is almost a universal finding in patients with NAFLD and thus the FibroScan’s utility in fibrosis quantification in patients with NAFLD needs further evaluation.

### Table 4. Predictive Value of the Scoring System Obtained from the Validation Group (n = 253)

<table>
<thead>
<tr>
<th>Low Cutoff Point (&lt; −1.455)</th>
<th>Indeterminate (−1.455-0.676)</th>
<th>High Cutoff Point (&gt;0.676)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>144</td>
<td>70</td>
<td>39</td>
</tr>
<tr>
<td>No significant fibrosis (stage 0-2)</td>
<td>127</td>
<td>45</td>
<td>7</td>
</tr>
<tr>
<td>Significant fibrosis (stage 3-4)</td>
<td>17</td>
<td>25</td>
<td>32</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>77%</td>
<td>43%</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>71%</td>
<td>96%</td>
<td></td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>52%</td>
<td>82%</td>
<td></td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>88%</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio (+)</td>
<td>2.652</td>
<td>11.058</td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio (−)</td>
<td>0.324</td>
<td>0.591</td>
<td></td>
</tr>
<tr>
<td>Interpretation</td>
<td>Absence of significant fibrosis (88% certainty)</td>
<td>Presence of significant fibrosis (82% certainty)</td>
<td></td>
</tr>
</tbody>
</table>

NOTES. Prevalence of advanced fibrosis of 29% in the validation group.

### Table 5. Predictive Values of the Cutoff Points for Different Prevalences of Advanced Fibrosis (n = 733)

<table>
<thead>
<tr>
<th>Prevalence of Significant Fibrosis (%)</th>
<th>Lower Cutoff Value (&lt; −1.455)</th>
<th>Higher Cutoff Value (&gt;0.676)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PPV (95% CI)</td>
<td>NPV (95% CI)</td>
</tr>
<tr>
<td>5</td>
<td>19 (14-25)</td>
<td>98 (97-100)</td>
</tr>
<tr>
<td>10</td>
<td>33 (26-39)</td>
<td>97 (95-99)</td>
</tr>
<tr>
<td>15</td>
<td>42 (35-49)</td>
<td>96 (93-98)</td>
</tr>
<tr>
<td>20</td>
<td>49 (42-56)</td>
<td>94 (91-97)</td>
</tr>
<tr>
<td>25</td>
<td>55 (47-62)</td>
<td>93 (90-96)</td>
</tr>
<tr>
<td>30</td>
<td>59 (52-66)</td>
<td>92 (88-95)</td>
</tr>
<tr>
<td>35</td>
<td>63 (56-70)</td>
<td>90 (87-94)</td>
</tr>
<tr>
<td>40</td>
<td>66 (59-73)</td>
<td>89 (85-93)</td>
</tr>
<tr>
<td>45</td>
<td>68 (62-75)</td>
<td>88 (84-92)</td>
</tr>
<tr>
<td>50</td>
<td>71 (64-77)</td>
<td>87 (83-90)</td>
</tr>
</tbody>
</table>
Our results suggest that by applying the NAFLD fibrosis score, a liver biopsy to determine severity of fibrosis would be required in only 25% of patients with NAFLD, that is, those identified as “indeterminate”. Most importantly, considering that most patients with NAFLD seen in clinical practice do not have advanced fibrosis (73% (534/733) of our cohort), the lower cutoff point was particularly accurate in ruling out the presence of advanced fibrosis; the NPV was 93% and 88% in the estimation and validation groups, respectively, and ranged from 87% to 98% for the prevalence of advanced fibrosis of 5% to 50%. Among our 733 patients, 439 (60%) had a negative diagnosis of advanced fibrosis (score below −1.455), and thus a liver biopsy would have been avoided by applying the NAFLD fibrosis score; of these 439, 400 (91%) indeed had stage 0-2 fibrosis.

Our study has several unique features. First, we included the largest cohort of patients with liver biopsy–proven NAFLD ever reported. Second, patients were untreated, consecutively biopsied individuals with NAFLD seen in different areas of the world, and thus, the population includes a wide ranging prevalence of fibrosis severity. Third, our predictive model consists of objective clinical and readily available laboratory variables that are routinely determined in patients with NAFLD in clinical practice, and no additional tests are necessary. Fourth, given the uneven distribution of fibrosis in the liver in patients with NAFLD, fibrosis severity predicted by our model may be a better reflection of fibrosis severity in the whole organ than fibrosis stage determined by liver biopsy.

Our study has some limitations. First, we included patients from different centers in the world that have a particular interest in studying NAFLD, and thus, some referral bias cannot be ruled out. Second, although we used a well-defined and acceptable scoring system to stage liver fibrosis, liver biopsies were read by independent liver pathologists at each center, and we were not able to quantify the effect on our results of some intraobserver and interobserver variability in fibrosis staging. However, quantification of fibrosis in patients with NAFLD has the lowest intraobserver and interobserver variability as compared to any other histological features. Further, it has been suggested that the level of pathologist experience has more influence on agreement than the characteristics of the biopsy specimens. Third, 90% of our patient population was Caucasian and 98% were 21 years of age or older. Fourth, we acknowledge that the severity of fibrosis in our study was determined by means of a percutaneous liver biopsy, which is prone to sampling error.

Due to these limitations, our results need to be validated in independent patient populations by other investigators. Also, further studies are necessary to determine the potential utility of our model in children and adolescents and in non–Caucasian patients with NAFLD as well as in patients with persistently normal aminotransferases, and in following fibrosis progression. In addition, further studies are needed to determine the potential benefits of diagnosing advanced fibrosis with our model, such as reinforcement of lifestyle measures and enrollment of patients with advanced fibrosis into screening programs for early detection of HCC and esophageal varices.

In summary, we demonstrate that a NAFLD fibrosis score constructed from routine clinical and laboratory variables can accurately predict the presence or absence of advanced fibrosis in NAFLD, rendering liver biopsy unnecessary in the vast majority of patients. It has to be determined, however, whether the addition of serum markers of fibrosis or imaging modalities increases the diagnostic accuracy of the NAFLD fibrosis score.

### References


