

Published in final edited form as:

Gastroenterology. 2013 October ; 145(4): 782–9.e4. doi:10.1053/j.gastro.2013.06.057.

Simple Noninvasive Systems Predict Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease

PAUL ANGULO¹, ELISABETTA BUGIANESI², EINAR S. BJORNSSON³, PHUNCHAI CHARATCHAROENWITTHAYA⁴, PETER R. MILLS⁵, FRANCISCO BARRERA⁶, SVANHILDUR HAFLIDADOTTIR³, CHRISTOPHER P. DAY^{7,§}, and JACOB GEORGE^{6,§}

¹Division of Digestive Diseases and Nutrition, University of Kentucky Medical Center, Lexington, Kentucky ²Division of Gastroenterology and Hepatology, Department of Medical Sciences, University of Torino, Torino, Italy ³Section of Gastroenterology and Hepatology, National University Hospital, Reykjavik, Iceland ⁴Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand ⁵Gartnavel General Hospital, Glasgow, UK ⁶Storr Liver Unit, Westmead Millennium Institute, University of Sydney and Department of Gastroenterology and Hepatology, Westmead Hospital, Westmead, New South Wales, Australia ⁷Institute of Cellular Medicine, Faculty of Medical Sciences, Newcastle University, Newcastle, United Kingdom

Abstract

BACKGROUND & AIMS—Some patients with nonalcoholic fatty liver disease (NAFLD) develop liver-related complications and have higher mortality than other patients with NAFLD. We determined the accuracy of simple, noninvasive scoring systems in identification of patients at increased risk for liver-related complications or death.

METHODS—We performed a retrospective, international, multicenter cohort study of 320 patients diagnosed with NAFLD, based on liver biopsy analysis through 2002 and followed through 2011. Patients were assigned to mild-, intermediate-, or high-risk groups based on cutoff values for 2 of the following: NAFLD fibrosis score, aspartate aminotransferase/platelet ratio index, FIB-4 score, and BARD score. Outcomes included liver-related complications and death or liver transplantation. We used multivariate Cox proportional hazard regression analysis to adjust for relevant variables and calculate adjusted hazard ratios (aHRs).

RESULTS—During a median follow-up period of 104.8 months (range, 3–317 months), 14% of patients developed liver-related events and 13% died or underwent liver transplantation. The aHRs for liver-related events in the intermediate-risk and high-risk groups, compared with the low-risk group, were 7.7 (95% confidence interval [CI]: 1.4–42.7) and 34.2 (95% CI: 6.5–180.1), respectively, based on NAFLD fibrosis score; 8.8 (95% CI: 1.1–67.3) and 20.9 (95% CI: 2.6–165.3) based on the aspartate aminotransferase/platelet ratio index; and 6.2 (95% CI: 1.4–27.2) and 6.6 (95% CI: 1.4–31.1) based on the BARD score. The aHRs for death or liver transplantation in the intermediate-risk and high-risk groups compared with the low-risk group were 4.2 (95% CI: 1.3–13.8) and 9.8 (95% CI: 2.7–35.3), respectively, based on the NAFLD fibrosis scores. Based

© 2013 by the AGA Institute

Address requests for reprints to: Paul Angulo, MD, Division of Digestive Diseases and Nutrition, University of Kentucky Medical Center, 800 Rose Street, Room MN649, Lexington, Kentucky 40536. paul.angulo@uky.edu; fax: 859-257-8860.

§Authors share co-senior authorship.

Supplementary Material Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at <http://dx.doi.org/10.1053/j.gastro.2013.06.057>.

Conflicts of interest

The authors disclose no conflicts.

on aspartate aminotransferase/platelet ratio index and FIB-4 score, only the high-risk group had a greater risk of death or liver transplantation (aHR = 3.1; 95% CI: 1.1–8.4 and aHR = 6.6; 95% CI: 2.3–20.4, respectively).

CONCLUSIONS—Simple noninvasive scoring systems help identify patients with NAFLD who are at increased risk for liver-related complications or death. NAFLD fibrosis score appears to be the best indicator of patients at risk, based on HRs. The results of this study require external validation.

Nonalcoholic fatty liver disease (NAFLD) encompasses a wide spectrum of liver pathology ranging from simple steatosis to nonalcoholic steatohepatitis (NASH), the latter characterized by steatosis plus features of cellular injury, such as inflammation and hepatocyte ballooning.¹ Some patients with NAFLD develop liver fibrosis, with a proportion progressing to cirrhosis and its complications of liver failure, portal hypertension, and hepatocellular carcinoma (HCC).^{2–4} Currently, cirrhotic-stage NAFLD represents the third⁵ or fourth⁶ most common indication for liver transplantation in the United States, and the second most common indication for liver transplantation in large transplantation centers.⁷ In addition, the prevalence of NAFLD-related cirrhosis has markedly increased in recent years as the underlying liver disease among patients transplanted for HCC in the United States.⁸ These data reflect the high prevalence of NAFLD in the general population, putting a substantial proportion of individuals at risk for NAFLD-associated morbidity and mortality.⁹

The long-term prognosis for individuals with NAFLD is not the same across the spectrum of the disease. Steatosis not associated with cellular injury or fibrosis follows a relatively benign clinical course, with an overall mortality similar to the general population of the same age and sex.^{10,11} For instance, <1% of patients with simple steatosis progressed to cirrhosis or died from liver-related complication after a mean follow-up of 15 years in a pooled analysis of several reported series.¹² However, patients with NASH, particularly those with increased fibrosis, have a worse prognosis as compared with an age- and sex-matched population.¹¹ The prevalence of cirrhosis and death related to liver complications is about 11% and 7%, respectively, in patients with NASH during the first 15 years of follow-up.¹² It has become clear that a subgroup of patients with NAFLD are at a higher risk for development of liver-related complications and death from liver-related causes. Unfortunately, other than presenting with overt cirrhosis or having a liver biopsy demonstrating advanced liver fibrosis, there is no accurate way to predict which subgroup of patients with NAFLD are at a higher risk for development of adverse long-term outcomes, including liver-related complications, liver transplantation, or death.

Several noninvasive scoring systems composed of routinely measured clinical and laboratory variables have been proposed to distinguish between patients with NAFLD with and without advanced liver fibrosis, including the NAFLD fibrosis score (NAFLD-FS),¹³ the AST/platelet ratio index (APRI),¹⁴ the FIB-4 score,¹⁵ and the BARD score.¹⁶ Most have been extensively validated for their accuracy in distinguishing between NAFLD patients with and without advanced fibrosis.^{17–19} It remains unknown, however, whether these scores can be used to identify the subgroup of patients with NAFLD who are at a higher risk for liver-related morbidity and mortality. We sought to determine the accuracy of these 4 scores in predicting the long-term outcomes of patients with NAFLD, including liver-related complications, transplantation, or overall mortality.

Patients and Methods

This was a retrospective, international, multicenter cohort study of 320 patients with well-characterized and liver biopsy-confirmed NAFLD. They were untreated, consecutively

biopsied patients that met the eligibility criteria as described here, and were recruited before 2002 from the following medical centers: University of Kentucky Medical Center, Lexington, KY; Westmead Hospital, Sydney Australia; Newcastle Hospitals National Health Service Foundation Trust in Newcastle-upon-Tyne, UK; National University Hospital, Reykjavik Iceland; Siriraj Hospital, Mahidol University, Bangkok, Thailand; Gartnavel General Hospital, Glasgow, UK; and Division of Gastroenterology and Hepatology University of Torino, Torino, Italy. The year 2002 was chosen in order to have up to a decade of follow-up for the last patient recruited into the study. The first subject included underwent a liver biopsy in 1984. Patients were retrospectively identified by reviewing the pathology database at each center of subjects with the pathology diagnosis of steatosis, steatohepatitis, or fatty liver. After an extensive review of the patients' medical records, only those with an unequivocal diagnosis of NAFLD were included in the analysis. The liver biopsy was performed to confirm the diagnosis of NAFLD after appropriate exclusion of liver disease of other etiology, such as 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, laboratory, radiographic, and/or histological criteria. Serology for viral hepatitis B and C was investigated in all subjects and all tested negative. Serology for hepatitis C virus was investigated either before the liver biopsy in those biopsied after 1991 when serology for hepatitis C virus became available, or on subsequent visits for those biopsied before hepatitis C virus testing was available. All patients had a negative history of alcohol 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 many cases, by also interviewing close relatives. No subjects underwent bariatric surgery before or during the study period, and none received treatment with vitamin E or a glitazone.

Extensive clinical and laboratory data were collected at the time the liver biopsy was performed. A complete medical history and physical examination were completed in all patients. The ethnicity (Hispanic or Latino, and not Hispanic or Latino) and race (white, Asian, black or African American, American Indian/Alaska native, Native Hawaiian or other Pacific Islander) of the patients were determined based on the categories proposed by the US Department of Health and Human Services Public Health Service.²⁰ Body mass index (BMI) was calculated using the formula: weight (in kilograms)/height (in meters)². 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 (alanine aminotransferase [ALT] and aspartate aminotransferase [AST] levels, total bilirubin, albumin, alkaline phosphatase, and γ -glutamyl transpeptidase); complete blood count; total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and triglycerides; fasting glucose; serum ferritin; transferrin saturation; viral serology for hepatitis B and C infection; autoantibodies; α 1 antitrypsin levels and phenotype; and ceruloplasmin levels. Components of the metabolic syndrome²¹ were recorded, including central obesity (waist circumference >102 cm for men and >88 cm for women; or 90 cm Asian men and 80 cm in Asian women), obesity (BMI \geq 30 or \geq 25 in Asians) and overweight (BMI 25–29.9 or 23–24.9 in Asians), hyperglycemia (fasting blood glucose \geq 6.1 mmol/L or previously diagnosed type 2 diabetes), hypertriglyceridemia (triglycerides \geq 1.7 mmol/L or under treatment for this lipid abnormality), hypertension (blood pressure \geq 130/85 mm Hg or treatment of previously diagnosed hypertension), and low high-density lipoprotein cholesterol (<1.04 mmol/L in men or <1.3 mmol/L in women). The presence of diabetes mellitus (fasting glucose \geq 7.0 mmol/L or treatment with antidiabetic drugs) was also recorded. Data on the use of statins during the study period were collected from review of medical records.

Four validated noninvasive scoring systems that were originally created to distinguish between patients with and without advanced (stage 3–4) liver fibrosis were calculated using the original reported formulas.^{13–16} They were the NAFLD-FS formula: $-1.675 + 0.037 \times \text{age (years)} + 0.094 \times \text{BMI} + 1.13 \times \text{hyperglycemia or diabetes (yes = 1, no = 0)} + 0.99 \times \text{AST/ALT ratio} - 0.013 \times \text{platelet (} \times 10^9/\text{L)} - 0.66 \times \text{albumin (g/dL)}$; the APRI formula: $\text{AST (} \times \text{upper limit of normal)} / \text{platelet (} 10^9/\text{L)} \times 100$; the FIB-4 score formula: $[\text{age (years)} \times \text{AST (U/L)} \text{platelet}(10^9/\text{L}) \times \text{ALT (U/L)}]$; the BARD score, scale 0–4: BMI $\geq 28 = 1$ point, AST to ALT ratio $\geq 0.8 = 2$ points; diabetes mellitus = 1 point. The values for the upper limit of normal for AST were set according to the International Federation of Clinical Chemistry, that is, 35 U/L for men and 30 U/L for women. The values for the upper limit of normal for ALT were 19 U/L in women and 30 U/L in men.²² Three risk categories (low, intermediate, high) for each score were created and based on the 2 cut points described in the original publications.^{13–15} They are -1.455 and 0.676 for the NAFLD-FS¹³; 0.5 and 1.5 for the APRI¹⁴; and 1.30 and 2.67 for the FIB-4 score¹⁵; a scale of 0/1, 2/3, and 4 was created for the BARD score.

Liver Histology

Liver biopsies were routinely stained with H&E, Masson's trichrome, and special stains for iron and copper. Liver biopsies were read by a single liver pathologist in each participating center. The stage of fibrosis (stage 0 to stage 4) was recorded based on the scoring system proposed by Kleiner et al.²³ Liver biopsy features, such as grade of steatosis, inflammation and cellular ballooning,²³ and presence of NASH²⁴ were also recorded. All biopsies were of appropriate size and included enough portal tracts for a confident pathological grading and staging of the histological features.

Of the 320 liver biopsies, 221 were reviewed again by a local liver pathologist for staging and grading. In the remaining 99 liver biopsies, data were obtained from the final biopsy report maintained in the medical records, which had enough, and accurate, descriptions of the severity and distribution of the histological lesions.

Follow-up

Patients were followed-up at variable intervals ranging from 3 to 12 months after the diagnosis of NAFLD. The follow-up was extended until the end of 2009 in 3 participating centers and until the end of 2011 in 4 centers. Patients were considered lost to follow-up when their health status was unknown within the last 24 months. At each visit, a complete medical history and physical examination were performed along with routine laboratory workup to follow their liver disease. Patients with cirrhosis underwent endoscopy screening for gastroesophageal varices and screening for HCC at regular intervals after standard of care recommendations or guidelines in place at specific times as proposed by liver societies.^{25,26} However, the decision to perform an upper gastrointestinal endoscopy and imaging of the abdomen in noncirrhotic patients was left up to the treating physician in each center. We did an extensive review of the medical records of all subjects; development of liver-related events was captured from the medical records when such events had been captured in the records by the treating physician. Given the lack of approved treatment for NAFLD, treatment recommendations specifically for NAFLD during the study period were similar in all centers and consisted of the standard recommendation to achieve and maintain appropriate body weight with increased physical activity and diet changes. However, there was no particular treatment regimen with a specific diet composition or specific type of physical activity used in all centers. At the end of the follow-up, the BMI was essentially unchanged and varied from 33.0 at baseline to 33.5 at the end of the follow-up ($P = .2$).

Statistical Analysis

Data are presented as median (interquartile range) or number (percentage) of patients with a condition, or otherwise specified. Standard parametric and nonparametric statistics were used for comparison of variables. The outcomes analyzed were liver-related events (ascites, gastroesophageal varices/bleeding, portosystemic encephalopathy, hepatopulmonary syndrome, HCC, spontaneous bacterial peritonitis, hepatorenal syndrome) and overall mortality (death from any cause or liver transplantation, whichever occurred first). The diagnostic accuracy of the 4 scoring systems to distinguish between patients with and without increased risk for the outcomes was investigated by determining the area under the receiver operating characteristic (ROC) curves (c statistics). We use a leave-one-out validation procedure or jackknife with all areas under the ROC curve computed and the mean area under the ROC curve calculated for each score. Cumulative liver-related events and overall mortality/liver transplantation during follow-up among the 3 risk categories were calculated using Kaplan–Meier analysis and compared by log-rank testing. Adjusted hazard rate ratio (aHR) estimates (relative risk) for outcomes were calculated by Cox proportional hazard regression analysis to control for the effect of potential risk factors (confounder), while taking into consideration varying lengths of follow-up. The Cox hazard regression models were adjusted by variables chosen a priori that were not included already in the scores, such as age, sex, race, BMI, diabetes, hypertension, use of statins for at least 6 months, and fibrosis stage. Time at risk (T_0) was from the date of liver biopsy to the date of outcome or last follow-up. Patients known to have, or who developed, any of the outcomes within 3 months of the liver biopsy procedure were eliminated a priori from the database. Only 11 of the 320 patients developed a liver-related event within 3 months of liver biopsy and they were not included in the analysis of liver-related outcomes. $P < .05$ was considered to indicate statistical significance. Statistical analyses were performed using IBM SPSS Statistics version 20.0 (IBM Corporation, Armonk, NY) software. The study was approved by appropriate regulatory bodies at all centers.

Results

Baseline Characteristics

Table 1 describes the baseline characteristics of the 320 patients. Median age was 52 years (interquartile range, 43–61 years) with a similar distribution of men and women. There was a predominance of white race and overweight or obese individuals, and about one third to more than a half of patients suffered from diabetes, hypertension, or dyslipidemia. Mean ALT and AST were about twice normal, and there was a uniform distribution of patients across the stages of fibrosis. Mean values of each of the 4 scores increased significantly from fibrosis stage 0 to stage 1–2 to stage 3–4 (Figure 1). During the study duration, 43 of 309 (14%) patients developed a total of 60 liver-related events and 41 of 320 patients (13%) died (36 patients) or underwent liver transplantation (5 patients).

Long-term Follow-up

Median duration of follow-up of the 320 patients was 104.8 months (range, 3–317 months). Complete follow-up was achieved in 80% of patients, with 20% considered lost to follow-up. Type and number of outcomes are described in Table 2. Development of gastroesophageal varices with or without bleeding, ascites, and portosystemic encephalopathy were the predominant liver-related events. Cardiovascular events followed by non-liver malignancy, including breast ($n = 5$), gastric ($n = 2$), colon ($n = 1$), pancreas ($n = 1$), kidney ($n = 1$) and cancer with unknown primary ($n = 1$), and complications of cirrhosis were the most common causes of death.

The diagnostic accuracy of the 4 scores to distinguish between patients with and without risk for the outcomes as indicated by ROC analysis is described in Table 3. For either outcome analyzed, the NAFLD-FS had the highest area under the ROC curve as compared with the other 3 scores.

Long-term Outcomes Based on the Score Risk Categories

The cumulative probability (unadjusted) obtained by Kaplan–Meier analysis of liver-related events and death/ liver transplantation was significantly different among the 3 risk categories for each of the 4 scores, as illustrated in Table 4 and Supplementary Figures 1A, B, C, and D and Supplementary Figures 2A, B, C, and D.

Kaplan–Meier analysis of the data showed a significantly shorter cumulative survival free of the outcome death/liver transplantation in patients with more advanced fibrosis when comparing fibrosis stage 0 vs 1/2 vs 3 /4 (log rank = 24.2; $P < .001$). Results were the same when patients were grouped into stage 0–2 vs 3–4 (log rank = 24.6; $P < .001$). There was no significant difference among grades of steatosis (grade 1 vs 2 vs 3; log rank = 0.52; $P = .8$) or between patients with and without definitive NASH (log rank = 0.40; $P .5$). As for the outcome of liver-related events, again there was a significant difference among stages of fibrosis (0 vs 1–2 vs 3–4; log rank = 51; $P < .001$) and between stage 0–2 vs 3–4 (log rank 7.7; $P .005$); but among grades of steatosis (log rank = 4.0; $P .26$) or between those with and without definitive NASH (log rank = 3.2; $P .07$).

The multivariate aHR for the outcomes of the 3 risk categories of each score are described in Table 5 (and detailed in Supplementary Tables 1 and 2). As compared with the low-risk category, both the intermediate-risk and high-risk categories of the NAFLD-FS, APRI, and BARD score were significantly associated with a higher likelihood to develop liver-related events; as for the FIB-4 score, only the high-risk category increased the likelihood to develop liver-related events as compared with the low-risk category. Both the intermediate-risk and high-risk categories of the NAFLD-FS significantly increased the likelihood to reach the outcome of death/liver transplantation as compared with the low-risk category; as for the APRI score and FIB-4 score, only the high-risk category increased significantly the likelihood of this outcome, whereas there was no difference among the 3 risk categories of the BARD score for the outcome of death/liver transplantation.

Discussion

Our study demonstrates that simple baseline noninvasive scores allow appropriate identification of patients with NAFLD at a higher risk of developing liver-related complications, or the outcome of death/liver transplantation. Given the ready availability of the data and the simplicity of the calculation, along with the relatively high accuracy in separating patients' risks, these scores seems to be valuable and practical tools that can be used clinically for patient counseling and monitoring. The ability of these scores to predict outcomes of liver-related events and death/liver transplantation was controlled in a multivariate Cox regression analysis for the stage of fibrosis on liver biopsy and for the use of statins as well as other variables not included in the individual scores that can potentially affect long-term outcomes.

The accuracy of the scores in separating patients' risk for long-term outcomes can be explained by the variables included in the scores. Low albumin level included in the NAFLD-FS is an indirect measurement of hepatic synthetic reserve and low platelet count included in all, but the BARD score is an indication of more advanced liver disease and portal hypertension.²⁷ The AST and ALT, included in all scores either as an AST/ALT ratio or AST/platelet ratio, are good indicators of more advanced fibrosis and cirrhosis not only in

NAFLD,²⁸ but also in other liver diseases, such as hepatitis C infection.²⁹ Having low values of albumin and platelets and high AST/ALT ratio or AST/platelet ratio will increase the scores, allowing the identification of patients with a higher risk for liver-related complications and liver-related death or need for liver transplantation. Other variables included in the scores, such as diabetes or hyperglycemia, older age, and greater BMI are well-known risk factors for mortality from cardiovascular disease³⁰ and malignancy.³¹ Suffering from diabetes, being older, and having more adiposity would increase the scores, allowing for the identification of patients with NAFLD with a higher risk of developing and dying from non liver-related complications, such as cardiovascular disease and cancer.

Although the study was not aimed at comparing the accuracy of one score vs another, the NAFLD-FS had the highest area under the ROC curve for either outcome, and it was the only 1 of the 4 with an aHR for both intermediate-risk and high-risk categories significantly greater than for the low-risk category in both outcomes analyzed (Table 5). The most common causes of death in our cohort were, by frequency order, cardiovascular, non-liver cancer, and cirrhosis complications. However, if both complications of cirrhosis and liver transplantation are counted together (Table 2), then liver-related causes would be the number one in our cohort.

The study has several strengths, including the large patient population with the diagnosis of NAFLD established by liver biopsy; the ability to adjust the accuracy of the scores in predicting outcomes by the severity of liver fibrosis; and the long follow-up of the cohort averaging almost a decade, and almost 3 decades in some cases. Having the diagnosis of NAFLD confirmed by a liver biopsy avoids the uncertainty of results and the conclusions of studies using ultrasonography for NAFLD diagnosis, given the well-known poor sensitivity of ultrasound in detecting fat infiltration of the liver when less than a third of the liver parenchyma has steatosis³² and in individuals with a BMI >40.³³ In addition, considering the close follow-up of the patients during the course of the disease, the study was able to determine the risk of not only overall mortality and liver transplantation, but also the development of liver-related events, which, on their own, contribute significantly to the disease morbidity.

We recognize that our study has some limitations, most of which are inherent in retrospective studies, including selection bias. There was no specific protocol developed a priori to routinely monitor patients for the development of liver-related events, such as gastroesophageal varices and HCC. Development of liver-related events was captured from the medical records when such events were recorded. It is possible that additional cases of varices, HCC, and other liver-related events were developed by some subjects but not identified. The figures reported in our article are most likely an underestimate. There was no unique specific protocol for NAFLD treatment and subjects were recommended to achieve appropriate weight control with lifestyle modifications, such as diet and increased physical activity. The intensity and composition of that lifestyle intervention was left up to the treating physician without uniformity among centers. This lifestyle intervention, however, was disappointing, with a BMI at the end of the study similar to that at baseline. Another limitation of our study is the use of a racial/ethnicity grouping,²⁰ which might not be applicable worldwide, and the need for external cross validation of our results. In addition, given the small number of Hispanics and African Americans included, our results cannot be extrapolated to these racial groups. Despite these limitations, however, we believe the results are reasonably robust and demonstrate that currently available noninvasive clinical decision aids (ie, scores) can predict long-term outcomes in patients with NAFLD.

In summary, our study demonstrates that simple noninvasive scoring systems, which are free and easy to calculate, can identify patients with NAFLD at higher risk for development of

liver-related complications and higher overall mortality. The scores can assist practicing clinicians in patient counseling and monitoring. Among the 4 scores analyzed, the NAFLD-FS seems the most accurate based on both ROC analysis and separating patient risk for long-term outcomes of liver-related complications and death or liver transplantation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Title and abstract edited by Gastroenterology Science Editor, Kristine Novak, PhD.

Funding

This study was supported by a National Institute of Health R01 DK82426 grant (to P. Angulo) and The European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. HEALTH-F2-2009-241762 for the project FLIP (to E. Bugianesi). J. George is supported by grants from the Sydney Medical Foundation and grants from the National Health and Medical Research Council (632630 and 1049857). These sponsors played no role in the study design or the collection, analysis, and interpretation of data.

Abbreviations used in this paper

| | |
|-----------------|-----------------------------------|
| aHR | adjusted hazard ratio |
| ALT | alanine aminotransferase |
| APRI | AST/platelet ratio index |
| AST | aspartate aminotransferase |
| BMI | body mass index |
| HCC | hepatocellular carcinoma |
| NAFLD | nonalcoholic fatty liver disease |
| NAFLD-FS | NAFLD fibrosis score |
| NASH | nonalcoholic steatohepatitis |
| ROC | receiver operating characteristic |

References

1. Kleiner DE, Brunt EM. Nonalcoholic fatty liver disease: pathologic patterns and biopsy evaluation in clinical research. *Semin Liver Dis.* 2012; 32:3–13. [PubMed: 22418883]
2. Caldwell SH, Oelsner DH, Iezzoni JC, et al. Cryptogenic cirrhosis: clinical characterization and risk factors for underlying disease. *Hepatology.* 1999; 29:664–666. [PubMed: 10051466]
3. Mendes FD, Suzuki A, Sanderson SO, et al. Prevalence and indicators of portal hypertension in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol.* 2012; 10:1028–1033. [PubMed: 22610002]
4. Baffy G, Brunt EM, Caldwell SH. Hepatocellular carcinoma in non-alcoholic fatty liver disease: an emerging menace. *J Hepatol.* 2012; 56:1384–1391. [PubMed: 22326465]
5. Charlton MR, Burns JM, Pedersen RA, et al. Frequency and outcomes of liver transplantation for nonalcoholic steatohepatitis in the United States. *Gastroenterology.* 2011; 141:1249–1253. [PubMed: 21726509]
6. Afzali A, Berry K, Ioannou GN. Excellent posttransplant survival for patients with nonalcoholic steatohepatitis in the United States. *Liver Transpl.* 2012; 18:29–37. [PubMed: 21932374]

7. Agopian VG, Kaldas FM, Hong JC, et al. Liver transplantation for nonalcoholic steatohepatitis: the new epidemic. *Ann Surg.* 2012; 256:624–633. [PubMed: 22964732]
8. Gedaly R, Davenport DL, Daily MF, et al. Rising incidence of nonalcoholic steatohepatitis among patients transplanted for hepatocellular carcinoma in the United States. *Hepatology.* 2012; 56(Suppl):40A.
9. Bellentani S, Marino M. Epidemiology and natural history of nonalcoholic fatty liver disease (NAFLD). *Ann Hepatol.* 2009; 8(Suppl 1):S4–S8. [PubMed: 19381118]
10. Dam-Larsen S, Becker U, Franzmann MB, et al. Final results of a long-term, clinical follow-up in fatty liver patients. *Scand J Gastroenterol.* 2009; 44:1236–1243. [PubMed: 19670076]
11. Ekstedt M, Franzen LE, Mathiensen UI, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology.* 2006; 44:865–873. [PubMed: 17006923]
12. Angulo P. Long-term mortality in NAFLD: Is liver histology of any prognostic significance? *Hepatology.* 2010; 51:373–375. [PubMed: 20101746] Erratum in: *Hepatology* 2010;51:1868.
13. Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology.* 2007; 45:846–854. [PubMed: 17393509]
14. Wai CT, Greenon JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology.* 2003; 38:518–526. [PubMed: 12883497]
15. Shah AG, Lydecker A, Murray K, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol.* 2009; 7:1104–1112. [PubMed: 19523535]
16. Harrison SA, Oliver D, Arnold HL, et al. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. *Gut.* 2008; 57:1441–1447. [PubMed: 18390575]
17. Musso G, Gambino R, Cassader M, et al. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. *Ann Med.* 2011; 43:617–649. [PubMed: 21039302]
18. Demir M, Lang S, Nierhoff D, et al. Stepwise combination of simple noninvasive fibrosis scoring systems increases diagnosis accuracy in nonalcoholic fatty liver disease. *J Clin Gastroenterol.* Feb 24, 2013 [Epub ahead of print].
19. Grandison GA, Angulo P. Can NASH be diagnosed, graded and staged noninvasively? *Clin Liver Dis.* 2012; 16:567–585. [PubMed: 22824481]
20. US Department of Health and Human Services, Public Health Service. Non-Competing Continuation Progress Report (PHS 2590). Available at: <http://grants.nih.gov/grants/funding/2590/phs2590.pdf>.
21. Grundy SM, Cleeman JI, Daniels SR, et al. American Heart Association; National Heart, Lung, and Blood Institute. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation.* 2005; 112:2735–2752. [PubMed: 16157765]
22. Prati D, Taioli E, Zanella A, et al. Updated definitions of healthy ranges for serum alanine aminotransferase levels. *Ann Intern Med.* 2002; 137:1–10. [PubMed: 12093239]
23. Kleiner DE, Brunt EM, Van Natta M, et al. Nonalcoholic Steatohepatitis Clinical Research Network. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* 2005; 416:1313–1321. [PubMed: 15915461]
24. Brunt EM, Kleiner DE, Wilson LA, et al. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. *Hepatology.* 2011; 53:810–820. [PubMed: 21319198]
25. Franchis R. Revising consensus in portal hypertension: Report of the Baveno V consensus workshop on methodology of diagnosis and therapy in portal hypertension. *J Hepatol.* 2010; 53:762–768. [PubMed: 20638742]
26. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology.* 2011; 53:1020–1022. [PubMed: 21374666]
27. Schuppan D, Afdhal NH. Liver cirrhosis. *Lancet.* 2008; 371(9615):838–851. [PubMed: 18328931]

28. Angulo P, Keach JC, Batts KP, et al. Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis. *Hepatology*. 1999; 30:1356–1362. [PubMed: 10573511]
29. Sheth SG, Flamm SL, Gordon FD, et al. AST/ALT ratio predicts cirrhosis in patients with chronic hepatitis C virus infection. *Am J Gastroenterol*. 1998; 93:44–48. [PubMed: 9448172]
30. Lavie CJ, Milani RV, Ventura HO. Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. *J Am Coll Cardiol*. 2009; 53:1925–1932. [PubMed: 19460605]
31. Renehan AG, Tyson M, Egger M, et al. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008; 371(9612):569–578. [PubMed: 18280327]
32. Ryan CK, Johnson LA, Germin BI, et al. One hundred consecutive hepatic biopsies in the workup of living donors for right lobe liver transplantation. *Liver Transpl*. 2002; 8:1114–1122. [PubMed: 12474149]
33. Mottin CC, Moretto M, Padoin AV, et al. The role of ultrasound in the diagnosis of hepatic steatosis in morbidly obese patients. *Obes Surg*. 2004; 14:635–637. [PubMed: 15186630]

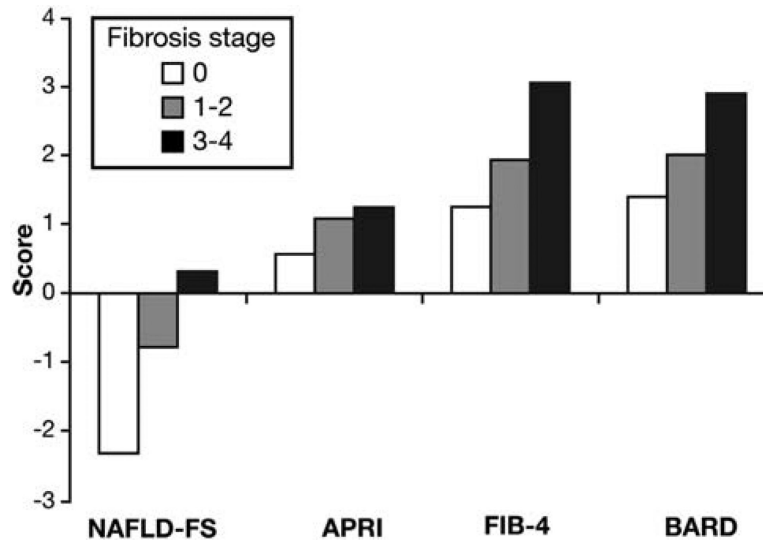


Figure 1. Values of the scores by fibrosis stage. Bars represent the mean values ($P < .001$ for each score).

Table 1

Baseline Characteristics of the Patient Population

| Variable | Total (n = 320) |
|--|------------------|
| Age, y, median (IQR) [100] | 52 (43 to 61) |
| Sex, n (%) [100] | |
| Female | 182 (57) |
| Male | 138 (43) |
| Ethnic group, n (%) [100] | |
| Hispanic | 2 (0.6) |
| Non-Hispanic | 318 (99.4) |
| Race group, n (%) [100] | |
| White | 294 (92) |
| Asian | 16 (5) |
| Black or African American | 6 (2) |
| American Indian/Alaska Native | 4 (1) |
| Native Hawaiian or other Pacific Islander | 0 |
| Body mass index, median (IQR) [100] | 33 (29.4 to 36) |
| BMI category, n (%) [100] | |
| Normal | 15 (4.7) |
| Overweight | 70 (21.9) |
| Obese | |
| BMI 30–34.9 | 135 (42.2) |
| BMI 35–39.9 | 62 (19.4) |
| BMI 40 | 38 (11.9) |
| Waist circumference, cm, median (IQR) [23] | 106 (98 to 115) |
| Diabetes, yes, n (%) [100] | 116 (36.2) |
| Hypertension, yes, n (%) [100] | 152 (47.5) |
| Hypertriglyceridemia, yes, n (%) [98] | 190 (59.4) |
| Hypercholesterolemia, yes, n (%) [97] | 85 (26.6) |
| Low-HDL cholesterol, yes, n (%) [91] | 177 (55.3) |
| Statins use, yes [100] | 32 (10) |
| ALT, IU/L, median (IQR) [100] | 61 (38 to 85) |
| AST, IU/L, median (IQR) [100] | 50 (37 to 78) |
| AST/ALT ratio, median (IQR) [100] | 1.0 (0.6 to 1.1) |
| Total bilirubin, $\mu\text{mol/L}$, median (IQR) [98] | 12 (8.5 to 20.5) |
| Albumin, g/L, median (IQR) [100] | 43 (40 to 45) |
| Alkaline phosphatase, IU/L, median (IQR) [63] | 160 (127 to 227) |
| γ -Glutamyltransferase, IU/L, median (IQR) [36] | 123 (93 to 168) |
| INR, median (IQR) [45] | 1.0 (1.0 to 1.0) |
| Platelet, $\times 10^9/L$, median (IQR) [100] | 218 (165 to 265) |
| Glucose, mmol/L, median (IQR) [100] | 5.9 (5.3 to 7.6) |
| Triglycerides, mmol/L, median (IQR) [98] | 2.0 (1.4 to 2.8) |

| Variable | Total (n = 320) |
|--|------------------------|
| Total cholesterol, <i>mmol/L</i> , median (IQR) [97] | 5.3 (4.6 to 6.1) |
| HDL-cholesterol, <i>mmol/L</i> , median (IQR) [91] | 1.1 (0.9 to 1.3) |
| LDL-cholesterol, <i>mmol/L</i> , median (IQR) [43] | 2.6 (2.0 to 3.2) |
| Ferritin, <i>pmol/L</i> , median (IQR) [75] | 292 (133 to 593) |
| NAFLD-FS, median (IQR) [100] | -0.904 (-2.07 to 0.46) |
| APRI, median (IQR) [100] | 0.74 (0.47 to 1.26) |
| FIB-4, median (IQR) [100] | 1.70 (1.04 to 2.89) |
| BARD, median (IQR) [100] | 2 (1 to 3) |
| Steatosis, grade, n (%) [100] | |
| 1 | 134 (42) |
| 2 | 137 (43) |
| 3 | 49 (15) |
| NASH, definitive, n (%) [100] | 158 (49) |
| Fibrosis stage, n (%) [100] | |
| 0 | 92 (29) |
| 1/2 | 65 (20) |
| 3/4 | 163 (51) |

NOTE. Numbers in brackets are percentage of subjects with the variable measured.

HDL, high-density lipoprotein; INR, international normalized ratio; IQR, interquartile range; LDL, low-density lipoprotein.

Table 2

Liver-Related Events and Causes of Death

| Outcome | n | Source |
|-----------------------------------|----|--|
| Liver-related events | | |
| Ascites | 17 | Australia 5; Glasgow 1; Iceland 1; Newcastle 2; United States 8 |
| Gastroesophageal varices/bleeding | 20 | Australia 4; Glasgow 1; Iceland 2; Newcastle 10; Thailand 1; United States 2 |
| Portosystemic encephalopathy | 15 | Australia 2; Glasgow 1; Newcastle 4; United States 8 |
| Spontaneous bacterial peritonitis | 2 | Australia 1; United States 1 |
| Hepatocellular cancer | 4 | Newcastle 2; United States 2 |
| Hepatopulmonary syndrome | 1 | United States 1 |
| Hepatorenal syndrome | 1 | United States 1 |
| Death or OLT | | |
| Cardiovascular disease | 13 | Australia 2; Glasgow 1; Iceland 2; Thailand 1; United States 7 |
| Cirrhosis complications | 10 | Australia 5; Iceland 1; United States 4 |
| Liver transplantation | 5 | Newcastle 2; United States 3 |
| Non-liver cancer deaths | 11 | Iceland 3; Glasgow 1; United States 7 |
| Other | 2 | Glasgow 1; United States 1 |

NOTE. Some patients developed more than one liver-related event.

OLT, Orthotopic liver transplantation.

Table 3

Area Under the ROC Curve of Four Scoring Systems to Predict Outcomes

| Score | Area under the ROC curve \pm SE (95% CI) [<i>P</i> value] | |
|----------------------|--|---|
| | Liver-related events | Death/liver transplantation |
| NAFLD fibrosis score | 0.86 \pm 0.03 (0.80–0.92) [$<.001$] | 0.70 \pm 0.04 (0.62–0.78) [$<.001$] |
| APRI | 0.80 \pm 0.03 (0.73–0.86) [$<.001$] | 0.63 \pm 0.05 (0.53–0.72) [.006] |
| FIB-4 | 0.81 \pm 0.03 (0.76–0.87) [$<.001$] | 0.67 \pm 0.05 (0.58–0.76) [.01] |
| BARD | 0.73 \pm 0.04 (0.66–0.80) [$<.001$] | 0.66 \pm 0.04 (0.58–0.74) [.001] |

NOTE. *P* value in square bracket is the comparison of the area under the ROC curve of each model with the reference line (null hypothesis) of .5.

Table 4

Cumulative Probability (Unadjusted) of Outcomes by Score Risk Category

| Outcome | NAFLD-FS | | | APRI | | | FIB-4 | | | BARD | | |
|------------------------|------------|--|---------|------------|--|---------|------------|--|---------|------------|--|---------|
| | Events (n) | Unadjusted cumulative (%) ^a | P value | Events (n) | Unadjusted cumulative (%) ^a | P value | Events (n) | Unadjusted cumulative (%) ^a | P value | Events (n) | Unadjusted cumulative (%) ^a | P value |
| Liver events (n = 309) | | | <.001 | | | <.001 | | | <.001 | | | <.001 |
| Low risk | 2/125 | 1.6 | | 1/90 | 1.1 | | 3/111 | 2.7 | | 2/119 | 1.7 | |
| Intermediate risk | 12/115 | 10.4 | | 21/162 | 13 | | 4/112 | 3.6 | | 23/119 | 19.3 | |
| High risk | 29/69 | 42 | | 21/57 | 36.8 | | 36/86 | 41.9 | | 18/71 | 25.4 | |
| Mortality (n = 320) | | | <.001 | | | <.001 | | | <.001 | | | .01 |
| Low risk | 5/125 | 4 | | 9/93 | 9.7 | | 6/111 | 5.4 | | 5/119 | 4.2 | |
| Intermediate risk | 20/120 | 16.7 | | 17/169 | 10 | | 14/117 | 12 | | 21/124 | 16.9 | |
| High risk | 16/75 | 21.3 | | 15/58 | 25.9 | | 21/92 | 22.8 | | 15/77 | 19.5 | |

NOTE. Risk category: low risk, NAFLD-FS <-1.455; APRI <0.5; FIB-4 <1.30; BARD 0/1; moderate risk, NAFLD-FS -1.455 to 0.676; APRI 0.5 to 1.5; FIB-4 1.30 to 2.67; BARD 2/3; high risk, NAFLD-FS >0.676; APRI >1.5; FIB-4 >2.67; BARD 4.

^aEstimated using Kaplan-Meier analysis.

Table 5

Multivariate Adjusted Hazard Ratios and 95% CIs of Outcome by Risk Score Category

| Outcome | NAFLD-FS | | APRI | | FIB-4 | | BARD | |
|------------------------|----------------------------|---------|----------------------------|---------|----------------------------|---------|----------------------------|---------|
| | HR (95% CI) ^{a,b} | P value | HR (95% CI) ^{a,c} | P value | HR (95% CI) ^{a,d} | P value | HR (95% CI) ^{a,e} | P value |
| Liver events (n = 309) | | | | | | | | |
| Low risk | 1.0 | | 1.0 | | 1.0 | | 1.0 | |
| Intermediate risk | 7.7 (1.4–42.7) | <.02 | 8.8 (1.1–67.3) | <.04 | 0.92 (0.2–4.8) | .9 | 6.2 (1.4–27.2) | <.02 |
| High risk | 34.2 (6.5–180.9) | <.001 | 20.9 (2.6–165.3) | .004 | 14.6 (4.1–52.6) | <.001 | 6.6 (1.4–31.1) | <.02 |
| Mortality (n = 320) | | | | | | | | |
| Low risk | 1.0 | | 1.0 | | 1.0 | | 1.0 | |
| Intermediate risk | 4.2 (1.3–13.8) | <.02 | 1.1 (0.4–2.7) | .9 | 2.3 (0.8–6.6) | .1 | 1.8 (0.6–5.1) | .3 |
| High risk | 9.8 (2.7–35.3) | <.001 | 3.1 (1.1–8.4) | .03 | 6.9 (2.3–20.4) | .001 | 1.6 (0.5–4.9) | .4 |

^aEstimated using Cox proportional hazard regression analysis.^bAdjusted by sex, hypertension, race, use of statins, and fibrosis stage.^cAdjusted by age, sex, diabetes, body mass index, hypertension, race, use of statins, and fibrosis stage.^dAdjusted by sex, diabetes, body mass index, hypertension, race, use of statins, and fibrosis stage.^eAdjusted by age, sex, hypertension, race, use of statins, and fibrosis stage.