Performance of Plasma Biomarkers and Diagnostic Panels for Nonalcoholic Steatohepatitis and Advanced Fibrosis in Patients With Type 2 Diabetes

Fernando Bril,[1,2] Michael J. McPhaul,[3] Michael P. Caulfield,[3] Virginia C. Clark,[4] Consuelo Soldevilla-Pico,[4] Roberto J. Firpi-Morell,[4] Jinping Lai,[5] Dov Shiffman,[3] Charles M. Rowland,[3] and Kenneth Cusi[1,6]

[1]Division of Endocrinology, Diabetes and Metabolism, University of Florida, Gainesville, FL
[2]Internal Medicine, Department of Medicine, University of Florida, Gainesville, FL
[3]Quest Diagnostics Nichols Institute, San Juan Capistrano, CA
[4]Division of Gastroenterology, Hepatology and Nutrition, University of Florida, Gainesville, FL
[5]Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL
[6]Division of Endocrinology, Diabetes and Metabolism, Malcom Randall VA Medical Center, Gainesville, FL

Corresponding author: Kenneth Cusi, kenneth.cusi@medicine.ufl.edu

The contents of this article do not reflect the views of the Department of Veterans Affairs or the U.S. government.

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc19-1071/-/DC1.

This article is part of a special article collection available at https://care.diabetesjournals.org/collection/nafld-in-diabetes.

OBJECTIVE

The 2019 Standards of Medical Care in Diabetes suggested that patients with nonalcoholic fatty liver disease (NAFLD) should be evaluated for liver fibrosis. However, the performance of noninvasive clinical models/scores and plasma biomarkers for the diagnosis of nonalcoholic steatohepatitis (NASH) and advanced fibrosis has not been carefully assessed in patients with type 2 diabetes mellitus (T2DM).

RESEARCH DESIGN AND METHODS

In this cross-sectional study, patients (n = 213) had a liver MRS, and those with a diagnosis of NAFLD underwent a percutaneous liver biopsy. Several noninvasive clinical models/scores and plasma biomarkers were measured to identify NASH and advanced fibrosis (NASH: ALT, cytokeratin-18, NashTest 2, HAIR, BARD, and OWLiver; advanced fibrosis: AST, fragments of propeptide of type III procollagen [PRO-C3], FIB-4, APRI, NAFLD fibrosis score, and FibroTest).

RESULTS

None of the noninvasive tools assessed for the diagnosis of NASH in patients with T2DM had an optimum performance (all areas under the curve [AUCs] <0.80). Of note, none of the panels or biomarkers was able to outperform plasma ALT (AUC 0.78 [95% CI 0.71–0.84]). Performance was better to diagnose advanced fibrosis, in which plasma PRO-C3, AST, and APRI showed better results than the other approaches (AUC 0.90 [0.85–0.95], 0.85 [0.80–0.91], and 0.86 [0.80–0.91], respectively). Again, none of the approaches did significantly better than plasma AST. Sequential use of plasma AST and other noninvasive tests may help in limiting the number of liver biopsies required to identify patients with advanced fibrosis.

CONCLUSIONS

Performance of noninvasive clinical models/scores and plasma biomarkers for the diagnosis of NASH or advanced fibrosis was suboptimal in patients with T2DM. Combination of multiple tests may provide an alternative to minimize the need for liver biopsies to detect fibrosis in these patients.

About 70% of patients with type 2 diabetes mellitus (T2DM) have nonalcoholic fatty liver disease (NAFLD) (1,2), and ~20–30% have the more severe form of the disease with lobular inflammation and hepatocyte ballooning (nonalcoholic steatohepatitis [NASH]). Although long-term prospective studies are lacking, these patients are also believed to be at higher risk of disease progression to advanced fibrosis and cirrhosis (3–5). Recent *Standards of Medical Care in Diabetes* (6) suggested that patients with NAFLD based on liver ultrasound or elevated plasma aminotransferases should be evaluated for liver fibrosis. Unfortunately, it is still unclear what the best strategy to accurately diagnose liver fibrosis in these patients is. Many patients are only diagnosed once advanced fibrosis has already developed (7,8). Therefore, in order to modify the natural history of the disease, an early diagnosis is essential in high-risk populations, such as those with T2DM.

Many primary care physicians only suspect NASH in the presence of elevated plasma aminotransferases, despite ample evidence that a significant number of patients with NASH have "normal" plasma aminotransferase levels (i.e., <40 IU/L) (9,10). Patients with NAFLD may even have a negative liver ultrasound, as this technique cannot detect steatosis unless it is rather significant (11). Further, other imaging techniques (i.e., FibroScan or magnetic resonance elastography) may not be readily available outside hepatology clinics, and patients and clinicians shy away from percutaneous liver biopsies (the gold standard for the diagnosis of NASH) in the absence of U.S. Food and Drug Administration–approved agents.

However, due to increasing awareness about the health risks associated with NASH in T2DM and recent findings showing that weight loss and pharmacological treatments may induce resolution of NASH (12,13), there is a renewed interest in noninvasive diagnostic tools to identify and monitor these patients. Advanced fibrosis (≥F3) is the most relevant target for early diagnosis and treatment because it has been associated in a number of studies with future development of cirrhosis and increased overall mortality (12,13). Several clinical models/scores and plasma biomarkers have been assessed for the diagnosis of NASH or

advanced fibrosis, with mixed results (14–18). Several of them are based on the presence of diabetes, hyperglycemia, or hyperinsulinemia to identify patients with more severe liver disease (14–18). As a result, if these models/scores were to be applied to only patients with diabetes (as seen, for example, in an endocrinology clinic), they could potentially overestimate the prevalence of NASH or advanced fibrosis due to spectrum bias. In line with this, recent studies have reported that noninvasive models appear to underperform in patients with T2DM (19–21).

In the current study, we aimed to assess the performance of several noninvasive clinical models/scores and plasma biomarkers for the diagnosis of definite NASH and advanced fibrosis (stage F3 or above) in patients with T2DM.

RESEARCH DESIGN AND METHODS
Patients
Subjects were recruited from the general population and from hepatology and endocrinology clinics at the University of Florida in Gainesville, FL, and at the University of Texas Health Science Center at San Antonio in San Antonio, TX. Only patients with a diagnosis of T2DM were included in this study. The only glucose-lowering drugs allowed were metformin, sulfonylureas, or insulin, as long as patients were on a stable dose for at least 3 months prior to enrollment. Other main exclusion criteria included: any liver disease other than NASH (i.e., hepatitis B or C, autoimmune hepatitis, hemochromatosis, Wilson disease, or drug-induced hepatitis), significant alcohol consumption (≥30 g/day for males and ≥20 g/day for females), type 1 diabetes, or use of prohibited medications (i.e., vitamin E, pioglitazone, weight loss medications, amiodarone, glucocorticoids, methotrexate, olanzapine, and protease inhibitors). Patients were also excluded if they had missing information regarding results from liver proton MRS (^{1}H-MRS) or the liver biopsy.

The study was approved by the institutional review boards at the University of Florida and University of Texas Health Science Center at San Antonio, and written informed consent was obtained from each patient prior to participation. Patients included in this analysis

have been previously reported on in studies that focused on assessing the performance of plasma fragments of propeptide of type III procollagen (PRO-C3) (22), the FibroMax test (20), and OWLiver test (19) in patients with T2DM. In the current work, we have compared head-to-head the performance of the plasma PRO-C3 testing, FibroMax test, and OWLiver test to a variety of published algorithms to assess the presence of NASH and to detect and characterize fibrosis in a cohort of patients with T2DM. In addition, we explored the potential utility of combining the results from more than one test as an avenue to improve the performance of individual tests.

Study Design
This is a cross-sectional study in which all patients underwent a two-step diagnostic approach for NASH, including a liver ^{1}H-MRS and a percutaneous liver biopsy if they were found to have NAFLD by imaging. For the prediction of definite NASH, we calculated/measured the following noninvasive approaches: *1*) plasma alanine aminotransferase (ALT); *2*) plasma cytokeratin 18 (CK-18); *3*) BARD score (defined as the sum of BMI ≥28 = 1 point, AST/ALT ratio ≥0.80 = 2 points, and diabetes = 1 point); *4*) NashTest 2 (a proprietary score based on serum α_2-macroglobulin, apolipoprotein A1, haptoglobin, total bilirubin, γ-glutamyl transpeptidase [GGT], AST, cholesterol, and triglycerides); *5*) OWLiver (a proprietary, BMI-dependent logistic regression algorithm based on serum levels of a panel of 20 triglycerides); *6*) HAIR score (defined as the sum of hypertension = 1 point, plasma ALT >40 units/L = 1 point, and insulin resistance index [log fasting insulin + log fasting glucose] >5.0 = 1 point); and *7*) a model specifically developed in our cohort from demographic, clinical, and routine biochemical data.

For the prediction of advanced fibrosis, we calculated/measured the following noninvasive approaches: *1*) plasma AST; *2*) APRI (defined as [AST in units/L]/[40 units/L as the upper limit of normal]/[platelets in 10^9/L]); *3*) FibroTest (a proprietary score based on serum α_2-macroglobulin, apolipoprotein A1, haptoglobin, total bilirubin, and GGT); *4*) FIB-4 (defined as [age × AST]/[platelets ×

$\sqrt{ALT}$]); *5*) NAFLD fibrosis score (defined as −1.675 + 0.037 × [age in years] + 0.094 × [BMI in kg/m^2] + 1.13 [all patients had diabetes] + 0.99 × [AST/ALT ratio] − 0.013 × [platelets in 10^9/L] − 0.66 × [albumin in g/dL]); *6*) plasma PRO-C3; and *7*) a model specifically developed in our cohort from demographic, clinical, and routine biochemical data.

For determination of the NashTest 2 and FibroTest (BioPredictive algorithms), samples were blindly provided to Quest Diagnostics (San Juan Capistrano, CA) to measure haptoglobin, α_2-macroglobulin, apolipoprotein A1, bilirubin, GGT, AST, triglycerides, and total cholesterol. Samples were also blindly provided to One Way Liver, S.L. (Derio, Spain), for measurement of their specific score (i.e., OWLiver) and to Nordic Bioscience (Herlev, Denmark) for measurement of plasma PRO-C3. Routine as well as metabolic laboratory tests (insulin, C-peptide, glucose, and free fatty acids) were run in our laboratory as previously reported (23,24). Sulfonylureas were held the morning of the fasting tests, while basal insulin was held the evening before and the morning of the fasting tests.

Measurements of Intrahepatic Triglyceride Content

Intrahepatic triglyceride content was measured by liver ^{1}H-MRS in a 3-Tesla MRI scanner. Three areas of 30 × 30 × 30 mm were selected in the liver, avoiding any large vessel. A single experienced observer analyzed the spectra using commercial software (NUTS; Acorn NMR Inc., Livermore, CA). Intrahepatic triglyceride content was calculated as fat fraction (area under the curve [AUC] fat peak/[AUC fat peak + water peak]). Measurements were corrected for T1 and T2 relaxation using methods previously described (11). A liver fat content of >5.56% was considered diagnostic of NAFLD (12,13).

Percutaneous Liver Biopsy

Liver biopsies were performed under ultrasound guidance. Histological characteristics for the diagnosis of definite NASH were assessed using standard criteria (25). Briefly, a diagnosis of definite NASH was made if the biopsies showed presence of zone 3 accentuation of macrovesicular steatosis (any grade), hepatocellular ballooning (of any degree), and

lobular inflammatory infiltrates (of any amount). Fibrosis stages were defined as previously established (26): stage 0 if no fibrosis was present, stage 1 for perisinusoidal or periportal fibrosis, stage 2 for presence of both perisinusoidal and portal/periportal, stage 3 for bridging fibrosis, and stage 4 for cirrhosis. Advanced fibrosis was considered as the presence of stages 3 or 4. All liver biopsies were read by an expert pathologist (J.L.), who was unaware of patients' characteristics. Mean length of liver biopsies was 17 mm, and only 7% of the specimens were <10 mm. Mean number of portal tracts was 9.

Laboratory Assays

Assays used in the calculation of the BioPredictive algorithms were run on routine automated platforms using standard reagents. Total cholesterol, triglycerides, AST, GGT, and total bilirubin were run on Beckman Coulter AU series instruments (Brea, CA). Haptoglobin, α_2-macroglobulin, and apolipoprotein A1 were run on a Siemens BNII instrument (Siemens Healthcare Diagnostics, Tarrytown, NY). The serum lipidomic profiles for the OWLiver test were obtained using methods previously described (19). Plasma PRO-C3 was analyzed using competitive ELISAs as previously described (27). In all cases, samples were blindly analyzed without access to any of the associated clinical data.

Statistical Analysis

Data were summarized as number (percentages) for categorical variables and as means ± SD for numeric variables. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for the biomarker panels were assessed considering liver histology as the gold standard reference. Patients without NAFLD by ^{1}H-MRS, and in whom other causes of liver disease had been excluded, were used as control subjects (i.e., as not having definite NASH and not having advanced fibrosis) to assess the specificities of tests. A biopsy was not performed as it was considered unethical and the chances of disease minimal (12,13). Cohort-specific models based on demographic and routine clinical and biochemical characteristics (i.e., age, sex, BMI, HOMA of insulin resistance [HOMA-IR], fasting plasma glucose, hemoglobin A$_{1c}$, fasting plasma insulin,

triglycerides, HDL cholesterol, blood pressure, platelets, albumin, ALT, AST, and CK-18) were also developed by multivariate logistic regression analysis with forward selection. A significance level of <0.20 in the univariate analysis was defined to allow a variable into the model, and only significant variables were kept in the final models. Receiver operating characteristic (ROC) curves were plotted and the AUC calculated to represent their performance to predict binary outcomes (NASH and advanced fibrosis). For the cohort-specific models, AUCs were validated using a bootstrap procedure to adjust for overfitting. Two hundred bootstrap samples were randomly generated while setting aside "out-of-bag" samples (~32% of subjects were not selected in each bootstrap sample). For these, the median of the 200 out-of-bag AUC estimates is reported with 95% CIs. Comparisons between AUCs were performed with the *roccomp* command (test of equality of ROC areas) in Stata. Predictive mean matching imputation with five imputed data sets was used as a sensitivity analysis to account for missing data. Missing values were considered to be missing at random, and imputed AUCs were combined according to Rubin's rules. A two-tailed value of $P < 0.05$ was considered to indicate statistical significance. Analyses were performed with Stata 11.1 (StataCorp LP, College Station, TX) and graphs with Prism 6.0 (GraphPad Software, Inc., La Jolla, CA).

RESULTS

Patients' Characteristics

A total of 213 patients were included in this study. Based on results from the liver ^{1}H-MRS and liver biopsy, patients were divided into three groups: No NAFLD, No definite NASH, and Definite NASH. As can be observed in Table 1, there were no significant differences among the three groups regarding age, sex, diabetes control (hemoglobin A$_{1c}$ or fasting plasma glucose), or use of glucose-lowering drugs. The No NAFLD group included a higher proportion of African American patients and a lower proportion of Hispanic patients compared with the other two groups. Patients with NAFLD (with or without NASH) had higher BMI than patients without NAFLD (34.5 ± 4.7 vs. 31.2 ± 4.5 kg/m^2; $P < 0.001$), but

Table 1—Demographic and clinical characteristics of patients

| | No NAFLD (*n* = 51) | NAFLD | | *P* value |
		No NASH (*n* = 72)	Definite NASH (*n* = 90)	
Age, years	60 ± 8	57 ± 9	57 ± 8	0.08
Sex, male %	86	82	83	0.81
Ethnicity, *n* (%)				
Caucasian	32 (63)	38 (53)	58 (65)	
Hispanic	8 (16)	27 (37)	27 (30)	0.011
African American	11 (21)	7 (10)	4 (4)	
Other	0 (0)	0 (0)	1 (1)	
BMI, kg/m^2	31.2 ± 4.5	34.2 ± 4.7	34.7 ± 4.7	<0.001
Total body fat, %	34 ± 6	36 ± 7	36 ± 7	0.17
Hemoglobin A$_{1c}$, %	7.1 ± 1.3	6.9 ± 1.1	7.2 ± 1.2	0.31
Fasting plasma glucose, mg/mL	150 ± 50	137 ± 36	153 ± 42	0.06
Fasting plasma insulin, μU/mL	8 ± 6	14 ± 10	19 ± 13	<0.001
Diabetes medications, %				
Metformin	75	67	79	0.25
Sulfonylurea	44	44	38	0.73
Insulin	29	31	18	0.12
Intrahepatic triglyceride content, %	3 ± 1	12 ± 7	15 ± 8	<0.001
AST, units/L	22 ± 7	32 ± 18	48 ± 29	<0.001
ALT, units/L	24 ± 11	43 ± 33	64 ± 41	<0.001
NAFLD activity score	—	2.3 ± 0.9	4.8 ± 1.3	<0.001
Steatosis grade	—	1.2 ± 0.8	1.9 ± 0.7	<0.001
Inflammation grade	—	1.0 ± 0.5	1.6 ± 0.6	<0.001
Ballooning grade	—	0.1 ± 0.3	1.3 ± 0.5	<0.001
Fibrosis stage	—	0.6 ± 0.9	1.8 ± 1.0	<0.001

Data are mean ± SD unless otherwise indicated. *P* values represent comparison among the three groups with ANOVA.

we observed no differences in BMI between patients in the No NASH and Definite NASH groups. Fasting plasma insulin showed a stepwise increase among the three groups (No NAFLD 8 ± 6 vs. No definite NASH 14 ± 10 vs. Definite NASH 19 ± 13 μU/mL; *P* < 0.001). Patients with NASH had higher plasma ALT and AST, as well as worse liver histology. Among patients with a liver biopsy, 105 (65%) had fibrosis stages 0 or 1 (F0 to F1), 26 (16%) had fibrosis stage 2 (F2), and 31 (19%) had advanced fibrosis (F3 to F4).

Performance of Noninvasive Tests for the Diagnosis of Definite NASH
In Fig. 1*A*, ROC curves assessing the performance of different clinical scores and biomarkers for the diagnosis of definite NASH were plotted. As can be observed, none of the noninvasive strategies was successful to predict the presence of definite NASH in our cohort of patients with T2DM (all AUCs <0.80). Comparisons among each pair of ROC curves were performed, and *P* values can be found in Supplementary Table 1. Plasma ALT, CK-18, and our own cohort-specific model performed better than the other strategies, with no significant differences among these three approaches. Of note, the final cohort-specific model (0.0941 × [HOMA-IR] + 0.0039 × [CK-18] − 1.8647) had an AUC after bootstrapping (*n* = 200) of 0.74 (95% CI 0.64–0.83). In Table 2, we have summarized sensitivity, specificity, PPV, and NPV for all of these noninvasive tools based on their predefined as well as cohort-specific cutoff points. We observed high sensitivity for the diagnosis of NASH when the predefined cutoff points were used for BARD, HAIR, and NashTest 2, but these came at the expense of very low specificity. Plasma ALT, CK-18, and our cohort-specific model had modest sensitivity and specificity (all in the ~60–80% range). Supplementary Table 2 allows for a head-to-head comparison between the specificities of the clinical scores/biomarkers after fixing their sensitivity at 95%. As can be observed, at this

sensitivity, plasma ALT had the highest specificity of all of the predictive tools (38% [95% CI 16–51]).

Performance of Noninvasive Tests for the Diagnosis of Advanced Fibrosis
As can be observed in Fig. 1*B*, the performance of noninvasive clinical scores or biomarkers was overall better for the diagnosis of advanced fibrosis than for definite NASH. Plasma AST, APRI, our cohort-specific model, and PRO-C3 performed significantly better than the FibroTest and NAFLD fibrosis score (Supplementary Table 3). No significant differences were observed among these four better approaches when they were compared with each other. FIB-4 had an intermediate performance, being only significantly worse than PRO-C3 (*P* = 0.008), but not different from AST, APRI, or our cohort-specific model. Of note, our final cohort-specific model was 0.0034 × (CK-18) + 0.0588 × (fasting insulin) − 0.0116 × (platelets) − 1.3336 × (sex) + 0.4469 × (HbA$_{1c}$) − 3.82 (where for sex, 1 = male and 0 = female) and had an AUC after bootstrapping (*n* = 200) of 0.81 (0.65–0.92). Table 3 summarizes sensitivity, specificity, PPV, and NPV for all of these noninvasive tools. As can be observed, all noninvasive tests had high NPVs (>90%). Interestingly, when predefined cutoff points were used for these scores (in several of them, creating an indeterminate zone of patients not classified), FIB-4 showed the best combination of PPV and NPV (80% and 94%, respectively). As observed in Supplementary Table 2, plasma PRO-C3 had the highest specificity after sensitivity was fixed at 95% (71% [95% CI 53–85]), closely followed by plasma AST (58% [48–79]) and APRI (57% [38–77]).

Performance of Noninvasive Tests in Specific Subgroups of Patients and Sensitivity Analysis
As sensitivity analyses, we assessed the performance of all of these noninvasive tools in different subgroups of patients. Among females, the NAFLD fibrosis score (AUCs: 0.83 [95% CI 0.68–0.99] vs. 0.59 [0.46–0.72]; *P* = 0.016) and our own fibrosis prediction model (AUCs: 0.96 [0.89–1.00] vs. 0.84 [0.74–0.93]; *P* = 0.049) performed significantly better than in male individuals. In addition, plasma ALT showed a trend toward better performance in males for the diagnosis of NASH (AUCs: 0.80 [0.73–0.86]

Non-invasive Diagnosis of Definite NASH

Non-invasive Diagnosis of Advanced Fibrosis
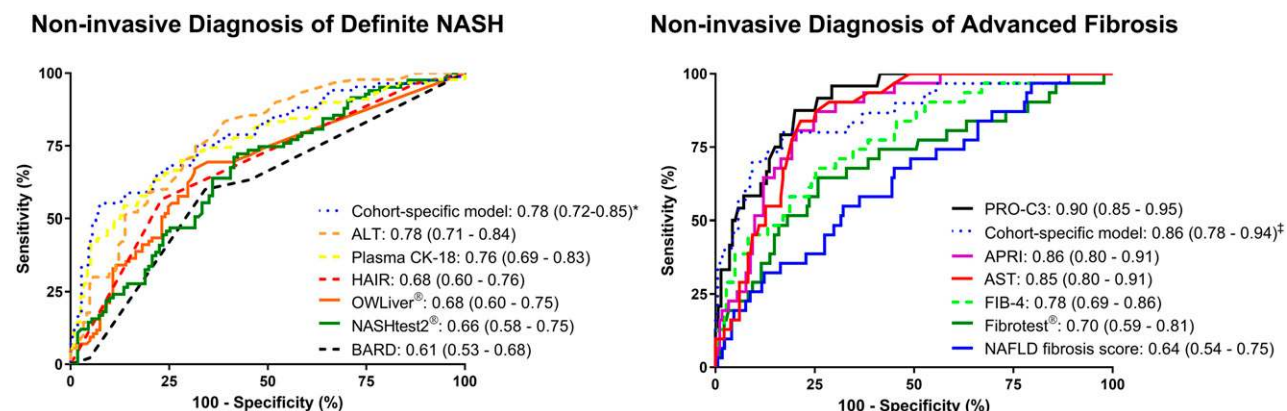
Figure 1—Performance of the different noninvasive clinical models or plasma biomarkers for the diagnosis of definite NASH (*A*) or advanced fibrosis (*B*). Data in parentheses are 95% CI. *AUC was 0.74 (0.64–0.83) after validation using bootstrap (*n* = 200); ‡AUC was 0.81 (0.65–0.92) after validation using bootstrap (*n* = 200).

vs. 0.64 [0.45–0.84]; *P* = 0.13). No significant differences were observed in the performance of these methods among different age-groups or ethnic groups, presence or absence of obesity, or diabetes control (data not shown). Of note, as only one African American patient had advanced fibrosis, the performance of these tests for the diagnosis of advanced fibrosis was not tested in this ethnic group. Of note, as a sensitivity analysis, analyses were repeated after imputing missing data for PRO-C3 (*n* = 49), HAIR (*n* = 26), NashTest 2 (*n* = 19), NAFLD fibrosis score (*n* = 11), OWLiver (*n* = 7), and CK-18 (*n* = 5). No differences were observed in the overall results after predictive mean matching was applied, except for a small reduction in the performance of PRO-C3: 0.83 (95% CI 0.76–0.90).

Use of Multiple Tests for the Diagnosis of Definite NASH and Advanced Fibrosis

Because none of the noninvasive models or biomarkers individually predicted definite NASH or advanced fibrosis with high accuracy, we assessed whether combining these models and biomarkers could improve the identification of patients with NASH or advanced fibrosis. We combined models and biomarkers either sequentially (starting with one test and progressing based on the results) or using parallel testing (using all tests at the same time).

Sequential Testing

We assessed which tool allowed for exclusion of more patients (not definite NASH or not advanced fibrosis) with an NPV of 100%. None of the tools was able to exclude a significant number of patients with an NPV of 100% for definite NASH. For advanced fibrosis, plasma AST <26 units/L excluded 93 (44%) patients correctly (Fig. 2 and Supplementary Fig. 1). In the remaining population, PRO-C3 (<10 ng/mL) excluded an additional 19% of the cohort with an NPV of 100% (Fig. 2A). This resulted in only 37% of the initial

cohort requiring a liver biopsy, of whom 41% would have advanced fibrosis. No patients with advanced fibrosis were missed in diagnosis with this approach. The application of this approach using FIB-4 (<0.87) after AST was able to reduce the number of biopsies to 48% of the entire cohort (Fig. 2B). Results combining the other testing algorithms with plasma AST can be found in Supplementary Fig. 1. As can be observed, APRI and NAFLD fibrosis score showed results similar to those of FIB-4, but FibroTest was not able to exclude any additional patients after plasma AST.

Parallel Testing

The combination of all models/biomarkers (*n* = 6 different tools) for the diagnosis of definite NASH resulted in an AUC of 0.82 (95% CI 0.75–0.88). If at least 4 out of the 6 tests were positive for definite NASH (using the cohort-specific cutoff points described in Table 2), the PPV was 71% (60–81). If only 0–3 tests were positive, then the NPV was 80% (70–88). Sensitivity and specificity were 76%

Table 2—Performance of noninvasive tests for the diagnosis of definite NASH

	Predefined cutoff points					Cohort-specific cutoff points				
	Cutoff	Sensitivity	Specificity	PPV	NPV	Cutoff	Sensitivity	Specificity	PPV	NPV
Cohort-specific model*	<−0.941 >−0.123	74 (61–84)	74 (63–83)	70 (58–81)	77 (66–86)	0.255	55 (44–66)	93 (86–97)	86 (73–94)	73 (64–80)
ALT	40 units/L	66 (55–75)	72 (64–80)	63 (53–73)	74 (65–82)	31 units/L	83 (74–90)	61 (52–70)	61 (52–70)	83 (74–90)
CK-18	N/A	—	—	—	—	241 units/L	63 (52–73)	80 (71–86)	70 (59–80)	74 (66–81)
HAIR	2	96 (90–99)	12 (7–20)	47 (39–55)	81 (54–96)	3	57 (45–68)	77 (68–85)	66 (54–77)	69 (60–77)
OWLiver	0.500	34 (24–45)	87 (79–92)	64 (49–78)	65 (57–72)	0.016	66 (55–76)	69 (60–77)	60 (49–70)	74 (65–82)
NashTest 2	0.250	92 (83–96)	29 (21–38)	49 (41–57)	82 (67–92)	0.410	71 (60–80)	58 (48–67)	56 (46–65)	73 (62–82)
BARD	2	98 (92–100)	5 (2–10)	43 (36–50)	75 (35–97)	2	98 (92–100)	5 (2–10)	43 (36–50)	75 (35–97)

Data are % (95% CI). N/A, not applicable. *Model was = 0.0941 × (HOMA-IR) + 0.0039 × (CK-18) − 1.8647.

Table 3—Performance of noninvasive tests for the diagnosis of advanced fibrosis

	Predefined cutoff points					Cohort-specific cutoff points				
	Cutoff	Sensitivity	Specificity	PPV	NPV	Cutoff	Sensitivity	Specificity	PPV	NPV
PRO-C3	20.0 ng/mL	50 (29–71)	96 (91–98)	67 (41–87)	92 (86–96)	13.2 ng/mL	88 (68–97)	80 (72–86)	43 (29–58)	97 (93–100)
Cohort-specific model‡	<−2.613, >−1.015$	88 (68–97)	86 (78–92)	57 (40–73)	97 (91–99)	−1.369	80 (61–92)	83 (77–88)	45 (32–60)	96 (92–98)
APRI	<0.500, >1.500^	31 (9–61)	99 (95–100)	67 (22–96)	94 (90–97)	0.423	84 (66–94)	75 (68–81)	36 (25–48)	96 (92–99)
AST	40 units/L	77 (59–90)	81 (74–86)	41 (28–54)	96 (91–98)	38 units/L	84 (66–94)	79 (72–84)	40 (28–53)	97 (92–99)
FIB-4	<1.450, >3.250#	33 (10–65)	99 (95–100)	80 (28–100)	94 (88–97)	1.666	68 (49–83)	75 (69–81)	31 (21–44)	93 (88–97)
FibroTest	<0.300, >0.700¥	17 (2–48)	98 (93–100)	40 (5–85)	92 (86–96)	0.353	64 (45–81)	74 (67–80)	30 (19–42)	92 (87–96)
NAFLD fibrosis score	<−1.455, >0.676*	91 (59–100)	40 (26–56)	26 (13–43)	95 (75–100)	−0.053	68 (49–83)	55 (47–63)	21 (14–31)	90 (83–95)

Data are % (95% CI) unless otherwise specified. *144 patients not classified. $68 not classified. ^48 not classified. #84 not classified. ¥83 not classified. ‡Model was = $0.0034 \times (CK\text{-}18) + 0.0588 \times (\text{fasting insulin}) - 0.0116 \times (\text{platelets}) - 1.3336 \times (\text{sex}) + 0.4469 \times (HbA_{1c}) - 3.82$ (where for sex, 1 = male and 0 = female).

and 75%, respectively. When the same strategy was assessed for advanced fibrosis, the AUC for the combination of all models/biomarkers ($n = 6$) was 0.91 (0.85–0.97). Similarly, using a cutoff point of five positive tests (based on cohort-specific cutoff points described in Table 3) (Supplementary Fig. 2), this provided a sensitivity of 71% (49–87), specificity of 94% (88–97), PPV of 68% (46–85), and NPV of 95% (89–98). Removing the two proprietary tests (FibroTest and PRO-C3) from this approach significantly reduced the AUC to 0.85 (0.75–0.94; $P = 0.003$), but NPV remained in the same range of 94% (88–98).

CONCLUSIONS

Due to the increased risk of advanced liver disease in patients with T2DM and NAFLD (4,28), there is an urgent need for accurate noninvasive diagnostic tools to identify these patients, as several glucose-lowering medications have been shown to be safe and effective to treat patients with NAFLD. In the current work, we compared head-to-head the most important noninvasive clinical scores and plasma biomarkers for the diagnosis of definite NASH or advanced fibrosis in a cohort of patients with T2DM. Our results suggest that the most frequently used noninvasive scores/biomarkers for the diagnosis of definite NASH and/or advanced fibrosis may not be significantly better than plasma ALT or AST concentration, respectively, to assist in the management of patients with T2DM.

Therefore, due to their low cost and availability, it is likely that ALT and AST are still the most important stand-alone tests that can help health care providers decide which patients to biopsy. Complementary and more advanced diagnostic imaging methods, such as transient elastography or magnetic resonance elastography, are not widely available. Therefore, it is likely that unless a combination of greater clinician awareness, a multidisciplinary approach (i.e., primary care physicians, endocrinologists, and hepatologists), and development of novel, noninvasive algorithms is implemented, many patients with T2DM will continue to be at risk for NASH or advanced fibrosis and go undiagnosed for long periods of time due to the low sensitivity of current noninvasive blood biomarkers.

The use of sequential and parallel multiple testing was also assessed in our study. Unfortunately, no major significant improvements in the overall performance were observed when these strategies were tested. While we observed some trends toward improvement in sensitivity and specificity when these diagnostic approaches were combined, larger studies are required to assess whether these improvements are cost-effective or if, on the contrary, we should still rely on plasma ALT and AST until a better diagnostic tool is available.

Several reasons may explain their underperformance in patients with T2DM. First, patients with T2DM may only represent a relatively small part of the whole spectrum of NAFLD severity, resulting in a potential spectrum effect. However, this explanation is unlikely due
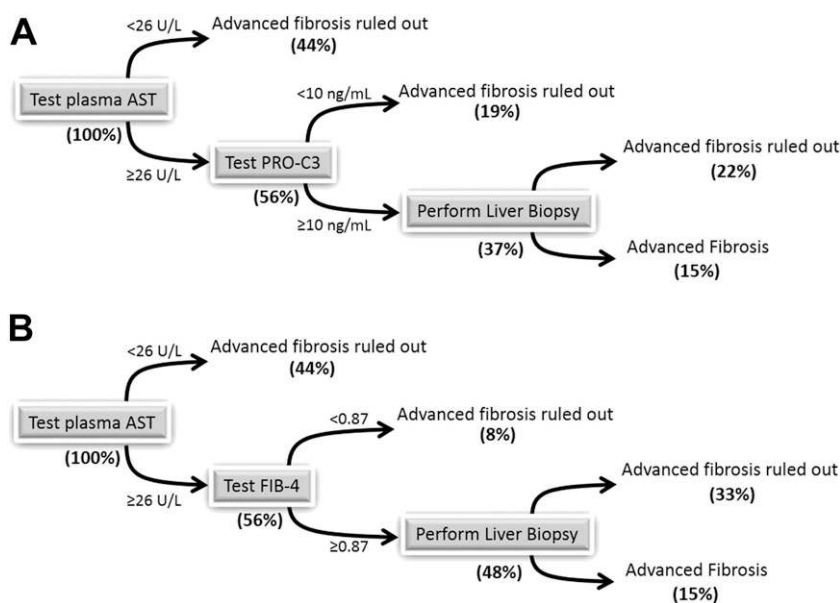
Figure 2—Sequential testing for the diagnosis of advanced fibrosis optimized for an NPV of 100% (i.e., avoiding any false negative) assessing PRO-C3 (A) and FIB-4 (B) after use of AST. Results with APRI and NAFLD fibrosis score were similar to FIB-4 and can be found in Supplementary Fig. 1.

to the wide range of liver disease included in this study (from absence of NAFLD to severe NASH with advanced fibrosis). Second, glucose-lowering agents may significantly affect liver fat accumulation and/or measurements used to calculate these scores (i.e., plasma ALT, AST, BMI, etc.). In order to avoid any bias from different use of glucose-lowering medications, only metformin, sulfonylureas, and insulin were allowed. Moreover, use of these medications was similar among all groups. In addition, this is a multiethnic cohort, which could have also affected the results, as many of these diagnostic tools have been developed and validated in Caucasian populations (14–18). While our sensitivity analyses did not observe significant differences in the performance of the tests among the different ethnic groups, this study was not specifically designed to assess these differences. Therefore, it is still possible that these noninvasive tests may perform differently among different ethnic groups. Finally, observational studies such as this are always prone to some selection bias that can affect generalizability of the study. Overall, the prevalence of NAFLD in our cohort was 76%, similar to the prevalence usually reported in the literature (~60–75%) (5), with a recent meta-analysis reporting prevalence rates from 29.6% to 87.1% (29). The prevalence of advanced fibrosis among these patients was 19% in our cohort, which is similar to the 17.2% reported by Koehler et al. (30) as well as the 17.7% reported by Kwok et al. (31) in patients with NAFLD and T2DM using transient elastography in population-based studies. These results suggest that this cohort of patients with T2DM is somehow comparable to the general population with T2DM. However, as this study was focused on assessing noninvasive tests for the diagnosis of NASH and/or advanced fibrosis, the main emphasis was actually placed on having all stages of the disease well represented, which is the reason why patients were recruited from different sources.

In summary, results from our work suggest that noninvasive clinical scores and plasma biomarkers have significant limitations as stand-alone tests for the diagnosis of definite NASH and/or advanced fibrosis in patients with T2DM. While they may offer diagnostic guidance

for a small number of patients with extreme values, they will not provide definitive results for most of the population. Moreover, the use of combinations of these blood-based approaches was not shown in our study to meaningfully improve the performance of these tests or their discrimination capacity. The most important implication of these findings is that they should be used with caution in patients with diabetes, taking into the account the characteristics of the target population (i.e., in whom the prevalence of advanced fibrosis may be less than in hepatology clinics). In this setting, the greatest value of the tests may lie in ruling out advanced fibrosis (i.e., high NPV). New, noninvasive, and affordable diagnostic tools for definite NASH and advanced fibrosis are very much needed in order to tackle this epidemic. Future work should focus on developing a combination of imaging and noninvasive clinical scores and plasma biomarkers to optimize the noninvasive diagnosis of advanced fibrosis. Until then, liver biopsy remains the gold standard for the diagnosis of definite NASH and advanced fibrosis.

Duality of Interest. M.J.M., M.P.C., D.S., and C.M.R. are full-time employees and stockholders at Quest Diagnostics. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. F.B. was responsible for patient recruitment and follow-up, study design, data acquisition and interpretation, statistical analysis, and writing and editing of the manuscript. M.J.M. and M.P.C. were responsible for study design and data acquisition, aided in discussion of data analysis, provided inputs, and approved the manuscript. V.C.C., C.S.-P., and R.J.F.-M. were responsible for data acquisition and critical revision of the manuscript. J.L. was responsible for reading of liver biopsies and critical revision of the manuscript. D.S. and C.M.R. were responsible for statistical analysis, interpretation of results, and critical revision of the manuscript. K.C. was responsible for study design and funding, data acquisition and interpretation, and critical revision and editing of the manuscript. F.B. and K.C. are the guarantors of this work and, as such, had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Bril F, Cusi K. Nonalcoholic fatty liver disease: the new complication of type 2 diabetes mellitus.

Endocrinol Metab Clin North Am 2016;45:765–781
2. Williamson RM, Price JF, Glancy S, et al.; Edinburgh Type 2 Diabetes Study Investigators. Prevalence of and risk factors for hepatic steatosis and nonalcoholic fatty liver disease in people with type 2 diabetes: the Edinburgh Type 2 Diabetes Study. Diabetes Care 2011; 34:1139–1144
3. El-Serag HB, Tran T, Everhart JE. Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma. Gastroenterology 2004; 126:460–468
4. Hazlehurst JM, Woods C, Marjot T, Cobbold JF, Tomlinson JW. Non-alcoholic fatty liver disease and diabetes. Metabolism 2016;65:1096–1108
5. Bril F, Cusi K. Management of nonalcoholic fatty liver disease in patients with type 2 diabetes: a call to action. Diabetes Care 2017;40: 419–430
6. American Diabetes Association. 4. Comprehensive medical evaluation and assessment of comorbidities: *Standards of Medical Care in Diabetes—2019*. Diabetes Care 2019;42(Suppl. 1):S34–S45
7. Bugianesi E, Vanni E, Marchesini G. NASH and the risk of cirrhosis and hepatocellular carcinoma in type 2 diabetes. Curr Diab Rep 2007;7:175–180
8. Caldwell S, Marchesini G. Cryptogenic vs. NASH-cirrhosis: the rose exists well before its name..... J Hepatol 2018;68:391–392
9. Portillo-Sanchez P, Bril F, Maximos M, et al. High prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus and normal plasma aminotransferase levels. J Clin Endocrinol Metab 2015;100:2231–2238
10. Verma S, Jensen D, Hart J, Mohanty SR. Predictive value of ALT levels for non-alcoholic steatohepatitis (NASH) and advanced fibrosis in non-alcoholic fatty liver disease (NAFLD). Liver Int 2013;33:1398–1405
11. Bril F, Ortiz-Lopez C, Lomonaco R, et al. Clinical value of liver ultrasound for the diagnosis of nonalcoholic fatty liver disease in overweight and obese patients. Liver Int 2015;35:2139–2146
12. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. Hepatology 2018;67:328–357
13. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. Diabetologia 2016;59:1121–1140
14. 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
15. Ratziu V, Massard J, Charlotte F, et al.; LIDO Study Group; CYTOL study group. Diagnostic value of biochemical markers (FibroTest-FibroSURE) for the prediction of liver fibrosis in patients with non-alcoholic fatty liver disease. BMC Gastroenterol 2006;6:6
16. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. Gut 2008;57:1441–1447

17. Pearce SG, Thosani NC, Pan JJ. Noninvasive biomarkers for the diagnosis of steatohepatitis and advanced fibrosis in NAFLD. Biomark Res 2013;1:7

18. Dixon JB, Bhathal PS, O'Brien PE. Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. Gastroenterology 2001;121: 91–100

19. Bril F, Millán L, Kalavalapalli S, et al. Use of a metabolomic approach to non-invasively diagnose non-alcoholic fatty liver disease in patients with type 2 diabetes mellitus. Diabetes Obes Metab 2018;20:1702–1709

20. Bril F, McPhaul MJ, Caulfield MP, et al. Performance of the SteatoTest, ActiTest, Nash-Test and FibroTest in a multiethnic cohort of patients with type 2 diabetes mellitus. J Investig Med 2019;67:303–311

21. Bertot LC, Jeffrey GP, de Boer B, et al. Diabetes impacts prediction of cirrhosis and prognosis by non-invasive fibrosis models in non-alcoholic fatty liver disease. Liver Int 2018; 38:1793–1802

22. Bril F, Leeming DJ, Karsdal MA, et al. Use of plasma fragments of propeptides of type III, V, and VI procollagen for the detection of liver fibrosis in type 2 diabetes. Diabetes Care 2019; 42:1348–1351

23. Bril F, Portillo-Sanchez P, Liu IC, Kalavalapalli S, Dayton K, Cusi K. Clinical and histologic characterization of nonalcoholic steatohepatitis in African American patients. Diabetes Care 2018;41:187–192

24. Bril F, Barb D, Portillo-Sanchez P, et al. Metabolic and histological implications of intrahepatic triglyceride content in nonalcoholic fatty liver disease. Hepatology 2017;65:1132–1144

25. Sanyal AJ, Brunt EM, Kleiner DE, et al. Endpoints and clinical trial design for nonalcoholic steatohepatitis. Hepatology 2011;54:344–353

26. 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;41:1313–1321

27. Nielsen MJ, Nedergaard AF, Sun S, et al. The neo-epitope specific PRO-C3 ELISA measures true formation of type III collagen associated with liver and muscle parameters. Am J Transl Res 2013;5:303–315

28. Loomba R, Abraham M, Unalp A, et al.; Nonalcoholic Steatohepatitis Clinical Research Network. Association between diabetes, family history of diabetes, and risk of nonalcoholic steatohepatitis and fibrosis. Hepatology 2012; 56:943–951

29. Dai W, Ye L, Liu A, et al. Prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus: a meta-analysis. Medicine (Baltimore) 2017;96:e8179

30. Koehler EM, Plompen EP, Schouten JN, et al. Presence of diabetes mellitus and steatosis is associated with liver stiffness in a general population: the Rotterdam study. Hepatology 2016; 63:138–147

31. Kwok R, Choi KC, Wong GL, et al. Screening diabetic patients for non-alcoholic fatty liver disease with controlled attenuation parameter and liver stiffness measurements: a prospective cohort study. Gut 2016;65:1359–1368