ORIGINAL ARTICLE (DIRECT VIA EEO)

Liver INTERNATIONAL  WILEY

# Repeated measurements of non-invasive fibrosis tests to monitor the progression of non-alcoholic fatty liver disease: A long-term follow-up study

Wile Balkhed[1] ⬤  |  Fredrik O. Åberg[2] ⬤  |  Patrik Nasr[1] ⬤  |  Mattias Ekstedt[1] ⬤  |
Stergios Kechagias[1] ⬤

[1]Department of Health, Medicine, and Caring Sciences, Linköping University, Linköping, Sweden

[2]Transplantation and Liver Surgery Clinic, Helsinki University Hospital, Helsinki University, Helsinki, Finland

**Correspondence**
Stergios Kechagias, Department of Gastroenterology and Hepatology, Department of Health, Medicine, and Caring Sciences, Linköping University, SE-581 83 Linköping, Sweden.
Email: stergios.kechagias@liu.se

## Abstract

**Background and Aims:** The presence of advanced hepatic fibrosis is the prime marker for the prediction of liver-related complications in non-alcoholic fatty liver disease (NAFLD). Blood-based non-invasive tests (NITs) have been developed to evaluate fibrosis and identify patients at risk. Current guidelines propose monitoring the progression of NAFLD using repeated NITs at 2–3-year intervals. The aim of this study was to evaluate the association of changes in NITs measured at two time points with the progression of NAFLD.

**Methods:** We retrospectively included NAFLD patients with NIT measurements in whom the baseline hepatic fibrosis stage had been assessed by biopsy or transient elastography (TE). Subjects underwent follow-up visits at least 1 year from baseline to evaluate the progression of NAFLD. NAFLD progression was defined as the development of end-stage liver disease or fibrosis progression according to repeat biopsy or TE. The following NITs were calculated at baseline and follow-up: Fibrosis-4 (FIB-4), NAFLD fibrosis score (NFS), aspartate aminotransferase to platelet ratio index (APRI) and dynamic aspartate-to-alanine aminotransferase ratio (dAAR).

**Results:** One hundred and thirty-five patients were included with a mean follow-up of $12.6 \pm 8.5$ years. During follow-up, 41 patients (30%) were diagnosed with progressive NAFLD. Change in NIT scores during follow-up was significantly associated with disease progression for all NITs tested except for NFS. However, the diagnostic precision was suboptimal with area under the receiver operating characteristics 0.56–0.64 and positive predictive values of 0.28–0.36 at sensitivity fixed at 90%.

**Conclusions:** Change of FIB-4, NFS, APRI, and dAAR scores is only weakly associated with disease progression in NAFLD. Our findings do not support repeated measurements of these NITs for monitoring the course of NAFLD.

## 1 | INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) has become the most common chronic liver disease with an estimated global prevalence of 25%.[1] Although the majority of NAFLD patients will never experience liver-related complications, the progressive potential of NAFLD is indisputable with 5%–10% of subjects progressing to cirrhosis, end-stage liver disease (ESLD) or hepatocellular carcinoma (HCC).[2] In some parts of the world, NAFLD is now the primary aetiology of cirrhosis,[3] the main cause of HCC[4] and the second-leading cause of liver transplantation.[5]

Non-alcoholic fatty liver disease entails a spectrum of histopathological features that range from simple steatosis, via the establishment of lobular inflammation and hepatocellular injury (i.e. non-alcoholic steatohepatitis [NASH]) with or without fibrosis, to cirrhosis with the risk of developing ESLD or HCC.[6,7] Longitudinal studies have demonstrated that NAFLD patients with advanced fibrosis (i.e. fibrosis stage 3–4) are at the highest risk of developing cirrhosis-related complications which may lead to liver transplantation or death.[8–10] The presence of advanced fibrosis, particularly cirrhosis, alters clinical management, including the possible initiation of surveillance for gastroesophageal varices and HCC.

Given the high prevalence of NAFLD, detecting advanced fibrosis using liver biopsy is not plausible. Numerous methods have been developed for the non-invasive evaluation of liver fibrosis in NAFLD, particularly blood tests and elastography devices.[11,12] However, the availability of transient elastography (TE) is limited to specialized centres, which is a limitation for the identification of NAFLD patients with advanced fibrosis in larger populations. Several non-invasive scores, combining various parameters, have been developed to identify individuals with prevalent advanced fibrosis.[13] Of these, the most common scores are composed of widely available blood tests and therefore easily accessible for physicians. In general, these scores have excellent ability to exclude advanced fibrosis but their ability to detect advanced fibrosis is lower.[13] Current guidelines recommend that blood-based fibrosis scores should be calculated for every NAFLD patient, to rule out significant fibrosis.[14] If significant fibrosis cannot be ruled out, patients should be referred to a specialized centre for TE.[15,16] The optimal follow-up of NAFLD patients is yet to be determined. However, guidelines and position papers state that it is reasonable to monitor the progression of the hepatic disease by analysing blood-based non-invasive tests (NITs) for fibrosis monitoring at 2–3-year intervals.[12,15,16]

In this study, we evaluate if repeated measurements of the widely available NITs for assessment of fibrosis in NAFLD, fibrosis-4 (FIB-4), NAFLD fibrosis score (NFS), aspartate aminotransferase to platelet

**Lay Summary**

Patients with non-alcoholic fatty liver disease (NAFLD) may progress to cirrhosis and liver failure. Analysis of blood tests at 2–3-year intervals is recommended to monitor disease progression. We followed 135 NAFLD patients and evaluated if repeated measurements of four common tests can be used to distinguish patients with progressive NAFLD. We found that the precision of the tests was not optimal.

ratio index (APRI), and the newly developed dynamic aspartate-to-alanine aminotransferase ratio (dAAR)[17] improve the detection of NAFLD patients with progressive disease. Our specific aims were to (i) investigate the association of changes in NITs measured at two time points with the progression of disease severity in NAFLD and (ii) evaluate if analysis of NITs at baseline can predict future progression of NAFLD.

## 2 | MATERIALS AND METHODS

### 2.1 | Subjects and setting

In this retrospective cohort study, we included patients who had previously been diagnosed with NAFLD and had undergone an assessment of hepatic fibrosis either by histopathological evaluation or TE at the Department of Gastroenterology and Hepatology, University Hospital in Linköping, Sweden, between 1987 and 2019. At baseline, a diagnostic workup was performed including physical examination, laboratory investigation and an extensive review of patients' medical charts. Steatosis was diagnosed histopathologically or with imaging. After excluding other chronic liver diseases, medication associated with hepatic steatosis, and significant alcohol consumption (>210 g/week for men and >140 g/week for women) patients were diagnosed with NAFLD.

Eligible subjects underwent follow-up at least 1 year from baseline to evaluate the course of NAFLD. The follow-up visit included clinical assessment for the development of ESLD, biochemical investigations and, in those without signs of ESLD, assessment of hepatic fibrosis either by histopathology or by TE. Patients during follow-up had developed ascites, hepatic encephalopathy, variceal bleeding or bilirubin >3 mg/dl were considered having progressed to ESLD. Procedures for identification of included patients and reasons for exclusions are summarized in Figure 1.

**FIGURE 1** Details about patients showing reasons for exclusions ESLD, End-stage liver disease; TE, Transient elastography
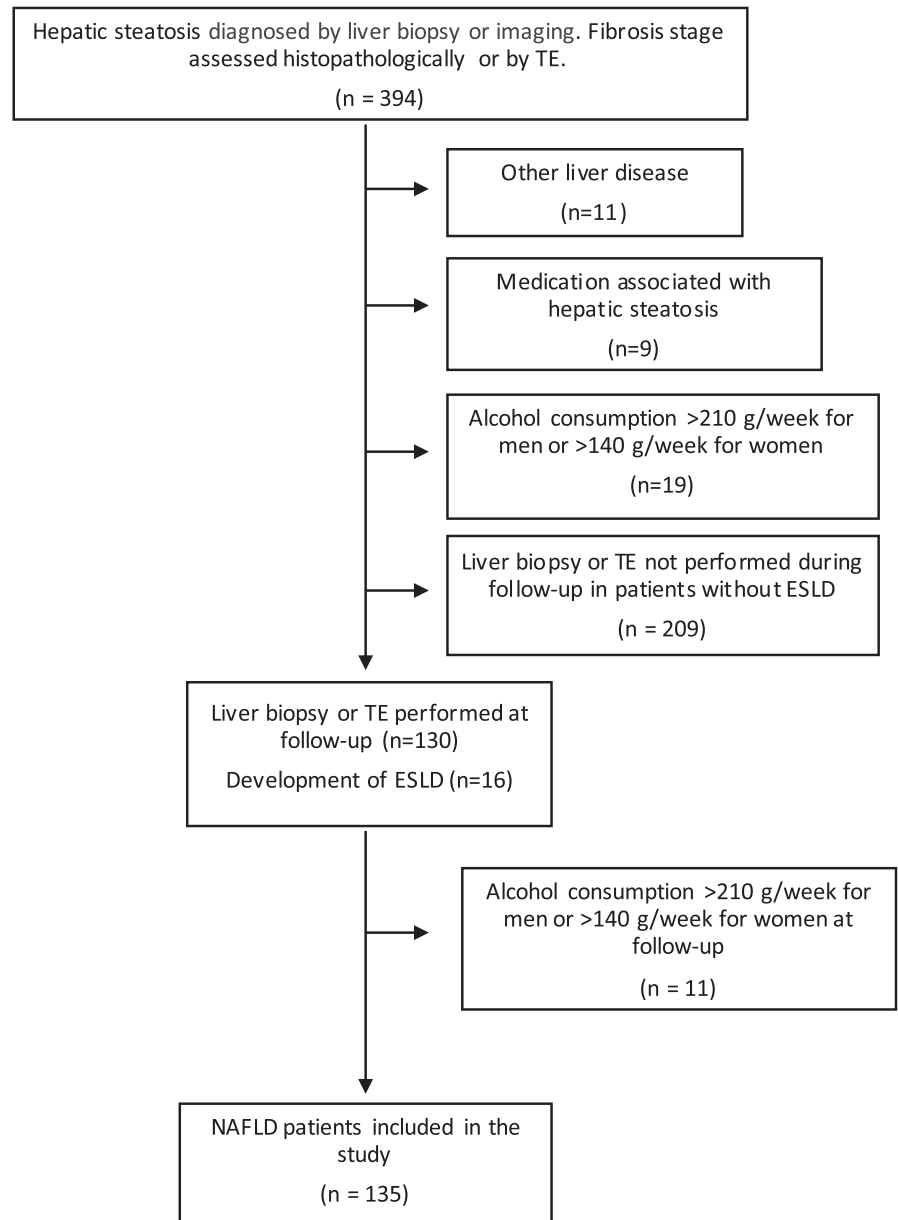


## 2.2 | Data collection

### 2.2.1 | General patient data

Baseline characteristics on patient age, gender, height and weight as well as diagnosis of type 2 diabetes mellitus (T2DM) were collected from medical charts. T2DM was defined as fasting plasma glucose >126 mg/dl or plasma glucose >199 mg/dl 2 h after a standard dose of oral glucose. Impaired glucose tolerance (IGT) was defined as plasma glucose >140 mg/dl but <199 mg/dl 2 h after a standard dose of oral glucose.

### 2.2.2 | Serum fibrosis algorithms

Subjects had blood drawn at baseline and follow-up. FIB-4,[18] NFS,[19] APRI,[20] and dAAR[17] were calculated according to previously published formulas (Table S1).

### 2.2.3 | Hepatic fibrosis assessment at baseline and at follow-up

Hepatic fibrosis assessment was performed in patients without signs of ESLD using either liver biopsy or TE. Liver biopsies were performed percutaneously with ultrasound guidance and 1.6 mm Biopince needles on an outpatient basis. Histopathological fibrosis stage was scored according to Kleiner et al.[21] Advanced fibrosis was defined as stage 3 or 4 (F3 or F4). TE was performed using FibroScan (Echosens) by experienced hepatologists (S.K., M.E., P.N.) according to a previously described approach.[22] Median liver stiffness <8 kPa by TE was considered to rule out compensated advanced NAFLD, while median liver stiffness >12 kPa was considered to confirm compensated advanced NAFLD.[23] A minimum of 10 valid readings, with at least a 60% success rate, and an interquartile range of ≤30% of the median value were required for reliable liver stiffness measurements.[24] Fibrosis assessment by liver biopsy or TE was performed

within 6 months of the clinical and biochemical examinations both at baseline and at follow-up.

## 2.2.4 | Progression of non-alcoholic fatty liver disease during follow-up

Patients fulfilling one of the following criteria during follow-up were considered to have progressive NAFLD: (i) development of ESLD, (ii) progression of fibrosis stage from F0–F2 to F3–F4, (iii) progression of fibrosis stage from F0 to F2, (iv) progression of liver stiffness by TE from <8 to >12 kPa and (v) fibrosis stage F0–F2 at baseline and liver stiffness by TE >12 kPa at the follow-up if liver biopsy had not been performed at the follow-up. In case of development of ESLD during follow-up, NAFLD was considered progressive irrespective of fibrosis stage or liver stiffness at baseline and follow-up. In patients undergoing both liver biopsy and TE at baseline or follow-up, histopathological fibrosis stage was preferred to classify the severity of NAFLD.

## 2.3 | Statistics

Categorical variables were compared using Pearson's $\chi^2$ test. Continuous variables were tested for normality of distribution using Shapiro Wilk's test. Normally distributed continuous variables were compared using $t$ test and non-normally distributed continuous variables were compared using the Mann–Whitney $U$ test. Correlations were assessed by Spearman correlation. Two-sided $p$ values were used. Missing points of data were excluded pairwise without imputation.

The association between baseline model score (FIB-4, APRI, NFS or dAAR) and progression of NAFLD during follow-up was analysed by univariable logistic regression. We repeated this using the delta values of the scores, that is the change in model score from baseline to the follow-up. We then performed multivariable logistic regression analysis separately for each model with the delta-score as the independent variable, disease progression as the outcome variable and baseline model score, age, sex and length of follow-up as covariates. Predictive performance of the various models was evaluated using areas under the receiver operative characteristics (AUROCs) and by sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). Cut-offs for a change from baseline to follow-up in the various scores were chosen based on fixed sensitivity at or above 90% and fixed specificity at or above 90%. A potential non-linear association between change in model score and likelihood of NAFLD progression was visualized by spline plots using restricted cubic splines.

## 3 | RESULTS

### 3.1 | Study population and clinical characteristics

One hundred and thirty-five patients were included with a mean age of 49.8 (±13.0, range 21–74) years at baseline and with a male

predominance (63%). Mean follow-up time was 12.6 (±8.5, range 1–33) years. No patient had been enrolled in any clinical trial for the treatment of progression of NAFLD or prevention of NASH. Eighteen percent of patients had a diagnosis of either T2DM or IGT at baseline, while 57% had T2DM or IGT at follow-up. Average BMI was 29.0 kg/m² at baseline and increased to 29.9 kg/m² at follow-up. The general characteristics of the patient cohort are shown in Table 1.

One hundred and fourteen patients underwent liver biopsy at baseline. Of these, 103 (90%) had F0–F2 and 11 (10%) had F3–F4. At follow-up, 75 patients underwent liver biopsy. Of these, 55 (73%) had F0–F2, while 20 (27%) had F3–F4. Twenty-one and 53 patients underwent hepatic fibrosis assessment by TE and had no biopsy at baseline and follow-up respectively. In all, 18 patients underwent hepatic fibrosis assessment solely by TE at both baseline and follow-up.

Of the 135 patients, 41 (30%) were classified as progressive NAFLD and 94 (70%) as non-progressive at follow-up. Follow-up time in patients with progressive NAFLD was 13.8 ± 7.7 years and in non-progressive NAFLD 12.1 ± 8.9 years ($p$ = .31). In the progressive group, 12 patients had developed ESLD (ascites: $n$ = 8, variceal bleeding: $n$ = 3, jaundice: $n$ = 1). Of these, 10 had undergone liver biopsy at baseline (F0–F2: $n$ = 7, F3: $n$ = 2, F4: $n$ = 1). Twenty-three patients were classified as progressive by histopathological fibrosis stage. Of these, 11 progressed from F0 to F2, while 12 progressed from F0–F2 to F3–F4. Four patients with fibrosis stage F0–F2 at baseline had liver stiffness by TE >12 kPa at follow-up and two patients progressed from liver stiffness by TE <8 kPa at baseline to >12 kPa at follow-up.

Non-alcoholic fatty liver disease patients with progressive disease had gained more weight and had a higher prevalence of T2DM at follow-up. Moreover, they had significantly higher aminotransferases and lower platelet count at follow-up while these parameters were not significantly different between the two groups at baseline (Table 2).

### 3.2 | Non-invasive test scores and disease progression

Non-invasive test scores at baseline did not differ significantly between the progressive and non-progressive groups. At follow-up, NIT scores were significantly different compared to baseline values for all NITs tested. Changes in NIT scores between baseline and follow-up were significantly different between the progressive and non-progressive group for all NITs tested except for NFS (Table 3).

Univariable logistic regression showed that NIT scores at baseline were not significantly associated with disease progression. However, change in NIT scores during follow-up was significantly associated with disease progression for all NITs tested except for NFS. Furthermore, multivariable logistic regression confirmed the significant independent association of change of NIT scores and progression of NAFLD during follow-up for all NITs except for NFS

**TABLE 1** Baseline characteristics of NAFLD patients (*n* = 135) [Mean ± SD or *n* (%)]

| | |
|---|---|
| Follow-up time (years) | 12.6 ± 8.5 |
| Sex (male) | 85 (63%) |
| Age (years) | 49.8 ± 13.0 |
| T2DM/IGT | 24 (18%) |
| Hyperlipidemia | 42 (31%) |
| Hypertension | 54 (40%) |
| Cardiovascular disease | 9 (7%) |
| Insulin | 5 (4%) |
| GLP-1 analogues | 1 (1%) |
| Oral glucose-lowering agents | 15 (11%) |
| Statins | 28 (21%) |
| Fibrates | 1 (1%) |
| Thiazolidinediones | 0 |
| Vitamin E | 0 |
| BMI (kg/m$^2$) | 29.0 ± 4.7 (*n* = 127) |
| ALT (U/L) | 81 ± 47 |
| AST (U/L) | 45 ± 21 |
| Platelet count (×10$^9$/L) | 227 ± 55 (*n* = 128) |
| Albumin (g/L) | 42 ± 4 (*n* = 132) |
| Histopathology (*n*) | 114 (84%) |
|     F0 | 53 |
|     F1 | 26 |
|     F2 | 24 |
|     F3 | 10 |
|     F4 | 1 |
| dAAR | 1.24 ± 0.88 |
| NFS | −1.93 ± 1.44 (*n* = 120) |
| APRI | 0.51 ± 0.29 (*n* = 128) |
| FIB-4 | 1.24 ± 0.75 (*n* = 128) |

Abbreviations: ALT, alanine aminotransferase; APRI, AST to platelet ratio index; AST, aspartate aminotransferase; BMI, body mass index; dAAR, dynamic aspartate-to-alanine aminotransferase ratio; F0–F4, fibrosis stage 0–4; FIB-4, fibrosis-4; IGT, impaired glucose tolerance; NFS, NAFLD fibrosis score; SD, standard deviation; T2DM, type 2 diabetes mellitus.

(Table S2). Spline plots showing the probability that the patient has progressed according to the change over time in NIT scores are shown in Figure 2. As seen, the relationship between a change in NIT scores and the likelihood of disease progression was almost linear for all NITs. Although the likelihood of disease progression was steepest for FIB-4 and least steep for NFS, the confidence intervals are wide, indicating a high degree of uncertainty.

Area under the receiver operating characteristics of baseline NIT scores for discriminating disease progression were small (0.55–0.58). AUROCs for change of NIT scores over time were only slightly higher (0.56–0.64) (Figure 3).

Analyses of diagnostic performance of delta values of NITs for prediction of disease progression during follow-up was performed at 90% sensitivity and 90% specificity. Specificities, sensitivities, NPVs and PPVs for change of FIB-4, NFS, APRI and dAAR are shown in Table 4.

## 3.3 | Subgroup analyses

Seventy-four patients had undergone liver biopsy both at baseline and follow-up with a mean time between biopsies of 17.1 (±5.5, range 5–33) years. This subgroup was used to investigate the correlation between change in NIT scores and change in fibrosis stage during follow-up. Thirty-eight patients increased in fibrosis stage, 10 decreased, while 26 had an unchanged fibrosis stage (Table S3). Change in dAAR score showed a weak but significant correlation ($\rho$ = .25, *p* = .037). No significant correlations were found for the other three NITs evaluated.

To investigate if the correlation between change in NIT scores and disease progression was dependent on baseline fibrosis stage, patients with histopathological fibrosis assessment at baseline were divided into two groups, F0–F1 and F2–F4 respectively. One hundred and fourteen patients had undergone liver biopsy at baseline. Of these 79 had F0–F1 at baseline, 23 (29%) of which were classified as progressive at follow-up. Thirty-five patients had F2–F4 at baseline and of these 14 (40%) were progressive.

In the F0–F1 group, both univariable and multivariable logistic regression showed a significant association between change in NIT scores and disease progression for dAAR, APRI and FIB-4 but not for NFS. In patients with F2–F4 at baseline, no significant association for any of the tested NITs was found with univariable or multivariable logistic regression (Table S4).

Twelve patients in the cohort developed ESLD. Results of sensitivity analysis including only ESLD as outcome showed similar results as the broader findings of the study. Baseline NITs were not significantly associated with future ESLD development, even when adjusted for time to follow-up. However, change in NITs between baseline and time for the liver-related event was significantly correlated to progression, in both adjusted and unadjusted analyses, for all NITs except NFS (Table S5). AUROCs were, however, mediocre (NFS, 0.58; dAAR, 0.71; APRI, 0.73; FIB-4, 0.80).

## 4 | DISCUSSION

The present long-term follow-up study evaluated the accuracy of four widely available serum fibrosis algorithms (FIB-4, NFS, APRI and dAAR), and their changes over time, to detect progression of NAFLD. Although a change of NIT score was associated with disease progression for all NITs evaluated except for NFS (Table 3), the diagnostic precision was suboptimal with AUROCs 0.56–0.64 (Figure 3B) and PPVs of 0.28–0.36 at sensitivity fixed at 90% (Table 4). This indicates that repeated measurements of these NITs are of limited clinical usefulness for monitoring the course of NAFLD.

**TABLE 2** Clinical, biochemical and histopathological characteristics at baseline and follow-up of NAFLD patients with the progressive and non-progressive disease [Mean ± SD or n (%)]

| | Progressive (n = 41) | Non-progressive (n = 94) | p value[a] |
|---|---|---|---|
| Follow-up time (years) | 13.8 ± 7.7 | 12.1 ± 8.9 | NS[b] |
| Sex (male) | 30 (72%) | 55 (59%) | NS[c] |
| Age (years) | | | |
| Baseline | 50.7 ± 12.5 | 49.4 ± 13.3 | NS[b] |
| Follow-up | 64.4 ± 12.2 | 61.5 ± 11.6 | NS[b] |
| T2DM/IGT | | | |
| Baseline | 8 (21%) | 16 (17%) | NS[c] |
| Follow-up | 28 (70%) | 49 (52%) | NS[c] |
| BMI (kg/m$^2$) | | | |
| Baseline | 29.0 ± 3.1 (n = 38) | 29.1 ± 5.3 (n = 89) | NS[b] |
| Follow-up | 31.2 ± 4.7 (n = 31) | 29.5 ± 5.0 (n = 86) | NS[b] |
| Change | 1.7 ± 3.6 (n = 29) | 0.2 ± 2.6 (n = 82) | .006[b] |
| Hyperlipidemia[a] | 13 | 29 | NS[c] |
| Hypertension[a] | 19 | 35 | NS[c] |
| Cardiovascular disease[a] | 2 | 7 | NS[c] |
| Insulin[d] | 3 | 2 | NS[c] |
| GLP-1 analogues[d] | 0 | 1 | NS[c] |
| Oral glucose-lowering agents[d] | 5 | 10 | NS[c] |
| Statins[d] | 10 | 18 | NS[c] |
| Fibrates[a] | 0 | 1 | NS[c] |
| Thiazolidinediones[a] | 0 | 0 | NS[c] |
| Vitamin E[d] | 0 | 0 | NS[c] |
| ALT (U/L) | | | |
| Baseline | 86 ± 47 | 79 ± 48 | NS[b] |
| Follow-up | 68 ± 37 | 48 ± 29 | <.001[b] |
| Change | −18 ± 59 | −31 ± 50 | NS[b] |
| AST (U/L) | | | |
| Baseline | 48 ± 24 | 44 ± 20 | NS[b] |
| Follow-up | 51 ± 28 | 38 ± 27 | <.001[b] |
| Change | 3 ± 35 | −6 ± 26 | NS[b] |
| Platelet count (×10$^9$/L) | | | |
| Baseline | 218 ± 49 (n = 39) | 231 ± 58 (n = 89) | NS[e] |
| Follow-up | 195 ± 60 | 234 ± 60 (n = 93) | .001[b] |
| Change | −27 ± 68 (n = 39) | 2 ± 51 (n = 88) | .011[b] |
| Albumin (g/L) | | | |
| Baseline | 41 ± 3 | 42 ± 4 (n = 91) | .027[e] |
| Follow-up | 41 ± 5 (n = 34) | 41 ± 4 (n = 93) | NS[e] |
| Change | 1 ± 5 (n = 34) | −1 ± 5 (n = 90) | NS[e] |
| Baseline fibrosis assessment | | | |
| Histopathology (n) | 37 | 77 | |
| F0 | 18 | 35 | |
| F1 | 5 | 21 | |
| F2 | 11 | 13 | |
| F3 | 2 | 8 | |
| F4 | 1 | | |

**TABLE 2** (Continued)

|  | Progressive (n = 41) | Non-progressive (n = 94) | p value[a] |
|---|---|---|---|
| TE (n) | 4 | 17 | |
| kPa | 5.6 ± 1.0 | 8.8 ± 5.4 | |
| <8 kPa | 4 | 11 | |
| >12 kPa | | 3 | |
| Follow-up fibrosis assessment | | | |
| Histopathology (n) | 28 | 47 | |
| F0 | | 21 | |
| F1 | | 19 | |
| F2 | 11 | 4 | |
| F3 | 10 | 3 | |
| F4 | 7 | | |
| TE (n) | 6 | 47 | |
| kPa | 16.5 ± 5.6 | 7.5 ± 4.3 | |
| <8 kPa | 0 | 31 | |
| >12 kPa | 6 | 5 | |

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; F0–F4, fibrosis stage 0–4; IGT, impaired glucose tolerance; NS, not significant; SD, standard deviation; T2DM, type 2 diabetes mellitus; TE, transient elastography.

[a]p values were calculated comparing the progressive and non-progressive groups.

[b]Mann–Whitney U test. Bold values indicate statistically significant differences.

[c]$\chi^2$-test.

[d]At baseline.

[e]Student's t test. Bold value indicates statistically significant difference.

Worsening of FIB-4 and NFS has previously been associated with histological progression of fibrosis in a study with serial biopsies a median of 6.6 years apart.[25] Of the 108 patients included, 42% had progression of fibrosis and there was a significant relationship between the change in fibrosis stage between biopsies and the change in both NFS and FIB-4. However, the correlation was weak ($r_s$ = .24 for both scores). Similarly, Siddiqui et al.[26] evaluated 216 biopsy-proven NAFLD patients with non-advanced fibrosis at baseline. A repeat liver biopsy a median of 2.6 years later showed progression to advanced fibrosis in 35 patients (16%). Changes in APRI, FIB-4, NFS and aspartate aminotransferase (AST)/alanine aminotransferase (ALT) ratio were significantly associated with disease progression, and the authors concluded that these NITs could detect fibrosis progression in NAFLD. It should be noted that their results are not readily comparable to ours because of different study designs. In our study, mean follow-up time was significantly longer (12.6 years), which enabled us to include several patients developing ESLD in the progressive group. Despite the difference in the definition of progressive NAFLD, and although Siddiqui et al. reported high NPVs for all NITs evaluated, PPVs at sensitivity fixed at 90% were suboptimal (35%–37%). This is in accordance with our results and questions the clinical utility of these NITs.

In a study conducted in a general population setting, Hagström et al.[27] reported that repeated measurements of FIB-4 can, in comparison with a single measurement, help to identify individuals who are at higher risk of developing severe liver disease. Although the mean follow-up time between measurements was short (2.4 years),

an increase in the FIB-4 over time was associated with increased risk for future ESLD, while a decrease in the FIB-4 was associated with reduced risk. However, even if there was a clear association between higher ESLD risk and FIB-4, still nearly half (48.4%) of the ESLD events occurred in subjects classified as low risk by FIB-4 on both occasions. This indicates the limited usefulness of repeated measurements of FIB-4 in the follow-up of individual NAFLD patients.

In our study, NIT scores at baseline were unable to predict the future progression of NAFLD. Contrary to our results, Younossi et al.[28] recently reported that baseline NIT scores and their changes over time predicted adverse clinical and patient-related outcomes. However, limitations of their study were the relatively short follow-up (median 16 months) and that they included only patients with advanced NASH (F3–F4). In addition, patients were enrolled on clinical trials which might limit the generalizability of their findings. Interestingly, Younossi et al. showed that the best NIT to predict future adverse outcomes was the enhanced liver fibrosis (ELF) test. Unfortunately, the ELF test was unavailable in our data.

Using a large Swedish cohort, Hagström et al.[29] reported that higher NIT scores at baseline were associated with an increased risk of cirrhosis in the general population. However, the predictive ability of all the NITs was modest, and the authors concluded that better scoring systems are needed to evaluate the risk for adverse liver events in the general population and in primary care.

Non-alcoholic fatty liver disease is in most cases a slowly progressive disease and most patients never experience liver-related

**TABLE 3** Serum fibrosis algorithms at baseline and follow-up in NAFLD patients with the progressive and non-progressive disease [Mean ± SD or n (%)]

| | Progressive (n = 41) | Non-progressive (n = 94) | p value[a] |
|---|---|---|---|
| **pdAAR** | | | |
| Baseline | 1.38 ± 0.78 | 1.18 ± 0.92 | NS[b] |
| Follow-up | 1.99 ± 1.28 | 1.23 ± 1.23 | **.001**[b] |
| Change | 0.61 ± 1.46 | 0.04 ± 1.20 | **.028**[c] |
| **Risk group** | | | |
| Baseline | | | NS[d] |
| Minimal | 31 (76%) | 72 (77%) | |
| Low | 6 (15%) | 18 (19%) | |
| Intermediate | 4 (10%) | 2 (2%) | |
| High | 0 | 2 (2%) | |
| Follow-up | | | NS[d] |
| Minimal | 18 (44%) | 63 (67%) | |
| Low | 11 (27%) | 18 (19%) | |
| Intermediate | 9 (22%) | 10 (10%) | |
| High | 3 (7%) | 3 (3%) | |
| Change | | | **.021**[d] |
| Increase | 18 (44%) | 21 (22%) | |
| Decrease | 5 (12%) | 8 (9%) | |
| **NFS** | | | |
| Baseline | −1.75 ± 1.39 (n = 37) | −2.00 ± 1.47 (n = 83) | NS[b] |
| Follow-up | −0.28 ± 1.45 (n = 30) | −1.00 ± 1.34 (n = 86) | **.015**[b] |
| Change | 1.28 ± 1.05 (n = 28) | 1.07 ± 1.25 (n = 76) | NS[b] |
| **Risk group** | | | |
| Baseline | | | NS[d] |
| Low | 26 (70%) | 56 (67%) | |
| Intermediate | 10 (27%) | 25 (30%) | |
| High | 1 (3%) | 2 (2%) | |
| Follow-up | | | **.038**[d] |
| Low | 5 (17%) | 34 (40%) | |
| Intermediate | 18 (60%) | 43 (50%) | |
| High | 7 (23%) | 9 (10%) | |
| Change | | | NS[d] |
| Increase | 18 (64%) | 30 (39%) | |
| Decrease | 0 | 3 (4%) | |
| **APRI** | | | |
| Baseline | 0.55 ± 0.30 (n = 39) | 0.49 ± 0.29 (n = 89) | NS[c] |
| Follow-up | 0.72 ± 0.53 | 0.45 ± 0.51 (n = 93) | **<.001**[c] |
| Change | 0.16 ± 0.61 (n = 39) | −0.03 ± 0.35 (n = 88) | **.021**[c] |
| **Risk group** | | | |
| Baseline | | | NS[d] |
| Low | 24 (62%) | 56 (63%) | |
| Intermediate | 15 (38%) | 32 (36%) | |
| High | 0 | 1 (1%) | |

**TABLE 3** (Continued)

| | Progressive (n = 41) | Non-progressive (n = 94) | p value[a] |
|---|---|---|---|
| Follow-up | | | **<.001**[d] |
| Low | 15 (37%) | 75 (81%) | |
| Intermediate | 23 (56%) | 16 (17%) | |
| High | 3 (7%) | 2 (2%) | |
| Change | | | **<.001**[d] |
| Increase | 15 (38%) | 7 (8%) | |
| Decrease | 4 (10%) | 22 (25%) | |
| **FIB-4** | | | |
| Baseline | 1.32 ± 0.73 (n = 39) | 1.21 ± 0.76 (n = 89) | NS[c] |
| Follow-up | 2.52 ± 2.02 | 1.62 ± 1.11 (n = 93) | **.001**[c] |
| Change | 1.18 ± 2.10 (n = 39) | 0.43 ± 0.79 (n = 88) | **.016**[c] |
| **Risk group** | | | |
| Baseline | | | NS[d] |
| Low | 24 (62%) | 59 (66%) | |
| Intermediate | 12 (31%) | 29 (33%) | |
| High | 3 (7%) | 1 (1%) | |
| Follow-up | | | **.002**[d] |
| Low | 7 (17%) | 44 (47%) | |
| Intermediate | 24 (59%) | 39 (42%) | |
| High | 10 (24%) | 10 (11%) | |
| Change | | | **.045**[d] |
| Increase | 20 (51%) | 27 (31%) | |
| Decrease | 3 (7%) | 4 (5%) | |

Abbreviations: APRI, AST to platelet ratio index; dAAR, dynamic aspartate-to-alanine aminotransferase ratio; FIB-4, fibrosis-4; NFS, NAFLD fibrosis score; NS, not significant; SD, standard deviation.

[a]p values were calculated comparing the progressive and non-progressive groups.

[b]Mann–Whitney U test. Bold value indicates statistically significant difference.

[c]$\chi^2$-test. Bold values indicate statistically significant differences.

[d]Student's t test. Bold values indicate statistically significant differences.

events.[2] A major strength of our study is the long follow-up time enabling us to identify a significant amount (30%) of patients with progressive NAFLD, including several developing ESLD. Another strength is that we used strict definitions of significant fibrosis progression. Previous studies[30] have shown that NAFLD patients with advanced fibrosis (F3–F4) have a worse prognosis. Thus, progression from F0–F2 to F3–F4 represents an important clinical event with respect to fibrosis progression. Recently, it was also shown that the probability of compensated advanced chronic liver disease (cACLD) is very high in NAFLD patients exhibiting liver stiffness by TE >12 kPa, while the probability of cACLD is very low in those with liver stiffness by TE <8 kPa.[23] Thus, progression of liver stiffness by TE from <8 to >12 kPa is also likely to represent a significant clinical event in NAFLD. Another important issue is whether a change of
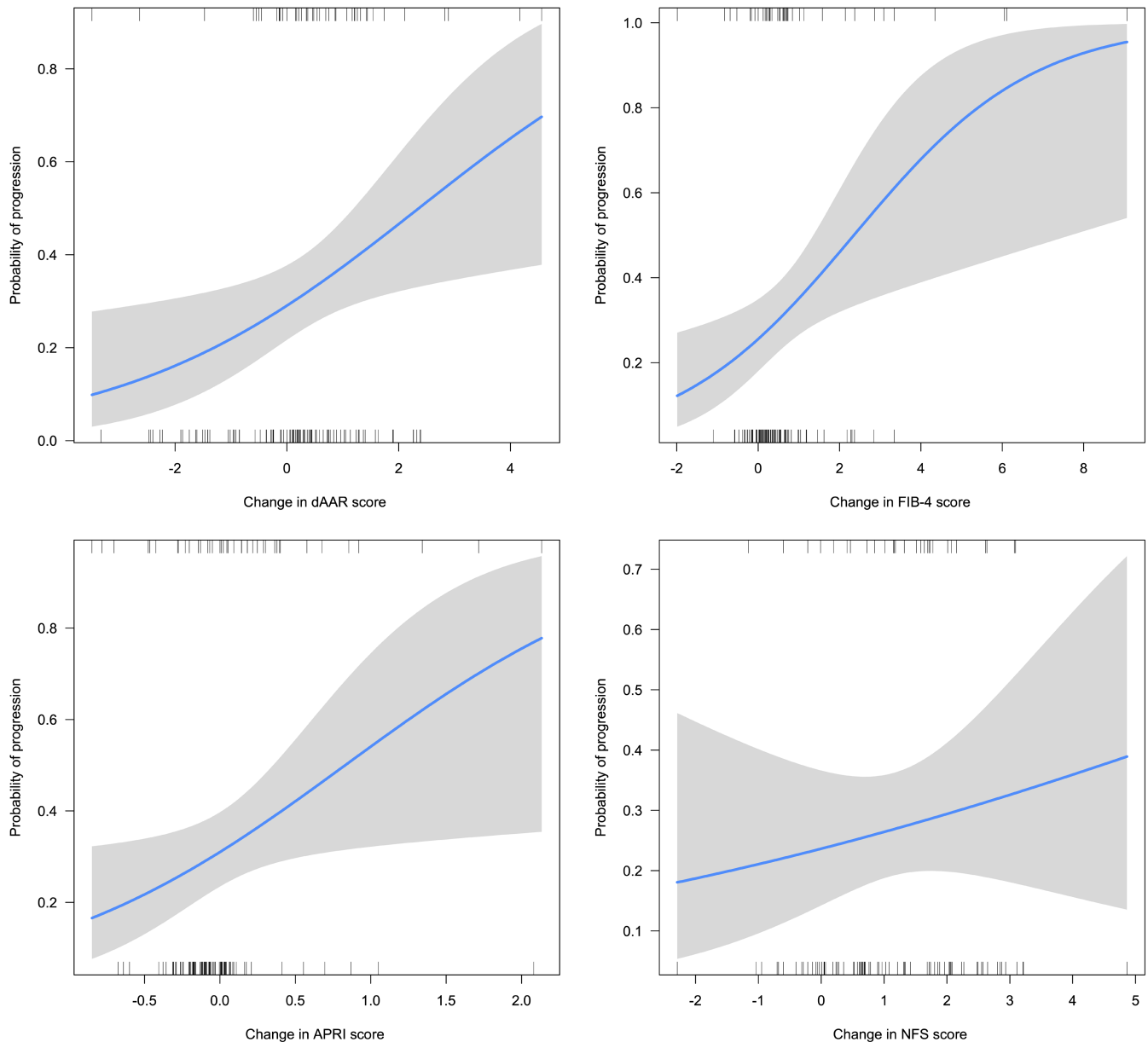
**FIGURE 2** Spline plots showing the probability that the patient has progressed according to the change over time in serum fibrosis algorithms (unadjusted)

NITs represents a 'true' progression of fibrosis stage or a progression that may have occurred 'falsely' owing to sampling errors because of inadequate sample acquisition or incorrect sample representation. Ratziu et al.[31] obtained two samples from NAFLD patients that underwent liver biopsy and showed that the probability of one stage or more difference was 41%. However, the probability of two stages or more difference was significantly lower (12%). Thus, we believe that progression from F0 to F2 or higher during follow-up is most likely to represent a true progression of the fibrosis stage.

Non-invasive tests have primarily been designed to identify or exclude NAFLD patients with advanced fibrosis using liver biopsy as the reference standard. Transition from compensated cirrhosis to ESLD does not necessarily imply progression of fibrosis meaning that a significant change in NIT score does not have to occur.

A limitation of the present study is that patients developing ESLD were considered to have progressive fibrosis without undergoing repeat liver biopsy. However, of the 12 patients developing ESLD, 10 patients had undergone liver biopsy at baseline and in 9 patients, cirrhosis was absent implying that most ESLD patients also had progressed in fibrosis stage during follow-up. Moreover, in a subgroup analysis including those patients who had undergone liver biopsy at both baseline and follow-up (n = 74), only change in dAAR value significantly correlated with progression in the fibrosis stage. The correlation coefficient was very low (0.25) further corroborating the limited clinical value of the evaluated NITs to predict fibrosis progression.

The optimal combination or sequential use of imaging and blood-based fibrosis biomarkers in the follow-up of NAFLD patients needs
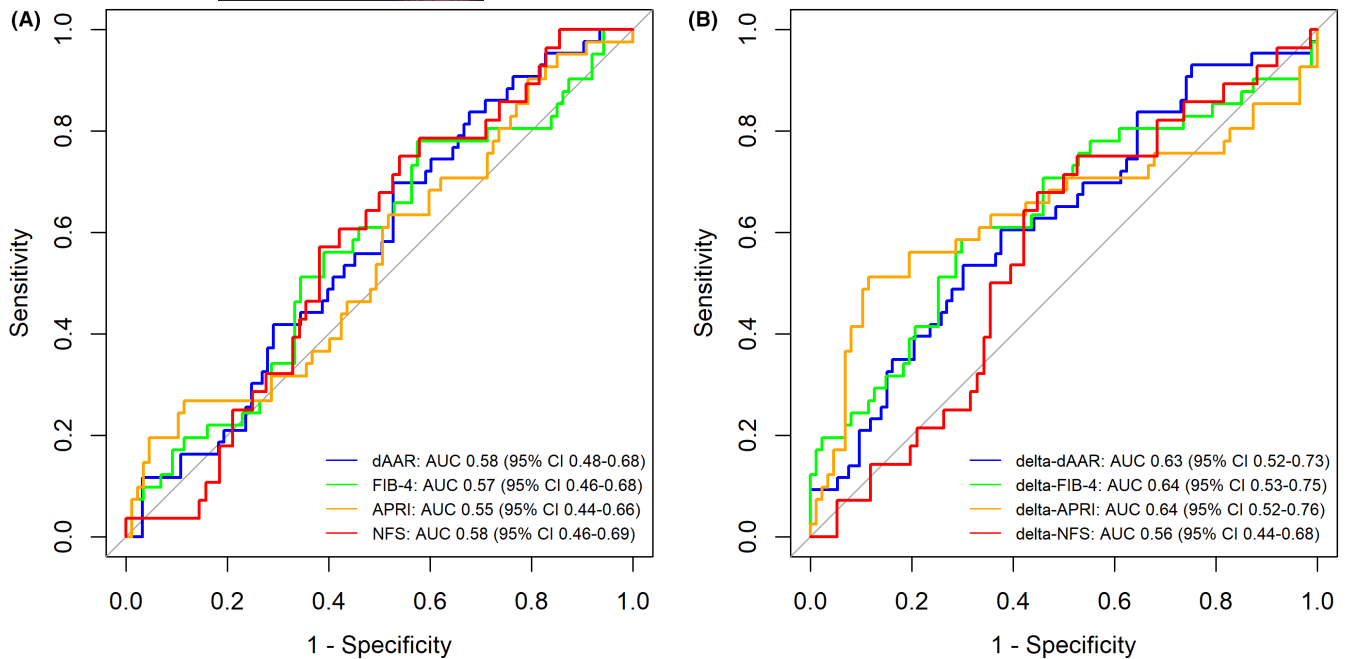
**FIGURE 3** Receiver-operating characteristic (ROC) curves for (A) baseline serum fibrosis algorithms values to predict progression of NAFLD (B) delta values between baseline and follow-up of serum fibrosis algorithms to predict progression of NAFLD. NAFLD, non-alcoholic fatty liver disease

**TABLE 4** Diagnostic performance of delta values of NIT scores at 90% sensitivity and 90% specificity

|  | Cut-off | Sensitivity | Specificity | Accuracy | PPV | NPV |
|---|---|---|---|---|---|---|
| Sensitivity fixed at ≥90%[a] | | | | | | |
| Delta-dAAR | −0.548 | 0.907 | 0.258 | 0.463 | 0.361 | 0.857 |
| Delta-NFS | −0.215 | 0.928 | 0.118 | 0.337 | 0.280 | 0.818 |
| Delta-FIB-4 | −0.197 | 0.902 | 0.126 | 0.375 | 0.327 | 0.733 |
| Delta-APRI | −0.470 | 0.902 | 0.035 | 0.313 | 0.306 | 0.429 |
| Specificity fixed at ≥90%[b] | | | | | | |
| Delta-dAAR | 1.423 | 0.209 | 0.903 | 0.684 | 0.500 | 0.712 |
| Delta-NFS | 2.903 | 0.071 | 0.934 | 0.702 | 0.286 | 0.732 |
| Delta-FIB-4 | 1.517 | 0.244 | 0.920 | 0.703 | 0.588 | 0.721 |
| Delta-APRI | 0.178 | 0.415 | 0.920 | 0.758 | 0.708 | 0.769 |

Abbreviations: APRI, AST to platelet ratio index; dAAR, dynamic aspartate-to-alanine aminotransferase ratio; FIB-4, fibrosis-4; NFS, NAFLD fibrosis score; NIT, non-invasive test; NPV, negative predictive value; PPV, positive predictive value.

[a]Fixed at sensitivity closest to 90% but ≥90%.

[b]Fixed at specificity closest to 90% but ≥90%.

to be defined. In the recently published EASL clinical practice guidelines,[16] it is recommended that repeated measurements of NITs can be used every 3 years in NAFLD patients with early stage, and every year in patients with advanced-stage NAFLD, to refine stratification of risk of liver-related events. However, our study does not support repeated measurement of FIB-4, NFS, APRI or dAAR for this purpose.

All four evaluated NITs in this study are based on indirect fibrosis markers which mirror common functional alterations in the liver, alterations that do not necessarily reflect extracellular matrix turnover and/or fibrogenic cell changes. A better understanding of

the pathophysiology of liver fibrosis has prompted investigators to use more refined markers to identify different fibrosis stages. These so-called direct serum markers, for example the ELF test,[32] may have better diagnostic precision to predict fibrosis progression in the clinical setting, although this has not been proven. Unfortunately, direct fibrosis markers were unavailable in our patient cohort.

In conclusion, change of FIB-4, NFS, APRI and dAAR scores is only weakly associated with disease progression in NAFLD. Our findings do not support repeated measurements of these NITs for monitoring the course of NAFLD in the clinical setting. There is a

need to evaluate these NITs and direct fibrosis markers in larger prospective follow-up studies.

## CONFLICT OF INTEREST

No disclosures to report pertaining to this work. The authors confirm that there is no known conflict of interest associated with this publication.

## ETHICS APPROVAL AND PATIENT CONSENT STATEMENT

All participating subjects gave written informed consent. The study design was approved by the ethics committee at the University Hospital in Linköping (02-454, amendments: 2011/468-32, 2012/229-32 and 2013/72-32).

## PERMISSION TO REPRODUCE MATERIAL FROM OTHER SOURCES

No material was reproduced from other sources.

## ORCID

*Wile Balkhed* https://orcid.org/0000-0003-4223-8968
*Fredrik O. Åberg* https://orcid.org/0000-0002-3833-0705
*Patrik Nasr* https://orcid.org/0000-0002-2928-4188
*Mattias Ekstedt* https://orcid.org/0000-0002-5590-8601
*Stergios Kechagias* https://orcid.org/0000-0001-7614-739X

## REFERENCES

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease: meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology.* 2016;64:73-84.
2. Nasr P, Ignatova S, Kechagias S, Ekstedt M. Natural history of non-alcoholic fatty liver disease: a prospective follow-up study with serial biopsies. *Hepatol Commun.* 2018;2:199-210.
3. Setiawan VW, Stram DO, Porcel J, Lu SC, Marchand LL, Noureddin M. Prevalence of chronic liver disease and cirrhosis by underlying cause in understudied ethnic groups: the multiethnic cohort. *Hepatology.* 2016;64:1969-1977.
4. Dyson J, Jaques B, Chattopadyhay D, et al. Hepatocellular cancer: the impact of obesity, type 2 diabetes and a multidisciplinary team. *J Hepatol.* 2014;60:110-117.
5. Goldberg D, Ditah C, Saeian K, et al. Changes in the prevalence of hepatitis C virus infection, nonalcoholic steatohepatitis, and alcoholic liver disease among patients with cirrhosis or liver failure on the waitlist for liver transplantation. *Gastroenterology.* 2017;152:1090-1099.
6. Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology.* 1999;116:1413-1419.
7. Sanyal AJ. AGA technical review on nonalcoholic fatty liver disease. *Gastroenterology.* 2002;123:1705-1725.
8. Ekstedt M, Hagström H, Nasr P, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology.* 2015;61:1547-1554.
9. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology.* 2015;149:389-397.
10. Hagström H, Nasr P, Ekstedt M, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. *J Hepatol.* 2017;67:1265-1273.
11. Lykiardopoulos B, Hagström H, Fredrikson M, et al. Development of serum marker models to increase diagnostic accuracy of advanced fibrosis in nonalcoholic fatty liver disease: the New LINKI algorithm compared with established algorithms. *PLoS One.* 2016;11:e0167776.
12. Boursier J, Tsochatzis EA. Case-finding strategies in non-alcoholic fatty liver disease. *JHEP Rep.* 2020;5(3):100219.
13. Wong VW, Adams LA, de Lédinghen V, Wong GL, Sookoian S. Noninvasive biomarkers in NAFLD and NASH - current progress and future promise. *Nat Rev Gastroenterol Hepatol.* 2018;15:461-478.
14. Castera L, Friedrich-Rust M, Loomba R. Noninvasive assessment of liver disease in patients with nonalcoholic fatty liver disease. *Gastroenterology.* 2019;156:1264-1281.e4.
15. 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. *J Hepatol.* 2016;64:1388-1402.
16. European Association for the Study of the Liver (EASL). European Association for the Study of the Liver EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. *J Hepatol.* 2021;75:659-689.
17. Åberg F, Danford CJ, Thiele M, et al. A dynamic aspartate-to-alanine aminotransferase ratio provides valid predictions of incident severe liver disease. *Hepatol commun.* 2021;5:1021-1035.
18. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology.* 2006;43:1317-1325.
19. 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.
20. Wai CT, Greenson 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.
21. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* 2005;41:1313-1321.
22. Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Biol.* 2003;29:1705-1713.
23. Papatheodoridi M, Hiriart JB, Lupsor-Platon M, et al. Refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease. *J Hepatol.* 2021;74:1109-1116.
24. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. *J Hepatol.* 2008;48:835-847.
25. McPherson S, Hardy T, Henderson E, Burt AD, Day CP, Anstee QM. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: implications for prognosis and clinical management. *J Hepatol.* 2015;62:1148-1155.
26. Siddiqui MS, Yamada G, Vuppalanchi R, et al. Diagnostic accuracy of noninvasive fibrosis models to detect change in fibrosis stage. *Clin Gastroenterol Hepatol.* 2019;17:1877-1885.
27. Hagström H, Talbäck M, Andreasson A, Walldius G, Hammar N. Repeated FIB-4 measurements can help identify individuals at risk of severe liver disease. *J Hepatol.* 2020;73:1023-1029.
28. Younossi ZM, Anstee QM, Wai-Sun Wong V, et al. The association of histologic and noninvasive tests with adverse clinical and patient-reported outcomes in patients with advanced fibrosis due to nonalcoholic steatohepatitis. *Gastroenterology.* 2021;160:1608-1619.
29. Hagström H, Talbäck M, Andreasson A, Walldius G, Hammar N. Ability of noninvasive scoring systems to identify individuals in

the population at risk for severe liver disease. *Gastroenterology.* 2020;158:200-214.

30. Dulai PS, Singh S, Patel J, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. *Hepatology.* 2017;65:1557-1565.

31. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology.* 2005;128:1898-1906. doi:10.1053/j.gastro.2005.03.084

32. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology.* 2004;127:1704-1713.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.