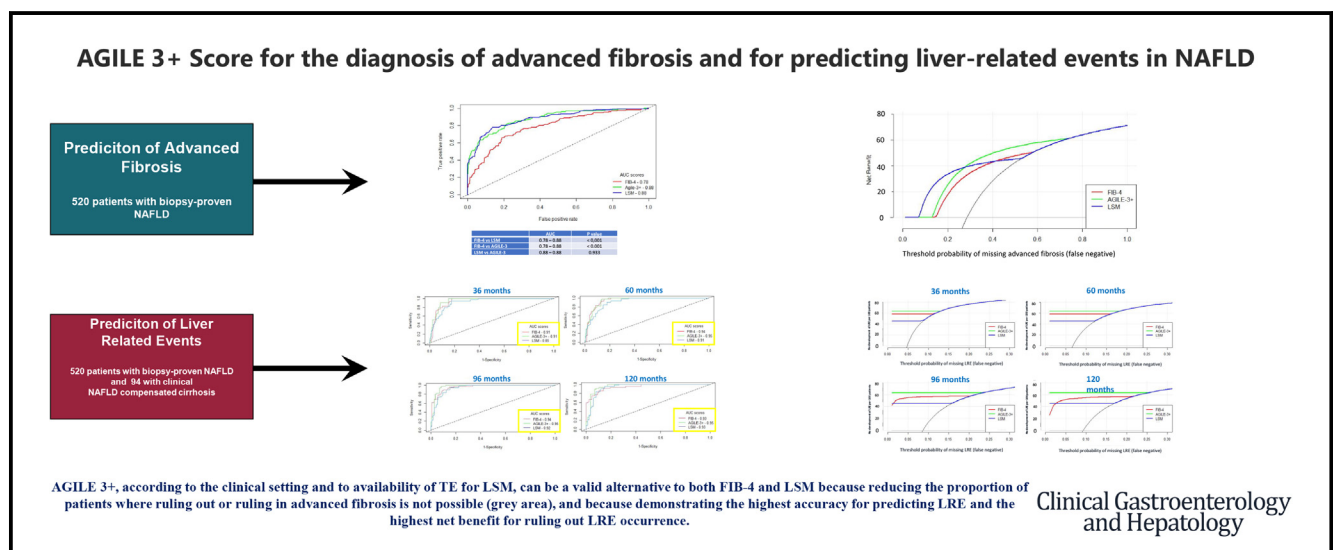


AGILE 3+ Score for the Diagnosis of Advanced Fibrosis and for Predicting Liver-related Events in NAFLD

Grazia Pennisi,^{1,*} Marco Enea,^{1,*} Alessandra Pandolfo,¹ Ciro Celsa,¹ Michela Antonucci,² Carlo Ciccioli,¹ Giuseppe Infantino,¹ Claudia La Mantia,¹ Stefanie Parisi,¹ Adele Tulone,¹ Vito Di Marco,¹ Antonio Craxì,¹ Calogero Cammà,¹ and Salvatore Petta¹

¹Sezione di Gastroenterologia, PROMISE, University of Palermo, Italy; and ²Dipartimento di Biomedicina, Neuroscienze e Diagnostica avanzata (BIND), University of Palermo, Palermo, Italy



BACKGROUND & AIMS:

We aimed to assess the diagnostic accuracy of AGILE 3+, a recently developed score based on the combination of aspartate aminotransferase/alanine aminotransferase ratio, platelet count, diabetes status, sex, age, and liver stiffness measurement (LSM) by transient elastography, when compared with Fibrosis-4 (FIB-4) and LSM, for the diagnosis of advanced fibrosis and for the prediction of liver-related events (LREs) occurrence in patients with NAFLD.

METHODS:

A total of 614 consecutive patients with biopsy-proven NAFLD or clinical diagnosis of NAFLD-related compensated cirrhosis were enrolled. LREs were recorded during follow-up. FIB-4, LSM by transient elastography (FibroScan device), and AGILE 3+ were measured. The diagnostic performance of noninvasive criteria for advanced fibrosis and for the prediction of LREs was assessed using the area under the receiver operating characteristic curve (AUROC) and decision curve analysis.

RESULTS:

In patients with biopsy-proven NAFLD (n = 520), LSM and AGILE 3+ had higher AUROC than FIB-4 (0.88 for LSM and AGILE 3+ vs 0.78 for FIB-4; $P < .001$) for advanced fibrosis, and AGILE 3+ exhibited a smaller indeterminate area in the test (25.2% for FIB-4 vs 13.1% for LSM vs

*Authors share co-first authorship.

Abbreviations used in this paper: ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic curve; DCA, decision curve analysis; FIB-4, Fibrosis-4; HCC, hepatocellular carcinoma; IQR, interquartile range; LD, liver decompensation; LRE, liver-related event; LSM, liver stiffness measurement; M, median; NAFLD, nonalcoholic fatty liver disease; NPV, negative predictive

value; PLT, platelets; PPV, positive predictive value; ROC, receiver operating characteristic; TE, transient elastography.

8.3% for AGILE 3+). Within the entire cohort of patients, AGILE 3+ had significantly higher AUROC for predicting LREs with respect to LSM (AUROC 36 months 0.95 vs 0.93; $P = .008$; 60 months 0.95 vs 0.92; $P = .006$; 96 months 0.97 vs 0.95; $P = .001$). Decision curve analysis showed that all scores had modest net benefit for ruling-out advanced fibrosis at the risk threshold of 5% to 10% where advanced fibrosis was absent. At the risk threshold of 5% of false negatives or false positives in LRE at 36, 60, 96, and 120 months, AGILE 3+ outperformed both FIB-4 and LSM for ruling out LRE.

CONCLUSIONS:

Depending on resource availability, clinical setting, and the risk scenarios, AGILE 3+ is an accurate and valid alternative to FIB-4 and LSM for the noninvasive assessment of disease severity and prognosis in patients with NAFLD.

Keywords: AGILE 3+; DCA; Fibrosis; NAFLD; NASH.

Nonalcoholic fatty liver disease (NAFLD), affecting roughly 25% of the general adult population¹ and paralleling the pandemic of obesity and diabetes, is the leading cause of chronic liver disease² and its complications like liver decompensation (LD)³ and hepatocellular carcinoma (HCC),⁴ as well as being an increasing indication for liver transplantation.⁵ NAFLD also increases the risk of extrahepatic events such as cardiovascular events⁶ and extrahepatic cancers.⁷ In this emerging scenario, natural history studies and meta-analyses have thoroughly demonstrated that the severity of liver fibrosis is the main determinant of prognosis in patients with NAFLD, due to the fact that a higher degree of liver fibrosis is indicative of a greater risk of developing liver-related (LREs) and extrahepatic events.^{8,9} Despite the relevant clinical burden of NAFLD, only a minority of patients develop a progressive disease leading to a higher risk of complications. Consequently, the noninvasive identification of patients with advanced fibrosis and the prediction of development of LREs are both highly useful but unfulfilled demands in the research agenda for NAFLD.

Easy-to-obtain and widely available noninvasive scores like Fibrosis-4 (FIB-4) and liver stiffness measurement (LSM) by transient elastography (TE) have acceptable accuracy¹⁰⁻¹³ for excluding advanced fibrosis, and their use alone, and in combination with each other, is suggested by clinical guidelines in patients with NAFLD or at high risk of NAFLD-related liver damage due to obesity and/or diabetes and/or metabolic syndrome.^{14,15} Available longitudinal cohort studies also suggest the utility of these scores/tools to stratify the risk of developing LRE.¹⁶⁻¹⁹ A recently developed and proposed score, namely AGILE 3+, based on the simultaneous combination of aspartate aminotransferase (AST)/alanine aminotransferase (ALT) ratio, platelet count, diabetes status, sex, age, and LSM, has shown a slight improvement in diagnostic ability in comparison with LSM and FIB-4 for the diagnosis of advanced fibrosis in NAFLD.²⁰

We aimed to assess the diagnostic accuracy of AGILE 3+ in comparison with FIB-4 and LSM in the diagnosis of advanced fibrosis and for the prediction of LRE event occurrence in patients with NAFLD.

Methods

Patients

We retrospectively analyzed data from 614 patients prospectively recruited at the Gastrointestinal and Liver Unit of the Palermo University Hospital and with histological diagnosis of NAFLD or clinical diagnosis of Child Pugh A5 cirrhosis related to NAFLD and without previous history of LD, portal thrombosis, esophageal varices band ligation, and HCC. Specifically, in patients without histology, cirrhosis was diagnosed by LSM >11.5 kPa for M probe²¹ or >11 kPa for XL probe,²² and the diagnosis of NAFLD required the presence of ultrasonography-assessed steatosis plus at least one criterion of the metabolic syndrome (obesity, diabetes, arterial hypertension, dyslipidemia).

Only patients with a follow-up of at least 6 months were included. Other causes of liver disease were ruled out, including alcohol intake (>30 g/day in men, >20 g/day in women) as evaluated by a questionnaire, viral causes (hepatitis B surface antigen, anti-hepatitis C virus, and anti-human immunodeficiency virus negativity), autoimmune hepatitis, primary biliary cholangitis, primary sclerosing cholangitis, hereditary hemochromatosis, and alpha-1 antitrypsin deficiency. Patients who took part in randomized controlled trials with new nonalcoholic steatohepatitis drugs during the follow-up were also excluded.

The study was carried out in accordance with the principles of the Helsinki Declaration and in accordance with local and national laws. Approval was obtained by the Ethical Committee of the University Hospital "Paolo Giaccone" in Palermo.

Clinical, Laboratory Assessment, and Histology

Clinical and anthropometric data, including body mass index and type 2 diabetes, were collected at the time of enrollment. A diagnosis of type 2 diabetes was made according to standards set by the American Diabetes Association,²³ using a value of fasting blood glucose ≥ 126 mg/dL and/or use of medication.

The same day as the liver biopsy, a 12-hour overnight fasting blood sample was drawn to determine serum levels of AST, ALT, platelets (PLT), albumin, total and high-density lipoprotein cholesterol, and triglycerides.

A liver-dedicated expert pathologist, who was unaware of patients' identity and history, coded and read histological slides. A minimum 15-mm length of the biopsy specimen or the presence of at least 10 complete portal tracts was required. Steatosis was assessed as the percentage of hepatocytes containing fat droplets (minimum 5%), and as a categorical variable. Kleiner classification²⁴ was used to stage fibrosis from 0 to 4.

Noninvasive Fibrosis Algorithms/Tools

The Fibrosis-4 (FIB-4) (age, AST, ALT, PLT) was calculated using the original reported formula, and patients were classified as low risk of advanced fibrosis if <1.30 , intermediate risk if 1.30 to 2.67 , and high risk if >2.67 .²⁵

Transient elastography (TE) was performed with FibroScan (Echosens, Paris, France), using the M and the XL probes when appropriate. In each center, LSM was assessed on the same day as the liver biopsy, before the procedure and after an overnight fast, by a trained operator who had previously performed at least 300 determinations in patients with chronic liver disease. As recently reported in literature,²⁶ we classified all LSM examinations into 3 reliability categories: "very reliable" (interquartile range [IQR]/median (M) ≤ 0.10), "reliable" ($0.10 < \text{IQR}/\text{M} \leq 0.30$, or $\text{IQR}/\text{M} > 0.30$ with $\text{LSM M} < 7.1$ kPa), and "poorly reliable" ($\text{IQR}/\text{M} > 0.30$ with $\text{LSM M} \geq 7.1$ kPa). Only patients with 10 valid measurements were included, and "poorly reliable" results were excluded from the analysis. LSM was assessed after overnight fasting. $\text{LSM} < 7.9$ kPa was defined as indicating a low risk of F3 to F4 fibrosis; $\text{LSM } 7.9$ to 9.6 kPa as an intermediate risk; and $\text{LSM} \geq 9.6$ kPa as a high risk.²¹

The AGILE 3+ score (age, sex, AST, ALT, PLT, diabetes status, and LSM) was calculated using the original reported formula, and patients were classified as low risk of advanced fibrosis if <0.45 , intermediate risk if 0.45 to 0.68 , and high risk if >0.68 .²⁰

Liver-related Events

LREs were recorded during the entire follow-up, and they were defined as development of LD (occurrence of ascites and/or bleeding varices and/or encephalopathy and/or jaundice) or of HCC.

Ultrasound examination for HCC surveillance was carried out annually in patients with F0 to F2 fibrosis, and every 6 months in patients with F3 fibrosis or cirrhosis, in accordance with international guidelines.²⁷

In patients with cirrhosis, upper gastrointestinal endoscopy was performed at baseline and repeated as

What You Need to Know

Background

Noninvasive scores/tools like Fibrosis-4 (FIB-4) and liver stiffness measurement (LSM) have acceptable accuracy for ruling out advanced fibrosis and seem to sufficiently predict the occurrence of liver-related events (LREs). A recently developed score, namely AGILE 3+, which is based on the simultaneous combination of aspartate aminotransferase/alanine aminotransferase ratio, platelet count, diabetes status, sex, age, and LSM, has been developed and proposed for the diagnosis of advanced fibrosis in nonalcoholic fatty liver disease (NAFLD).

Findings

In patients with NAFLD, AGILE 3+ could be preferred to LSM and FIB-4 because it reduces the proportion of patients where ruling out or ruling in of advanced fibrosis is not possible (gray area), and because it has demonstrated the highest accuracy for predicting LREs and the highest net benefit for ruling out LRE occurrence.

Implications for Patient Care

Depending on clinical setting (general practitioners, diabetes care centers, hepatology referral centers), resource status, and availability of transient elastography for LSM, AGILE 3+ can be a valid alternative to both FIB-4 and LSM for assessing NAFLD severity and risk of liver-related complications.

recommended by clinical guidelines. Patients with progression to medium or large (F2 or F3) esophageal varices were treated with β -blockers or underwent elastic banding, whereas no prophylaxis was scheduled for patients with small (F1) varices.

Patients found to be developing hepatic events during follow-up were evaluated for available therapies and/or for liver transplantation, as appropriate.^{27,28}

Statistics

Data for continuous variables was expressed as mean and standard deviation or median and IQRs, and data for categorical variables were expressed as frequency and percentage. Differences between continuous data were assessed by the Student *t* test or by the Mann Whitney *U* test. Differences between categorical variables were assessed by the χ^2 test.

The accuracy of each score for detection of advanced fibrosis (F3–F4) was assessed in the subgroup of patients with histological diagnosis of NAFLD using receiver operating characteristic (ROC) curves described as the area under the ROC (AUROC) with 95% confidence intervals (95% CIs). Cutoff points of LSM, NAFLD fibrosis score, and FIB-4 for the advanced fibrosis model were derived from literature. Specifically, for LSM, cutoffs of

<7.9 kPa and of ≥ 9.6 kPa were used to rule-out and rule-in, respectively, severe fibrosis²¹; for FIB-4, cutoffs of <1.30 and of >2.67 were used to rule-out and rule-in, respectively, severe fibrosis²⁵; for AGILE 3+, cutoffs of <0.45 and of >0.68 were used to rule-out and rule-in, respectively, severe fibrosis.²⁰ Accordingly, false-negative and false-positive rates of the single test, and of their combination, as well as sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. The accuracy of the same scores and of the same cutoffs for detection of occurrence of LREs was assessed in the entire cohort of patients with biopsy-proven NAFLD or with clinical diagnosis of compensated NAFLD cirrhosis using time-dependent ROC curves described as AUROC with 95% CIs. Kaplan-Meier curves were also used to depict the time-dependent risk of developing LREs according to noninvasive score cutoffs.

As AUROC focuses only on the predictive accuracy of a model, not taking into account the cases where a false-negative result is more harmful than a false-positive result, we also performed a decision curve analysis (DCA) for identifying the threshold probabilities where use of noninvasive criteria translate into maximum net benefit of detecting advanced fibrosis or LREs.^{29,30}

DCA-evaluated prediction models in comparison with default strategies of performing liver biopsy in all patients, or none, or performing clinical-instrumental follow-up for detecting LREs in all patients, or none, allowing for an assessment of overall yield of prediction rules. DCA estimates a “net benefit” for each prediction rule, defined as

$$\text{net benefit} = \text{sensitivity} \times \text{prevalence} - (1 - \text{specificity}) \times (1 - \text{prevalence}) \times w$$

where w is the odds of correct diagnosis (ie, advanced fibrosis or LRE in this case) across different threshold probabilities. In this setting, net benefit represents a composite of the benefit gained by performing liver biopsy for proven advanced fibrosis or follow-up with clinical instruments for LRE in patients classified as high-risk according to noninvasive scores (true positive) and risk/discomfort incurred due to liver biopsy or follow-up with clinical instruments in those without advanced fibrosis or those who will not develop LRE but who were classified as high-risk according to noninvasive scores (false positive). Threshold probability represents a theoretical risk-level where the expected benefit of treatment is equal to the expected benefit of avoiding treatment (eg, benefit of liver biopsy or of follow-up with clinical instruments equals risk of not performing them). Thus, net benefit is assessed across a range of threshold probabilities to identify the best diagnostic strategy for different risk scenarios.

All data was analyzed using R studio. DCA was implemented in R using code derived from Zhang et al.³¹

In addition to the base packages in R, tidy verse, survival, survminer, boot, reshape2, and readxl packages were used.

Results

Patient Features and Outcomes

Baseline characteristics of the 614 patients with NAFLD stratified for advanced fibrosis are shown in [Supplementary Table 1](#). As expected, when comparing patients with advanced fibrosis with those without, the former tended to be older and had categorically higher mean body mass index, higher prevalence of diabetes, lower PLT and serum albumin values, and higher mean values of FIB-4, LSM, and AGILE 3+ ($P < .01$ for all comparisons).

The diagnosis of NAFLD was supported by histology in 520 cases (84.7%), whereas 94 patients (15.3%) had clinical diagnosis of compensated NAFLD cirrhosis. Specifically, 80 of the advanced fibrosis/cirrhosis patients had a liver stiffness measurement >15 kPa,³² and 14 were between 11 kPa and 15 kPa (11 of these cases also had esophageal varices, and the other 3 had signs of portal hypertension: splenomegaly and portal vein ectasia). [Supplementary Table 2](#) shows the characteristics of patients with cirrhosis by clinical criteria compared with patients with biopsy-proven NAFLD.

Diagnostic Accuracy of Noninvasive Scores/Tools for Advanced Fibrosis

The assessment of the accuracy of noninvasive scores/tools for the diagnosis of advanced fibrosis was done in the 520 patients with histological diagnosis of NAFLD.

[Table 1](#) reports the distribution of patients according to low, intermediate, and high-risk cutoffs of advanced fibrosis for each score/tool. Proportions of patients classified as having low, intermediate, and high risk of advanced fibrosis were 67.9%, 25.2%, and 6.9%, respectively, using FIB-4; 6.9%, 52.9%, and 13.1%, respectively, using LSM; and 74.4%, 8.3%, and 17.3%, respectively, using AGILE 3+.

[Figure 1](#) and [Table 1](#) show the AUROCs of the scores for the diagnosis of advanced fibrosis. Specifically, all the scores had a good accuracy in terms of AUROC (FIB-4, 0.78; LSM, 0.88; AGILE 3+, 0.88), and both LSM and AGILE 3+ performed better than FIB-4 ($P < .001$ for both comparisons). AGILE 3+ had the highest sensitivity (91%), FIB-4 the highest specificity (98%), and LSM the highest PPV (90%) and the highest NPV (93%) ([Table 1](#)).

Net benefit of FIB-4, LSM, and AGILE 3+ scores for ruling out advanced fibrosis at 5%, 10%, and 25% threshold probabilities of missing advanced fibrosis is displayed in the upper panel of [Table 2](#). At the risk threshold of 5% of missing advanced fibrosis, all scores/

Table 1. Accuracy of FIB-4, LSM, and AGILE-3 for the Diagnosis of Advanced Liver Fibrosis in Patients With NAFLD From the Training Cohort

	Proportion of cases	AUROC	Sensitivity	Specificity	PPV	NPV
FIB-4 <1.30	67.9	0.78	0.81 (0.77–0.85)	0.65 (0.57–0.73)	0.58 (0.50–0.66)	0.85 (0.81–0.89)
FIB-4 1.30–2.67	25.2		0.20 (0.14–0.27)	0.98 (0.97–0.99)	0.83 (0.67–0.94)	0.75 (0.71–0.79)
FIB-4 >2.67	6.9					
LSM <8	52.9	0.88	0.87 (0.80–0.92)	0.69 (0.64–0.73)	0.53 (0.46–0.59)	0.93 (0.89–0.96)
LSM 8–9.6	13.1		0.78 (0.70–0.84)	0.84 (0.79–0.87)	0.90 (0.87–0.93)	0.66 (0.58–0.73)
LSM >9.6	34.0					
AGILE 3+ <0.48	74.4	0.88	0.91 (0.87–0.93)	0.66 (0.58–0.73)	0.74 (0.65–0.81)	0.87 (0.83–0.90)
AGILE 3+ 0.48–0.65	8.3		0.51 (0.43–0.59)	0.96 (0.94–0.98)	0.84 (0.75–0.91)	0.83 (0.79–0.86)
AGILE 3+ >0.65	17.3					

AUROC, Area under the receiver operating curve; FIB-4, Fibrosis-4; LSM, liver stiffness measurement; NAFLD, nonalcoholic fatty liver disease; NPV, negative predictive value; PPV, positive predictive value.

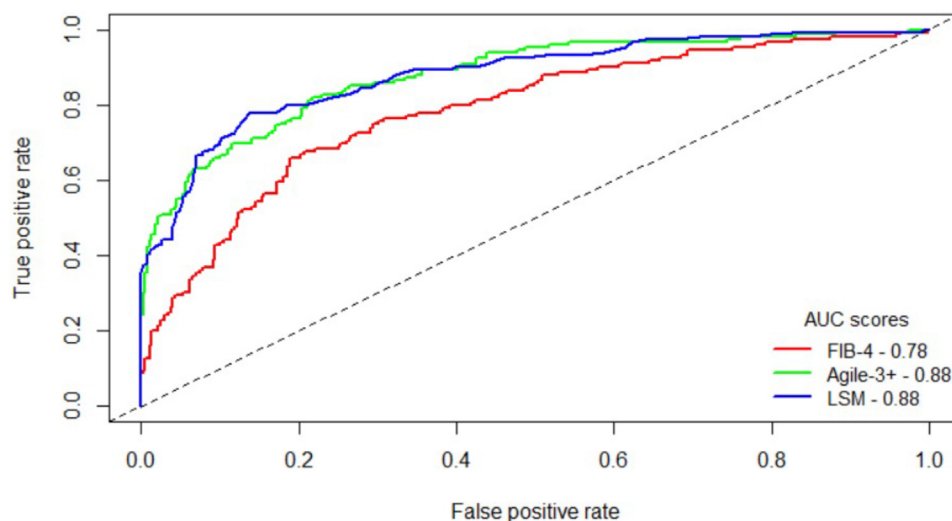
tools showed no benefit for ruling out advanced fibrosis compared with the strategy of performing liver biopsy in all patients. At the risk thresholds of 10% and 25%, the observed net benefit was modest, and LSM outperformed both FIB-4 and AGILE 3+. Overall, net benefit for ruling out advanced fibrosis was higher for LSM compared with both FIB-4 and AGILE 3+ at threshold probabilities between 5% and 25%, as reported in [Figure 2](#), showing the number of liver biopsies avoided per 100 patients through a range of different threshold probabilities of missing advanced fibrosis. Results obtained for ruling in advanced fibrosis are shown in [Supplementary Figure 1](#) and the upper panel of [Supplementary Table 3](#).

When also considering the strategy suggested in the EASL guideline of using FIB-4 as a first test, pertaining to

LSM-only patients at intermediate or high risk of advanced fibrosis by FIB-4, we showed that net benefit for ruling out advanced fibrosis was higher for both LSM and AGILE 3+ compared with the sequential FIB-4 to LSM strategy at threshold probabilities between 5% and 25% ([Supplementary Figure 2](#)).

Diagnostic Accuracy of Noninvasive Scores/Tools for Predicting the Occurrence of Liver-related Events

The assessment of the accuracy of noninvasive scores/tools for predicting the occurrence of LREs was applied to the entire cohort of 614 patients with

**Figure 1.** ROC curves for prediction of F3 to F4 fibrosis using FIB-4, LSM, and AGILE 3+.

	AUC	P value
FIB-4 vs LSM	0.78 – 0.88	< .001
FIB-4 vs AGILE-3	0.78 – 0.88	< .001
LSM vs AGILE-3	0.88 – 0.88	.933

Table 2. Net Benefit by DCA for Ruling Out Advanced Liver Fibrosis and Development of LREs

Outcome: advanced fibrosis		Number of liver biopsy avoided per 100 patients (net benefit)		
		FIB-4	LSM	AGILE 3+
Threshold probabilities of missing advanced fibrosis	5%	0	0	0
	10%	0	14	0
	25%	28	37	35

Outcome: LREs		Number of follow-up visits avoided per 100 patients (net benefit)			
		Time, mos	FIB-4	LSM	AGILE-3
Threshold probabilities of missing LREs	5%	36	58	45	63
		60	58	45	63
		96	55	45	63
		120	52	45	63
	20%	36	76	76	76
		60	66	67	67
		96	57	57	63
		120	57	54	63

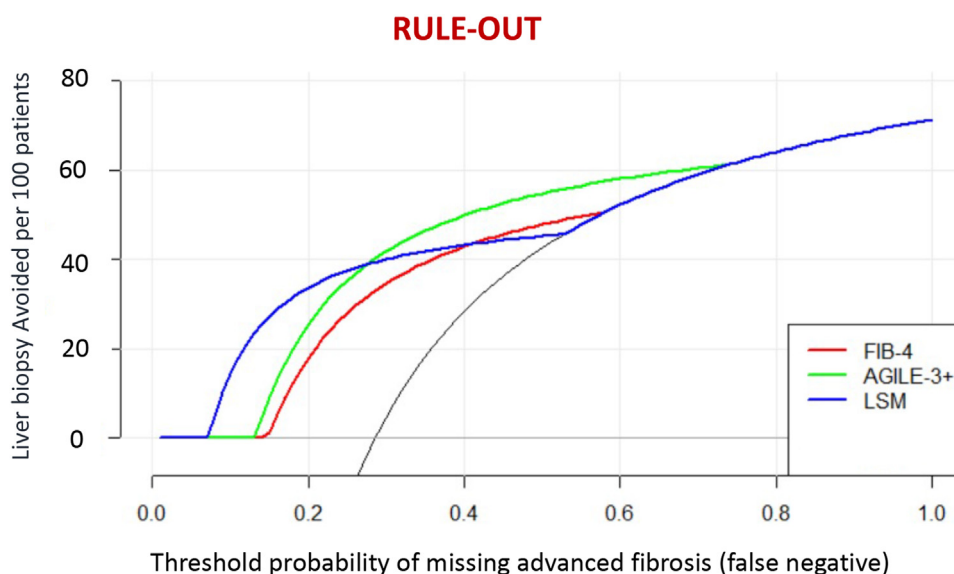
DCA, Decision curve analysis; FIB-4, Fibrosis-4; LSM, liver stiffness measurement; LRE, liver-related event.

histological diagnosis of NAFLD or with clinical diagnosis of compensated NAFLD-related cirrhosis.

During a median follow-up of 63.9 months, 57 patients developed LRE (44 LD and 13 HCC). Only 0.6%, 0%, and 0% of patients at low risk of advanced fibrosis by FIB-4, LSM, and AGILE 3+ respectively, developed LRE, compared with 5.7%, 1.4%, and 2.2% of those at intermediate risk, and to 46%, 21.1%, and 30.9% of those at high risk. Kaplan-Meier curves depicting the time-dependent probability of not developing LREs across the different risk classes of each score are reported in [Supplementary Figure 3](#). The 36-, 60-, 96-, and 120-month AUROC of the scores were good (>0.90 for all) and are reported in [Figure 3](#). Specifically, comparison

of AUROCs showed a significantly better performance of AGILE 3+ when compared with LSM at 36 (AUC, 0.95 vs 0.93; $P = .008$), 60 (AUC, 0.95 vs 0.92; $P = .006$), and 96 months (AUC, 0.97 vs 0.95; $P = .001$).

Net benefit of FIB-4, LSM, and AGILE 3+ scores for ruling out occurrence of LRE at 5% and 20% threshold probabilities of missing LRE is displayed in the bottom panel of [Table 2](#). At the risk threshold of 5% of missing occurrence of LRE at 36, 60, 96, and 120 months, AGILE 3+ outperformed both FIB-4 and LSM; at the risk thresholds of 20% of missing occurrence of LRE, all scores had similar net benefit at 36 and 60 months, whereas AGILE 3+ again outperformed both FIB-4 and LSM at 96 and 120 months. Overall, AGILE 3+ exhibited

**Figure 2.** Net benefits by DCAs of FIB-4, LSM, and AGILE 3+ for ruling out advanced liver fibrosis.

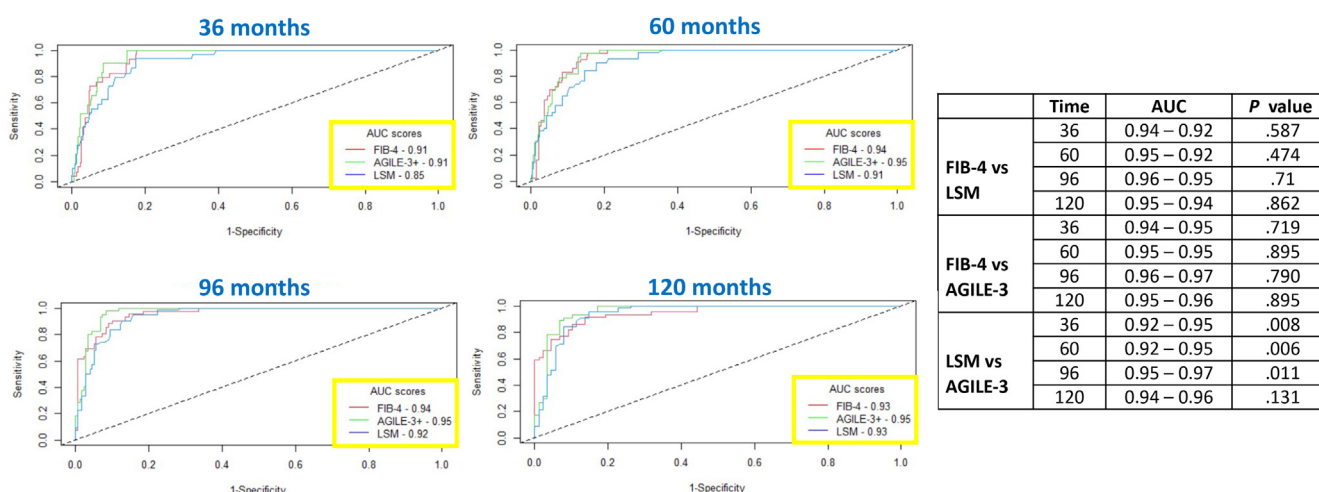


Figure 3. ROC curves for prediction of LREs using FIB-4, LSM, and AGILE 3+.

a higher net benefit for ruling out occurrence of LRE compared with both FIB-4 and LSM at a threshold probability between 5% and 20%, as displayed in Figure 4, A, reporting the number of no development of LREs per 100 patients through a range from 0% to 30% of threshold probabilities of missing occurrence of LREs. Results obtained for ruling in occurrence of LREs are shown in Figure 4, B and in the bottom panel of Supplementary Table 3.

Furthermore, when considering the EASL-driven sequential strategy of using FIB-4 as a first method of testing and then LSM in patients at intermediate or high risk of advanced fibrosis by FIB-4, we observed that at 5% and 20% threshold probabilities of missing occurrence of LRE at 36 and 60 months, the sequential strategy had a slightly higher net benefit than AGILE 3+, while a similar net benefit was found at 96 and 120 months (Supplementary Figure 4).

In a subgroup analysis performed on the 520 patients with histological diagnosis of NAFLD, we observed similar results in terms of both AUROCs and DCA (Supplementary Figure 5, A–B). Moreover, when comparing results in this sub-set FIB-4, LSM, and AGILE 3+ to fibrosis stage by histology, noninvasive scores were not inferior (Supplementary Figure 5).

Discussion

In this study carried out on a large cohort of patients with histological diagnosis of NAFLD or with clinical diagnosis of compensated NAFLD cirrhosis, we provided evidence that AGILE 3+ can be a useful alternative to FIB-4 and LSM because: (1) it leads to a lower proportion of patients in the uncertainty area, and (2) is characterized by higher accuracy compared with FIB-4 for the diagnosis of advanced fibrosis and compared with LSM for predicting occurrence of LREs. Furthermore, DCA showed that the net benefit for the ruling out of advanced fibrosis was modest for all tools, whereas the

net benefit for ruling out the occurrence of LRE clearly showed that AGILE 3+ was the best tool.

FIB-4 and LSM are the most readily available, used, and validated noninvasive scores/tools that, used alone or in combination according to clinical context and availability, can help in identifying patients with NAFLD with advanced fibrosis (ie, at risk of hepatic and extra-hepatic complications)^{10–13} and/or at risk of developing LRE.^{14–19} In this emerging clinical setting, a new score named AGILE 3+, and based on the simultaneous combination of clinical-biochemical variables with LSM, has been proposed for the noninvasive assessment of advanced liver fibrosis.²⁰

In the present study, we found that AGILE 3+, together with LSM, had the highest accuracy for the noninvasive diagnosis of advanced fibrosis, and, at the price of a NPV slightly lower than that of LSM, had the lowest proportion of patients in the uncertainty area of the test. This is a clinically relevant issue that ought to guide the management of NAFLD for each specific patient. Overall, we also confirmed the good accuracy of FIB-4 and LSM for the noninvasive diagnosis of advanced fibrosis as reported in recent meta-analyses.^{10,11} In our DCA, noninvasive scores/tools showed similar modest net benefit for ruling out advanced fibrosis at threshold probabilities of 5% and 10% of missing advanced fibrosis, confirming the need for liver biopsy for a correct identification of patients with advanced fibrosis, especially if the patient is worthy of inclusion in clinical trials.

Another relevant and currently debated topic in the research agenda of NAFLD is the noninvasive prediction of developing LRE. We first investigated the accuracy of AGILE 3+ for stratifying the risk of LRE occurrence in NAFLD, reporting that AGILE 3+ performed better when compared with FIB-4 and was significantly more accurate when compared with LSM. We also confirmed the diagnostic ability of both FIB-4 and LSM in stratifying the risk of LRE occurrence, our data being aligned with that reported in recent multicenter cohort studies showing a similarly good performance of these scores for predicting

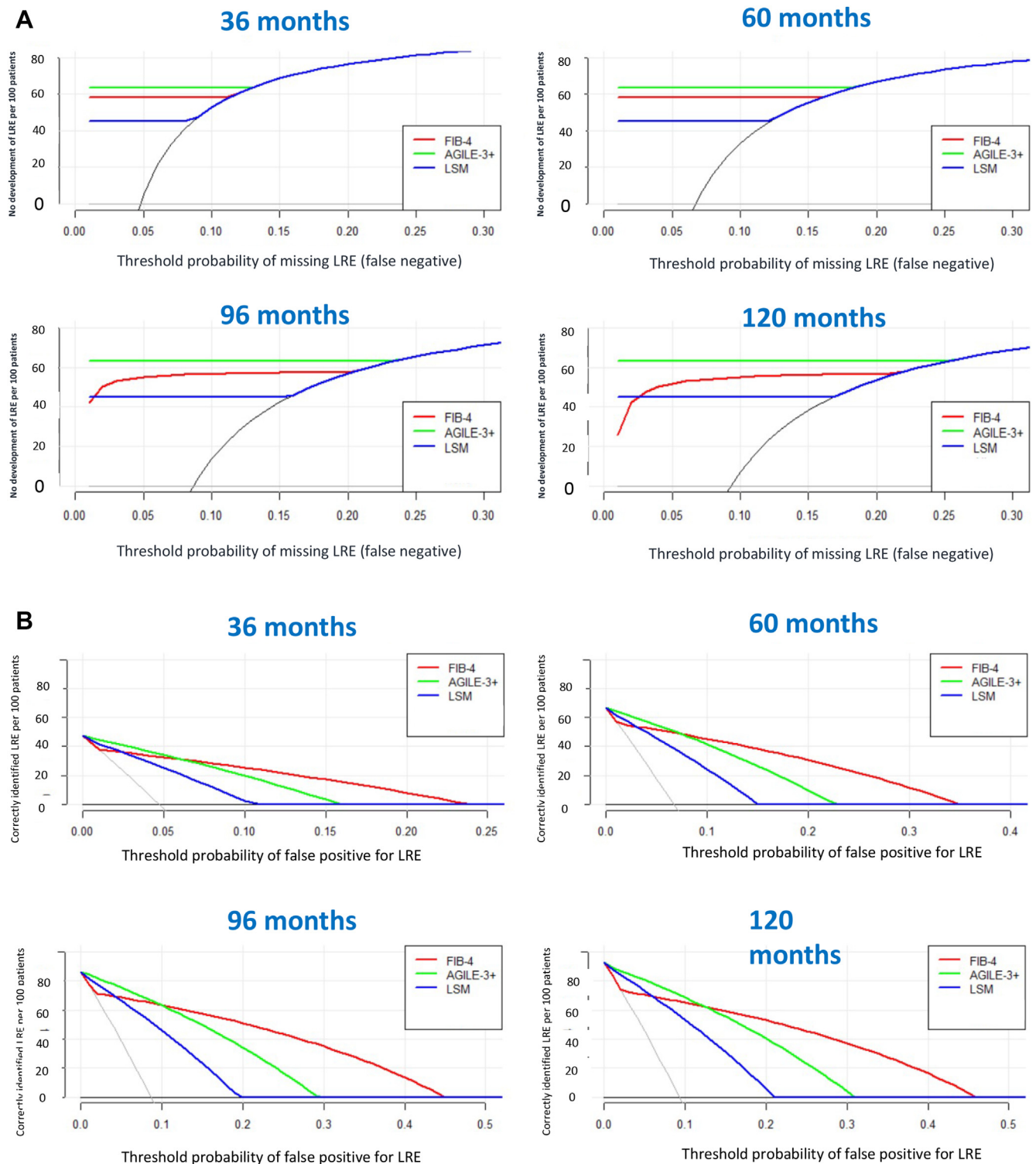


Figure 4. Net benefit of FIB-4, LSM, and AGILE 3+ by DCAs for ruling out (A) or ruling in (B) occurrence of LRE.

liver complications.¹⁶⁻¹⁹ In our DCA, especially in the early (36-month) period of follow-up, AGILE 3+ showed the highest net benefit for ruling out LRE occurrence across a threshold probability of missing occurrence of LREs ranging from 5% to 20%. Notably, these results were also confirmed after excluding patients with clinical diagnosis of NAFLD-compensated cirrhosis from the analyses, where we also demonstrated that noninvasive

scores were not inferior to fibrosis staging by histology for the purpose of predicting LRE occurrence.

From a clinical point of view, depending on: (1) the clinical setting (general practitioners, diabetes care centers, hepatology referral centers); (2) the resource status; and (3) the availability of TE for LSM, our study suggests that AGILE 3+ can be a valid alternative to both FIB-4 and LSM for assessing NAFLD severity and risk of

liver-related complications. Recent European guidelines for noninvasive tests for the assessment of chronic liver disease^{14,15} suggest using FIB-4 as the first test in non-specialized centers where patients are at low (general practitioners) or medium (diabetes care centers) risk of NAFLD-related advanced fibrosis, specifying LSM only for those patients who are at intermediate or high risk of advanced fibrosis as assessed by the FIB-4. In our study, however, we showed that LSM and AGILE 3+ alone outperformed the sequential strategy for diagnosing advanced fibrosis, whereas in the early period of follow-up, the FIB-4 to LSM strategy had a similar and slightly higher net benefit for ruling out LRE occurrence compared with AGILE 3+. The same guidelines also recommend to directly use LSM as first test in hepatological referral centers, suggesting the use, if available, of patented tests to confirm the presence of advanced fibrosis.¹⁵ In this scenario, AGILE 3+ could be preferred to LSM because of the following: (1) it reduces the proportion of patients where ruling out or ruling in advanced fibrosis is not possible (gray area), and (2) it demonstrates the highest accuracy for predicting LRE and the highest net benefit for ruling out LRE occurrence.

There are a number of limitations to the present study. The main limitation of the present study lies in its potentially limited validity of the results in different populations and settings. It is plausible that the obtained results could not be replicated in populations differing in age, biochemical profile, metabolic comorbidities, and severity of liver disease. Specifically, our analyses were done in a setting in which there was a high prevalence of advanced fibrosis (29%), whereas in populations at lower prevalence (such as the general population [about 1%–2% of advanced fibrosis] and diabetic patients [about 10% of advanced fibrosis]), it is expected that the PPV and the NPV of the tested noninvasive tools for advanced fibrosis will be lower and higher, respectively, compared with the population in our study. Moreover, the use of LSM for diagnosing cirrhosis may have resulted in an overestimation of severity of baseline liver disease. However, we are also confident that potential bias related to the correct definition of liver disease severity in the included patients with clinical diagnosis of NAFLD cirrhosis does not affect the interpretation of the results, because in our analyses, patients were stratified according to noninvasive scores/tools. Lack of an external validation cohort and potentially hidden alcohol abuse in some patients could further limit the interpretation of our results. On the other hand, the primary strength of our study is the use of the innovative DCA methodology.

Conclusion

In conclusion, we demonstrated that, according to resource availability, clinical setting, and the risk scenarios, AGILE 3+ is an accurate and valid alternative to

FIB-4 and LSM for the noninvasive assessment of disease severity and prognosis in patients with NAFLD.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at <https://doi.org/10.1016/j.cgh.2022.06.013>.

References

1. Le MH, Yeo YH, Li X, et al. 2019 Global NAFLD prevalence: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2021. Online ahead of print.
2. Harris R, Harman DJ, Card TR, et al. Prevalence of clinically significant liver disease within the general population, as defined by non-invasive markers of liver fibrosis: a systematic review. *Lancet Gastroenterol Hepatol* 2017;2:288–297.
3. Gu W, Hortlik H, Erasmus HP, et al. Trends and the course of liver cirrhosis and its complications in Germany: nationwide population-based study (2005 to 2018). *Lancet Reg Health Eur* 2021;12:100240.
4. Vitale A, Svegliati-Baroni G, Ortolani A, et al. Italian Liver Cancer (ITA.LI.CA) group. Epidemiological trends and trajectories of MAFLD-associated hepatocellular carcinoma 2002–2033: the ITA.LI.CA database. *Gut* 2021. Online ahead of print.
5. Haldar D, Kern B, Hodson J, et al. European Liver and Intestine Transplant Association (ELITA). Outcomes of liver transplantation for non-alcoholic steatohepatitis: a European Liver Transplant Registry study. *J Hepatol* 2019;71:313–322.
6. Mantovani A, Csermely A, Petracca G, et al. Non-alcoholic fatty liver disease and risk of fatal and non-fatal cardiovascular events: an updated systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* 2021;6:903–913.
7. Mantovani A, Petracca G, Beatrice G, et al. Non-alcoholic fatty liver disease and increased risk of incident extrahepatic cancers: a meta-analysis of observational cohort studies. *Gut* 2022; 71:778–788.
8. Taylor RS, Taylor RJ, Bayliss S, et al. Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. *Gastroenterology* 2020;158:1611–1625.e12.
9. Pennisi G, Enea M, Romero-Gomez M, et al. Liver-related and extrahepatic events in patients with non-alcoholic fatty liver disease: a retrospective competing risks analysis. *Aliment Pharmacol Ther* 2022;55:604–615.
10. Mózes FE, Lee JA, Selvaraj EA, et al. LITMUS Investigators. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. *Gut* 2022;71:1006–1019.
11. Selvaraj EA, Mózes FE, Jayaswal ANA, et al. LITMUS Investigators. Diagnostic accuracy of elastography and magnetic resonance imaging in patients with NAFLD: a systematic review and meta-analysis. *J Hepatol* 2021;75:770–785.
12. Petta S, Wai-Sun Wong V, Bugianesi E, et al. Impact of obesity and alanine aminotransferase levels on the diagnostic accuracy for advanced liver fibrosis of noninvasive tools in patients with non-alcoholic fatty liver disease. *Am J Gastroenterol* 2019;114:916–928.
13. Petta S, Wong VW, Cammà C, et al. Serial combination of non-invasive tools improves the diagnostic accuracy of severe liver

- fibrosis in patients with NAFLD. *Aliment Pharmacol Ther* 2017; 46:617–627.
14. 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.
 15. 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.
 16. Petta S, Sebastiani G, Viganò M, et al. Monitoring occurrence of liver-related events and survival by transient elastography in patients with nonalcoholic fatty liver disease and compensated advanced chronic liver disease. *Clin Gastroenterol Hepatol* 2021;19:806–815.e5.
 17. Younes R, Caviglia GP, Govaere O, et al. Long-term outcomes and predictive ability of non-invasive scoring systems in patients with non-alcoholic fatty liver disease. *J Hepatol* 2021; 75:786–794.
 18. Decraecker M, Dutartre D, Hiriart JB, et al. Long-term prognosis of patients with metabolic (dysfunction)-associated fatty liver disease by non-invasive methods. *Aliment Pharmacol Ther* 2022;55:580–592.
 19. Boursier J, Hagström H, Ekstedt M, et al. Non-invasive tests accurately stratify patients with NAFLD based on their risk of liver-related events. *J Hepatol* 2022;76:1013–1020.
 20. Younossi ZS, Newsome P, Chan W, et al. Development and validation of Agile 3+: novel FibroScan based score for the diagnosis of advanced fibrosis in patients with nonalcoholic fatty liver disease. *J Hepatol* 2021;75:S205–S293.
 21. Wong VW, Vergniol J, Wong GL, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology* 2010;51:454–462.
 22. Wong VW, Vergniol J, Wong GL, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. *Am J Gastroenterol* 2012;107:1862–1871.
 23. Introduction: standards of medical care in diabetes-2020. *Diabetes Care* 2020;43(Suppl 1):S1–S2.
 24. 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;41:313–321.
 25. McPherson S, Stewart SF, Henderson E, et al. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut* 2010;59:1265–1269.
 26. Boursier J, Zarski JP, de Ledinghen V, et al. Multicentric Group from ANRS/HC/EP23 FIBROSTAR Studies. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology* 2013;57:1182–1191.
 27. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018;69:182–236.
 28. Cillo U, Burra P, Mazzaferro V, et al. I-BELT (Italian Board of Experts in the Field of Liver Transplantation). A multistep, consensus-based approach to organ allocation in liver transplantation: toward a “blended principle model. *Am J Transplant* 2015;15:2552–2561.
 29. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; 26:565–574.
 30. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res* 2019;3:18.
 31. Zhang Z, Rousson V, Lee WC, et al. written on behalf of AME Big-Data Clinical Trial Collaborative Group. Decision curve analysis: a technical note. *Ann Transl Med* 2018;6:308.
 32. Wong VW, Irls M, Wong GL, et al. Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease. *Gut* 2019;68:2057–2064.

Correspondence

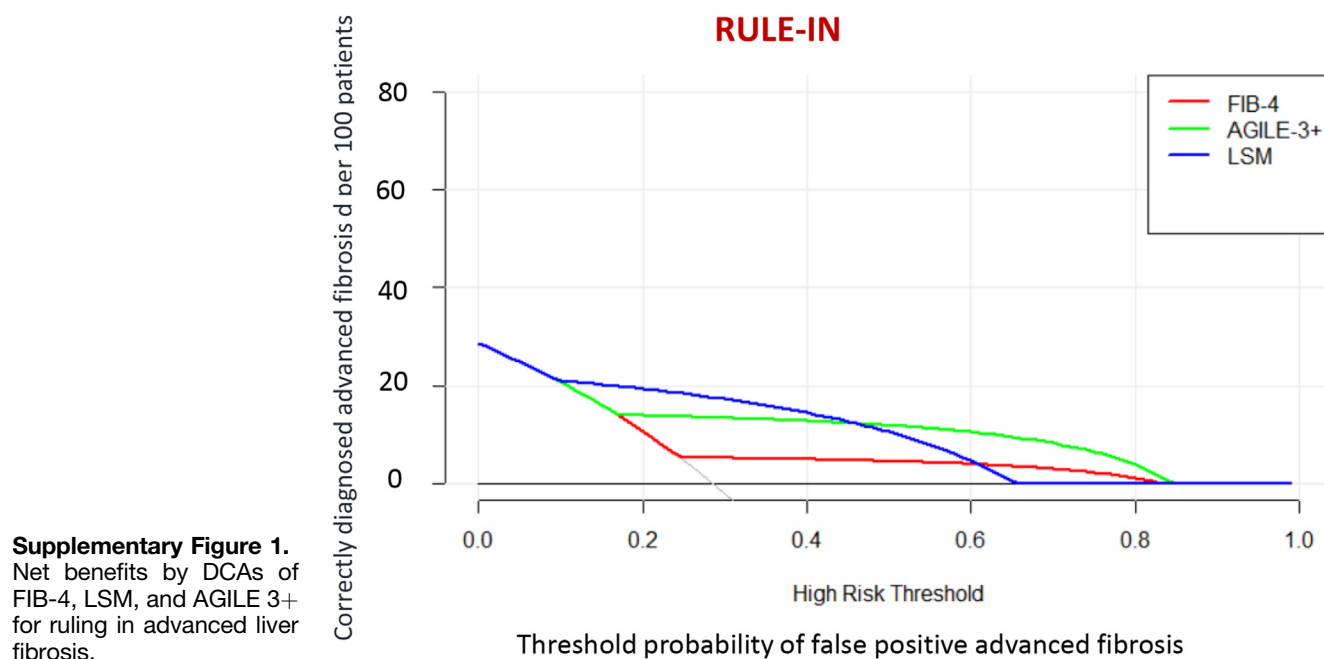
Address correspondence to: Dr Salvatore Petta, Gastroenterology and Hepatology, DiBiMIS, Policlinico Universitario Paolo Giaccone, Piazza delle Cliniche, 2, 90127 Palermo, Italy. e-mail: salvatore.petta@unipa.it; tel: +39 091-655-2274; fax +39 091-655-2156.

CRedit Authorship Contributions

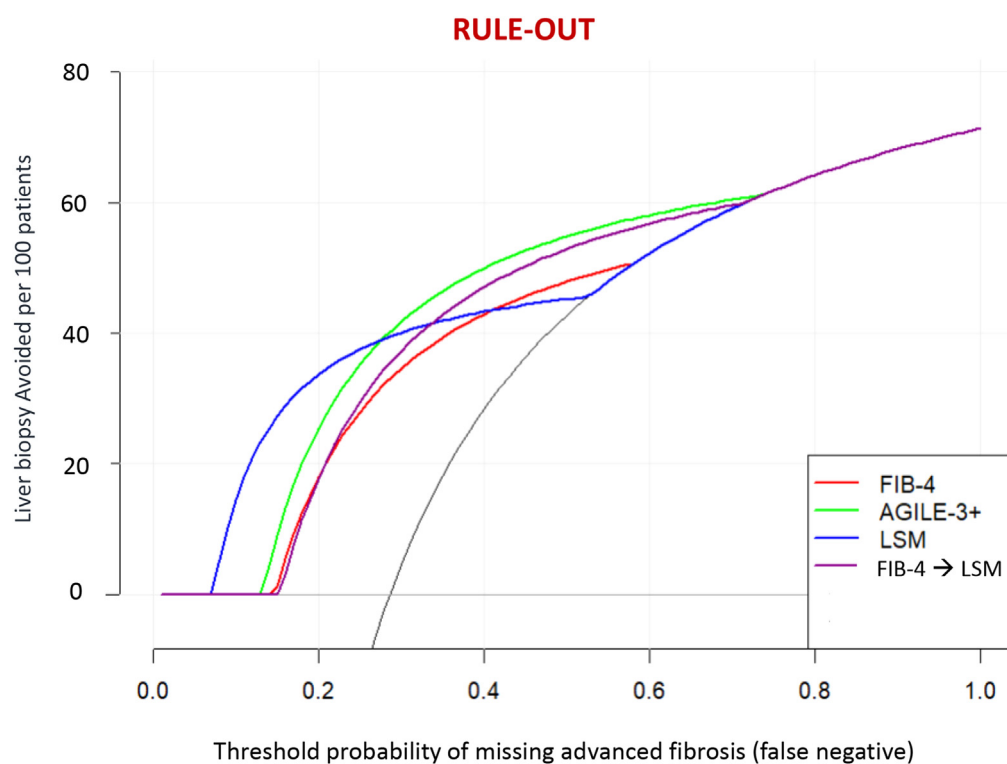
Grazia Pennisi (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Marco Enea (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Alessandra Pandolfo (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Ciro Celsa (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Michela Antonucci (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Carlo Ciccioli (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Giuseppe Infantino (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Claudia La Mantia (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Stefanie Parisi (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Adele Tulone (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Vito Di Marco (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Antonio Craxi (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Calogero Cammà (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)
 Salvatore Petta (Methodology; Formal analysis; Data curation; Visualization; Supervision; Writing – original draft)

Conflicts of interest

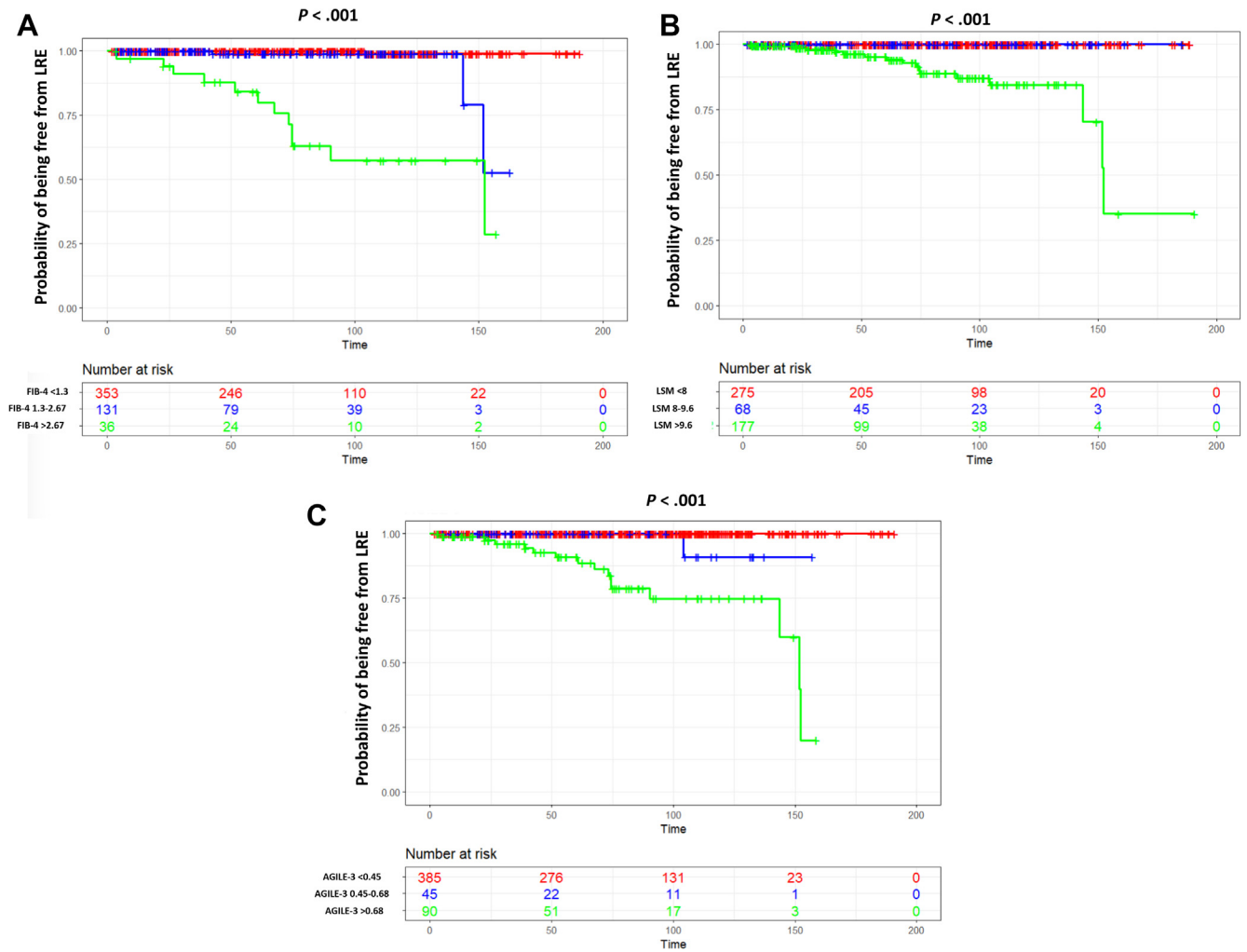
The authors disclose no conflicts.



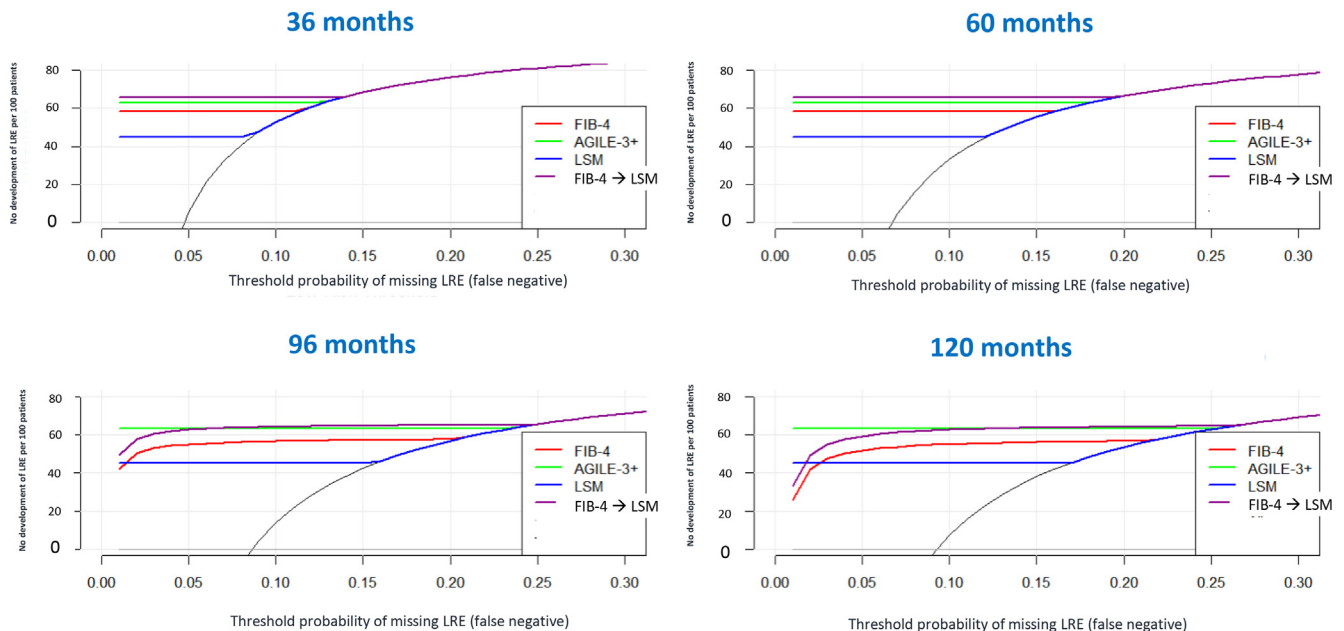
Supplementary Figure 1. Net benefits by DCAs of FIB-4, LSM, and AGILE 3+ for ruling in advanced liver fibrosis.



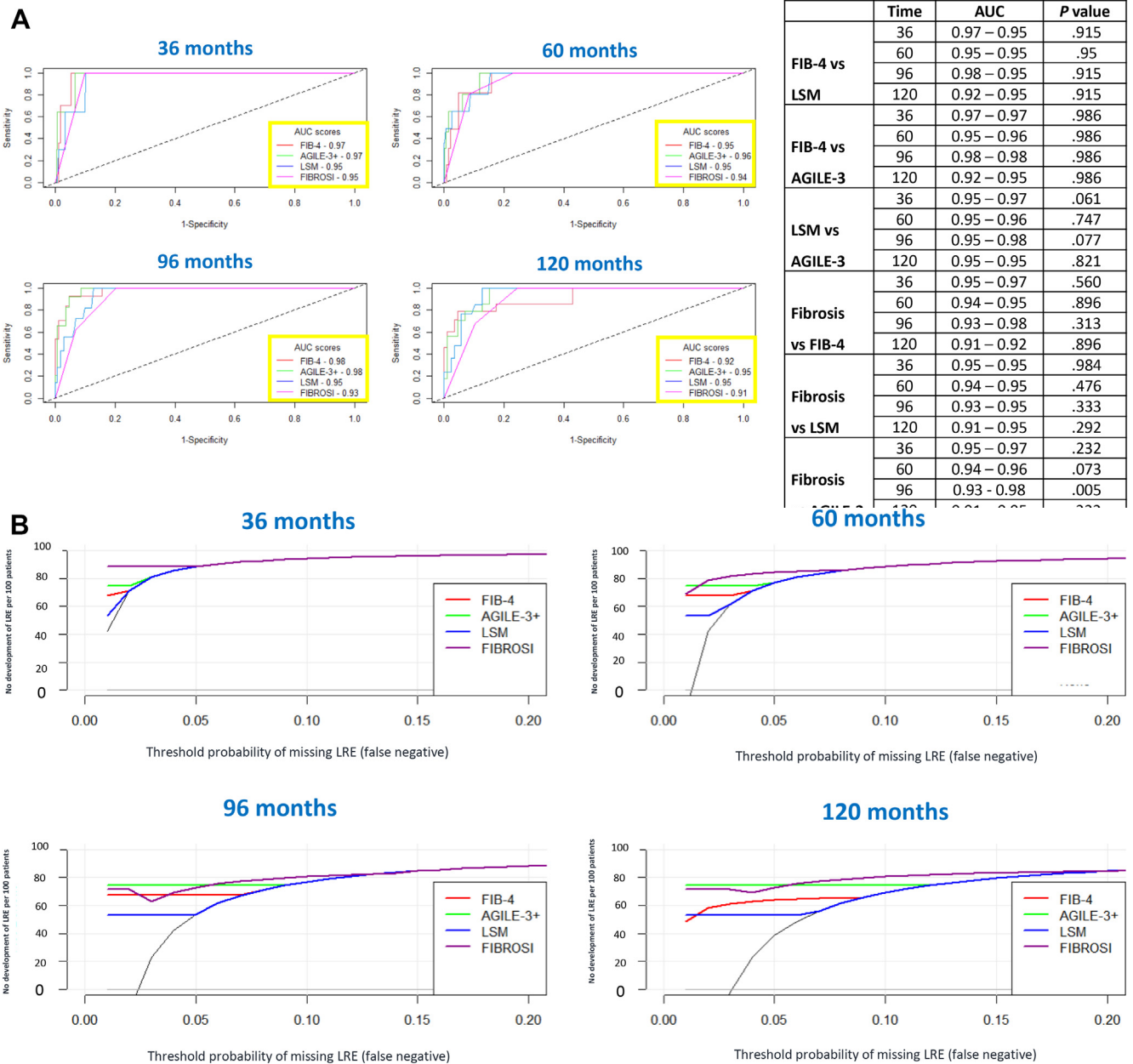
Supplementary Figure 2. Net benefits by DCAs of FIB-4, LSM, AGILE 3+ and the sequential FIB-4 to LSM strategy for ruling out advanced liver fibrosis.



Supplementary Figure 3. Kaplan-Meier curves depicting the probabilities of being free of LREs according to risk classes of FIB-4 (A), LSM (B), and AGILE 3+ (C).



Supplementary Figure 4. Net benefit of FIB-4, LSM, AGILE 3+ and the sequential FIB-4 to LSM strategy by DCAs for ruling out occurrence of LRE.



Supplementary Figure 5. A, ROC curves for prediction of LREs; and B, net benefit of histological fibrosis stage, FIB-4, LSM, and AGILE 3+ in the sub-group of patients with biopsy-proven NAFLD.

Supplemental Table 1. Baseline Demographic, Metabolic and Laboratory Features of Patients With NAFLD From the Training Cohort Stratified According to Advanced Fibrosis

	F0–F2 (n = 371)	F3–F4 (n = 243) ^a	P value
Age, y	44.73 ± 12.82	59.68 ± 10.92	< .001
Male gender	69%	57%	.002
Diabetes	21%	63%	< .001
BMI, kg/m ²	28.63 (26.14–32.05)	31.83 (28.69–35.10)	< .001
Albumin, g/dL	4.60 (4.40–4.81)	4.30 (4.00–4.50)	< .001
PLT, 10 ³ mmc	240 (205 - 278)	166 (117–236.5)	< .001
AST, U/L	36 (27 - 47)	40 (29–53)	.007
ALT, U/L	65 (43 - 93)	47 (33–68.50)	< .001
Total cholesterol, mg/dL	202 (175.50–231.50)	171 (147–194)	< .001
Triglycerides, mg/dL	125 (86.50–172.50)	126 (92.5–165.5)	.922
FIB-4	0.86 (0.59–1.19)	2.18 (1.31–3.69)	< .001
LSM	6.40 (5.20–8.70)	18.80 (11.60–29.15)	< .001
AGILE 3+	0.06 (0.03–0.17)	0.89 (0.56–0.97)	< .001
Follow-up, mos	78.75 ± 44.23	55.72 ± 40.64	.525

Note: Data are presented as mean ± standard deviations, median and interquartile range, or percentage of cases.

Note: P-values of the differences were adjusted by using the Benjamini-Hochberg correction for multiple tests.

ALT, Alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; NAFLD, nonalcoholic fatty liver disease; PLT, platelets.

^aThis group includes patients with clinical diagnosis of compensated NAFLD cirrhosis

Supplemental Table 2. Baseline Demographic, Metabolic and Laboratory Features of Patients With NAFLD From the Training Cohort Stratified According to Histological Diagnosis of NAFLD or Clinical Diagnosis of Compensated NAFLD Cirrhosis

	Clinical NAFLD compensated cirrhosis (n = 94)	Histological NAFLD (n = 520)	P value
Age, y	64.8 ± 8.1	48.0 ± 13.4	< .001
Male gender	60%	65%	.353
Diabetes	79%	30%	< .001
BMI, kg/m ²	32.4 (29.5 - 35.8)	29.4 (26.5 - 33.1)	< .001
Albumin, g/dL	4 (3.7 - 4.5)	4.5 (4.3 - 4.8)	< .001
PLT, 10 ³ mmc	116.5 (84 - 149.5)	233 (192.8 - 274)	< .001
AST, U/L	36 (27.2 - 46.7)	37 (28 - 51)	.192
ALT, U/L	35 (24.2 - 45)	62 (42 - 89.2)	< .001
Total cholesterol, mg/dL	165.5 (143 - 182.8)	194 (168 - 226)	< .001
Triglycerides, mg/dL	115 (84 - 152.8)	126.5 (88 - 172.2)	.102
FIB-4	3.46 (2.47 - 5.15)	0.97 (0.65 - 1.47)	< .001
LSM	26.0 (17.6 - 39.3)	7.8 (5.7 - 11.1)	< .001
AGILE 3+	0.97 (0.93 - 0.99)	0.11 (0.03 - 0.46)	< .001
Follow-up, mos	41.7 ± 30.6	74.6 ± 44.4	.321

Note: Data are presented as mean ± standard deviations, median and interquartile range, or percentage of cases.

Note: P-values of the differences were adjusted by using the Benjamini-Hochberg correction for multiple tests.

ALT, Alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; NAFLD, nonalcoholic fatty liver disease; PLT, platelets.

Supplementary Table 3. Net Benefit by DCA for Ruling In Advanced Liver Fibrosis and Development of LREs

Outcome: advanced fibrosis		Correctly diagnosed advanced fibrosis per 100 patients (net benefit)			
		FIB-4	LSM	AGILE 3+	
Threshold probabilities of missing advanced fibrosis	5%	25	25	25	
	10%	21	21	21	
	25%	5	18	14	
Outcome: LREs		Correctly identified LREs per 100 patients (net benefit)			
		Time, mos	FIB-4	LSM	AGILE 3+
Threshold probabilities of missing LREs	5%	36	3	2	3
		60	5	5	5
		96	7	7	7
		120	7	7	8
	20%	36	0	0	0
		60	3	0	1
		96	5	0	3
		120	5	0	4

DCA, Decision curve analysis; FIB-4, Fibrosis-4; LSM, liver stiffness measurement; LRE, liver-related event.