



Poor Performance of Non-invasive Tests for Advanced Fibrosis in Nonalcoholic Fatty Liver Disease: A Multicentric Asian Study

Umang Arora¹ · Ritik M. Goyal¹ · Kevin K. J. Teh² · Yiyang Pei² · George B. B. Goh² · Su Lin³ · Rahul Kumar^{4,5} · Sagnik Biswas¹ · Manas Vaishnav¹ · Shekhar Swaroop¹ · Piyush Pathak¹ · Sabreena Sheikh¹ · Vishwesh Bharadiya¹ · Anshuman Elhence¹ · Shivanand Gamanagatti⁶ · Rajni Yadav⁷ · Prasenjit Das⁷ · Sandeep Aggarwal⁸ · Narendra Choudhary⁹ · Prajna Anirvan¹⁰ · Shivaram P. Singh¹⁰ · Arka De¹¹ · Ajay Duseja¹¹ · Shalimar¹

Received: 8 September 2022 / Accepted: 31 July 2023 / Published online: 21 September 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Background Non-invasive tests (NITs) are useful to assess advanced fibrosis (AF) in nonalcoholic fatty liver disease (NAFLD). Data from Asian countries suggest that these tests have poor performance. We aimed to assess diagnostic accuracy of established thresholds of biomarker-based NITs and Transient Elastography (TE) in identifying AF and evaluated the utility of a two-step test approach.

Methods Biopsy-proven 641 NAFLD patients (55.2% males, median age 42 years) were included from three different centers of Asia. AF (\geq F3) was identified as per histological staging (24.8%).

Results TE had the highest area under the receiver operating characteristic curve (AUROC) 0.82 (0.79–0.86), and all other biomarker-based NITs had low AUROC ($<$ 0.7). NITs performed poorly at established thresholds. The combination of NITs utilizing liver stiffness measurement (LSM) and biomarkers, Agile 3+ and FAST, demonstrated acceptable diagnostic accuracy (AUROC 0.82 and 0.78, respectively), but none were superior to LSM alone. LSM measured using appropriate M and XL probes remained accurate regardless of body mass index (BMI); NFS and APRI scores were less accurate at higher BMI ranges. A two-step approach using NFS rule-out criteria ($<$ –2.97 to rule out) followed by LSM ($<$ 7.3 kPa to rule out and \geq 12.7 kPa to rule in) correctly classified 62.4% of patients, with only 10.2% of patients incorrectly classified.

Conclusion NITs have not been validated to identify AF in the Asian NAFLD population, and internationally accepted thresholds yield high false-negative rates. LSM and LSM-based combination tests remain the most accurate.

Keywords Fatty liver · NITs · Liver biopsy · Hepatitis · NASH · Cirrhosis

Introduction

Nonalcoholic fatty liver disease (NAFLD) is characterized by excess hepatic fat in the absence of significant alcohol consumption and alternate etiologies for hepatic steatosis [1]. It has burgeoned to become the most common liver disease, with an estimated global prevalence of 25% [2]. Nearly 20% of all NAFLD patients suffer from nonalcoholic steatohepatitis (NASH), characterized by hepatocyte injury and an increased risk for liver cirrhosis, hepatocellular carcinoma, liver failure, and cardiovascular death [2]. Although

one-third of all Asians are estimated to suffer from NAFLD, it remains under-recognized, and less than 10% of diagnosed cases are referred to a specialist [3–5].

Patients with NASH, particularly those with advanced fibrosis (AF) \geq F3, are at risk of adverse outcomes. Therefore, this sub-population needs referral and management by a specialist. Liver biopsy remains the gold standard for identifying NASH and fibrosis. However, the invasive nature, sampling variability, inter- and intra-observer reporting variability, and associated risks and costs prohibit its widespread use in NAFLD. Non-invasive tests (NITs) for the assessment of fibrosis include biomarker-based NITs like NAFLD fibrosis score (NFS), fibrosis-4 index (FIB4), aspartate aminotransferase (AST)-platelet ratio index (APRI), AST-alanine aminotransferase (ALT) ratio (AAR), enhanced liver fibrosis (ELF), and body mass index-AAR-diabetes mellitus

An editorialcommenting on this article is available at <https://doi.org/10.1007/s10620-023-08094-x>.

Extended author information available on the last page of the article

(BARD) score, or mechanical liver stiffness using magnetic resonance-based or ultrasound-based transient elastography (TE) techniques [6]. Biomarker results are influenced by factors including age, comorbidities, the prevalence of fibrosis, and NASH [7]. Moreover, most NITs have been developed and validated predominantly in Caucasians [8]. Reports from Asia and other geographic regions suggest poor diagnostic accuracy of biomarker-based NITs [9–11]. Therefore, the present study was designed to evaluate the diagnostic accuracy of NITs for identifying AF in biopsy-proven NAFLD patients. We aimed to identify optimal thresholds for the Asian population and assess the role of two-step screening with a blood-based biomarker followed by TE in resource-limited settings.

Patients and Methods

We invited authors with prior publications/experience in managing NAFLD from South Asia to be part of this study. A total of 9 centers were invited to contribute data, of whom 6 agreed to participate. Two centers did not include any patients with advanced fibrosis (Medanta Medicity, Gurgaon, Haryana, India and SCB Hospital, Cuttack, Orissa,

India), while another center had not performed TE for their patients (First Affiliated Hospital of Fujian Medical University, Fuzhou, Fujian, China), their details are provided in Fig. 1.

Finally, we retrospectively analyzed data from three centers which included four cohorts—three from India and one from Singapore. Each cohort consisted of a prospectively maintained database of consecutive adult patients (age > 18 years) with biopsy-proven NAFLD. The first cohort (All India Institute of Medical Sciences, AIIMS, liver clinic cohort) comprised patients recruited from a gastroenterology out-patient clinic, and the second cohort consisted of patients referred to the surgical department for bariatric surgery at the same center (AIIMS bariatric cohort). The third cohort included patients following up in the hepatology department at the Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India, while the fourth cohort included patients following up at Singapore General Hospital, Singapore. Investigations to exclude other causes of liver disease were performed, which included hepatitis-B surface-antigen, anti-hepatitis C antibody, anti-nuclear antibody, anti-smooth muscle antibody, and anti-liver-kidney-microsomal antibody and serum ceruloplasmin. Any patients who consumed alcohol

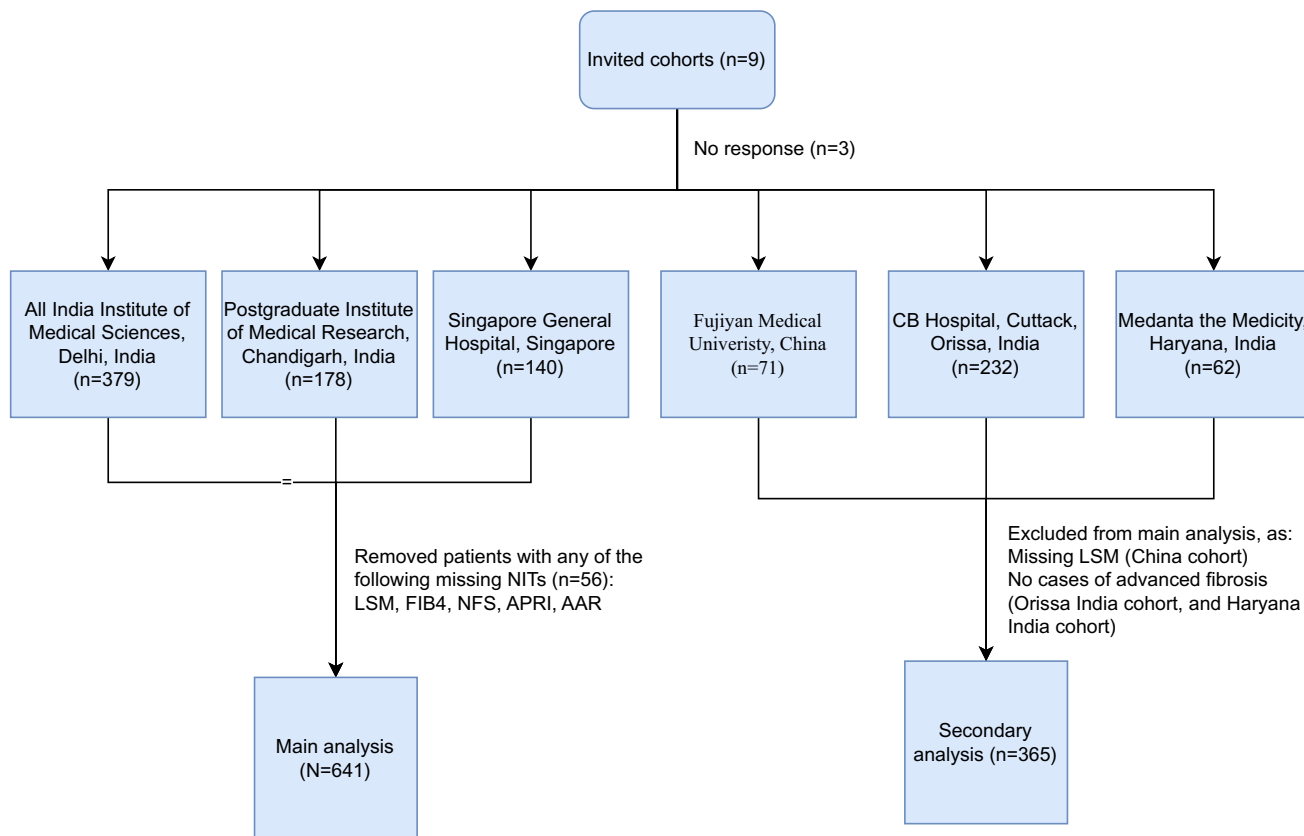


Fig. 1 Patient flow in the study cohort

in amounts beyond the cut-off recommendation for NAFLD (14 units/week for females and 21 units/week for males) were excluded [12]. Additionally, any patient with human immunodeficiency virus (HIV) infection, pregnancy, heart failure, ascites, active malignancy, valvular heart disease, or liver transplant were excluded.

Liver Biopsy and Fibrosis

An ultrasound-guided percutaneous plugged liver biopsy was performed by treating physician or interventional radiologist at each individual center. In patients undergoing bariatric surgery (AIIMS bariatric cohort), liver biopsy was performed intraoperatively from the left lobe using a 16G Trucut biopsy needle (Bard, Maxcore). The biopsy specimens were fixed using 10% buffered neutral formalin, processed, and paraffin-embedded. Hematoxylin–eosin staining along with Masson’s trichrome was used to stain sections of 4 microns thickness which were evaluated by expert pathologists at each center blinded to clinical and laboratory parameters. Samples were considered adequate if at least 6 portal triads could be identified. NASH Clinical Research Network (CRN) scoring system was used to assess the stage of fibrosis, and NASH activity score > 5 points to define NASH [13]. Fibrosis seen in NAFLD patients was graded from F0 to F4. F0 represents the absence of fibrosis, F1 includes periportal or perisinusoidal fibrosis, F2 perisinusoidal and portal/periportal, F3 bridging fibrosis, and F4 cirrhosis [13]. AF includes F3 and F4 stages of fibrosis and is the key prognostic indicator for liver-related events, overall mortality, and occurrence of HCC in NAFLD. [14]

Non-invasive Tests

All patients underwent TE and blood investigations. TE was performed after overnight fasting using FibroScan touch 502 (Echosens, France) at each center by trained operators blinded to clinical and histological details. TE provides two outputs: liver stiffness measurement (LSM), which represents tissue stiffness and fibrosis, and the controlled attenuation parameter (CAP), which represents steatosis. Measurements were taken from right lobe of liver using a window in the intercostal space with the patient in supine position and the right arm abducted. The “M-probe” was used for those with BMI < 30 kg/m² and the “XL-probe” for BMI > 30 kg/m². [15] Ten valid measurements were taken for LSM and CAP and values reported as median with interquartile range (IQR), with measurements considered reliable if IQR/median was < 0.3.

Blood investigations included hemogram, liver, and renal biochemical tests, lipid profile, fasting blood glucose, and insulin levels. These were used to calculate the five established blood-based NITs: NFS [9], FIB4 [16], BARD score

[17], APRI [18], and AAR (Supplementary Table 1). Agile 3+ score is a recently validated score for diagnosis of AF in NAFLD that utilizes a combination of TE and blood tests, and was included in the primary analysis. A similar score (FAST score) has been used to diagnose fibrotic NASH (i.e., NAS score ≥ 4 and significant fibrosis ≥ F2)—a concept distinct from AF. All blood investigations were performed within 4 weeks of TE.

Statistical Analysis

We limited the analysis to those patients from the cohorts who had data available for estimating each of the five blood-based tests as well as LSM by TE. The qualitative data were presented as proportions and compared using the Chi-square test. The quantitative data were assessed for normality using the Shapiro–Wilk test and presented as mean (standard deviation, SD) or median (interquartile range, IQR) as appropriate. Continuous data were compared using Students’ t-test for the mean (SD) or Mann–Whitney U test for median (IQR). The diagnostic accuracy of each NIT was assessed using the area under the receiver operating characteristic curve (AUROC). Two diagnostic cut-offs corresponding to the 90% sensitivity and specificity thresholds for AF were calculated to form the rule-out and rule-in criteria for each test. We also estimated sensitivity and specificity at thresholds recommended by professional hepatology societies. Univariate analysis and multivariate logistic regression were performed using a regression model to assess factors predicting AF. AUROC for the available NITs was compared using the DeLong method. For deriving a combination of two tests in series with the best accuracy, we calculated the proportions of patients who were correctly classified (combining true positives (TP) and true negatives (TN)), incorrectly classified (False negatives [FN] and False Positives [FP]), or indeterminate among total cases. A *p* value of < 0.05 was considered significant. Analysis was performed using Stata v14 [Stata Corp, Texas].

Results

A total of 641 patients were available for analysis (323 patients from two cohorts from New Delhi, India, 178 patients from Chandigarh, India, and 140 patients from Singapore) (Fig. 1). Baseline clinical, comorbidity, and laboratory parameters along with NIT measurement, stratified by the center, are shown in Table 1. The study population comprised 55.2% males, mean (\pm SD) age 42.9 ± 12.8 years, and mean BMI 30.5 ± 8.55 kg/m². The AIIMS bariatric cohort expectedly had significantly greater median BMI compared to other cohorts, (42.7 (37.9–47.9) kg/m² vs 26.6 (24.2 vs 29.8) kg/m², *p* < 0.001), while the Singapore cohort included

Table 1 Clinical, laboratory parameters, and non-invasive tests (transient elastography, biomarker based, and combination tests) across all cohorts ($N=641$)

	AIIMS Liver Clinic ($n=205$)	AIIMS Bariatric cohort ($n=118$)	PGI Chandigarh ($n=178$)	Singapore ($n=140$)	Total ($N=641$)
Male, n (%)	142 (69.3%)	31 (26.3%)	117 (65.7%)	64 (45.7%)	354 (55.2%)
Age, years	39 (32–47)	39 (33–48)	38 (31–47)	55 (47–62)	42 (34–52)
BMI, mean (SD), kg/m^2	27.56 (6.27)	43.5 (8.08)	27.28 (4.61)	28.19 (4.72)	30.55 (8.55)
Diabetes, n (%)	50 (24.4%)	47 (39.8%)	31 (17.4%)	88 (62.9%)	216 (33.7%)
Platelets, $\times 10^9/\mu\text{L}$	186 (134–246)	249.5 (208–319)	219.5 (161–264)	247.5 (206–300.5)	224 (164–274)
Bilirubin, mg/dL	0.7 (0.5–1.1)	0.6 (0.4–0.7)	0.7 (0.5–1.04)	0.73 (0.58–0.94)	0.7 (0.5–0.9)
Aspartate aminotransferase (AST), IU/L	47 (33–69)	25.5 (20–40)	68.64 (47–92)	48 (36–61.5)	47 (32.44–72)
Alanine aminotransferase (ALT), IU/L	70 (43–99)	29 (22–43)	86.85 (61–115.05)	66.5 (46–91.5)	66 (39–94.85)
Alkaline Phosphatase (ALP), IU/L	195.5 (117–270.5)	118 (81–177)	131 (100–197)		152 (100–228)
Albumin, mean (SD), g/dL	4.63 (0.55)	4.09 (0.46)	4.2 (0.65)	4.23 (0.35)	4.32 (0.57)
Triglycerides, mg/dL	146 (105–203)	137 (106–161)			142 (105.5–182.5)
Low-density lipoprotein (LDL), mg/dL	109.5 (87.5–137)	105 (84.5–130)			106 (86–134.5)
Steatosis grade (S0–S3, %)					
S0	14 (6.8%)	20 (16.9%)	3 (1.7%)	1 (0.7%)	38 (5.9%)
S1	93 (45.4%)	55 (46.6%)	40 (22.5%)	47 (33.6%)	235 (36.7%)
S2	65 (31.7%)	34 (28.8%)	84 (47.2%)	71 (50.7%)	254 (39.6%)
S3	33 (16.1%)	9 (7.6%)	51 (28.7%)	21 (15%)	114 (17.8%)
NASH, n (%)	54 (26.3%)	19 (16.1%)	29 (16.3%)	80 (57.1%)	182 (28.4%)
Fibrosis grade (F0–F4, %)					
F0	71 (34.6%)	36 (30.5%)	77 (43.3%)	8 (5.7%)	192 (30%)
F1	53 (25.9%)	35 (29.7%)	60 (33.7%)	53 (37.9%)	201 (31.4%)
F2	32 (15.6%)	22 (18.6%)	19 (10.7%)	16 (11.4%)	89 (13.9%)
F3	20 (9.8%)	15 (12.7%)	18 (10.1%)	46 (32.9%)	99 (15.4%)
F4	29 (14.1%)	10 (8.5%)	4 (2.2%)	17 (12.1%)	60 (9.4%)
Advanced fibrosis (\geq F3, %)	49 (23.9%)	25 (21.2%)	22 (12.4%)	63 (45%)	159 (24.8%)
NAFLD activity score (0–8)	3 (3–5)	3 (3–4)	3 (2–4)	5 (4–5)	4 (3–5)
Liver stiffness measurement (LSM), kPa	7.4 (5.6–11.4)	8.3 (5.9–10.5)	6.85 (5.4–9.9)	11.4 (7.05–14.7)	8 (5.8–11.8)
Controlled attenuation parameter (CAP) ($n=609$), dB/m	319 (283–349)	332 (304–366)	300 (268–331)	329.5 (296–359)	318 (283–349)
FIB4	1.18 (0.85–2.07)	0.71 (0.55–1.33)	1.34 (0.84–2.33)	1.33 (0.99–1.96)	1.16 (0.75–1.94)
NAFLD fibrosis score (NFS),	– 2.26 (– 3.24 to – 1.22)	– 0.89 (– 1.78–0.6)	– 2.4 (– 3.3 to – 1.25)	– 1.52 (– 2.73 to – 0.57)	– 1.88 (– 2.94 to – 0.7)
AST-platelet ratio index (APRI)	0.7 (0.42–1.06)	0.28 (0.18–0.46)	0.78 (0.56–1.32)	0.49 (0.37–0.65)	0.58 (0.35–0.95)
AST-ALT ratio	0.73 (0.56–0.94)	0.86 (0.71–1.07)	0.78 (0.63–0.95)	0.74 (0.61–0.91)	0.78 (0.61–0.95)
BARD	1 (0–2)	3 (2–4)	1 (0–2)	2 (1–3)	2 (1–3)
FAST score ($n=609$)	0.51 (0.33–0.67)	0.28 (0.13–0.55)	0.57 (0.41–0.68)	0.63 (0.49–0.76)	0.53 (0.33–0.68)
Agile 3+	0.21 (0.08–0.56)	0.25 (0.09–0.54)	0.16 (0.07–0.39)	0.53 (0.23–0.82)	0.24 (0.09–0.6)

All values are presented as median (IQR) unless specified

SD standard deviation, IQR interquartile range

older patients (54.2 ± 11.1 vs 39.8 ± 10.5 years, $p < 0.001$). On liver biopsy, the proportion of patients with fibrosis grades F0, F1, F2, F3, and F4 were 30%, 31.4%, 13.9%, 15.4%, and 9.4%, respectively, thus there were 159 patients (24.8%) with AF (\geq F3 fibrosis). Histological evidence of NASH was present in 28.4% of patients, with a higher prevalence among patients with AF compared to F0-F2 fibrosis (53.3% vs 15.7%, $p < 0.001$).

Diagnostic Accuracy of Existing NITs in Identifying Fibrosis

We measured AUROC for all tests in identifying AF in each cohort (Table 2) Overall, the diagnostic performance of LSM was superior to all blood biomarker-based NITs for the detection of AF ($p < 0.001$, Delong test), and most biomarker-based tests had poor diagnostic accuracy with an AUROC < 0.7 (Fig. 2). Diagnostic accuracy was high (> 0.8) for only LSM, and Agile3+ score—a combination score that utilizes LSM values. On pairwise comparison (Supplementary Table 2), LSM was superior to the FAST score ($p = 0.011$, Delong test), but had similar diagnostic accuracy to Agile 3+ . Among the biomarker-based NITs, NFS and FIB4 had the highest AUROC (0.69 and 0.67) but were not significantly different ($p = 0.461$, Delong test). We performed univariate and multivariate analysis to identify factors associated with AF in this population and found age, diabetes, $\text{BMI} \geq 25 \text{ kg/m}^2$, AST, and platelet counts to be significant (Supplementary Table 3).

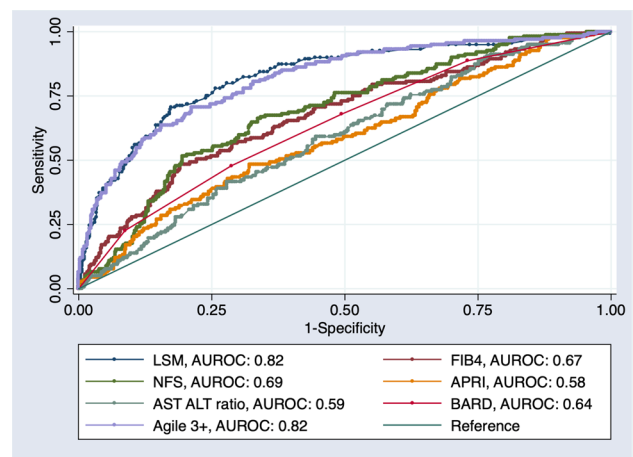


Fig. 2 Area under ROC curve (AUROC) of various non-invasive tests for detecting advanced fibrosis (AF). *LSM* Liver stiffness measurement, *NFS* NAFLD Fibrosis score, *APRI* AST-Platelet Ratio Index, *AAR* AST-ALT ratio, *LSM* liver stiffness measurement, Agile 3+

Diagnostic Accuracy of NITs for Various BMI Categories

We compared the AUROC of each NIT for diagnosis of AF with BMI groups both below and above thresholds of 25, 30, 35, and 40 kg/m^2 (Table 3). The most accurate NITs, LSM, and Agile 3+ remained accurate over these BMI categories.

Existing NITs Underperform at Recommended Thresholds

We reviewed existing major society guidelines [19–21] and compared the sensitivity and specificity data from our study population and that described by other cohorts. We found

Table 2 Area under ROC curve (AUROC) for non-invasive tests for detecting advanced fibrosis (AF) in all cohorts ($N = 641$)

	AIIMS Liver Clinic ($n = 205$)	AIIMS Bariatric cohort ($n = 118$)	PGI Chandigarh ($n = 178$)	Singapore ($n = 140$)	Total ($N = 641$)
Liver stiffness measurement (LSM) by TE	0.88 (0.81–0.94)	0.82 (0.72–0.92)	0.72 (0.61–0.84)	0.76 (0.68–0.84)	0.82 (0.79–0.86)
FIB4	0.73 (0.64–0.81)	0.63 (0.51–0.75)	0.66 (0.54–0.79)	0.7 (0.61–0.78)	0.67 (0.62–0.72)
NAFLD fibrosis score (NFS)	0.75 (0.67–0.83)	0.57 (0.45–0.7)	0.65 (0.53–0.78)	0.73 (0.64–0.81)	0.69 (0.64–0.74)
AST-platelet ratio index (APRI)	0.64 (0.55–0.74)	0.71 (0.59–0.83)	0.64 (0.52–0.76)	0.61 (0.51–0.7)	0.58 (0.53–0.64)
AST-ALT ratio	0.65 (0.56–0.75)	0.51 (0.38–0.65)	0.55 (0.43–0.68)	0.64 (0.55–0.73)	0.59 (0.54–0.64)
BARD	0.67 (0.58–0.76)	0.58 (0.45–0.71)	0.62 (0.48–0.76)	0.67 (0.58–0.76)	0.64 (0.59–0.69)
Agile 3+	0.87 (0.81–0.93)	0.75 (0.64–0.86)	0.74 (0.63–0.85)	0.77 (0.69–0.85)	0.82 (0.78–0.85)
	AIIMS Liver Clinic ($n = 205$)	AIIMS Bariatric cohort ($n = 118$)	PGI Chandigarh ($n = 178$)	Singapore ($n = 108$)	Total ($N = 609$)
FAST score ($n = 609$)	0.8 (0.72–0.88)	0.82 (0.72–0.91)	0.79 (0.71–0.87)	0.73 (0.64–0.83)	0.78 (0.74–0.83)

Table 3 Area under ROC curve (AUROC) for non-invasive tests for detecting advanced fibrosis (AF) in various BMI strata (N=641)

	<25 kg/m ² (n = 172)	≥25 kg/m ² (n = 469)	p value	<30 kg/m ² (n = 394)	≥30 kg/m ² (n = 247)	p value	<35 kg/m ² (n = 504)	≥35 kg/m ² (n = 137)	p value	<40 kg/m ² (n = 544)	≥40 kg/m ² (n = 97)	p value
Liver stiffness measurement (LSM) by TE	0.84 (0.75–0.92)	0.82 (0.77–0.86)	0.707	0.82 (0.77–0.87)	0.82 (0.76–0.88)	0.994	0.82 (0.77–0.86)	0.85 (0.78–0.93)	0.411	0.83 (0.78–0.87)	0.83 (0.73–0.93)	0.411
FIB4	0.76 (0.65–0.87)	0.67 (0.62–0.72)	0.163	0.68 (0.62–0.75)	0.7 (0.63–0.77)	0.669	0.67 (0.62–0.73)	0.7 (0.6–0.8)	0.645	0.69 (0.64–0.74)	0.6 (0.46–0.73)	0.645
NAFLD fibrosis score (NFS)	0.78 (0.68–0.87)	0.65 (0.59–0.7)	0.027*	0.71 (0.65–0.77)	0.63 (0.55–0.7)	0.087	0.72 (0.67–0.78)	0.61 (0.5–0.72)	0.062	0.73 (0.68–0.78)	0.53 (0.38–0.68)	0.062
AST-platelet ratio index (APRI)	0.63 (0.51–0.76)	0.6 (0.54–0.66)	0.63	0.57 (0.5–0.64)	0.66 (0.59–0.74)	0.072	0.54 (0.48–0.6)	0.76 (0.66–0.86)	<0.001*	0.57 (0.51–0.62)	0.69 (0.55–0.83)	<0.001*
AST-ALT ratio	0.64 (0.54–0.75)	0.57 (0.51–0.63)	0.227	0.59 (0.52–0.65)	0.57 (0.5–0.65)	0.227	0.6 (0.55–0.66)	0.56 (0.44–0.67)	0.486	0.61 (0.55–0.66)	0.5 (0.36–0.64)	0.153
BARD	0.66 (0.55–0.78)	0.61 (0.55–0.66)	0.402	0.63 (0.57–0.7)	0.61 (0.54–0.69)	0.402	0.67 (0.61–0.72)	0.6 (0.49–0.71)	0.31	0.66 (0.61–0.72)	0.57 (0.43–0.71)	0.199
Agile 3+	0.86 (0.78–0.94)	0.8 (0.76–0.85)	0.19	0.82 (0.76–0.87)	0.81 (0.75–0.87)	0.19	0.82 (0.77–0.86)	0.8 (0.71–0.89)	0.676	0.83 (0.79–0.87)	0.72 (0.6–0.85)	0.111
FAST score (n = 609)	0.76 (0.64–0.89)	0.79 (0.74–0.83)	0.733	0.8 (0.74–0.86)	0.79 (0.72–0.85)	0.733	0.77 (0.71–0.82)	0.86 (0.78–0.93)	0.049*	0.78 (0.73–0.83)	0.85 (0.75–0.94)	0.197

*Significant at $p < 0.05$ (Delong test)

that thresholds of FIB4 and NFS that rule out NAFLD had showed low sensitivity in our cohort (66% and 65%, respectively). The 90% sensitivity and 90% specificity cut-offs in the study population for LSM were 7.3 and 12.7 kPa, for FIB4 0.66 and 2.45, and for NFS – 2.97 and 0.29, respectively (Table 4).

Two-Step Approach Combining Biomarker-Based NIT and LSM with TE

We assessed the diagnostic accuracy for identifying AF using a combination of two tests that included a biomarker-based NIT followed by TE (Table 5). As a single test, LSM performed the best with 55.4% patients correctly classified for AF, while 10.1% patients were incorrectly classified, and 34.5% patients remained unclassified. The biomarker-based NITs (FIB4, NFS) had a large proportion of patients (64–68%) in the unclassified category. The combination score Agile 3+ did not improve diagnostic classification

compared to LSM alone, with respect to the proportion of correctly or incorrectly classified patients. We applied a two-step diagnostic model for biomarker-based NITs followed by LSM. The first approach utilized both rule-out and rule-in criteria of FIB4 or NFS followed by LSM only in patients not classified by either FIB4 or NFS. However, there was an increase in the number of false positives based on the rule-in criteria of FIB4 or NFS (from 10.1% in LSM alone to 16.5%). A second algorithm was used where 90% sensitivity threshold for FIB4 or NFS was used to rule out AF followed by TE, and this approach yielded similar correctly classified patients with a lower number of incorrectly classified. The highest numerical percentage of correctly classified cases (62.4%), as well as the lowest percentage among indeterminate cases (27%), was found to be with a sequence of NFS (rule out criteria only) followed by LSM (i.e., NFS < – 2.97 to rule out, and subsequently apply LSM criteria to all cases with NFS ≥ – 2.97) (Fig. 3). One-fourth (155, 24.8%) of patients had an NFS of < – 2.97 and thus

Table 4 Area under ROC curve (AUROC), sensitivity, specificity of various non-invasive tests for advanced fibrosis (except for FAST and Agile 4) as per cut-offs described in the literature (N= 641)

Non-invasive tests	Existing literature		AUROC	Study population	
	Recommended cut-offs	Sensitivity, specificity at the cut-offs in reported literature		Sensitivity, specificity at recommended cut-offs	90% sensitivity and 90% specificity cut-offs
Liver stiffness measurement (LSM)	< 7.4, to rule out [24]	< 7.4: 90%, 60% [24]	0.85 [24]	< 7.4: 89%, 57%	< 7.3, to rule out
	< 8.2, to rule out [28]	< 8.2: 90%, 61% [29]	0.86 [29]	< 8.2: 88%, 64%	
	> 12.1, to rule in [24]	55%, 95% [24]		57%, 89%	≥ 12.7 to rule in
FIB4	< 1.30, to rule out [21, 30, 28]	84%, 68% [10] 74%, 71% [25]	0.84 [10] 0.80 [10]	66%, 61%	< 0.66, to rule out
	> 2.67, to rule in [21]	> 2.67:	0.87 [26]	23%, 92% at > 2.67	≥ 2.45, to rule in
	> 3.25, to rule in [30, 28]	33,98% [25]	0.85 [30]	17%, 96% at > 3.25	
		> 3.25: 38%, 96% [10] 48%, 95% [26] 18, 100% [30]	0.71 [31] 0.70 [32]		
NAFLD Fibrosis score (NFS)	< – 1.45, to rule out [21], [30], [28]	82%, 77% [9] 78%, 58% [27] 90%, 60% [33] 39%, 81% [34] 92%, 63% [26]	0.82 [9] 0.81 [27] 0.85 [33] 0.64 [34] 0.86 [30]	65%, 66%	< – 2.97, to rule out
	≥ 0.67, to rule in [21], [30, 28]	51%, 98% [9] 33%, 98% [27] 67%, 97% [33] 0%, 99% [34] 9, 98% [30] 33, 96% [26]	0.65 [32]	14%, 93%	≥ 0.29, to rule in
FAST* (n = 609)	< 0.35, to rule out [23] > 0.67, to rule in [23]	89%, 64% [23] 49%, 92% [23]	0.80 [23]	90%, 34% 57%, 82%	≤ 0.35, to rule out > 0.72, to rule in
Agile 3+	< 0.45, to rule out [22]** > 0.68, to rule in [22]	87%, 77% [22] 69%, 91% [27]	0.90 [22]	70%, 78% 54%, 95%	≤ 0.16, to rule out ≥ 0.70, to rule in

*FAST score was developed and validated to identify fibrotic NASH which is defined as NAS ≥ 4 and fibrosis ≥ F2, instead of advanced fibrosis (≥ F3 fibrosis)

**85% sensitivity threshold was evaluated in the publication

Table 5 Diagnostic accuracy of individual non-invasive tests and two-step including biomarker-based test followed by liver stiffness measurement

	Rule in and out criteria	Number of patients classified (<i>n</i>)	Sensitivity (%)	Specificity (%)	PPV(%)	NPV(%)	Correctly classified (%)	Intermediate Unclassified, (%)	Incorrectly classified (%)
LSM	<7.3, ≥12.7	TP: 87, TN: 268, FP:49, FN:16	84.5	84.5	64	94.4	55.4	34.5	10.1
FIB4	<0.66, ≥2.45	TP: 44, TN: 98, FP:48, FN:15	74.6	67.1	25.3	86.7	22.2	68	9.8
NFS	< -2.97, ≥0.29	TP: 28, TN: 139, FP:47, FN:16	63.6	74.7	22	89.7	26.1	64.1	9.8
Agile 3+	<0.16, ≥0.7	TP: 81, TN: 245, FP:48, FN:17	82.7	83.6	12.8	93.5	50.9	39	10.1
FIB4 (rule in and rule out) → LSM	FIB4 <0.66 or if FIB4 0.66-2.45: LSM <7.3, FIB4 ≥2.09 or if FIB4 0.61-2.08: LSM ≥12.5	TP: 103, TN: 281, FP:81, FN:25	80.5	77.6	17.4	91.8	59.9	23.6	16.5
FIB4 (rule out criteria only) → LSM	FIB4 <0.66 or if FIB4 ≥0.61: LSM <7.3; For FIB4 ≥0.66: LSM ≥12.7	TP: 85, TN: 303, FP:43, FN:28	75.2	87.6	10	91.5	60.5	28.4	11.1
NFS (rule in and rule out) → LSM	NFS < -2.97 or if NFS -2.97 - 0.29: LSM <7.3, NFS ≥0.29 or if NFS -2.97 - 0.29: LSM ≥12.7	TP: 88, TN: 299, FP:79, FN:25	77.9	79.1	17	92.3	60.4	23.4	16.2
NFS (rule out criteria only) → LSM	NFS < -2.97 or if NFS ≥ -2.97: LSM <7.3; For NFS ≥0.29: LSM ≥12.7	TP: 83, TN: 317, FP:43, FN:25	76.9	88.1	9.7	92.7	62.4	27	10.6

LSM liver stiffness measurement (transient elastography), NFS NAFLD fibrosis score, EFS Expanded FIB4 score, TP true positive, TN true negative, FP false positive, FN false negative, PPV positive predictive value, NPV negative predictive value

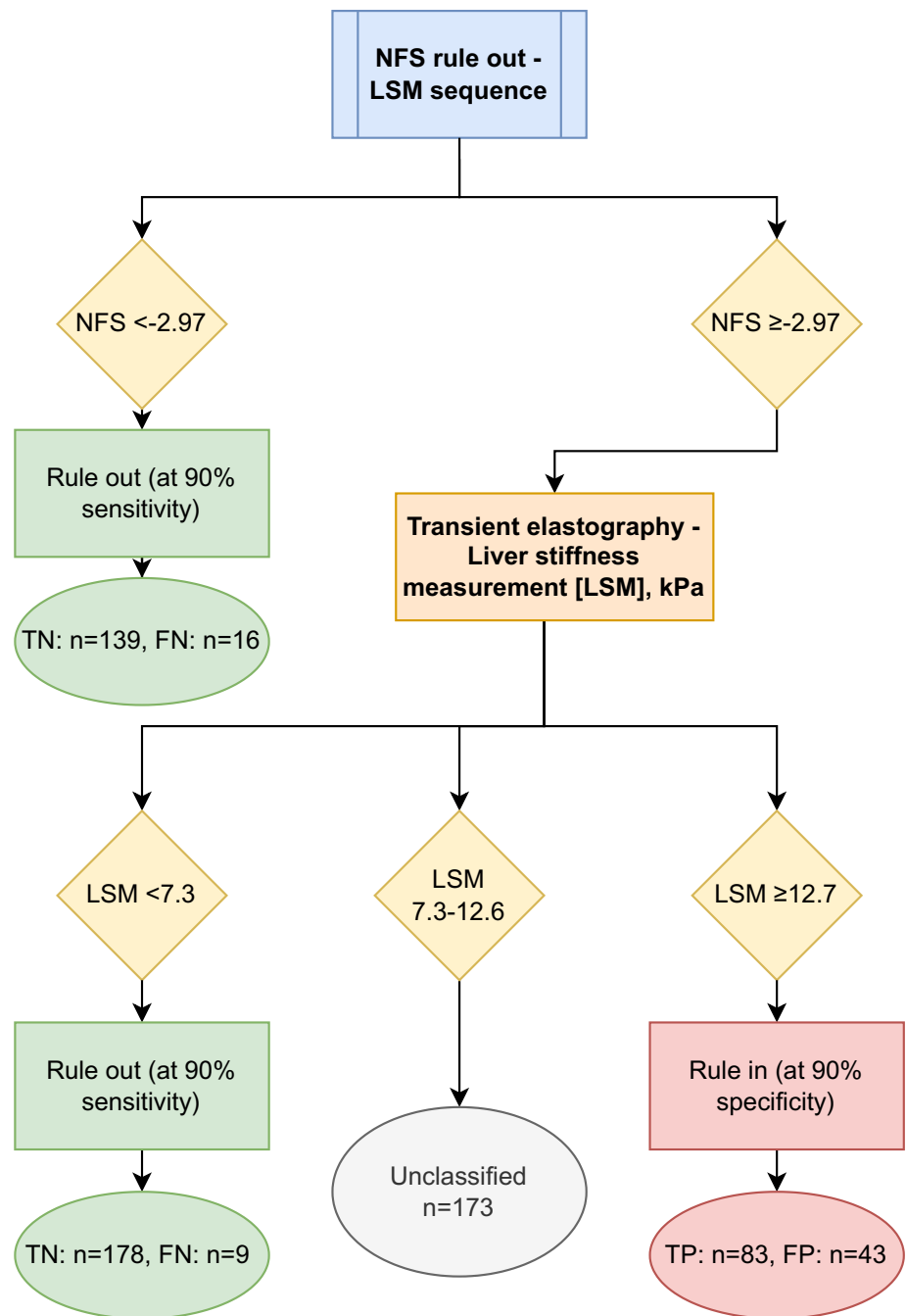
may avoid referral for TE as per this algorithm. Similarly, for FIB4 (rule out at <0.66) followed by LSM criteria as above, correctly classified and indeterminate cases were 60.5% and 28.4%, respectively.

At currently recommended threshold of FIB4 < 1.3 (rule out) and ≥2.67 (rule in) for example, the classification yielded number of TP: 37, TN: 296, FP: 39, FN: 57, and corresponding correctly classified, indeterminate, and incorrectly classified proportions of 52%, 15%, and 33%, respectively. Similarly, the recommended NFS criteria

at < -1.45 (rule out) and 0.67 (rule in) yielded TP: 23 TN: 320, FP: 33, FN: 56, and corresponding correctly classified, indeterminate, and incorrectly classified proportions of 55%, 14%, and 31%, respectively.

The clinical characteristics as well as measures of NITs of the 3 cohorts which could not be included due to lack of test-positive (no patients with advanced fibrosis) or non-availability of TE data are shown in Supplementary Table 4. Available biomarker-based NITs in the cohort from China (*n* = 67) demonstrated poor diagnostic accuracy (Supplementary Table 5).

Fig. 3 Algorithm for the non-invasive assessment of advanced fibrosis in patients of NAFLD, using NFS rule out criteria followed by liver stiffness measurement (LSM)



Discussion

Advanced fibrosis is an important prognostic factor in NAFLD, correlating with liver-related outcomes and mortality. Considering the invasive nature of liver biopsies and the increasing global prevalence of NAFLD, use of NITs has become widespread to predict AF and refer patients to specialists. Despite abundant information from the West suggesting sufficient diagnostic accuracy of existing NITs, there is sparse validation of the same from the East. We found poor diagnostic accuracy of currently used biomarker-based NITs (AUROC <0.7)

at cut-offs recommended by existing society guidelines in an analysis of four cohorts from Asia. LSM measured by TE (and Agile 3+ score) is the most accurate diagnostic tests for identifying AF. A two-step model using either FIB4 or NFS (rule out criteria only)-LSM sequence accurately classified nearly two-thirds of patients with AF.

Factors Affecting Diagnostic Accuracy of Biomarker-Based NITs

All currently used biomarker-based NITs had poor diagnostic accuracy in the four cohorts from 3 centers (two from India and one from Singapore) in the main analysis, as well as the Chinese cohort. These tests are dependent on several factors. First, NITs have been shown to have poor accuracy in Asian populations. Even within a particular region, these tests were less accurate among patients of South Asian descent [30]. A post hoc analysis found similar diagnostic accuracy for FIB4 and NFS between White and Asian populations (AUROC between 0.73 and 0.80) but did not analyze South Asians separately [28]. Second, obesity and BMI influence NITs, as shown by a study on 584 patients with a mean BMI of 45 kg/m² that reported AUROC for NFS, FIB4, and APRI to be below 0.62 [29]. Our results suggest worse performance of certain NITs (NFS and APRI) among patients with higher BMI categories (Table 3). Third, increasing age is associated with an increase in liver stiffness, which further confounds the expected presence of NASH-related fibrosis, and a higher FIB4 threshold has been proposed for those over 65 years of age [35]. Fourth, the NITs are calculated from parameters (AST, ALT, and platelet counts) with significant intra-individual variation and thus measurement error [36]. Lastly, population prevalence and pre-test probability immensely affect the positive and negative predictive values (PPV, NPV) of a diagnostic test—this is particularly true for AF in NASH, where most tests show a high NPV and a low PPV [37, 38]. In a population-based study from Hong Kong among 922 individuals with a 2% prevalence of AF, a negative FIB4 or NFS had a high NPV (98%) for AF, but sensitivity was low (7–13%). [39]

Choice Among NITs

LSM has the highest AUROC, in patients from Asia (0.91), Europe and America (0.89), and is relatively unaffected by BMI, similar to the findings of our study [40]. This BMI independence of LSM is contingent on using appropriate XL probes patients with BMI \geq 30 kg/m² [41]. Our results support that LSM is more accurate than all biomarker-based NITs in detecting AF but the cost and limited availability of TE may preclude its universal implementation among all NAFLD patients. Second-best choice among NITs were FIB4 and NFS, with low accuracy (AUROC < 0.7). Another panel, the easy liver fibrosis test (eLIFT) has been developed for ‘bedside use’ based on simple parameters (age, gender, AST, GGT, platelet count, and prothrombin time) for population screening [42]. However, it contains multiple parameters and is thus difficult to use bedside. Its accuracy was similar compared to FIB4 (AUROC 0.78 vs 0.79, $p=0.42$). GGT

was not measured in our study and thus eLIFT could not be assessed. Recommended FIB4 cut-offs in Asians have poor sensitivity and specificity, and the 90% sensitivity and specificity cut-offs should be 0.66 and 2.45. The use of the existing FIB4 cut-offs defined in literature will lead to low positive predictive value (10.5%) in Asian cohorts; although even with modified cut-offs, the PPV remains low (25.3%). Importantly, there was a high proportion of incorrectly classified patients (> 30%) for AF when existing guideline criteria were adopted for Asian patients for FIB4 and NFS.

Two-Step Diagnostic Algorithm

Each diagnostic test that utilizes a rule-out and rule-in cut-off, which creates a gray zone of unclassified individuals where the diagnosis remains indeterminate. All biomarker-based NITs have an unacceptably high proportion of patients (> 50%) that cannot be classified using the rule-in or rule-out criteria alone. A combination of 2 NITs can help further assess and reduce the unclassified proportion, the proportion of cases not allotted a diagnostic label [43, 44]. Two-step application of FIB4 followed by ELFTM improves the detection of AF five-fold and reduces referrals by 80% [26, 44]. Boursier et al. compared sequential application of NITs and found that an algorithm with TE or FIB4 as the first test followed by FibroMeter^{VCTE} carried high diagnostic accuracy, correctly classifying 90% of their population for the presence of AF [43, 45]. Unlike the present study, their population was derived from French centers and most likely included a predominant Caucasian ethnicity. We did not assess ELF, FibroMeter, and FibroTest that require proprietary fibrosis markers beyond routine blood tests. The GO ASIA working group has also shown utility of a two-step approach of NFS or FIB4 followed by LSM in an Asian setting [46].

In our population, we found the highest diagnostic accuracy with a sequence of NFS followed by TE to identify AF. NFS (rule out)-LSM sequence could correctly classify 62.4% of the population, while 27% remained unclassified. The advantage of this two-step algorithm over LSM alone is a reduction of unclassified cases from 34.5 to 27%. Thus, an NFS of > -2.97 can be used as a referral basis for TE in resource-limited settings—reducing referral for approximately 25% of patients with NFS below this threshold and emphasizing that NFS or other biomarker-based NITs have low specificity for ruling in AF. An algorithm of FIB4 followed by shear-wave elastography [SWE] has been reported to be cost-effective, with cost-per-individual and effectiveness being 103.93\$ and 0.92, followed by FIB4-LSM (135.84\$, 0.88), outperforming LSM (226.95\$, 0.84), FIB4 (296.2\$, 0.76), and NFS (541.1\$, 0.67) [47].

How to Approach an ‘Unclassified’ Reading?

A classification that yielded an indeterminate result, such as NFS above -2.97 along with LSM between 7.3 and 12.7 , was found in 27% of patients using the NFS (rule-out) followed by LSM algorithm. Such test results pose difficulty to the physician while planning further evaluation and management. A few options exist—first, a liver biopsy may be recommended to confirm or refute AF. In real life, less than a quarter of gastroenterologists perform a liver biopsy in patients with NAFLD and more than half initiate therapy with vitamin E without biopsy confirmation of NASH [48]. Liver biopsy rates in clinical practice outside referral centers are likely to be even lower. Second, the physician may follow-up after a period of time, given the slow progression of fibrosis in these patients. An increase in test parameters meeting rule-in criteria may clarify the treatment plan. Alternatively, an increase in test values over time itself can indicate an increased risk of adverse liver events. For example, an increase by 1 point value in FIB4 or transitioning from low-risk to high-risk value was associated with an increased risk of liver events with a hazards ratio of 1.81 and 7.99 , respectively [49]. Third option is to consider that the patient has AF and initiate lifestyle changes for therapeutic benefit, since this will likely reduce the cardiovascular risk and prevent the development of type-2 diabetes mellitus, in addition to reducing liver-related events.

Combination Tests

FibroScan-based tests such as Agile 3+ , Agile 4, and FAST score have been used to accurately diagnose AF (\geq F3 fibrosis), cirrhosis (F4 fibrosis), and fibrotic NASH, respectively. [22, 23] They help to reduce the number of patients in unclassified category. Sanyal et al. found Agile 3+ to improve AUROC for diagnosis of AF to 0.90 which was superior to LSM (vs 0.86 , $p < 0.001$), and reduce proportion of indeterminate patients from 23 to 13% . [22] However, we found LSM and Agile 3+ to have similar diagnostic accuracy (0.82 each, $p = 0.438$) as well as a similar proportion of unclassified patients (34% and 39%), respectively. Thus, in our Asian cohort, Agile 3+ , an equation derived and validated from a predominantly Caucasian population did not yield incremental benefit over using LSM alone.

Strengths and Limitations

Our study analyzed a large sample size from 4 different cohorts to determine the diagnostic accuracy of NITs in detecting the presence of AF in biopsy-proven NAFLD. However, it had a few limitations: ours was a retrospective study that may have been prone to selection bias in choosing patients for liver biopsy. Second, liver biopsy is an imperfect

gold standard, with inter-observer and intra-observer variability, particularly due to different pathologists at each center interpreting fibrosis stage. We did not assess NITs that depend on proprietary tests, since they have limited availability in resource-limited settings.

In conclusion, our study found that the established biomarker-based non-invasive tests perform poorly in Asian settings. LSM is the single most accurate non-invasive test to detect the presence of advanced fibrosis in NAFLD patients. Combination tests of LSM and blood investigations such as Agile 3+ and FAST scores have high AUROC but do not provide any additional diagnostic accuracy compared to LSM alone. Sequential application of NFS (rule-out)-LSM can correctly classify advanced fibrosis in two-thirds of cases. This reduces the proportion of unclassified cases and potentially avoid costs of TE in one-fourth of cases in whom NFS can rule out AF with 90% specificity.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10620-023-08085-y>.

Declarations

Conflict of interest All authors declare that there are no conflicts of interest to disclose.

References

1. Barritt AS, Watkins S, Gitlin N, Klein S, Lok AS, Loomba R et al. Patient Determinants for Histologic Diagnosis of NAFLD in the Real World: A TARGET-NASH Study. *Hepatol Commun*. 2021;5:938–946.
2. Kumar R, Priyadarshi RN, Anand U. Non-alcoholic Fatty Liver Disease: Growing Burden, Adverse Outcomes and Associations. *J Clin Transl Hepatol*. 2020 Mar 28;8:76–86.
3. Shalimar, Elhence A, Bansal B, Gupta H, Anand A, Singh TP, et al. Prevalence of Non-alcoholic Fatty Liver Disease in India: A Systematic Review and Meta-analysis. *J Clin Exp Hepatol*. 2022 May;12:818–29.
4. Riazi K, Azhari H, Charette JH, Underwood FE, King JA, Afshar EE et al. The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol*. 2022 Sep;7:851–861.
5. Blais P, Husain N, Kramer JR, Kowalkowski M, El-Serag H, Kanwal F. Nonalcoholic fatty liver disease is underrecognized in the primary care setting. *Am J Gastroenterol*. 2015 Jan;110:10–14.
6. Castera L. Non-invasive tests for liver fibrosis in NAFLD: Creating pathways between primary healthcare and liver clinics. *Liver International*. 2020 Feb;40:77–81.
7. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. *J Hepatol*. 2018 Feb;68:305–315.
8. Papagianni M, Sofogianni A, Tziomalos K. Non-invasive methods for the diagnosis of nonalcoholic fatty liver disease. *World J Hepatol*. 2015 Apr 8;7:638–648.
9. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology*. 2007 Apr;45:846–854.


10. Sun W, Cui H, Li N, Wei Y, Lai S, Yang Y et al. Comparison of FIB-4 index, NAFLD fibrosis score and BARD score for prediction of advanced fibrosis in adult patients with non-alcoholic fatty liver disease: A meta-analysis study. *Hepatol Res.* 2016 Aug;46:862–870.
11. Cassinotto C, Boursier J, de Lédinghen V, Lebigot J, Lapuyade B, Cales P et al. Liver stiffness in nonalcoholic fatty liver disease: A comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. *Hepatology.* 2016 Jun;63:1817–1827.
12. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology.* 2012 Jun;55:2005–2023.
13. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* 2005 Jun;41:1313–1321.
14. Powell EE, Wong VWS, Rinella M. Non-alcoholic fatty liver disease. *Lancet.* 2021 Jun 5;397:2212–2224.
15. Rout G, Kedia S, Nayak B, Yadav R, Das P, Acharya SK et al. Controlled Attenuation Parameter for Assessment of Hepatic Steatosis in Indian Patients. *J Clin Exp Hepatol.* 2019 Feb;9:13–21.
16. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology.* 2006 Jun;43:1317–1325.
17. 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 Oct 1;57:1441–1447.
18. Catanzaro R, Aleo A, Sciuto M, Zanoli L, Balakrishnan B, Marotta F. FIB-4 and APRI scores for predicting severe liver fibrosis in chronic hepatitis HCV patients: a monocentric retrospective study. *Clin Exp Hepatol.* 2021 Mar;7:111–116.
19. Shiha G, Ibrahim A, Helmy A, Sarin SK, Omata M, Kumar A et al. Asian-Pacific Association for the Study of the Liver (APASL) consensus guidelines on invasive and non-invasive assessment of hepatic fibrosis: a 2016 update. *Hepatol Int.* 2017 Jan;11:1–30.
20. 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 Sep;75:659–689.
21. Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases: *Hepatology*, Vol. XX, No. X, 2017. *Hepatology.* 2018 Jan;67:328–57.
22. Sanyal AJ, Foucquier J, Younossi ZM, Harrison SA, Newsome PN, Chan WK, et al. Enhanced diagnosis of advanced fibrosis and cirrhosis in individuals with NAFLD using FibroScan-based Agile scores. *Journal of Hepatology.* 2023 Feb;78(2):247–59.
23. Newsome PN, Sasso M, Deeks JJ, Paredes A, Boursier J, Chan WK et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol.* 2020 Apr 1;5:362–373.
24. Mózes FE, Lee JA, Selvaraj EA, Jayaswal ANA, Trauner M, Boursier J et al. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. *Gut.* 2022 May 1;71:1006–1019.
25. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol.* 2009 Oct;7:1104–1112.
26. Sumida Y, Yoneda M, Hyogo H, Itoh Y, Ono M, Fujii H et al. Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. *BMC Gastroenterol.* 2012 Jan;5:2.
27. McPherson S, Stewart SF, Henderson E, Burt AD, Day CP. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut.* 2010 Sep;59:1265–1269.
28. Wong VWS, Tak WY, Goh GBB, Cheng PN, Lawitz EJ, Younossi ZM, et al. Performance of Noninvasive Tests of Fibrosis Among Asians, Hispanic, and non-Hispanic Whites in the STELLAR Trials. *Clin Gastroenterol Hepatol.* 2022 Jan 22;
29. Alqahtani SA, Golabi P, Paik JM, Lam B, Moazez AH, Elariny HA et al. Performance of Noninvasive Liver Fibrosis Tests in Morbidly Obese Patients with Nonalcoholic Fatty Liver Disease. *Obes Surg.* 2021;31:2002–2010.
30. De Silva S, Li W, Kemos P, Brindley JH, Mecci J, Samsuddin S et al. Non-invasive markers of liver fibrosis in fatty liver disease are unreliable in people of South Asian descent. *Frontline Gastroenterol.* 2018 Apr;9:115–121.
31. Kumari B, Kumar R, Sharma S, Banerjee A, Kumar V, Kumar P et al. Diagnostic Accuracy of FIB-4 and FIB-5 Scores as Compared to Fibroscan for Assessment of Liver Fibrosis in Patients With Non-Alcoholic Fatty Liver Disease. *Cureus.* 2021 Aug;13:e17622.
32. Zambrano-Huaila R, Guedes L, Stefano JT, de Souza AAA, Marciano S, Yvamoto E et al. Diagnostic performance of three non-invasive fibrosis scores (Hepamet, FIB-4, NAFLD fibrosis score) in NAFLD patients from a mixed Latin American population. *Ann Hepatol.* 2020 Nov 1;19:622–626.
33. Musso G, Gambino R, Cassader M, Pagano G. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. *Ann Med.* 2011 Dec;43:617–649.
34. Wong VWS, Wong GLH, Chim AML, Tse AML, Tsang SWC, Hui AY et al. Validation of the NAFLD fibrosis score in a Chinese population with low prevalence of advanced fibrosis. *Am J Gastroenterol.* 2008 Jul;103:1682–1688.
35. McPherson S, Hardy T, Dufour JF, Petta S, Romero-Gomez M, Allison M et al. Age as a Confounding Factor for the Accurate Non-Invasive Diagnosis of Advanced NAFLD Fibrosis. *Am J Gastroenterol.* 2017 May;112:740–751.
36. Banfi G, Fabbro MD. Biological Variation in Tests of Hemostasis. *Semin Thromb Hemost.* 2009 Feb;35:119–126.
37. Hernaez R, Thrift AP. High Negative Predictive Value, Low Prevalence, and Spectrum Effect: Caution in the Interpretation. *Clin Gastroenterol Hepatol.* 2017 Sep;15:1355–1358.
38. Hallsworth K, Adams LA. Lifestyle modification in NAFLD/NASH: Facts and figures. *JHEP Rep.* 2019 Nov 5;1:468–479.
39. Mahady SE, Macaskill P, Craig JC, Wong GLH, Chu WCW, Chan HLY et al. Diagnostic Accuracy of Noninvasive Fibrosis Scores in a Population of Individuals With a Low Prevalence of Fibrosis. *Clin Gastroenterol Hepatol.* 2017 Sep;15:1453-1460.e1.
40. Cao Y tian, Xiang L lan, Qi F, Zhang Y juan, Chen Y, Zhou X qiao. Accuracy of controlled attenuation parameter (CAP) and liver stiffness measurement (LSM) for assessing steatosis and fibrosis in non-alcoholic fatty liver disease: A systematic review and meta-analysis. *eClinicalMedicine.* 2022 Sep;51:101547.
41. Wong VWS, Irls M, Wong GLH, Shili S, Chan AWH, Merrouche W et al. Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease. *Gut.* 2019 Nov;68:2057–2064.

42. Boursier J, de Ledinghen V, Leroy V, Anty R, Francque S, Salmon D et al. A stepwise algorithm using an at-a-glance first-line test for the non-invasive diagnosis of advanced liver fibrosis and cirrhosis. *J Hepatol*. 2017 Jun;66:1158–1165.
43. Boursier J, de Ledinghen V, Zarski JP, Fouchard-Hubert I, Gallois Y, Oberti F et al. Comparison of eight diagnostic algorithms for liver fibrosis in hepatitis C: new algorithms are more precise and entirely noninvasive. *Hepatology*. 2012;55:58–67.
44. Srivastava A, Gailer R, Tanwar S, Trembling P, Parkes J, Rodger A et al. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. *J Hepatol*. 2019 Aug;71:371–378.
45. Boursier J, Guillaume M, Leroy V, Irlès M, Roux M, Lannes A et al. New sequential combinations of non-invasive fibrosis tests provide an accurate diagnosis of advanced fibrosis in NAFLD. *J Hepatol*. 2019 Aug 1;71:389–396.
46. Chan WK, Treeprasertsuk S, Goh GBB, Fan JG, Song MJ, Charatcharoenwitthaya P et al. Optimizing Use of Nonalcoholic Fatty Liver Disease Fibrosis Score, Fibrosis-4 Score, and Liver Stiffness Measurement to Identify Patients With Advanced Fibrosis. *Clin Gastroenterol Hepatol*. 2019 Nov;17:2570–2580.e37.
47. Congly SE, Shaheen AA, Swain MG. Modelling the cost effectiveness of non-alcoholic fatty liver disease risk stratification strategies in the community setting. *PLOS ONE*. 2021 May 21;16:e0251741.
48. Rinella ME, Lominadze Z, Loomba R, Charlton M, Neuschwander-Tetri BA, Caldwell SH et al. Practice patterns in NAFLD and NASH: real life differs from published guidelines. *Ther Adv Gastroenterol*. 2016 Jan;9:4–12.
49. 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 Nov;73(5):1023–1029.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Umang Arora¹ · Ritik M. Goyal¹ · Kevin K. J. Teh² · Yiyang Pei² · George B. B. Goh² · Su Lin³ · Rahul Kumar^{4,5} · Sagnik Biswas¹ · Manas Vaishnav¹ · Shekhar Swaroop¹ · Piyush Pathak¹ · Sabreena Sheikh¹ · Vishwesh Bharadiya¹ · Anshuman Elhence¹ · Shivanand Gamanagatti⁶ · Rajni Yadav⁷ · Prasenjit Das⁷ · Sandeep Aggarwal⁸ · Narendra Choudhary⁹ · Prajna Anirvan¹⁰ · Shivaram P. Singh¹⁰ · Arka De¹¹ · Ajay Duseja¹¹ · Shalimar¹ 

✉ Shalimar
drshalimar@aiims.edu

Umang Arora
umangarora@gmail.com

Ritik M. Goyal
ritik5678@gmail.com

Kevin K. J. Teh
kevin.teh.k.j@singhealth.com.sg

Yiyang Pei
yiyang.pei@mohh.com.sg

George B. B. Goh
goh.boon.bee@singhealth.com.sg

Su Lin
sumer5129@fjmu.edu.cn

Rahul Kumar
rahul.kumar@singhealth.com.sg

Sagnik Biswas
Sagnik_biswas@yahoo.co.in

Manas Vaishnav
manas91vaishnav@gmail.com

Shekhar Swaroop
s.swaroop123456@gmail.com

Piyush Pathak
drppathak1985@gmail.com

Sabreena Sheikh
simu201089@gmail.com

Vishwesh Bharadiya
vishweshbharadiya000@gmail.com

Anshuman Elhence
anshu27790@gmail.com

Shivanand Gamanagatti
shiv223@rediffmail.com

Rajni Yadav
drrajniyadav@gmail.com

Prasenjit Das
prasenaaiims@gmail.com

Sandeep Aggarwal
sandeep_aiims@yahoo.co.in

Narendra Choudhary
docnarendra@gmail.com

Prajna Anirvan
prajnaanirvan@gmail.com

Shivaram P. Singh
spsingh.cuttack@gmail.com

Arka De
arkascore@gmail.com

Ajay Duseja
ajayduseja@yahoo.co.in

¹ Department of Gastroenterology and Human Nutrition, All India Institute of Medical Sciences, Delhi 110029, India

² Department of Gastroenterology and Hepatology, Singapore General Hospital, Singapore, Singapore

- ³ Department of Liver Research Center, The First Affiliated Hospital of Fujian Medical University, Fuzhou, Fujian, China
- ⁴ Department of Gastroenterology and Hepatology, Changi General Hospital, Singapore, Singapore
- ⁵ Medicine Academic Clinical Programme, SingHealthDuke-NUS Academic Medical Centre, Singapore, Singapore
- ⁶ Department of Radiology, All India Institute of Medical Sciences, Delhi, India
- ⁷ Department of Pathology, All India Institute of Medical Sciences, Delhi, India
- ⁸ Department of Surgical Disciplines, All India Institute of Medical Sciences, Delhi, India
- ⁹ Department of Hepatology, Medanta the Medicity, Gurgaon, Haryana, India
- ¹⁰ Srirama Chandra Bhanja Medical College and Hospital, Cuttack, Orissa, India
- ¹¹ Department of Hepatology, Post Graduate Institute of Medical Education and Research, Chandigarh, India