Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Body mass index and fibrosis 4 index performance in metabolic dysfunction- associated steatotic liver disease

Xiaoyue Yang
The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China
Ming Xu
The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China
Yiting Qi
The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China
Xinmin Chu
The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China
Kaiguang Zhang
The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China
Jun Xing
xingjunjoy@hotmail.com

The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China

Research Article

Keywords: MASLD, FIB-4, BMI, effect modification, liver fibrosis, diagnostic performance, obesity, risk stratification

Posted Date: November 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11609816/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background and Aims: The performance of the fibrosis-4 index (FIB-4) may vary with body mass index (BMI) in metabolic dysfunction-associated steatotic liver disease (MASLD). We assessed whether BMI modified the associations of FIB-4 and its components with elevated liver stiffness and compared diagnostic performance across BMI categories.

Methods: This retrospective cross-sectional study included 20,893 adults with MASLD who underwent vibration-controlled transient elastography during health examinations. Participants were classified as normal-weight, overweight, or having obesity according to Chinese BMI criteria. The primary outcome was liver stiffness measurement (LSM) ≥8.0 kPa. Associations and BMI interactions were assessed using multivariable logistic regression. Diagnostic performance was assessed using the area under the receiver operating characteristic curve (AUC) and a FIB-4 threshold of 1.30.

Results: The prevalence of LSM ≥8.0 kPa was 5.9%, 11.4%, and 29.6% in the normal-weight, overweight, and obesity groups, respectively. BMI interactions were significant for age and platelet count (both P<0.001), but not for aspartate or alanine aminotransferase (P=0.063 and P=0.167). The adjusted odds ratio per unit increase in FIB-4 was 2.02 (95% confidence interval 1.79–2.27) in the normal-weight group and 1.30 (1.19–1.41) in the obesity group (P for interaction <0.001). AUC decreased from 0.702 to 0.529, and sensitivity at FIB-4 ≥1.30 decreased from 55.5% to 20.0%.

Conclusions: BMI modified the associations of age and platelet count with elevated liver stiffness. FIB-4 had lower discrimination in participants with obesity, and the 1.30 threshold identified only one in five participants with LSM ≥8.0 kPa in this group.

Lay Summary

In adults with fatty liver disease, a commonly used blood-based score was less able to identify raised liver stiffness in people with obesity than in those of normal weight. Four in five participants with obesity who had raised liver stiffness had a score below the usual threshold for further testing. These findings suggest that body size may need to be considered when interpreting the score, although further studies are needed before changing clinical practice.

Introduction

Metabolic dysfunction-associated steatotic liver disease (MASLD) affects more than 30% of adults worldwide [1,2]. Fibrosis stage is strongly associated with long-term outcomes[3,4]. Non-invasive assessment is therefore important for identifying individuals who may require further investigation, specialist referral, or closer monitoring[5].

Non-invasive assessment is therefore important for identifying individuals who may require further investigation, specialist referral, or closer monitoring[5–7]. A value < 1.30 is generally used to identify individuals at low risk of advanced fibrosis, with age-specific considerations[5,6]. Reduced FIB-4 performance in populations with obesity has been reported[7–9]. However, comparisons of AUC alone do not establish whether the associations of individual components with fibrosis vary by BMI. ALT may be influenced by hepatic steatosis and metabolic injury[10,11], and platelet counts may be affected by obesity-associated inflammation[12]. Differences in age and metabolic disease burden may also affect the information conveyed by these variables[13]. Examining component associations alongside the performance of the composite score may help explain differences across BMI groups.

We examined whether BMI modified the associations of FIB-4 components with LSM ≥ 8.0 kPa and assessed the association and diagnostic performance of the composite score across BMI categories. We also evaluated classification at a fixed FIB-4 threshold of 1.30. This threshold is commonly used in advanced fibrosis pathways, whereas our primary outcome was VCTE-defined significant fibrosis.

Patients and methods

Study setting and design

This was a single-center retrospective cross-sectional study based on routinely collected data from a health screening program. The study protocol was approved by the Medical Research Ethics Committee of the First Affiliated Hospital of the University of Science and Technology of China (approval No. 2026-RE-108).

Study population

Adults aged ≥ 18 years who underwent abdominal ultrasonography and VCTE during routine health examinations at the First Affiliated Hospital of University of Science and Technology of China between 2023 and 2025 were screened. MASLD was defined by the multi-society Delphi consensus criteria: hepatic steatosis plus at least one cardiometabolic risk factor[14]. Hepatic steatosis was identified by CAP ≥ 248 dB/m on VCTE. Exclusion criteria: significant alcohol intake (≥ 30 g/day men, ≥ 20 g/day women), viral hepatitis (HBsAg-positive or anti-HCV-positive), other chronic liver diseases (autoimmune hepatitis, primary biliary cholangitis, Wilson disease, hemochromatosis), decompensated cirrhosis (ascites, variceal bleeding, hepatic encephalopathy), prior liver transplantation, active malignancy, incomplete data for FIB-4 or VCTE, or BMI < 18.5 kg/m². Duplicate records were excluded. The final sample comprised 20,893 participants (Fig. 1).

Clinical and laboratory assessment

Body mass index (BMI) was calculated as weight divided by height squared (kg/m²) and categorized according to Chinese criteria as normal-weight (18.5 ≤ BMI < 24 kg/m²), overweight (24 ≤ BMI < 28 kg/m²), or obesity (BMI ≥ 28 kg/m²)[15]. Hypertension was defined on the basis of a history of previously diagnosed hypertension. An elevated blood pressure reading at the health examination alone was not sufficient to classify a participant as having hypertension. Diabetes was defined as fasting glucose ≥ 7.0 mmol/L, glycated hemoglobin ≥ 6.5%, or use of glucose-lowering medication. Fasting venous blood was drawn for glucose, lipids (total cholesterol, HDL-C, LDL-C, triglycerides), liver enzymes (AST, ALT, GGT), total bilirubin, albumin, and hematologic parameters (WBC, hemoglobin, platelet count). All assays used standard automated methods. FIB-4 was calculated as: FIB − 4 = (Age × AST) / (Platelet count × $\sqrt{\mathrm{ALT}}$)[16]. VCTE was performed with FibroScan® (Echosens, Paris, France) by experienced operators blinded to clinical data. Exams with ≥ 10 valid measurements and IQR/median ≤ 30% were considered reliable [17]. Significant fibrosis was operationally defined as LSM ≥ 8.0 kPa, and advanced fibrosis as LSM ≥ 9.4 kPa. VCTE provides a non-invasive assessment of fibrosis but does not establish histological stages[18].

Statistical analysis

Continuous variables were summarized as median (IQR), and categorical variables as n (%). Comparisons across BMI groups used Kruskal–Wallis tests for continuous variables and chi-square tests for categorical variables. BMI-stratified multivariable logistic regression was used to examine associations with LSM ≥ 8.0 kPa. The prespecified covariates were age, sex, fasting glucose, HDL-C, LDL-C, triglycerides, total cholesterol, AST, ALT, total bilirubin, albumin, WBC, hemoglobin, platelet count, hypertension, diabetes, and CAP. FIB-4 was evaluated in separate models; age, ALT, and platelet count were not entered as separate covariates in these models. Results are presented as ORs with 95% CIs.

Multiplicative interaction terms between BMI category and each variable were assessed using likelihood ratio tests, with BMI entered as a three-level categorical variable. Stratum-specific ORs were estimated. Bonferroni correction was applied to the 16 component and clinical-variable interaction tests, using a significance threshold of P < 0.003. The BMI × FIB-4 interaction was evaluated separately. Restricted cubic splines with four knots were fitted within each BMI category to estimate the association between FIB-4 and the probability of LSM ≥ 8.0 kPa, with confidence intervals derived by bootstrap resampling. Discrimination of continuous FIB-4 was assessed by AUC, and AUCs were compared using the DeLong test[19]. Sensitivity, specificity, PPV, and NPV were calculated at FIB-4 ≥ 1.30. The false-negative rate was defined as the proportion of participants with LSM ≥ 8.0 kPa who had

FIB-4 < 1.30. Calibration of predicted probabilities from BMI-stratified logistic regression models was examined using loess-smoothed plots and decile-based groups. Brier scores summarized overall prediction error. Sensitivity analyses used alternative LSM thresholds (≥ 8.2, ≥ 8.5, ≥ 9.4, and ≥ 11.0 kPa), WHO BMI categories (< 25, 25 ≤ BMI < 30, and ≥ 30 kg/m²) within the cohort after exclusion of underweight participants, and exclusion of platelet counts < 100 or > 400 ×10⁹/L. Analyses were performed using Python version 3.14. Unless otherwise specified, statistical significance was defined as a two-sided P value < 0.05.

Results

Cohort characteristics

The cohort included 20,893 participants with MASLD: 5,853 with normal weight, 7,739 with overweight, and 7,301 with obesity (Fig. 1). Participants with higher BMI were younger and had higher blood pressure, CAP, LSM, AST, ALT, WBC, and hemoglobin (all P < 0.001). Fasting glucose, lipid concentrations, and other laboratory measures also differed across groups (Table 1). The prevalence of LSM ≥ 8.0 kPa increased from 5.9% to 11.4% and 29.6% across the three BMI groups (P < 0.001). In contrast, median FIB-4 was lower in participants with obesity than in those with normal weight (0.75 vs 0.96).

Table 1

Baseline characteristics of the study cohort stratified by BMI category.

Variable	Normal-weight	Overweight	Obesity	P
Demographics				
Age, years	49.00 (37.00, 58.00)	48.00 (36.00, 56.00)	41.00 (33.00, 52.00)	< 0.001
Sex, Male	4822 (82.4%)	6266 (81.0%)	6285 (86.1%)	< 0.001
Anthropometrics				
BMI, kg/m2	23.14 (22.47, 23.60)	25.90 (24.99, 26.85)	29.71 (28.70, 31.30)	< 0.001
SBP, mmHg	126.00 (115.00, 138.00)	127.00 (117.00, 138.00)	131.00 (121.00, 142.00)	< 0.001
DBP, mmHg	79.00 (72.00, 86.00)	80.00 (73.00, 87.00)	84.00 (76.00, 92.00)	< 0.001
Metabolic Profile				
Fasting glucose, mmol/L	5.21 (4.90, 5.74)	5.19 (4.86, 5.63)	5.32 (4.96, 5.86)	< 0.001
HDL-C, mmol/L	1.02 (0.91, 1.19)	1.08 (0.95, 1.26)	1.00 (0.88, 1.14)	< 0.001
LDL-C, mmol/L	3.09 (2.52, 3.62)	3.06 (2.56, 3.58)	3.13 (2.59, 3.66)	< 0.001
Triglycerides, mmol/L	2.09 (1.75, 2.74)	1.55 (1.11, 2.23)	1.88 (1.35, 2.71)	< 0.001
Total cholesterol, mmol/L	5.02 (4.45, 5.68)	4.95 (4.38, 5.58)	4.96 (4.40, 5.56)	< 0.001
Liver Function				
AST, U/L	22.00 (19.00, 27.00)	22.00 (19.00, 27.00)	25.00 (20.00, 33.00)	< 0.001
ALT, U/L	24.00 (17.62, 35.00)	24.00 (17.00, 35.00)	34.00 (22.50, 55.00)	< 0.001
Total bilirubin, umol/L	14.60 (11.70, 18.60)	14.10 (11.30, 18.00)	14.10 (11.20, 18.10)	< 0.001
Albumin, g/L	46.00 (44.40, 47.60)	45.70 (44.20, 47.30)	46.00 (44.50, 47.60)	< 0.001
Hematology				
WBC, x10^9/L	6.08 (5.26, 7.04)	6.06 (5.21, 7.06)	6.62 (5.66, 7.77)	< 0.001
Hemoglobin, g/L	151.00 (142.00, 158.00)	151.00 (142.00, 158.00)	154.00 (146.00, 162.00)	< 0.001
Platelets, x10^9/L	226.00 (192.00, 263.00)	228.00 (192.00, 265.00)	239.00 (203.00, 278.00)	< 0.001
Non-invasive Tests				
FIB-4 index	0.96 (0.65, 1.38)	0.92 (0.62, 1.31)	0.75 (0.53, 1.11)	< 0.001
LSM, kPa	5.90 (5.10, 6.70)	6.20 (5.40, 7.10)	6.90 (5.90, 8.30)	< 0.001
CAP, dB/m	261.00 (253.00, 269.00)	274.00 (262.00, 284.00)	301.00 (290.00, 314.00)	< 0.001
LSM categories, kPa				
<7.0	4806 (82.1%)	5514 (71.2%)	3705 (50.7%)	< 0.001

Data are presented as median (IQR) for continuous variables and n (%) for categorical variables. P values were derived from Kruskal–Wallis tests for continuous variables and chi-squared tests for categorical variables. BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CAP, controlled attenuation parameter; FIB-4, fibrosis-4 index; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; LSM, liver stiffness measurement. LSM categories do not indicate histologically confirmed fibrosis stages.

Variable	Normal-weight	Overweight	Obesity	P
Demographics				
7.0-7.9	699 (11.9%)	1340 (17.3%)	1437(19.7%)	--
8.0-9.3	267 (4.6%)	602 (7.8%)	1143 (15.7%)	--
9.4–10.9	40 (0.7%)	205 (2.6%)	593 (8.1%)	--
>=11.0	41 (0.7%)	78 (1.0%)	423 (5.8%)	--
Comorbidities				
Hypertension	977 (16.7%)	1055 (13.6%)	1294 (17.7%)	< 0.001
Diabetes	386 (6.6%)	240 (3.1%)	230 (3.2%)	< 0.001

Data are presented as median (IQR) for continuous variables and n (%) for categorical variables. P values were derived from Kruskal–Wallis tests for continuous variables and chi-squared tests for categorical variables. BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CAP, controlled attenuation parameter; FIB-4, fibrosis-4 index; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; LSM, liver stiffness measurement. LSM categories do not indicate histologically confirmed fibrosis stages.

BMI-stratified multivariable associations with significant fibrosis

In BMI-stratified multivariable models (Table 2), the OR for age per year was 1.041 (95% CI 1.030–1.051) in normal-weight participants, 1.035 (1.028–1.042) in overweight participants, and 1.020 (1.014–1.025) in participants with obesity. Platelet count was inversely associated with LSM ≥ 8.0 kPa in the normal-weight and overweight groups; the estimate was closer to 1.0 in the obesity group.

Table 2

BMI-stratified multivariable logistic regression for VCTE-defined significant fibrosis (LSM ≥ 8.0 kPa).

Variable	Normal-weight (n = 5853)		Overweight (n = 7739)		Obesity (n = 7301)	
	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
Age	1.041 (1.030–1.051)	< 0.001	1.035 (1.028–1.042)	< 0.001	1.020 (1.014–1.025)	< 0.001
Fasting glucose	1.106 (1.036–1.180)	0.003	1.118 (1.062–1.177)	< 0.001	1.083 (1.043–1.125)	< 0.001
HDL-C	1.449 (0.790–2.657)	0.231	0.854 (0.561–1.300)	0.462	0.715 (0.496–1.032)	0.073
LDL-C	1.414 (0.993–2.014)	0.055	1.132 (0.883–1.450)	0.329	0.986 (0.834–1.166)	0.870
Total cholesterol	0.641 (0.452–0.910)	0.013	0.890 (0.697–1.135)	0.347	1.056 (0.891–1.251)	0.529
Triglycerides	1.160 (1.047–1.285)	0.004	1.008 (0.932–1.091)	0.834	0.972 (0.921–1.026)	0.303
AST	1.035 (1.020–1.049)	< 0.001	1.018 (1.008–1.028)	< 0.001	1.009 (1.003–1.016)	0.006
ALT	0.998 (0.990–1.006)	0.573	1.009 (1.004–1.014)	< 0.001	1.009 (1.006–1.013)	< 0.001
Total bilirubin	1.023 (1.006–1.039)	0.007	1.007 (0.995–1.020)	0.255	1.002 (0.992–1.011)	0.726
Albumin	1.084 (1.037–1.134)	< 0.001	1.029 (0.999–1.059)	0.057	0.999 (0.978–1.020)	0.891
WBC	0.998 (0.920–1.083)	0.963	1.047 (0.993–1.104)	0.087	1.028 (0.994–1.063)	0.102
Hemoglobin	0.995 (0.986–1.005)	0.314	1.001 (0.995–1.007)	0.771	1.001 (0.996–1.006)	0.624
Platelets	0.996 (0.993–0.998)	< 0.001	0.996 (0.995–0.998)	< 0.001	0.999 (0.998–1.000)	0.270
Hypertension	1.590 (1.197–2.112)	0.001	1.105 (0.901–1.356)	0.339	1.133 (0.975–1.315)	0.103
Diabetes	1.001 (0.679–1.475)	0.997	1.125 (0.790–1.603)	0.514	1.414 (1.040–1.922)	0.027
CAP	1.032 (1.025–1.040)	< 0.001	1.029 (1.024–1.033)	< 0.001	1.033 (1.030–1.036)	< 0.001

BMI-stratified models included age, sex, fasting glucose, HDL-C, LDL-C, triglycerides, total cholesterol, AST, ALT, total bilirubin, albumin, WBC, hemoglobin, platelet count, hypertension, diabetes, and CAP. Sex was included as an adjustment variable but is not displayed. ORs are expressed per one-unit increase in the units reported in Table 1, except hypertension and diabetes, which compare presence with absence. FIB-4 was evaluated in separate models. Results are presented as ORs with 95% CIs.

ALT was not associated with LSM ≥ 8.0 kPa in normal-weight participants (OR 0.998; 95% CI 0.990–1.006), but was positively associated in overweight participants (1.009; 1.004–1.014) and participants with obesity (1.009; 1.006–1.013). These within-group results do not by themselves demonstrate a difference between BMI groups; formal interaction results are reported in Table 3.

Table 3

Effect modification by BMI category: multiplicative interaction analysis.

Variable	Normal-weight OR (95% CI)	Overweight OR (95% CI)	Obesity OR (95% CI)	P for interaction
Age	1.041 (1.029,1.052)	1.035 (1.028,1.043)	1.020 (1.014,1.026)	< 0.001
Fasting glucose	1.106 (1.036,1.181)	1.119 (1.063,1.178)	1.087 (1.046,1.129)	0.121
HDL-C	1.441 (0.780,2.662)	0.876 (0.572,1.341)	0.765 (0.526,1.113)	0.246
LDL-C	1.395 (0.980,1.985)	1.130 (0.881,1.450)	0.988 (0.834,1.169)	0.025
Total cholesterol	0.654 (0.461,0.927)	0.892 (0.699,1.139)	1.055 (0.890,1.251)	0.295
Triglycerides	1.156 (1.043,1.281)	1.008 (0.932,1.090)	0.973 (0.922,1.027)	0.298
AST	1.035 (1.020,1.050)	1.018 (1.008,1.028)	1.009 (1.003,1.016)	0.063
ALT	0.998 (0.990,1.006)	1.009 (1.004,1.014)	1.009 (1.006,1.013)	0.167
Total bilirubin	1.022 (1.006,1.039)	1.007 (0.994,1.020)	1.002 (0.992,1.011)	0.005
Albumin	1.086 (1.029,1.145)	1.030 (0.993,1.068)	0.997 (0.971,1.024)	0.892
WBC	0.996 (0.918,1.081)	1.046 (0.992,1.103)	1.030 (0.996,1.065)	0.076
Hemoglobin	0.993 (0.982,1.005)	0.999 (0.991,1.007)	0.998 (0.992,1.004)	0.281
Platelets	0.996 (0.993,0.998)	0.996 (0.995,0.998)	1.000 (0.998,1.001)	< 0.001
Hypertension	1.574 (1.182,2.096)	1.089 (0.885,1.341)	1.106 (0.950,1.288)	< 0.001
Diabetes	0.994 (0.674,1.467)	1.128 (0.792,1.608)	1.388 (1.020,1.887)	0.807
CAP	1.033 (1.023,1.043)	1.029 (1.024,1.034)	1.033 (1.030,1.036)	0.092

Interaction terms between BMI category and each variable were tested using likelihood ratio tests, with BMI entered as a three-level categorical variable. The 16 interaction tests were interpreted using a Bonferroni-corrected significance threshold of P < 0.003. Stratum-specific ORs with 95% CIs are shown.

BMI modifies the association between clinical variables and significant fibrosis

After correction for multiple testing, BMI interactions were significant for age and platelet count among the FIB-4 components (both P for interaction < 0.001; Table 3). In the interaction models, the age ORs per year were 1.041, 1.035, and 1.020 across the normal-weight, overweight, and obesity groups. The corresponding platelet-count ORs per 10^9/L were 0.996, 0.996, and 1.000. Interactions were not significant for AST (P = 0.063) or ALT (P = 0.167).

Among the other clinical variables, hypertension showed a significant BMI interaction (P < 0.001). LDL-C (P = 0.025) and total bilirubin (P = 0.005) did not meet the Bonferroni-corrected threshold. None of the remaining variables showed a significant interaction after correction, including diabetes (P = 0.807). These results indicated BMI-related heterogeneity in selected component associations rather than a consistent change in all FIB-4 components.

The composite FIB-4–fibrosis association is modified by BMI

The adjusted spline curves showed different associations between FIB-4 and the probability of LSM ≥ 8.0 kPa across BMI categories (Fig. 2). Within the overlapping range shown, predicted probabilities were higher in participants with obesity than in those with normal weight. At FIB-4 = 1.30, the model-based probabilities were 5.1%, 11.3%, and 27.2% in the normal-weight, overweight, and obesity groups, respectively. In the linear FIB-4 models, the OR per one-unit increase was 2.02 (95% CI 1.79–2.27), 1.80 (1.64–1.99), and 1.30 (1.19–1.41), respectively (P for interaction < 0.001).

Diagnostic performance of FIB-4 across BMI categories

Discrimination of FIB-4 decreased across BMI groups (Fig. 3, Table 4). AUC was 0.702 (95% CI 0.674–0.732) in normal-weight participants, 0.623 (0.603–0.642) in overweight participants, and 0.529 (0.515–0.543) in participants with obesity (all pairwise P < 0.001, DeLong test). At FIB-4 ≥ 1.30, sensitivity was 55.5%, 41.9%, and 20.0%, whereas specificity was 73.2%, 76.6%, and 84.9%, respectively. PPV was 11.6%, 18.8%, and 35.7%, and NPV was 96.3%, 91.1%, and 71.6%. The variation in predictive values should be interpreted in the context of the different outcome prevalences.

Table 4

Discrimination of FIB-4 and diagnostic performance at a threshold of 1.30 across BMI categories.

BMI group	N	LSM ≥ 8.0 kPa, n (%)	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Detected (n)	Missed (n)	Missed rate
Normal-weight	5853	348 (5.9%)	0.702 (0.674, 0.732)	55.5%	73.2%	11.6%	96.3%	193	155	44.5%
Overweight	7739	885 (11.4%)	0.623 (0.603, 0.642)	41.9%	76.6%	18.8%	91.1%	371	514	58.1%
Obesity	7301	2159 (29.6%)	0.529 (0.515, 0.543)	20.0%	84.9%	35.7%	71.6%	432	1727	80.0%

AUC was calculated from continuous FIB-4 values. Sensitivity, specificity, PPV, and NPV were calculated using FIB-4 ≥ 1.30 as a positive result. Missed cases were participants with LSM ≥ 8.0 kPa and FIB-4 < 1.30. The missed rate is the false-negative rate, calculated as missed cases divided by all participants with LSM ≥ 8.0 kPa in each BMI group. AUCs were compared using the DeLong test. AUC, area under the receiver operating characteristic curve; NPV, negative predictive value; PPV, positive predictive value.

Calibration plots for the BMI-stratified probability models are shown in Fig. 4. In participants with obesity, the decile-based predicted probabilities were concentrated within a relatively narrow range. Brier scores were 0.054, 0.099, and 0.207 across the normal-weight, overweight, and obesity groups. Because outcome prevalence differed substantially, these unscaled scores should not be interpreted as a direct comparison of calibration across groups. The false-negative rates at FIB-4 < 1.30 were 44.5% (155/348), 58.1% (514/885), and 80.0% (1,727/2,159), respectively (Table 4).

Sensitivity analyses

The findings were consistent across sensitivity analyses. Using WHO BMI categories, discrimination declined across groups and the BMI × FIB-4 interaction remained significant (P < 0.001). FIB-4 had the lowest discrimination in the obesity group at each alternative LSM threshold (≥ 8.2, ≥ 8.5, ≥ 9.4, and ≥ 11.0 kPa). Excluding extreme platelet counts did not materially change the findings. Detailed results are provided in the Supplementary Materials.

Discussion

In this hospital-based health-screening cohort of 20,893 adults with MASLD, FIB-4 showed lower discrimination and sensitivity for LSM ≥ 8.0 kPa in participants with obesity. The association per one-unit increase in FIB-4 was also weaker in this group. Among participants with obesity who met the outcome, 80.0% had FIB-4 < 1.30, compared with 44.5% in normal-weight participants. Component analyses identified BMI interactions for age and platelet count, but not for AST or ALT. Previous studies have reported reduced FIB-4 performance in populations with obesity[7,8]. Our findings extend those observations by examining component associations alongside the composite score, although they do not establish a mechanism for the differences in performance.

The association between age and LSM ≥ 8.0 kPa was weaker at higher BMI. In the interaction models, the OR per year decreased from 1.041 in normal-weight participants to 1.020 in participants with obesity. Earlier metabolic liver injury may reduce the

relative information conveyed by chronological age, but this explanation cannot be tested in a cross-sectional analysis. The obesity group was also younger, and differences in age distribution may contribute to the observed performance of a score that includes age.

The inverse association of platelet count with the outcome was attenuated in participants with obesity. Obesity-associated inflammation may influence platelet counts independently of fibrosis[12], but the present study did not assess the pathways responsible. ALT was positively associated with the outcome in the overweight and obesity groups, yet the BMI × ALT interaction was not significant. It would therefore be premature to attribute the poorer performance of FIB-4 to a BMI-dependent change in the ALT association. Although higher ALT lowers FIB-4 when its other components are held constant, the contribution of this mathematical relationship to the observed group differences remains uncertain.

The fixed-threshold results have implications for further assessment of individuals with obesity. Guidelines commonly use FIB-4 < 1.30 to identify a low risk of advanced fibrosis[5,6], whereas our primary outcome was LSM ≥ 8.0 kPa, an operational definition of VCTE-assessed significant fibrosis. The two outcomes should not be treated as interchangeable. In the obesity group, 1,727 participants met the LSM outcome despite FIB-4 < 1.30. Lower sensitivity accompanied by higher specificity describes a change in classification at this threshold; the decline in AUC also indicates reduced discrimination. These results do not, on their own, establish systematic underprediction by a probability model. As treatment options for fibrotic MASH develop[20], identifying individuals who require further assessment remains important, but this study did not determine MASH status or treatment eligibility. The findings are consistent with reports of reduced FIB-4 performance in populations with obesity and diabetes[8,9,21,22].

The large sample allowed us to examine component associations and FIB-4 performance across three BMI categories. Assessing both helped distinguish differences in within-group associations from formal evidence of interaction. The findings for the composite score were consistent across analyses using alternative LSM thresholds, WHO BMI categories, and platelet-count exclusions.

Several limitations should be acknowledged. First, the cross-sectional design precludes causal inference and does not show that BMI-informed interpretation improves clinical outcomes. Second, fibrosis was assessed by VCTE rather than liver biopsy. The LSM categories are operational definitions, and measurement uncertainty, particularly in participants with obesity, may have affected the findings. Third, this was a single-center health-screening cohort composed predominantly of men, using BMI criteria for Chinese adults. Selection into the cohort, including exclusion for missing data, may limit generalizability. Fourth, waist circumference and body composition were unavailable. Adjustment for CAP does not exclude residual confounding by steatosis or other metabolic factors. Fifth, the BMI groups differed in age and outcome prevalence, both of which are relevant to interpretation of test performance. Finally, calibration was assessed for the BMI-stratified logistic regression models, rather than for FIB-4 itself. These models require validation in an independent cohort to determine whether their predictive performance is maintained in other populations.

Conclusion

FIB-4 was less sensitive and discriminated less well for LSM ≥ 8.0 kPa in participants with obesity than in those with normal weight. BMI modified the associations of age and platelet count with elevated liver stiffness; no significant interactions were found for AST or ALT. Among participants with obesity and LSM ≥ 8.0 kPa, 80.0% had a FIB-4 score below 1.30. Whether accounting for BMI can improve selection for further assessment of elevated liver stiffness warrants prospective evaluation.

Abbreviations

FIB-4, fibrosis-4 index; MASLD, metabolic dysfunction-associated steatotic liver disease; BMI, body mass index; LSM, liver stiffness measurement; AST, aspartate aminotransferase; ALT, alanine aminotransferase; VCTE, vibration-controlled transient elastography; CAP, controlled attenuation parameter; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; GGT, gamma-glutamyl transferase; WBC, white blood cell count; IQR, interquartile range; OR, odds ratio; CI, confidence interval; AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; WHO, World Health Organization; MASH, metabolic dysfunction-associated steatohepatitis.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Medical Research Ethics Committee of the First Affiliated Hospital of the University of Science and Technology of China (approval No. 2026-RE-108). The requirement for informed consent was waived by the Medical Research Ethics Committee of the First Affiliated Hospital of the University of Science and Technology of China due to the retrospective nature of the study. This study was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Availability of data and materials

The data supporting the findings of this study are not publicly available to protect participant privacy. De-identified data may be made available by the corresponding author upon reasonable request, subject to institutional approval and applicable data protection requirements.

Competing Interests

The authors declare no conflicts of interest that pertain to this work.

Funding

This work was supported by the Research Start-up Fund (Grant No. RC2021012), the USTC Research Fund for the New Medicine (Grant No. YD9110002074) and the Clinical Medicine Transformation Special Project of Anhui Province (Grant No. 202527c10020035).

Authors' contributions

XY: Conceptualization, Methodology, Formal analysis, Writing — original draft. MX: Investigation, Data curation, Formal analysis. YQ: Investigation, Data curation. XC: Data curation, Resources; KZ: Supervision, Project administration, Writing — review & editing. JX: Conceptualization, Supervision, Resources, Funding acquisition, Writing — review & editing.

Acknowledgements

We thank all participants who contributed to this study and the clinical staff involved in data collection and patient care.

References

1. Younossi, Z. M. *et al.* Global epidemiology of nonalcoholic fatty liver disease-meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 64, 73–84(2016).
2. Riazi, K. *et al.* The prevalence and incidence of NAFLD worldwide: A systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* 7, 851–861(2022).
3. Dulai, P. S. *et al.* Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. *Hepatology* 65, 1557–1565(2017).
4. Angulo, P. *et al.* Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* 149, 389–397.e10(2015).
5. Rinella, M. E. *et al.* AASLD practice guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology* 77, 1797–1835(2023).

6. Kanwal, F. *et al.* Clinical care pathway for the risk stratification and management of patients with nonalcoholic fatty liver disease. *Gastroenterology* 161, 1657–1669(2021).
7. Kjaergaard, M. *et al.* Using the ELF test, FIB-4 and NAFLD fibrosis score to screen the population for liver disease. *J Hepatol* 79, 277–286(2023).
8. Vali, Y. *et al.* Biomarkers for staging fibrosis and non-alcoholic steatohepatitis in non-alcoholic fatty liver disease (the LITMUS project): A comparative diagnostic accuracy study. *Lancet Gastroenterol Hepatol* 8, 714–725(2023).
9. Shaheen, A.-A. *et al.* Diabetes and obesity reduce FIB-4 accuracy in MASLD referral pathways. *JHEP Rep* 8, 101735(2026).
10. Maximos, M. *et al.* The role of liver fat and insulin resistance as determinants of plasma aminotransferase elevation in nonalcoholic fatty liver disease. *Hepatology* 61, 153–160(2015).
11. Rinella, M. E. Nonalcoholic fatty liver disease: A systematic review. *JAMA* 313, 2263–2273(2015).
12. Samocha-Bonet, D. *et al.*Platelet counts and platelet activation markers in obese subjects. *Mediators Inflamm* 2008, 834153 (2008).
13. Castera, L., Friedrich-Rust, M. & Loomba, R. Noninvasive assessment of liver disease in patients with nonalcoholic fatty liver disease. *Gastroenterology* 156, 1264–1281.e4(2019).
14. Rinella, M. E. *et al.* A multisociety delphi consensus statement on new fatty liver disease nomenclature. *J Hepatol* 79, 1542–1556(2023).
15. Zhou, B.-F. & Cooperative Meta-Analysis Group of the Working Group on Obesity in China. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults–study on optimal cut-off points of body mass index and waist circumference in chinese adults. *Biomed Environ Sci* 15, 83–96(2002).
16. Sterling, R. K. *et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43, 1317–1325(2006).
17. Castera, L., Forns, X. & Alberti, A. Non-invasive evaluation of liver fibrosis using transient elastography. *J Hepatol* 48, 835–847(2008).
18. Wong, V. W.-S. *et al.* Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology* 51, 454–462(2010).
19. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 44, 837–845(1988).
20. Harrison, S. A. *et al.* A phase 3, randomized, controlled trial of resmetirom in NASH with liver fibrosis. *N Engl J Med* 390, 497–509(2024).
21. Kim, R. G., Deng, J., Reaso, J. N., Grenert, J. P. & Khalili, M. Noninvasive fibrosis screening in fatty liver disease among vulnerable populations: Impact of diabetes and obesity on FIB-4 score accuracy. *Diabetes Care* 45, 2449–2451(2022).
22. Mózes, F. E. *et al.* Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: An individual patient data meta-analysis. *Gut* 71, 1006–1019(2022).

Figures

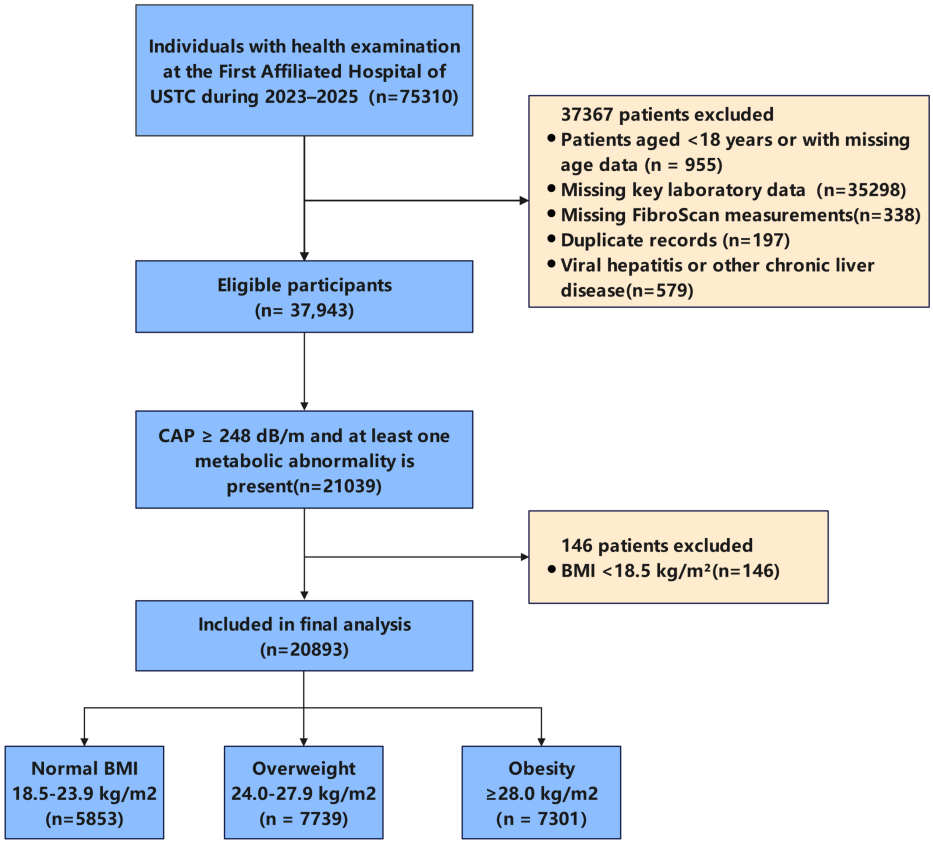


Figure 1

Study flow diagram. CAP, controlled attenuation parameter; MASLD, metabolic dysfunction-associated steatotic liver disease; VCTE, vibration-controlled transient elastography.

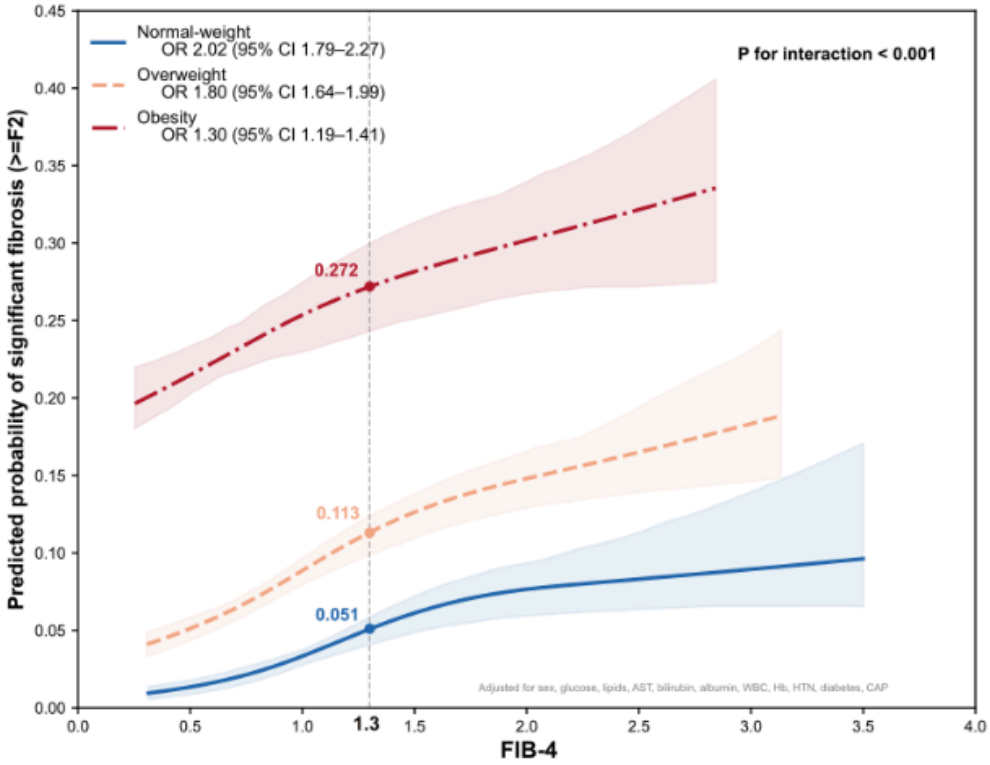


Figure 2

Model-based probability of LSM ≥8.0 kPa by FIB-4 across BMI categories, estimated using restricted cubic splines with four knots. Shaded areas represent bootstrap 95% CIs. The vertical dashed line marks FIB-4=1.30; the labels at this value show the model-based probabilities. Curves were adjusted for sex, fasting glucose, HDL-C, LDL-C, triglycerides, total cholesterol, AST, total bilirubin, albumin, WBC, hemoglobin, hypertension, diabetes, and CAP. The displayed ORs per one-unit increase in FIB-4 and their 95% CIs refer to the linear models; the BMI × FIB-4 interaction P value is also shown.

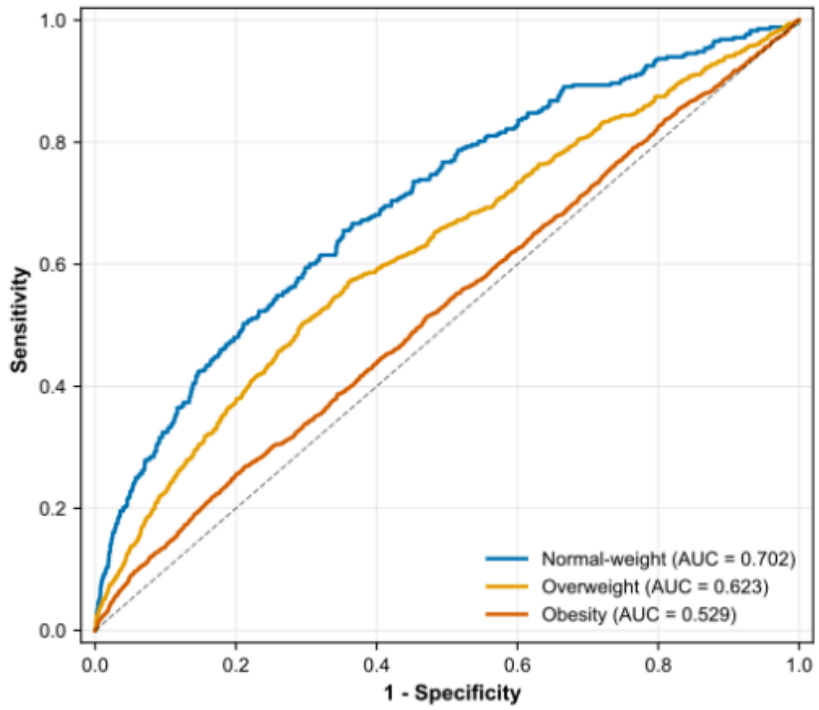


Figure 3

Receiver operating characteristic curves for continuous FIB-4 in detecting LSM ≥8.0 kPa across BMI categories. AUCs were compared using the DeLong test; all pairwise comparisons had P<0.001.

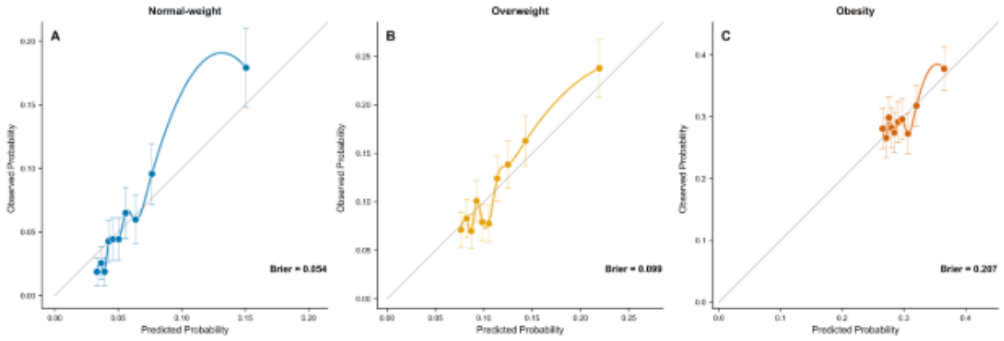


Figure 4

Observed versus predicted probabilities of LSM ≥8.0 kPa from BMI-stratified logistic regression models. (A) Normal-weight, (B) Overweight, and (C) Obesity. The diagonal line represents perfect calibration, the solid curve shows loess-smoothed calibration, and points represent decile-based groups. Brier scores are shown in each panel. Axis ranges differ between panels. These plots assess the fitted probability models rather than the raw FIB-4 score.