Original Article
Yonsei Med J 2026 Feb;67(2):104-111
https://doi.org/10.3349/ymj.2024.0318
Yonsei Medical Journal
YMJ
pISSN: 0513-5796 · eISSN: 1976-2437

Steatosis-Associated Fibrosis Estimator Score in Asian Patients with Metabolic Dysfunction-Associated Steatotic Liver Disease

Kunhee Kim[1]*, Hye Won Lee[1,2,3]*, Jae Seung Lee[1,2,3], Mi Na Kim[1,2,3], and Jun Yong Park[1,2,3]

[1]Department of Internal Medicine, Yonsei University College of Medicine, Seoul;
[2]Institute of Gastroenterology, Yonsei University College of Medicine, Seoul;
[3]Yonsei Liver Center, Severance Hospital, Seoul, Korea.

Purpose: Recently, the steatosis-associated fibrosis estimator (SAFE) score was developed to predict significant fibrosis in primary care. We externally validated the SAFE score in Asian patients with metabolic dysfunction-associated steatotic liver disease (MASLD).
Materials and Methods: We validated the SAFE score in 6229 patients who underwent transient elastography (TE) from 2012 to 2022. The sensitivities, specificities, negative predictive values, and positive predictive values of SAFE scores (two cutoffs: <0 and ≥100) for predicting fibrosis stage ≥2 were calculated.
Results: Based on TE results, the SAFE score had an area under the receiver operating characteristic curve of 0.753 (95% confidence interval 0.737–0.769), outperforming the Fibrosis-4 index (0.672) and the nonalcoholic fatty liver disease fibrosis score (0.663). Non-obese and obese patients had similar sensitivities (77.0% vs. 78.4%) and specificities (61.5% vs. 51.8%) for SAFE score <0, and similar sensitivities (50.0% vs. 50.0%) and specificities (90.1% vs. 85.4%) for SAFE score ≥100. Sensitivity of the SAFE score for ≥100 increased with age, from 16.1% (age 19–30) to 79.7% (age ≥61), whereas specificity for ≥100 decreased.
Conclusion: We externally validated the good performance of the SAFE score in Asian patients. The SAFE score has potential as an initial assessment to identify a low-risk population in a primary care setting.

Key Words: Metabolic dysfunction-associated steatotic liver disease, steatosis-associated fibrosis estimator score, fibrosis, validation

INTRODUCTION

Metabolic dysfunction-associated steatotic liver disease (MASLD) has a global prevalence of 25%–32%, and the prevalence is gradually increasing worldwide.[1,2] The prevalence of MASLD in South Korea is 30.3%–31.5%, and the incidence is 42.8 per 1000 person-years.[3-5] MASLD can evolve into more serious conditions, including metabolic dysfunction-associated steatohepatitis, liver cirrhosis, or hepatocellular carcinoma (HCC).[1] Therefore, it is important to assess the risk and prevent disease progression of MASLD. Among various prognostic factors of MASLD, fibrosis stage is one of the most important factors associated with poor clinical outcomes. Therefore, predicting fibrosis extent is important for timely referral of patients to a tertiary center.

Compared with patients with lower fibrosis stages [Fibrosis (F) 0 or F1], those with higher stages (≥F2) have an increased risk of severe liver disease, liver-related mortality, and overall mortality.[2-5] Currently, liver biopsy is the gold standard to assess liver fibrosis. However, due to the invasiveness of liver biopsy, non-invasive methods to evaluate liver fibrosis have been studied extensively, including the Fibrosis-4 (FIB-4) index, the nonalcoholic fatty liver disease (NAFLD) fibrosis score (NFS),

Received: October 8, 2024 **Revised:** June 29, 2025
Accepted: July 15, 2025 **Published online:** December 16, 2025
Corresponding author: Jun Yong Park, MD, PhD, Department of Internal Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
E-mail: drpjy@yuhs.ac

*Kunhee Kim and Hye Won Lee contributed equally to this work.
•The authors have no potential conflicts of interest to disclose.

and transient elastography (TE).[6-8] Although FIB-4 and NFS are recommended for fibrosis screening, they were not originally developed in MASLD patients. Additionally, TE is not routinely available in many primary care settings.[9]

Given the limitations, the steatosis-associated fibrosis estimator (SAFE) score was recently developed to detect significant fibrosis in primary care.[10] The SAFE score can predict significant fibrosis (≥F2) with an area under the receiver operating characteristic curve (AUROC) of 0.80–0.83. Although the SAFE score showed good performance, it was developed primarily in Caucasian populations and needs validation in Asian populations. In this study, 90% of the study population had a body mass index (BMI) over 25 kg/m^2, and there was insufficient representation of Asian patients, who have a higher rate of lean MASLD. Hence, it is essential to validate the SAFE score in Asia, especially among non-obese patients with MASLD.

Thus, we aimed to externally validate the SAFE score in Asian patients with MASLD. We also investigated the utility of the SAFE score in lean or non-obese MASLD patients. Additionally, we explored the feasibility of the SAFE score as a screening tool to identify at-risk MASLD patients in primary care.

MATERIALS AND METHODS

Study population

Between 2012 and 2022, a total of 6229 patients diagnosed with MASLD at Severance Hospital, Seoul, Korea were enrolled in this study (Fig. 1). The exclusion criteria were as follows: 1) age <19 years; 2) prior liver transplantation (LT) or prior HCC diagnosis before baseline; 3) insufficient data to calculate NFS, FIB-4, or SAFE score; 4) interquartile range (IQR) to median liver stiffness (LS) ratio >0.3. Patients with hepatitis B virus or hepatis C virus infection and excessive alcohol drinkers (male ≥210 g/week and female ≥140 g/week) were not included in the cohort.

This study was approved by the Institutional Review Board of Severance Hospital (IRB no. 4-2023-0082) and conducted according to the ethical guidelines of the 2013 Declaration of Helsinki and 2018 Declaration of Istanbul. The need for informed consent was waived because of the retrospective design of the study.

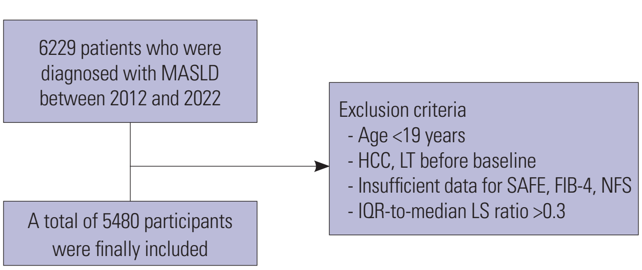

Fig. 1. Flowchart of the study population. MASLD, metabolic dysfunction-associated steatotic liver disease; HCC, hepatocellular carcinoma; LT, liver transplantation; SAFE, steatosis-associated fibrosis estimator; FIB-4, Fibrosis-4; NFS, NAFLD fibrosis score; IQR, interquartile range; LS, liver stiffness.

Diagnosis of MASLD and definition of lean, overweight, and obese patients

Steatotic liver disease was diagnosed based on image studies. MASLD was defined as steatotic liver disease and any of the cardiometabolic criteria without other causes of steatosis.[11] All patients underwent TE using FibroScan® (Echosens, Paris, France). TE was performed by expert operators and the results were noted in kilopascals (kPa). All examinations were conducted using a M or XL probe, with patients lying in a dorsal decubitus position and right arm in abduction.[12] Only LS values with at least 10 successful measurements with a success rate of >60% were recorded.[13] Clinically significant fibrosis (≥F2) was defined as LS values ≥8 kPa by TE.[14,15] Histologically confirmed fibrosis of stage F2 or greater was considered significant fibrosis.

Patients were categorized by BMI as follows, according to the definitions for Asian individuals: lean (BMI <23 kg/m^2), overweight or obese (BMI ≥23 kg/m^2), non-obese (BMI <25 kg/m^2), and obese (BMI ≥25 kg/m^2).[16]

Calculation of SAFE, FIB-4, and NFS scores and cutoff values of SAFE score

We used the following formulas to calculate the SAFE score, FIB-4, and NFS:

The SAFE score was calculated using a predefined formula incorporating age, BMI (capped at 40 kg/m^2 if >40), diabetes status (0=no, 1=yes), and log-transformed values of aspartate aminotransferase (AST), alanine aminotransferase (ALT), serum globulin, and platelet count. The complete formula is presented below.

SAFE=(2.97×age)+(5.99×BMI)+(62.85×diabetes)+[154.85×Ln (AST)]–[58.23×Ln (ALT)]+[195.48×Ln (globulins)]–[141.61×Ln (platelets)]–75.[10]

FIB-4[8] and NFS[6] were defined as in previous studies.

Impaired fasting glucose was defined as fasting glucose of 110 mg/dL or higher, according to the original article.[6]

Cutoff values of SAFE scores were set as follows: SAFE <0 (low risk), 0≤SAFE<100 (indeterminate risk), 100≤SAFE (high risk).[10]

Statistical analysis

Categorical variables are presented as numbers and percentages; continuous variables are shown as medians and IQRs, as normality assessed by the Shapiro–Wilk test was not satisfied. The baseline characteristics of the patients were compared using the Student t-test, analysis of variance, and the Pearson chi-square test. Two-sided *p*-values less than 0.05 were considered statistically significant. We used the DeLong test to calculate the 95% confidence interval (CI) for the AUROC.[17] Since SAFE <0 is used to rule out MASLD, sensitivity, specificity, and negative predictive value (NPV) were used to assess diagnostic accuracy. Similarly, since SAFE ≥100 is used to rule in MASLD, sensitivity, specificity, and positive predictive value (PPV) were used to assess diagnostic accuracy. All statistical analyses were conducted using R ver. 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Baseline characteristics

After excluding 749 patients, a total of 5480 patients were finally analyzed. The median patient age was 52.0 years, and 3236 (59.1%) were male (Table 1). 3959 (72.2%) patients were obese (BMI ≥25 kg/m^2), and 1521 (27.8%) were not obese. Based on TE data, patients with clinically significant fibrosis (n=1228, 22.4%), compared with those without significant fibrosis, had a lower proportion of males (53.2% vs. 60.7%, p=0.001), and higher rates of diabetes mellitus (DM) (48.9% vs. 35.7%, p<0.001), hypertension (HTN) (44.9% vs. 33.6%, p<0.001), and obesity (83.4% vs. 69.0%, p<0.001), a higher BMI (28.7 kg/m^2 vs. 26.6 kg/m^2, p<0.001), and higher levels of total protein, albumin, globulin, ALT, AST, γ-glutamyl transferase, creatinine, and platelet count. Moreover, we analyzed the baseline variables according to biopsy availability (Supplementary Table 1, only online). Patients who underwent biopsy had higher AST and ALT values, as well as higher LS and SAFE scores.

Risk groups based on cutoff values of SAFE score (0 and 100 points)

A total of 5480 patients were divided into three groups: 1) SAFE score <0 (47.1%, n=2583); 2) 0 ≤SAFE score <100 (31.1%, n=1702); and 3) SAFE score ≥100 (21.8%, n=1195). The proportions of DM (19.8% vs. 48.3% vs. 65.7%, p<0.001), HTN (23.9% vs. 42.1% vs. 55.3%, p<0.001), and obesity (66.9% vs. 74.7% vs. 80.3%, p<0.001) increased with higher SAFE scores. Age, BMI, total protein, globulin, ALT, AST, and γ-glutamyl transferase also increased with higher SAFE scores. In contrast, the proportion of males, albumin, and platelet count decreased with higher SAFE scores (Supplementary Table 2, only online).

SAFE score validation using TE data

Among 2568 (47.3%) patients with a SAFE score <0, 10.4% (n=262) had ≥F2 (Table 2). The sensitivity, specificity, and NPV of a SAFE score <0 for ruling out clinically significant fibrosis were 78.2%, 54.8%, and 89.6%, respectively. Among 1165 (21.5%) patients with a SAFE score ≥100, 52% (n=613) had clinically significant fibrosis (≥F2). The sensitivity, specificity, and PPV of a SAFE score ≥100 for ruling in clinically significant fibrosis were 50.0%, 86.8%, and 52.6%, respectively.

The performance of the SAFE score was compared between non-obese patients (BMI <25 kg/m^2, n=1521, 27.8%) and obese patients (BMI ≥25 kg/m^2, n=3959, 72.2%). Non-obese and obese patients had similar sensitivity (77.0% vs. 78.4%) and specificity (61.5% vs. 51.8%) for a SAFE score <0, and similar sensitivity (50.0% vs. 50.0%) and specificity (90.1% vs. 85.4%) for a SAFE score ≥100. The NPV for a SAFE score <0 was higher in non-obese than obese patients (94.5% vs. 87.1%), and the PPV for a SAFE score ≥100 was higher in obese than in non-obese patients (44.0% vs. 54.8%). The sensitivity of a SAFE score <0 increased with age (41.9% vs. 54.5% vs. 76.6% vs. 89.3% vs. 97.3%

Table 1. Baseline Characteristics

	Total (n=5480)	Transient elastography		p
		LSM <8 kPa (n=4252)	LSM ≥8 kPa (n=1228)	
Age (yr)	52.0 (40.0–61.0)	52.0 (40.0–60.0)	52.5 (39.0–63.0)	0.097
Sex				0.001
Male	3236 (59.1)	2583 (60.7)	653 (53.2)	
Female	2244 (40.9)	1669 (39.3)	575 (46.8)	
DM	2118 (38.6)	1518 (35.7)	600 (48.9)	<0.001
HTN	1782 (35.8)	1348 (33.6)	434 (44.9)	<0.001
BMI (kg/m^2)	27.0 (24.8–29.7)	26.6 (24.5–29.1)	28.7 (26.0–32.0)	<0.001
Obesity	3959 (72.2)	2935 (69.0)	1024 (83.4)	<0.001
Total protein (g/dL)	7.3 (7.0–7.5)	7.2 (7.0–7.5)	7.3 (7.0–7.6)	<0.001
Albumin (g/dL)	4.5 (4.3–4.7)	4.5 (4.3–4.7)	4.5 (4.3–4.7)	<0.001
Globulin (g/dL)	2.7 (2.5–2.9)	2.7 (2.5–2.9)	2.8 (2.6–3.1)	<0.001
Total bilirubin (mg/dL)	0.7 (0.6–1.0)	0.7 (0.6–1.0)	0.7 (0.6–0.9)	0.057
ALT (U/L)	38.0 (23.0–64.0)	36.0 (22.0–59.0)	52.0 (30.0–87.0)	<0.001
AST (U/L)	32.0 (23.0–48.0)	29.0 (22.0–41.0)	49.5 (34.0–73.0)	<0.001
GGT (IU/L)	42.0 (26.0–69.0)	37.0 (24.0–61.0)	59.0 (40.0–96.0)	<0.001
Creatinine (mg/dL)	0.8 (0.7–0.9)	0.8 (0.7–0.9)	0.8 (0.7–0.9)	0.001
PLT (×10^3/μL)	244.0 (205.0–287.0)	248.0 (211.0–290.0)	225.5 (183.0–275.3)	<0.001

LSM, liver stiffness measurement; DM, diabetes mellitus; HTN, hypertension; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gammaglutamyl transferase; PLT, platelet count.
Values are n (%) or median (interquartile range).

https://doi.org/10.3349/ymj.2024.0318

Table 2. Diagnostic Accuracy of the SAFE Score for Predicting ≥F2 according to TE Results

	SAFE <0					SAFE ≥100				
	n	Fibrosis	Sensitivity (%)	Specificity (%)	NPV (%)	n	Fibrosis	Sensitivity (%)	Specificity (%)	PPV (%)
All	2568	268	78.2	54.8	89.6	1165	613	50.0	86.8	52.6
BMI										
Lean (<23)	292	23	68.9	63.7	92.1	68	33	44.6	91.7	48.5
Overweight or obese (≥23)	2276	245	78.8	53.8	89.2	1097	580	50.3	86.3	52.9
Non-obese (<25)	851	47	77.0	61.5	94.5	232	102	50.0	90.1	44.0
Obese (≥25)	1717	221	78.4	51.8	87.1	933	511	50.0	85.4	54.8
Sex										
Male	1693	196	69.9	58.5	88.4	549	253	38.8	88.4	46.1
Female	875	72	87.5	49.0	91.8	616	360	62.6	84.4	58.4
DM										
No	2060	212	66.2	68.1	89.7	405	217	34.6	93.1	53.6
Yes	508	56	90.7	30.5	89.0	760	396	66.0	75.5	52.1
Age groups										
19–30 years	444	90	41.9	87.6	79.7	36	25	16.1	97.3	69.4
31–40 years	624	86	54.5	80.7	86.2	57	40	21.2	97.5	70.2
41–50 years	685	51	76.6	68.8	92.6	119	82	37.6	96.0	68.9
51–60 years	614	31	89.3	49.2	95.0	301	167	57.6	88.7	55.5
≥61 years	201	10	97.3	18.7	95.0	652	299	79.7	65.4	45.9

SAFE, steatosis-associated fibrosis estimator; TE, transient elastography; NPV, negative predictive value; PPV, positive predictive value; BMI, body mass index; DM, diabetes mellitus.
Clinically significant fibrosis was defined as liver stiffness measurement by TE ≥8 kPa.

for age of 19–30, 31–40, 41–50, 51–60, 61 years or more), whereas the specificity decreased with age (87.6% vs. 80.7% vs. 68.8% vs. 49.2% vs. 18.7%).

SAFE score validation using biopsy data

Among 5480 patients, liver biopsy was performed in 9.0% (n=492). Of the 156 (31.7%) patients with a SAFE score <0, 20 (12.8%) showed significant fibrosis (≥F2). The sensitivity, specificity, and NPV of a SAFE score <0 for ruling out significant fibrosis (≥F2) were 89.7%, 45.6%, and 87.2%, respectively. Among 190 (38.6%) patients with a SAFE score ≥100, 126 (66.3%) had significant fibrosis. The sensitivity, specificity, and NPV of a SAFE score ≥100 for diagnosing significant fibrosis were 64.9%, 78.5%, and 66.3%, respectively.

We further compared the performance between non-obese (BMI <25 kg/m^2, n=100, 20.3%) and obese (BMI ≥25 kg/m^2, n=392, 79.7%) patients. Non-obese and obese patients had similar performance in sensitivity (95.7% vs. 87.8%), specificity (42.6% vs. 46.3%), and NPV (92.0% vs. 86.3%) for a SAFE score <0 and similar sensitivity (63.0% vs. 65.5%), specificity (77.8% vs. 78.7%), and PPV (70.7% vs. 65.1%) for a SAFE score ≥100. Male patients, compared with female patients, had a similar sensitivity (82.0% vs. 93.2%) but higher specificity (58.2% vs. 27.3%) for a SAFE score <0, and a lower sensitivity (55.7% vs. 69.2%) but higher specificity (90.4% vs. 61.2%) for a SAFE score ≥100. For a SAFE score <0, sensitivity increased with age (50.0% vs. 83.3% vs. 97.6% for age of 19–30, 31–50, 51 years or more), whereas specificity decreased with age (76.5% vs. 48.0% vs. 15.9%).

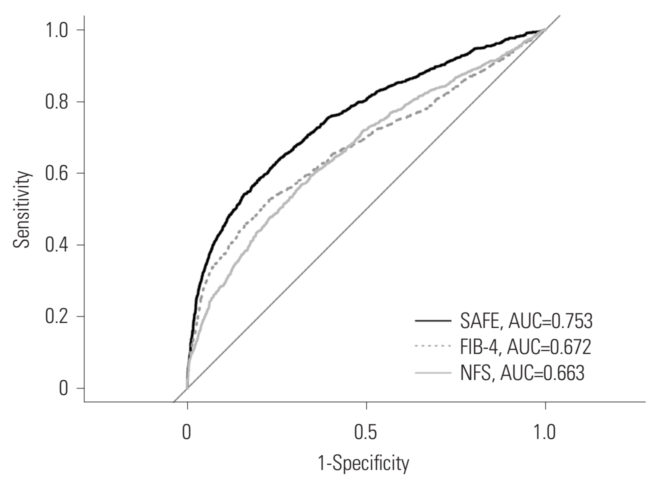

Fig. 2. Comparison of the areas under the curve for SAFE, FIB-4, and NFS in predicting ≥F2 using transient elastography results. SAFE, steatosis-associated fibrosis estimator; FIB-4, Fibrosis-4; NFS, NAFLD fibrosis score; AUC, area under the curve.

SAFE score performance predicting ≥F2 compared to FIB-4 or NFS

For predicting ≥F2 using TE data, the AUROC of the SAFE score was 0.753 (95% CI 0.737–0.769), which was significantly higher than those of the FIB-4 (0.672, 95% CI 0.654–0.691) and NFS (0.663, 95% CI 0.645–0.681) (Fig. 2). As shown in Table 3, non-obese patients had a higher AUROC for the SAFE score (0.786, 95% CI 0.749–0.822) than obese patients (0.738, 95% CI 0.720–0.757, *p*=0.024), while AUROCs were similar between lean

Table 3. Areas Under the Receiver Operating Characteristic Curve for Each Model Using TE Results

	SAFE	FIB-4	NFS
All	0.753 (0.737–0.769)	0.672 (0.654–0.691)	0.663 (0.645–0.681)
BMI			
Lean (<23)	0.738 (0.669–0.807)	0.717 (0.642–0.792)	0.661 (0.582–0.740)
Overweight or obese (≥23)	0.752 (0.735–0.769)	0.672 (0.652–0.692)	0.661 (0.642–0.679)
Non-obese (<25)	0.786 (0.749–0.822)	0.758 (0.717–0.799)	0.693 (0.650–0.735)
Obese (≥25)	0.738 (0.720–0.757)	0.665 (0.644–0.686)	0.650 (0.630–0.671)
Sex			
Male	0.706 (0.683–0.730)	0.616 (0.590–0.643)	0.627 (0.602–0.652)
Female	0.806 (0.785–0.827)	0.734 (0.708–0.760)	0.701 (0.675–0.726)
DM			
No	0.728 (0.705–0.751)	0.645 (0.618–0.672)	0.642 (0.616–0.668)
Yes	0.770 (0.746–0.793)	0.687 (0.660–0.714)	0.637 (0.610–0.664)
Age groups			
19–30 years	0.732 (0.686–0.778)	0.652 (0.601–0.702)	0.635 (0.582–0.688)
31–40 years	0.763 (0.725–0.802)	0.650 (0.605–0.694)	0.685 (0.639–0.730)
41–50 years	0.804 (0.770–0.837)	0.741 (0.702–0.780)	0.715 (0.676–0.755)
51–60 years	0.820 (0.791–0.848)	0.774 (0.742–0.806)	0.714 (0.681–0.747)
≥61 years	0.812 (0.786–0.838)	0.775 (0.747–0.804)	0.710 (0.679–0.740)

TE, transient elastography; SAFE, steatosis-associated fibrosis estimator; FIB-4, Fibrosis-4; NFS, NAFLD fibrosis score; BMI, body mass index; DM, diabetes mellitus.
95% confidence intervals were calculated using DeLong's method.

(0.738, 95% CI 0.669–0.807) and overweight or obese (0.752, 95% CI 0.735–0.769) patients (p=0.695). AUROC increased with age: 19–30 years (0.732, 95% CI 0.686–0.778), 31–40 years (0.763, 95% CI 0.725–0.802), 41–50 years (0.804, 95% CI 0.770–0.837), 51–60 years (0.820, 95% CI 0.791–0.848), and ≥61 years (0.812, 95% CI 0.786–0.838). Pairwise p-values are listed in Supplementary Table 3 (only online).

Based on biopsy data, the AUROC of the SAFE score was 0.790 (95% CI 0.749–0.831), which was similar to those of the FIB-4 (0.800, 95% CI 0.759–0.841), and NFS (0.802, 95% CI 0.761–0.843) (Fig. 3). Non-obese (0.792, 95% CI 0.702–0.881) and obese (0.790, 95% CI 0.743–0.837) patients had similar AUROCs for the SAFE score (p=0.972). AUROC was higher in lean patients (0.901, 95% CI 0.813–0.989) than in overweight or obese patients (0.784, 95% CI 0.740–0.828, p=0.022). By age, AUROC varied: 19–30 years (0.616, 95% CI 0.467–0.765), 31–50 years (0.772, 95% CI 0.698–0.846), and ≥51 years (0.710, 95% CI 0.638–0.781).

We also analyzed the accuracy of the recently proposed metabolic dysfunction-asssociated fIbrosis-5 (MAF-5) score.[18] AUROC for predicting significant fibrosis with MAF 5 was 0.772 (95% CI 0.755–0.790) by TE and 0.580 (95% CI 0.520–0.640) by biopsy. Since the AUROC of SAFE for TE prediction falls within the CI of MAF-5, their performance can be considered comparable in this context. However, for biopsy-based prediction, the SAFE score demonstrated superior performance. These findings further support the applicability of the SAFE score in Asian patients.

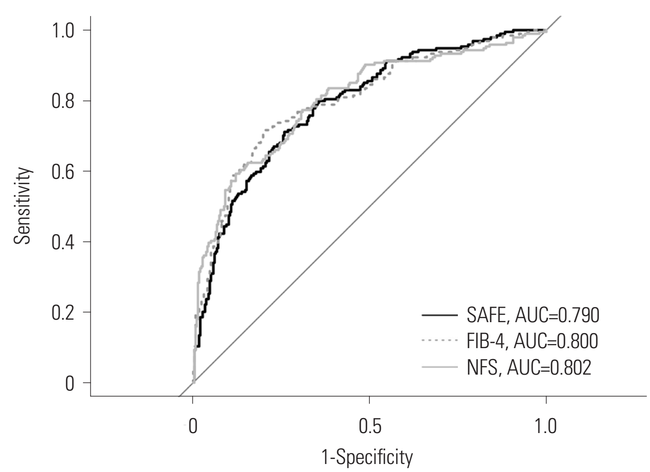

Fig. 3. Comparison of the receiver operating characteristic curves for SAFE, FIB-4, and NFS in predicting significant fibrosis using histologic data. SAFE, steatosis-associated fibrosis estimator; FIB-4, Fibrosis-4; NFS, NAFLD fibrosis score; AUC, area under the curve.

DISCUSSION

In this study, we validated the performance of the SAFE score in Asian patients with MASLD which were less obese than Western population. The SAFE score showed good performance for predicting significant fibrosis. This performance was consistent at both non-obese and obese patients with MASLD. The SAFE score outperformed the NFS and FIB-4 in predicting clinically significant fibrosis. Interestingly, the sensitivity and AUROC of the SAFE score tended to be lower in the younger group, which may require adjustment of the cutoff value in this age group. When we applied SAFE score to our liver biopsy cohort, the

https://doi.org/10.3349/ymj.2024.0318

AUROC of the SAFE score was similar to those of the NFS and FIB-4. These results suggest that SAFE score, which focused more on metabolic risk factors, might be utilized as a screening tool to identify high risk group in MASLD patients.

The AUROC of the SAFE score for predicting clinically significant fibrosis was 0.753, which was similar to the AUROC of the previous validation study.[19] Based on its predictive power >0.7, the SAFE score is suitable for predicting significant fibrosis in Asian populations. The sensitivity of a SAFE score <0 for ruling out clinically significant fibrosis was 78.2% , which was lower than that of the previous study, which ranged from 90% to 98%.[10] However, the sensitivity for predicting significant fibrosis based on liver biopsy (89.7%) was similar to the previous data.[10] Since the original SAFE score was originated from biopsy results, this difference could be a result of discrepancies between diagnostic procedures. Despite the difference of study population, the sensitivity, specificity, NPV, and PPV for predicting clinically significant fibrosis were similar to those reported previously in a validation study.[19] For predicting F2 fibrosis based on biopsy results, the SAFE score was not superior to other scoring systems. However, since biopsy is an invasive procedure, there could be a selection bias toward patients with higher disease severity (high SAFE score, LS, AST, and ALT), as shown in Supplementary Table 2 (only online). Since the SAFE score was originally intended to screen for significant fibrosis in low-risk rather than high-risk groups, this result might be expected.

In a recent study, SAFE score was validated with good performance in Asian patients with NAFLD.[20] In this study, the cohort included fewer low-risk patients and more high-risk patients. In our study, 47.1% of patients were in the low-risk group. This reflects that the even Asian can have different at-risks of disease progression according to food, lifestyle, or metabolic risks. When we consider that SAFE score was developed to detect low-risk NAFLD, our cohort seems to be appropriate to validate this score. In addition, unlike previous studies, we focused on lean and non-obese patients, because original paper included patients who were mostly obese (BMI ≥30 kg/m^2) and had few lean patients.[10] Asians have a higher proportion of MASLD with normal body weight compared with Westerners.[21] Among people with non-obese or lean NAFLD, 39.0% had steatohepatitis, 29.2% had significant fibrosis, and 3.2% had liver cirrhosis.[22] Additionally, lean patients with MASLD tend to have more severe liver disease and poorer outcomes compared with obese patients with MASLD.[23] Thus, it is important to validate the SAFE score in lean or non-obese patients, especially in Asian patients.

In our study, clinically significant fibrosis was more prevalent in obese patients compared to non-obese patients (25.9% vs. 13.4%). This is consistent with a previous report showing that lean NAFLD patients have less severe histologic findings than obese subjects.[22] The AUROC of SAFE score was slightly higher in lean or non-obese patients compared to overweight or obese patients. For predicting significant fibrosis based on TE, the AUROC was higher in females than males, whereas for predicting significant fibrosis based on biopsy results, the AUROC was higher in males than females, although statistically insignificant. These differences may be explained by the fact that, among patients undergoing liver biopsy, males had less severe fibrosis stages than females. When patients were divided into age groups 19–30, 31–40, 41–50, 51–60, and ≥61 years, sensitivity increased and specificity decreased with age. Since the SAFE score is used to assess the risk of significant fibrosis, the 16.1% sensitivity of a SAFE score cutoff ≥100 in patients 19–30 years of age is concerning. Moreover, the AUROC for predicting clinically significant fibrosis in the younger age groups was lower than that in the older age groups. Therefore, minor adjustments of parameters or cutoff values might be needed in younger patients. As in the previous validation study, patients aged 19−40 years had lower sensitivity (18.9%) compared to those aged 41−60 years (49.0%), and ≥61 years (79.7%).[19]

One possible explanation for the low AUROC and sensitivity observed in the younger age group, especially patients with 19−30 years could be lower proportion of DM patients within this age range. For example, among patients aged 19–30 years, only 21.9% (119/544) had DM, whereas among all patients the DM prevalence was 38.7% (2087/5393), consistent with previous reports.[24] When we calculated the sensitivity and specificity for non-DM and DM patients using the TE data, sensitivity was lower in non-DM patients (65.6% vs. 90.5% in SAFE <0, 33.4% vs. 65.7% in SAFE ≥100), and specificity was higher in non-DM patients (68.3% vs. 30.6% in SAFE <0, 93.2% vs. 75.6% in SAFE ≥100). Moreover, when we calculated AUROC of SAFE score in DM and non-DM patients using the TE data, AUROC was higher in DM patients (0.768) than in non-DM patients (0.724). Therefore, since patients aged 19−30 years include a higher proportion of non-DM individuals who have lower AUROC and sensitivity but higher specificity, these age related differences may account for the overall lower sensitivity and AUROC. These results suggest a potential need for separate modeling or different cut off values by age group, and they provide an explanation for the lower AUROC observed in younger patients.

Primary risk assessment is important for identifying patients with MASLD who are at-risk for disease progression. The American Association for the Study of Liver Diseases and European Association for the Study of Liver Diseases guidelines recommend a two-step algorithm using FIB-4 and TE to identify patients at risk.[25,26] FIB-4 has reasonable specificity but low sensitivity, and is mainly composed of liver-oriented variables.[27] Additionally, in patients with an indeterminate or high FIB-4, additional risk stratification using TE should be considered. Although TE is the best-validated imaging technique for fibrosis risk stratification, TE is not always available to use in primary care setting. Therefore, the SAFE score can be used as a new and simple method that includes more metabolic variables such as BMI and DM. We also analyzed the proportion of patients with ≥F2 according to both previous referral criteria and SAFE-based criteria: in both referred and non-referred groups, the SAFE-identified referral group had a higher proportion with ≥F2 (Supplementary Table 4, only online). The American Diabetes Association also emphasizes initial fibrosis risk stratification and appropriate referral of at-risk DM patients to a gastroenterologist or hepatologist for further workup.[28] The SAFE score is expected to be the primary tool for risk stratification of patients with MASLD and/or DM.

This study has several limitations. First, these are retrospective data from a single tertiary center. Most MASLD patients were referred from health check-ups or by primary care physicians. Thus, this cohort is suitable for validating the SAFE score to detect low-risk MASLD patients. Third, BMI, DM, and age are components of the SAFE score, possibly resulting in overfitting if the values were fitted to the individual subgroups. However, we used subgroups to validate the performance of the SAFE score, rather than for fitting the model. We focused on demonstrating that the SAFE score has non-inferior performance in non-obese patients. Moreover, biopsy results were available for 492 patients, resulting in wide CIs. TE has been validated as a non-invasive test for predicting fibrosis,[9,29] and not all MASLD patients undergo liver biopsy; therefore, the results of this study remain clinically relevant.

In conclusion, we externally validated the good performance of the SAFE score in Asian patients. The SAFE score has potential as an initial assessment to identify a low-risk population in a primary care setting. Further studies are needed to validate the use of the SAFE score with long-term follow-up.

ACKNOWLEDGEMENTS

This study was supported by the National Research Foundation of Korea (2020M3A9E4038694).

AUTHOR CONTRIBUTIONS

Conceptualization: Jun Yong Park. Data curation: Hye Won Lee. Formal analysis: Kunhee Kim. Funding acquisition: Jun Yong Park. Investigation: Jae Seung Lee. Methodology: Mi Na Kim. Project administration: Jun Yong Park. Resources: Hye Won Lee and Jun Yong Park. Software: Kunhee Kim. Supervision: Jun Yong Park. Validation: Hye Won Lee. Visualization: Kunhee Kim. Writing—original draft: Kunhee Kim. Writing—review & editing: Hye Won Lee and Jun Yong Park. Approval of final manuscript: all authors.

ORCID iDs

Kunhee Kim	https://orcid.org/0009-0008-9280-3455
Hye Won Lee	https://orcid.org/0000-0002-3552-3560
Jae Seung Lee	https://orcid.org/0000-0002-2371-0967
Mi Na Kim	https://orcid.org/0000-0001-5381-230X
Jun Yong Park	https://orcid.org/0000-0001-6324-2224

REFERENCES

1. Powell EE, Wong VW, Rinella M. Non-alcoholic fatty liver disease. Lancet 2021;397:2212-24.
2. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology 2015;149:389-97.e10.
3. Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. Hepatology 2017;65:1557-65.
4. Hagström H, Nasr P, Ekstedt M, Hammar U, Stål P, Hultcrantz R, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017;67:1265-73.
5. Ekstedt M, Hagström H, Nasr P, Fredrikson M, Stål P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology 2015;61:1547-54.
6. 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;45:846-54.
7. Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, Mal F, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 2003;29:1705-13.
8. 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;43:1317-25.
9. Kwok R, Tse YK, Wong GL, Ha Y, Lee AU, Ngu MC, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease--the role of transient elastography and plasma cytokeratin-18 fragments. Aliment Pharmacol Ther 2014;39:254-69.
10. Sripongpun P, Kim WR, Mannalithara A, Charu V, Vidovszky A, Asch S, et al. The steatosis-associated fibrosis estimator (SAFE) score: a tool to detect low-risk NAFLD in primary care. Hepatology 2023;77:256-67.
11. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. Ann Hepatol 2024;29:101133.
12. Zhang S, Mak LY, Yuen MF, Seto WK. Screening strategy for non-alcoholic fatty liver disease. Clin Mol Hepatol 2023;29(Suppl):S103-22.
13. Kim DS, Kim BK, Lee JS, Lee HW, Park JY, Kim DY, et al. Noninvasive risk assessment of hepatic decompensation in patients with hepatitis B virus-related liver cirrhosis. J Gastroenterol Hepatol 2023;38:1372-80.
14. Harring M, Golabi P, Paik JM, Shah D, Racila A, Cable R, et al. Sarcopenia among patients with nonalcoholic fatty liver disease (NAFLD) is associated with advanced fibrosis. Clin Gastroenterol Hepatol 2023;21:2876-88.e5.
15. Lee HW, Park SY, Kim SU, Jang JY, Park H, Kim JK, et al. Discrimination of nonalcoholic steatohepatitis using transient elastography in patients with nonalcoholic fatty liver disease. PLoS One 2016;11:e0157358.
16. Pan WH, Yeh WT. How to define obesity? Evidence-based multiple action points for public awareness, screening, and treatment: an extension of Asian-Pacific recommendations. Asia Pac J Clin Nutr 2008;17:370-4.

17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.
18. van Kleef LA, Francque SM, Prieto-Ortiz JE, Sonneveld MJ, Sanchez-Luque CB, Prieto-Ortiz RG, et al. Metabolic dysfunction-associated fibrosis 5 (MAF-5) score predicts liver fibrosis risk and outcome in the general population with metabolic dysfunction. Gastroenterology 2024;167:357-67.e9.
19. van Kleef LA, de Knegt RJ, Ayada I, Pan Q, Brouwer WP. The steatosis-associated fibrosis estimator (SAFE) score: validation in the general US population. Hepatol Commun 2023;7:e0075.
20. Li G, Lin H, Sripongpun P, Liang LY, Zhang X, Wong VWS, et al. Diagnostic and prognostic performance of the SAFE score in nonalcoholic fatty liver disease. Liver Int 2024;44:15-26.
21. Young S, Tariq R, Provenza J, Satapathy SK, Faisal K, Choudhry A, et al. Prevalence and profile of nonalcoholic fatty liver disease in lean adults: systematic review and meta-analysis. Hepatol Commun 2020;4:953-72.
22. Ye Q, Zou B, Yeo YH, Li J, Huang DQ, Wu Y, et al. Global prevalence, incidence, and outcomes of non-obese or lean non-alcoholic fatty liver disease: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2020;5:739-52.
23. Nabi O, Lapidus N, Boursier J, de Ledinghen V, Petit JM, Kab S, et al. Lean individuals with NAFLD have more severe liver disease and poorer clinical outcomes (NASH-CO study). Hepatology 2023;78:272-83.
24. Rathmann W, Giani G. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004;27:2568-9; author reply 2569.
25. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, Abdelmalek MF, Caldwell S, Barb D, et al. AASLD practice guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology 2023;77:1797-835.
26. 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;75:659-89.
27. Siddiqui MS, Yamada G, Vuppalanchi R, Van Natta M, Loomba R, Guy C, et al. Diagnostic accuracy of noninvasive fibrosis models to detect change in fibrosis stage. Clin Gastroenterol Hepatol 2019;17:1877-85.e5.
28. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. 1. Improving care and promoting health in populations: standards of care in diabetes—2023. Diabetes Care 2023;46(Supple 1):S10-8.
29. Pennisi G, Enea M, Falco V, Aithal GP, Palaniyappan N, Yilmaz Y, et al. Noninvasive assessment of liver disease severity in patients with nonalcoholic fatty liver disease (NAFLD) and type 2 diabetes. Hepatology 2023;78:195-211.