



Accuracy of Noninvasive Scoring Systems in Assessing Liver Fibrosis in Patients with Nonalcoholic Fatty Liver Disease: A Systematic Review and Meta-Analysis

Sangsoo Han¹, Miyoung Choi², Bora Lee³, Hye-Won Lee⁴, Seong Hee Kang⁵, Yuri Cho⁶, Sang Bong Ahn⁷, Do Seon Song⁸, Dae Won Jun⁹, Jieun Lee¹⁰, and Jeong-Ju Yoo¹¹

¹Department of Emergency Medicine, Soonchunhyang University Bucheon Hospital, Bucheon, ²Clinical Evidence Research, National Evidence-based Healthcare Collaborating Agency (NECA), ³Department of Statistics, Graduate School of Chung-Ang University, ⁴Department of Internal Medicine, Yonsei University College of Medicine, Seoul, ⁵Department of Internal Medicine, Wonju Severance Christian Hospital, Yonsei University Wonju College of Medicine, Wonju, ⁶Center for Liver and Pancreatobiliary Cancer, National Cancer Center, Goyang, ⁷Department of Internal Medicine, Nowon Eulji Medical Center, Eulji University College of Medicine, Seoul, ⁸Department of Internal Medicine, St. Vincent's Hospital, College of Medicine, The Catholic University of Korea, Suwon, ⁹Department of Internal Medicine, Hanyang University College of Medicine, Seoul, ¹⁰Department of Internal Medicine, Soonchunhyang University College of Medicine, Cheonan, and ¹¹Department of Internal Medicine, Soonchunhyang University Bucheon Hospital, Bucheon, Korea

Article Info

Received August 24, 2021

Revised November 18, 2021

Accepted November 23, 2021

Published online February 23, 2022

Corresponding Author

Dae Won Jun

ORCID <https://orcid.org/0000-0002-2875-6139>

E-mail noshin@hanyang.ac.kr

Jeong-Ju Yoo

ORCID <https://orcid.org/0000-0002-7802-0381>

E-mail puby17@naver.com

Sangsoo Han and Miyoung Choi contributed equally to this work as first authors.

Background/Aims: Several noninvasive scoring systems have been developed to determine the risk of advanced fibrosis in nonalcoholic fatty liver disease (NAFLD). We examined the diagnostic accuracy of the fibrosis-4 (FIB-4) score and NAFLD fibrosis score (NFS) in patients with biopsy-proven NAFLD.

Methods: For this meta-analysis, various databases including PubMed (MEDLINE), EMBASE, OVID Medline and the Cochrane Library were systematically searched. After the acquired abstracts were reviewed by two investigators, manuscripts were chosen for a full-text examination.

Results: Thirty-six studies evaluating biopsy-proven NAFLD were selected for meta-analysis. A total of 14,992 patients were analyzed. The lower cutoff sensitivity of the FIB-4 score predicting histological fibrosis stage 3 or more ($\geq F3$) was 69%, with specificity of 64%, positive likelihood ratio (LR+) of 1.96, and negative likelihood ratio (LR-) of 0.47. The low baseline sensitivity of the NFS score predicting $\geq F3$ was 70%, with a specificity of 61%, LR+ of 1.83, and LR- of 0.48. The area under the receiver operating characteristic curve (AUC) values of the FIB-4 score predicting $\geq F3$ and $\geq F2$ were 76% and 68%, respectively. The AUC values of the NFS score predicting $\geq F3$ and $\geq F2$ were 74% and 60%, respectively.

Conclusions: The FIB-4 or NFS test can be used to predict the degree of liver fibrosis in NAFLD, and the diagnostic accuracy resulted as relatively high in fibrosis stages of F3 or higher. (**Gut Liver, Published online February 23, 2022**)

Key Words: Liver fibrosis; Meta-analysis; Nonalcoholic fatty liver disease; Predictive value of tests

INTRODUCTION

With a prevalence of 25% to 40% in the general population, nonalcoholic fatty liver disease (NAFLD) is the most common liver disease worldwide, a pressing health concern associated with insulin resistance and metabolic syndrome.^{1,2} NAFLD affects nearly 100 million individu-

als in the United States and occurs in 90% of the obese population.^{1,3} Due to such burden of the disease, the early identification of patients with high morbidity and mortality associated with NAFLD is essential.

NAFLD can be categorized into various stages, from simple steatosis without fibrosis to nonalcoholic steatohepatitis related cirrhosis.¹ The severity of NAFLD is

Copyright © Gut and Liver.



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

determined by three factors: steatosis, inflammation, and fibrosis. Among these factors, the degree of hepatic fibrosis is the most essential factor in clinical settings, allowing clinicians to estimate the long-term prognosis in patients with NAFLD, such as the development of hepatocellular carcinoma, liver-related death or cardiovascular mortality.^{1,4} In fact, while simple steatosis is considered a non-progressive condition, nonalcoholic steatohepatitis or significant fibrosis (SF) is regarded as one of the main causes of liver transplantation.⁵ Therefore, it is vital to promptly identify advanced fibrosis (AF; \geq stage 3 fibrosis) or SF (\geq stage 2 fibrosis) in such patients.⁶

Liver biopsy is the gold standard for staging and identifying fibrosis in NAFLD patients.⁷ However, it is not suitable for a routine screening use, due to various reasons including its invasive nature, potential complications, possibility of sampling error, and high cost.^{7,8} Therefore, a simple, inexpensive and noninvasive panel to identify and quantify liver fibrosis is necessary. Likewise, though techniques such as magnetic resonance elastography or transient elastography have been recently developed, their high expense prevents their use as routine screening tests. Thus, noninvasive fibrosis scoring systems based on serologic tests such as the fibrosis-4 (FIB-4) index and NAFLD fibrosis score (NFS) have been developed and widely used as screening tools to assess the degree of fibrosis.^{9,10} However, a comprehensive study on such scoring system is crucial, since not only was the FIB-4 score developed only for patients with viral hepatitis, the accuracy of these serologic scoring systems in NAFLD patients also differ among studies. The aim of this systematic review and meta-analysis is to evaluate the diagnostic accuracy of noninvasive fibrosis scoring systems (FIB-4 and NFS), compared to that of the corresponding liver histologic data, to predict AF and SF in patients with NAFLD.

MATERIALS AND METHODS

This meta-analysis adhered to the protocol previously registered with PROSPERO (International Prospective Register of Systematic Reviews, CRD42021241243). We administered this systematic review and meta-analysis following guidelines provided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Diagnostic Test Accuracy.

1. Inclusion criteria, exclusion criteria, and study outcome

Studies that documented the accuracy of FIB-4 and NFS, evaluated by the corresponding liver histology results

in NAFLD patients, were considered eligible for inclusion. The following criteria were required for studies to be selected: (1) patients with NAFLD; (2) reports of the accuracy of FIB-4 and NFS based on liver histology results. The state of fatty liver was determined by histologic characteristics. Eligible study designs were randomized controlled trials, cross-sectional studies, and cohort studies, both prospective and retrospective. Studies were excluded by the following criteria: (1) case reports; (2) case series, in which less than five patients in total were involved; (3) reviews; (4) cell or animal studies; (5) chronic viral hepatitis, such as hepatitis B or hepatitis C; (6) human immunodeficiency virus; (7) significant alcohol consumption; (8) fatty liver defined by imaging or serologic criteria, without any histology result provided; or (9) non-English studies.

The primary outcome of this meta-analysis was the diagnostic accuracy of the FIB-4 score and NFS, compared to the corresponding liver histology in patients with NAFLD.

2. Search strategy

We searched PubMed (MEDLINE), EMBASE, the Cochrane Library, Korean Medical Database, and Korean Studies Information Service System to identify studies published in English between January 1, 1997, and October 31, 2020. The keywords used in the Patient/Problem, Intervention, Comparison, and Outcome model are provided in the Supplementary Material. The search words were NAFLD index words, FIB-4-related index words or NFS-related index words. We combined free-text words and controlled terms such as Medical Subject Headings and Emtree according to the databases. The search strategy and the following result of each database are provided in the Supplementary material. The entire search process was administered by a professional librarian (M.C.).

3. Study selection and data extraction

During the process of study selection, two reviewers (S.H. and J.J.Y.) first independently extracted relevant titles and abstracts. After an independent examination of the full-text articles, any resulting disparity between the two reviewers was resolved by a discussion with a third reviewer (H.W.L. or S.H.K.).

In addition, the two reviewers thoroughly examined the remaining procedures, such as screening full-text articles and assessing the risk of bias. The extraction of study characteristics and outcomes was conducted independently and documented in a standardized format by the two reviewers. Any discrepancy was settled by a discussion with Y.C. and S.B.A.

4. Methodological quality and risk of bias assessment

To determine the risk of bias, we utilized the Cochrane risk of bias tool, with any relevant information provided in the Supplementary Material. Again, any discrepancy was resolved by a discussion with additional reviewers (D.S.S. and D.W.J.). Risk of bias was evaluated using two tools, QUADAS¹¹ and QUIPS¹² tools. The overall outcome of the risk of bias is provided in the Supplementary material. Publication bias was evaluated through a funnel plot.

5. Statistical analysis

The process of meta-analysis with sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio is as the following: (1) transform the proportion into a quantity (Freeman-Tukey variant of the arcsine square root transformed the proportion); (2) calculate the pooled prevalence as the back-transformation of the weighted mean of the transformed prevalence using DerSimonian-Laird weights assuming the random-effect model; (3) calculate the confidence interval with the Clopper-Pearson interval. To further analyze the heterogeneity within the studies, researchers conducted a meta-regression to understand the influence of other factors on diagnostic accuracy. RevMan 5 (Cochrane Library) or the meta package in R version 4.1.0 (The R Foundation for Statistical Computing, Vienna, Austria) were utilized in the statistical analyses.

RESULTS

1. Characteristics of included studies

A thorough database search of titles and abstracts resulted in 86 relevant studies. Out of the 86, 50 studies were excluded due to inappropriate patient population (n=1), inappropriate outcome measurement (n=42), overlapping population (n=6), or insufficient data (n=1). Finally, 36 studies were eligible for inclusion in this review (Fig. 1). Detailed characteristics of the studies included in this meta-analysis are provided in Table 1.¹³⁻⁴⁸ A total of 14,992 patients were analyzed, with a mean age of 48.57±6.13. Studies were conducted in various countries in the world (Asia 16, Europe 8, America 9, two or more continents 2, and Australia 1). The median co-morbidity rates of diabetes, hypertension, and dyslipidemia were 36.2%, 43.5%, and 54.3%, respectively. Median aspartate aminotransferase and alanine aminotransferase levels were 43 U/L and 62 U/L, respectively.

2. FIB-4 for predicting AF (≥ F3)

Analyzing the diagnostic accuracy of FIB-4 for predicting AF involved 13,764 patients from 32 studies (Table 2). As a lower cutoff value for predicting AF, a value from 1.02 to 1.45 was most frequently used (20 studies), and as a higher cutoff value, 2.67 was predominantly used (18 stud-

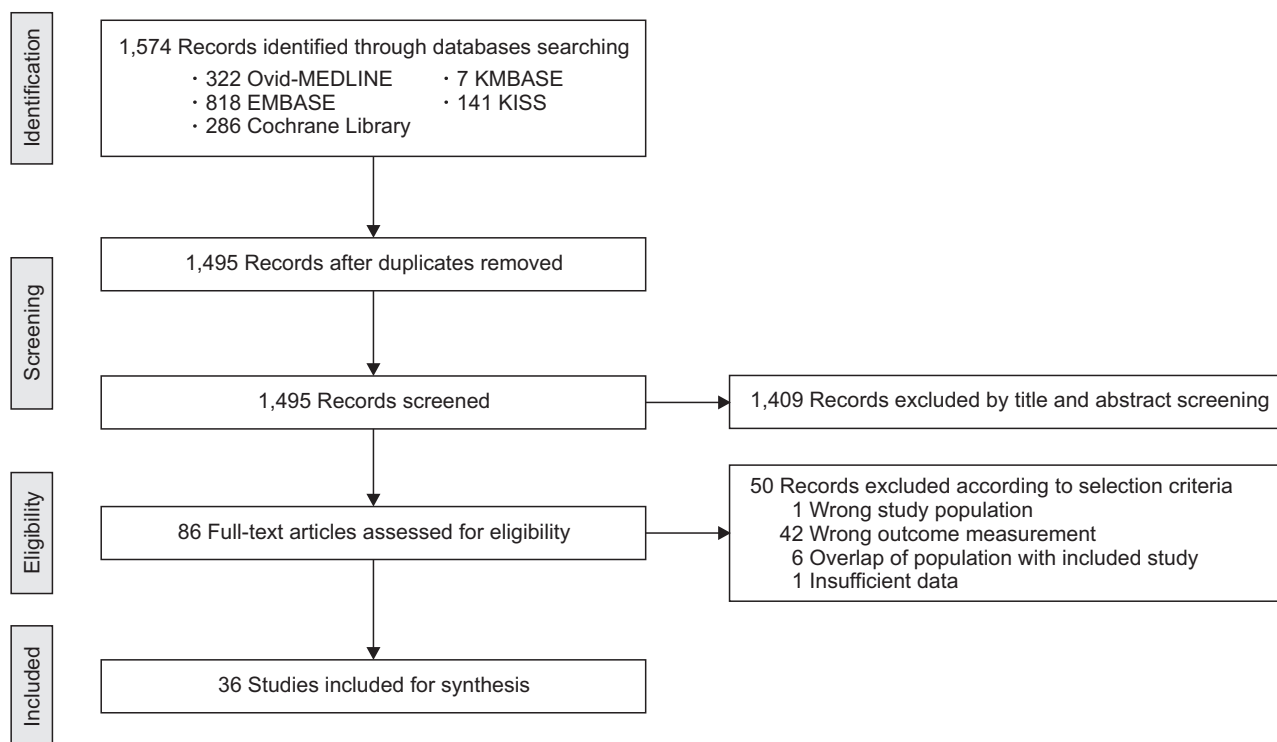


Fig. 1. Flowchart showing the process of study inclusion and exclusion in the systematic review. KMBASE, Korean Medical Database; KISS, Korean Studies Information Service System.

Table 1. Characteristics and Results of the Included Studies

Author (year)	Location	Method	No. of samples	Mean age, yr	Male, %	Clinical characteristics						
						DM, %	HTN, %	Dyslipidemia, %	Mean BMI, kg/m ²	Mean waist, cm	Mean AST, U/L	Mean ALT, U/L
Aida (2015) ¹³	Japan	FIB-4	148	61	36	-	-	-	26.9	-	42	52
Anstee (2019) ¹⁴	Global	FIB-4/NFS	3,123	59	42	68	-	-	-	-	44	45
Balakrishnan (2020) ¹⁵	USA	FIB-4/NFS	99	44.7	33.9	43.5	-	-	32.1	-	59	108
Boursier (2019) ¹⁶	France	FIB-4/NFS	938	56.5	58.5	51.1	-	-	31.8	-	39	56
Chan (2015) ¹⁷	Malaysia	NFS	147	50.5	54.4	52.4	89.1	-	29.3	98.2	41	71
Chan (2019) ¹⁸	Asia	FIB-4/NFS	583	50.9	52.9	52.3	55.1	74.5	28.9	96.5	38	63
Cui (2015) ¹⁹	USA	FIB-4	102	51	41.2	25.5	-	-	31.7	-	42.3	58
Demir (2013) ²⁰	Germany	NFS	120	43.8	47.2	19.9	41.9	-	37	-	36.8	56.6
de Carli (2019) ²¹	Brazil	NFS	266	36.5	20.4	10.5	-	-	44.2	123.9	24.9	32.3
Goh (2015) ²²	USA	NFS	238	52	29.3	100	-	42	37.1	-	53.5	64
	USA	NFS	263	46	46	0	-	15.2	35.2	-	58.5	78
Joo (2017) ²³	Korea	FIB-4/NFS	315	55	50.8	37.8	38.4	-	27	94.6	36	43
Jun (2017) ²⁴	Korea	FIB-4/NFS	328	36.4	70.7	33	14.6	-	28.6	96.4	91.3	98.5
Kakisaka (2018) ²⁵	Japan	FIB-4	63	54.9	58	-	-	-	28.1	-	62	94
Kao (2020) ²⁶	Taiwan	FIB-4	73	35.3	31.5	16.9	26.8	-	41	118.3	38.2	55
Kaya (2020) ²⁷	Turkey	FIB-4/NFS	463	46	47.5	37.8	34.8	-	31.7	104	42	66
Kim (2013) ²⁸	USA	FIB-4/NFS	142	52.8	26.8	27.5	45.1	-	36.32	-	47.2	60.4
Labenz (2018) ²⁹	Germany	FIB-4/NFS	261	51	52.5	29.9	-	37.5	30.9	-	48	60
Lang (2020) ³⁰	Germany	FIB-4/NFS	95	50	46.2	10.8	56.9	-	30	105	32.5	50.5
Lum (2020) ³¹	Singapore	FIB-4/NFS	263	50.4	52.5	49	-	66.5	30.4	113.2	-	-
McPherson (2013) ³²	UK	FIB-4/NFS	70	54	56	43	-	-	32.9	105	28	28
	UK	FIB-4/NFS	235	48	63	40	-	-	34.4	110	59	95
Meneses (2020) ³³	Spain	FIB-4/NFS	50	49	30	26	52	28	44.3	135	21	25
Nasr (2016) ³⁴	Sweden	FIB-4/NFS	58	60.4	71	53	93	-	28	102	34	60
Ooi (2017) ³⁵	Australia	FIB-4/NFS	101	49	33.7	34.7	79.2	73.3	41.9	-	-	-
	Australia	FIB-4/NFS	53	43	30.2	24.5	45.3	19.2	46.6	-	-	-
Patel (2018) ³⁶	USA	FIB-4/NFS	114	41.8	79	30	-	-	33.9	-	49	84
	USA	FIB-4/NFS	151	60	95	70	-	-	33.7	-	48.2	54.6
Pérez-Gutiérrez (2013) ³⁷	Mexico	FIB-4/NFS	243	48.6	49	21.5	-	-	-	-	57.6	73
Petta (2015) ³⁸	Italy	FIB-4/NFS	179	45.4	67.5	19.5	24	-	29.3	-	45.7	80.3
	Italy	FIB-4/NFS	142	43.9	71.8	15.4	11.9	-	27.4	-	42.2	75.6
Petta (2019) ³⁹	Global	FIB-4/NFS	968	50.1	62.9	37	39.4	-	29.3	-	46.1	76.1
Siddiqui (2020) ⁴⁰	USA	FIB-4/NFS	1,904	50.3	37	39	58	62	34.4	-	51.2	69.8
Singh (2020) ⁴¹	USA	FIB-4/NFS	1,134	51.1	35.4	100	74.5	70.8	35.5	-	27	28
Treeprasertsuk (2016) ⁴²	Thailand	FIB-4/NFS	139	40.9	47	38	-	-	36.1	-	38	56
Wong (2008) ⁴³	Hong Kong	NFS	128	46	59	57	48	-	28.5	95	43	75
Wong (2010) ⁴⁴	Hong Kong	FIB-4/NFS	246	51	54.9	36.2	40.2	-	28	94	-	75
Xun (2012) ⁴⁵	China	FIB-4/NFS	152	37.1	79.6	32.2	-	-	26.1	-	61	100
Yang (2019) ⁴⁶	China	FIB-4/NFS	453	36.56	58.9	30.2	34.8	-	26.93	-	74.12	135.11
Yoneda (2013) ⁴⁷	Japan	FIB-4/NFS	235	59.9	-	46	-	63.8	26.9	-	24.7	23.7
Zhou (2019) ⁴⁸	China	FIB-4/NFS	207	41.8	73	24.6	20.3	46.6	27	91.2	45.7	49

DM, diabetes mellitus; HTN, hypertension; BMI, body mass index; AST, aspartate aminotransferase; ALT, alanine aminotransferase; FIB-4, Fibrosis-4; NFS, nonalcoholic fatty liver disease (NAFLD) fibrosis score.

ies). Regarding the FIB-4 index, pooled sensitivity was 0.42 (95% confidence interval [CI], 0.33 to 0.51) and pooled specificity was 0.93 (95% CI, 0.91 to 0.95). Pooled diagnostic odds ratio (DOR) with 95% CI was 10.83 (7.55 to 15.54) with I^2 of 85% ($p < 0.01$). Summary statistics of FIB-4 at various thresholds for prediction of AF and forest plots are presented in Supplementary Table 1 and Supplementary

Fig. 1. The area under the receiver operating characteristic curve (AUC) of summary receiver operating characteristic (SROC) was 0.76 (95% CI, 0.74 to 0.81) (Fig. 2A).

3. NFS for predicting AF ($\geq F3$)

Analyzing the diagnostic accuracy of NFS for predicting AF involved 13,337 patients from 33 studies (Table 2). As

Table 2. Summary Sensitivities, Specificities, Diagnostic Odds Ratio, Positive Likelihood Ratio, Negative Likelihood Ratio, and AUC of FIB-4 Index and NFS at Various Diagnostic Thresholds for the Prediction of Advanced Fibrosis and Significant Fibrosis

Variable	Cutoff	No. of samples (No. of patients)	Summary statistics (95% CI)				LR+	LR-	AUC
			Sensitivity	Specificity	DOR	LR+			
Advanced fibrosis									
FIB-4	All	32 (13,764)	0.42 [0.33-0.51]	0.93 [0.91-0.95]	10.83 [7.55-15.54]	6.65 [5.01-8.83]	0.61 [0.52-0.71]	0.76 [0.74-0.81]	
	1.02 to 1.45	20 (10,304)	0.69 [0.59-0.77]	0.64 [0.57-0.71]	4.11 [2.24-7.55]	1.96 [1.49-2.57]	0.47 [0.33-0.67]	0.73 [0.67-0.80]	
	1.515 to 2.09	5 (2,408)	0.74 [0.58-0.85]	0.80 [0.72-0.86]	11.97 [7.48-19.16]	3.82 [3.00-4.88]	0.32 [0.20-0.51]	0.82 [0.78-0.87]	
	2.67	18 (8,731)	0.34 [0.27-0.42]	0.95 [0.92-0.96]	9.32 [6.26-13.87]	6.46 [4.66-8.95]	0.69 [0.62-0.77]	0.74 [0.70-0.76]	
	3.25	9 (2,721)	0.39 [0.22-0.59]	0.95 [0.93-0.97]	14.23 [5.92-34.19]	8.97 [4.87-16.50]	0.63 [0.45-0.86]	0.76 [0.77-0.88]	
NFS	All	33 (13,337)	0.38 [0.28-0.50]	0.94 [0.90-0.96]	10.16 [7.18-14.37]	6.60 [4.85-8.97]	0.64 [0.54-0.76]	0.74 [0.71-0.79]	
	-1.98 to -1.036	23 (10,158)	0.70 [0.57-0.80]	0.61 [0.53-0.69]	3.77 [2.16-6.56]	1.83 [1.46-2.28]	0.48 [0.33-0.70]	0.72 [0.66-0.79]	
	-0.126 to 0.19	2 (1,962)	0.61 [0.28-0.86]	0.87 [0.79-0.92]	10.79 [3.65-31.84]	4.79 [3.13-7.35]	0.44 [0.20-0.97]	0.80 [0.78-0.88]	
	4.39 to 4.8	31 (11,471)	0.31 [0.23-0.41]	0.95 [0.93-0.97]	10.17 [7.06-14.63]	7.29 [5.34-9.93]	0.71 [0.63-0.81]	0.73 [0.71-0.81]	
Significant fibrosis									
FIB-4	All	6 (547)	0.42 [0.16-0.73]	0.93 [0.56-0.99]	9.71 [2.43-38.70]	6.05 [1.25-29.29]	0.62 [0.40-0.95]	0.68 [0.65-0.76]	
	0.66 to 0.89	4 (434)	0.69 [0.55-0.80]	0.61 [0.44-0.76]	3.58 [1.96-6.53]	1.79 [1.26-2.55]	0.50 [0.35-0.70]	0.66 [0.61-0.77]	
	1.4 to 1.9	2 (136)	0.65 [0.53-0.75]	0.66 [0.51-0.79]	3.69 [1.61-8.49]	1.94 [1.21-3.09]	0.52 [0.35-0.78]	0.74 [0.54-0.85]	
NFS	2.67 to 3.25	2 (151)	0.06 [0.02-0.22]	0.98 [0.94-1.00]	4.25 [0.57-31.68]	10.82 [0.47-257.67]	0.93 [0.85-1.02]	0.67 [0.61-0.73]	
	All	5 (539)	0.25 [0.02-0.82]	0.76 [0.37-0.94]	1.16 [0.37-3.57]	1.12 [0.50-2.49]	0.96 [0.69-1.33]	0.60 [0.52-0.69]	
	-3.168 to -1.455	2 (335)	0.57 [0.28-0.81]	0.68 [0.37-0.88]	2.82 [1.53-5.18]	1.78 [1.11-2.84]	0.63 [0.42-0.94]	0.61 [0.58-0.72]	
0.676	3 (279)	0.04 [0.01-0.25]	0.92 [0.77-0.98]	0.52 [0.07-3.59]	0.54 [0.08-3.49]	1.03 [0.96-1.12]	0.53 [0.42-0.73]		
1.292	2 (154)	0.81 [0.59-0.92]	0.27 [0.16-0.44]	1.65 [0.50-5.36]	1.12 [0.87-1.44]	0.67 [0.26-1.73]	0.67 [0.57-0.75]		

AUC, area under curve; FIB-4, fibrosis-4; NFS, nonalcoholic fatty liver disease (NAFLD) fibrosis score; CI, confidence interval; DOR, diagnostic odds ratio; LR+, positive likelihood ratio; LR-, negative likelihood ratio.

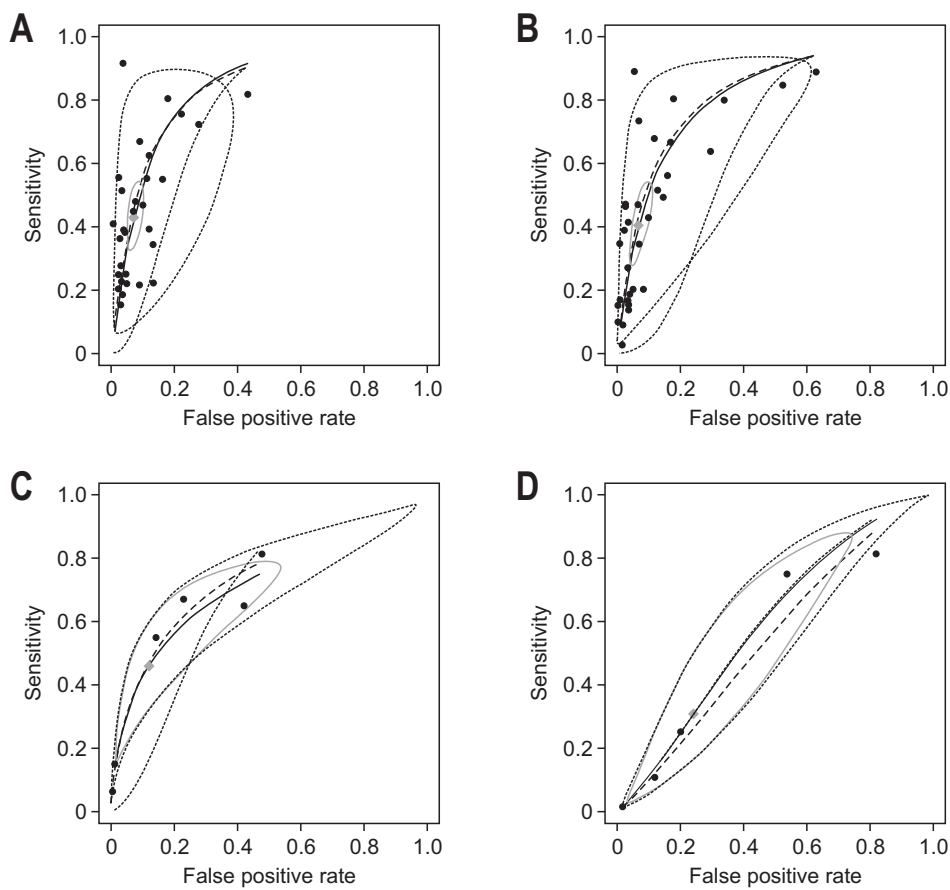


Fig. 2. The summary receiver operating characteristic curve. (A) FIB-4 index for predicting advanced fibrosis (\geq F3), (B) NFS for predicting advanced fibrosis (\geq F3), (C) FIB-4 index for predicting significant fibrosis (\geq F2), and (D) NFS for predicting significant fibrosis (\geq F2) in patients with NAFLD. FIB-4, fibrosis-4; NFS, nonalcoholic fatty liver disease (NAFLD) fibrosis score.

a lower cutoff value for predicting AF, a value from -1.98 to -1.03 was most frequently used (23 studies), and as a higher cutoff value, 4.39 to 4.8 was predominantly used (31 studies). For NFS, pooled sensitivity was 0.38 (95% CI, 0.28 to 0.50) and pooled specificity was 0.94 (95% CI, 0.90 to 0.96) (Table 2). Pooled DOR with 95% CI was 10.16 (7.18 to 14.37) with I^2 of 85% ($p < 0.01$), indicating heterogeneity of the meta-analysis (Table 2, Supplementary Fig. 2). Summary statistics of FIB-4 at various thresholds for prediction of AF and forest plots are presented in Supplementary Table 2 and Supplementary Fig. 2. The AUC of SROC was 0.74 (95% CI, 0.71 to 0.79) (Fig. 2B).

4. FIB-4 and NFS for predicting SF (\geq F2)

Studies on SF were relatively scarce, compared to studies on AF (32 studies vs 6 studies) (Table 2). Also, while the cutoff for AF was consistent for each study, the cutoff for SF differed significantly between studies. In regard to the FIB-4 index, the pooled sensitivity was 0.42 (95% CI, 0.16 to 0.73) and pooled specificity was 0.93 (95% CI, 0.56 to 0.99). The pooled DOR with 95% CI was 9.71 (2.43 to 38.70) with I^2 of 0% ($p = 0.54$). The AUC of SROC was 0.68 (95% CI, 0.65 to 0.76) (Fig. 2C). Summary statistics of FIB-4 at various thresholds for prediction of SF and forest plots

are presented in Supplementary Table 3 and Supplementary Fig. 3.

In regard to NFS, the pooled sensitivity was 0.25 (95% CI, 0.02 to 0.82) and pooled specificity was 0.76 (95% CI, 0.37 to 0.94). The pooled DOR with 95% CI was 1.16 (0.37 to 3.57) with I^2 of 0% ($p = 0.54$), indicating homogeneity. The AUC of SROC was 0.60 (95% CI, 0.52 to 0.69) (Fig. 2D). Summary statistics of NFS at various thresholds for prediction of SF and forest plots are presented in Supplementary Table 4 and Supplementary Fig. 4.

5. Difference in diagnosis accuracy according to region and body mass index

Finally, we analyzed whether the accuracy of NFS or FIB-4 differs according to the study area and body mass index (BMI) (Table 3). Regions could be classified into three categories, Asia/Europe/America, except for the two studies that were conducted globally on two or more continents. We found that the accuracy of FIB-4 or NFS to predict F3 relatively increased in Europe (FIB-4: pooled DOR 16.37 , NFS: pooled DOR 21.94) compared with Asia (FIB-4: pooled DOR 6.09 , NFS: pooled DOR 6.22) or America (FIB-4: pooled DOR 6.23 , NFS: pooled DOR 3.70). For BMI, individual patient BMI data could not be obtained,

Table 3. Summary Sensitivities, Specificities, Diagnostic Odds Ratio, Positive Likelihood Ratio, Negative Likelihood Ratio and AUC of FIB-4 Index and NFS According to Region or Body Mass Index for Prediction of Advanced Fibrosis

Variable	Cutoff	No. of samples (No. of patients)	Summary statistics (95% CI)					
			Sensitivity	Specificity	DOR	LR+	LR-	AUC
Region								
FIB-4 (F3)	Asia	12 (3,388)	0.48 (0.36-0.60)	0.86 (0.80-0.91)	6.09 (4.44-8.36)	3.63 (2.77-4.75)	0.59 (0.49-0.72)	0.74 (0.71-0.77)
	Europe	6 (1,978)	0.61 (0.43-0.76)	0.91 (0.83-0.95)	16.37 (7.93-33.74)	6.95 (3.91-12.32)	0.42 (0.27-0.64)	0.80 (0.78-0.82)
	America	8 (4,155)	0.51 (0.37-0.65)	0.85 (0.71-0.93)	6.23 (2.51-15.47)	3.52 (1.72-7.22)	0.56 (0.42-0.75)	0.73 (0.68-0.78)
NFS (F3)	Asia	12 (3,407)	0.38 (0.24-0.54)	0.90 (0.83-0.95)	6.22 (4.52-8.57)	4.20 (3.07-5.76)	0.67 (0.55-0.82)	0.70 (0.65-0.74)
	Europe	7 (2,098)	0.55 (0.39-0.71)	0.94 (0.85-0.98)	21.94 (10.03-47.94)	10.23 (4.43-23.64)	0.46 (0.33-0.65)	0.83 (0.80-0.86)
	America	8 (4,447)	0.50 (0.33-0.68)	0.78 (0.64-0.87)	3.70 (1.77-7.74)	2.33 (1.44-3.77)	0.62 (0.45-0.88)	0.72 (0.69-0.76)
Body mass index								
FIB-4 (F3)	<30 kg/m ²	12 (3,870)	0.50 (0.38-0.61)	0.88 (0.82-0.92)	7.67 (5.46-10.78)	4.33 (3.21-5.85)	0.56 (0.46-0.68)	0.75 (0.71-0.82)
	30 to <35 kg/m ²	10 (1,331)	0.51 (0.38-0.64)	0.88 (0.78-0.93)	7.92 (3.73-16.81)	4.33 (2.39-7.85)	0.54 (0.41-0.71)	0.74 (0.72-0.76)
	≥35 kg/m ²	4 (675)	0.60 (0.48-0.71)	0.82 (0.69-0.90)	7.31 (5.29-10.12)	3.49 (2.26-5.40)	0.47 (0.39-0.57)	0.77 (0.74-0.80)
NFS (F3)	<30 kg/m ²	12 (3,889)	0.35 (0.23-0.50)	0.93 (0.88-0.96)	8.17 (5.54-12.05)	5.60 (3.82-8.21)	0.68 (0.57-0.82)	0.71 (0.67-0.75)
	30 to <35 kg/m ²	9 (4,583)	0.51 (0.35-0.66)	0.84 (0.75-0.91)	5.96 (2.77-12.85)	3.41 (1.98-5.86)	0.57 (0.41-0.78)	0.73 (0.70-0.67)
	≥35 kg/m ²	6 (2,205)	0.55 (0.33-0.76)	0.83 (0.61-0.94)	6.37 (3.82-10.63)	3.38 (1.81-6.32)	0.53 (0.37-0.75)	0.80 (0.74-0.85)

AUC, area under curve; FIB-4, fibrosis-4; NFS, nonalcoholic fatty liver disease (NAFLD) fibrosis score; CI, confidence interval; DOR, diagnostic odds ratio; LR+, positive likelihood ratio; LR-, negative likelihood ratio.

so the average BMI for each study was used. In all studies, BMI was 25 kg/m² or higher (minimum, 26.1 kg/m²), and BMI was classified into three groups: <30 kg/m², 30 to <35 kg/m², and 35 kg/m² or higher. As a result of stratification analysis, BMI values did not significantly affect the accuracy of NFS or FIB-4. Meta-regression analysis was added for the effect of BMI on DOR of FIB-4 or NFS, but it was not significant as well (Table 4).

DISCUSSION

This systematic review and meta-analysis of 36 relevant studies indicated that the FIB-4 index and NFS can be effectively used to predict the degree of liver fibrosis in NAFLD. Additionally, our results demonstrated that the diagnostic accuracy of FIB-4 and NFS is relatively higher in predicting AF than in SF. Our study holds significance in its ability to assist clinicians in deciding treatments for NAFLD patients, by accurately predicting the degree of liver fibrosis.

Out of all panels based on serological markers, the NFS system was the most frequently studied. NFS is a scoring system based on 733 NAFLD patients diagnosed by liver biopsy.¹⁰ In previous studies, the diagnostic ability of NFS for hepatic fibrosis was AUC 0.82-0.88. Given such precedent, two cutoff values (<-1.455 [low probability, negative predictive value 88% to 93%] and >0.676 [high probability, positive predictive value 82% to 90%]) were proposed.¹⁰ Existing meta-studies on the diagnostic predictive ability of NFS in NAFLD have reported up to AUC 0.73-0.86, which is consistent with the results of our study.⁴⁹⁻⁵¹ However, in the NFS scoring system, there are cases in which either high or low probability for advanced liver fibrosis cannot be classified (indeterminate probability); in such cases, a liver biopsy may be necessary.⁵²

On the other hand, the FIB-4 index was created by Sterling *et al.*⁹ using a cohort of 832 human immunodeficiency virus/hepatitis C virus co-infected patients. Compared to NFS, FIB-4 has been widely validated for liver diseases of relatively diverse etiologies.⁵³⁻⁵⁵ In previous meta-studies, the diagnostic AUC of FIB-4 for AF in NAFLD was 0.76. Like NFS, FIB-4 has two cutoffs: FIB-4 <1.30 can be interpreted as low possibility of AF, while FIB-4 >2.67 can be interpreted as high possibility of AF. In addition to NFS and FIB-4, other noninvasive markers exist.

However, in regard to NAFLDs, NFS and FIB-4 showed higher diagnostic abilities than other noninvasive panels have. Notably, both markers demonstrated a diagnostic ability of AF similar to that of magnetic resonance elastography.⁵⁶ Therefore, most of the NAFLD guidelines recommend NFS and FIB-4 as screening tools for diagnosing AF.⁵⁷⁻⁵⁹

First, our study found that FIB-4 has better diagnostic performance than NFS in predicting AF (AUC 0.76 vs AUC 0.74). Similar results have been manifested in previous studies. According to a meta-analysis conducted on 13,046 NAFLD subjects in 2017, AUROC of FIB-4 and NFS for the prediction of AF were 0.80 and 0.78, respectively, indicating that FIB-4 index has higher diagnostic accuracy than NFS.⁴⁹ Also, in a meta-analysis of 5,735 NAFLD patients in 2021, AUROC of FIB-4 index and NFS were 0.76 (95% CI, 0.74 to 0.77) and 0.73 (95% CI, 0.71 to 0.75), respectively.⁶⁰

Secondly, according to our data, the ability to predict SF was inferior to the AF predictive abilities of both markers. We can assume diagnostic accuracy of each scoring system is higher for more severe types of patients, as supported by previous studies. According to a study by Xiao *et al.*,⁴⁹ AUROC values of FIB-4 and NFS increased according to SF, AF, and liver cirrhosis (0.75, 0.80 and 0.85 for FIB-4 and 0.72, 0.78 and 0.83 for NFS).

Table 4. Meta-Regression for Diagnostic Odds Ratio of Each Measurement

Variable	FIB-4 for AF		NFS for AF	
	Coefficient (95% CI)	p-value	Coefficient (95% CI)	p-value
Mean age, yr	0.063 [0.022 to 0.104]	0.002	0.095 [0.048 to 0.143]	<0.001
Proportion of male, %	-0.019 [-0.037 to -0.001]	0.034	-0.022 [-0.043 to -0.001]	0.040
DM, %	0.005 [-0.011 to 0.021]	0.561	0.013 [-0.001 to 0.028]	0.065
HTN, %	0.028 [-0.001 to 0.045]	0.055	0.021 [-0.003 to 0.046]	0.089
Dyslipidemia, %	-0.009 [-0.067 to 0.048]	0.747	0.005 [-0.028 to 0.039]	0.758
Mean BMI, kg/m ²	0.004 [-0.086 to 0.095]	0.919	0.062 [-0.032 to 0.156]	0.196
Mean waist, cm	-0.009 [-0.075 to 0.056]	0.774	0.001 [-0.070 to 0.071]	0.997
Mean AST	-0.020 [-0.041 to 0.001]	0.060	-0.023 [-0.047 to 0.001]	0.055
Mean ALT	-0.014 [-0.026 to -0.002]	0.017	-0.018 [-0.032 to -0.005]	0.007

FIB-4, fibrosis-4; AF, advanced fibrosis; NFS, nonalcoholic fatty liver disease (NAFLD) fibrosis score; CI, confidence interval; DM, diabetes mellitus; HTN, hypertension; BMI, body mass index; AST, aspartate aminotransferase; ALT, alanine aminotransferase.

Third finding of our study is that the ability of FIB-4 or NFS to predict AF is lower than that of previous reports. This is interpreted because of the diversity of the patient population used in our study analysis. In fact, when comparing existing studies or meta-analyses, the AUC of FIB-4 or NFS gradually decreases as the number of patients or studies analyzed in the study increases. When analyzing 145 patients, AUC of FIB-4 and NFS was 0.86 and 0.81, respectively.⁶¹ In 1,038 patients, AUC of FIB-4 was 0.849.⁵⁰ When analyzed in 5,735 patients from 37 studies, the AUC for FIB-4 and NFS was 0.76 and 0.73, which is almost consistent with our study results.⁶⁰ In our study, the AUCs of FIB-4 and NFS were 0.76 and 0.74, respectively.

Finally, our study found that accuracy of FIB-4 or NFS to predict F3 relatively increased in Europe compared with Asia or America. This is interpreted because the proportion of Caucasians in the cohort study in which FIB or NFS was developed was relatively high, 79% for FIB-4 and 90% for NFS.^{9,10,62} Therefore, the accuracy of FIB or NFS is relatively low in Asian countries or in the multiracial America continent.

Advantages of these markers in NAFLD include low cost, quick diagnosis, and easy repeatability. In primary care, the use of noninvasive markers can increase early detection of AF, decrease avoidable referral of patients with mild diseases and ensure cost-effectiveness.⁶³ Therefore, many experts recommend implementing a two-tier approach to improve resource utilization.^{18,57,64} On the other hand, compared to transient elastography or magnetic resonance elastography, disadvantages such as the low ability to diagnose AF should always be noted.⁶⁰

The strength of our meta-analysis is its focus on comparing the diagnostic accuracies of two noninvasive and routinely usable scoring systems to predict the degree of liver fibrosis in NAFLD patients. To the best of our knowledge, this meta-analysis has the largest sample size, amongst those that compare the diagnostic accuracy of noninvasive scoring systems for AF and SF in NAFLD patients. However, there are some limitations to consider. First, the biggest limitation of this study is that it lacks new information compared to the existing meta-analysis related to FIB-4 or NFS. Before starting the study, we reviewed the existing meta-analysis literature and found that more literature than expected was not included in the analysis. Therefore, in order to obtain more accurate results, it was determined that accurate inclusion criteria should be re-applied, and as a result, a wider range of papers could be analyzed than existing meta-analysis papers. Eventually, the results of our study are not very different from the existing meta-analyses, and thus provide few new revelations. However, our study has clinical significance as the most

extensive analytical research on this subject. Second, limiting our review to manuscripts published only in English may have caused publication bias. We conducted a funnel plot and Egger's test, later to confirm that our study has no publication bias with $p=0.135$ in FIB-4 for AF. However, there was a publication bias in FIB-4 for SF, and NSF for AF and SF with $p<0.05$, which may limit the credibility of our results. Third, only two noninvasive scoring systems were the focus of our analysis. Other noninvasive scoring systems such as BMI, aspartate aminotransferase/alanine aminotransferase ratio, diabetes (BARD) score, and aspartate aminotransferase to platelet ratio tests were not considered.^{50,65} Fourth, our research focused on liver fibrosis and did not consider the degree of hepatic steatosis. Furthermore, the included studies did not provide sufficient information on the patients' duration of NAFLD or past treatments, which can affect the incidence and severity of liver fibrosis.

In summary, both FIB-4 index and NFS were useful in predicting the degree of liver fibrosis in NAFLD. The diagnostic accuracy of these scoring systems was higher in predicting AF than in SF. Thus, the FIB-4 index and NFS may be considered as alternative diagnostic methods to liver biopsy when predicting the level of fibrosis in NAFLD.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2021R1G1A1007886), and in part by the Soonchunhyang University Research Fund.

AUTHOR CONTRIBUTIONS

Study concept and design: D.W.J. Provision of study materials or patients: M.C., H.W.L. Collection and assembly of data: S.H.K., Y.C., S.B.A., D.S.S. Data analysis and interpretation: B.L., S.H., J.J.Y. Manuscript writing: S.H., J.L., J.J.Y. Final approval of manuscript: all authors.

ORCID

Sangsoo Han	https://orcid.org/0000-0002-9709-3332
Miyoung Choi	https://orcid.org/0000-0002-2424-9965
Bora Lee	https://orcid.org/0000-0002-6322-5712
Hye-Won Lee	https://orcid.org/0000-0002-3552-3560
Seong Hee Kang	https://orcid.org/0000-0002-0783-509X
Yuri Cho	https://orcid.org/0000-0002-4488-5352
Sang Bong Ahn	https://orcid.org/0000-0001-7419-5259
Do Seon Song	https://orcid.org/0000-0002-8278-4878
Dae Won Jun	https://orcid.org/0000-0002-2875-6139
Jieun Lee	https://orcid.org/0000-0001-9494-2493
Jeong-Ju Yoo	https://orcid.org/0000-0002-7802-0381

SUPPLEMENTARY MATERIALS

Supplementary materials can be accessed at <https://doi.org/10.5009/gnl210391>.

REFERENCES

- Rinella ME. Nonalcoholic fatty liver disease: a systematic review. *JAMA* 2015;313:2263-2273.
- Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease: meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 2016;64:73-84.
- Williams CD, Stengel J, Asike MI, et al. Prevalence of non-alcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. *Gastroenterology* 2011;140:124-131.
- Sesti G, Sciacqua A, Fiorentino TV, Perticone M, Succurro E, Perticone F. Association between noninvasive fibrosis markers and cardio-vascular organ damage among adults with hepatic steatosis. *PLoS One* 2014;9:e104941.
- Charlton M. Nonalcoholic fatty liver disease: a review of current understanding and future impact. *Clin Gastroenterol Hepatol* 2004;2:1048-1058.
- Farrell GC, Larter CZ. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. *Hepatology* 2006;43(2 Suppl 1):S99-S112.
- Nalbantoglu IL, Brunt EM. Role of liver biopsy in nonalcoholic fatty liver disease. *World J Gastroenterol* 2014;20:9026-9037.
- Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* 2005;128:1898-1906.
- Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006;43:1317-1325.
- Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007;45:846-854.
- Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
- Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158:280-286.
- Aida Y, Abe H, Tomita Y, et al. Serum immunoreactive collagen IV detected by monoclonal antibodies as a marker of severe fibrosis in patients with non-alcoholic fatty liver disease. *J Gastrointest Liver Dis* 2015;24:61-68.
- Anstee QM, Lawitz EJ, Alkhoury N, et al. Noninvasive tests accurately identify advanced fibrosis due to NASH: baseline data from the STELLAR trials. *Hepatology* 2019;70:1521-1530.
- Balakrishnan M, Seth A, Cortes-Santiago N, et al. External validation of four point-of-care noninvasive scores for predicting advanced hepatic fibrosis in a predominantly Hispanic NAFLD population. *Dig Dis Sci* 2021;66:2387-2393.
- Boursier J, Guillaume M, Leroy V, et al. New sequential combinations of non-invasive fibrosis tests provide an accurate diagnosis of advanced fibrosis in NAFLD. *J Hepatol* 2019;71:389-396.
- Chan WK, Nik Mustapha NR, Mahadeva S. A novel 2-step approach combining the NAFLD fibrosis score and liver stiffness measurement for predicting advanced fibrosis. *Hepatol Int* 2015;9:594-602.
- Chan WK, Treeprasertsuk S, Goh GB, 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;17:2570-2580.
- Cui J, Ang B, Haufe W, et al. Comparative diagnostic accuracy of magnetic resonance elastography vs. eight clinical prediction rules for non-invasive diagnosis of advanced fibrosis in biopsy-proven non-alcoholic fatty liver disease: a prospective study. *Aliment Pharmacol Ther* 2015;41:1271-1280.
- Demir M, Lang S, Schlattjan M, et al. NIKEL: a new inexpensive and non-invasive scoring system to exclude advanced fibrosis in patients with NAFLD. *PLoS One* 2013;8:e58360.
- de Carli MA, de Carli LA, Correa MB, Junqueira G Jr, Tovo CV, Coral GP. Performance of noninvasive scores for the diagnosis of advanced liver fibrosis in morbidly obese with nonalcoholic fatty liver disease. *Eur J Gastroenterol Hepatol*

- 2020;32:420-425.
22. Goh GB, Pagadala MR, Dasarathy J, et al. Clinical spectrum of non-alcoholic fatty liver disease in diabetic and non-diabetic patients. *BBA Clin* 2014;3:141-145.
 23. Joo SK, Kim W, Kim D, et al. Steatosis severity affects the diagnostic performances of noninvasive fibrosis tests in non-alcoholic fatty liver disease. *Liver Int* 2018;38:331-341.
 24. Jun DW, Kim SG, Park SH, et al. External validation of the non-alcoholic fatty liver disease fibrosis score for assessing advanced fibrosis in Korean patients. *J Gastroenterol Hepatol* 2017;32:1094-1099.
 25. Kakisaka K, Suzuki Y, Fujiwara Y, et al. Evaluation of ballooned hepatocytes as a risk factor for future progression of fibrosis in patients with non-alcoholic fatty liver disease. *J Gastroenterol* 2018;53:1285-1291.
 26. Kao WY, Chang IW, Chen CL, et al. Fibroscan-based score to predict significant liver fibrosis in morbidly obese patients with nonalcoholic fatty liver disease. *Obes Surg* 2020;30:1249-1257.
 27. Kaya E, Bakir A, Kani HT, Demirtas CO, Keklikkiran C, Yilmaz Y. Simple noninvasive scores are clinically useful to exclude, not predict, advanced fibrosis: a study in Turkish patients with biopsy-proven nonalcoholic fatty liver disease. *Gut Liver* 2020;14:486-491.
 28. Kim D, Kim WR, Talwalkar JA, Kim HJ, Ehman RL. Advanced fibrosis in nonalcoholic fatty liver disease: non-invasive assessment with MR elastography. *Radiology* 2013;268:411-419.
 29. Labenz C, Huber Y, Kalliga E, et al. Predictors of advanced fibrosis in non-cirrhotic non-alcoholic fatty liver disease in Germany. *Aliment Pharmacol Ther* 2018;48:1109-1116.
 30. Lang S, Farowski F, Martin A, et al. Prediction of advanced fibrosis in non-alcoholic fatty liver disease using gut microbiota-based approaches compared with simple non-invasive tools. *Sci Rep* 2020;10:9385.
 31. Lum J, Cheah M, Leow WQ, et al. Clinical profile of non-alcoholic fatty liver disease in nonobese patients. *J Gastroenterol Hepatol* 2021;36:257-261.
 32. McPherson S, Anstee QM, Henderson E, Day CP, Burt AD. Are simple noninvasive scoring systems for fibrosis reliable in patients with NAFLD and normal ALT levels? *Eur J Gastroenterol Hepatol* 2013;25:652-658.
 33. Meneses D, Oliveira A, Corripio R, et al. Performance of noninvasive liver fibrosis scores in the morbid obese patient, same scores but different thresholds. *Obes Surg* 2020;30:2538-2546.
 34. Nasr P, Hilliges A, Thorelius L, Kechagias S, Ekstedt M. Contrast-enhanced ultrasonography could be a non-invasive method for differentiating none or mild from severe fibrosis in patients with biopsy proven non-alcoholic fatty liver disease. *Scand J Gastroenterol* 2016;51:1126-1132.
 35. Ooi GJ, Burton PR, Doyle L, et al. Modified thresholds for fibrosis risk scores in nonalcoholic fatty liver disease are necessary in the obese. *Obes Surg* 2017;27:115-125.
 36. Patel YA, Gifford EJ, Glass LM, et al. Identifying nonalcoholic fatty liver disease advanced fibrosis in the Veterans Health Administration. *Dig Dis Sci* 2018;63:2259-2266.
 37. Pérez-Gutiérrez OZ, Hernández-Rocha C, Candia-Balboa RA, et al. Validation study of systems for noninvasive diagnosis of fibrosis in nonalcoholic fatty liver disease in Latin population. *Ann Hepatol* 2013;12:416-424.
 38. Petta S, Vanni E, Bugianesi E, et al. The combination of liver stiffness measurement and NAFLD fibrosis score improves the noninvasive diagnostic accuracy for severe liver fibrosis in patients with nonalcoholic fatty liver disease. *Liver Int* 2015;35:1566-1573.
 39. Petta S, Wai-Sun Wong V, Bugianesi E, et al. Impact of obesity and alanine aminotransferase levels on the diagnostic accuracy for advanced liver fibrosis of noninvasive tools in patients with nonalcoholic fatty liver disease. *Am J Gastroenterol* 2019;114:916-928.
 40. Siddiqui MS, Yamada G, Vuppalanchi R, et al. Diagnostic accuracy of noninvasive fibrosis models to detect change in fibrosis stage. *Clin Gastroenterol Hepatol* 2019;17:1877-1885.
 41. Singh A, Gosai F, Siddiqui MT, et al. Accuracy of noninvasive fibrosis scores to detect advanced fibrosis in patients with type-2 diabetes with biopsy-proven nonalcoholic fatty liver disease. *J Clin Gastroenterol* 2020;54:891-897.
 42. Treeprasertsuk S, Piyachaturawat P, Soontornmanokul T, Wisedopas-Klaikaew N, Komolmit P, Tangkijavanich P. Accuracy of noninvasive scoring systems to assess advanced liver fibrosis in Thai patients with nonalcoholic fatty liver disease. *Asian Biomed* 2016;10:s49-s55.
 43. Wong VW, Wong GL, Chim AM, et al. Validation of the NAFLD fibrosis score in a Chinese population with low prevalence of advanced fibrosis. *Am J Gastroenterol* 2008;103:1682-1688.
 44. Wong VW, Vergniol J, Wong GL, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology* 2010;51:454-462.
 45. Xun YH, Fan JG, Zang GQ, et al. Suboptimal performance of simple noninvasive tests for advanced fibrosis in Chinese patients with nonalcoholic fatty liver disease. *J Dig Dis* 2012;13:588-595.
 46. Yang M, Jiang L, Wang Y, et al. Step layered combination of noninvasive fibrosis models improves diagnostic accuracy of advanced fibrosis in nonalcoholic fatty liver disease. *J Gastrointest Liver Dis* 2019;28:289-296.
 47. Yoneda M, Imajo K, Eguchi Y, et al. Noninvasive scoring systems in patients with nonalcoholic fatty liver disease with normal alanine aminotransferase levels. *J Gastroenterol*

- 2013;48:1051-1060.
48. Zhou YJ, Ye FZ, Li YY, et al. Individualized risk prediction of significant fibrosis in non-alcoholic fatty liver disease using a novel nomogram. *United European Gastroenterol J* 2019;7:1124-1134.
 49. Xiao G, Zhu S, Xiao X, Yan L, Yang J, Wu G. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: a meta-analysis. *Hepatology* 2017;66:1486-1501.
 50. Sun W, Cui H, Li N, 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;46:862-870.
 51. Jaruvongvanich V, Wijarnpreecha K, Ungprasert P. The utility of NAFLD fibrosis score for prediction of mortality among patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis of cohort study. *Clin Res Hepatol Gastroenterol* 2017;41:629-634.
 52. Bril F, Sninsky JJ, Baca AM, et al. Hepatic steatosis and insulin resistance, but not steatohepatitis, promote atherogenic dyslipidemia in NAFLD. *J Clin Endocrinol Metab* 2016;101:644-652.
 53. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: clinical prediction rules and blood-based biomarkers. *J Hepatol* 2018;68:305-315.
 54. Sonneveld MJ, Brouwer WP, Chan HL, et al. Optimisation of the use of APRI and FIB-4 to rule out cirrhosis in patients with chronic hepatitis B: results from the SONIC-B study. *Lancet Gastroenterol Hepatol* 2019;4:538-544.
 55. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology* 2007;46:32-36.
 56. Imajo K, Kessoku T, Honda Y, et al. Magnetic resonance imaging more accurately classifies steatosis and fibrosis in patients with nonalcoholic fatty liver disease than transient elastography. *Gastroenterology* 2016;150:626-637.
 57. Kang SH, Lee HW, Yoo JJ, et al. KASL clinical practice guidelines: management of nonalcoholic fatty liver disease. *Clin Mol Hepatol* 2021;27:363-401.
 58. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the Management of Non-Alcoholic Fatty Liver Disease. *Obes Facts* 2016;9:65-90.
 59. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. *Gastroenterology* 2012;142:1592-1609.
 60. Mózes FE, Lee JA, Selvaraj EA, et al. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. *Gut*. Epub 2021 May 17. <https://doi.org/10.1136/gutjnl-2021-324243>.
 61. 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;59:1265-1269.
 62. Torriani FJ, Rodriguez-Torres M, Rockstroh JK, et al. Peginterferon Alfa-2a plus ribavirin for chronic hepatitis C virus infection in HIV-infected patients. *N Engl J Med* 2004;351:438-450.
 63. Srivastava A, Jong S, Gola A, et al. Cost-comparison analysis of FIB-4, ELF and fibroscan in community pathways for non-alcoholic fatty liver disease. *BMC Gastroenterol* 2019;19:122.
 64. Salva-Pastor N, Chávez-Tapia NC, Uribe M, Nuño-Lámbarri N. The diagnostic and initial approach of the patient with non-alcoholic fatty liver disease: role of the primary care provider. *Gastroenterol Hepatol Bed Bench* 2019;12:267-277.
 65. Tapper EB, Krajewski K, Lai M, et al. Simple non-invasive biomarkers of advanced fibrosis in the evaluation of non-alcoholic fatty liver disease. *Gastroenterol Rep (Oxf)* 2014;2:276-280.