

VOLUME 30 SUPPLEMENT September 2024

pISSN 2287-2728
eISSN 2387-285X

CLINICAL and MOLECULAR HEPATOLOGY

The forum for latest knowledge of hepatobiliary diseases

KASL guidelines for NIT in CLD

Optimal cut-offs of NIT for NAFLD

Diagnostic accuracy of FIB-4 in NAFLD patients with
T2DM

Prediction of HCC recurrence using VCTE

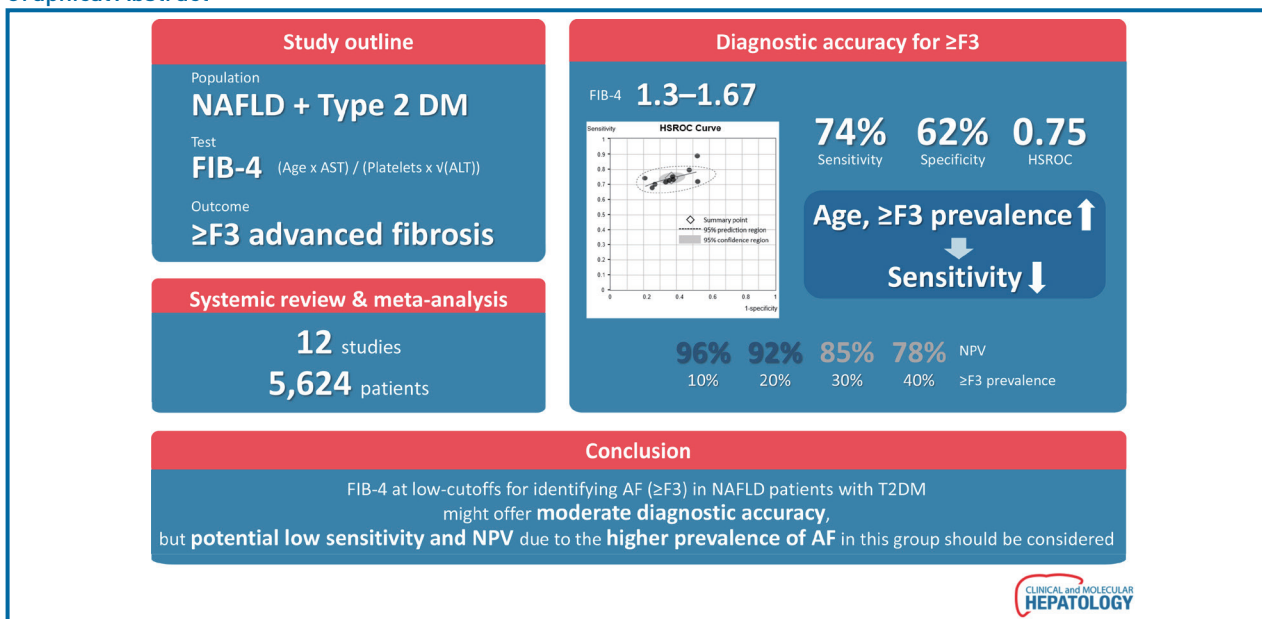
HCC prediction using VCTE-determined LSM

Diagnostic accuracy of the Fibrosis-4 index for advanced liver fibrosis in nonalcoholic fatty liver disease with type 2 diabetes: A systematic review and meta-analysis

Ji Won Han^{1,2,*}, Hee Yeon Kim^{1,3,*}, Jung Hwan Yu⁴, Mi Na Kim⁵, Young Eun Chon⁶, Ji Hyun An⁷, Young-Joo Jin⁴, Miyoung Choi⁸, Seung Up Kim⁵, Han Ah Lee⁹, and Dae Won Jun¹⁰

¹The Catholic University Liver Research Center, College of Medicine, The Catholic University of Korea, Seoul; ²Department of Internal Medicine, College of Medicine, Seoul St. Mary's Hospital, The Catholic University of Korea, Seoul; ³Department of Internal Medicine, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul; ⁴Department of Internal Medicine, Inha University Hospital, Inha University School of Medicine, Incheon; ⁵Department of Internal Medicine, Yonsei University College of Medicine, Seoul; ⁶Department of Gastroenterology, CHA Bundang Medical Center, CHA University, Seongnam; ⁷Department of Gastroenterology and Hepatology, Hanyang University College of Medicine, Guri; ⁸Clinical Evidence Research, National Evidence-based Healthcare Collaborating Agency (NECA), Seoul; ⁹Department of Internal Medicine, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul; ¹⁰Department of Internal Medicine, Hanyang University College of Medicine, Seoul, Korea

Graphical Abstract



Study Highlights

- This study is the first to provide a comprehensive analysis of the diagnostic accuracy of FIB-4 specifically in NAFLD patients with T2DM, focusing on its effectiveness at varying cutoff points. The meta-analysis highlights the potential of FIB-4, particularly at low cutoffs, for excluding advanced fibrosis with moderate sensitivity and specificity. Importantly, the study emphasizes the need to consider AF prevalence when interpreting negative predictive values and explores the potential of adjusting cutoff values for improved diagnostic accuracy. The findings contribute valuable insights for optimizing noninvasive screening strategies in this high-risk population.

Background/Aims: The Fibrosis-4 index (FIB-4) is a noninvasive test widely used to rule out advanced liver fibrosis (AF) in patients with nonalcoholic fatty liver disease (NAFLD). However, its diagnostic accuracy in NAFLD patients with type 2 diabetes mellitus (T2DM) is controversial due to the high prevalence of AF in this population.

Methods: Research focusing on the diagnostic accuracy of FIB-4 for liver fibrosis as validated by liver histology in NAFLD patients with T2DM was included, and 12 studies (n=5,624) were finally included in the meta-analysis. Sensitivity, specificity, hierarchical summary receiver operating characteristic (HSROC), positive predictive values (PPVs), and negative predictive values (NPVs) at low cutoffs (1.3–1.67) and high cutoffs (2.67–3.25) for ruling in and out AF were calculated.

Results: At low cutoffs, the meta-analysis revealed a sensitivity of 0.74, specificity of 0.62, and HSROC of 0.75. At high cutoffs, the analysis showed a sensitivity of 0.33, specificity of 0.92, and HSROC of 0.85, suggesting FIB-4 as useful for identifying or excluding AF. In subgroup analyses, high mean age and F3 prevalence were associated with lower sensitivity. The calculated NPV and PPV were 0.82 and 0.49 at low cutoffs, whereas the NPV was 0.28 and the PPV was 0.70 at high cutoffs. There were insufficient estimated NPVs <0.90 at a hypothesized prevalence of AF >30% at an FIB-4 cutoff range of 1.3–1.67.

Conclusions: Collectively, FIB-4 has moderate diagnostic accuracy for identifying or excluding AF in NAFLD patients with T2DM, but more evidence must be accumulated due to the limited number of currently reported studies and their heterogeneity. (*Clin Mol Hepatol* 2024;30(Suppl):S147-S158)

Keywords: FIB-4; Diabetes; NAFLD; Meta-analysis

INTRODUCTION

Among noninvasive liver fibrosis assessments, the Fibrosis-4 index (FIB-4) is a representative serum marker for the screening liver fibrosis, originally developed using cohorts of patients with chronic hepatitis C and human immunodeficiency virus co-infections.^{1,2} It underpins an algorithm that excludes patients unlikely to have advanced liver fibrosis ($\geq$ F3) (AF) based on its high sensitivity. Subse-

quently, other noninvasive tests such as vibration-controlled transient elastography (VCTE) or magnetic resonance elastography (MRE) can be applied. Commonly used threshold values for FIB-4 for the exclusion of AF include 1.3 or 1.45.³ In patients with nonalcoholic fatty liver disease (NAFLD), a lower threshold (FIB-4 <1.3) identifies a low-risk group for AF, where follow-up observation is deemed sufficient, while it is recommended that intermediate- (FIB-4 1.3–2.67) or high-risk (FIB-4 >2.67) pa-

Corresponding author : Han Ah Lee

Department of Internal Medicine, Chung-Ang University Hospital, Chung-Ang University College of Medicine, 102 Heukseok-ro, Dongjak-gu, Seoul 06973, Korea
Tel: +82-2-6299-1408, Fax: +82-2-6299-2469, E-mail: amelia86@naver.com
<https://orcid.org/0000-0003-4082-1121>

Dae Won Jun

Department of Internal Medicine, Hanyang University College of Medicine, 222-1, Wangsimni-ro Seongdong-gu, Seoul 04763, Korea
Tel: +82-2-2220-8338, Fax: +82-2-2298-9183, E-mail: noshin@hanyang.ac.kr
<https://orcid.org/0000-0002-2875-6139>

*Equally contributed.

Editor: Heejoon Jang, Seoul National University, Korea

Received : May 4, 2024 / **Revised :** Jul. 23, 2024 / **Accepted :** Jul. 24, 2024

Abbreviations:

FIB-4, Fibrosis-4 index; AF, advanced liver fibrosis; VCTE, vibration-controlled transient elastography; MRE, magnetic resonance elastography; NAFLD, nonalcoholic fatty liver disease; T2DM, type 2 diabetes mellitus; PPV, positive predictive value; NPV, negative predictive value; NAFLD, non-alcoholic fatty liver disease; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; BMI, body mass index; HbA1c, hemoglobin A1c; ALT, alanine aminotransferase

tients undergo further fibrosis assessments such as VCTE, MRE, or liver biopsy.⁴

However, recent studies have indicated that NAFLD patients with type 2 diabetes mellitus (T2DM) have significantly higher rates of fibrosis progression compared to those without T2DM over periods of 4, 8, and 12 years (with cumulative incidence rates of 24% vs. 20%, 60% vs. 50%, and 93% vs. 76%, respectively).⁵ Within the low-risk group (FIB-4 <1.3), T2DM was identified as an independent risk factor for liver-related complications, and a higher prevalence of liver stiffness greater than 8 kPa with VCTE was observed in diabetic patients (77%) compared to non-diabetic patients (40%).⁶ Similarly, even among NAFLD patients with T2DM who were classified as low risk based on FIB-4, a substantial proportion (51.5%) had >8 kPa liver stiffness measured by VCTE, raising concerns about the reduced sensitivity of FIB-4 in this population.⁷

Drawing on existing research, a recent proposal in the United States suggested re-evaluating FIB-4 every 1–2 years in NAFLD patients with T2DM and every 2–3 years in those without,⁸ although research into the diagnostic performance and utility of FIB-4 in this population remains limited. A recent meta-analysis of individual patient data, which collected data from 1,780 biopsy-proven NAFLD patients with T2DM across five studies, found the diagnostic AUC of FIB-4 for AF to be 0.75, with sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) at a threshold of 1.3 of 73%, 62%, 62%, and 72%, respectively,⁹ reflecting the potential insufficient sensitivity and NPV of FIB-4 alone for ruling out AF patients in this population.

However, that study was not able to compile data from all existing studies due to the nature of individual patient data analysis and only collected data from studies published between 2009–2018. Due to the heterogeneity of the reported diagnostic accuracy of FIB-4 in this population, a systemic review and meta-analysis collecting various studies with larger samples and recently reported data is needed. The present study incorporated 12 studies to investigate the diagnostic accuracy of FIB-4 using data from biopsy-confirmed AF in NAFLD patients with T2DM and showed not only the diagnostic values of FIB-4 at low and high cut-offs, but also analyzed potential factors able to affect the accuracy of this representative noninvasive marker. Importantly, we focused on whether an FIB-4 cutoff of 1.3 might

be appropriate to rule out patients with AF in terms of sensitivity and NPV in this subpopulation.

MATERIALS AND METHODS

This meta-analysis was conducted in accordance with a protocol previously registered with PROSPERO (the International Prospective Register of Systematic Reviews [ID no. CRD42024454538]). The execution of this systematic review and meta-analysis was in line with the recommendations outlined in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Extension for Diagnostic Test Accuracy.

Inclusion criteria, exclusion criteria, and study outcome

Patients with non-alcoholic fatty liver disease (NAFLD) with T2DM are considered NAFLD with T2DM because NAFLD specifically includes metabolic risk factors like T2DM in its definition, reflecting the same underlying metabolic disturbances present in NAFLD with T2DM.¹⁰ Therefore, studies focusing on the diagnostic accuracy of FIB-4 for liver fibrosis as validated by liver histology in NAFLD patients with T2DM were considered suitable for inclusion in this study. Selection criteria are as follows: (1) diagnosed NAFLD cases and (2) patients who underwent FIB-4 with liver biopsy. Eligible study formats for inclusion were randomized controlled trials, cross-sectional analyses, and cohort studies either prospective or retrospective in nature. We excluded (1) individual case studies, (2) case series involving fewer than five patients, (3) literature reviews, (4) studies of patients with chronic viral infections like hepatitis B or C, (5) studies of patients with excessive alcohol intake, (6) studies of patients with diagnoses of fatty liver based on imaging or serological methods without histological confirmation, (7) studies of patients with type I diabetes, (8) studies not published in English, and (9) potential inclusion of patients with excessive alcohol consumption (defined as <30 g/day in men and <20 g/day in women). The primary outcome in this meta-analysis was the diagnostic accuracy of FIB-4 compared against liver histology in NAFLD patients with T2DM.

Search strategy

We conducted paper searches across multiple databases, including PubMed (MEDLINE), EMBASE, the Cochrane Library, and the Korean Medical Database for studies in English published from January 1, 2008, to October 20th, 2023. Our search used the Population, Index Test, Comparison, and Outcome (PICO) model, with specific keywords detailed in Supplementary Table 1. Search terms included “NAFLD,” “FIB-4,” and “diabetes.” We used a combination of unstructured text and specialized indexing terms, including Medical Subject Headings in PubMed and Emtree terms in EMBASE, to enhance the accuracy and comprehensiveness of our literature search. Details on our search methodology and the individual results from each database are documented in Supplementary Table 2.

Selection of studies and extraction of data

During the initial phase of study selection, reviewers J.W.H. and M.N.K. independently screened titles and abstracts to extract those considered relevant. Subsequent to their individual review of the full-text articles, any differences in their evaluations were reconciled through consultation with a third reviewer, either D.W.J. or S.U.K. In addition, these reviewers conducted subsequent reviews, encompassing a detailed examination of the entire texts and an assessment of the possibility of bias. Both reviewers independently extracted data on study attributes and outcomes, organizing this information in a uniform manner, and any inconsistencies were addressed and resolved through discussions with D.W.J. and S.U.K.

Quality and risk-of-bias assessment

To assess the potential for bias, we applied the Cochrane risk-of-bias assessment tool; pertinent details are presented in Supplementary Figure 1A. In cases of disagreement, discussions were held with reviewers (J.W.H. and M.N.K.) to reach a consensus. The evaluation of bias risk used the QUADAS-2 tool.¹¹ Summaries of these bias assessments are available in Supplementary Figure 1B. Publication bias was evaluated using funnel plots and Egger's test (Supplementary Fig. 2).

Statistical analysis

The diagnostic meta-analysis procedure, focusing on sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), PPV, NPV, and diagnostic odds ratio (DOR), was conducted using RevMan 5 (Cochrane, London, UK), MetaDisc version 1.4 (Clinical BioStatistics Unit-Hospital Ramon y Cajal, Madrid, Spain), and R 4.4.1 (R Core Team). Summary receiver operating characteristic (SROC), hierarchical summary receiver operating characteristic (HSROC) curves, and forest plots were generated, and the meta-regression analysis was performed. To identify the heterogeneity observed across the studies, a random-effects model was applied, and the I^2 statistic was calculated.

RESULTS

Study characteristics

A review of titles and abstracts in the chosen databases identified 28 related studies. Of these, 16 were excluded due to a non-relevant patient group ($n=3$) or index test ($n=2$), absence of comparative tests ($n=2$), and a deficit of data ($n=9$). In the end, 12 studies qualified for the subsequent analysis, which is detailed in Supplementary Figure 3.

These 12 studies observing AF in patients with T2DM and NAFLD as determined by FIB-4 and liver biopsy were included in this systematic review and meta-analysis as shown in Table 1.^{9,12-22} The studies comprised one multinational study, one study from Japan, and the remaining studies from non-Asian regions, with two being prospective cohort studies and the remaining 10 being cross-sectional studies. The risk of bias was as outlined in Supplementary Figure 2. The low and high cutoff values for FIB-4 used across these studies ranged from 1.3–1.67 and 2.67–3.25, respectively. The mean age of subjects ranged from 50–61.9 years. Few Asian patients were included in each study except the Japanese study, although some studies did not report the percentage of Asian patients. The mean body mass index (BMI) ranged from 27.4–37 kg/m², with the lowest values in the study from Japan.¹⁶ Mean serum levels of hemoglobin A1c (HbA1c) and alanine aminotransferase

Table 1. Study characteristics

Ref.	Publication year	Region	Design	Patient no.	Asian (%)	HbA1c (%)	Age (years)	BMI (kg/m ²)	ALT (U/L)	F3-4 ^a no.	F3-4 ^a prev. (%)	FIB-4 cutoff
Bertot et al. ¹²	2018	Australia	Cross-sectional	124	4.0	NR	58	37	52	60	48.4	3.25
Bril et al. ¹³	2019	U.S.	Cross-sectional	191	0.7	7.2	59.6	34.4	53.6	27	14.1	1.3
Alkayali et al. ¹⁴	2020	Turkey	Cross-sectional	166	NR	6.5	50	32.4	57	49	29.5	1.3/2.67
Bril et al. ¹⁵	2020	U.S.	Cross-sectional	162	0.6	7.1	57.7	34.5	54.7	31	19.1	1.67/3.25
Ishiba et al. ¹⁶	2020	Japan	Cross-sectional	311	100	6.7	60	27.4	64	54	17.4	1.3/1.67
Singh et al. ¹⁷	2020	U.S.	Cross-sectional	1,157	0.8	6.7	51.1	35.5	28	367	31.7	1.45/2.67
Bril et al. ¹⁸	2022	U.S.	Cross-sectional	169	1.8	6.5	54	33.9	46	20	11.8	1.3
Almera et al. ¹⁹	2023	U.S.	Prospective	130	21.4	6.9	61.9	31.9	43	39	30.0	1.3/2.67
Boursier et al. ²⁰	2023	France	Cross-sectional	523	NR	NR	60.1	32.4	55	127	24.3	1.3/2.67
Castera et al. ²¹	2023	France	Prospective	319	NR	7.5	59	32.0	49	124	38.0	0.91/1.93
Pennisi et al. ⁹	2023	Europe, China, Korea, Japan	Cross-sectional (IPDMA)	1,780	NR	NR	56.7	31.4	53	822	46.2	0.97/1.3/2.67
Singh et al. ²²	2023	U.S.	Cross-sectional	592	NR	NR	55.9	NR	NR	201	34.0	1.45/2.67

Data are presented as number or mean.

FIB-4, fibrosis-4; NR, not recorded; IPDMA, individual patient data meta-analysis; prev., prevalence; Ref., reference; U.S., United States of America.

^aDiagnosed by liver biopsy.

(ALT) were 6.5–7.5%, and 28–64 U/L, respectively. The prevalence of AF measured by liver biopsy varied widely across the reviewed literature, ranging from 11.4–48.4%. Notably, the study by Pennisi et al.⁹ reported a prevalence of AF at 46.2%, and Bertot et al.¹² reported 48.4%, which were higher than the prevalence rates of 11.8–34.0% reported in other studies.

Diagnostic accuracy of FIB-4 for excluding and including AF

To evaluate diagnostic accuracy, a meta-analysis was ini-

tially conducted focusing on FIB-4 lower cutoff values of 1.3–1.67 (low cutoffs), and 5,181 patients across 10 studies were included in the analysis (Table 2). There was no significant publication bias among these studies (Supplementary Fig. 2A). The results revealed a sensitivity of 0.74, specificity of 0.62, PLR of 1.98, NLR of 0.42, DOR of 4.69, SROC of 0.75, and HSROC of 0.75 (Fig. 1A and Supplementary Fig. 4). Significantly, the low heterogeneity in sensitivity, NLR, and DOR across studies underscores the reliability of our findings. These findings suggest that FIB-4 at lower cutoff values is useful for diagnosing AF, and that lower cutoff values of FIB-4 can be employed to exclude AF.

Table 2. Results of meta-analysis for diagnostic accuracy of FIB-4 in patients with NAFLD accompanied by T2DM for AF

Cutoff	Study (patient) number	SROC/HSROC	Sensitivity (95% CI)	I ² (P)	Specificity (95% CI)	I ² (P)
1.3–1.67	10 (5,181)	0.75/0.75	0.74 (0.72–0.76)	0.22 (0.241)	0.62 (0.60–0.64)	0.90 (<0.001)
2.67–3.25	9 (4,945)	0.82/0.85	0.33 (0.31–0.36)	0.94 (<0.001)	0.92 (0.91–0.93)	0.81 (<0.001)
Cutoff	PLR (95% CI)	I ² (P)	NLR (95% CI)	I ² (P)	DOR (95% CI)	I ² (P)
1.3–1.67	1.98 (1.77–2.20)	70.9 (<0.001)	0.42 (0.39–0.46)	0.0 (0.820)	4.69 (4.11–5.35)	0.0 (0.482)
2.67–3.25	4.51 (3.51–5.81)	60.6 (0.009)	0.67 (0.58–0.77)	0.91 (<0.001)	7.49 (5.06–11.1)	0.74 (<0.001)

AF, advanced liver fibrosis; CI, confidence interval; DOR, diagnostic odds ratio; FIB-4, fibrosis-4; HSROC, hierarchical summary receiver operating characteristic; NAFLD, nonalcoholic fatty liver disease; NLR, negative likelihood ratio; PLR, positive likelihood ratio; SROC, summary receiver operating characteristic; T2DM, type 2 diabetes mellitus.

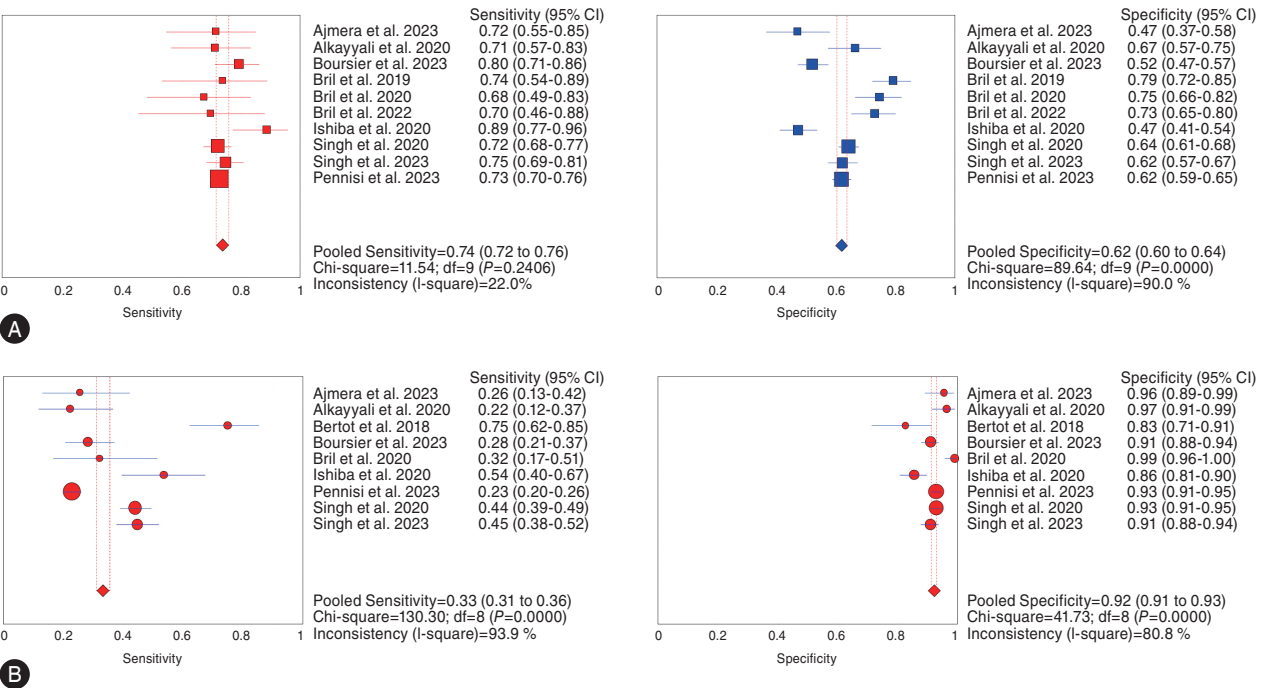


Figure 1. Results of meta-analysis for sensitivity and specificity. (A) At low FIB-4 cutoffs (1.3–1.67) for excluding AF. (B) At high FIB-4 cutoffs (2.67–3.25) for including AF. FIB-4, fibrosis-4; AF, advanced liver fibrosis; CI, confidence interval.

Separately, to investigate the diagnostic accuracy at FIB-4 cutoff values of 2.67–3.25 (high cutoffs), additional analysis was conducted using 4,945 patients across nine studies (Table 2). There was no significant publication bias among these studies (Supplementary Fig. 2B). The results showed a sensitivity of 0.33, specificity of 0.92, PLR of 4.51, NLR of 0.67, DOR of 7.49, SROC of 0.82, and HSROC of 0.85 (Supplementary Fig. 5). However, there was significant heterogeneity among included studies. Nevertheless, in patients with T2DM, FIB-4 at higher cutoff values might serve as a diagnostic tool for AF, exhibiting high specificity.

Influences of various factors on the diagnostic accuracy

Supplementary Table 3 shows the results of a meta-regression analysis evaluating the influence of various factors on diagnosis of AF according to FIB-4 cutoff points. Factors analyzed included the percentage of Asian participants, HbA1c level, age, BMI, ALT level, and the prevalence of AF determined by biopsy. The results, indicated by coefficients, confidence intervals, and *P*-values, revealed mostly not significant associations between these variables and the likelihood of diagnosing AF, with a few exceptions; Asian, for instance, displayed a slight but significant nega-

tive relationship with AF diagnosis at the high cutoff point, while age presented a small negative association at the same threshold.

Table 3 shows the subgroup analyses for sensitivity and specificity of low cutoff FIB-4 values. We could not complete a subgroup analysis with regard to race or study region because just one study was performed in Asia. The subgroup analyses showed that high mean age and F3 prevalence were associated with lower sensitivity, whereas low mean HbA1c, age, and BMI were associated with lower specificity, although the findings of this subgroup analysis should be interpreted with caution due to the limited number of studies and the existence of study heterogeneity. Nevertheless, these findings suggest the complexity of diagnosing AF and highlight the potential need for incorporating additional factors to improve accuracy in future studies.

Sensitivity of FIB-4 for screening AF according to the cutoff points

Next, we assessed the diagnostic accuracy of FIB-4 at low cutoff values of 1.3, 1.45, and 1.67 (Table 4). At a 1.3 cutoff, the analysis yielded an SROC of 0.74, sensitivity of 0.74, specificity of 0.60, PLR of 1.93, NLR of 0.43, and DOR of 4.70. For the 1.45 cutoff, the sensitivity was also

Table 3. Subgroup analysis for sensitivity and specificity at low FIB-4 cutoffs

	Sensitivity	<i>P</i> -value	Specificity	<i>P</i> -value
Mean HbA1c ≥ 7 vs. < 7 (%)	0.69 (0.54–0.81)/0.74 (0.70–0.78)	0.528	0.74 (0.68–0.79)/0.62 (0.60–0.65)	0.009
Mean age ≥ 60 vs. < 60 (years)	0.73 (0.71–0.75)/0.80 (0.75–0.85)	0.01	0.65 (0.63–0.67)/0.50 (0.46–0.54)	0.005
Mean BMI ≥ 34 vs. < 34 (kg/m ²)	0.72 (0.68–0.76)/0.74 (0.72–0.77)	0.406	0.68 (0.65–0.71)/0.59 (0.56–0.61)	0.006
Mean ALT ≥ 50 vs. < 50 (U/L)	0.74 (0.72–0.77)/0.72 (0.68–0.76)	0.406	0.61 (0.59–0.63)/0.64 (0.61–0.67)	0.103
F3 prevalence ≥ 30 vs. < 30 (%)	0.73 (0.71–0.75)/0.78 (0.73–0.82)	0.046	0.62 (0.60–0.64)/0.61 (0.58–0.64)	0.587

FIB-4, fibrosis-4; BMI, body mass index; ALT, alanine aminotransferase.

Table 4. Results of meta-analysis for diagnostic accuracy according to FIB-4 cutoffs

Cutoff	SROC	Sensitivity (95% CI)	Specificity (95% CI)	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)
$< 1.0^a$	0.71/0.75	0.89/0.90	0.33/0.38	1.33/1.44	0.33/0.29	3.99/4.96
1.3	0.74	0.74 (0.72–0.77)	0.60 (0.58–0.62)	1.93 (1.66–2.25)	0.43 (0.38–0.47)	4.70 (3.66–6.03)
1.45	0.74	0.74 (0.72–0.76)	0.61 (0.58–0.63)	1.94 (1.74–2.16)	0.42 (0.38–0.47)	4.66 (4.07–5.33)
1.67	0.75	0.74 (0.72–0.76)	0.62 (0.60–0.64)	1.98 (1.77–2.20)	0.42 (0.39–0.46)	4.69 (4.11–5.35)

^aData from Castera et al.²¹ (cutoff, 0.91) and Pennisi et al.⁹ (cutoff, 0.97).

CI, confidence interval; DOR, diagnostic odds ratio; FIB-4, fibrosis-4; NLR, negative likelihood ratio; PLR, positive likelihood ratio; SROC, summary receiver operating characteristic.

0.74, and other values also remained consistent with a minor improvement in specificity to 0.61 and an adjusted PLR of 1.94. The highest threshold of 1.67 led to a slight increase in SROC to 0.75 and specificity to 0.62 along with slight changes in both PLR and NLR, indicating a marginal enhancement in diagnostic precision, although the sensitivity was also not affected. Among reviewed articles, two applied extremely lower cutoffs of FIB-4 (0.91²¹ and 0.97⁹), which showed respective sensitivities of 0.89 and 0.90. Considering a high prevalence of AF in NAFLD patients with T2DM, these findings suggest that either the currently recommended FIB-4 cutoff of 1.3 or larger cutoffs might have an insufficient sensitivity <0.8. Although lowering the cutoff to <1.0 might maximize sensitivity, the usefulness and diagnostic accuracy of such an approach should be investigated in future studies.

NPV and PPV of FIB-4 at low and high cutoffs

We opted to focus on NPV due to its critical role in excluding AF in patients classified as low-risk by FIB-4. This approach ensures that cases of AF are not missed, improving the precision of the diagnostic process. In the reviewed literature, the values for NPV ranged from 0.73–0.95, dem-

onstrating variability across studies. The calculated NPV from meta-analysis for FIB-4 low cutoffs was 0.82 (Fig. 2A). Among reviewed articles, two applied extremely low cutoffs of FIB-4 (0.91²¹ and 0.97⁹), which led to NPVs of 0.84 and 0.79, respectively. In addition, the calculated PPV of low cutoffs was 0.49. We also evaluated NPV and PPV at FIB-4 high cutoffs. PPV ranged from 0.44 to 0.91, and calculated PPV from meta-analysis was 0.70 (Fig. 2B). NPV was 0.28. Although there was significant heterogeneity among the studies, which necessitates cautious interpretation, this variability might be attributable to differences in the presence of AF in NAFLD patients with T2DM across regions and races, considering that both NPV and PPV are significantly affected by prevalence.

Estimated NPV of FIB-4 at low cutoffs according to the prevalence of AF in NAFLD patients with T2DM

Among reviewed articles, one applying extremely low cutoffs of FIB-4 yielded an NPV of 0.84,¹⁵ which was not different from the calculated NPV of 0.82 using the FIB-4 cutoff range of 1.3–1.67. Table 5 shows the potential impact

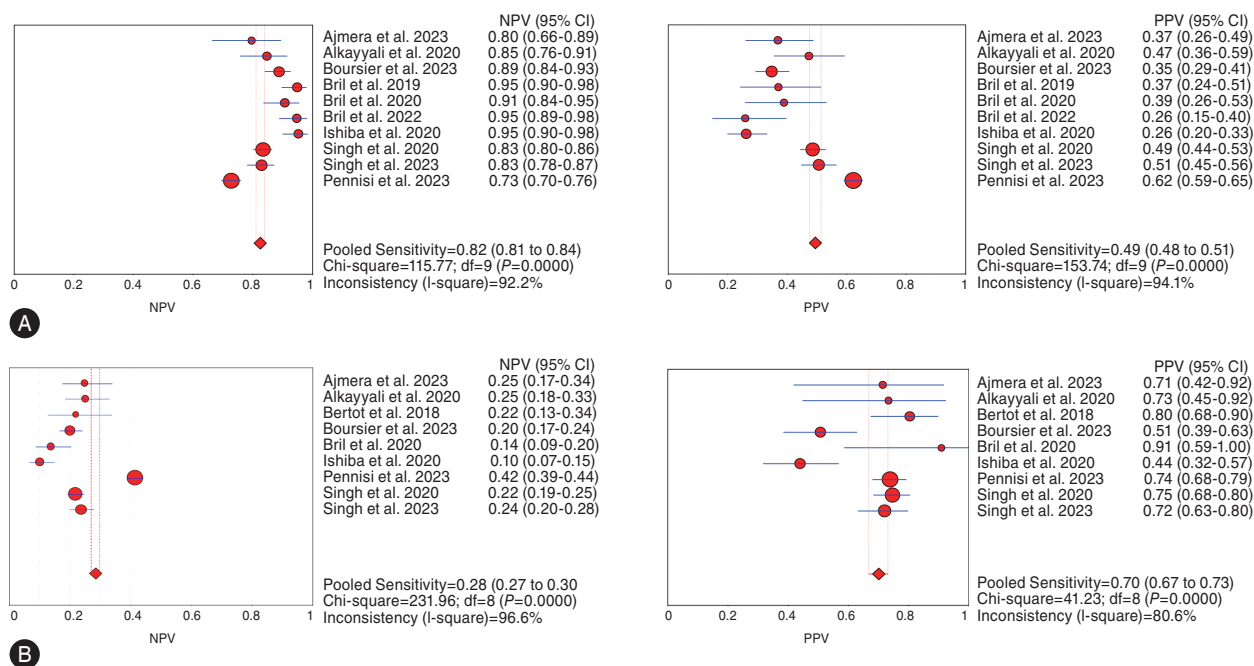


Figure 2. Results of meta-analysis for PPV and NPV. (A) At low FIB-4 cutoffs (1.3–1.67) for excluding AF. (B) At high FIB-4 cutoffs (2.67–3.25) for including AF. FIB-4, fibrosis-4; NPV, negative predictive value; AF, advanced liver fibrosis; CI, confidence interval.

of varying prevalence rates of AF on the NPV and PPV derived from low FIB-4 cutoffs and minimum or maximum sensitivity and specificity values from included studies in the meta-analysis. With an increase in AF prevalence, there is a noticeable decline in the NPV for FIB-4 at low cutoffs. This downward trend in NPV was evident across a range of prevalence rates (from 0.01–0.5); for instance, at a prevalence rate of 0.01, the NPV decreases from 1.00 to 0.70 at a prevalence rate of 0.5, reflecting a significant reduction in the ability of FIB-4 to accurately identify individuals without AF as the prevalence increases. In particular, there were insufficient estimated NPVs below 0.90 when we hypothesized a prevalence of AF >30% at FIB-4 values of 1.3–1.67. Conversely, PPVs exhibited an upward trend, starting from 0.02 at a prevalence rate of 0.01 and escalating to 0.66 at a prevalence rate of 0.5, indicating an enhanced capacity to correctly identify individuals with AF as the prevalence rises. This analysis highlights the critical impact of AF prevalence on the NPV of FIB-4 for excluding AF, emphasizing how the NPV decreases with higher prevalence rates.

DISCUSSION

In this systemic review and meta-analysis, we performed an in-depth analysis to assess the diagnostic accuracy of FIB-4 for excluding or including AF in NAFLD patients with T2DM. A recent individual patient data meta-analysis of five studies showed the sensitivity, specificity, NPV, PPV, and AUC of noninvasive tests, including FIB-4, for diagnos-

ing AF,⁹ but our study focused on the diagnostic accuracy of FIB-4 and potential related factors. Importantly, we further analyzed the potential influence of AF prevalence on the NPV at a low cutoff, which raises the necessity of further investigation to uncover the real prevalence of AF in this population. Of note, we first simulated NPV and PPV according to various prevalence rates of AF in NAFLD patients with T2DM, which could provide useful information for screening AF using FIB-4.

This diagnostic meta-analysis for FIB-4 with a cutoff range of 1.3–1.67 (low cutoffs) demonstrated a sensitivity of 0.74, specificity of 0.62, and a diagnostic AUC of 0.75. Notably, the consistency across studies in terms of sensitivity, NLR, and DOR was sufficient, lending credibility to the results. In a recent meta-analysis involving general NAFLD population, a FIB-4 value between 1.02 and 1.45 to ensure diagnostic accuracy for AF had a sensitivity of 0.69, specificity of 0.54, and a diagnostic AUC of 0.73,²³ which are values comparable to those from our analysis. These findings suggest that FIB-4, when used with lower cutoff values, might not be significantly different between diabetic and general NAFLD populations in terms of sensitivity and specificity.

For the high FIB-4 cutoff range of 2.67–3.25, the meta-analysis revealed a sensitivity of 0.33, specificity of 0.92, and diagnostic AUC of 0.85, although there was significant inconsistency across studies. Recent data on general NAFLD population for AF using FIB-4 values of 2.67–3.25 showed a sensitivity of 0.34–0.39, specificity of 0.95, and diagnostic AUC of 0.74–0.75.²³ The high variability across synthesized studies calls for cautious interpretation, and it

Table 5. Estimated NPV and PPV of FIB-4 at low cutoffs according to the prevalence of AF in NAFLD patients with T2DM

Calculated sensitivity and specificity					Minimum sensitivity and specificity					Maximum sensitivity and specificity				
AF prev.	Sen	Spe	NPV	PPV	AF prev.	Sen	Spe	NPV	PPV	AF prev.	Sen	Spe	NPV	PPV
0.01	0.74	0.62	1.00	0.02	0.01	0.68	0.47	0.99	0.01	0.01	0.8	0.79	1.00	0.04
0.03			0.99	0.06	0.03			0.98	0.04	0.03			0.99	0.11
0.05			0.98	0.09	0.05			0.97	0.06	0.05			0.99	0.17
0.1			0.96	0.18	0.1			0.93	0.12	0.1			0.97	0.30
0.2			0.91	0.33	0.2			0.85	0.24	0.2			0.94	0.49
0.3			0.85	0.45	0.3			0.77	0.35	0.3			0.90	0.62
0.4			0.78	0.56	0.4			0.69	0.46	0.4			0.86	0.72
0.5			0.70	0.66	0.5			0.59	0.56	0.5			0.80	0.79

AF, advanced liver fibrosis; FIB-4, fibrosis-4; NAFLD, nonalcoholic fatty liver disease; NPV, negative predictive value; PPV, positive predictive value; Sen, sensitivity; Spe, specificity; T2DM, type 2 diabetes mellitus.

remains uncertain whether additional tests for diagnosing AF in NAFLD patients with T2DM can be conclusively ruled out based on current evidence. However, in this population, the high cutoff of FIB-4 might be a highly specific test for diagnosing AF, comparable to the general NAFLD population. While the specificity of the test was high in high-risk patients with FIB-4 >2.67, current evidence cannot eliminate the need for further noninvasive tests for fibrosis or liver biopsy to confirm fibrosis. Furthermore, complementing FIB-4 with related factors like race or age, as observed in the meta-regression analysis, could enhance the power of FIB-4 for ruling-in AF in this population. Overall, our findings suggest that the diagnostic AUC of FIB-4 is equally effective in NAFLD patients with T2DM as it is in general NAFLD patients.

We found that FIB-4 at low cutoffs had a sensitivity of 0.74 and NPV of 0.82; thus, it should be complemented by the adjustment of cutoff values or supplementation of other related factors of liver fibrosis due to an insufficient accuracy to rule out AF in NAFLD patients with T2DM. Both European and American guidelines have suggested that VCTE or other imaging-based tests be considered as initial tests in the tertiary setting, particularly in high-risk patients like those with T2DM.^{24,25} However, due to the cost-effectiveness and the limited availability of VCTE, additional advancement in serum-based noninvasive tests is necessary to achieve a sensitivity or NPV >0.9 for the initial screening of AF in this population. Furthermore, there is controversy regarding the suitability of the commonly used cutoff value of 1.3 for excluding AF in this population. In this regard, one study suggested lowering the FIB-4 cutoff to 1.0 to maintain a sensitivity >90% in patients with T2DM.¹⁹ Other studies have confirmed that lowering the FIB-4 cutoff to 0.91 resulted in a sensitivity of 0.89,²¹ while lowering it to 0.97 resulted in a sensitivity of 0.9.⁹ Therefore, further research is necessary to determine the appropriate cutoff values, and the potential lowered cutoff in this population should also be considered in real clinical practice. Moreover, there are few studies to investigate T2DM-specific scores adding various clinical factors to the FIB-4 to minimize undetected population by initial screening for AF,^{22,26} although they need to be validated in future studies. As previously reported, the combination with other serum tests such as M2BPGi could also enhance diagnostic accuracy.²⁷ Sequential combinations with imaging-based tests

such as VCTE should also be considered due to its better diagnostic accuracy.⁹ A recent guideline suggested that FIB-4 can be reassessed periodically every 2–3 years in patients without T2DM, but 1–2 years in patients with T2DM.⁸ Therefore, frequent, repeated assessment in this population could minimize the potential non-detection of AF patients.

NPV and PPV are influenced by sensitivity, which underscores the need to consider lower cutoff points to rule out AF in NAFLD patients with T2DM. However, they can be critically influenced by AF prevalence, although the sensitivity and specificity can be maintained regardless. Recent prospective cohort studies in the United States have shown a higher prevalence of AF detected via MRE, around 14% in patients with T2DM.¹⁹ Other cohort studies have confirmed a significantly higher prevalence of AF in patients with T2DM compared to those without (33% vs. 19%) through biopsy.⁵ A recent meta-analysis targeting general NAFLD population also showed that the NPV of FIB-4 at a cutoff of 1.3 varies with the prevalence of AF.²⁸ Thus, it is crucial to recognize that NPV and PPV are highly dependent on disease prevalence which can vary significantly across different populations and studies. As shown in Table 5, high prevalence can inflate PPV, whereas NPV can be overly optimistic in low-prevalence settings, masking the true likelihood of false negatives. Therefore, it is important to contextualize these values within the specific prevalence rates of the study population. Additionally, due to this variability, sensitivity and specificity could be highlighted as more stable measures, providing a clearer indication of a diagnostic performance independent of prevalence. In addition, studies investigating the prevalence of AF in NAFLD patients with T2DM should be performed, and large datasets of various races and regions should be accumulated to determine accurate cutoff points and appropriate strategies to screen AF patients using FIB-4.

Despite including a significant number of studies investigating biopsy-confirmed NAFLD patients with T2DM, there are several limitations in this study. Our findings may not be applicable to east Asian regions with varying rates of NAFLD and T2DM because most included studies were from Western countries analyzing non-Asian participants. In addition, selection bias may have arisen from the clinicians' discretionary decisions to perform liver biopsies, and the true prevalence of AF may be affected by this factor,

which is important for the NPV of FIB-4. Nevertheless, we incorporated current evidence and calculated the overall diagnostic accuracy of FIB-4 in this population.

Collectively, FIB-4 for identifying or excluding AF in NAFLD patients with T2DM might offer moderate diagnostic accuracy, although a limited number and heterogeneity of current published studies were observed. Concerns remain about potential low sensitivity and NPV due to the higher prevalence of AF in this group, and there is a pressing need for more research on the ideal cutoffs. Additionally, investigating the prevalence of AF across ethnicities and regions is mandatory for assessing the effectiveness of initial FIB-4 screening approaches.

Authors' contribution

JW Han, JH Yu, HA Lee, and DW Jun were responsible for the concept and design of the study, the data acquisition, analysis and interpretation of the data, and manuscript drafting. M Choi helped with the statistical analysis and data interpretation. MN Kim, YE Chon, HY Kim, JH An, YJ Jin, and SU Kim helped with the data interpretation.

Acknowledgements

The authors thank the Clinical Practice Guideline Committee for Noninvasive Tests (NIT) to Assess Liver Fibrosis in Chronic Liver Disease of the Korean Association for the Study of the Liver (KASL) for providing the opportunity to conduct this research.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2022R11-1A1A01063636 to J.W.H.). This research was also supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI23C1489 to J.W.H.).

Conflicts of Interest

The authors have no conflict of interest to declare.

SUPPLEMENTARY MATERIAL

Supplementary material is available at Clinical and Molecular Hepatology website (<http://www.e-cmh.org>).

REFERENCES

1. 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;43:1317-1325.
2. 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.
3. Stern C, Castera L. Identification of high-risk subjects in non-alcoholic fatty liver disease. *Clin Mol Hepatol* 2023;29(Suppl): S196-S206.
4. Reinson T, Buchanan RM, Byrne CD. Noninvasive serum biomarkers for liver fibrosis in NAFLD: current and future. *Clin Mol Hepatol* 2023;29(Suppl):S157-S170.
5. Huang DQ, Wilson LA, Behling C, Kleiner DE, Kowdley KV, Dasarthy S, et al. Fibrosis progression rate in biopsy-proven nonalcoholic fatty liver disease among people with diabetes versus people without diabetes: A multicenter study. *Gastroenterology* 2023;165:463-472.e5.
6. Boursier J, Hagström H, Ekstedt M, Moreau C, Bonacci M, Cure S, et al. Non-invasive tests accurately stratify patients with NAFLD based on their risk of liver-related events. *J Hepatol* 2022;76:1013-1020.
7. Gracen L, Hayward KL, Irvine KM, Valery PC, Powell EE. Low accuracy of FIB-4 test to identify people with diabetes at low risk of advanced fibrosis. *J Hepatol* 2022;77:1219-1221.
8. 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;77:1797-1835.
9. 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.
10. Kim GA, Moon JH, Kim W. Critical appraisal of metabolic dysfunction-associated steatotic liver disease: Implication of Janus-faced modernity. *Clin Mol Hepatol* 2023;29:831-843.
11. 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.
12. Bertot LC, Jeffrey GP, de Boer B, MacQuillan G, Garas G, Chin J, et al. Diabetes impacts prediction of cirrhosis and prognosis

- by non-invasive fibrosis models in non-alcoholic fatty liver disease. *Liver Int* 2018;38:1793-1802.
13. Bril F, Leeming DJ, Karsdal MA, Kalavalapalli S, Barb D, Lai J, et al. Use of plasma fragments of propeptides of type III, V, and VI procollagen for the detection of liver fibrosis in type 2 diabetes. *Diabetes Care* 2019;42:1348-1351.
14. Alkayyali T, Qutranji L, Kaya E, Bakir A, Yilmaz Y. Clinical utility of noninvasive scores in assessing advanced hepatic fibrosis in patients with type 2 diabetes mellitus: a study in biopsy-proven non-alcoholic fatty liver disease. *Acta Diabetol* 2020;57:613-618.
15. Bril F, McPhaul MJ, Caulfield MP, Clark VC, Soldevilla-Pico C, Firpi-Morell RJ, et al. Performance of plasma biomarkers and diagnostic panels for nonalcoholic steatohepatitis and advanced fibrosis in patients with type 2 diabetes. *Diabetes Care* 2020;43:290-297.
16. Ishiba H, Sumida Y, Seko Y, Tanaka S, Yoneda M, Hyogo H, et al. Type IV collagen 7S is the most accurate test for identifying advanced fibrosis in NAFLD with type 2 diabetes. *Hepatol Commun* 2020;5:559-572.
17. Singh A, Gosai F, Siddiqui MT, Gupta M, Lopez R, Lawitz E, 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.
18. Bril F, Godinez Leiva E, Lomonaco R, Shrestha S, Kalavalapalli S, Gray M, et al. Assessing strategies to target screening for advanced liver fibrosis among overweight and obese patients. *Metab Target Organ Damage* 2022;2:11.
19. Ajmera V, Cepin S, Tesfai K, Hofflich H, Cadman K, Lopez S, et al. A prospective study on the prevalence of NAFLD, advanced fibrosis, cirrhosis and hepatocellular carcinoma in people with type 2 diabetes. *J Hepatol* 2023;78:471-478.
20. Boursier J, Canivet CM, Costentin C, Lannes A, Delamarre A, Sturm N, et al. Impact of type 2 diabetes on the accuracy of noninvasive tests of liver fibrosis with resulting clinical implications. *Clin Gastroenterol Hepatol* 2023;21:1243-1251.e12.
21. Castera L, Laouenan C, Vallet-Pichard A, Vidal-Trécan T, Manchon P, Paradis V, et al. High prevalence of NASH and advanced fibrosis in type 2 diabetes: A prospective study of 330 outpatients undergoing liver biopsies for elevated ALT, using a low threshold. *Diabetes Care* 2023;46:1354-1362.
22. Singh A, Garg R, Lopez R, Alkhouri N. Diabetes liver fibrosis score to detect advanced fibrosis in diabetics with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol* 2022;20:e624-e626.
23. Han S, Choi M, Lee B, Lee HW, Kang SH, Cho Y, et al. Accuracy of noninvasive scoring systems in assessing liver fibrosis in patients with nonalcoholic fatty liver disease: A systematic review and meta-analysis. *Gut Liver* 2022;16:952-963.
24. American Diabetes Association Professional Practice Committee. 4. Comprehensive medical evaluation and assessment of comorbidities: Standards of care in diabetes-2024. *Diabetes Care* 2024;47(Suppl 1):S52-S76.
25. Archer AJ, Belfield KJ, Orr JG, Gordon FH, Abeysekera KW. EASL clinical practice guidelines: non-invasive liver tests for evaluation of liver disease severity and prognosis. *Frontline Gastroenterol* 2022;13:436-439.
26. Bazick J, Donithan M, Neuschwander-Tetri BA, Kleiner D, Brunt EM, Wilson L, et al. Clinical model for NASH and advanced fibrosis in adult patients with diabetes and NAFLD: Guidelines for Referral in NAFLD. *Diabetes Care* 2015;38:1347-1355.
27. Moon SY, Baek YH, Jang SY, Jun DW, Yoon KT, Cho YY, et al. Proposal of a novel serological algorithm combining FIB-4 and serum M2BPGi for advanced fibrosis in nonalcoholic fatty liver disease. *Gut Liver* 2024;18:283-293.
28. Mózes FE, Lee JA, Selvaraj EA, Jayaswal ANA, Trauner M, Boursier J, et al. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. *Gut* 2022;71:1006-1019.

Supplementary Table 1. The key research question and PICO

Population	Patients with non-alcoholic fatty liver disease accompanied by type 2 diabetes
Index test	Fibrosis-4 (FIB-4) index
Comparison	Liver histology
Outcome	Advanced fibrosis (≥F3)

PICO, Population, Index Test, Comparison, and Outcome.

Supplementary Table 2. Search methodology from each database

Ovid MEDLINE

Keywords			
P	1	exp Non-alcoholic Fatty Liver Disease/ OR (Non-alcoholic Fatty Liver Disease OR NAFLD OR (Steatohepatitid* adj Nonalcoholic) OR (fatty liver adj2 (non-alcoholic or nonalcoholic))).tw,kw	
	2	exp Diabetes Mellitus/ OR (Diabetes Mellitus OR diabete*).tw,kw.	
I	3	(fibrosis-4 or FIB-4).tw,kw.	
P&I	4	1 AND 2 AND 3	

Embase

Keywords			
P	1	'nonalcoholic fatty liver'/exp OR ('nonalcoholic fatty liver' OR NAFLD OR 'Steatohepatitides, Nonalcoholic' OR ('fatty liver' NEAR/2 (non-alcoholic or nonalcoholic))):ab,ti,kw	
	2	'diabetes mellitus'/exp OR ('Diabetes Mellitus' OR diabete*):ab,ti,kw	
I	3	(fibrosis-4 or FIB-4):ab,ti,kw	
P&I	4	#1 AND #2 AND #3	

Cochrane library

Keywords			
P	1	[mh "Non-alcoholic Fatty Liver Disease"] OR ("Non-alcoholic Fatty Liver Disease" OR NAFLD OR "Steatohepatitides, Nonalcoholic" OR ("fatty liver" NEAR/2 (non-alcoholic or nonalcoholic))):ab,ti,kw	
	2	[mh "Diabetes Mellitus"] OR ("Diabetes Mellitus" OR diabete*):ab,ti,kw	
I	3	(fibrosis-4 or FIB-4):ab,ti,kw	
P&I	4	#1 AND #2 AND #3	

KoreaMed

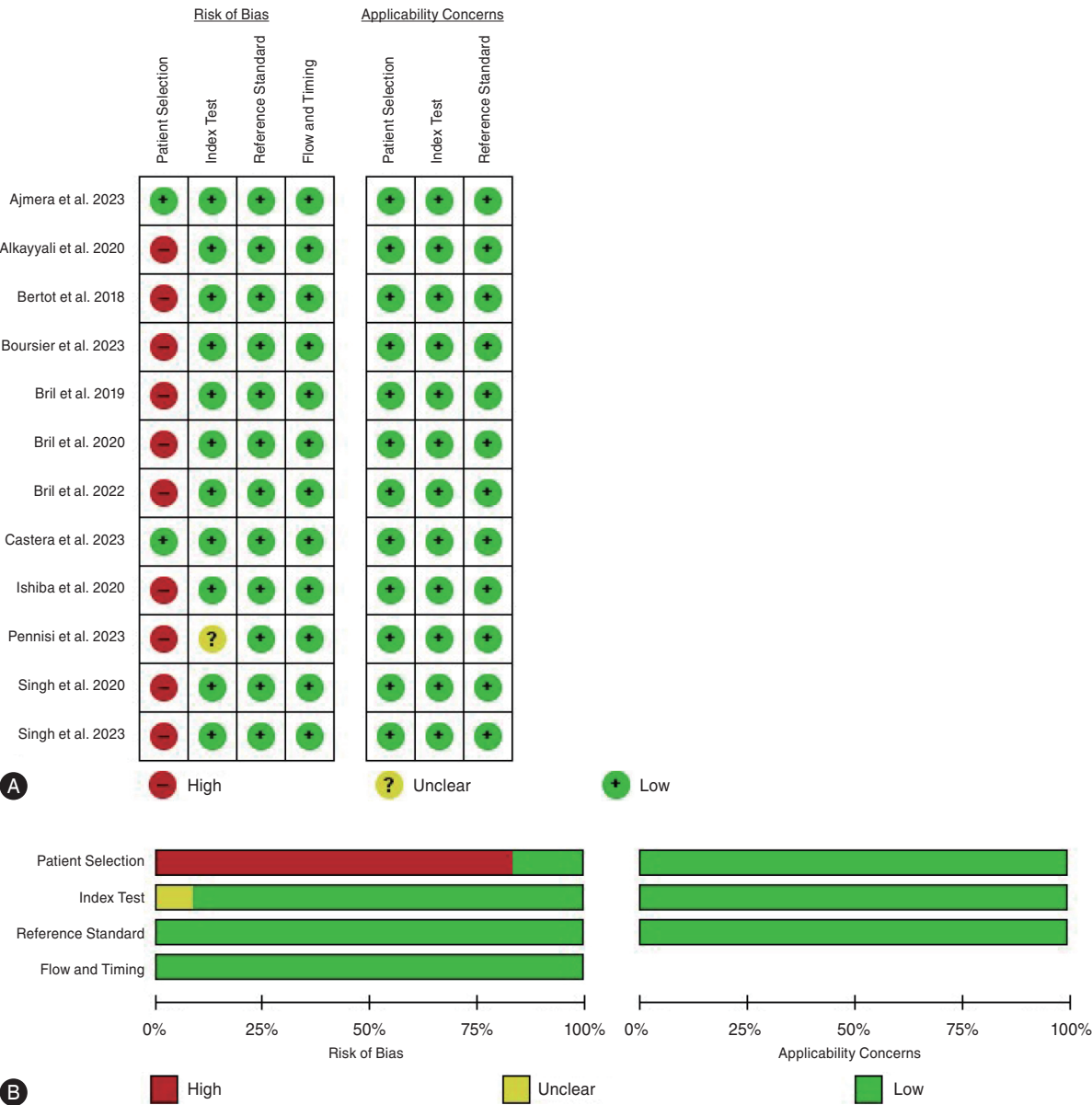
Keywords			
P	1	("Non-alcoholic Fatty Liver Disease"[MH]) OR ("Non-alcoholic Fatty Liver Disease" [ALL] OR "NAFLD"[ALL] OR "Steatohepatitides, Nonalcoholic" [ALL] OR ("fatty liver" [ALL] AND ("non-alcoholic" [ALL] or "nonalcoholic" [ALL])))	
	2	("Diabetes Mellitus" [MH]) OR ("Diabetes Mellitus" [ALL] or "Diabetes" [ALL])	
I	3	("fibrosis-4"[ALL] OR "FIB-4"[ALL])	
P&I	4	1 AND 2 AND 3	

I, index test; P, population.

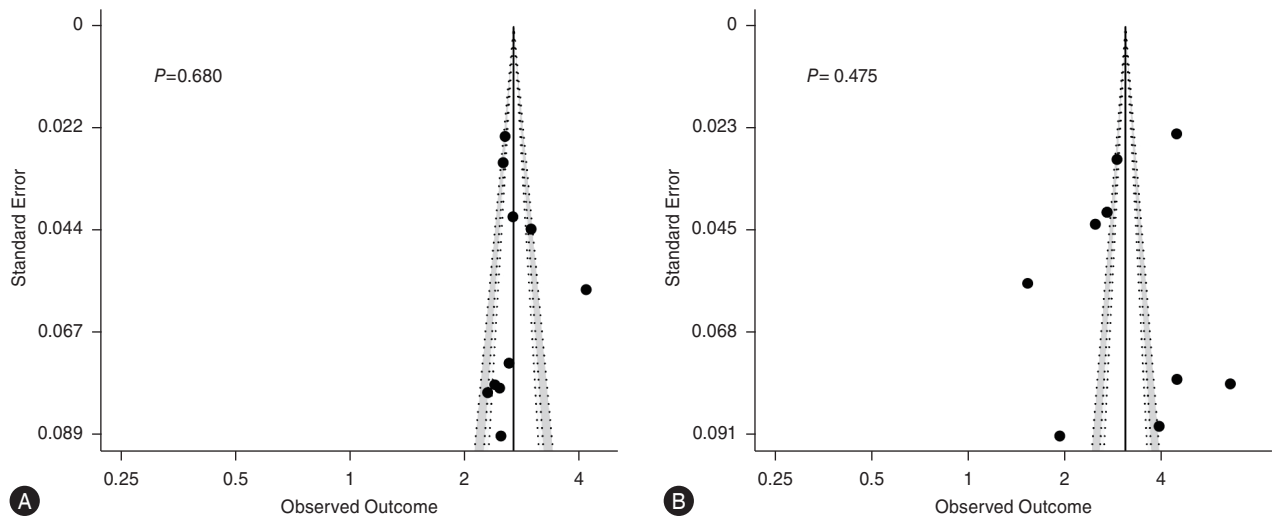
Supplementary Table 3. Meta-regression for diagnostic odds ratio of each cutoff

	Low cutoff excluding AF			High cutoff including AF		
	Coefficient	95% CI	P-value	Coefficient	95% CI	P-value
Asian (%)	0.30	0.79–2.30	0.169	−0.02	0.96–0.99	0.029
HbA1c (%)	0.35	0.23–8.66	0.619	−0.29	0.00–8.58	0.925
Age, years	−0.001	0.94–1.06	0.974	−0.07	0.87–0.99	0.048
BMI, kg/m ²	0.001	0.90–1.11	0.99	0.24	0.95–1.70	0.088
ALT, U/L	0.003	0.99–1.02	0.67	−0.03	0.88–1.06	0.393
F3 prevalence (by Bx.)	−0.01	0.97–1.01	0.194	−0.01	0.97–1.01	0.194

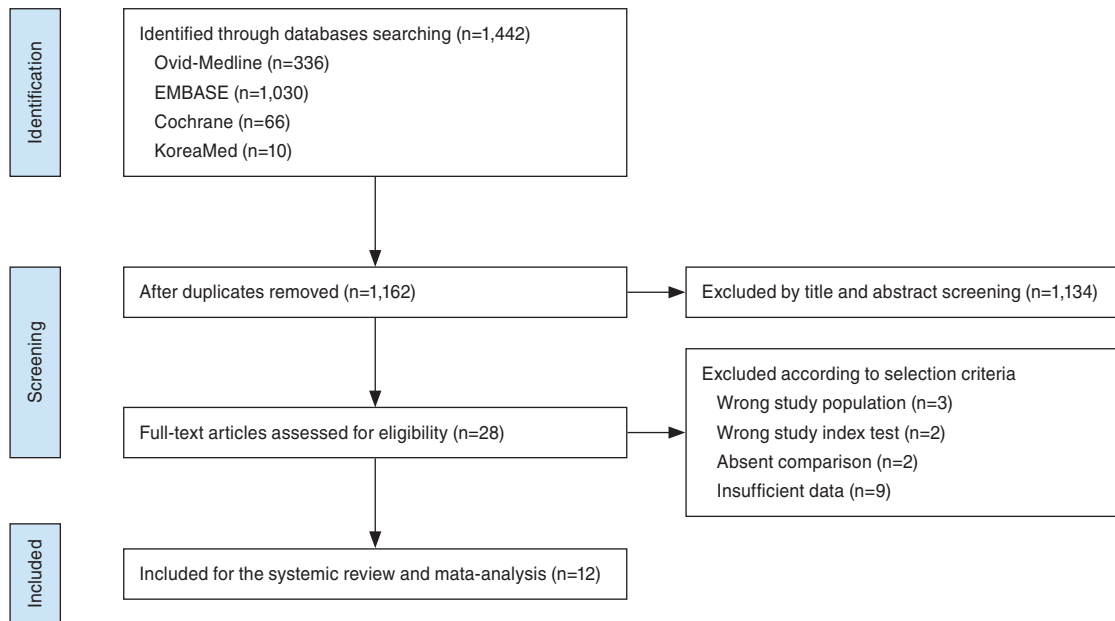
AF, advanced liver fibrosis; BMI, body mass index; ALT, alanine aminotransferase; Bx., biopsy.



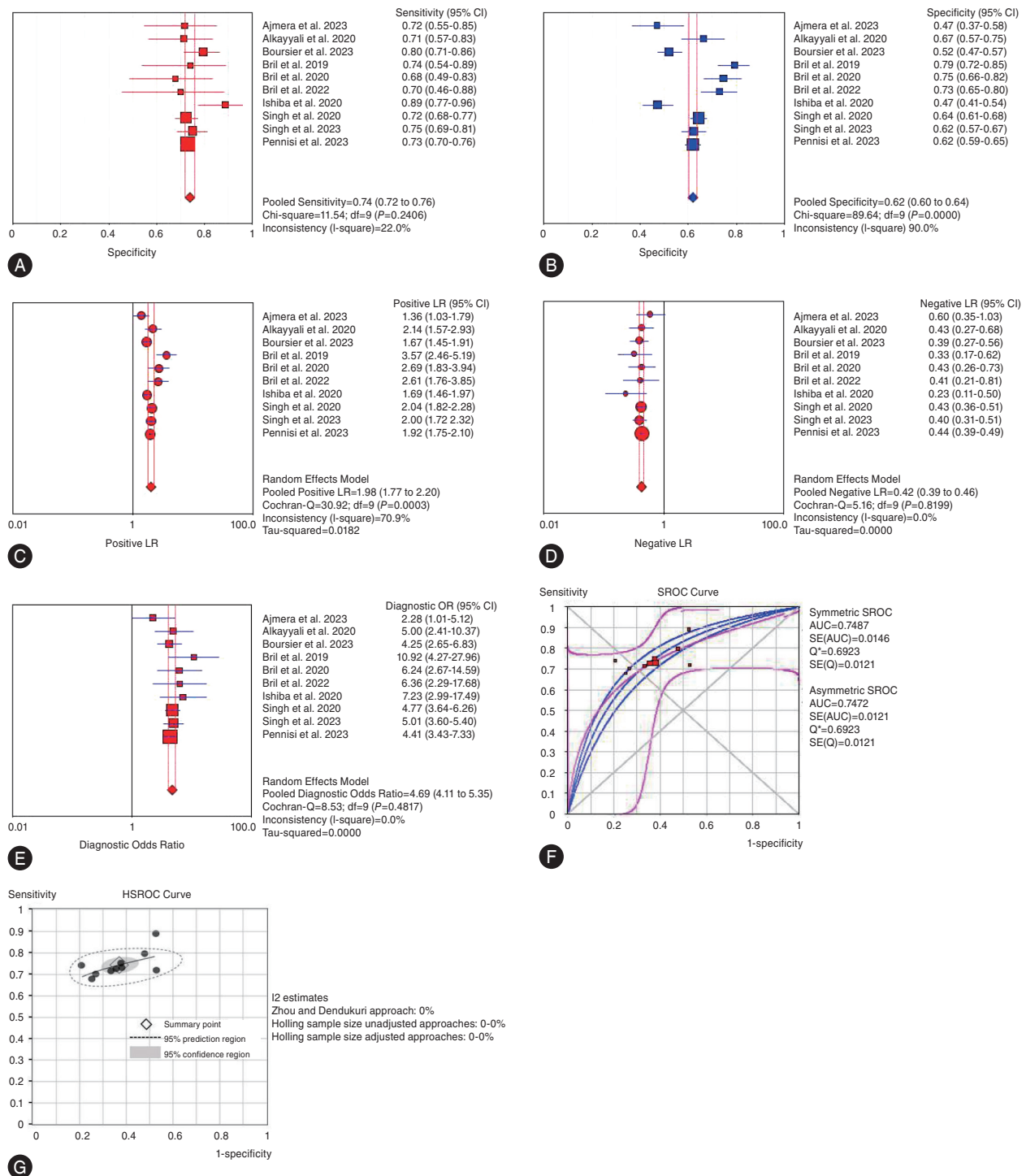
Supplementary Figure 1. Quality assessment. (A) Individual results enrolled studies for risk of bias and applicability concerns using the revised QUADAS-2. (B) Summary of the quality assessment.



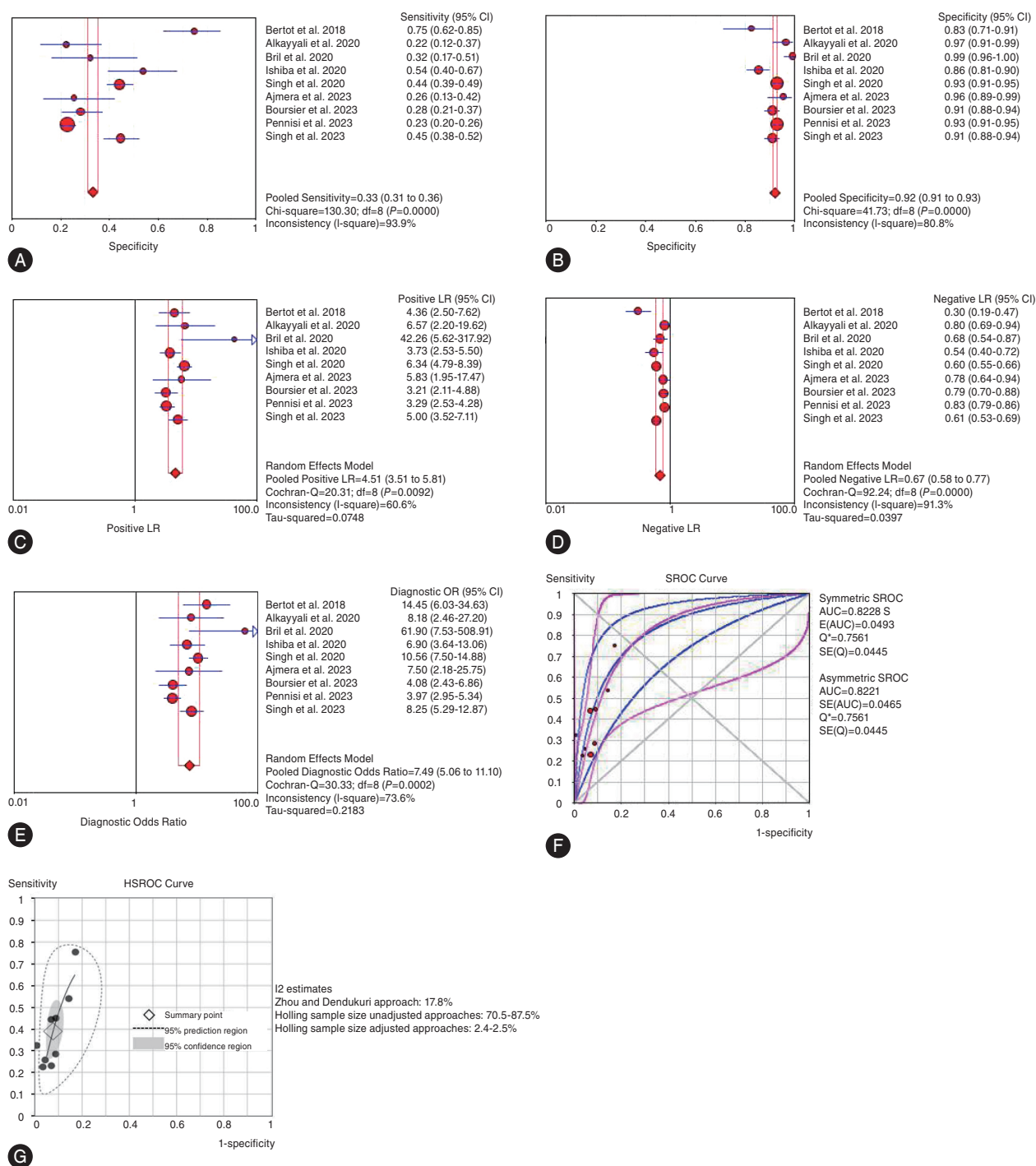
Supplementary Figure 2. Funnel plots representing publication bias. (A) A funnel plot for sensitivity of low cut-off points (FIB-4 1.3–1.67, n=10). (B) A funnel plot for specificity of high cut-off points (FIB-4 2.67–3.25, n=9). FIB-4, fibrosis-4.



Supplementary Figure 3. Flowchart depicting the criteria for including and excluding studies in the systematic review process.



Supplementary Figure 4. Results of meta-analysis for diagnostic accuracy of FIB-4 for advanced liver fibrosis in low-risk group. (A) Sensitivity. (B) Specificity. (C) Positive likelihood ratio (LR). (D) Negative LR. (E) Diagnostic odds ratio (OR). (F) Summary receiver operating curve (SROC). (G) Hierarchical summary receiver operating characteristic (HSROC). FIB-4, fibrosis-4; CI, confidence interval; AUC, area under the curve.



Supplementary Figure 5. Results of meta-analysis for diagnostic accuracy of FIB-4 for advanced liver fibrosis in high-risk group. (A) Sensitivity. (B) Specificity. (C) Positive likelihood ratio (LR). (D) Negative LR. (E) Diagnostic odds ratio (OR). (F) Summary receiver operating curve (SROC). (G) Hierarchical summary receiver operating characteristic (HSROC). FIB-4, fibrosis-4; CI, confidence interval; AUC, area under the curve.