Received: 23 August 2024 | Accepted: 12 November 2024

DOI: 10.1097/HEP.0000000000001190

ORIGINAL ARTICLE

OPEN

Diagnostic accuracy of 2D-SWE ultrasound for liver fibrosis assessment in MASLD: A multilevel random effects model meta-analysis

Madalina-Gabriela Indre[1,2,3,4,5] | **Dan-Corneliu Leucuta**[4] | **Monica Lupsor-Platon**[4,5] | **Laura Turco**[1] | **Silvia Ferri**[2] | **Ahmed Hashim**[2,6] | **Olga Hilda Orasan**[4] | **Bogdan Procopet**[4,5] | **Horia Stefanescu**[5] | **Maria Cristina Morelli**[1] | **Fabio Piscaglia**[2,3] | **Federico Ravaioli**[2,3]

[1]Internal Medicine Unit for the Treatment of Severe Organ Failure, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

[2]Division of Internal Medicine, Hepatobiliary and Immunoallergic Diseases, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy

[3]Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

[4]Faculty of Medicine, "Iuliu Hatieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania

[5]Hepatology Department, Regional Institute of Gastroenterology and Hepatology "Octavian Fodor", Cluj-Napoca, Romania

[6]Cambridge Liver Unit, Addenbrooke's Hospital, Cambridge University Hospitals, Cambridge, United Kingdom

Correspondence
Federico Ravaioli, Department of Medical and Surgical Sciences, University of Bologna, Bologna, 40138, Italy.
Email: f.ravaioli@unibo.it

Abstract

Background and Aims: Metabolic dysfunction–associated steatotic liver disease (MASLD) imposes significant health care burdens. Early detection of advanced fibrosis and cirrhosis in MASLD is essential due to their unfavorable outcomes. This multilevel random-effects meta-analysis aimed to provide the best evidence for the diagnostic accuracy of 2-dimensional shear wave elastography in detecting liver fibrosis in biopsy-proven MASLD.

Approach and Results: This study involves systematic search in PubMed/MEDLINE, Embase, Scopus, Web of Science, LILACS, and Cochrane Library electronic databases for full-text articles published in any language up to February 26, 2024. Included studies reported liver stiffness measurement by 2-dimensional shear wave elastography and used histological diagnosis as the gold standard. A linear mixed-effects multiple thresholds model was employed, and summary estimates for sensitivity, specificity (Sp), and summary area under the receiver operator characteristic curve were computed. Twenty observational studies (SuperSonic Imagine, General Electric Healthcare, and Canon Medical Systems) fulfilled the inclusion criteria, comprising

Abbreviations: 2D-SWE, 2-dimensional shear wave elastography; AASLD, American Association for the Study of Liver Diseases; cACLD, compensated advanced chronic liver disease; EFSUMB, European Federation of Societies, for Ultrasound in Medicine and Biology; F0, no fibrosis; F1, mild fibrosis; F2, significant fibrosis; F3, advanced fibrosis; F4, cirrhosis; LB, liver biopsy; LS, liver stiffness; LSM, liver stiffness measurement; MASLD, metabolic dysfunction–associated steatotic liver disease; NPV, negative predictive value; PPV, positive predictive value; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; QIBA, Quantitative Imaging Biomarkers Alliance; sAUC, summary area under the receiver operator characteristic curve; Se, sensitivity; Sp, specificity; SRU, Society of Radiologists in Ultrasound; SSI, SuperSonic Imagine; T2DM, type 2 diabetes mellitus; US, ultrasound; VCTE, vibration-controlled transient elastography; WFUMB, World Federation for Ultrasound in Medicine and Biology.

Madalina-Gabriela Indre and Dan-Corneliu Leucuta contributed equally and shared the co-first authorship.

Fabio Piscaglia and Federico Ravaioli contributed equally and shared the co-last authorship

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.hepjournal.com.

2223 participants with biopsy-proven MASLD. The prevalence of mild fibrosis (F1), significant fibrosis (F2), advanced fibrosis (F3), and cirrhosis (F4) was 30.0%, 18.5%, 17.9%, and 10.9%, respectively. The summary area under the receiver operator characteristic curve [95% CI] in detecting ≥F1, ≥F2, ≥F3, and F4 for all ultrasound machines considered together were 0.82 [0.16–0.98], 0.82 [0.76–0.88], 0.86 [0.77–0.93], and 0.89 [0.80–0.95], respectively. The optimal cutoff values were 6.432 kPa for ≥F1, 8.174 kPa for ≥F2, 9.418 kPa for ≥F3, and 11.548 kPa for F4, respectively.

Conclusions: Our meta-analysis identified optimized cutoffs for fibrosis staging by 2-dimensional shear wave elastography in etiology-specific chronic liver diseases (MASLD), with excellent diagnostic performance, underscoring the potential for standardizing cutoff values.

Keywords: 2-dimensional shear-wave elastography, cirrhosis, liver imaging, metabolic dysfunction–associated steatotic liver disease, noninvasive tests

INTRODUCTION

NAFLD, newly redefined as metabolic dysfunction–associated steatotic liver disease (MASLD),[1] represents a real challenge for the health care system.[2–4] Liver fibrosis is associated with long-term outcomes in patients with MASLD,[4] and fibrosis progression to advanced stages (F3-F4) leads to an increased risk of liver-related mortality, all-cause mortality, and higher incidence of major cardiovascular events.[5–7] This prompts the need for adequate identification of liver fibrosis stages, especially the advanced ones.[8–11]

Liver biopsy (LB) is the gold standard for assessing liver fibrosis, but its invasive nature, risk of sampling errors, and interobserver and intraobserver variability limit its use.[12,13] To address these limitations, various noninvasive tests, including vibration-controlled transient elastography (VCTE; FibroScan, Echosens), have been developed and validated for measuring liver stiffness.[14,15] However, challenges may arise in settings with limited access or among patients with higher body mass index, potentially affecting performance and accessibility.[9] These considerations highlight the need for ongoing exploration of complementary liver stiffness measurement (LSM) approaches.[16]

Two-dimensional shear-wave elastography (2D-SWE) is a promising alternative technique for LSM.[17] Most new ultrasound (US) machines now come equipped with elastography modules, making 2D-SWE potentially valuable for assessing patients with MASLD in primary, secondary, and tertiary care settings, especially when supported by well-validated cutoff values. One advantage of 2D-SWE is its ability to integrate LSM with US-based structural liver evaluation in a multiparametric approach. Some devices also include features for assessing liver steatosis and inflammation, though further validation is needed to confirm their accuracy in routine clinical practice.[18–20] However, variations in hardware and software among different manufacturers of 2D-SWE elastography can lead to inconsistencies in cutoff values.[21,22] If thoroughly validated, 2D-SWE could be integrated into referral pathways for identifying or predicting liver fibrosis,[23] similar to current practices with VCTE.[24] While most patients at risk for MASLD undergo US scans, only a fraction receive LSM with VCTE. The rapid advancement of US technology incorporating 2D-SWE suggests that future assessments for LSM in MASLD patients could favor this method over VCTE. Establishing reliable thresholds for fibrosis assessment to maximize the global application of 2D-SWE is urgently globally warranted.

Therefore, we aimed to perform a systematic review and multilevel random effects model meta-analysis to provide the best evidence for the diagnostic accuracy of 2D-SWE in patients with biopsy-proven MASLD and to propose reliable cutoffs for different stages of liver fibrosis that could present adequate applicability in the clinical practice.

METHODS

Registration of review protocol

A systematic review and a multilevel random effects model meta-analysis were conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, http://links.lww.com/HEP/J662) guidelines.[25–27] The protocol was registered on OFS (https://doi.org/10.17605/OSF.IO/9WV8F).

Search strategy and selection criteria

We systematically searched PubMed/MEDLINE, Embase, Scopus, Web of Science, LILACS, and Cochrane Library electronic databases for full-text articles published in any language from inception to February 1, 2023, and subsequently updated the search to February 26, 2024 (Supplemental Materials, http://links.lww.com/HEP/J663, Supplemental Table S1, http://links.lww.com/HEP/J663).

We included studies meeting the following prespecified criteria: (i) population—biopsy-proven NAFLD/NASH (MASLD/MASH); (ii) target condition—liver fibrosis assessment through LB (reference standard); and (iii) index test—LSM by 2D-SWE. We considered any US machine equipped with 2D-SWE.

Data extraction and quality assessment

Data from the studies considered eligible were independently extracted by 3 authors (Madalina-Gabriela Indre, Dan-Corneliu Leucuta, and Federico Ravaioli).

In the case of multiple publications, we included the one that reported the most comprehensive information. When needed, we contacted the corresponding authors to obtain additional information (Supplemental Table S2, http://links.lww.com/HEP/J663).

We evaluated the presence of potential bias in the patient selection process, blinding to the histological diagnosis or the index test, description of the reference standard, and inclusion of all patients in the analysis (Supplemental Table S3, http://links.lww.com/HEP/J663, Supplemental Figures S1, S2, http://links.lww.com/HEP/J663). Two authors (Madalina-Gabriela Indre and Federico Ravaioli) independently assessed the methodological quality using the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool.[28] Any disagreements were resolved through discussion with a third author (Fabio Piscaglia).

Data synthesis and analysis

The primary outcome of the meta-analysis was the diagnostic accuracy of 2D-SWE in identifying different stages of liver fibrosis in patients with biopsy-proven MASLD, compared to LB, which was the gold standard. Subsequently, we performed subgroup analysis for the type of US manufacturer, population characteristics, and the presence of comorbidities, such as type 2 diabetes mellitus (T2DM) and obesity.

More detailed information about the selection criteria, data extraction, and quality assessment were reported in Supplemental Materials, http://links.lww.com/HEP/J663.

Statistical analysis

A meta-analysis was carried out for index tests and target conditions with over 3 studies with sufficient information for generating classification tables. Index tests without enough studies to perform a meta-analysis were analyzed using narrative synthesis. When the authors provided a single cutoff for a specific liver fibrosis stage or referred to it as "optimal," "best," or at the "Youden index," the threshold was defined as "best cutoff." A cutoff specified for sensitivity (Se) $\geq$90% was considered as a "rule-out," while a cutoff indicated for Sp $\geq$90% was considered as the "rule-in." We used a bivariate logit-normal random effects model to estimate the mean Se, Sp, variances, and covariance. Summary area under receiver operator characteristic curves (sAUCs) were generated with 95% confidence and 95% prediction region. The 95% confidence region is based on the CI around the summary point and indicates that, based on the available data, we expect the "real value" to be within that region 95% of the time. A linear mixed-effects model was used to model the multiple threshold data of individual studies.[14,29,30] The multiple thresholds model is a multilevel random effects model that enables the calculation of summary sensitivities and specificities of different cutoffs and the calculation of the positive predictive value (PPV) and negative predictive value (NPV), given the prevalence of the target condition. Se and Sp were combined at every recommended cutoff to produce a multiple threshold sAUC. PPV and NPV were obtained for predefined prevalences (1%, 2.5%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, and 70%), and cutoffs required to achieve minimum acceptable criteria were determined. We explored the heterogeneity across the eligible studies by performing Se analyses by study setting (Asian vs. non-Asian cohort), presence of obesity (body mass index $\geq$30 kg/m^2 or $\geq$25 kg/m^2 in Asian cohort),[31–33] presence of T2DM, ALT levels, and type of US manufacturer. We chose not to create funnel plots because it is widely recognized that tests based on funnel plot asymmetry are unreliable in distinguishing between publication bias and other reasons for asymmetry, such as including multiple thresholds, in systematic reviews of diagnostic test accuracy studies.[34]

The analysis was performed using the statistical software R with the *mada* and *diagmeta* packages (Version 4.3.2; R Foundation for Statistical Computing).

RESULTS

The electronic search identified 8390 records. After eliminating 2353 duplicates, the titles and abstracts of 6037 publications were evaluated. Of these, 116 full-text articles were assessed for eligibility. After excluding

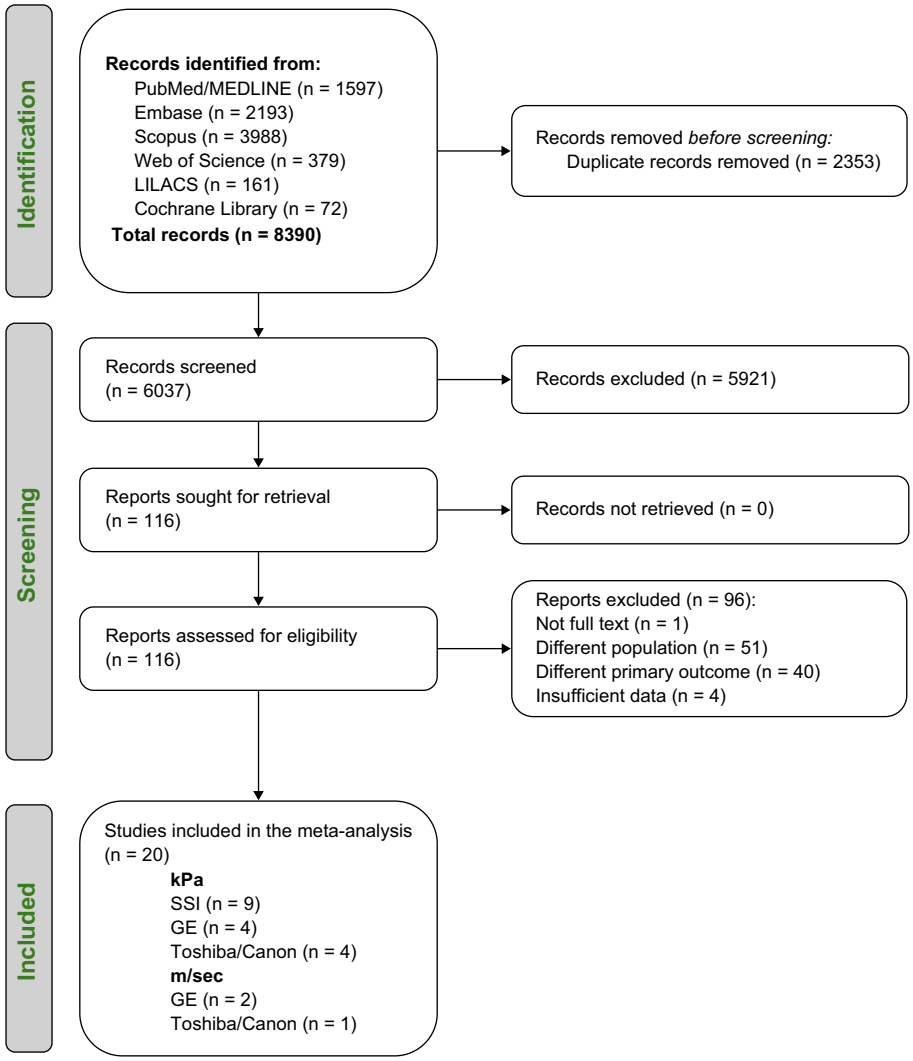

FIGURE 1 PRISMA flow diagram. Abbreviations: GE, General Electric Healthcare; n, number; SSI, SuperSonic Imagine.

96 publications for various reasons (Supplemental Table S2, http://links.lww.com/HEP/J663), 20 studies met the inclusion criteria for the meta-analysis. The PRISMA flow diagram (Figure 1) visually represents the selection process.

Study characteristics

The main characteristics of the included studies are presented in Table 1. Additional details concerning the histology interpreter, US examiner, unreliable 2D-SWE measurements, and the percentage of liver steatosis on histology are illustrated in Supplemental Tables S4–S6, http://links.lww.com/HEP/J663.

The 20 observational studies included 2223 participants with biopsy-proven MASLD. The prevalence of mild fibrosis (F1), significant fibrosis (F2), advanced fibrosis (F3), and cirrhosis (F4) was 30.0% (n = 666), 18.5% (n = 411),17.8% (n = 397), and 10.9% (n = 242), respectively, and 22.8% (n = 507) of patients were classified as having no fibrosis (F0).

The mean age (±SD) ranged from 31.5 (±10.1) to 61 (±14.9) years, and the percentage of females ranged from 35.2% to 80%. The mean body mass index ranged from 24.9 (±3.75) kg/m^2 to 45.46 (±6.26) kg/m^2, with a prevalence of obesity for both non-Asian and Asian populations. Fourteen (70%) studies were conducted in Asia,[36–38,41,43,45–51,53,54] 4 (20%) in America,[39,40,44,52] and 2 (10%) in Europe.[35,42] Two (10%) studies were multicentric,[35,49] and 18 (90%) were single center. Most of the studies (90%, n = 18) included cohorts of patients with biopsy-proven MASLD from tertiary health care centers, whereas 2 (10%) studies[38,41] included patients from bariatric surgery clinics with available LB data. Fifty percent of the studies were retrospective and 50% presented a cross-sectional design. Supplemental Table S7, http://links.lww.com/HEP/J663, reports the summary diagnostic performance of 2D-SWE for detecting fibrosis stages in

TABLE 1 Main characteristics of the included studies (n = 20)

Author (Ref)	Country	Population (n)	M (n)	Study design, setting	Condition	Age (y), mean ± SD	Sex, F, n (%)	BMI (kg/m²), mean ± SD	ALT (IU/l), mean ± SD	T2DM, n (%)
Measurement in kPa										
Supersonic Imagine (SSI)										
Cassinotto et al[35]	France	291	5	Prospective multicentric cross-sectional	b-p MASLD	56.7 ± 12	119 (40.9)	32.1 ± 6	71.2 ± 50.7	153 (52.6)
Lee et al[36]	Korea	94	10	Prospective single-center	b-p MASLD	55.5 ± 3.87	53 (56.4)	27.13 ± 1.12	50.1 ± 11.6	37 (39.4)
Takeuchi et al[37]	Japan	71	5	Prospective single-center	b-p MASLD	50.8 ± 15.7	25 (35.2)	29.2 ± 5.1	87.9 ± 90.3	—
Jamialahmadi et al[38]	Iran	90	10	Observational single-center bariatric surgery	b-p MASLD	38.5 ± 11.1	72 (80)	45.46 ± 6.26	—	25 (27.8)
Ozturk et al[39]	USA	116	10	Retrospective single-center	b-p MASLD	50.6 ± 11.8	62 (53.4)	31.4 ± 5.1	67.1 ± 43.2	38 (32.8)
Sharpton et al[40]	USA	114	3 to 5	Prospective cross-sectional single-center	b-p MASLD	54.7 ± 14.27	62 (54.4)	31.5 ± 3.98	48.3 ± 26.3	—
Chen et al[41]	China	100	—	Retrospective single-center bariatric surgery	b-p MASLD	31.55 ± 10.12	69 (69)	38 ± 14.56	49.0 ± 42.5	37 (37)
Mendoza et al[42]	Switzerland	88	3	Prospective single-center	b-p MASLD	53.2 ± 12.7	33 (57.5)	31.6 ± 7.1	84.4 ± 58.4	39 (44.3)
Zhou et al[43]	China	116	5	Retrospective single-center	b-p MASLD	46.4 ± 16.45	61 (52.6)	—	127.5 ± 183.1	23 (19.8)
General Electric Healthcare (GE)										
Furlan et al[44]	USA	62	3–12	Prospective cross-sectional single-center	b-p MASLD	50 ± 13	36 (58)	34.8 ± 7.2	56 ± 33	22 (35)
Kuroda et al[45]	Japan	202	10	Prospective cross-sectional single-center	b-p MASLD	55.2 ± 12.9	103 (51)	29.27 ± 5.67	21.9 ± 6.5	82 (40.6)
Imajo et al[46]	Japan	201	10	Prospective single-center	b-p MASLD	61 ± 14.93	95 (47.3)	27.17 ± 4.18	52.2 ± 31.7	127 (61.9)
Kalaiyarasi et al[47]	Singapore	16	5	Pilot, cross-sectional single-center	b-p MASLD	49.2 ± 11.9	11 (68.8)	29.6 ± 3.7	90.5 ± 50.6	11 (68.8)
Canon Medical Systems/Toshiba										
Lee et al[48]	Korea	102	9	Prospective single-center	b-p MASLD	47 ± 24.82	59 (57.8)	25.6 ± 5.49	74.0 ± 70.7	17 (16.7)
Jang et al[49]	Korea	132	5	Prospective cross-sectional multicentric	b-p MASLD	39.67 ± 20.24	69 (52.3)	24.9 ± 3.75	38.5 ± 47.6	30 (22.7)
Laroia et al[50]	India	50	10	Prospective cross-sectional single-center	b-p MASH	46.76 ± 10.7	19 (38)	—	—	—
Kim et al[51]	Korea	60	10	Retrospective single-center	b-p MASLD b-p MASH	50.9 ± 13.4	32 (53.3)	29.9 ± 4.3	101.1 ± 69.6	37 (61.7)

Measurement in m/s										
General Electric Healthcare (GE)										
Zhang et al[52]	USA	100	10	Prospective cross-sectional single-center	b-p MASLD	51.8 ± 12.9	54 (54)	31.6 ± 4.7	64.7 ± 36.9	—
Ogino et al[53]	Japan	107	6	Retrospective single-center	b-p MASLD	51 ± 14	42 (39.3)	29.0 ± 4.3	71.1 ± 45.7	—
Canon Medical Systems/Toshiba										
Sugimoto et al[54]	Japan	111	10	Prospective cross-sectional single-center	b-p MASLD	53 ± 18	54 (48.6)	27.7 ± 4.3	79.7 ± 52.8	—

Abbreviations: —, not available; BMI, body mass index; b-p, biopsy-proven; F, women; IU/l, international units per liter; M, measurements; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease; T2DM, type 2 diabetes mellitus.

biopsy-proven MASLD by applying a bivariate logit-normal random effects model.

Study quality assessment

The methodological quality was assessed with the QUADAS-2 tool,[28] as summarized in Supplemental Table S3, http://links.lww.com/HEP/J663, Supplemental Figures S1, S2, http://links.lww.com/HEP/J663. Nineteen (95%) studies had no risk of bias or applicability concerns. The studies that did not explicitly state whether consecutive or random sampling was made were judged as having a "high risk" of bias in the patient selection domain of the QUADAS-2 tool. This included 3 (15%) studies out of 20. Regarding the index test domain, the QUADAS-2 tool was judged as "high risk" for all the included studies, as a prespecified threshold for 2D-SWE is not currently being used.[19] The flow and timing domain of the QUADAS-2 tool was judged as having an "unclear risk" of bias in 19 (95%) studies, as the "intention to diagnose" approach was specified in one study.

Diagnostic performance of 2D-SWE by kPa for different stages of liver fibrosis by multiple threshold data meta-analysis

Mild fibrosis (≥F1; F0 vs. F1-F4)

The diagnostic accuracy in detecting mild fibrosis (≥F1) was investigated in 9 (45%) studies that reported the results in kPa. The sAUC [95% CI], Se [95% CI], and Sp [95% CI], when all US machines were considered together, were 0.82 [0.16–0.98], 0.76 [0.66–0.83], and 0.76 [0.60–0.86], respectively, with an optimal cutoff for depicting ≥F1 of 6.432 kPa (Figure 2A, Table 2).

Significant fibrosis (≥F2; F0-F1 vs. F2-F4)

The diagnostic accuracy in detecting significant fibrosis (≥F2) was investigated in 16 (80%) studies that reported the results in kPa. The sAUC [95% CI], Se [95% CI], and Sp [95% CI], when all US machines were considered together, were 0.82 [0.76–0.88], 0.76 [0.65–0.84], and 0.76 [0.65–0.84], respectively, with an optimal cutoff for depicting ≥F2 of 8.174 kPa (Figure 2B, Table 2).

Advanced fibrosis (≥F3; F0-F2 vs. F3-F4)

The diagnostic accuracy in detecting advanced fibrosis (≥F3) was investigated in 17 (85%) studies that

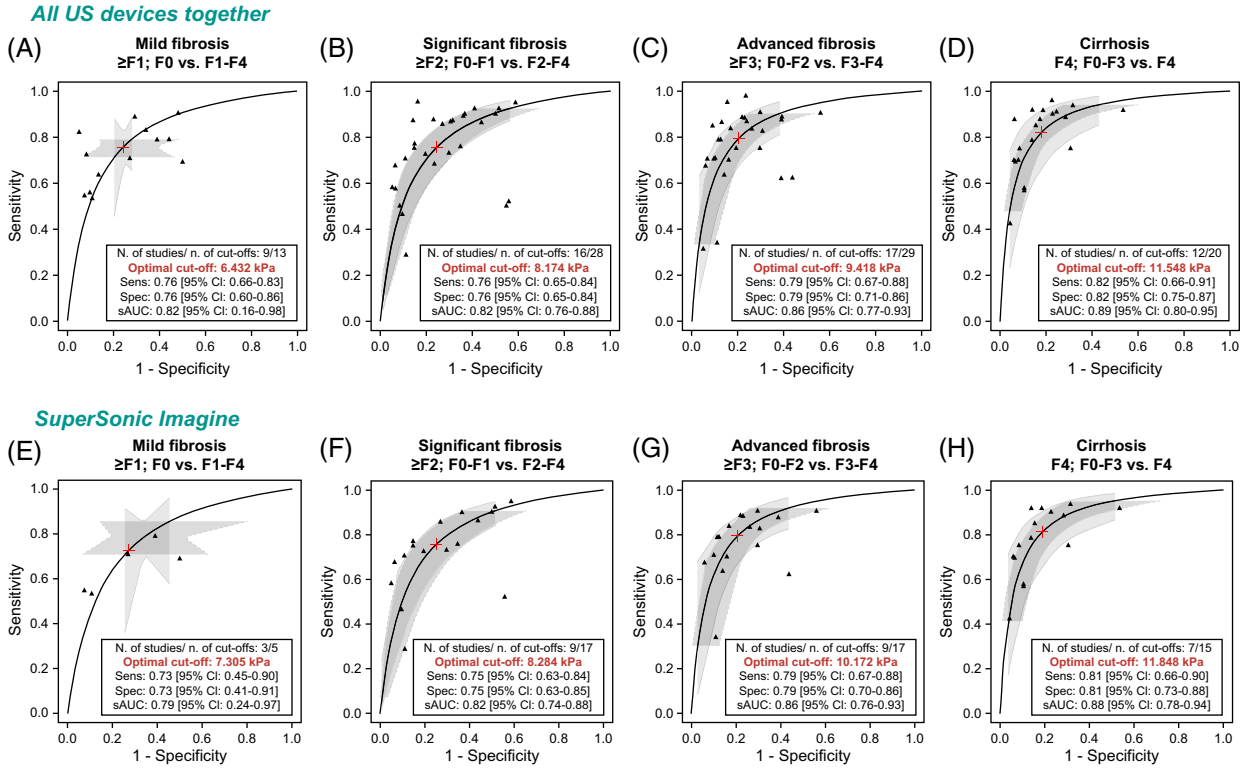

FIGURE 2 sAUC by multilevel random effects (CICS) model for multiple thresholds data for all US devices together and SuperSonic Imagine (kPa). (A) Mild fibrosis by all US devices; (B) significant fibrosis by all US devices; (C) advanced fibrosis by all US devices; (D) cirrhosis by all US devices; (E) Mild fibrosis by Supersonic Image; (F) significant fibrosis by Supersonic Image; (G) advanced fibrosis by Supersonic Image; (H) cirrhosis by all Supersonic Image. Created in BioRender. Taru, M. (2024) https://BioRender.com/o82f827. Abbreviations: CICS, common random intercept and common slope; F1, mild fibrosis; F2, significant fibrosis; F3, advanced fibrosis; F4, fibrosis; N, number; sAUC, summary area under the receiver operator characteristic curve; Sens, sensitivity; Spec, specificity; US, ultrasound.

reported the results in kPa. The sAUC [95% CI], Se [95% CI], and Sp [95% CI], when all US machines were considered together, were 0.86 [0.77–0.93], 0.79 [0.67–0.88], and 0.79 [0.71–0.86], respectively, with an optimal cutoff for depicting ≥F3 of 9.418 kPa (Figure 2C, Table 2).

Cirrhosis (F4; F0-F3 vs. F4)

The diagnostic accuracy in detecting cirrhosis (F4) was investigated in 12 (60%) studies that reported the results in kPa. The sAUC [95% CI], Se [95% CI], and Sp [95% CI], when all US machines were considered together, were 0.89 [0.80–0.95], 0.82 [0.66–0.91], and 0.82 [0.75–0.87], respectively, with an optimal cutoff for depicting F4 of 11.548 kPa (Figure 2D, Table 2).

Diagnostic performance of 2D-SWE by m/s for different stages of liver fibrosis by multiple threshold data meta-analyses

The diagnostic accuracy in detecting ≥F2, ≥F3, and F4 was investigated in 3 (15%), 3 (15%), and 3 (15%) studies that reported the results in m/s, respectively. The sAUC [95% CI], Se [95% CI], and Sp [95% CI] were 0.82 [0.65–0.94], 0.76 [0.42–0.93], and 0.76 [0.43–0.93] for ≥F2; 0.84 [0.51–0.95], 0.77 [0.63–0.87], and 0.77 [0.70–0.83] for ≥F3; 0.89 [0.38–0.99], 0.82 [0.56–0.94], and 0.82 [0.75–0.88] for F4, respectively. The optimal cutoffs in detecting ≥F2, ≥F3, and F4 were 1.481, 1.633, and 1.748 m/s, respectively (Supplemental Figure S3, http://links.lww.com/HEP/J663, Table 2).

Se analysis of the diagnostic performance of 2D-SWE (kPa) for different stages of liver fibrosis

In Figure 3, Supplemental Figures 3–5, http://links.lww.com/HEP/J663, Supplemental Tables S4, S5, and S8, http://links.lww.com/HEP/J663, we explored heterogeneity across studies by performing Se analyses. Moreover, we performed an analysis excluding the bariatric surgery population that yielded similar results (Supplemental Table S8, http://links.lww.com/HEP/J663). A summary of the optimal cutoff values is shown in Figure 4.

TABLE 2 Summary diagnostic performance of 2D-SWE for the detection of different stages of liver fibrosis in biopsy-proven MASLD by bivariate logit-normal random effects model

Fibrosis stage	Studies, n (%)	Patients (n)	Prevalence (%)	Cutoff range	sAUC [95% CI]	Se [95% CI]	Sp [95% CI]
Measurement in kPa							
All US devices together							
Mild fibrosis, ≥F1; F0 vs. F1-F4	9 (45)	1489	78.2	5.85–8.9	0.84 [0.81–0.87]	0.76 [0.69–0.82]	0.80 [0.69–0.87]
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	16 (80)	936	49.1	5.7–11.9	0.84 [0.80–0.87]	0.78 [0.70–0.84]	0.76 [0.69–0.82]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	17 (85)	556	29.2	6.75–13.07	0.87 [0.84–0.90]	0.81 [0.73–0.86]	0.80 [0.74–0.84]
Cirrhosis, F4; F0-F3 vs. F4	12 (60)	208	10.9	6.75–15.73	0.93 [0.91–0.95]	0.90 [0.84–0.94]	0.82 [0.77–0.87]
SuperSonic Imagine							
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	9 (45)	533	49.4	6.6–11.57	0.81 [0.77–0.84]	0.72 [0.63–0.81]	0.77 [0.65–0.86]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	9 (45)	294	27.2	6.75–13.07	0.85 [0.82–0.88]	0.79 [0.71–0.86]	0.77 [0.69–0.83]
Cirrhosis, F4; F0-F3 vs. F4	7 (35)	105	9.7	6.75–15.73	0.93 [0.90–0.95]	0.91 [0.80–0.96]	0.81 [0.74–0.87]

Abbreviations: 2D-SWE, 2-dimensional shear-wave elastography; F1, mild fibrosis; F2, significant fibrosis; F3, advanced fibrosis; F4, cirrhosis; MASLD, metabolic dysfunction–associated steatotic liver disease; n, number; sAUC, summary area under the receiver operator characteristic curve; Se, sensitivity; Sp, specificity; US, ultrasound.

Asian population

The diagnostic accuracy in detecting ≥F1, ≥F2, ≥F3, and F4 was investigated in 8 (40%), 11 (55%), 12 (60%), and 9 (45%) Asian studies that reported the results in kPa, respectively. The sAUC [95% CI], Se [95% CI], and Sp [95% CI] were 0.83 [0.73–0.90], 0.76 [0.62–0.86], and 0.76 [0.57–0.88] for ≥F1; 0.80 [0.69–0.88], 0.74 [0.60–0.84], and 0.74 [0.63–0.83] for ≥F2; 0.84 [0.43–0.96], 0.78 [0.60–0.89], and 0.78 [0.68–0.85] for ≥F3; 0.90 [0.78–0.97], and 0.83 [0.62–0.94], and 0.83 [0.75–0.89] for F4, respectively. The optimal cutoffs in detecting ≥F1, ≥F2, ≥F3, and F4 were 6.953, 8.561, 10.182, and 12.466 kPa, respectively (Table 3, Figure 3).

Population with T2DM

The diagnostic accuracy in detecting ≥F2, ≥F3, and F4 for the population with T2DM was investigated in 9 (45%), 10 (50%), and 6 (30%) studies that reported the results in kPa, respectively. The sAUC [95% CI], Se [95% CI], and Sp [95% CI] were 0.84 [0.77–0.90], 0.77 [0.64–0.87], and 0.77 [0.63–0.87] for ≥F2; 0.88 [0.81–0.93], 0.81 [0.68–0.90], and 0.81 [0.69–0.89] for ≥F3; 0.88 [0.81–0.93], 0.81 [0.67–0.90], and 0.81 [0.73–0.87] for F4, respectively. The optimal cutoffs in detecting ≥F2, ≥F3, and F4 were 7.961, 9.010, and 11.760 kPa, respectively (Table 3, Figure 3).

Type of US manufacturer

SuperSonic Imagine

The diagnostic accuracy in detecting ≥F1, ≥F2, ≥F3, and F4 by SuperSonic Imagine (SSI) was investigated in 3 (15%), 9 (45%), 9 (45%), and 7 (35%) studies that reported the results in kPa, respectively. The sAUC [95% CI], Se [95% CI], and Sp [95% CI] were 0.79 [0.24–0.97], 0.73 [0.45–0.90], and 0.73 [0.41–0.91] for ≥F1; 0.82 [0.74–0.88], 0.75 [0.63–0.84], and 0.75 [0.63–0.85] for ≥F2; 0.86 [0.76–0.93], 0.79 [0.67–0.88], and 0.79 [0.70–0.86] for ≥F3; 0.88 [0.78–0.94], 0.81 [0.66–0.90], and 0.81 [0.73–0.88] for F4, respectively. The optimal cutoffs in detecting ≥F1, ≥F2, ≥F3, and F4 were 7.305, 8.284, 10.172, and 11.848 kPa, respectively (Table 3, Figure 2).

Canon Medical Systems/Toshiba

The diagnostic accuracy in detecting ≥F2 and ≥F3 by Canon Medical Systems/Toshiba was investigated in 4 (20%) studies. The sAUC [95% CI], Se [95% CI], and Sp [95% CI] were 0.82 [0.26–0.96], 0.75 [0.45–0.91], and 0.76 [0.61–0.86] for ≥F2; 0.86 [0.21–0.98], 0.80

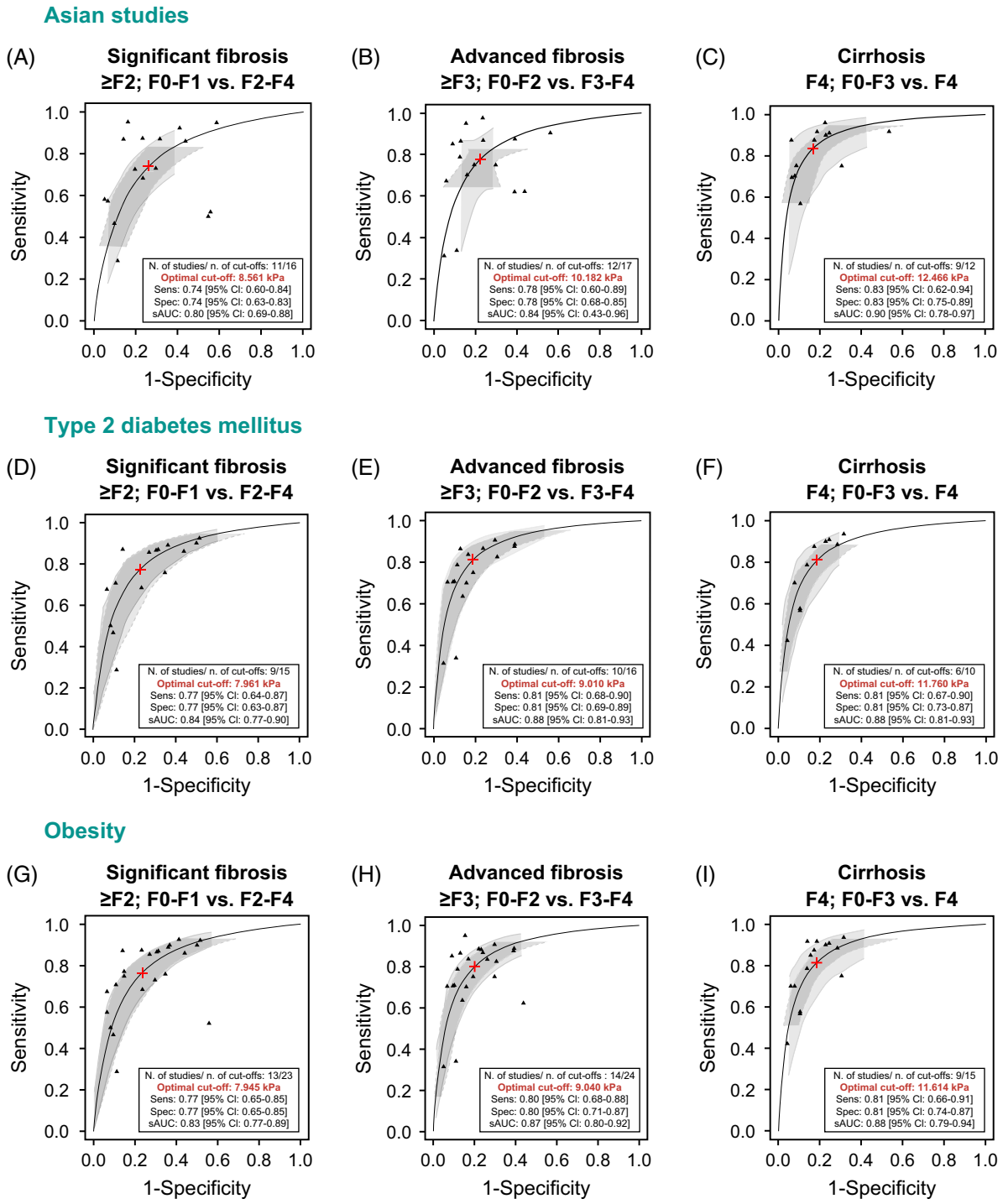

FIGURE 3 sAUC by multilevel random effects (CICS) model for multiple thresholds data for sensitivity analysis for different subpopulations (kPa). (A) significant fibrosis in Asian studies; (B) advanced fibrosis in Asian studies; (C) cirrhosis in Asian studies; (D) significant fibrosis in type 2 diabetes mellitus; (E) advanced fibrosis in type 2 diabetes mellitus; (F) cirrhosis In type 2 diabetes mellitus; (G) significant fibrosis in obesity; (H) advanced fibrosis in obesity; (I) cirrhosis in obesity. Created in BioRender. Taru, M. (2024) https://BioRender.com/a51n542. Abbreviations: CICS, common random intercept and common slope; F2, significant fibrosis; F3, advanced fibrosis; F4, cirrhosis; N, number; sAUC, summary area under the receiver operator characteristic curve; Sens, sensitivity; Spec, specificity; T2DM, type 2 diabetes mellitus.

[0.51–0.94], and 0.80 [0.69–0.87] for ≥F3, respectively. The optimal cutoffs in detecting ≥F2 and ≥F3 were 7.226 and 7.831 kPa, respectively (Table 3, Supplemental Figure S4, http://links.lww.com/HEP/J663).

General Electric Healthcare

The diagnostic accuracy in detecting ≥F2 and ≥F3 by General Electric Healthcare was investigated in 3 (15%) and 4 (20%) studies, respectively. The sAUC [95% CI],

Optimal cut-off values for identifying different stages of liver fibrosis in biopsy-proven MASLD				
Groups	**Mild fibrosis**	**Significant fibrosis**	**Advanced fibrosis**	**Cirrhosis**
	≥F1; F0 vs. F1-F4	≥F2; F0-F1 vs. F2-F4	≥F3; F0-F2 vs. F3-F4	F4; F0-F3 vs. F4
All ultrasound devices together (kPa)	**6.4** kPa sAUC: 0.82 [95% CI: 0.16-0.98]	**8.2** kPa sAUC: 0.82 [95% CI: 0.76-0.88]	**9.4** kPa sAUC: 0.86 [95% CI: 0.77-0.93]	**11.6** kPa sAUC: 0.89 [95% CI: 0.80-0.95]
All ultrasound devices together (m/s)	N/A	**1.48** m/s sAUC: 0.82 [95% CI: 0.65-0.94]	**1.63** m/s sAUC: 0.84 [95% CI: 0.51-0.95]	**1.74** m/s sAUC: 0.89 [95% CI: 0.38-0.99]
SuperSonic Imagine (kPa)	**7.3** kPa sAUC: 0.79 [95% CI: 0.24-0.97]	**8.3** kPa sAUC: 0.82 [95% CI: 0.74-0.88]	**10.2** kPa sAUC: 0.86 [95% CI: 0.76-0.93]	**11.8** kPa sAUC: 0.88 [95% CI: 0.78-0.94]
General Electric Healthcare (kPa)	N/A	**7.2** kPa sAUC: 0.88 [95% CI: 0.74-0.95]	**8.1** kPa sAUC: 0.90 [95% CI: 0.77-0.96]	N/A
Canon Medical Systems/Toshiba (kPa)	N/A	**7.2** kPa sAUC: 0.82 [95% CI: 0.26-0.96]	**7.8** kPa sAUC: 0.86 [95% CI: 0.21-0.98]	N/A
Optimal cut-off values for ruling-out advanced fibrosis (≥F3) at different pre-test probabilities in biopsy-proven MASLD				
Pre-test probability	5%	10%	30%	60%
All ultrasound devices together **6.5** kPa	NPV=0.99	NPV=0.98	NPV=0.94	NPV=0.81
SuperSonic Imagine **7.4** kPa	NA	NA	NPV=0.94	NPV=0.81
All ultrasound devices together **1.34** m/s	NPV=0.99	NPV=0.98	NPV=0.93	NPV=0.79

FIGURE 4 Optimal 2D-SWE cutoff values for differentiating different stages of liver fibrosis in biopsy-proven MASLD. Abbreviations: MASLD, metabolic dysfunction–associated steatotic liver disease; N/A, not available; NPV, negative predictive value; PPV, positive predictive value; sAUC, summary area under the receiver operator characteristic curve.

Se [95% CI], and Sp [95% CI] were 0.88 [0.74–0.95], 0.81 [0.56–0.94], and 0.81 [0.56–0.94] for ≥F2; 0.90 [0.77 – 0.96], 0.84 [0.60 – 0.95], and 0.84 [0.62 – 0.94] for ≥F3, respectively. The optimal cutoffs in detecting ≥F2 and ≥F3 were 7.204 and 8.060 kPa, respectively (Table 3, Supplemental Figure S4, http://links.lww.com/HEP/J663).

Se analysis of the diagnostic performance of 2D-SWE (kPa) in different clinical scenarios

We evaluated the accuracy of 2D-SWE cutoffs in different disease prevalences representing different clinical scenarios (1%, 2.5%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, and 70%). We also calculated the NPV and PPV for different stages of liver fibrosis (pretest probability) in different subpopulations. The results are found in Table 4 and Supplemental Tables S10–S14, http://links.lww.com/HEP/J663. An Se analysis of the diagnostic performance of 2D-SWE based on ALT levels (stratified by quartiles) is shown in Supplemental Table S15, http://links.lww.com/HEP/J663 and Supplemental Figure S6, http://links.lww.com/HEP/J663. Finally, Supplemental Table S16, http://links.lww.com/HEP/J663 and Supplemental Figure S7, http://links.lww.com/HEP/J663 present the diagnostic performance of VCTE in identifying various stages of liver fibrosis in biopsy-proven MASLD, utilizing a multilevel random effects model.

DISCUSSION

Our study employs a multi-threshold approach within the meta-analysis to provide robust evidence on the applicability of 2D-SWE for identifying various stages of liver fibrosis and to contribute to the standardization of cutoff values. This standardization has the potential to advance the field and offer practical cutoff value for clinical practice (Figure 4).

The 2017 EFSUMB Guidelines indicate that 2D-SWE measurement bias can vary based on the software used.[19] However, technological advancements have reduced variance in fibrosis staging across US systems, largely due to the efforts of the Quantitative Image Biomarker Alliance (QIBA).[55,56] Current guidelines and position papers highlight the need for standardized thresholds for 2D-SWE.[19,20,22,57–59] The Society of Radiologists in Ultrasound (SRU) consensus statement suggests that unique cutoff values for each manufacturer may not be necessary.[22] Despite this, discrepancies in cutoff values exist, particularly among patients

TABLE 3 Diagnostic performance of 2D-SWE in detecting different stages of liver fibrosis in biopsy-proven MASLD (all US devices together) by multilevel random effects model

Fibrosis stage	No. studies/ no. cutoffs	Cutoff	Se [95% CI]	Sp [95% CI]	sAUC for Se given Sp [95% CI]	sAUC for Sp given Se [95% CI]
All US devices together						
Measurement in kPa						
Mild fibrosis, ≥F1; F0 vs. F1-F4	9/13	6.432	0.76 [0.66–0.83]	0.76 [0.60–0.86]	0.82 [0.16–0.98]	0.82 [0.14–0.98]
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	16/28	8.174	0.76 [0.65–0.84]	0.76 [0.65–0.84]	0.82 [0.76–0.88]	0.82 [0.76–0.87]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	17/29	9.418	0.79 [0.67–0.88]	0.79 [0.71–0.86]	0.86 [0.77–0.93]	0.86 [0.79–0.92]
Cirrhosis, F4; F0-F3 vs. F4	12/20	11.548	0.82 [0.66–0.91]	0.82 [0.75–0.87]	0.89 [0.80–0.95]	0.89 [0.83–0.93]
Measurement in m/s						
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	3/6	1.481	0.76 [0.42–0.93]	0.76 [0.43–0.93]	0.82 [0.65–0.94]	0.82 [0.61–0.92]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	3/6	1.633	0.77 [0.63–0.87]	0.77 [0.70–0.83]	0.84 [0.51–0.95]	0.84 [0.57–0.94]
Cirrhosis, F4; F0-F3 vs. F4	3/6	1.748	0.82 [0.56–0.94]	0.82 [0.75–0.88]	0.89 [0.38–0.99]	0.89 [0.34–0.97]

Abbreviations: 2D-SWE, 2-dimensional shear-wave elastography; F1, mild fibrosis; F2, significant fibrosis; F3, advanced fibrosis; F4, cirrhosis; MASLD, metabolic dysfunction–associated steatotic liver disease; sAUC, summary area under the receiver operator characteristic curve; Se, sensitivity; Sp, specificity; US, ultrasound.

with compensated advanced chronic liver disease (cACLD).[60] Proposed 2D-SWE cutoff values from various guidelines are presented in Supplemental Table S17, http://links.lww.com/HEP/J663, although their interpretation has been cautious pending further validation.

This meta-analysis specifically aims to standardize and validate 2D-SWE thresholds. The SRU and WFUMB recommend the "rule of four" for acoustic radiation force impulse techniques in diagnosing cACLD.[22,58] In patients with MASLD, lower cutoffs may be applicable, with follow-up testing recommended for values between 7 and 9 kPa.[22] According to the AASLD practice guidelines, 2D-SWE values <7–8 kPa could rule out cACLD, while cutoffs between 13 and 16 kPa could confirm cACLD and values of 17–20 kPa could predict the presence of varices and decompensation.[20]

When compared to VCTE, 2D-SWE (particularly SSI) demonstrates similar accuracy,[19] with potentially greater applicability, as most modern US machines include elastography software.[61] In our meta-analysis, the cutoff values derived from the multilevel random effects model align with current guidelines: the optimal cutoff values are 6.4, 8.2, 9.4, and 11.6 kPa for stages ≥F1, ≥F2, ≥F3, and F4, respectively; 6.5 kPa to rule out ≥F3, and 13.6 kPa to rule in ≥F3. For SSI machines specifically, accounting for 9 studies (45% of the total), the optimal cutoff values are 7.3, 8.3, 10.2, and 11.8 kPa for ≥F1, ≥F2, ≥F3, and F4, respectively; 7.4 kPa to rule out ≥F3 and 14 kPa to rule in ≥F3.

In addition, we reported thresholds based on different pretest probabilities, as cutoff values can vary depending on the prevalence of the condition within the target population.[20] The prevalence of advanced fibrosis in high-risk MASLD populations ranges from 4% to 33%,[57] with our study population showing a prevalence of 17.9%. We provided cutoff values for different stages of liver fibrosis at various pretest probabilities, reflecting clinical scenarios in both primary care and tertiary hospital settings.

Overall diagnostic performance of 2D-SWE

We identified optimized cutoff values for each fibrosis stage with excellent sAUC, Se, and Sp. When pooling data from all US manufacturers, 2D-SWE showed excellent Se and Sp for detecting mild fibrosis (≥F1), significant fibrosis (≥F2), advanced fibrosis (≥F3), and cirrhosis (F4), with optimal cutoffs ranging from 6.432 to 11.548 kPa. When examining how ethnicity and comorbidities affect the accuracy of 2D-SWE, the optimal cutoffs showed minor differences, but these were not significant. In the Asian population, the optimal cutoffs tended to be higher for every stage of liver fibrosis. This

TABLE 4 Sensitivity analysis of the diagnostic performance of 2D-SWE (kPa) for different stages of liver fibrosis by multilevel random effects model

Fibrosis stage	No. studies/ no. cutoffs	Cutoff	Se [95% CI]	Sp [95% CI]	sAUC for Se given Sp [95% CI]	sAUC for Sp given Se [95% CI]
Measurement in kPa						
Asian population						
Mild fibrosis, ≥F1; F0 vs. F1-F4	8/10	6.953	0.76 [0.62–0.86]	0.76 [0.57–0.88]	0.83 [0.73–0.90]	0.83 [0.70–0.91]
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	11/16	8.561	0.74 [0.60–0.84]	0.74 [0.63–0.83]	0.80 [0.69–0.88]	0.80 [0.71–0.88]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	12/17	10.182	0.78 [0.60–0.89]	0.78 [0.68–0.85]	0.84 [0.43–0.96]	0.84 [0.47–0.96]
Cirrhosis, F4; F0-F3 vs. F4	9/12	12.466	0.83 [0.62–0.94]	0.83 [0.75–0.89]	0.90 [0.78–0.97]	0.90 [0.81–0.95]
Population with T2DM						
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	9/15	7.961	0.77 [0.64–0.87]	0.77 [0.63–0.87]	0.84 [0.77–0.90]	0.84 [0.76–0.90]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	10/16	9.010	0.81 [0.68–0.90]	0.81 [0.69–0.89]	0.88 [0.81–0.93]	0.88 [0.81–0.92]
Cirrhosis, F4; F0-F3 vs. F4	6/10	11.760	0.81 [0.67–0.90]	0.81 [0.73–0.87]	0.88 [0.81–0.93]	0.88 [0.83–0.91]
Population with obesity						
Mild fibrosis, ≥F1; F0 vs. F1-F4	7/11	6.792	0.74 [0.61–0.84]	0.74 [0.56–0.87]	0.81 [0.72–0.87]	0.81 [0.69–0.89]
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	13/23	7.945	0.77 [0.65–0.85]	0.77 [0.65–0.85]	0.83 [0.77–0.89]	0.83 [0.76–0.88]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	14/24	9.040	0.80 [0.68–0.88]	0.80 [0.71–0.87]	0.87 [0.80–0.92]	0.87 [0.81–0.91]
Cirrhosis, F4; F0-F3 vs. F4	9/15	11.614	0.81 [0.66–0.91]	0.81 [0.74–0.87]	0.88 [0.79–0.94]	0.88 [0.82–0.93]
Type of ultrasound manufacturer						
SuperSonic Imagine						
Mild fibrosis, ≥F1; F0 vs. F1-F4	3/5	7.305	0.73 [0.45–0.90]	0.73 [0.41–0.91]	0.79 [0.24–0.97]	0.79 [0.12–0.97]
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	9/17	8.284	0.75 [0.63–0.84]	0.75 [0.63–0.85]	0.82 [0.74–0.88]	0.82 [0.73–0.88]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	9/17	10.172	0.79 [0.67–0.88]	0.79 [0.70–0.86]	0.86 [0.76–0.93]	0.86 [0.76–0.92]
Cirrhosis, F4; F0-F3 vs. F4	7/15	11.848	0.81 [0.66–0.90]	0.81 [0.73–0.88]	0.88 [0.78–0.94]	0.88 [0.81–0.93]
Canon Medical Systems/Toshiba						
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	4/6	7.226	0.75 [0.45–0.91]	0.76 [0.61–0.86]	0.82 [0.26–0.96]	0.82 [0.38–0.96]
Advanced fibrosis, ≥F3; F0-F2 vs. F3-F4	4/6	7.831	0.80 [0.51–0.94]	0.80 [0.69–0.87]	0.86 [0.21–0.98]	0.86 [0.33–0.97]
General Electric Healthcare						
Significant fibrosis, ≥F2; F0-F1 vs. F2-F4	3/5	7.204	0.81 [0.56–0.94]	0.81 [0.56–0.94]	0.88 [0.74–0.95]	0.88 [0.76–0.95]
Advanced fibrosis, ≥F3; F0-F2 vs. F3–F4	4/6	8.060	0.84 [0.60–0.95]	0.84 [0.62–0.94]	0.90 [0.77–0.96]	0.90 [0.81–0.97]

Abbreviations: 2D-SWE, 2-dimensional shear-wave elastography; F1, mild fibrosis; F2, significant fibrosis; F3, advanced fibrosis; F4, cirrhosis; sAUC, summary area under the receiver operator characteristic curve; Se, sensitivity; Sp, specificity; T2DM, type 2 diabetes mellitus.

variation could be explained by population differences. However, after conducting subanalyses, it can be concluded that these values did not vary significantly.

After optimizing the thresholds, ≥F3 could be ruled out with excellent Se and NPV for all US devices together (7.4 kPa), T2DM subpopulation (7.7 kPa), the subgroup with prevalent obesity (7 kPa), non-Asian studies (7.5 kPa), and thresholds that are comparable to the 8 kPa cutoff proposed for LSM-VCTE.[15] Similarly, the rule-in cutoffs for ≥F3 with excellent Sp and PPV by all US devices together (13.6 kPa), among the T2DM subpopulation (10.6 kPa), the subpopulation with prevalent obesity (11.7 kPa) are comparable to the 12 kPa cutoff proposed for LSM-VCTE.[15]

Type of US manufacturer

The optimized thresholds provided by SSI for diagnosing different stages of liver fibrosis were 7.3 kPa for ≥F1, 8.3 kPa for ≥F2, 10.2 kPa for ≥F3, and 11.9 kPa for F4. These values were similar to those obtained when all US devices were analyzed together. However, cutoffs provided by Canon Medical Systems/Toshiba (7.3 kPa for ≥F2, 7.8 kPa for ≥F3) and General Electric Healthcare (7.2 kPa for ≥F2, 8.1 kPa for ≥F3) tended to be lower, keeping in with some previous comparative findings.[21] The difference in cutoff values might be attributed to specific characteristics of the US machine, including transducer technology, signal processing methods, shear-wave frequency, and calibration standards.[19] The threshold variation could also be explained by the fact that a considerable proportion of the studies included in the meta-analysis (9 out of 20, 45%) reported SSI. Nevertheless, machines from all companies demonstrate excellent performance in diagnosing different stages of liver fibrosis, albeit with slight variations in Se and Sp. Therefore, these differences may arise from the technical aspects of the machines, underscoring the need for standardized reporting and broad validation across US platforms.[22]

Comparison with similar published meta-analyses

Our meta-analysis stands out from previous studies in its application of a multilevel random effects model in biopsy-proven MASLD, a novel approach that has not been used before. While other studies have evaluated the diagnostic performance of 2D-SWE in mixed liver disease etiologies, or small biopsy-proven MASLD cohorts, none have taken this unique approach. Only Selvaraj et al[14] and Herrmann et al[62] have meta-analyzed data from 4 SSI studies, with comparable results regarding sAUC values (Supplemental Table S18, http://links.lww.com/HEP/J663).

Strengths and limitations

Our meta-analysis has significant strengths. First, we applied a multilevel random effects model, demonstrating robustness and improved efficacy compared to the standard bivariate model.[63,64] Second, we conducted a systematic literature search without any restrictions on sex, race, ethnicity, or language of publication. Third, we identified the optimized cutoff values for fibrosis staging from the vastest cohort of biopsy-proven MASLD. Despite these strengths, certain limitations must be acknowledged. The heterogeneity in study designs and US manufacturers may account for variability in the results. In addition, our reliance on published data restricted our capacity to control for potential confounders, including individual patient characteristics and factors that could have influenced 2D-SWE measurements, such as hypervolemia or chronic renal disease. Addressing these limitations would necessitate an individual patient data meta-analysis. To minimize this risk, we restricted our inclusion to studies that utilized histology as the reference standard.

The methodological quality of the included studies was not always optimal. Most of the studies did not follow the "intention to diagnose" approach when assessed using the QUADAS-2 tool. Thus, selection bias in flow and timing could not be excluded. The lack of an "intention to diagnose" approach across studies demonstrates the relatively low quality of some publications on this topic. Therefore, researchers and publishers should consider this approach for future publications on 2D-SWE.

CONCLUSIONS

In conclusion, this meta-analysis utilizing a multilevel random effects model provides strong evidence supporting the effectiveness of 2D-SWE in accurately identifying stages of liver fibrosis in patients with metabolic-associated steatotic liver disease. The findings, derived from over 2000 patients with biopsy-confirmed MASLD, emphasize the importance of incorporating patient-specific factors and variations in machine performance to achieve reliable non-invasive assessments. Standardized cutoff values established in this study can guide clinicians in making informed decisions and enhance patient risk stratification.

DATA AVAILABILITY STATEMENT

The full search strategy and key results used to generate data that inform the conclusions of this systematic review and multilevel random effects model meta-analysis can be found in Supplemental Materials, http://links.lww.com/HEP/J663.

AUTHOR CONTRIBUTIONS

Madalina-Gabriela Indre: study concept and design, protocol development, data collection and extraction, analysis and interpretation of data, technical and material support, study supervision, drafting the manuscript, and preparation of the final version of the manuscript. Monica Lupsor-Platon: protocol development, interpretation of data, technical and material support, preparation of the final version of the manuscript, and critical manuscript revision for important intellectual content. Laura Turco: interpretation of data, technical and material support, critical revision of the manuscript for important intellectual content, and preparation of the final version of the manuscript. Dan-Corneliu Leucuta: protocol development, search strategy design and adaptation for specific databases, data collection and extraction, initial risk of bias assessment, analysis and interpretation of data, technical and material support, statistical analysis, and preparation of the final version of the manuscript. Ahmed Hashim, Silvia Ferri, Olga Hilda Orasan, Bogdan Procopet, Horia Stefanescu, and Maria Cristina Morelli: interpretation of data, technical and material support, and critical revision of the manuscript for important intellectual content. Fabio Piscaglia: study concept and design, technical and material support, study supervision, data analysis and interpretation, critical revision of the manuscript for important intellectual content, and preparation of the final version of the manuscript. Federico Ravaioli: study concept and design, data analysis and interpretation, study supervision, technical and material support, statistical analysis, critical revision of the manuscript for important intellectual content, and preparation of the final version of the manuscript. All authors had full access to all the study data, reviewed the manuscript, and were ultimately responsible for the decision to submit it for publication.

FUNDING INFORMATION

Madalina-Gabriela Indre is financed by The Study Loans and Scholarships Agency, The Ministry of Education, Romania. The work was partially funded by UEFISCDI, the Romanian Executive Agency for Higher Education, Research, Development, and Innovation Funding, grant number PN-III-P4-PCE-2021-1140.

CONFLICTS OF INTEREST

Bogdan Procopet advises Boehringer Ingelheim International. He is on the speakers' bureau for AbbVie. Horia Stefanescu is on the speakers' bureau for General Electric Healthcare and MGT. Fabio Piscaglia advises and is on the speakers' bureau for AstraZeneca, Eisai, MSD, Siemens Healthineers, and Roche. He consults for and is on the speakers' bureau for Bracco. He consults for, advises, and is on the speakers' bureau for Nerviano. He consults for Signant Health. He advises BMS. He is on the speakers' bureau for Bayer, Esaote, Exact Sciences, Gilead, GE, Ipsen, and Samsung. The remaining authors have no conflicts to report.

ORCID

Madalina-Gabriela Indre https://orcid.org/0009–0005–1457–0076
Dan-Corneliu Leucuta https://orcid.org/0000–0003–4218–8622
Monica Lupsor-Platon https://orcid.org/0000–0001–7918–1956
Laura Turco https://orcid.org/0000–0001–8148–3769
Silvia Ferri https://orcid.org/0000–0002–6033–8594
Olga Hilda Orasan https://orcid.org/0000–0003–4435–142X
Horia Stefanescu https://orcid.org/0000–0002–4034–5471
Fabio Piscaglia https://orcid.org/0000–0001–8264–1845
Federico Ravaioli https://orcid.org/0000–0002–1142–8585

REFERENCES

1. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. A multi-society Delphi consensus statement on new fatty liver disease nomenclature. Hepatology. 2023;78:1966–86.
2. Riazi K, Azhari H, Charette JH, Underwood FE, King JA, Afshar EE, et al. The prevalence and incidence of NAFLD worldwide: A systematic review and meta-analysis. Lancet Gastroenterol Hepatol. 2022;7:851–61.
3. Le MH, Yeo YH, Li X, Li J, Zou B, Wu Y, et al. 2019 Global NAFLD prevalence: A systematic review and meta-analysis. Clin Gastroenterol Hepatol. 2022;20:2809–817 e28.
4. Hashida R, Nakano D, Kawaguchi M, Younossi ZM, Kawaguchi T. Changing from NAFLD to MASLD: The Implications for health related quality of life data. J Hepatol. 2024;80:e249–51.
5. Sanyal AJ, Van Natta ML, Clark J, Neuschwander-Tetri BA, Diehl A, Dasarathy S, et al. Prospective study of outcomes in adults with nonalcoholic fatty liver disease. N Engl J Med. 2021;385:1559–69.
6. Ng CH, Lim WH, Hui Lim GE, Hao Tan DJ, Syn N, Muthiah MD, et al. Mortality outcomes by fibrosis stage in nonalcoholic fatty liver disease: A systematic review and meta-analysis. Clin Gastroenterol Hepatol. 2023;21:931–39.e5
7. Tracey GS, Bjorn R, Hannes H, Johan S, Jonas FL. Non-alcoholic fatty liver disease and incident major adverse cardiovascular events: Results from a nationwide histology cohort. Gut. 2022;71:1867.
8. Lazarus JV, Mark HE, Anstee QM, Arab JP, Batterham RL, Castera L, et al. Advancing the global public health agenda for NAFLD: A consensus statement. Nat Rev Gastroenterol Hepatol. 2022;19:60–78.
9. Taru MG, Neamti L, Taru V, Procopciuc LM, Procopet B, Lupsor-Platon M. How to identify advanced fibrosis in adult patients with non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) using ultrasound elastography—A review of the literature and proposed multistep approach. Diagnostics (Basel). 2023;13:788.
10. Allen AM, Therneau TM, Ahmed OT, Gidener T, Mara KC, Larson JJ, et al. Clinical course of non-alcoholic fatty liver disease and the implications for clinical trial design. J Hepatol. 2022;77:1237–45.

11. Harrison SA, Taub R, Neff GW, Lucas KJ, Labriola D, Moussa SE, et al. Resmetirom for nonalcoholic fatty liver disease: A randomized, double-blind, placebo-controlled phase 3 trial. Nat Med. 2023;29:2919–28.
12. Zhou JH, Cai JJ, She ZG, Li HL. Noninvasive evaluation of nonalcoholic fatty liver disease: Current evidence and practice. World J Gastroenterol. 2019;25:1307–26.
13. Brunt EM, Clouston AD, Goodman Z, Guy C, Kleiner DE, Lackner C, et al. Complexity of ballooned hepatocyte feature recognition: Defining a training atlas for artificial intelligence-based imaging in NAFLD. J Hepatol. 2022;76:1030–41.
14. Selvaraj EA, Mózes FE, Jayaswal ANA, Zafarmand MH, Vali Y, Lee JA, et al. Diagnostic accuracy of elastography and magnetic resonance imaging in patients with NAFLD: A systematic review and meta-analysis. J Hepatol. 2021;75:770–85.
15. Berzigotti A, Tsochatzis E, Boursier J, Castera L, Cazzagon N, Friedrich-Rust M, et al. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis—2021 update. J Hepatol. 2021;75:659–89.
16. Boursier J, Vergniol J, Guillet A, Hiriart JB, Lannes A, Le Bail B, et al. Diagnostic accuracy and prognostic significance of blood fibrosis tests and liver stiffness measurement by FibroScan in non-alcoholic fatty liver disease. J Hepatol. 2016;65:570–8.
17. Mulazzani L, Salvatore V, Ravaioli F, Allegretti G, Matassoni F, Granata R, et al. Point shear wave ultrasound elastography with Esaote compared to real-time 2D shear wave elastography with supersonic imagine for the quantification of liver stiffness. J Ultrasound. 2017;20:213–25.
18. Liu F, Bi M, Jing X, Ding H, Zeng J, Zheng R, et al. Multiparametric US for identifying metabolic dysfunction–associated steatohepatitis: A prospective multicenter study. Radiology. 2024;310:e232416.
19. Dietrich C, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, et al. EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version) (EFSUMB-Leitlinien und Empfehlungen zur klinischen Anwendung der Leberelastographie, Update 2017 (Langversion)). Ultraschall Med. 2017;38:e16–47.
20. Kaplan DE, Ripoll C, Thiele M, Fortune BE, Simonetto DA, Garcia-Tsao G, et al. AASLD Practice Guidance on risk stratification and management of portal hypertension and varices in cirrhosis. Hepatology. 2024;79:1180–211.
21. Piscaglia F, Salvatore V, Mulazzani L, Cantisani V, Colecchia A, Di Donato R, et al. Differences in liver stiffness values obtained with new ultrasound elastography machines and Fibroscan: A comparative study. Dig Liver Dis. 2017;49:802–8.
22. Barr RG, Wilson SR, Rubens D, Garcia-Tsao G, Ferraioli G. Update to the Society of Radiologists in Ultrasound Liver Elastography Consensus Statement. Radiology. 2020;296:263–74.
23. Kjaergaard M, Lindvig KP, Thorhauge KH, Andersen P, Hansen JK, Kastrup N, et al. Using the ELF test, FIB-4 and NAFLD fibrosis score to screen the population for liver disease. J Hepatol. 2023;79:277–86.
24. Kanwal F, Shubrook JH, Adams LA, Pfotenhauer K, Wai-Sun Wong V, Wright E, et al. Clinical care pathway for the risk stratification and management of patients with nonalcoholic fatty liver disease. Gastroenterology. 2021;161:1657–69.
25. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 2009;6:e1000097.
26. Cohen JF, Deeks JJ, Hooft L, Salameh JP, Korevaar DA, Gatsonis C, et al. Preferred reporting items for journal and conference abstracts of systematic reviews and meta-analyses of diagnostic test accuracy studies (PRISMA-DTA for Abstracts): Checklist, explanation, and elaboration. BMJ. 2021;372:n265.
27. Salameh JP, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): Explanation, elaboration, and checklist. BMJ. 2020;370:m2632.
28. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529–36.
29. Steinhauser S, Schumacher M, Rücker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol. 2016;16:97
30. Schneider A, Linde K, Reitsma JB, Steinhauser S, Rücker G. A novel statistical model for analyzing data of a systematic review generates optimal cutoff values for fractional exhaled nitric oxide for asthma diagnosis. J Clin Epidemiol. 2017;92:69–78.
31. Organization WH. The Asia-Pacific Perspective: Redefining Obesity and Its Treatment; 2000.
32. Haam J-H, Kim BT, Kim EM, Kwon H, Kang JH, Park JH, et al. Diagnosis of obesity: 2022 update of clinical practice guidelines for obesity by the Korean Society for the Study of Obesity. J Obes Metab Syndrome. 2023;32:121.
33. WHO Consultation. Obesity: Preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser. 2000;894:i–xii; 1–253.
34. Bürkner PC, Doebler P. Testing for publication bias in diagnostic meta-analysis: A simulation study. Stat Med. 2014;33:3061–77.
35. Cassinotto C, Boursier J, de Lédinghen V, Lebigot J, Lapuyade B, Cales P, et al. Liver stiffness in nonalcoholic fatty liver disease: A comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. Hepatology. 2016;63:1817–27.
36. Lee MS, Bae JM, Joo SK, Woo H, Lee DH, Jung YJ, et al. Prospective comparison among transient elastography, supersonic shear imaging, and ARFI imaging for predicting fibrosis in nonalcoholic fatty liver disease. PLoS One. 2017;12:e0188321.
37. Takeuchi H, Sugimoto K, Oshiro H, Iwatsuka K, Kono S, Yoshimasu Y, et al. Liver fibrosis: Noninvasive assessment using supersonic shear imaging and FIB4 index in patients with non-alcoholic fatty liver disease. J Med Ultrason (2001). 2018;45:243–9.
38. Jamialahmadi T, Nematy M, Jangjoo A, Goshayeshi L, Rezvani R, Ghaffarzadegan K, et al. Measurement of liver stiffness with 2D-shear wave elastography (2D-SWE) in bariatric surgery candidates reveals acceptable diagnostic yield compared to liver biopsy. Obes Surg. 2019;29:2585–92.
39. Ozturk A, Mohammadi R, Pierce TT, Kamarthi S, Dhyani M, Grajo JR, et al. Diagnostic accuracy of shear wave elastography as a non-invasive biomarker of high-risk non-alcoholic steatohepatitis in patients with non-alcoholic fatty liver disease. Ultrasound Med Biol. 2020;46:972–80.
40. Sharpton SR, Tamaki N, Bettencourt R, Madamba E, Jung J, Liu A, et al. Diagnostic accuracy of two-dimensional shear wave elastography and transient elastography in nonalcoholic fatty liver disease. Therap Adv Gastroenterol. 2021;14:17562848211050436.
41. Chen G, Yang JC, Zhang GX, Cheng Z, Du X. Evaluation of six noninvasive methods for the detection of fibrosis in Chinese patients with obesity and nonalcoholic fatty liver disease. Obes Surg. 2022;32:3619–26.
42. Mendoza YP, Rodrigues SG, Delgado MG, Murgia G, Lange NF, Schropp J, et al. Inflammatory activity affects the accuracy of liver stiffness measurement by transient elastography but not by two-dimensional shear wave elastography in non-alcoholic fatty liver disease. Liver Int. 2022;42:102–11.
43. Zhou J, Yan F, Xu J, Lu Q, Zhu X, Gao B, et al. Diagnosis of steatohepatitis and fibrosis in biopsy-proven nonalcoholic fatty liver diseases: Including two-dimension real-time shear wave elastography and noninvasive fibrotic biomarker scores. Quant Imaging Med Surg. 2022;12:1800–14.

44. Furlan A, Tublin ME, Yu L, Chopra KB, Lippello A, Behari J. Comparison of 2D shear wave elastography, transient elastography, and MR elastography for the diagnosis of fibrosis in patients with nonalcoholic fatty liver disease. AJR Am J Roentgenol. 2020;214:W20–w26.
45. Kuroda H, Fujiwara Y, Abe T, Nagasawa T, Oguri T, Noguchi S, et al. Two-dimensional shear wave elastography and ultrasound-guided attenuation parameter for progressive non-alcoholic steatohepatitis. PLoS One. 2021;16:e0249493.
46. Imajo K, Honda Y, Kobayashi T, Nagai K, Ozaki A, Iwaki M, et al. Direct comparison of US and MR elastography for staging liver fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2022;20:908–917.e11.
47. Kalaiyarasi K, Sanchalika A, Hsien Min L, Wei Ming Y, Vishalkumar S, Kuo Chao Y, et al. Transient elastography is the best-performing non-invasive test of liver fibrosis in obese Asian patients with non-alcoholic fatty liver disease: A pilot, cross-sectional study. Medicina (Kaunas). 2024;60:169.
48. Lee DH, Cho EJ, Bae JS, Lee JY, Yu SJ, Kim H, et al. Accuracy of two-dimensional shear wave elastography and attenuation imaging for evaluation of patients with nonalcoholic steatohepatitis. Clin Gastroenterol Hepatol. 2021;19:797–805.e7.
49. Jang JK, Lee ES, Seo JW, Kim YR, Kim SY, Cho YY, et al. Two-dimensional shear-wave elastography and US attenuation imaging for nonalcoholic steatohepatitis diagnosis: A cross-sectional, multicenter study. Radiology. 2022;305:118–26.
50. Laroia ST, Vellore Srinivasan S, Yadav K, Rastogi A, Kumar S, Kumar G, et al. Performance of shear wave elastography: A single centre pilot study of mixed etiology liver disease patients with normal BMI. Australas J Ultrasound Med. 2021;24:120–36.
51. Kim JW, Lee CH, Kim B-H, Lee YS, Hwang SY, Park BN, et al. Ultrasonographic index for the diagnosis of non-alcoholic steatohepatitis in patients with non-alcoholic fatty liver disease. Quant Imaging Med Surg. 2021;12:1815–29.
52. Zhang YN, Fowler KJ, Boehringer AS, Montes V, Schlein AN, Covarrubias Y, et al. Comparative diagnostic performance of ultrasound shear wave elastography and magnetic resonance elastography for classifying fibrosis stage in adults with biopsy-proven nonalcoholic fatty liver disease. Eur Radiol. 2022;32:2457–69.
53. Ogino Y, Wakui N, Nagai H, Matsuda T. Comparison of strain elastography and shear wave elastography in diagnosis of fibrosis in nonalcoholic fatty liver disease. J Med Ultrason (2001). 2023;50:187–95.
54. Sugimoto K, Moriyasu F, Oshiro H, Takeuchi H, Abe M, Yoshimasu Y, et al. The role of multiparametric US of the liver for the evaluation of nonalcoholic steatohepatitis. Radiology. 2020;296:532–40.
55. Hall TJ, Milkowski A, Garra B, Carson P, Palmeri M, Nightingale K, et al. RSNA/QIBA: Shear wave speed as a biomarker for liver fibrosis staging. IEEE. 2013:397–400.
56. Palmeri M, Nightingale K, Fielding S, Rouze N, Deng Y, Lynch T, et al. RSNA QIBA ultrasound shear wave speed phase II phantom study in viscoelastic media. IEEE. 2015:1–4.
57. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, Abdelmalek MF, Caldwell S, Barb D, et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology. 2023;77:1797–835.
58. Ferraioli G, Barr RG, Berzigotti A, Sporea I, Wong VW, Reiberger T, et al. WFUMB guideline/guidance on liver multiparametric ultrasound: Part 1. Update to 2018 guidelines on liver ultrasound elastography. Ultrasound Med Biol. 2024;50:1071–87.
59. Ferraioli G, Barr RG, Berzigotti A, Sporea I, Wong VW, Reiberger T, et al. WFUMB guidelines/guidance on liver multiparametric ultrasound. Part 2: Guidance on liver fat quantification. Ultrasound Med Biol. 2024;50:1088–98.
60. Ferraioli G, De Silvestri A, Lissandrin R, Maiocchi L, Tinelli C, Filice C, et al. Evaluation of inter-system variability in liver stiffness measurements (Bewertung der Intersystem-Variabilität bei Lebersteifigkeitsmessungen). Ultraschall Med. 2019;40:64–75.
61. Cassinotto C, Charrie A, Mouries A, Lapuyade B, Hiriart JB, Vergniol J, et al. Liver and spleen elastography using supersonic shear imaging for the non-invasive diagnosis of cirrhosis severity and oesophageal varices. Dig Liver Dis. 2015;47:695–701.
62. Herrmann E, de Lédinghen V, Cassinotto C, Chu WCW, Leung VYF, Ferraioli G, et al. Assessment of biopsy-proven liver fibrosis by two-dimensional shear wave elastography: An individual patient data-based meta-analysis. Hepatology. 2018;67:260–72.
63. Rosenberger KJ, Chu H, Lin L. Empirical comparisons of meta-analysis methods for diagnostic studies: A meta-epidemiological study. BMJ Open. 2022;12:e055336.
64. Hamza TH, Arends LR, van Houwelingen HC, Stijnen T. Multivariate random effects meta-analysis of diagnostic tests with multiple thresholds. BMC Med Res Methodol. 2009;9:73.

How to cite this article: Indre M-G, Leucuta D-C, Lupsor-Platon M, Turco L, Ferri S, Hashim A, et al. Diagnostic accuracy of 2D-SWE ultrasound for liver fibrosis assessment in MASLD: A multilevel random effects model meta-analysis. Hepatology. 2025;82:454–469. https://doi.org/10.1097/HEP.0000000000001190