

BMJ Open Diagnostic accuracy of point shear wave elastography and transient elastography for staging hepatic fibrosis in patients with non-alcoholic fatty liver disease: a meta-analysis

Weixi Jiang,^{1,2} Sirun Huang,^{1,2} Hua Teng,^{1,2} Peipei Wang,^{1,2} Meng Wu,^{1,2} Xia Zhou,^{1,2} Haitao Ran^{1,2}

To cite: Jiang W, Huang S, Teng H, *et al.* Diagnostic accuracy of point shear wave elastography and transient elastography for staging hepatic fibrosis in patients with non-alcoholic fatty liver disease: a meta-analysis. *BMJ Open* 2018;**8**:e021787. doi:10.1136/bmjopen-2018-021787

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-021787>).

Received 19 January 2018
Accepted 20 July 2018



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Ultrasound Department, Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

²Ultrasound Department, Chongqing Key Laboratory of Ultrasound Molecular Imaging, Chongqing, China

Correspondence to
Professor Haitao Ran;
ranhaitao666@163.com

ABSTRACT

Objective This study aimed to assess the accuracy of staging liver fibrosis in patients with non-alcoholic fatty liver disease (NAFLD) using point shear wave elastography (pSWE) and transient elastography (TE).

Setting Relevant records on NAFLD were retrieved from PubMed, Embase, Web of Science and the China National Knowledge Infrastructure databases up to 20 December 2017. A bivariate mixed-effects model was conducted to combine sensitivity, specificity, positive likelihood ratio, negative likelihood ratio and area under the summary receiver operating characteristic curve (AUC) between pSWE and TE. A sensitivity analysis was implemented to explore the source of heterogeneity.

Participants Patients with NAFLD who had a liver stiffness measurement using pSWE and TE before liver biopsy were enrolled according to the following criteria: 2×2 contingency tables can be calculated via the reported number of cases; sensitivity and specificity were excluded according to the following criteria: history of other hepatic damage, such as chronic hepatitis C, concurrent active hepatitis B infection, autoimmune hepatitis, suspicious drug usage and alcohol abuse.

Results Nine pSWE studies comprising a total of 982 patients and 11 TE studies comprising a total of 1753 patients were included. For detection of significant fibrosis, advanced fibrosis and cirrhosis, the summary AUC was 0.86 (95% CI 0.83 to 0.89), 0.94 (95% CI 0.91 to 0.95) and 0.95 (95% CI 0.93 to 0.97) for pSWE, and the summary AUC was 0.85 (95% CI 0.82 to 0.88), 0.92 (95% CI 0.89 to 0.94) and 0.94 (95% CI 0.93 to 0.97) for TE, respectively. The proportion of failure measurement was over tenfold as common with TE using an M probe compared with pSWE.

Conclusion pSWE and TE, providing precise non-invasive staging of liver fibrosis in NAFLD, are promising techniques, particularly for advanced fibrosis and cirrhosis.

INTRODUCTION

As a result of the sprawling obesity epidemic, non-alcoholic fatty liver disease (NAFLD) has now become a worldwide chronic hepatic disease that affects approximately

Strengths and limitations of this study

- This is the first meta-analysis to assess the diagnostic performance of point shear wave elastography (pSWE) using bivariate mixed-effect model for conditions ranging from significant fibrosis to cirrhosis in patients with non-alcoholic fatty liver disease.
- The methodological quality of each study was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 for prospective observational studies.
- Sensitivity analysis was conducted to remove low-quality research, which did not change the direction of effect for any outcome.
- Because of the larger number of available samples included in this meta-analysis than previously published research, the results appear to be convincing and credible.
- The major limitation of this meta-analysis is its failure to make a direct comparison of pSWE and transient elastography (TE) due to the absence of sufficient data from studies that performed pSWE and TE simultaneously on the same patient population.
- Additionally, some heterogeneity was still present after exclusion of the outlier.

90% of the obese population and as much as 15%–40% of the general population.^{1 2} Although NAFLD is considered to be a component of metabolic syndrome that is related to type 2 diabetes and insulin resistance,^{3 4} a large amount of evidence indicates that patients with NAFLD may develop non-alcoholic steatohepatitis in greater than 30% of cases, fibrosis in approximately 25% of cases, cirrhosis in 10%–20% and even to hepatocellular carcinoma in 4% of cases.^{5 6} Therefore, diagnosing disease progression plays an essential role in treatment and prognosis, which are dependent

on the histological severity of the liver fibrosis.⁷ Traditionally, liver biopsy is regarded as the 'gold standard' in the diagnosis of NAFLD. However, this technique is limited by its invasiveness, the possibility of complications and expense.⁸⁻⁹ Hence, it seems appropriate to reappraise the diagnostic performance of other emerging non-invasive tests.¹⁰ Ultrasound elastography, aiming to replace liver biopsy, has recently emerged as a non-invasive tool suitable for the evaluation of liver fibrosis. Because ultrasound elastography is affected by the degree of liver fibrosis in chronic hepatic disease, it can be adopted to measure the mechanical properties of liver tissues. Elastography can be performed with an ultrasonic transducer in combination with shear wave techniques, such as transient elastography (TE), point shear wave elastography (pSWE) and two-dimensional shear wave elastography.¹¹⁻¹² pSWE is integrated into a conventional ultrasound device, which was known as acoustic radiation force imaging elastography quantification when first introduced by Siemens,¹³ and has been used by several ultrasound manufacturers. One benefit of this technique is that it uses energy from a transmitted pulse, and thus permits the quantitative measurement of liver stiffness. Additionally, the flexible mobility of the region of interest guarantees an easy assessment of any area of the liver.¹⁴⁻¹⁵ TE is performed with the FibroScan device (Echosens, Paris, France), which is composed of a low-frequency vibrator (50 Hz) and an ultrasonic transducer operating at 5 MHz mounted on the axis of the vibrator. A senseless vibration is generated by the vibrator and induces a shear wave propagating to the liver. A pulse-echo acquisition is released at the same time to receive and measure their speed of shear wave, which directly reflects the stiffness of tissue.¹⁶ In the last 4 years, several meta-analyses of measurements of liver stiffness in patients with viral hepatitis and patients with other chronic liver diseases using pSWE and TE have been published.¹¹⁻¹⁷⁻¹⁹ However, the unique aetiology differing from viral hepatitis makes the meta-analysis, which solely includes NAFLD requisite.¹¹⁻²⁰ A meta-analysis that was published in 2015 summarised nothing but the values of pSWE in relation to the evaluation of significant liver fibrosis ($F \geq 2$).¹⁸ Another systematic review, which performed web-based literature search of all publications in Medline, PubMed and Embase, was conducted in 2013 to compare the diagnostic accuracy between TE and plasma cytokeratin-18 fragments in patients with NAFLD.²¹ Because higher quality research with much larger samples comparing pSWE and TE for hepatic fibrosis assessment is now available, it appears it is time to estimate the performance characteristics of these two non-invasive diagnostic tests.²²⁻²⁴

METHODS

Patient and public involvement

The patients and the public were not involved in setting the research question or the outcome measures, and no

patients were involved in developing plans for the design or implementation of the study. We are unable to disseminate the results to study participants because of the nature of a meta-analysis.

Search strategy

Relevant records on NAFLD were retrieved from the following four databases up to 20 December 2017: PubMed, Embase, Web of Science and the China National Knowledge Infrastructure. Different combinations of the following free words and medical subject headings (MeSH) words were used as search terms (the following is an example for PubMed): (non-alcoholic fatty liver disease, NAFLD, OR NASH; point shear wave elastography, pSWE, VTTQ OR acoustic radiation force impulse elastography; FibroScan, transient elastography OR TE). (((((((('Non-alcoholic Fatty Liver Disease' (MeSH)) OR Non alcoholic Fatty Liver Disease) OR NAFLD) OR Nonalcoholic Fatty Liver Disease) OR Nonalcoholic Steatohepatitis) OR Nonalcoholic Fatty Liver) OR Steatohepatitis, Nonalcoholic) OR Fatty Liver,)) OR NASH)) AND (((('Elasticity Imaging Techniques' (MeSH)) OR Acoustic Radiation Force Impulse Imaging) OR Point Shear Wave Elastography) OR pSWE) OR VTQ) OR VTTQ), (((((transient elastography) OR transient sonoelastography)) OR fibroscan)) AND (((((((('Non-alcoholic Fatty Liver Disease' (MeSH)) OR non alcoholic fatty liver disease) OR nafld) OR nonalcoholic fatty liver disease) OR nonalcoholic steatohepatitis) OR nonalcoholic fatty liver) OR steatohepatitis, nonalcoholic) OR Fatty Liver,). Additional articles were obtained by reviewing the references of the retrieved articles (see online supplementary appendix 1).

Study selection and data extraction

Two reviewers independently filtered the titles and abstracts of the original articles that referred to the diagnostic capabilities of pSWE or TE in staging hepatic fibrosis in patients with NAFLD. The following required information was extracted: (1) the first author's name, (2) country, (3) number of cases, (4) age, (5) gender, (6) cut-off values, (7) sensitivities (SE), (8) specificities (SP), (9) area under the summary receiver operating characteristic curve (AUC) in the different stages of fibrosis and (10) a biopsy standard. Moreover, information related to the clinical setting, funding and technology failure in the two methods were extracted when available. Missing data were noted as not reported if it cannot be extracted from original articles. Next, 2x2 contingency tables were calculated using the reported number of cases, SE and SP. Studies were excluded according to the following criteria: no full-length text, reviews, paediatric and animal research; failure to perform liver biopsy as the 'gold-standard'; failure to extract a 2x2 contingency table; and enrolment of patients with history of other hepatic damage, such as chronic hepatitis C, concurrent active hepatitis B infection, autoimmune hepatitis, suspicious drug usage and alcohol abuse. Indefinite data and

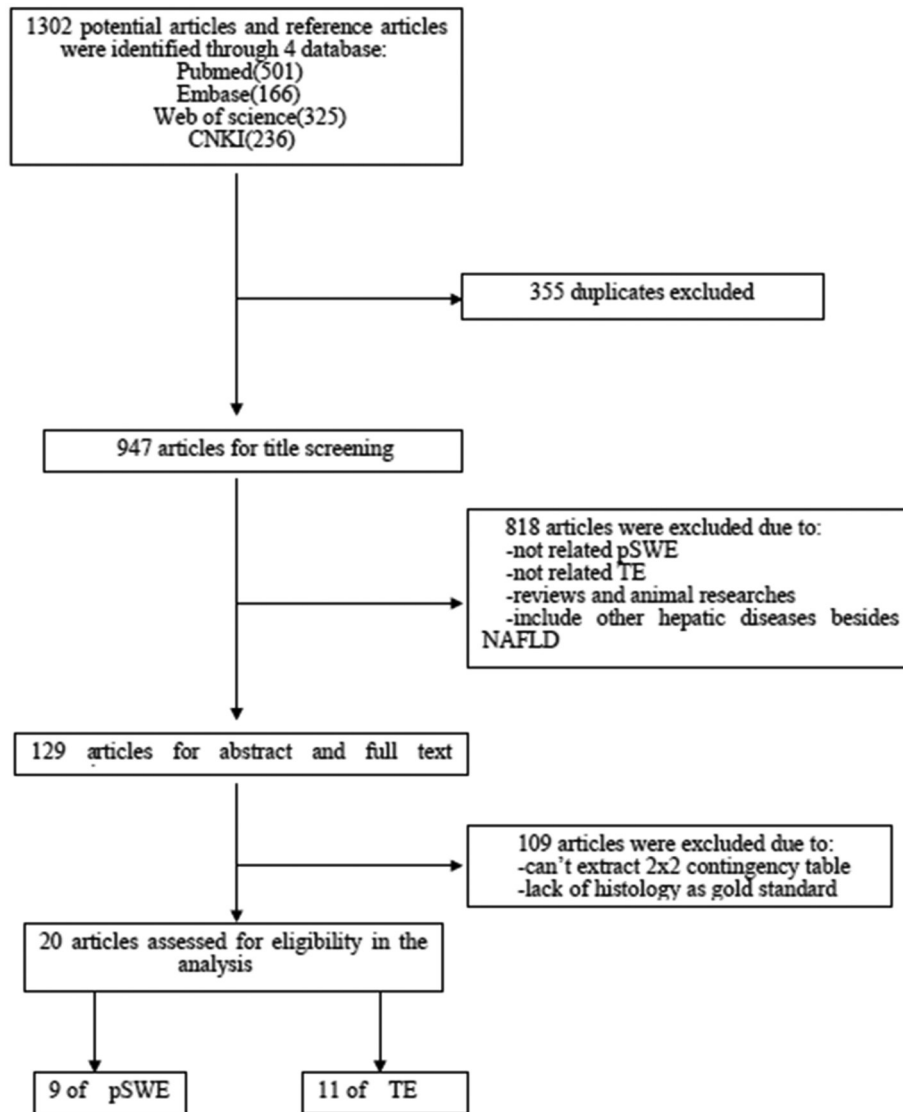


Figure 1 Flow chart illustrating the electronic database searches and selection of studies in the meta-analysis. CNKI, China National Knowledge Infrastructure; NAFLD, non-alcoholic fatty liver disease; pSWE, point shear wave elastography; TE, transient elastography.

discrepancies between the two reviewers were resolved by consensus reached through further discussion.

Quality assessment

The reliabilities of the studies were evaluated using the Quality Assessment of Diagnostic Accuracy Studies-2, which consists of four key factors. We also assessed whether the pSWE and TE were performed according to the manufacturer's instructions, which require at least 10 successful measurements with a success rate (the ratio of the number of successful acquisitions over the total number of acquisitions) over 60% and an IQR interval (IQR/M=the difference between the 75th and the 25th percentile, essentially the range of the middle 50% of the data) less than 30%.

Data synthesis and statistical analysis

The bivariate mixed-effect meta-analysis model was used to calculate the pooled parameters and their corresponding

95% CIs. Based on the 2×2 contingency tables, comparisons of the pooled SE, SP, positive likelihood ratio (LR+), negative likelihood ratio (LR-) and areas under the summary receiver operating characteristic curve between pSWE and TE in the detection of liver fibrosis were conducted to evaluate the performance of the two diagnoses. The Q test and the I^2 statistic were implemented to detect heterogeneity among the included studies. The existence of substantial heterogeneity was proven if the p value was less than 0.05 for a χ^2 test paired with an I^2 value over 50%. Furthermore, we employed SE analysis once significant heterogeneity was detected in our meta-analysis. Moreover, Deeks' test was performed to judge the reporting bias, and a $p < 0.05$ indicated symmetry and a high likelihood in this meta-analysis. All calculations were performed using Midas modules in the STATA V.14.0 statistical software.

Table 1 Main characteristics of pSWE studies in the meta-analysis

Authors (year)	Country	n	Age	Gender, male (%)	Clinical setting	Funding	Histological score	Cut-off values (m/s)	SE	SP	AUROC	Technology failure
Attia et al ²⁴ (2016)	Germany	61	50±12	31 (51)	Single-centre	Yes	Kleiner	F≥2 1.18 F≥3 1.45 F=4 1.95	0.81 1.00 0.86	0.86 0.96 1.00	0.88 0.99 0.97	1/61
Yoneda et al ²⁰ (2010)	Japan	54	50.6±13.7	26 (46)	Single-centre	No	Brunt	F≥3 1.77 F=4 1.90	1.00 1.00	0.91 0.96	0.97 0.98	NR
Cassinotto et al ²³ (2016)	France	291	56.7±12	172 (59)	Multicentre	No	Kleiner	F≥2 1.32 F≥3 1.53 F=4 2.04	0.56 0.59 0.44	0.91 0.90 0.90	0.77 0.84 0.84	2/291
Osaki et al ²² (2010)	Japan	23	NR	11 (48)	Single-centre	No	Brunt	F≥3 1.47	1.00	0.75	0.94	NR
Palmeri et al ³³ (2011)	USA	172	NR	65 (38)	Single-centre	Yes	Kleiner	F≥3	0.90	0.90	0.90	NR
Cui et al ²² (2016)	USA	114	48.9±15.4	57 (50)	Single-centre	Yes	Kleiner	F≥2 1.29 F≥3 1.34 F=4 2.48	0.82 0.95 0.78	0.78 0.74 0.93	0.85 0.95 0.86	3/114
Fierbinteanu Braticevici et al ³¹ (2013)	Romania	64	51	28 (43)	Single-centre	No	Kleiner	F≥2 1.16 F≥3 1.48 F=4 1.64	0.85 0.86 0.92	0.90 0.95 0.92	0.94 0.98 0.98	0/64
Zhang Da-kun and Yang ³⁵ (2014)	China	67	34.7±13.2	46 (69)	Single-centre	No	Brunt	F≥2 1.30 F≥3 1.36 F=4 1.40	0.61 0.69 1.00	0.81 0.82 0.75	0.71 0.77 0.85	0/67
Li et al ¹⁷ (2016)	China	136	35.8±2.6	88 (65)	Single-centre	Yes	Brunt	F≥2 1.30 F≥3 1.36 F=4 1.41	0.61 0.68 1.00	0.80 0.82 0.74	0.72 0.76 0.86	0/136

AUROC, area under the summary receiver operating characteristic curve; NR, not reported; pSWE, point shear wave elastography; SE, sensitivity; SP, specificity.

Table 2 Main characteristics of TE studies in the meta-analysis

Authors (year)	Country	n	Age	Gender, male (%)	Clinical setting	Funding	Histological score	Cut-off values (m/s)	SE	SP	AUROC	Technology failure
Attia et al ²⁴ (2016)	Germany	61	50±12	31 (51)	Single-centre	Yes	Kleiner	F≥2 7.0 F≥3 11.8 F=4 15	0.85 0.79 1.00	0.80 0.94 0.93	0.88 0.88 0.97	4/61
Yoneda et al ¹⁶ (2008)	Japan	54	50.6±13.7	26 (46)	Single-centre	No	Brunt	F≥3 9.9 F=4 16	1.00 1.00	0.93 0.98	0.99 0.99	NR
Cassinotto et al ²³ (2016)	France	291	56.7±12	172 (59)	Multicentre	No	Kleiner	F≥2 9.8 F≥3 12.5 F=4 16.1	0.60 0.57 0.65	1.00 0.90 0.90	0.82 0.86 0.87	42/291
Gaia et al ²⁹ (2011)	Italy	290	48	52 (72)	Single-centre	No	Brunt	F≥2 7.0 F≥3 8.0 F=4 10.5	0.76 0.65 0.78	0.80 0.80 0.96	0.80 0.76 0.94	21/290
Yoneda et al ²⁰ (2010)	Japan	102	51.8±13.7	40 (39)	Multicentre	Yes	Brunt	F≥2 6.7 F≥3 9.8 F=4 17.5	0.88 0.85 1.00	0.74 0.81 0.97	0.87 0.90 0.99	5/102
Lupsor et al ²⁵ (2010)	Romania	72	42	51 (71)	Single-centre	Yes	Brunt	F≥2 6.8 F≥3 10.4	0.67 1.00	0.84 0.97	0.79 0.98	7/72
Kumar et al ²⁶ (2013)	India	141	39.1	90 (64)	Single-centre	No	Kleiner	F≥2 7.7 F≥3 9.0 F=4 11.8	0.94 0.85 0.97	0.61 0.88 0.90	0.85 0.94 0.96	21/141
Myers et al ²⁸ (2010)	Canada	50	NR	34 (68)	Multicentre	Yes	Kleiner	F≥3 10.3 F=4 11.1	0.70 1.00	0.76 0.82	0.84 0.92	NR
Myers et al ²⁷ (2012)	Canada	276	NR	173 (63)	Multicentre	Yes	Kleiner	F=4 16.0	1.00	0.91	0.85	44/276
Wong et al ⁷ (2009)	France/China	274	51±11	135 (55)	Multicentre	Yes	Kleiner	F≥2 7.0 F≥3 9.6 F=4 11.5	0.79 0.75 0.76	0.76 0.92 0.91	0.84 0.93 0.95	28/274
Imajo et al ³⁰ (2016)	Japan	142	57.5±14.6	81 (61)	Single-centre	Yes	Brunt	F≥2 11.0 F≥3 11.4 F=4 14.0	0.65 0.86 1.00	0.89 0.84 0.76	0.82 0.88 0.92	15/142

AUROC, area under summary receiver operating characteristic curve; NR, not reported; SE, sensitivity; SP, specificity; TE, transient elastography.

Table 3 Risk of bias and concerns regarding the applicability of pSWE studies included in the analysis (QUADAS-2 criteria)

Study	Risk of bias				Applicability concerns		
	Patient selection	Index text	Reference standard	Flow and timing	Patient selection	Index text	Reference standard
Yoneda <i>et al</i> ²⁰ (2010)	Low	Unclear	Low	High	Low	Low	High
Attia <i>et al</i> ²⁴ (2016)	High	Unclear	Low	High	Low	Low	High
Cassinotto <i>et al</i> ²³ (2016)	High	Low	Low	Low	Low	Low	Unclear
Osaki <i>et al</i> ³² (2010)	High	Unclear	High	High	Low	Unclear	High
Palmeri <i>et al</i> ³³ (2011)	High	Low	High	High	Low	Unclear	High
Cui <i>et al</i> ²² (2016)	High	Low	High	Low	Unclear	Low	High
Fierbinteanu Braticevici <i>et al</i> ³¹ (2013)	Unclear	Low	Low	Low	Low	Low	High
Zhang Da-kun and Yang ³⁵ (2014)	Low	Unclear	Low	Low	Low	Low	Unclear
Li Yudan ³⁴ (2017)	Low	Unclear	Low	Low	Low	Low	Unclear

pSWE, point shear wave elastography; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies-2.

RESULTS

Included studies

We initially identified 1302 potential articles up to 20 December 2017. After removal of 355 duplicates, 947 articles remained for title screening, and 818 of these articles were excluded for the following reasons: not related to pSWE, not related to TE, reviews, paediatric or animal research, and the inclusion of various hepatic diseases other than NAFLD. Due to the restricted number of studies of new XL probe and the heterogeneity probably caused by two different probes (XL probe producing lower value most of the time than M probe), TE studies about XL probe were excluded. Another 109 articles were excluded for the following reasons: inability to extract 2×2 contingency tables or lack of histology as the gold standard. A total of 20 studies were obtained in the final search. Eleven studies met the inclusion criteria

for TE,^{7 16 20 23–30} and nine studies met the criteria for pSWE.^{20 22–24 31–35} The flow chart of study selection is presented in [figure 1](#). The list of 100 excluded articles is presented in online supplementary appendix 2.

Study characteristics and quality assessment

Nine pSWE studies comprising a total of 982 patients and eleven TE studies comprising a total of 1735 patients were included. The main characteristics of the pSWE studies and TE studies are presented in [tables 1 and 2](#). The outcomes of the quality assessments of the 20 studies are detailed in [tables 3 and 4](#).

Technology failure ratio

In this study, we also calculated the technology failure ratio to judge the feasibility of two elastographic tests. For TE, the percentage of failed measurements was 11.3%

Table 4 Risk of bias and concerns regarding the applicability of the TE studies included in the analysis (QUADAS-2 criteria)

Study	Risk of bias				Applicability concerns		
	Patient selection	Index text	Reference standard	Flow and timing	Patient selection	Index text	Reference standard
Yoneda <i>et al</i> ²⁰ (2010)	Low	Unclear	Low	High	Low	Low	High
Attia <i>et al</i> ²⁴ (2016)	High	Unclear	Unclear	High	Low	Low	High
Cassinotto <i>et al</i> ²³ (2016)	High	Low	Low	Low	Low	Low	High
Gaia <i>et al</i> ²⁹ (2011)	Unclear	High	High	Low	Low	Low	High
Yoneda <i>et al</i> ¹⁶ (2008)	High	Low	Low	High	Low	Low	High
Lupsor <i>et al</i> ²⁵ (2010)	High	Low	Low	High	Low	Unclear	High
Kumar <i>et al</i> ²⁶ (2013)	Unclear	Low	High	Unclear	Low	Unclear	High
Myers <i>et al</i> ²⁸ (2010)	High	Low	High	Low	Low	Low	High
Myers <i>et al</i> ²⁷ (2012)	High	Low	High	Low	High	Low	High
Wong <i>et al</i> ⁷ (2010)	High	Low	Low	Unclear	Low	Low	High
Imajo <i>et al</i> ³⁰ (2016)	Low	Low	Low	Low	Low	Low	High

QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies-2; TE, transient elastography.

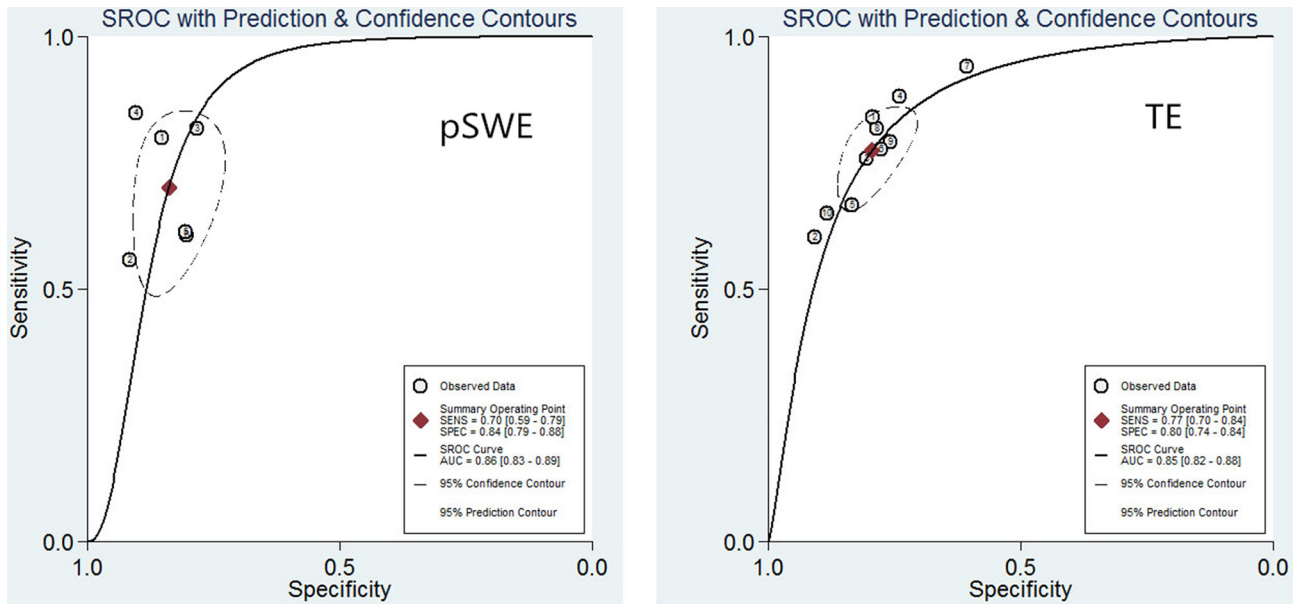


Figure 2 Summary receiver operating characteristic (SROC) curves of the performance of point shear wave elastography (pSWE) and transient elastography (TE) in the diagnosis of significant fibrosis. The area demarcated by the dotted line represents 95% CI and prediction interval. AUC, area under the summary receiver operating characteristic curve; SENS, sensitivity; SPEC, specificity.

(187/1649) as reported in nine studies, which meant that it was impossible to obtain any valid data. For pSWE, the percentage of failed measurements was 0.8% (6/733) as mentioned in six studies.

Diagnostic accuracy

Forest plots of the data from 20 studies in terms of SE, SP, LR+ and LR- with their CIs and heterogeneities and summary receiver operating characteristic curves for the pSWE and TE assays in the diagnosis of NAFLD are

presented below. (The detailed information of forest plots is shown in online supplementary figures 1–6).

Diagnosis of significant fibrosis

The outcomes employed to appraise the diagnostic accuracy of pSWE for significant fibrosis ($F \geq 2$) came from six studies. The summary SE was 0.70 (95% CI 0.59 to 0.79), the summary SP was 0.84 (95% CI 0.79 to 0.88), the summary LR+ was 4.34 (95% CI 3.26 to 5.77), the summary LR- was 0.36 (95% CI 0.26 to 0.49) and the

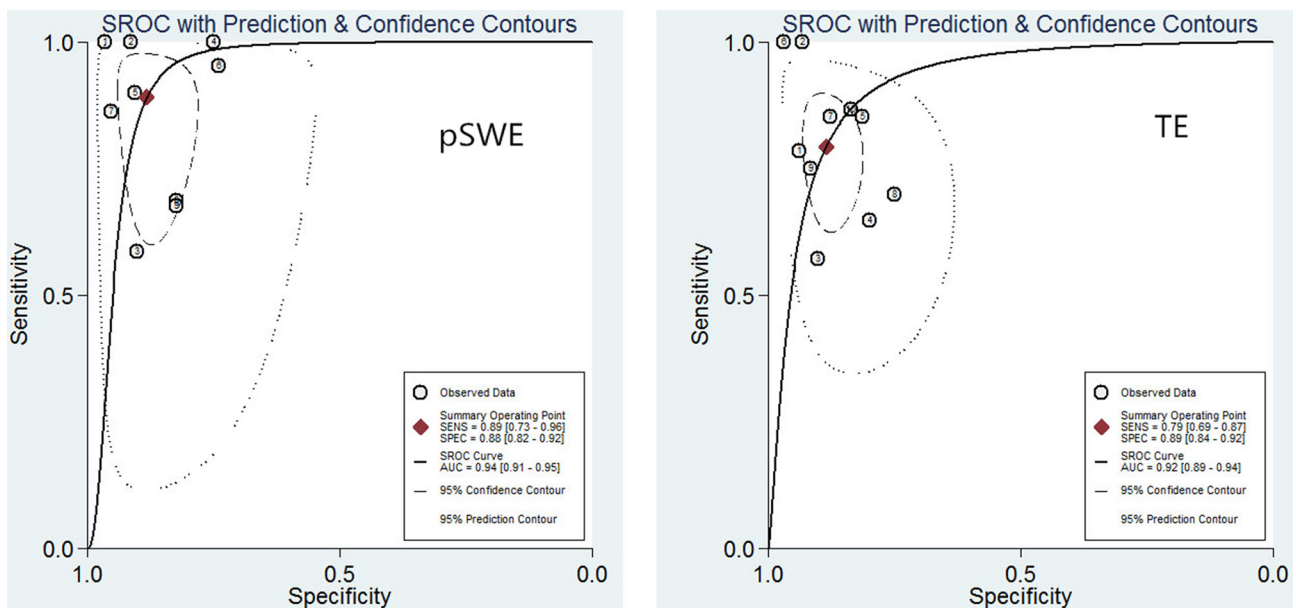


Figure 3 Summary receiver operating characteristic (SROC) curves of the performance of point shear wave elastography (pSWE) and transient elastography (TE) in the diagnosis of advanced fibrosis. The area demarcated by the dotted line represents 95% CI and prediction interval. AUC, area under the summary receiver operating characteristic curve; SENS, sensitivity; SPEC, specificity.

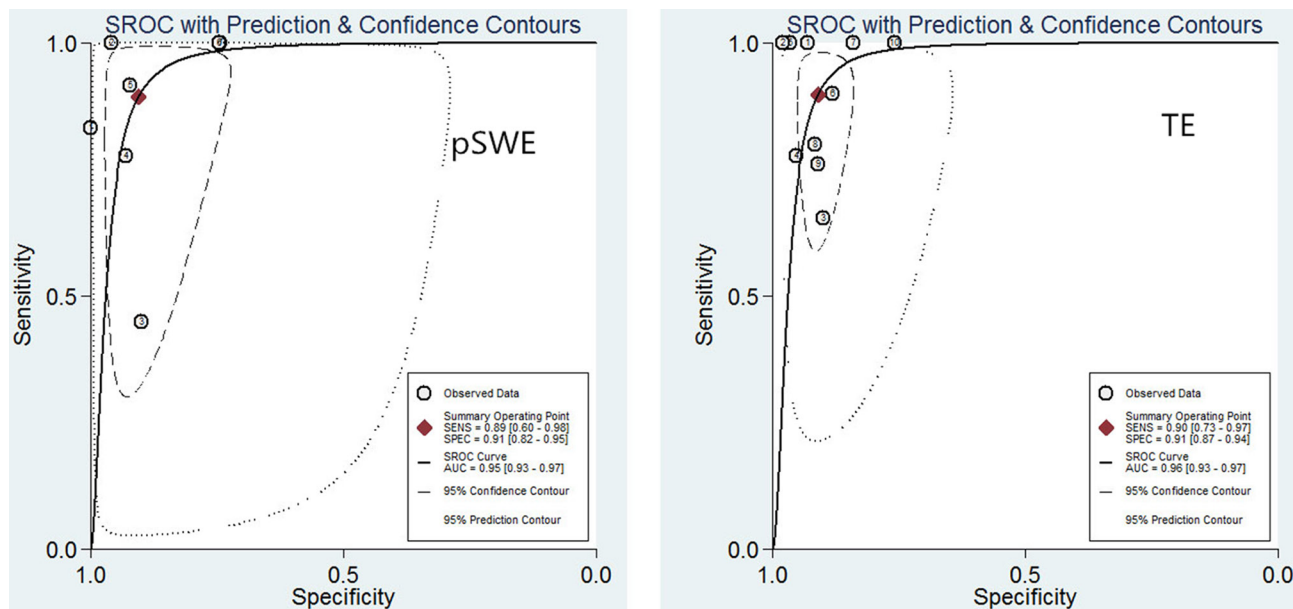


Figure 4 Summary receiver operating characteristic (SROC) curves of the performance of point shear wave elastography (pSWE) and transient elastography (TE) in the diagnosis of cirrhosis. The area demarcated by the dotted line represents 95% CI and prediction interval. AUC, area under the summary receiver operating characteristic curve; SENS, sensitivity; SPEC, specificity.

corresponding AUC was 0.86 (95% CI 0.83 to 0.89). Regarding TE, data were collected from 10 studies. The summary SE was 0.77 (95% CI 0.70 to 0.84), the summary SP was 0.80 (95% CI 0.74 to 0.84), the summary LR+ was 3.78 (95% CI 3.05 to 4.67), the summary LR- was 0.28 (95% CI 0.22 to 0.27) and the corresponding AUC was 0.85 (95% CI 0.82 to 0.88) (figure 2).

Diagnosis of advanced fibrosis

The outcomes employed to appraise the diagnostic accuracy of pSWE for advanced fibrosis ($F \geq 3$) came from nine studies. The summary SE was 0.89 (95% CI 0.73 to 0.96), the summary SP was 0.88 (95% CI 0.82 to 0.92), the summary LR+ was 7.58 (95% CI 4.85 to 11.86), the summary LR- was 0.12 (95% CI 0.05 to 0.33) and the corresponding AUC was 0.94 (95% CI 0.91 to 0.95). Regarding TE, data were collected from 11 studies. The summary SE was 0.79 (95% CI 0.69 to 0.87), the summary SP was 0.89 (95% CI 0.84 to 0.92), the summary LR+ was 6.92 (95% CI 4.78 to 10.01), the summary LR- was 0.23 (95% CI 0.15 to 0.36) and the corresponding AUC was 0.92 (95% CI 0.89 to 0.94) (figure 3).

Diagnosis of cirrhosis

The outcomes employed to appraise the diagnostic accuracy of pSWE for cirrhosis ($F=4$) came from seven studies. The summary SE was 0.89 (95% CI 0.60 to 0.98), the summary SP was 0.91 (95% CI 0.82 to 0.95), the summary LR+ was 9.50 (95% CI 4.89 to 18.74), the LR- was 0.12 (95% CI 0.03 to 0.53) and the corresponding AUC was 0.95 (95% CI 0.93 to 0.97). Regarding TE, the data from 11 studies were collected. The summary SE was 0.90 (95% CI 0.73 to 0.97), the summary SP was 0.91 (95% CI 0.87 to 0.94), the summary LR+ was 9.98 (95% CI 6.68 to 14.91),

the summary LR- was 0.11 (95% CI 0.04 to 0.33) and the corresponding AUC was 0.96 (95% CI 0.93 to 0.97) (figure 4).

SE analysis

Due to the observable heterogeneity observed in the overall pooled parameters, we implemented an SE analysis to explore the source of heterogeneity by removing each study individually and then calculating the pooled statistics again. The SE analysis revealed that the study by Cassinotto *et al*²³ might be the source of the statistical heterogeneity in this meta-analysis. After this outlier was removed, the I^2 value for heterogeneity was conspicuously decreased with regard to the SE, SP, LR+ and LR- in the different stages of fibrosis.

The calculations after removing the outlier are presented in table 5.

Publication bias

A Deeks' test was conducted to assess the final set of studies for possible reporting bias. Significant publication biases were observed when pSWE was performed to detect cirrhosis and when TE was used to detect cirrhosis ($p \leq 0.05$), which suggests asymmetry in the data and a high likelihood of publication bias (figures 5–6).

DISCUSSION

Even though pSWE has been considered as a refined name in the new guideline from the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB), there were still numerous studies that used Acoustic Radiation Force Impulse (ARFI) to identify pSWE in the past; thus, we conducted literature search

Table 5 Diagnostic performance with heterogeneity of pSWE and TE compared with the outlier excluded

Pooled estimates	pSWE F2				pSWE F3				pSWE F4				TE F2				TE F3				TE F4				
	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	Overall	Outlier excluded	
	6	5	9	8	7	6	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10
Studies (n)	6	5	9	8	7	6	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10	9	10
SE (I ²)	0.70 (76.37%)	0.74 (64.46%)	0.89 (85.96%)	0.91 (72.73%)	0.89 (77.24%)	0.92 (90.85%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)	0.81 (3.66%)	0.77 (74.93%)
SP (I ²)	0.84 (39.75%)	0.90 (0.00%)	0.88 (82.75%)	0.88 (79.11%)	0.91 (90.93%)	0.91 (90.85%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)	0.77 (0.00%)	0.80 (42.72%)
LR+ (I ²)	4.34 (0.00%)	3.99 (0.00%)	7.58 (70.26%)	7.65 (69.44%)	9.5 (75.95%)	10.74 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)	3.46 (0.00%)	3.78 (0.00%)
LR- (I ²)	0.36 (65.18%)	0.32 (67.72%)	0.12 (86.90%)	0.10 (75.65%)	0.12 (80.71%)	0.09 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)	0.25 (0.00%)	0.28 (47.94%)
AUC	0.86	0.84	0.94	0.95	0.95	0.97	0.85	0.85	0.85	0.85	0.97	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

AUC, area under the summary receiver operating characteristic curve; LR+, positive likelihood ratio; LR-, negative likelihood ratio; pSWE, point shear wave elastography; SE, sensitivity; SP, specificity; TE, transient elastography.

using the terms ‘point shear-wave elastography’ and ‘acoustic force impulse elastography’. In our meta-analysis, 9 studies that were all performed by VTQ (Siemens) evaluated pSWE and 11 studies evaluated TE, and all of these studies fulfilled the criterion of including NAFLD. First, we compared the technology failure ratios between pSWE and TE. The pooled analysis revealed that the proportion of failed measurements was more than tenfold greater with TE using the M probe compared with pSWE, and this difference is much greater than previous descriptions.¹⁹ This result might be explained by the difficulty in the recognition of liver damage in overweight and obese patients. Several recently published articles have proven that the proportion of successful measurements with TE decreases remarkably when it is performed on overweight and obese patients (body mass index >25 kg/m²),²⁷⁻³⁶ whereas obesity seems to have less influence on pSWE. However, notably, most patients with NAFLD have high body mass index and subcutaneous fat thickness. To resolve this issue, the manufacturer of the TE device developed a new XL probe to cater to obese patients. The use of a larger probe and lower ultrasound frequency of 2.5 MHz enables the examination of deeper liver tissues at 35–75 mm from the skin surface and reduces the failure ratio.³⁷⁻³⁹ Therefore, it would be worth comparing the performance of TE using an XL probe and pSWE if sufficient valid data are available.

Second, our meta-analysis calculated the pooled estimates of SE, SP, LR+, LR- and summary receiver operating characteristic curve for pSWE and TE in different stages of liver fibrosis using a bivariate effect model. The outcome suggested that both pSWE and TE provide excellent diagnostic accuracies in staging advanced fibrosis and cirrhosis, with summary AUCs reaching 0.94 and 0.95 for ≥F3 and =F4, respectively, on pSWE, and reaching 0.92 and 0.94 for ≥F3 and =F4 on TE, respectively. Additionally, these techniques provide good diagnostic accuracies in staging significant fibrosis, with an AUC of 0.86 for ≥F2 on pSWE and AUC of 0.85 for ≥F2 on TE. These results accord with, or even slightly exceed, the acknowledged information that has previously been published.¹¹⁻¹⁸ Moreover, after pooling the SE and SP of the two elastographic methods, we found that these two parameters were high for the screening of advanced fibrosis (≥F3) and cirrhosis (F=4). We also pooled the likelihood ratios, which was particularly useful because this method enables examination of the incremental values of the diagnostic test.⁴⁰ Under these circumstances, if the LR+ is more than 10, we may conclude that the gold standard is absolutely positive when the diagnostic methods yield positive results. If the LR- is less than 0.1, we can regard the subject as a healthy subject without hesitation when the diagnostic methods yield negative results. Our outcomes demonstrate that the two methods exhibited the best performance in the staging of cirrhosis, that is, the phase that predisposes patients to liver failure and hepatocellular carcinoma, with an LR+ of 10.74 on pSWE and 10.47 on TE after removing the outlier. Based on

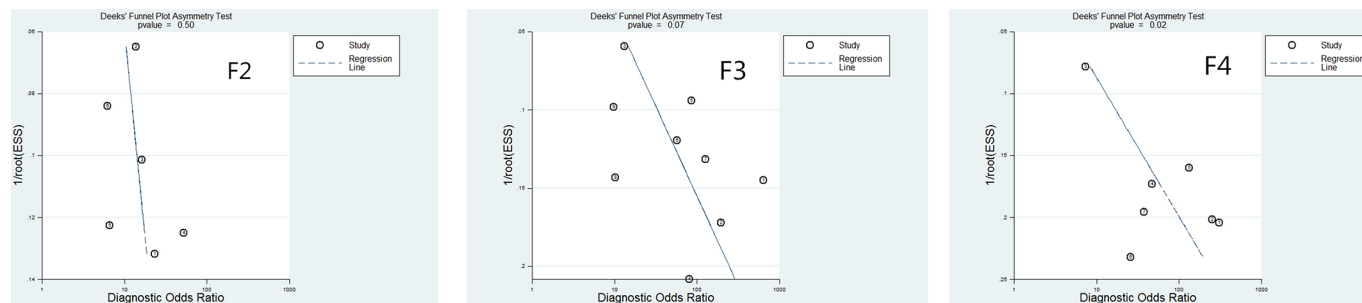


Figure 5 Result of a Deeks' test to assess the publication bias of point shear wave elastography in the diagnosis of significant fibrosis, advanced fibrosis and cirrhosis. A non-zero slope coefficient is suggestive of a significant small study bias ($p < 0.05$).

the above discussion, the conclusions that the two tests provide accurate non-invasive staging of hepatic fibrosis in NAFLD and that both have higher diagnostic accuracies in the ratings of advanced fibrosis ($\geq F3$) and cirrhosis ($F=4$) than in significant fibrosis ($F \geq 2$) can safely be drawn. These observations are similar to those of previously published studies.^{19 20} Moreover, compared with the previous research, our outcomes seem to suggest that there is no apparent distinction between NAFLD, viral hepatitis and other chronic liver diseases in terms of the diagnostic performance of the staging of hepatic fibrosis.¹⁹

Additionally, the SE analysis in our meta-analysis was conducted to explore the potential source of heterogeneity, and revealed that the study by Cassinotto *et al*²³ was probably the source of the statistical heterogeneity. The reason for this heterogeneity might be that these authors did not adopt the optimal cut-off values, which might have led to maximising the SE at the cost of reducing the SP or vice versa in this study.

Despite the prospective outcome, there are several limitations to our study that need to be addressed. First, very different cut-off values were used in the included studies, and our study fails to make a comparison of pSWE and TE due to the absence of sufficient data from studies that performed pSWE and TE simultaneously on the same patient population (only three such studies were involved in this meta-analysis^{20 23 24}). We therefore did not perform a statistical comparison of the pooled parameters to determine which one is the better method; instead, we describe their diagnostic performance separately. The inclusion of additional research

that directly compares pSWE and TE in the detection of NAFLD is needed. Second, although some scholars have long claimed to believe that liver stiffness as measured by elastographic techniques seems to be affected by steatosis, inflammation and fibrosis, all of which are histological features of NAFLD,^{41 42} two studies in 2016 reported that there is no conclusive evidence demonstrating that steatosis or steatohepatitis independently alters pSWE or TE measurements.^{24 43} Therefore, we believe that more incontrovertible meta-analyses aimed at detecting the correlation between elastographic methods and NAFLD scores are essential. Third, significant publication bias was identified in our study, which indicates the possibilities of the presence of a significant reporting bias. Missing data like incomplete technology failure may limit the persuasiveness of the results. Fourth, some heterogeneity was still present after exclusion of the outlier, which probably was caused by the quality of histology (interobserver variability; centre effects; sampling error), and research based on the role of funding and on study quality might improve heterogeneity slightly. In addition, that we did not write a study protocol before conducting this meta-analysis is a shortcoming of our study. Despite the existence of the limitations mentioned above, our study has several notable strengths that can be summarised as follows: (1) this is the first meta-analysis to assess the diagnostic performance of pSWE for conditions ranging from significant fibrosis to cirrhosis in patients with NAFLD; (2) we performed a comparison that evaluated the technology failure ratios of the two tests, and a higher rate was found for TE; and (3) because of the greater number of available samples

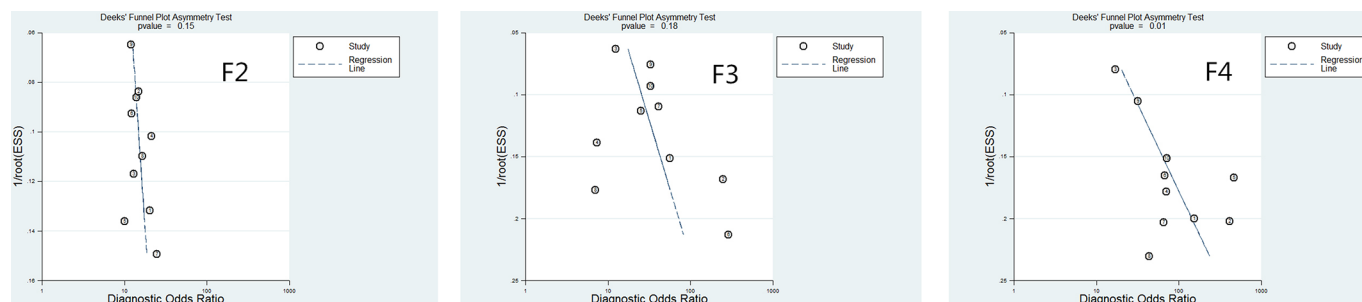


Figure 6 Result of a Deeks' test to assess the publication bias of transient elastography in the diagnosis of significant fibrosis, advanced fibrosis and cirrhosis. A non-zero slope coefficient is suggestive of a significant small study bias ($p < 0.05$).

included in this meta-analysis, our results appear to be convincing and credible.

In conclusion, both pSWE and TE are feasible imaging techniques that enable the non-invasive staging of liver fibrosis in patients with NAFLD, particularly those with advanced fibrosis and cirrhosis. pSWE appears to perform more stably and with higher mission success rate than TE due to the effect of subcutaneous fat thickness and high body mass index on TE. In the future, additional large clinical studies remain to be performed to enable further discussions of the direct comparison of pSWE and TE in terms of diagnostic efficiency in NAFLD.

Contributors HR is the guarantor of the integrity of the entire study. HR carried out the study concept and study design. SH and HT performed the literature research. PW, MW and XZ performed the data acquisition. WJ performed the manuscript preparation, data analysis/interpretation, statistical analysis and manuscript editing. HR performed the manuscript revision/review.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Extra data can be accessed via the Dryad Data Repository at <http://datadryad.org/> with doi:10.5061/dryad.b7f1j3.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Williams CD, Stengel J, Asike MI, *et al.* Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. *Gastroenterology* 2011;140:124–31.
- Wong VW, Chu WC, Wong GL, *et al.* Prevalence of non-alcoholic fatty liver disease and advanced fibrosis in Hong Kong Chinese: a population study using proton-magnetic resonance spectroscopy and transient elastography. *Gut* 2012;61:409–15.
- Watanabe S, Hashimoto E, Ikejima K, *et al.* Evidence-based clinical practice guidelines for nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *J Gastroenterol* 2015;50:364–77.
- Marchesini G, Brizi M, Bianchi G, *et al.* Nonalcoholic fatty liver disease: a feature of the metabolic syndrome. *Diabetes* 2001;50:1844–50.
- Rinella ME. Nonalcoholic fatty liver disease: a systematic review. *JAMA* 2015;313:2263–73.
- Yasui K, Hashimoto E, Komorizono Y, *et al.* Characteristics of patients with nonalcoholic steatohepatitis who develop hepatocellular carcinoma. *Clin Gastroenterol Hepatol* 2011;9:428–33. quiz e50.
- Wong VW, Vergniol J, Wong GL, *et al.* Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology* 2010;51:454–62.
- Bedossa P, Dargère D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;38:1449–57.
- Spengler EK, Loomba R. Recommendations for Diagnosis, Referral for Liver Biopsy, and Treatment of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. *Mayo Clin Proc* 2015;90:1233–46.
- Nobili V, Vizzutti F, Arena U, *et al.* Accuracy and reproducibility of transient elastography for the diagnosis of fibrosis in pediatric nonalcoholic steatohepatitis. *Hepatology* 2008;48:442–8.
- Guo Y, Parthasarathy S, Goyal P, *et al.* Magnetic resonance elastography and acoustic radiation force impulse for staging hepatic fibrosis: a meta-analysis. *Abdom Imaging* 2015;40:818–34.
- Singh S, Venkatesh SK, Wang Z, *et al.* Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of individual participant data. *Clin Gastroenterol Hepatol* 2015;13:440–51.
- Dietrich CF, Bamber J, Berzigotti A, *et al.* EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (Long Version). *Ultraschall Med* 2017;38:e16–e47.
- Koizumi Y, Hirooka M, Kisaka Y, *et al.* Liver fibrosis in patients with chronic hepatitis C: noninvasive diagnosis by means of real-time tissue elastography--establishment of the method for measurement. *Radiology* 2011;258:610–7.
- Righi S, Fiorini E, De Molo C, *et al.* ARFI elastography in patients with chronic autoimmune liver diseases: A preliminary study. *J Ultrasound* 2012;15:226–31.
- Yoneda M, Yoneda M, Mawatari H, *et al.* Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with nonalcoholic fatty liver disease (NAFLD). *Dig Liver Dis* 2008;40:371–8.
- Li Y Y-S, Huang Z-Z, *et al.* Systematic review with meta-analysis: the diagnostic accuracy of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B. *Alimentary pharmacology & therapeutics*. 2016;43:458–69.
- Liu H, Fu J, Hong R, *et al.* Acoustic radiation force impulse elastography for the non-invasive evaluation of hepatic fibrosis in non-alcoholic fatty liver disease patients: a systematic review & meta-analysis. *PLoS One* 2015;10:e0127782.
- Bota S, Herkner H, Sporea I, *et al.* Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. *Liver Int* 2013;33:1138–47.
- Yoneda M, Suzuki K, Kato S, *et al.* Nonalcoholic fatty liver disease: US-based acoustic radiation force impulse elastography. *Radiology* 2010;256:640–7.
- Kwok R, Tse YK, Wong GL, *et al.* Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease--the role of transient elastography and plasma cytochrome-18 fragments. *Aliment Pharmacol Ther* 2014;39:254–69.
- Cui J, Heba E, Hernandez C, *et al.* Magnetic resonance elastography is superior to acoustic radiation force impulse for the Diagnosis of fibrosis in patients with biopsy-proven nonalcoholic fatty liver disease: a prospective study. *Hepatology* 2016;63:453–61.
- Cassinotto C, Boursier J, de Lédinghen V, *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.
- Attia D, Bantel H, Lenzen H, *et al.* Liver stiffness measurement using acoustic radiation force impulse elastography in overweight and obese patients. *Aliment Pharmacol Ther* 2016;44:366–79.
- Lupsor M, Badea R, Stefanescu H, *et al.* Performance of unidimensional transient elastography in staging non-alcoholic steatohepatitis. *J Gastrointest Liver Dis* 2010;19:53–60.
- Kumar R, Rastogi A, Sharma MK, *et al.* Liver stiffness measurements in patients with different stages of nonalcoholic fatty liver disease: diagnostic performance and clinicopathological correlation. *Dig Dis Sci* 2013;58:265–74.
- Myers RP, Pomier-Layrargues G, Kirsch R, *et al.* Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. *Hepatology* 2012;55:199–208.
- Myers RP, Elkashab M, Ma M, *et al.* Transient elastography for the noninvasive assessment of liver fibrosis: a multicentre Canadian study. *Can J Gastroenterol* 2010;24:661–70.
- Gaia S, Carenzi S, Barilli AL, *et al.* Reliability of transient elastography for the detection of fibrosis in non-alcoholic fatty liver disease and chronic viral hepatitis. *J Hepatol* 2011;54:64–71.
- Imajo K, Kessoku T, Honda Y, *et al.* Magnetic resonance imaging more accurately classifies steatosis and fibrosis in patients with nonalcoholic fatty liver disease than transient elastography. *Gastroenterology* 2016;150:626–37.
- Fierbinteanu Braticevici C, Sporea I, Panaitescu E, *et al.* Value of acoustic radiation force impulse imaging elastography for non-invasive evaluation of patients with nonalcoholic fatty liver disease. *Ultrasound Med Biol* 2013;39:1942–50.
- Osaki A, Kubota T, Suda T, *et al.* Shear wave velocity is a useful marker for managing nonalcoholic steatohepatitis. *World J Gastroenterol* 2010;16:2918–25.
- Palmeri ML, Wang MH, Rouze NC, *et al.* Noninvasive evaluation of hepatic fibrosis using acoustic radiation force-based shear stiffness in patients with nonalcoholic fatty liver disease. *J Hepatol* 2011;55:666–72.
- Li Yudan DC. Diagnostic value of acoustic radiation force impulse imaging and APRI ratio index for quantitative for evaluating the

- degree of liver fibrosis in non-alcoholic fatty liver disease patients. *Chin J Med Ultrasound* 2017;544–8.
35. Zhang Da-kun CM, Yang LIU. The diagnostic value of ARFI imaging and APRI for quantitative evaluating the degree of liver fibrosis in NAFLD patients. *J Chin Clin Med Imaging* 2014;186–7.
 36. Wong GL, Wong VW, Chim AM, *et al.* Factors associated with unreliable liver stiffness measurement and its failure with transient elastography in the Chinese population. *J Gastroenterol Hepatol* 2011;26:300–5.
 37. Wong VW, Vergniol J, Wong GL, *et al.* Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. *Am J Gastroenterol* 2012;107:1862–71.
 38. Fraquelli M, Rigamonti C, Casazza G, *et al.* Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut* 2007;56:968–73.
 39. Friedrich-Rust M, Hadji-Hosseini H, Kriener S, *et al.* Transient elastography with a new probe for obese patients for non-invasive staging of non-alcoholic steatohepatitis. *Eur Radiol* 2010;20:2390–6.
 40. Stengel D, Bauwens K, Sehouli J, *et al.* A likelihood ratio approach to meta-analysis of diagnostic studies. *J Med Screen* 2003;10:47–51.
 41. Kleiner DE, Brunt EM, Van Natta M, *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313–21.
 42. Friedrich-Rust M, Romen D, Vermehren J, *et al.* Acoustic radiation force impulse-imaging and transient elastography for non-invasive assessment of liver fibrosis and steatosis in NAFLD. *Eur J Radiol* 2012;81:e325–31.
 43. Harris N, Nadebaum D, Christie M, *et al.* Acoustic radiation force impulse accuracy and the impact of hepatic steatosis on liver fibrosis staging. *J Med Imaging Radiat Oncol* 2016;60:587–92.