

Meta-analysis of the accuracy of transient elastography in measuring liver stiffness to diagnose esophageal varices in cirrhosis

Fan Cheng, MPH^a, Hongyan Cao, PhD^{a,b}, Jinchun Liu, MD, PhD^c, Lijun Jiang, MM^c, Hongjuan Han, MS^{a,b}, Yanbo Zhang, PhD^{a,d,*}, Dongxing Guo, PhD^{a,b,*}

Abstract

Background: To assess the diagnostic performance of transient elastography (TE) in detecting the presence and size of esophageal varices (EV) in cirrhotic patients.

Methods: We searched PubMed, Web of Science, Wiley Online Library, Science Direct, China National Knowledge Infrastructure, WeiPu, WanFang database, and Baidu Scholar to identify studies that evaluated the diagnostic accuracy of TE in liver stiffness measurement, compared with esophagogastroduodenoscopy (EGD), for the detection of the presence and degree of EV in cirrhosis.

Results: We included 32 studies in the presence of any EV (grade 1–3; n = 4082), 27 studies on substantial EV (grade 2–3; n = 5221) and 5 studies on large EV (grade 3). The pooled sensitivity, specificity, and diagnostic odds ratio (DOR) were 0.8 (95% CI, 0.78–0.86), 0.68 (95% CI, 0.62–0.74), and 10 (95% CI, 7–14) for any EV; 0.81 (95% CI, 0.77–0.85), 0.72 (95% CI, 0.66–0.77), and 11 (95% CI, 8–15) for substantial EV; and 0.92 (95% CI, 0.83–0.96), 0.78 (95% CI, 0.70–0.85), and 40 (95% CI, 15–107) for large EV. Subgroup analysis revealed that the heterogeneity among studies on any EV could potentially be explained by study location, proportion of Child A, and time interval between TE and EGD; for substantial EV, the proportion of Child A, etiology of cirrhosis, and the time interval between TE and EGD were important heterogeneity factors. Publication bias was found among studies evaluating diagnostic performance of TE for any EV.

Conclusion: TE is a good tool for detecting the presence and degree of EV; however, in determination of the liver stiffness cutoff values means that TE is only cautiously used in clinical practice.

Abbreviations: AUROC = area under the receiver operating characteristics curve, CI = confidence interval, DOR = diagnostic odds ratio, EGD = esophagogastroduodenoscopy, EV = esophageal varices, FN = false negative, FP = false positive, LSM = liver stiffness measurement, NLR = negative likelihood ratio, PLR = positive likelihood ratio, QUADAS-2 = scores quality assessment of diagnostic accuracy studies-2 scores, TE = transient elastography, TN = true negative, TP = true positive.

Keywords: esophageal varices, liver stiffness, meta-analysis, transient elastography

Editor: Hua Yang.

Compliance with ethical standards.

This work was supported by a grant from the National Natural Science Foundation of China (71403156).

The authors have no conflicts of interest to disclose.

^a Department of Health Statistics, School of Public Health, ^b Department of Mathematics, School of Basic Medicine, Shanxi Medical University, ^c Department of Gastroenterology, The First Affiliated Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, ^d Molecular Imaging Precision Medicine Collaborative Innovation Center, Shanxi Medical University, Taiyuan, Shanxi, China.

* Correspondence: Dongxing Guo, Department of Health Statistics, School of Public Health, Shanxi Medical University and Department of Mathematics, School of Basic Medicine, Shanxi Medical University. No. 56 South Xinjian Road, Taiyuan, Shanxi 030001, China (e-mail: guodongxingmath@163.com), Yanbo Zhang, Department of Health Statistics, School of Public Health, Shanxi Medical University. No. 56 South Xinjian Road, Taiyuan, Shanxi 030001, China and Shanxi Medical University Molecular Imaging Precision Medicine, Collaborative Innovation Center, Taiyuan, Shanxi 030001, China (e-mail: sxmuzbyb@126.com).

Copyright © 2018 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medicine (2018) 97:28(e11368)

Received: 29 November 2017 / Accepted: 11 June 2018

<http://dx.doi.org/10.1097/MD.00000000000011368>

1. Introduction

Esophageal varices (EV) develops in nearly 50% of cirrhotic patients.^[1] Bleeding occurs in almost 25% of cirrhotic patients at 2 years,^[2] and the mortality risks from a first bleeding event is 20% to 35%.^[3] Thus, early diagnosis of cirrhosis is essential in preventing the development of disease and determining the immediate treatment. Esophagogastroduodenoscopy (EGD) is currently regarded as the gold standard for estimating the presence and degree of the EV in cirrhotic patients.^[1] However, EGD is invasive and probably increases the morbidity and mortality rate of complications.^[4] Other limitations, including patient discomfort and high medical costs, also cause a decline in patient compliance. The development of a noninvasive tool for evaluating the presence and extent of EV would be greatly helpful for the diagnosis and treatment of cirrhotic patients with EV.

Studies have suggested that liver stiffness values are potentially associated with EV.^[5] Liver stiffness can be measured by transient elastography (TE), which is a noninvasive method used to predict the presence and extent of EV in cirrhotic patients. TE provides high diagnostic accuracy for detecting the presence and severity of EV, and overcomes the limitations associated with therapeutic burden and patient discomfort.^[6] Recently, TE has been widely introduced into clinical practice; however, the predictive results of different studies have been shown to be inaccurate and unsta-

ble.^[5,7] Hence, this meta-analysis was conducted to systematically and comprehensively evaluate the diagnostic performance of TE for detection of the presence and size of the EV in cirrhotic patients.

2. Methods

2.1. Search strategy

We systematically conducted a literature search in PubMed, Web of Science, Wiley Online Library, Science Direct, China National Knowledge Infrastructure, WeiPu, and WanFang database to identify all studies conducted from January 1, 2006 to May 31, 2017 that evaluated the diagnostic accuracy of TE for the presence and size of EV. We used the medical subject heading search terms including “transient elastography,” “FibroScan,” “TE,” “FS,” “stiffness,” “esophageal varices,” and “cirrhosis.” No other search limitations were made. According to the inclusion and exclusion criteria described below, 2 workers preliminarily screened the titles and abstracts of literature independently to exclude irrelevant articles, and carefully evaluated the full text of the remaining articles. Next, additional articles were identified by manually searching the bibliographies of the key articles in Baidu Scholar. Finally, all eligible studies were included in this meta-analysis. During the retrieval process, any disagreements between the 2 researchers were resolved by discussion with a third author. This study followed the PRISMA guidelines.^[8]

2.2. Selection criteria

The present meta-analysis included all prospective and retrospective full-text articles comparing the diagnostic performance of TE with EGD that satisfied the following inclusion criteria: EGD was used as a gold standard for the identification of EV in cirrhotic patients (grade 0: none; grade 1: small, straight varices; grade 2: medium-sized, enlarged tortuous varices occupying less than one-third of the lumen; grade 3: large-sized, coil-shaped varices occupying more than one-third of the lumen);^[9] liver stiffness was measured by TE to evaluate EV in cirrhotic patients older than 18 years, and was confirmed by liver biopsy or other clinical, biochemical, or imaging methods; sufficient information was available to precisely calculate true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values of TE for the diagnosis of targeted disease; an optimal cutoff value was selected to maximize the sensitivity and specificity. There were no language restrictions imposed, but the included studies had sample sizes of at least 30.

We excluded studies that included patients younger than 18 years, patients with ascites, obesity, comorbid HIV, or liver transplantation that may have affected the TE result, and studies that provided unnecessary data such that no one 2×2 table could be created. When considering subsequent publications that were extensions of previously published cohorts, we only included the studies with the larger number of patients. For a number of studies with the same cohort but different results, we took the full approach to avoid missing important findings.

2.3. Data extraction

Using a data table formulated in advance, specific information was extracted independently by 2 researchers, including: basic characteristics: the first author, publication year, location of the study, study time period, and study design; patient characteristics: sample size, number of males, mean/medium age, body mass index, proportion of Child A, and etiology of cirrhosis; information regarding the index test (and reference standard): cutoff value,

sensitivity, specificity (TP, FP, TN, and FN values could be computed), area under the receiver operating characteristics curve (AUROC), the presence or absence of blinding, and time interval between TE and EGD; information on study outcomes: EV grades, prevalence rates of EV, and Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) scores.

2.4. Quality assessment

The methodological quality of studies was evaluated by 2 researchers using the QUADAS-2 questionnaire, which is a revised 14-item tool designed to evaluate the internal and external validity of diagnostic accuracy studies included in meta-analyses.^[10] Each term was judged as “yes” (one point) if it was reported, as “no” (zero points) if not reported, or as “unclear” (zero points) if the available information was not enough to allow a conclusion to be made. High-quality studies were defined as those with scores of 10 points or more. Disagreements between the 2 researchers were settled by discussion, with the final judgment made by a third researcher.

2.5. Outcome measures

The primary assessment was the diagnostic accuracy of TE in liver stiffness measurement (LSM) performed to detect the presence of EV (grade 1–3) and substantial EV (grade 2–3), compared with the gold standard of EGD. Considering potential heterogeneity among the studies, we conducted preplanned subgroup analyses based on 6 different variables to explore the sources of heterogeneity. The detection of large EV (grade 3) was the secondary assessment; we only reported the association between LSM detected by TE and large EV, without exploring the sources of heterogeneity.

2.6. Statistical analysis

The meta-analysis of diagnostic test accuracy was achieved by establishing a bivariate mixed effect model. All statistical analyses were performed with the Midas command in Stata 14.0 (StataCorp, College Station, TX).

2.6.1. Indicators of diagnostic accuracy. Sensitivity (the proportion of those with the disease who have TP results) and specificity (the proportion of those without the disease who have TN results) are pooled summary statistics used to describe the accuracy of diagnostic tests. We calculated the positive likelihood ratio (PLR) as the sensitivity divided by the FP rate, and the negative likelihood ratio (NLR) as the FN rate divided by the specificity. We also calculated the diagnostic odds ratio (DOR), which is a comprehensive evaluation indicators defined as the ratio of the PLR to the NLR; the DOR indicates how much greater the possibility of contracting the disease is in subjects with a positive test result compared with subjects with a negative test result. We also plotted a summary ROC curve by considering the derived estimates of sensitivity, specificity, and respective variances, and the AUROC was regarded as an optimal measure of the diagnostic test to avoid the influence of different cutoff values. In addition, we assumed pretest probabilities of 25%, 50%, and 75%, and estimated the post-test probability with a positive or negative test result using Fagan plots.^[11]

2.6.2. Analysis of heterogeneity. The between-study heterogeneity of the pooled sensitivity and specificity was evaluated initially by Forest plots, then by statistical methods including the *Q* test and

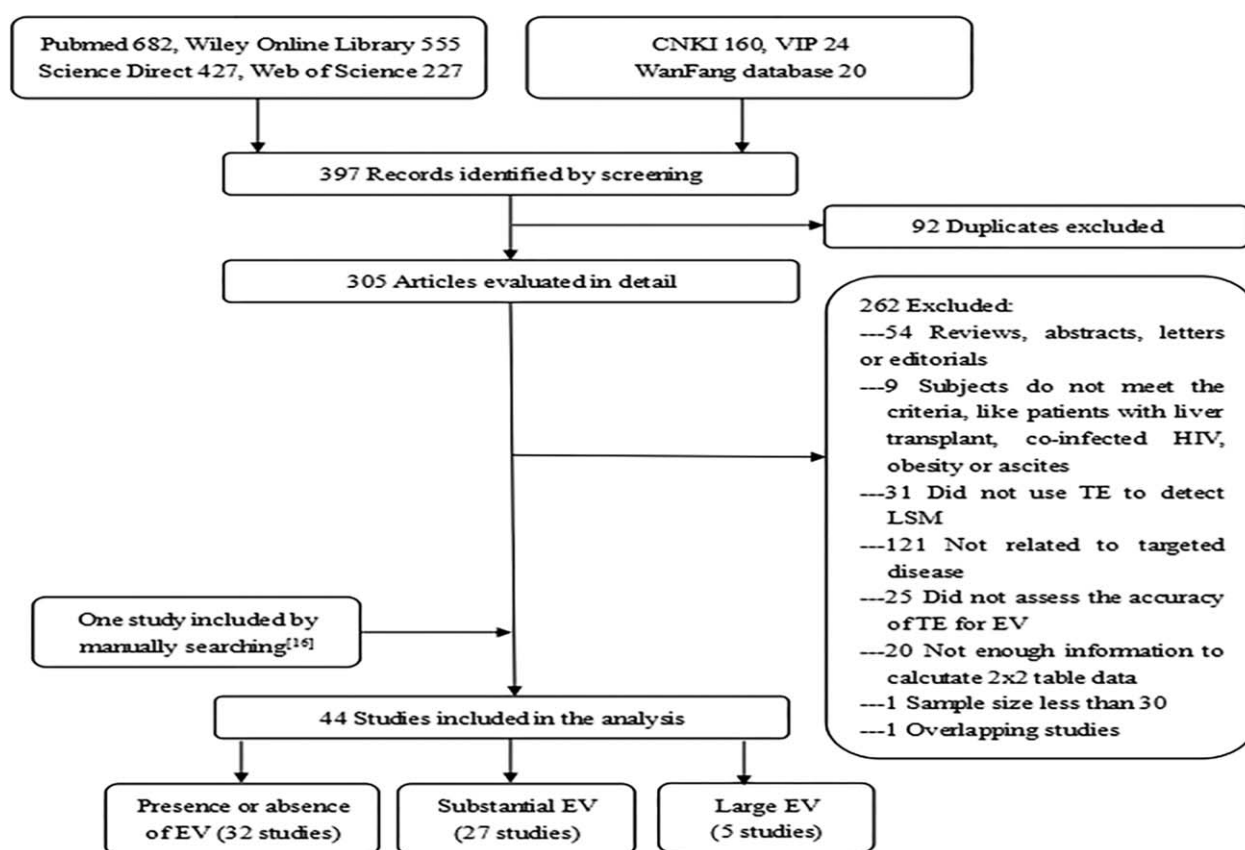


Figure 1. Flowchart of literature search and selection.

I^2 statistic. A P value of $<.10$ or an I^2 value of $>50\%$ was taken as an indicator of significant heterogeneity. If there was no threshold effect estimated by the Spearman correlation coefficient, sources of heterogeneity were explored using subgroup analysis based on the predefined characteristics of all studies.^[12,13]

2.6.3. Sensitivity analysis and publication bias. We also performed a sensitivity analysis to control for heterogeneity by excluding each of the studies one by one or excluding all studies with a methodological quality score below 10 points to verify the robustness of the diagnostic test accuracy.^[14] The publication bias was evaluated using the asymmetry test of the Deeks funnel plot, with $P < .10$ suggesting substantial asymmetry and significant publication bias.^[15]

3. Results

3.1. Literature search

A total of 396 full-text articles were identified by screening. After removing duplicates, 305 remaining studies were evaluated in detail. Of these, 54 articles were excluded because they were reviews, abstracts, letters, or others; 152 studies were excluded because they did not relate to TE or EV; 45 studies were excluded because the diagnostic accuracy of TE for EV was not assessed, or there were insufficient information to calculate 2×2 tabular data; 11 studies were excluded due to inappropriate subjects, inadequate quantity, or overlapping records. A final total of 44 studies in which the diagnostic accuracy of TE for the detection of presence and size of EV was assessed were included in this meta-analysis. One included study^[16] was found by manual searching. Thirty-two included studies involved the diagnosis of the

presence or absence of EV, 27 involved substantial EV, and 5 involved large EV (Fig. 1).

3.2. Characteristics and quality assessment of included studies

Liver stiffness cutoff values were 6.1 to 29.7 kPa for any EV, 12.2 to 48.0 kPa for substantial EV, and 17.7 to 34.6 kPa for large EV. Of the included 44 studies, 18 were performed in the European population, 21 in the Asian population (of which 16 were in China), 3 in the African population, and 2 in the North American population. The follow-up period in the earliest study started in 2002, and the study period of the latest study ended in 2016; the earliest study was a prospective study conducted in France, and the latest study was a retrospective study conducted in China. The follow-up duration of one study carried out in Romania was not reported. All articles published from 2006 to 2017 were prospective and retrospective studies. There were 7294 cirrhosis patients in total (mean age 54.26 years). Viral hepatitis was the leading cause of cirrhosis, with an etiology of HBV or HCV in only 13 studies. The etiology of cirrhosis in 2 studies was unknown. Details of the included studies are shown in Tables 1 and 2.

The methodological quality assessment revealed that 3 of the 44 included studies achieved a score of 9 or lower; in general, the quality of all articles was moderate. On assessment of each QUADAS-2 item, 22 studies did not provide enough proof as to whether the TE results were interpreted by assessors blinded to the EGD results, while 18 studies did not state whether the EGD results were interpreted by assessors blinded to the TE results. Seven studies did not describe the gold standard and index tests in detail. The time interval between detection of LSM by TE and performance of EGD was not clearly noted in 23 studies. Seven

Table 1
Basic information of included studies assessing the diagnostic accuracy of transient elastography.

Author, Year, Location	Time intervals of study	N (male)	Medium/Mean age, years	BMI, kg/m ²	Child A (%)	Etiology	Design	Blind	Time interval EGD and TE	Grade 1–3/2–3/3 prevalence (%)	QUADAS scores
Kazemi, ^[17] 2006, France	2002.11–2004.6	165 (111)	56	24.75	NA	HCV/59.4%, ALD, HBV, other	Retrospective	Y	The same day	44.9/28.5/	14
Foucher, ^[18] 2006, France	2003.6–2004.9	144 (102)	53	25.2	34.0	HBV, HCV, ALD, other	Prospective	U	Unclear	759.0/	9
Vizzutti, ^[19] 2007, Italy	2005.3–2006.7	47 (30)	57.8	22.0	59.6	HCV	Prospective	U	Within 1 week	65.2/	10
Bureau, ^[20] 2008, France	2005.11–2006.8	89	55	NA	33.7	ALD, VHC, other	Prospective	U	Unclear	71.9/48.3/	11
Jung, ^[21] 2008, Korea	2006.12–2007.11	112 (88)	53.3	NA	NA	ALD/50%, HBV, HCV, other	Prospective	Y	Within 1 week	73.2/81.3/	13
Castera, ^[22] 2009, France	2003.6–2007.4	70 (42)	54.1	25.9	100	HCV	Prospective	Y	Unclear	54.3/18.6/	10
Nguyen-Khac, ^[23] 2010, France	2005–2008	183 (118)	55.2	NA	63.4	ALD/56.3%, HCV, HBV, other	Prospective	Y	Unclear	722.4/	14
Malik, ^[24] 2010, USA	Over 3-year	124 (87)	53	28.7	NA	HCV/70%, other	Prospective	U	Within 3 days	50.8/	10
Meng (unpublished), 2010, China	2009.9–2010.3	61	NA	NA	NA	Mixed	Retrospective	U	Unclear	767.2/	8
Stefanescu, ^[25] 2011, Romania	NA	137 (77)	56	26.36	64.9	HCV, ALD	Prospective	U	Unclear	84.9/	10
Pritchett, ^[26] 2011, Canada	2004.11–2008.7	211 (152)	54.2	28.7	NA	HCV/72.5%, ALD, HBV, other	Prospective	U	Unclear	737.4/	11
Stefanescu, ^[6] 2011, Romania	2009.1–2010.8	231 (135)	55.66	27.3	75.8	VHC/49.8%, ALD, VHC+ALD	Prospective	Y	Within 6 months	68.0/29.4/	12
Sporrea, ^[27] 2011, Romania	NA	1000	NA	NA	NA	NA	NA	Y	Within 3 months	735.3/	13
Chen, ^[28] 2012, China	2007.6–2010.8	222 (187)	42.7	NA	48.6	HBV	Prospective	U	Unclear	256.6/	13
Li, ^[29] 2012, China	2010.1–2011.12	158 (126)	47.1	NA	57.6	HBV	NA	U	Unclear	57.0/25.9/11.4	11
Li, ^[30] 2012, China	2011.4–2012.4	201	50.86	NA	NA	VHC/74.1%, ALD, other	Retrospective	U	Unclear	743.8/	11
Wang, ^[31] 2012, China	2008.11–2011.1	126 (93)	54.5	24.8	100	HBV	Prospective	Y	Unclear	38.1/10.3/	12
Colocchia, ^[32] 2012, Italy	2009.9–2011.1	100 (71)	54	25	68.0	HCV	NA	Y	Within 3 days	53.0/	14
Wang, ^[33] 2012, China	2008.11–2009.1	46 (30)	54	24	NA	HCV	NA	Y	Within 2 weeks	65.2/41.3/	14
Sporrea, ^[9] 2013, Romania	NA	697 (399)	57	NA	NA	HCV/69.6%, HBV, ALD	Retrospective	Y	Within 3 days	739.2/	12
Sharma, ^[34] 2013, India	2011.9–2012.3	174 (154)	49.3	24.6	31.6	VHC/73.6%, ALD	Retrospective	Y	Up to 6 months	73.3/	12
Calvanus, ^[35] 2013, Italy	2008.1–2011.3	96 (67)	63.2	27.0	100	HCV	Prospective	Y	Within 1 week	56.3/27.1/	13
Chaojint, ^[16] 2013, Thailand	2009.1–2009.12	52 (35)	55.7	NA	90.4	ALD/44.3%, HCV, HBV, other	Prospective	Y	The same day	71.6/	14
Saad, ^[7] 2013, Egypt	2011.4–2011.8	32 (20)	51	<35.0	71.9	HBV/42.3%, ALD, HCV, other	Prospective	Y	The same day	726.9/	14
Li, ^[36] 2014, China	2010.1–2011.12	260 (176)	49.4	NA	59.2	HCV	NA	Y	Within days	62.5/31.3/	12
Liu, ^[37] 2014, China	2010.2–2011.2	100 (70)	45	NA	NA	HBV/60.8%, HCV, ALD, other	NA	U	Unclear	59.2/26.5/9.2	11
Saiz, ^[38] 2014, Vienna	2009.1–2010.4	59	57.4	24.8	NA	ALD/72.9%, VHC, ALD+VHC	Prospective	U	Unclear	67.8/	10
Gabriella, ^[39] 2014, Hungary	2009.1–2012.12	74 (29)	60.5	NA	NA	HCV, HBV, PBC, other	NA	Y	After 2 days	47.3/	12
Augustin, ^[40] 2014, Spain	2010.1–2012.4	54 (30)	62	26.5	NA	HCV/89%, other	Prospective	Y	Within 1 month	20.4/	13
Hassan, ^[41] 2014, Egypt	2012.1–2013.2	65 (39)	52.4	NA	80.0	HCV	Prospective	Y	Within 1 month	76.9/49.2/	13
Hu, ^[42] 2015, China	2007.7–2012.8	200 (141)	45.1	NA	NA	HCV	Prospective	Y	The same week	55.0/34.5/	13
Brintan, ^[5] 2015, Romania	2009–2012	60 (39)	57.03	26.99	65.0	HBV/83.5%, HCV	Prospective	U	Unclear	78.3/53.3/	12
Al-Hamoudi, ^[43] 2015, Egypt	2010.1–2013.1	123 (59)	58.8	NA	NA	ALD/55%, HCV, HBV, ALD+HBV	NA	Y	Unclear	57.8/	12
Stefanescu, ^[44] 2015, Romania	2011.11–2013.8	90 (50)	55.73	26.71	62.2	HCV/54.5%, HBV, other	Retrospective	U	Within 6 months	73.7/29.0/16.7	11
Zeng, ^[45] 2016, China	2010.10–2014.7	396 (300)	43.6	NA	NA	HCV	Retrospective	U	Within 2 weeks	73.7/29.0/16.7	11
Wong, ^[46] 2016, China	2013.1–2013.9	144 (114)	58	24.4	NA	HCV	Prospective	Y	Within 4 weeks	21.5/	14
Shibata, ^[47] 2016, Japan	2013.4–2014.11	99 (43)	71	21.8	NA	HCV	Retrospective	Y	The same day	22.2/	13
Shibata, ^[48] 2016, Japan	2013.4–2015.12	230 (113)	63	22.7	NA	NAFLD/NASH, HBV, PBC, ALD, other	Retrospective	Y	Within 6 months	17.0/	13
Guo, ^[49] 2016, China	2014.1–2014.12	155	59	NA	NA	HCV/66.7%, ALD, PBC	NA	U	Unclear	65.2/	9
Chen, ^[50] 2016, China	2014.1–2015.1	154 (120)	50.14	20.37	100	HBV/73.4%, HCV, ALD, other	Retrospective	Y	Within 1 week	68.8/46.8/27.9	14
Huang (unpublished), 2016, China	2014.1–2016.1	120 (78)	50.86	NA	NA	VHC, ALD, other	NA	Y	Unclear	71.2/	13
Li, ^[51] 2016, China	2015.1–2015.9	132 (95)	49.8	NA	47.7	HBV/77.3%, HCV	Retrospective	Y	Within 3 days	75.8/	12
Llop, ^[52] 2017, Spain	2013.9–2015.9	161 (104)	57	NA	NA	HCV/85.1%, other	Prospective	U	Within 6 months	15.5/	11
Yan, ^[53] 2017, China	2014–2016	140 (100)	50.4	24.3	NA	HBV	Retrospective	U	Unclear	7/29.3	13

ALD = alcoholic liver disease; BMI = body mass index; EGD = esophagoendoscopy; HBV = hepatitis B virus; HCV = hepatitis C virus; NA = not available; NAFLD = nonalcoholic fatty liver disease; NASH = nonalcoholic steatohepatitis; PBC = primary biliary cirrhosis; QUADAS-2 = scores quality assessment of diagnostic accuracy studies-2 scores, TE = transient elastography; VHC = viral hepatitis C.
 * No data in the studies.

Table 2
The results of evaluation of transient elastography for different stages of esophageal varices.

Author	Cutoff value, kPa	Grade 1-3 EV		AUROC	Cutoff Value (kPa)	Grade 2-3 EV		AUROC	Cutoff value, kPa	Grade 3 EV	
		Sensitivity (%)	Specificity (%)			Sensitivity (%)	Specificity (%)			Sensitivity (%)	Specificity (%)
Kazemi ^[17]	13.9	95	43	0.840	19.0	91	60	0.830			
Vizzutti ^[18]	17.6	90	43	0.760							
Foucher ^[19]					27.5	88	53	0.730			
Bureau ^[20]	21.1	84	71	0.851	29.3	81	61	0.762			
Jung ^[21]	19.7	87	70	0.818	29.5	77	57	0.702			
Castera ^[22]	21.5	76	78	0.840	30.5	77	85	0.870			
Nguyen-Khac ^[23]					48.0	73.2	73.2	0.755			
Malik ^[24]	20.0	NA	NA	0.850							
Meng (unpublished)					38.0	53.66	85.71	0.675			
Stefanescu ^[25]	28.0	74.36	64.29	0.749							
Pritchett ^[26]					19.8	91	56	0.760			
Stefanescu ^[6]	19.0	84	32.39	0.656	38.0	55.56	75.32	0.687			
Sporea ^[27]					31.0	83	62	NA			
Chen ^[28]					17.1	90.2	43.6	0.730			
Lj ^[29]	23.3	80	60.3	0.798	31.5	78.1	71.8	0.823	34.6	94.4	75.0
Lj ^[30]					20.1	68.2	75.2	0.749			
Wang ^[31]	12.0	67	77	0.789	21.0	77	87	0.865			
Colecchia ^[32]	16.4	0.962	0.596	NA							
Wang ^[33]	13.4	90	69	0.850	14.6	90	63	0.830			
Sporea ^[9]					29.5	77.5	86.9	0.871			
Sharma ^[34]	27.3	91	72	0.908							
Calvaruso ^[35]	17.0	71	57	0.707	19.0	72	55	0.710			
Chaojin ^[16]					16.2	85	55	0.830			
Saad ^[7]	29.7	95	67	NA	38.2	100	77.3	NA			
Lj ^[36]	22.8	83.8	63.2	0.824	30.6	82.6	70.1	0.849	34.6	95.6	70.3
Liu ^[37]					30.0	83	62	0.760			
Satzl ^[38]	27.9	72.5	78.9	0.802							
Gabriella ^[39]	19.2	85	87	0.885							
Augustin ^[40]	25.0	80	72	NA							
Hassan ^[41]	18.2	82	73	0.790	22.4	84	72	0.801			
Hu ^[42]	20.25	86.4	72.2	0.843	25.55	84.1	72.5	0.855			
Bintintan ^[5]	15.0	95	100	0.960	28.8	87.10	82.76	0.900			
Al-Hamoudi ^[43]	19.9	83.4	50.0	0.704							
Stefanescu ^[44]					38	60	71	0.705			
Zeng ^[45]	9.9	60.8	71.1	0.765	12.2	91.8	73.9	0.884	17.7	78.8	75.3
Wong ^[46]	6.8	90.3	29.2	0.690							
Shibata ^[47]	7.9	86.4	83.1	0.894							
Shibata ^[48]	6.1	84.6	61.3	0.765							
Guo ^[49]	13.0	NA	NA	NA							
Chen ^[50]	26.2	83.3	81.8	0.854	30.9	77.6	100	0.917	34.4	93.2	92.7
Huang (unpublished)	17.1	66.2	83.7	0.880							
Lj ^[51]	22.3	55.5	90.9	0.762							
Llop ^[52]	20	76.0	70.6	NA							
Yan ^[53]					19.9	95.1	74.7	0.890			

AUROC = area under the receiver operating characteristic curve; EV = esophageal varices; NA = not available.

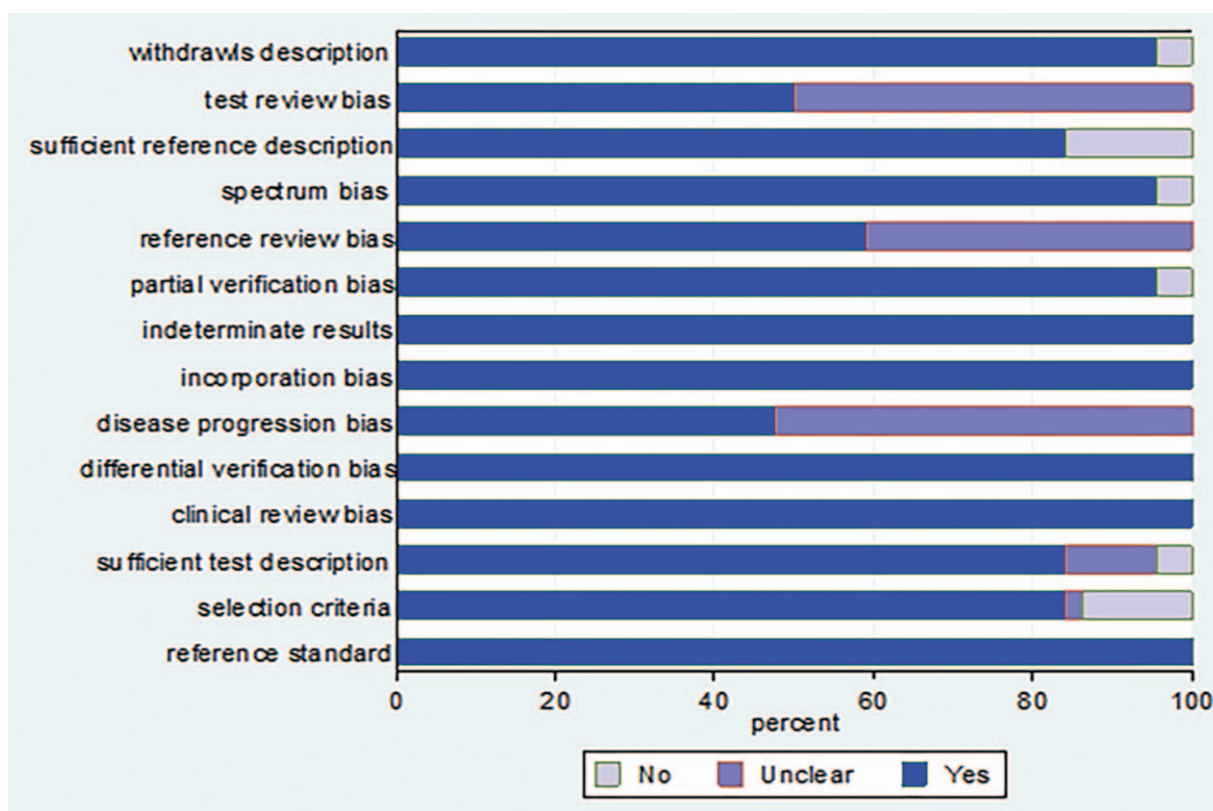


Figure 2. The graph of methodological quality of included studies.

studies did not describe inclusion and exclusion criteria of study subjects (Fig. 2).

3.3. Diagnosis of esophageal varices

Thirty-two studies ($n=4082$) evaluated the diagnostic performance of TE for detecting the presence or absence of EV. Several significant pooled indicators were calculated; the pooled sensitivity was 82% (95% CI, 78–86%), the pooled specificity was 68% (95% CI, 62–74%), the PLR was 2.6 (95% CI, 2.1–3.1), and the NLR was 0.26 (95% CI, 0.21–0.33). The DOR was 10 (95% CI, 7–14), and the AUROC was 0.83 (95% CI, 0.79–0.86) (Table 3). However, significant heterogeneity was found among 32 studies (the I^2 statistics were 86.31% and 86.60% for pooled sensitivity and specificity, respectively). Calculation of Spearman's correlation coefficient indicated that there were no threshold effects (coefficient=0.24, $P=.06$).

Six subgroup analyses performed based on different variables indicated that the sources of heterogeneity may be explained by the study location, the time interval between TE and EGD, and the proportion of Child A (Table 4). Studies conducted in Asia (not in China) showed a superior diagnostic performance of TE compared with those conducted in China (DOR [95% CI] for Asia-China vs Asia-non-China: 8 [4–13] vs 22 [12–41]). The time interval between EGD and TE also influenced the predictive values of TE for targeted disease (DOR [95% CI] for TE and EGD performed within 1 week vs more than 1 week apart vs unclear: 16 [11–23] vs 6 [3–10] vs 8 [4–17]). Interestingly, DOR for the diagnosis of the presence of EV did not significantly differ between the subgroup with a

proportion of Child A of more than 80% and the subgroup with no more than 80%, but the I^2 values for the pooled sensitivity and specificity were 50% and 63%, respectively, indicating no significant heterogeneity.

3.4. Diagnosis of substantial esophageal varices

Twenty-nine studies ($n=5221$) were included in the analysis to assess the diagnostic performance of TE for substantial EV. The pooled sensitivity (95% CI) and pooled specificity (95% CI) were 81% (77%–85%) and 72% (66%–77%), respectively. The PLR (95% CI) and NLR (95% CI) were 2.9 (2.4–3.4) and 0.26 (0.22–0.32), respectively. The DOR was 11 (95% CI, 8–15). The AUROC was 0.84 (95% CI, 0.80–0.87) (Table 3). There was considerable heterogeneity observed in this analysis (the I^2 values for the pooled sensitivity and specificity were 75.70 and 89.10, respectively), and the test results showed no evidence of a threshold effect (coefficient=0.46, $P=.22$). Similarly, additional subgroup analyses were performed according to 6 factors (Table 5): 3 factors (proportion of Child A, etiology of cirrhosis, and time interval between TE and EGD) partially affected the heterogeneity between studies. The DOR of TE in study populations with more than 80% vs no more than 80% vs unclear proportion of Child A were 20 (4–109) vs 8 (6–11) vs 10 (7–15), respectively, indicating that a proportion of Child A of more than 80% was optimal. In subgroup analysis conducted by etiology of cirrhosis, the comparison of diagnostic performance of TE was made across patients with viral and mixed etiologies (DOR 15 and 10, respectively); in particular, for HCV only, the DOR was 19. We also categorized all included studies into 3

Table 3

Summary indicators of the transient elastography for evaluating different Stages of esophageal varices.

Esophageal varices stage	N	Sen (95%CI)	Sp (95%CI)	\hat{F} (%)	PLR (95%CI)	NLR (95%CI)	DOR (95%CI)	AUROC (95%CI)	Pre-test probability	Post-test (+)	Post-test (-)
Grade 1–3	32	0.82 (0.78,0.86)	0.68 (0.62,0.74)	86.60	2.6 (2.1,3.1)	0.26 (0.21,0.33)	10 (7,14)	0.83 (0.79,0.86)	0.25	0.46	0.08
Grade 2–3	27	0.81 (0.77,0.85)	0.72 (0.66,0.77)	89.10	2.9 (2.4,3.4)	0.26 (0.22,0.32)	11 (8,15)	0.84 (0.80,0.87)	0.50	0.72	0.21
Grade 3	5	0.92 (0.83,0.96)	0.78 (0.70,0.85)	85.12	4.3 (2.9,6.2)	0.11 (0.05,0.23)	40 (15,107)	0.92 (0.89,0.94)	0.75	0.89	0.44
									0.50	0.59	0.03
									0.75	0.81	0.10
										0.93	0.24

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; DOR = diagnostic odds ratio; NLR = negative likelihood ratio; PLR = positive likelihood ratio; Sen = sensitivity; Spe = specificity; \hat{F} > 50% was considered substantial heterogeneity.

categories based on the time interval between TE and EGD. The results indicated that the DOR was better in the studies with a time interval between TE and EGD of within 1 week.

3.5. Diagnosis of large esophageal varices

Five studies provided sufficient information on the diagnostic accuracy of TE for the diagnosis of large EV. In this meta-analysis, the summary sensitivity and specificity were 92% (95% CI, 83–96%) and 78% (95% CI, 70–85%), respectively. The PLR and NLR were 4.3 (95% CI, 2.9–6.2) and 0.11 (95% CI, 0.05–0.23), respectively. The DOR and the AUROC were 40 (95% CI, 15–107) and 0.92 (95% CI, 0.89–0.94), respectively (Table 3). There was significant heterogeneity for the summary sensitivity and specificity in this analysis. The testing of the threshold effect suggested that there may be a correlation between sensitivity and specificity (coefficient=0.14, $P=.02$). Due to the small number of included studies, no further subgroup analysis was conducted to explore the sources of heterogeneity.

3.6. Sensitivity analyses and publication bias

Sensitivity analyses were performed by excluding 3 studies (Meng QQ, unpublished data, January 2010) with QUADAS-2 scores <10 points carried out in patients with any or substantial EV.^[19,49] For studies involving any EV, the pooled sensitivity, the pooled specificity, and the DOR were 0.82, 0.68, and 9, respectively, with an AUROC of 0.83, which suggested that the results did not differ substantially; for studies involving substantial EV, the pooled indicators corresponding to the summary sensitivity, specificity, DOR and AUROC were 0.82, 0.72, 11, and 0.84, respectively, which also indicated that there were no significant differences. A repeated analysis was also conducted after excluding each study sequentially, which showed that the results did not change significantly.

On implementing the asymmetry test of the Deeks funnel plot, a significant publication bias was found among studies evaluating TE for detecting the presence of any EV ($P=.03$) (Fig. 3A). There was no evidence of a great possibility of publication bias for TE for assessing substantial and large EV ($P=.88$ and $P=.26$, respectively) (Fig. 3B and C).

4. Discussion

In this meta-analysis, we identified 44 studies comparing TE (using FibroScan) with EGD for detecting the presence and degree of EV. Of these, 32 studies evaluated the diagnostic performance of TE for the presence or absence of EV (grade 1–3), 27 studies for substantial EV (grade 2–3), and 5 studies for large EV (grade 3). Calculating all summary indicators led to several important discoveries. First, TE had a relatively high diagnostic accuracy for assessing the presence of any EV, substantial EV, and large EV, with a summary AUROC of 0.83, 0.84, and 0.92, respectively. Second, the pooled sensitivity and specificity corresponding to the presence of any EV, substantial EV, and large EV were 0.82 and 0.68, 0.81 and 0.72, 0.92, and 0.78, respectively, indicating that the estimates of pooled sensitivity and specificity were considered to be good. In particular, the results for detecting large EV were almost excellent. However, the summary specificity for assessing any EV was not satisfactory. Third, TE showed a high DOR of 40 in identifying large EV; however, the DORs for any and substantial EV were less than ideal.

Table 4

Subgroup analysis reporting the diagnostic accuracy of transient elastography for the liver stiffness measurement for evaluating the presence or absence of esophageal varices.

Subgroup	N	Sensitivity	I^2	Specificity	I^2	PLR	NLR	DOR	AUROC
Location (a)									
Europe	12	0.84 (0.77,0.89)	78.96	0.63 (0.52,0.73)	82.55	2.3 (1.7,3.0)	0.25 (0.17,0.38)	9 (5,16)	0.82 (0.78,0.85)
Asia	16	0.81 (0.74,0.86)	89.82	0.71 (0.63, 0.79)	90.25	2.8 (2.1,3.7)	0.27 (0.20,0.37)	10 (6,17)	0.83 (0.80,0.86)
Others	4	0.82 (0.74,0.88)	32.70	0.68 (0.54,0.80)	66.57	2.6 (1.7,3.8)	0.26 (0.19,0.38)	10 (5,18)	0.84 (0.80,0.87)
Location (b)									
Asia-China	11	0.77 (0.69,0.84)	88.76	0.69 (0.57,0.79)	91.68	2.5 (1.8,3.5)	0.33 (0.24,0.45)	8 (4,13)	0.80 (0.76,0.83)
Asia-nonChina	5	0.88 (0.84,0.91)	0	0.75 (0.65,0.83)	82.73	3.5 (2.4,5.1)	0.16 (0.11,0.22)	22 (12,41)	0.89 (0.86,0.91)
Child A									
> 80% (100%)	4	0.75 (0.67,0.82)	50.26	0.75 (0.65,0.82)	63.89	3.0 (2.0,4.4)	0.33 (0.23,0.47)	9 (4,18)	0.81 (0.78,0.85)
≤ 80%	12	0.84 (0.76,0.90)	90.74	0.64 (0.51,0.75)	82.96	2.3 (1.6,3.4)	0.25 (0.15,0.41)	9 (4,20)	0.82 (0.79,0.85)
Unclear	16	0.83 (0.77,0.87)	86.15	0.69 (0.61,0.76)	89.29	2.7 (2.1,3.4)	0.25 (0.19,0.32)	11 (7,16)	0.84 (0.80,0.87)
Etiology									
Viral	13	0.81 (0.73,0.87)	87.55	0.67 (0.57,0.76)	89.06	2.5 (1.9,3.2)	0.28 (0.20,0.39)	9 (6,14)	0.81 (0.78,0.84)
Mixed	19	0.83 (0.78,0.87)	82.87	0.68 (0.60,0.76)	85.04	2.6 (2.0,3.4)	0.25 (0.19,0.33)	11 (6,17)	0.84 (0.81,0.87)
HCV only	8	0.84 (0.73,0.91)	89.82	0.68 (0.59,0.76)	66.52	2.7 (2.0,3.5)	0.24 (0.14,0.41)	11 (6,23)	0.80 (0.76,0.83)
Interval									
Within 1week	13	0.86 (0.79,0.91)	86.81	0.72 (0.63,0.79)	82.29	3.1 (2.4,4.0)	0.19 (0.14,0.27)	16 (11,23)	0.86 (0.83,0.89)
Over 1 week	8	0.80 (0.71,0.87)	88.50	0.58 (0.46,0.70)	91.57	1.9 (1.5,2.5)	0.34 (0.24,0.48)	6 (3,10)	0.78 (0.74,0.81)
Unclear	11	0.78 (0.71,0.83)	79.99	0.70 (0.59,0.80)	80.36	2.6 (1.8,3.9)	0.31 (0.22,0.45)	8 (4,17)	0.81 (0.78,0.85)
Cutoff value, kPa									
Cutoff > 18.7	18	0.80 (0.75,0.84)	81.94	0.69 (0.61,0.76)	83.12	2.6 (2.0,3.3)	0.29 (0.23,0.37)	9 (6,14)	0.82 (0.79,0.85)
Cutoff ≤ 18.7	14	0.85 (0.78,0.91)	90.20	0.66 (0.56,0.75)	90.13	2.5 (1.9,3.4)	0.22 (0.15,0.33)	12 (7,20)	0.84 (0.80,0.87)
Blinding									
Yes	21	0.84 (0.79,0.88)	82.51	0.68 (0.60,0.75)	88.61	2.6 (2.1,3.3)	0.24 (0.19,0.30)	11 (8,16)	0.85 (0.81,0.87)
Unclear	11	0.78 (0.71,0.84)	87.88	0.69 (0.58,0.78)	80.37	2.5 (1.8,3.6)	0.31 (0.21,0.46)	8 (4,16)	0.81 (0.77,0.84)

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; DOR = diagnostic odds ratio; EV = esophageal varices; NLR = negative likelihood ratio; PLR = positive likelihood ratio; Sen = sensitivity; Spe = specificity; TE = transient elastography. $I^2 > 50\%$ was considered substantial heterogeneity.

Table 5

Subgroup analysis reporting the diagnostic accuracy of transient elastography for the liver stiffness measurement for detecting substantial esophageal varices.

Subgroup	N	Sensitivity	I^2	Specificity	I^2	PLR	NLR	DOR	AUROC
Location									
Europe	11	0.78 (0.71,0.84)	78.89	0.71 (0.63,0.77)	90.93	2.7 (2.1,3.4)	0.30 (0.23,0.40)	9 (6,13)	0.81 (0.78,0.85)
Asia	13	0.81 (0.75, 0.86)	74.33	0.74 (0.63,0.83)	89.63	3.1 (2.2,4.5)	0.25 (0.19,0.33)	12 (7,21)	0.85 (0.82,0.88)
Child A									
> 80%	5	0.77 (0.69,0.84)	0.00	0.86 (0.54,0.97)	93.81	5.4 (1.4,21.4)	0.26 (0.17,0.40)	20 (4,109)	0.78 (0.74,0.81)
≤ 80%	11	0.80 (0.72,0.86)	80.52	0.67 (0.61,0.73)	82.40	2.4 (2.1,2.9)	0.30 (0.22,0.40)	8 (6,11)	0.79 (0.75,0.82)
Unclear	11	0.83 (0.76,0.88)	80.50	0.70 (0.62,0.76)	90.78	2.8 (2.2,3.4)	0.24 (0.18,0.33)	11 (8,17)	0.83 (0.80,0.86)
Etiology									
Viral	9	0.85 (0.80,0.90)	56.44	0.72 (0.63,0.80)	89.91	3.0 (2.2,4.1)	0.20 (0.15,0.29)	15 (9,25)	0.87 (0.84,0.90)
Mixed	16	0.79 (0.73,0.84)	77.66	0.73 (0.64,0.80)	89.01	2.9 (2.2,3.9)	0.29 (0.22,0.37)	10 (7,16)	0.83 (0.80,0.86)
HCV only	5	0.88 (0.79,0.93)	62.18	0.72 (0.63,0.80)	72.11	3.2 (2.2,4.5)	0.17 (0.09,0.32)	19 (7,50)	0.87 (0.83,0.89)
Interval									
Within 1week	7	0.83 (0.78,0.87)	26.71	0.75 (0.51,0.90)	89.58	3.3 (1.5,7.2)	0.23 (0.16,0.32)	15 (5,41)	0.84 (0.81,0.87)
Over 1 week	8	0.80 (0.70,0.87)	86.18	0.73 (0.65,0.79)	91.63	2.9 (2.3,3.7)	0.28 (0.18,0.42)	11 (6,19)	0.82 (0.78,0.85)
Unclear	12	0.79 (0.72,0.85)	78.37	0.71 (0.63,0.78)	87.99	2.7 (2.2,3.4)	0.29 (0.23,0.37)	9 (7,12)	0.82 (0.78,0.85)
Cutoff value, kPa									
Cutoff > 26.5	16	0.78 (0.72,0.82)	75.31	0.75 (0.67,0.82)	89.81	3.2 (2.3,4.3)	0.30 (0.24,0.37)	11 (7,17)	0.83 (0.80,0.86)
Cutoff ≤ 26.5	11	0.86 (0.80,0.90)	72.26	0.66 (0.58,0.73)	87.68	2.5 (2.1,3.1)	0.21 (0.16,0.29)	12 (8,18)	0.85 (0.81,0.87)
Blinding									
Yes	16	0.81 (0.76,0.85)	68.07	0.73 (0.64,0.81)	93.02	3.0 (2.2,4.2)	0.26 (0.20,0.33)	12 (7,19)	0.85 (0.81,0.88)
Unclear	11	0.81 (0.73,0.87)	83.51	0.69 (0.64,0.75)	68.83	2.7 (2.3,3.1)	0.27 (0.20,0.38)	10 (6,14)	0.79 (0.76,0.83)

AUROC = area under the receiver operating characteristic curve; CI = confidence interval; DOR = diagnostic odds ratio; NLR = negative likelihood ratio; PLR = positive likelihood ratio; Sen = sensitivity; Spe = specificity. $I^2 > 50\%$ was considered substantial heterogeneity.

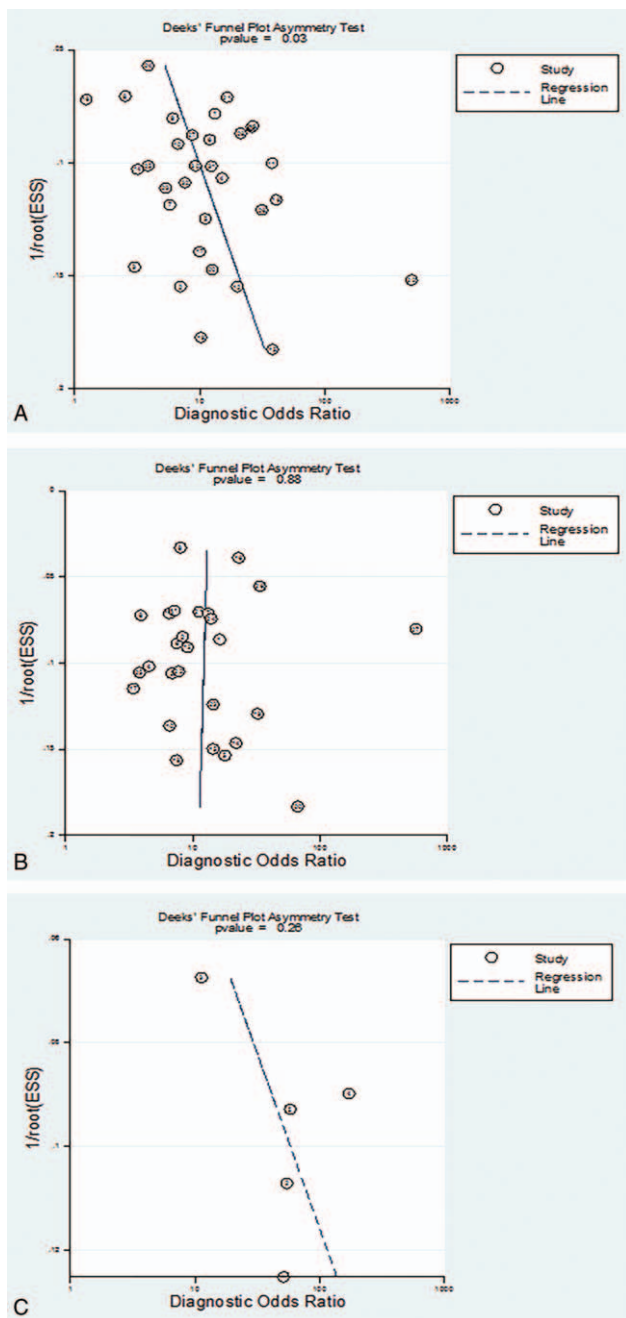


Figure 3. Deeks funnel plot of TE for the prediction of the presence and degree of EV. (A) TE detection of the presence of any EV. (B) TE detection of substantial EV. (C) TE detection of large EV. EV=esophageal varices, TE=transient elastography.

Our results suggest that TE could be used for diagnosing the presence or absence of EV. When pre-test probability equaled 50%, the probability of correctly diagnosing the presence of EV was 72% following a positive result. For later stage EV, the values were more convincing; the probabilities of correctly diagnosing substantial and large EV following a positive test result were 74% and 81%, respectively. If the test results were negative, there might still have been the presence of any EV present in 21% of patients, and large EV in 10%. When there was

a high pre-test probability of 75%, then the probability of an accurate diagnosis following a positive test result was more than 90% for all EV stages, but an incorrect diagnosis would still be possible in 24% to 44% of patients with a negative test result.

We observed considerable heterogeneity among studies assessing the diagnostic accuracy of TE for detecting the presence or absence of EV. To determine the source of heterogeneity, all studies were divided into 3 subgroups based on the study location (European, Asian, and others). However, there was no obvious difference in the diagnostic performance between subgroups. Furthermore, all studies performed in Asia were divided into 2 groups (non-Chinese and Chinese), as significant heterogeneity was found among all Asian studies. The DOR for detecting EV in Asian-non-Chinese studies was more than twice as high as that in Chinese studies, and the I^2 value for sensitivity was 0%, indicating that heterogeneity did not exist; this suggests that the heterogeneity of the studies may be explained by the different population groups. Although the DOR for identifying EV was 16 in studies in which the detection of EV was performed by TE within 1 week of EGD, and the DOR was 6 in studies where TE and EGD were performed more than 1 week apart, the heterogeneity of the pooled sensitivity and specificity was still very significant. This observed difference is possibly associated with differences in indicators of the body checked by TE on different days, as the situation of patients is dictated by the severity of disease and the clinical setting.^[54] Therefore, the effect may have been better if all studies were divided into 2 subgroups: studies with EV detection via TE and EGD performed on the same day, and those that reported a time interval between TE and EGD. Although there was no obvious difference between the DOR values on the basis of the proportion of Child A, there was a reduction in variation of the pooled sensitivity and specificity. That is, the proportion of Child A might be the origin of heterogeneity; however, additional exploration in future studies is required to confirm this.

There was also significant heterogeneity among included studies evaluating substantial EV. The source of heterogeneity was explored according to the proportion of Child A, etiology of cirrhosis, and the time interval between TE and EGD. Within each subgroup, we considered that the DORs for evaluating substantial EV of different levels of variables were comparable, and that the I^2 value for sensitivity or specificity was decreased. Thus, the variation in diagnostic performance of TE for substantial EV was potentially identified based on the proportion of Child A, etiology of cirrhosis, and time interval between TE and EGD; however, this result also should be considered with caution.

Although our meta-analysis is not the first to evaluate the diagnostic performance of TE for detecting the presence and degree of EV, we included more than twice the number of studies included in 4 previous meta-analyses published on the same subject.^[55-58] Li et al^[55] assessed the diagnostic accuracy of TE for detecting substantial EV in 20 included studies, and reported that summary sensitivity, specificity, PLR, NLR, DOR, and AUROC were 0.81, 0.71, 2.63, 0.27, 10.3, and 0.83 respectively, indicating that all pooled indicators were similar to ours. Pu et al,^[56] Qu et al,^[57] and Shi et al^[58] analyzed a total of 13, 20, and 12 studies, respectively, regarding the predictive accuracy of TE for detecting the presence of any and substantial EV, and reported results consistent with our research. Compared with other studies (including ours), the values of pooled specificity in the analysis by Shi et al^[58] were low (0.53 and 0.59 for any and substantial EV, respectively).

The first strength of our meta-analysis was the comprehensive assessment of the diagnostic performance of LSM measured by TE techniques for identifying the presence and size of EV. The association between LSM measured by TE and varying degrees of EV could be thoroughly analyzed by extracting a large amount of basic information. Most importantly, this is the first such review of the predictive performance of TE for detecting large EV; large EV carry a high risk of bleeding and mortality,^[2,3] which has a crucial impact on clinical practice and therapeutic decisions. Second, we exhaustively and vigorously searched numerous databases without language restrictions, resulting in the inclusion of a large number of studies in the final review. Furthermore, the latest publication year of studies included in this analysis was 2017,^[52,53] which makes the evaluation results more reliable and applicable to actual clinical practice. Third, as both sensitivity and specificity were independent of cutoff value, multiple subgroup analyses were carried out based on different variables to explore the sources of heterogeneity; we were able to potentially identify that the differences in diagnostic accuracy of TE for detecting the presence or absence of EV were mainly based on study location, time interval between TE and EGD, and proportion of Child A. For substantial EV, the potential heterogeneity primarily derived from the proportion of Child A, etiology of cirrhosis, and the time interval between TE and EGD.

Several limitations in our meta-analysis must be acknowledged. First, the limited number of included studies and the differences in TE cutoff values in each study made it difficult to define an optimal diagnostic threshold for accurate prediction of any and substantial EV. Second, assessment of methodological quality found that most studies provided insufficient information on whether the TE results were assessed by evaluators blinded to the EGD results, or vice versa, creating the risk of review bias. Furthermore, many studies did not state the time interval between TE and EGD, resulting in a risk of disease progression bias. Third, there was a significant publication bias in studies evaluating the accuracy of TE in detecting the presence of any EV. That is, the reproducibility and accuracy of the study conclusions would be limited because of negative results. We attempted to review all relevant research, including non-English language publications and postgraduate theses, to minimize the effect of publication bias (Meng QQ, unpublished data, January 2010; Huang LL, unpublished data, June 2016). Fourth, subgroup analyses that compared the observed differences in DOR values between several subgroups were conducted to explain the potential origins of heterogeneity; the decrease in I^2 values of the pooled sensitivity and specificity indicate that the present results should be interpreted cautiously.

In conclusion, this meta-analysis found that TE has a high degree of predictive accuracy for identification of the presence and size of EV in patients with liver cirrhosis, as compared with EGD. In the field of liver disease, this available and noninvasive TE technique represents a major advance; however, TE has not yet displaced EGD as the method of choice for diagnosing EV. Additional high-quality studies and more advanced data analysis techniques are required to further prove the performance of TE.

Author contributions

Conceptualization: Jinchun Liu, Lijun Jiang.

Data curation: Fan Cheng.

Formal analysis: Fan Cheng.

Methodology: Hongyan Cao.

Project administration: Yanbo Zhang, Dongxing Guo.

Software: Hongjuan Han.

Supervision: Yanbo Zhang, Dongxing Guo.

Validation: Hongyan Cao.

Visualization: Hongjuan Han.

Writing – original draft: Fan Cheng.

References

- Garcia-Tsao G, Sanyal AJ, Grace ND, et al. Prevention and management of gastroesophageal varices and variceal hemorrhage in cirrhosis. *Hepatology* 2007;46:922–38.
- Brocchi E, Caletti G, Brambilla G, et al. Prediction of the first variceal hemorrhage in patients with cirrhosis of the liver and esophageal varices. *N Engl J Med* 1988;319:983–9.
- Jensen DM. Endoscopic screening for varices in cirrhosis: findings, implications, and outcomes. *Gastroenterology* 2002;122:1620–30.
- Sharma VK, Nguyen CC, Crowell MD, et al. A national study of cardiopulmonary unplanned events after GI endoscopy. *Gastrointest Endosc* 2007;66:27–34.
- Bintintan A, Chria RI, Bintintan VV, et al. Value of hepatic elastography and Doppler indexes for predictions of esophageal varices in liver cirrhosis. *Med Ultrason* 2015;17:5–11.
- Stefanescu H, Grigorescu M, Badea R, et al. A new and simple algorithm for the noninvasive assessment of esophageal varices in cirrhotic patients using serum fibrosis markers and transient elastography. *J Gastrointest Liver Dis* 2011;20:57–64.
- Saad Y, Said M, Idris MO, et al. Liver stiffness measurement by fibroscan predicts the presence and size of esophageal varices in Egyptian patients with HCV related liver cirrhosis. *J Clin Diagn Res* 2013;7:2253–7.
- Moherl D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta analyses: the PRISMA statement. *J Chin Integr Med* 2009;7:889–96.
- Sporea I, Rațiu I, Bota S, et al. Are different cut-off values of liver stiffness assessed by transient elastography according to the etiology of liver cirrhosis for predicting significant esophageal varices? *Med Ultrason* 2013;15:111–5.
- Whiting P, Rutjes AWS, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004;8:1–234.
- Fagan TJ. The letter: nomogram for Bayes theorem. *BMJ* 1993;293:257.
- Annane D, Bellissant E, Bollaert PE, et al. Corticosteroids for severe sepsis and septic shock: a systematic review and meta-analysis. *BMJ* 2004;329:480.
- Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
- Li J, Qin L, Liu M. Basic methods for systematic review. *Chin J Evid Based Med* 2001;1:34–8.
- Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882–93.
- Chaojin T, Piratvisuth T. The prediction values of transient elastography for clinically significant esophageal varices in patients with liver cirrhosis. *Thai J Gastroenterol* 2013;14:66–70.
- Kazemi F, Kettaneh A, N'kontchou G, et al. Liver stiffness measurement selects patients with cirrhosis at risk of bearing large oesophageal varices. *J Hepatol* 2006;45:230–5.
- Vizzutti F, Arena U, Romanelli RG, et al. Liver stiffness measurement predicts severe portal hypertension in patients with HCV-related cirrhosis. *Hepatology* 2007;45:1290–7.
- Foucher J, Chanteloup E, Vergniol J, et al. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. *Gut* 2006;55:403–8.
- Bureau C, Metivier S, Peron JM, et al. Transient elastography accurately predicts presence of significant portal hypertension in patients with chronic liver disease. *Aliment Pharmacol Ther* 2008;27:1261–8.
- Jung HS, Kim YS, Kwon OS, et al. Usefulness of liver stiffness measurement for predicting the presence of esophageal varices in patients with liver cirrhosis. *Korean J Hepatol* 2008;14:342–50.
- Castera L, Bail BL, Roudot-Thoraval F, et al. Early detection in routine clinical practice of cirrhosis and oesophageal varices in chronic hepatitis C: comparison of transient elastography (FibroScan) with standard laboratory tests and non-invasive scores. *J Hepatol* 2009;50:59–68.
- Nguyen-Khac E, Saint-Leger P, Tramier B, et al. Noninvasive diagnosis of large esophageal varices by Fibroscan: strong influence of the cirrhosis etiology. *Alcohol Clin Exp Res* 2010;34:1146–53.
- Malik R, Lai M, Sadiq A, et al. Comparison of transient elastography, serum markers and clinical signs for the diagnosis of compensated cirrhosis. *J Gastroenterol Hepatol* 2010;25:1562–8.

- [25] Stefanescu H, Grigorescu M, Lupsor M, et al. Spleen stiffness measurement using FibroScan for the noninvasive assessment of esophageal varices in liver cirrhosis patients. *J Gastroenterol Hepatol* 2011;26:164–70.
- [26] Pritchett S, Cardenas A, Manning D, et al. The optimal cut-off for predicting large oesophageal varices using transient elastography is disease specific. *J Viral Hepat* 2011;18:75–80.
- [27] Sporea I, Rațiu I, Șirli R, et al. Value of transient elastography for the prediction of variceal bleeding. *World J Gastroenterol* 2011;17:2206–10.
- [28] Chen YP, Zhang Q, Dai L, et al. Is transient elastography valuable for high-risk esophageal varices prediction in patients with hepatitis-B-related cirrhosis? *J Gastroenterol Hepatol* 2012;27:533–9.
- [29] Li F, Yan T, Zhang J, et al. FibroScan can be used to diagnose the size of oesophageal varices in patients with HBV-related cirrhosis. *Chin J Exp Clin Virol* 2012;26:470–3.
- [30] Li TH, Liu F, Han T, et al. Noninvasive assessment of esophageal-gastric varices by spleen stiffness in liver cirrhosis patients. *Chin J Infect Dis* 2012;30:603–7.
- [31] Wang JH, Chuah SK, Lu SN, et al. Transient elastography and simple blood markers in the diagnosis of esophageal varices for compensated patients with hepatitis B virus-related cirrhosis. *J Gastroenterol Hepatol* 2012;27:1213–8.
- [32] Colecchia A, Montrone L, Scaioili E, et al. Measurement of spleen stiffness to evaluate portal hypertension and the presence of esophageal varices in patients with HCV-related cirrhosis. *Gastroenterology* 2012;143:646–54.
- [33] Wang HM, Lo GH, Chen WC, et al. Efficacy of transient elastography in screening for large esophageal varices in patients with suspicious or proven liver cirrhosis. *J Dig Dis* 2012;13:430–8.
- [34] Sharma P, Kumar A, Kimake V, et al. Spleen stiffness in patients with cirrhosis in predicting esophageal varices. *Am J Gastroenterol* 2013;108:1101–7.
- [35] Calvaruso V, Bronte F, Conte E, et al. Modified spleen stiffness measurement by transient elastography is associated with presence of large oesophageal varices in patients with compensated hepatitis C virus cirrhosis. *J Viral Hepat* 2013;20:867–74.
- [36] Li F, Yan T, Shao Q, et al. Clinical study of FibroScan efficiency for diagnosing size of oesophageal varices in liver cirrhosis patients. *Chin J Hepatol* 2014;22:600–3.
- [37] Liu YY, Ren WG. The role of Fibroscan in predicting the degree of esophageal varices and bleeding in cirrhosis patients. *Contemp Med* 2014;20:10–1.
- [38] Salzl P, Reiberger T, Ferlitsch M, et al. Evaluation of portal hypertension and varices by acoustic radiation force impulse imaging of the liver compared to transient elastography and AST to platelet ratio index. *Ultraschall Med* 2014;35:528–33.
- [39] Gabriella P, Andrea T, Ferenc P, et al. Transient elastography as a predictor of oesophageal varices in patients with liver cirrhosis. *Orv Hetil* 2014;155:270–6.
- [40] Augustin S, Millán L, González A, et al. Detection of early portal hypertension with routine data and liver stiffness in patients with asymptomatic liver disease: a prospective study. *J Hepatol* 2014;60:561–9.
- [41] Hassan EM, Omrana DA, El Beshlaweya ML, et al. Can transient elastography, Fib-4, Forns Index, and Lok Score predict esophageal varices in HCV-related cirrhotic patients? *Gastroenterol Hepatol* 2014;37:58–65.
- [42] Hu ZW, Li YY, Li C, et al. Using ultrasonic transient elastometry (FibroScan) to predict esophageal varices in patients with viral liver cirrhosis. *Ultrasound Med Biol* 2015;41:1530–7.
- [43] Al-Hamoudi WK, Abdelrahmana AA, Helmy A, et al. The role of Fibroscan in predicting the presence of varices in patients with cirrhosis. *Eur J Gastroenterol Hepatol* 2015;27:1307–12.
- [44] Stefanescu H, Radu C, Procopet B, et al. Non-invasive menage a trois for the prediction of high-risk varices: stepwise algorithm using lok score, liver and spleen stiffness. *Liver Int* 2015;35:317–25.
- [45] Zeng XH, Fang CX, Wang YM. Efficiency of FibroScan in assessing esophageal varices in patients with chronic hepatitis B. *Med J Chin PLA* 2016;44:144–8.
- [46] Wong GL, Kwok R, Chan HY, et al. Measuring spleen stiffness to predict varices in chronic hepatitis B cirrhotic patients with or without receiving non-selective beta-blockers. *J Dig Dis* 2016;17:538–46.
- [47] Shibata S, Joshita S, Umemura T, et al. Liver stiffness-spleen size-to-platelet ratio risk score identifies esophageal varices in Japanese patients with chronic hepatitis C. *Hepatol Res* 2016;46:884–9.
- [48] Shibata S, Umemura T, Yamazaki T, et al. Liver stiffness-spleen size-to-platelet ratio risk score detects esophageal varices in chronic liver disease. *Springerplus* 2016;5:1–6.
- [49] Guo CM, Wu J, Liu KL, et al. The predictive value of transient elastography for esophageal varices in hepatocirrhosis and its correlation with liver function classification. *Chin J Gastroenterol Hepatol* 2016;25:314–7.
- [50] Chen ZM, Huang DY, Liu HY, et al. Clinical value of transient elastography combined with aspartate aminotransferase to platelet ratio index and fibrosis index based on four factors for predicting the severity of esophageal varices in hepatic cirrhosis. *Acad J Guangzhou Med Univ* 2016;44:22–5.
- [51] Li WJ, Shang YY, Wang BX, et al. The diagnostic values of fibroscan for assessing esophageal varices an bleeding in cirrhosis patients. *J Pract Med* 2016;32:2893–6.
- [52] Llop E, Lopez M, Revilla J, et al. Validation of non invasive methods to predict the presence of gastroesophageal varices in a cohort of patients with compensated advanced chronic liver disease. *J Gastroenterol Hepatol* 2017;32:1867–72.
- [53] Yan S, Zhang HH, Guo QY, et al. Noninvasive models value to the diagnosis of esophageal varices in patients with chronic hepatitis B. *World Chin J Digestol* 2017;25:227–33.
- [54] Kapoor A, Dharel N, Sanyal AJ. Endoscopic diagnosis and therapy in gastroesophageal variceal bleeding. *Gastrointest Endosc Clin N Am* 2015;25:491–507.
- [55] Li T, Qu YD, Yang BH, et al. Evaluation of large esophageal varices in cirrhotic patients by transient elastography: a meta-analysis. *Rev Esp Enferm Dig* 2016;108:464–72.
- [56] Pu K, Shi JH, Wang X, et al. Diagnostic accuracy of transient elastography (FibroScan) in detection of esophageal varices in patients with cirrhosis: a meta-analysis. *World J Gastroenterol* 2017;23:345–56.
- [57] Qu YD, Li T, Ye Q, et al. A beginning or the end? A meta-analysis to assess the diagnostic accuracy of transient elastography for the prediction of esophageal varices. *Saudi J Gastroenterol* 2016;22:345–52.
- [58] Shi KQ, Fan YC, Pan ZZ, et al. Transient elastography: a meta-analysis of diagnostic accuracy in evaluation of portal hypertension in chronic liver disease. *Liver Int* 2013;33:62–71.