



Technical performance of shear wave elastography for measuring liver stiffness in pediatric and adolescent patients: a systematic review and meta-analysis

Dong Wook Kim¹ · Chan Park² · Hee Mang Yoon³ · Ah Young Jung³ · Jin Seong Lee³ · Seung Chai Jung³ · Young Ah Cho³

Received: 23 September 2018 / Revised: 26 October 2018 / Accepted: 20 November 2018 / Published online: 7 January 2019

© European Society of Radiology 2019

Abstract

Objective To assess the technical performance (proportion of technical failure and unreliable measurements) of shear wave elastography (SWE) for assessing liver stiffness in pediatric and adolescent patients.

Methods We searched Ovid-MEDLINE and EMBASE databases for eligible studies and selected original articles investigating transient elastography (TE), point shear wave elastography (pSWE), or two-dimensional SWE (2D-SWE) for measuring liver stiffness in pediatric and adolescent patients. A quantitative synthesis of studies reporting technical failures and/or unreliable measurements of TE, pSWE, or 2D-SWE is presented. Meta-analytic pooling was conducted using the random effects model. Meta-regression analysis was conducted to explore potential causes of heterogeneity.

Results Forty of 69 studies (58%) provided technical performance information. Technical failure data were reported in 3 TE, 6 pSWE, and 8 2D-SWE studies. Unreliable measurement data were provided in 21 TE, 4 pSWE, and 1 2D-SWE study. The pooled proportion of unreliable measurements of TE was 12.1%. Meta-regression analysis showed that the study population size and readers' blinding to pathologic results affected the study's heterogeneity. The pooled proportions of technical failure during pSWE and 2D-SWE were 4.1% and 2.2%, respectively, demonstrating no significant difference between the techniques.

Conclusions We reviewed the technical performance of SWE, especially the rate of unreliable measurements from TE studies and rates of technical failure from pSWE and 2D-SWE studies. Considering the importance of technical performance for clinical validation of SWE, numbers of and reasons for technical failure and unreliable measurements should be reported in future studies. Further efforts are necessary to standardize SWE reliability criteria.

Key Points

- Most TE studies reported rate of unreliable measurements, whereas pSWE and 2D-SWE studies were likely to report rates of technical failure.
- The pooled proportion of unreliable measurements of TE was 12.1%.
- The pooled proportions of technical failure during pSWE and 2D-SWE were 4.1% and 2.2%, respectively, demonstrating no significant difference between the techniques.

Keywords Elasticity imaging techniques · Liver · Meta-analysis · Pediatrics · Ultrasonography

Dong Wook Kim and Chan Park contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00330-018-5900-6>) contains supplementary material, which is available to authorized users.

✉ Hee Mang Yoon
espoirhm@gmail.com

¹ Department of Radiology, Taeon-gun Health Center and County Hospital, 1952-16, Seohae-ro, Pyeongcheon-ri, Taeon-eup, Taeon-gun, Chungcheongnam-do 32148, Republic of Korea

² Department of Radiology, Chonnam National University Hospital, 42, Jebong-ro, Dong-gu, Gwangju 61469, Republic of Korea

³ Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Republic of Korea

Abbreviations

2D-SWE	Two-dimensional shear wave elastography
ARFI	Acoustic radiation force impulse
pSWE	Point shear wave elastography
SWE	Shear wave elastography
TE	Transient elastography
US	Ultrasound

Introduction

Ultrasound (US) elastography is a non-invasive technique used to measure liver stiffness, for grading liver fibrosis or predicting portal hypertension, in patients with chronic liver disease. Shear wave elastography (SWE) is the US elastography technique used most widely for measuring liver stiffness. SWE may be performed with any of three techniques: transient elastography (TE), point shear wave elastography (pSWE), and two-dimensional shear wave elastography (2D-SWE). TE (FibroScan; Echosens) was the first technique developed and is used most widely. For TE, a 50-Hz mechanical impulse is delivered to the skin to generate a shear wave. The velocity of the shear wave generated by liver tissue is measured [1, 2]. pSWE and 2D-SWE are comparatively new techniques that use an acoustic radiation force impulse (ARFI) of 100–500 Hz to cause liver tissue deformation and generate a shear wave [3]. The operator freely chooses where to place the region-of-interest (ROI) under the guidance of gray-scale US images using conventional US probes. This approach cannot be used when performing TE. In pSWE, average shear wave speed within an ROI of fixed size is determined. Conversely, 2D-SWE allows the operator to modify the size of the ROI and to obtain color elasticity maps [3].

Although US elastography has been described by many studies as an effective diagnostic method for use in the adult population [1, 4–6], fewer studies have investigated use in pediatric and adolescent populations [7–42]. Furthermore, beyond diagnostic accuracy, the ability to obtain successful and reliable measurements is also important. Notably, the evaluation of children carries some disadvantages, such as small body size or the potential inability of the patient to hold his or her breath. Currently, a large population study [43] reported on the rates of technical failure and unreliable measurement when US elastography was used to assess an adult population. However, the issue has been investigated in pediatric and adolescent populations only in small-scale studies [7–42].

In this study, we systematically reviewed and conducted a meta-analysis to evaluate the technical performance of US elastography in pediatric and adolescent patients.

Materials and methods

This systematic review and meta-analysis was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [44].

Literature search

MEDLINE and EMBASE databases (up to July 12, 2017) were searched to find studies that were relevant to our research. The search terms used were as follows: ((children) OR (pediatric*) OR (paediatric*) OR (adolescent*)) AND ((liver) OR (hepatic)) AND ((elastography) OR (“transient elastography”) OR (TE) OR (fibroscan) OR (acoustic radiation force impulse imaging) OR (ARFI) OR (Virtual Touch tissue quantification) OR (VTQ) OR (Virtual Touch tissue imaging quantification) OR (VTIQ) OR (shear wave elastography) OR (shearwave elastography) OR (shear-wave elastography) OR (SWE) OR (Supersonic) OR (Aixplorer) OR (shear wave speed imaging)). Only English articles were evaluated. The bibliographies of all selected articles were screened to identify additional relevant publications.

Two reviewers independently performed literature search, study selection, and data extraction (D.W.K. and H.M.Y., with 2 and 5 years of experience in systematic reviews and meta-analyses, respectively). A third reviewer resolved all disagreements (Y.A.C., with 24 years of experience in pediatric radiology).

Exclusion criteria were as follows: (a) case reports or series including < 10 patients; (b) reviews, editorials, letters, comments, or conference abstracts; (c) studies using US elastography modalities other than shear wave techniques (e.g., strain elastography); and (d) partially overlapping study populations.

Original articles investigating shear wave elastography (TE, pSWE, or 2D-SWE) to measure liver stiffness in pediatric and adolescent patients (more than 95% of the population being under 20 years old) were selected for analysis. Retrieved studies were initially screened through their titles and abstracts for potential eligibility, and subsequently, the full texts were reviewed for final inclusion.

Data extraction

Data were extracted using a standardized form related to (a) study characteristics: authors, year of publication, institution, country of origin, duration of patient recruitment, and study design (prospective vs. retrospective); (b) demographic and clinical characteristics: patient number, male/female ratio, age (mean age and range), and etiology; (c) technical characteristics: type of shear wave elastography (TE, pSWE, or 2D-SWE), model, probe, number of measurements, representative values, number of readers, and presence of reader blinding; (d)

outcomes: rate of technical failure and/or unreliable measurement for each type of SWE.

Technical failure was defined as no or little value obtained for all acquisitions. However, based on the definition of the unreliable measurement for SWE, a consensus across the different techniques is unavailable. According to the manufacturer's recommendations [45], TE measurements are unreliable when they do not meet any of the following criteria: (a) < 10 valid measurements; (b) success rate (valid shots/total number of shots) < 60%; (c) interquartile range (IQR) \geq 30% of median liver stiffness value. However, a clear guideline for the unreliable measurement of pSWE and 2D-SWE is unavailable. Therefore, for all articles, we applied the original definition used in the individual study.

Data synthesis and analysis

The primary outcome of our systematic review and meta-analysis was the pooled proportion of technical failures and/or unreliable measurements of SWE. Meta-analytic pooling was performed using the inverse variance method to calculate weights [46–48]. Overall proportion was used to obtain random effects meta-analysis of single proportion. Logit transformation of proportion was performed. A confidence interval (CI) was obtained by Clopper–Pearson interval for individual studies, and a continuity correction of 0.5 in studies with zero cell frequencies was performed. Heterogeneity was evaluated using (1) Cochran's Q test for the summary estimates ($p < 0.05$ indicating heterogeneity) and (2) the Higgins inconsistency index (I^2) ($\geq 50\%$ indicating significant heterogeneity) [49, 50]. Funnel plots were used to visually assess publication bias, and Egger's test was used to determine statistical significance

($p < 0.10$ indicating significant bias) [51]. Meta-regression analysis was used for the unreliable measurement of TE, to explore potential heterogeneity causes. To this end, the following covariates were used: (a) the number of study population (< 100 vs. ≥ 100); (b) etiology (known chronic liver disease vs. others); (c) mean or median age (< 9 vs. ≥ 9 years old); (d) transducer (including or not a pediatric S probe); (e) presence of number of readers; and (f) readers' blinding to pathologic results.

One reviewer (D.W.K) performed all statistical analyses using "Meta" package R version 3.4.1 (R Foundation for Statistical Computing).

Results

Literature search and selection

Figure 1 presents a flowchart for our process of literature selection. A total of 1184 studies were obtained from Ovid-MEDLINE ($n = 405$) and EMBASE ($n = 779$), and 414 duplicates were excluded. Following titles and abstracts screening, further 693 studies were removed as follows: 48 review articles; 13 case reports/series; 262 conference abstracts; 14 studies were either a letter, an editorial, a comment, or a note; 339 studies were not in the field of interest; and 17 studies targeted adult populations. After a full-text review of the remaining 77 articles, 8 were excluded because of three case reports/series; two studies used other US elastography techniques (strain elastography); and three studies targeted adult populations. Ultimately, 69 studies satisfied our criteria.

Fig. 1 Flow diagram for study selection

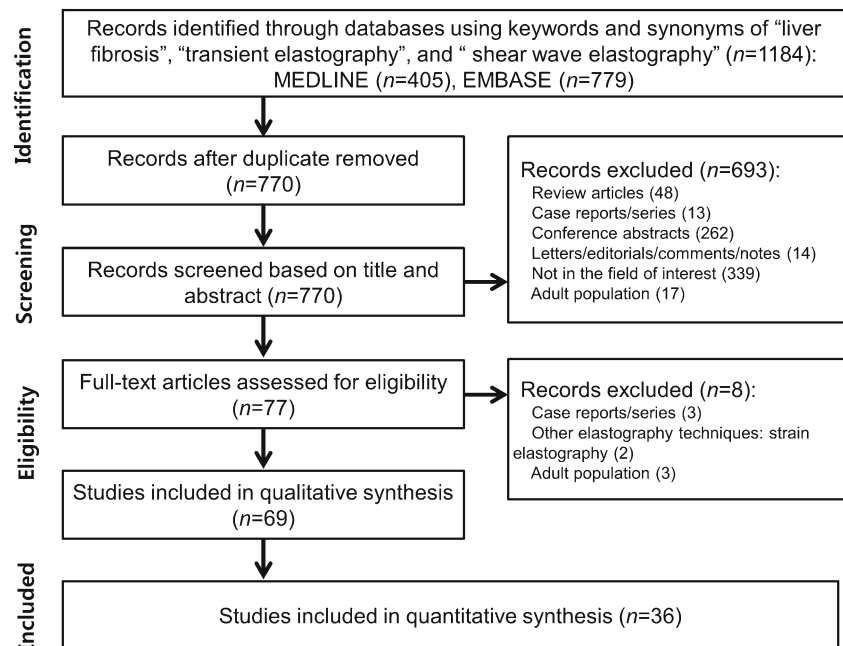


Table 1 Demographic characteristics for included study populations

Authors (year of publication)	Institution	Study period	Study design	Patients(<i>n</i>)	Mean age (range, years)	Male:female	Etiology
Alison M et al (2015) [7]	Robert Debré Children’s Hospital, France	2013.1–2013.9	Prospective	54	NA (0–21 days)	NA	Preterm neonates (< 36 weeks)
Alkhourri N et al (2013) [8]	Bambino Gesù’ Children’s Hospital, Italy	NA	Prospective	67	8.5 (5.5–11.3)	46:21	Biopsy-proven nonalcoholic fatty liver disease
Behrens CB et al (2013) [9]	University Hospital Giessen, Germany	2011.6–2011.12	Prospective	36	NA	19:17	Cystic fibrosis
Belei O et al (2016) [10]	“Louis Turcanu” Emergency Hospital & Emergency County Hospital Timisoara, Romania	NA	Prospective	54	8.6 (6–18)	NA	Chronic liver disease
Chen B et al (2016) [11]	British Columbia Children’s Hospital, Canada	2012.4–2014.10	NA	44	13.5 (5.7–17.7)	30:14	Patient who underwent the Fontan operation + healthy children
Cho Y et al (2015) [12]	Osaka City University Hospital, Japan	2013.4–2014.8	NA	214	NA (1.3–17.6)	121:93	Obese, liver disease, and healthy children
de Ledinghen V et al (2007) [13]	Hôpital Haut Lévéque, France	2004.2–2005.10	Prospective	116	NA	NA	Chronic liver disease
Dhiyani M et al (2015) [14]	Massachusetts General Hospital, USA	2010.6–2014.11	Prospective	24	17 (1–21)	13:11	Known or suspected diffuse liver disease
Dillman JR et al (2015) [15]	C. S. Mott Children’s Hospital, USA	2012.11–2014.7	Prospective	62	8.0 (4 weeks–18-years)	33:29	Known or suspected liver disease
Eiler J et al (2012) [16]	UKGM Giessen, Germany	2009.9–2011.11	NA	876	NA	NA	Healthy children
Engelmann G et al (2012) [17]	Heidelberg University Hospital, Germany	2009.9–2011.3	Prospective	975	9.1 (NA)	1:1*	NA
Fitzpatrick E et al (2013) [18]	King’s College Hospital, UK	2009.5–2011.3	Prospective	101	13.6 (median, 6–18)	62:39	Chronic liver disease
Fontanilla T et al (2014) [19]	Hospital Universitario Puerta de Hierro Majadahonda, Spain	2012.2–2012.5	Prospective	60	NA (1 day–14 years)	24:36	Healthy children
Franchi-Abella S et al (2016) [20]	Hôpital Bicêtre, France	NA	Prospective	96	NA	44:52	Liver disease + healthy children
Friedrich-Rust M et al (2008) [21]	Saarland University Hospital, Germany	2006	Prospective	39	11.6 (3–23)	20:19	Patients who underwent the Fontan operation
Garcovich M et al (2017) [22]	Gemelli Hospital, Italy	2015.5–2015.11	Prospective	69	12.6 (8–17)	37:32	Nonalcoholic steatohepatitis
Goldschmidt I et al (2013) [23]	Hannover Medical School, Germany	2009.2–2011.9	NA	117	8.9 (median, 0.4–18.5)	60:57	Liver-transplanted children
Goldschmidt I et al (2013) [24]	Hannover Medical School, Germany	2009.2–2011.7	NA	527	6 (median, 0.1–17.8)	298:229	NA
Hamidieh AA et al (2014) [25]	Shariati Hospital, Iran	2010.10–2012.2	NA	83	8 (median, 5–11 [IQR])	48:35	Major thalassaemia
Hanquinet S et al (2013) [26]	University Children’s Hospital of Geneva, Switzerland	NA	Prospective	109	NA	NA	Healthy children
Hukkinen M et al (2016) [27]	Helsinki University Hospital, Finland	2011–2015	Prospective	52	NA	NA	Intestinal failure
Kummer S et al (2017) [28]		2012.7–2014.8	Prospective	93	11.9 (3–18)	46:47	Type 1 diabetes mellitus

Table 1 (continued)

Authors (year of publication)	Institution	Study period	Study design	Patients(<i>n</i>)	Mean age (range, years)	Male:female	Etiology
Lee MJ et al (2013) [29]	University Hospital of Düsseldorf, Germany	2011.7–2011.8	Prospective	202	8.1 (NA)	92:110	Healthy children
Lewindon PJ et al (2016) [30]	Severance Children's Hospital, Republic of Korea	2011.4–2013.3	NA	173	8.8 (0.3–18)	NA	NA
Malbrunot-Wagner AC et al (2011) [31]	Royal Children's Hospital, Australia	2005.7–2010.2	Retrospective	18	14 (median, 9–18)	11:7	Cystic fibrosis
Nobili V et al (2008) [32]	Multicenter study, France	2007.7–2008.1	NA	52	13.6 (4–17)	32:20	Nonalcoholic steatohepatitis
Pawar SV et al (2016) [33]	Bambino Gesù Children's Hospital, Italy	NA	Prospective	100	NA (11–15)	NA	Obese or overweight high school students
Rubio A et al (2009) [34]	BYL Nair Hospital, India	NA	NA	21	NA	NA	HIV-1 chronically infected children
Sagir et al (2015) [35]	University Hospital of Nice, France	NA	Prospective	198	8 (6 weeks–18 years)	116:82	Chronic liver disease, metabolic disease, heart
Schenk JP et al (2014) [36]	University Hospital of Düsseldorf, Germany	2010.10–2012.3	NA	147	10.4 (NA)	80:67	disease, diseases without hepatic involvement, others
Shin HJ et al (2016) [37]	Severance Children's Hospital, Republic of Korea	2015.3–2016.8	Prospective	88	NA	45:43	Liver disease
Shin NY et al (2014) [38]	Severance Children's Hospital, Republic of Korea	2007.4–2010.7	Retrospective	51	NA	NA	Biliary atresia with Kasai operation + healthy children
Trout AT et al (2016) [39]	Cincinnati Children's Hospital Medical Center, USA	2014.7–2015.11	Prospective	55	14 (3.5–21.4)	31:24	Biliary atresia
Voutilainen S et al (2016) [40]	Helsinki University Hospital, Finland	2012.1–2015.3	Prospective	130	NA	NA	Chronic liver disease
Yoon HM et al (2017) [41]	Asan Medical Center, Republic of Korea	2012.1–2015.10	Retrospective	33	9.7 (3 months–16 years)	19:14	Abnormal liver function tests, 19; liver transplantation work-up, 3;
Zhou LY et al (2017) [42]	The First Affiliated Hospital of Sun Yat-sen University, China	2012.1–2015.5	Prospective	293	NA	NA	suspected Wilson's disease, 2; hyperbilirubinemia, 2; miscellaneous, 6
							Conjugated hyperbilirubinemia + healthy children

*Only sex ratio was reported

IQR, interquartile range, NA not available

Table 2 Technical characteristics of shear wave elastography used in included studies

Authors (year of publication)	Technique	Elastography system	Probe	Measurement number	Representative values	Reader	Blinding
Alison M et al (2015) [7]	2D-SWE	Aixplorer	L10-2	6	Mean	2 pediatric radiologists	Yes
Alkhoury N et al (2013) [8]	TE	FibroScan	S	10	Median	NA	NA
Behrens CB et al (2013) [9]	pSWE	Acuson S2000	4.5 MHz curved array	8	Median	3 radiologists	Yes
Belei O et al (2016) [10]	TE; pSWE; 2D-SWE	FibroScan; Acuson S2000; Aixplorer	M, XL; 4C1; SC6-1	10; 10; 5	Median	2 operators	NA
Chen B et al (2016) [11]	TE	FibroScan	S, M	10	Median	3 trained operators	NA
Cho Y et al (2015) [12]	TE	FibroScan	M	10	Median	NA	NA
de Ledinghen V et al (2007) [13]	TE	FibroScan	M	5–10	Median	NA	NA
Dhyani M et al (2015) [14]	2D-SWE	Aixplorer	SC6-1	NA	Median	1 sonographer	Yes
Dillman JR et al (2015) [15]	pSWE; 2D-SWE	Acuson S3000;	9L4; 9L4	6; 6	Mean, median, maximum	NA	NA
Eiler J et al (2012) [16]	pSWE	Acuson S3000	4C1	6	Mean	NA	NA
Engelmann G et al (2012) [17]	TE	FibroScan	S, M	10	Median	2 observers	NA
Fitzpatrick E et al (2013) [18]	TE	FibroScan	M	10	Median	NA	Yes
Fontanilla T et al (2014) [19]	pSWE	Acuson S2000	4C1, 9L4	20	Mean	NA	NA
Franchi-Abella S et al (2016) [20]	2D-SWE	Aixplorer	SL15-4, SC6-1	10	Mean	1 radiologist	NA
Friedrich-Rust M et al (2008) [21]	TE	FibroScan	NA	10	Median	NA	No
Garcovich M et al (2017) [22]	2D-SWE	Aixplorer	SC6-1	4	Mean	1 or 2 operators	Yes
Goldschmidt I et al (2013) [23]	TE	FibroScan	S, M	10	Median	NA	NA
Goldschmidt I et al (2013) [24]	TE	FibroScan	S, M	10 × 4 points	Median	NA	NA
Hamidi AA et al (2014) [25]	TE	FibroScan	S, M	10	Median	1 operator	Yes
Hanquinet S et al (2013) [26]	pSWE	Acuson S2000	4C1, 9L4	5	Mean	2 radiologists	NA
Hukkinen M et al (2016) [27]	TE	FibroScan	NA	10	Median	NA	NA
Kummer S et al (2017) [28]	TE; pSWE	FibroScan;	S, M; 6C1HD	10; 10	Median	NA	NA
Lee MJ et al (2013) [29]	pSWE	Acuson S2000	4C1, 9L4	3	Mean	1 pediatric radiologist	NA
Lewindon PJ et al (2016) [30]	TE	FibroScan	S, M	10	Median	NA	NA
Malbrunot-Wagner AC et al (2011) [31]	TE	FibroScan	M	10	Median	NA	NA
Nobili V et al (2008) [32]	TE	FibroScan	M	10	Median	2 operators	Yes
Pawar SV et al (2016) [33]	TE	FibroScan	M	10	Median	NA	NA
Rubio A et al (2009) [34]	TE	FibroScan	NA	10	Median	1	NA
Sagir et al (2015) [35]	TE; pSWE	FibroScan;	S, M; 6C1HD	10; 10	Median; mean	NA	NA
Schenk JP et al (2014) [36]	TE	Acuson S2000	S, M	10	Mean	NA	NA
Shin HJ et al (2016) [37]	2D-SWE	Aixplorer	SC6-1	15–30	Mean	2 pediatric radiologists	No
Shin NY et al (2014) [38]	TE	FibroScan	S, M	5	Median	NA	NA

Table 2 (continued)

Authors (year of publication)	Technique	Elastography system	Probe	Measurement number	Representative values	Reader	Blinding
Trout AT et al (2016) [39]	pSWE	Acuson S3000	6C1	10	Median	4 pediatric sonographers	No
Voutilainen S et al (2016) [40]	TE	FibroScan	S, M	10	Median	NA	NA
Yoon HM et al (2017) [41]	2D-SWE	Aixplorer	SC6-1	≥ 5	Median	2 radiologists	No
Zhou LY et al (2017) [42]	2D-SWE	Aixplorer	SL15-4	3	Mean	1	Yes

NA not available, pSWE point shear wave elastography, 2D-SWE two-dimensional shear wave elastography, TE transient elastography

Of the 69 studies, 29 (42.0%) lacked technical performance information. Furthermore, we were not able to obtain the rate of technical failure or unreliable measurements as per-patient level in four studies because they did not report the rate of successfully measured patients but only reported the rate of successful measurements among all measurements (per-measurement level; see [Supplementary References](#)). Therefore, our study included 36 articles [7–42] for quantitative synthesis.

Characteristics of included studies

Table 1 summarizes the demographic characteristics for the included study populations. Twenty-three [7–10, 13–15, 17–22, 26–29, 33, 35, 37, 39, 40, 42] were prospectively designed, and 3 [31, 38, 41] were retrospectively designed. The age of patients included in the studies ranged from newborn to 23 years. Table 2 summarizes the technical characteristics of the SWE in the included studies. Considering the techniques of SWE, 22 [8, 10–13, 17, 18, 23–25, 27, 28, 30–36, 38, 40], 10 [9, 10, 15, 16, 19, 26, 28, 29, 35, 39], and 9 studies [7, 10, 14, 15, 20, 22, 37, 41, 42] were performed using TE, pSWE, and 2D-SWE, respectively. One study included the three techniques [10], two studies included TE and pSWE [28, 35], and one study included pSWE and 2D-SWE [15]. Regarding the type of transducers for TE, 12 studies [8, 11, 17, 23–25, 28, 30, 35, 36, 38, 40] used a standard M and a pediatric S probe; 7 studies [10, 12, 13, 18, 31–33] did not use the pediatric S probe.

Systematic review of technical performance: quality of reporting

Regarding the TE studies, technical failure was reported in 3 of 22 studies (14%) [24, 33, 34] and unreliable measurement was reported in 21 of 22 studies (95%) [8, 10–13, 17, 18, 21, 23–25, 27, 28, 30–33, 35, 36, 38, 40]. Rates of technical failure ranged from 1.0 to 9.5%, and rate of unreliable measurements ranged from 0 to 28.9% (Supplementary Table 1). In the pSWE studies, technical failure was reported in 6 of 10 studies (60%) [15, 16, 19, 26, 29, 39] and unreliable measurement was reported in 4 of 10 studies (40%) [9, 10, 28, 35]. Rates of technical failure ranged from 0 to 21.0%. One of four studies [28] reported a rate of unreliable measurements of 5.4% (Supplementary Table 2); the other three studies [9, 10, 35] reported 0% as the rate of unreliable measurements. 2D-SWE rates of technical failure were reported in eight of nine studies (89%) [7, 10, 14, 15, 20, 22, 37, 42] and ranged from 0 to 22.6%. Indeed, no technical failure was found in five [7, 10, 14, 20, 37] of those eight studies. The unreliable measurement of 2D-SWE was reported in one out of nine studies [41] with rate of unreliable measurements of 3.1% (Supplementary Table 3). Table 3 summarizes each study's criterion for the reliable measurement. In terms of the unreliable measurement of TE, 16 [8, 10, 12, 17, 18, 23–25, 27, 28,

Table 3 Criteria for reliable measurement in included studies

Authors (year of publication)	Criteria
a. Transient elastography	
Alkhoury N et al (2013) [8]	Recommended criteria*
Belei O et al (2016) [10]	Recommended criteria*
Chen B et al (2016) [11]	IQR/median value < 30%
Cho Y et al (2015) [12]	Recommended criteria*
de Ledinghen V et al (2007) [13]	≥ 5 valid measurements
Engelmann G et al (2012) [17]	Recommended criteria*
Fitzpatrick E et al (2013) [18]	Recommended criteria*
Friedrich-Rust M et al (2008) [21]	10 valid measurements + a success rate ≥ 60%
Goldschmidt I et al (2013) [23]	Recommended criteria*
Goldschmidt I et al (2013) [24]	Recommended criteria*
Hamidieh AA et al (2014) [25]	Recommended criteria*
Hukkinen M et al (2016) [27]	Recommended criteria*
Kummer S et al (2017) [28]	Recommended criteria*
Lewindon PJ et al (2016) [30]	10 valid measurements + IQR/median value < 30%
Malbrunot-Wagner AC et al (2011) [31]	Recommended criteria*
Nobili V et al (2008) [32]	Recommended criteria*
Pawar SV et al (2016) [33]	Recommended criteria*
Sagir et al (2015) [35]	Recommended criteria*
Schenk JP et al (2014) [36]	Recommended criteria*
Shin NY et al (2014) [38]	≥ 5 valid measurements
Voutilainen S et al (2016) [40]	Recommended criteria*
b. Point shear wave elastography	
Behrens CB et al (2013) [9]	A success rate ≥ 60% + IQR/median value < 30%
Belei O et al (2016) [10]	A success rate ≥ 60%
Kummer S et al (2017) [28]	10 valid measurements + a success rate ≥ 60% + IQR/median value < 30%
Sagir et al (2015) [35]	A success rate ≥ 60% + IQR/median value < 30%
c. Two-dimensional wave elastography	
Yoon HM et al (2017) [41]	SD/mean value ≤ 0.2

IQR interquartile range, SD standard deviation

*Recommended criteria refer to the validated criteria recommended by the manufacturer as follows: (a) 10 valid measurements; (b) a success rate (valid shots/total number of shots) ≥ 60%; (c) interquartile range (IQR) < 30% of median liver stiffness value

31–33, 35, 36, 40] of the 21 studies followed the reliability criteria recommended by the manufacturer.

Meta-analysis

Meta-analysis was performed for the parameters consisting of more than five studies, including the unreliable measurements of TE, technical failure during pSWE, and technical failure during 2D-SWE.

Meta-analysis was performed for the unreliable measurement of TE from 16 studies [8, 10, 12, 17, 18, 23–25, 27, 28, 31–33, 35, 36, 40] with 3030 patients, which met the manufacturer's recommended criteria. A total of 397 unreliable

measurements occurred. Using a random effects model, the observed pooled proportion was 12.1% (95% CI, 9.4–15.5%) (Fig. 2). Significant heterogeneity was found in Cochran's *Q* test ($p < 0.01$) and Higgin's I^2 statistic ($I^2 = 79\%$). Significant publication bias was observed by funnel and Egger's tests ($p = 0.06$) (Supplementary Figure 1).

Use of a random effects model (Fig. 3) to analyze in 59 events from 1364 patients showed that the pooled proportion of technical failure during pSWE was 4.1% (95% CI, 1.5–10.7%). Furthermore, significant heterogeneity was observed (Cochran's *Q* test, $p < 0.01$; Higgin's I^2 statistic, $I^2 = 85\%$).

A pooled proportion of technical failure during 2D-SWE was 2.2% (95% CI, 0.6–7.6%) using a random effects model

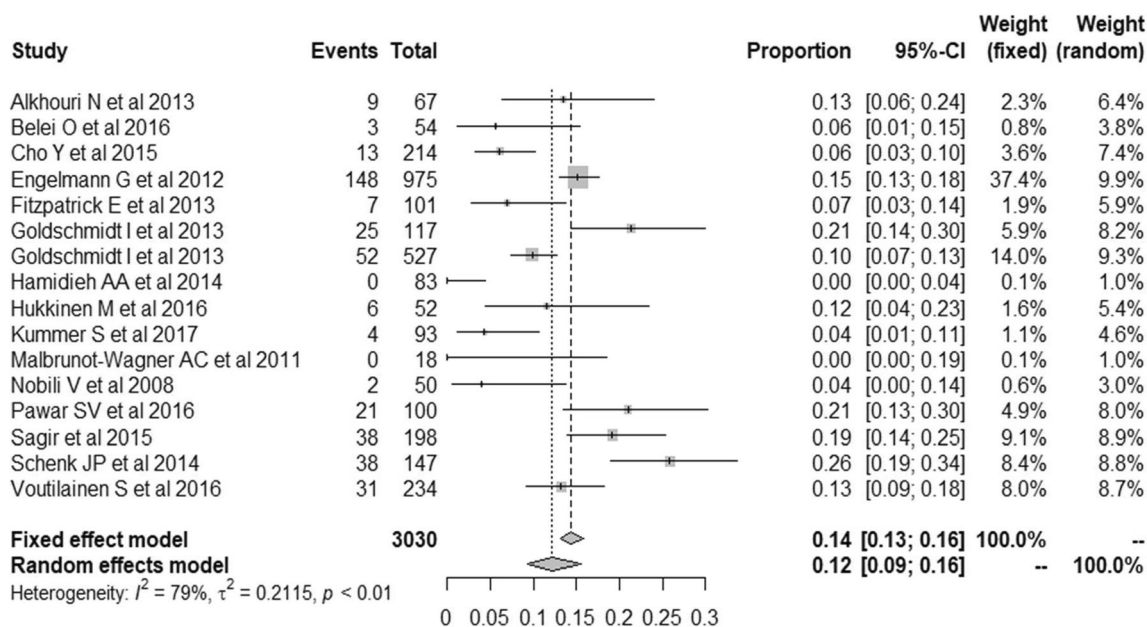


Fig. 2 Forest plots for pooled proportions of unreliable TE measurements

(Fig. 4) with significant heterogeneity (Cochran’s Q test, $p < 0.01$; Higgin’s I^2 statistic, $I^2 = 83\%$).

Comparison of technical failure rates between pSWE and 2D-SWE revealed a lower pooled proportion of technical failure during 2D-SWE than pSWE, albeit statistically not significant ($p = 0.61$).

Meta-regression analysis

Study heterogeneity was influenced by study population’s number ($p = 0.01$) and readers’ blinding to pathologic results ($p = 0.02$) in the meta-regression analysis for the unreliable measurement of TE. Other covariates including etiology ($p = 0.92$), age ($p = 0.63$), transducer ($p = 0.06$), and presence of number of readers ($p = 0.19$) did not influence the rate of unreliable measurements (Supplementary Table 4).

Discussion

According to our systematic review, 40 out of 69 studies investigating SWE in pediatric and adolescent patients provided technical performance information, especially the rate of unreliable measurements from TE studies and rates of technical failure from pSWE and 2D-SWE studies. In brief, the pooled proportion of unreliable measurements of TE was 12.1%, and the pooled proportion of technical failure during 2D-SWE tended to be lower than the pooled proportion of technical failure during pSWE (2.2% vs. 4%, $p = 0.61$).

The US Food and Drug Administration approved US elastography as a commercially available diagnostic device. However, US elastography remains to be validated by the clinical community as a quantitative biomarker for measurement of liver fibrosis in clinical practice. As to the clinical validation, technical success assessment and variability measurement are as important as the diagnostic performance

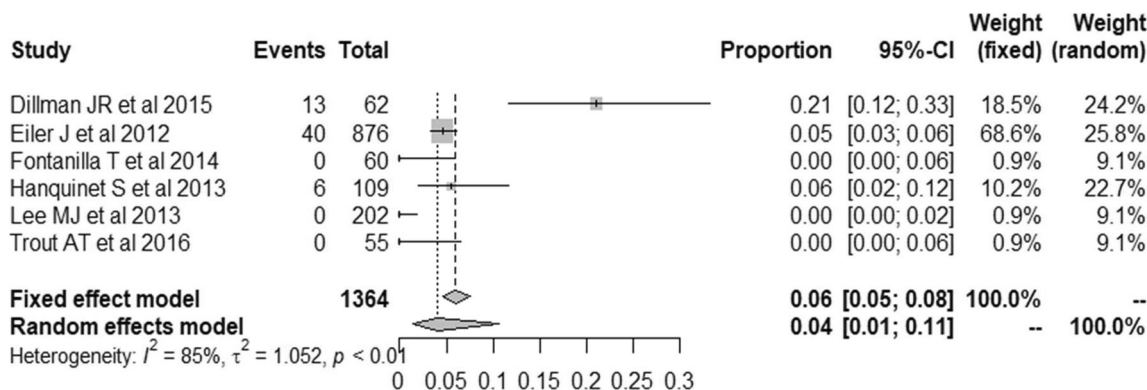


Fig. 3 Forest plots for pooled proportions of technical failure during pSWE

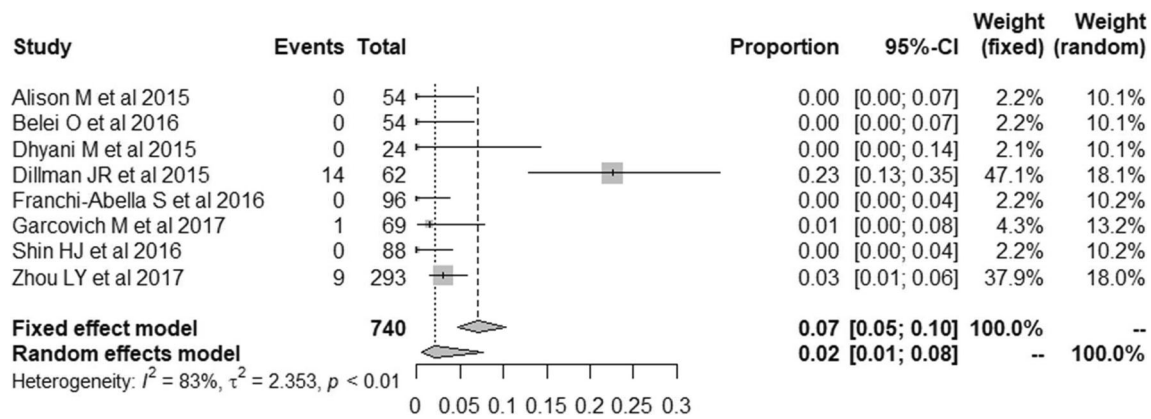


Fig. 4 Forest plots for pooled proportions of technical failure during 2D-SWE

assessment. However, we observed that 29 of 69 studies (42.0%) overlooked the technical performance, a considerable number, given its importance.

The contrast in technical performance reporting among techniques is likely linked to the standardization of reliability criteria. For TE, a manufacturer recommends criteria for reliable measurements, which are widely acceptable across the studies. Specifically, 10 successful acquisitions with a success rate $\geq 60\%$ and an IQR/median value $< 30\%$ are considered reliable. Unlike TE, standard reliability criteria are not available for pSWE and 2D-SWE. Therefore, only few pSWE and 2D-SWE studies reported the unreliable measurement with variable reliability criteria. Rather, they provided technical failure assessment. To this end, collaborative efforts between academia and industry to reach a consensus and standardize SWE measurements are currently underway. The Society of Radiologists in Ultrasound recently published a consensus statement on the technical aspects of US elastography, suggesting 10 measurements covering the same hepatic location with an IQR of 30% or less of the median value [4]. The World Federation for Ultrasound in Medicine and Biology recommends 5 to 10 measurements for pSWE and 4 measurements for 2D-SWE [1]. Recently, the 2017 European Federation of Societies for Ultrasound in Medicine and Biology guidelines recommend that 3 measurements for 2D-SWE are sufficient to obtain consistent results [52]. Of note, numerous measurements can lead to extended examination times, making young patients less tolerable [37]. Thus, conclusions that are valid for adult populations may not necessarily be extrapolated to pediatric and adolescent populations. In this regard, Shin et al [37] evaluated the optimal 2D-SWE acquisition number. They observed that three measurements are sufficient to measure liver stiffness in children over 6 years old. However, future studies are necessary to define optimal reliability criteria for pSWE and 2D-SWE and to determine the US elastography technique of choice in pediatric and adolescent population.

Varying technical principles among SWE techniques likely influence technical performance differences. We therefore

assumed that pSWE and 2D-SWE would have better technical performance than TE given its advantage in technical principles [3]. First, as opposed to TE, the ROI can be guided by conventional B-mode US while avoiding structures such as vessels, gallbladder, or focal lesions. Second, pSWE and 2D-SWE generate shear wave in the tissue of interest by ARFI, whereas TE applies a mechanical impulse to the skin surface to generate shear wave and the propagation of this shear wave is tracked. Thus, structures interposed between the skin surface and liver capsule (e.g., perihepatic ascites or fat) have a smaller impact on measuring liver stiffness in pSWE & 2D-SWE than TE. Last, color elasticity maps of 2D-SWE help avoid artifacts induced by the stress which tends to concentrate on the region near the boundary [42, 53].

Similarly to adults, pediatric and adolescent high BMI or obesity is associated with increased technical failure and unreliable measurement, due to excessive thickness of the subcutaneous adipose tissue [12, 15]. Moreover, in children, additional factors can affect the likelihood of technical failure and/or unreliable measurements. First, children with narrow intercostal space might have higher failure rate when measuring liver stiffness with a standard adult probe. For small children, the use of a pediatric S probe (frequency, 5 MHz; probe diameter, 5 mm) in TE can overcome the problem of a standard M probe (frequency, 3.5 MHz; probe diameter, 7 mm), especially for those children with a thorax perimeter smaller than 75 cm [10, 17, 24, 38]. Specifically, two different modes of S probe, S1 and S2, are available, with a measurement depth of 1.5–4 cm in S1 mode and 2–5 cm in S2 mode [36]. Second, children are less cooperative during an US elastography. Continuous movement or crying of patients causes a less accurate measurement [12, 14]. Therefore, young children (<24 months) who are vulnerable to agitation have lower success rates [17]. In this case, general anesthesia would reduce excessive movement and increase success rates, albeit its invasiveness. Last, breath holding can facilitate the obtainment of valid measurements. However, in young children, such practice is difficult if not impossible. So far, mixed

reports are available on whether breath holding affects successful measurements in pediatric patients. In fact, some studies showed that variability caused by breathing was not significant [19, 29, 37]. They rather worried that breath holding in children would cause an irregular and variable breathing rhythm, determining invalid measurements [19]. However, a recent study showed decreased liver stiffness with the free-breathing technique compared with the breath-holding technique in 2D-SWE [54]. Therefore, we performed the meta-regression over the factor of the age and the S probe usage. However, we observed that they did not affect study heterogeneity. Rather, studies including the S probe showed less reliable measurements than those not including it, without statistical significance (14.5% vs. 7.6%, $p = 0.06$). These conflicting results might depend on the heterogeneity of the population included in our study. Thus, future large-scale study allows the evaluation of factors influencing successful measurements in pediatric patients.

One limitation of our study is the small number of studies included. Therefore, for some SWE techniques, we were unable to perform a meta-analysis. Additionally, we were unable to compare rates of technical failure or unreliable measurement among techniques. Importantly though, we included all the available studies and overviewed the current evidence on aspects of the technical performance in pediatric population. An additional limitation is the significant heterogeneity of the available meta-analysis. Although the number of study population and reader blinding to the reference standard affected heterogeneity, other patient or technical factors may influence it. For example, the placement of ROI of pSWE and 2D-SWE may be user dependent that sufficient experience is required for obtaining consistent measurements [55]. Thus, a large-cohort study is warranted to evaluate factors that may affect technical performance.

In summary, TE studies seldom reported rates of technical failure, but rather reported rate of unreliable measurements with the pooled proportion of 12.1%. Conversely, pSWE and 2D-SWE studies often reported rates of technical failure showing comparable results, but rarely reported measurements of unreliability. Considering the importance of technical performance for clinical validation of SWE, the number of and reasons for technical failure and unreliable measurements should be reported in future studies. Additionally, further efforts are necessary to standardize the reliability criteria for SWE.

Funding This study was supported by a grant (2016-719) from the Asan Medical Center, Seoul, Korea.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Jin Seong Lee.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry An expert (Chong Hyun Suh), who has 5 years of experience with more than 30 papers in a systematic review and meta-analysis, assisted in statistics for our study.

Informed consent Written informed consent was not required for this study because of the nature of our study, which was a systemic review and meta-analysis.

Ethical approval Institutional Review Board approval was not required because of the nature of our study, which was a systemic review and meta-analysis.

Methodology

- A systemic review and meta-analysis performed at one institution

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Bamber J, Cosgrove D, Dietrich CF et al (2013) EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. Part 1: basic principles and technology. *Ultraschall Med* 34:169–184
2. Park SH, Kim SY, Suh CH et al (2016) What we need to know when performing and interpreting US elastography. *Clin Mol Hepatol* 22:406–414
3. Kennedy P, Wagner M, Castera L et al (2018) Quantitative elastography methods in liver disease: current evidence and future directions. *Radiology* 286:738–763
4. Barr RG, Ferraioli G, Palmeri ML et al (2015) Elastography assessment of liver fibrosis: society of radiologists in ultrasound consensus conference statement. *Radiology* 276:845–861
5. Ferraioli G, Filice C, Castera L et al (2015) WFUMB guidelines and recommendations for clinical use of ultrasound elastography: part 3: liver. *Ultrasound Med Biol* 41:1161–1179
6. European Association for Study of Liver; Asociacion Latinoamericana para el Estudio del Hgado (2015) EASL-ALEH clinical practice guidelines: non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol* 63:237–264
7. Alison M, Biran V, Tanase A et al (2015) Quantitative shear-wave elastography of the liver in preterm neonates with intra-uterine growth restriction. *PLoS One* 10:e0143220
8. Alkhoury N, Sedki E, Alisi A et al (2013) Combined paediatric NAFLD fibrosis index and transient elastography to predict clinically significant fibrosis in children with fatty liver disease. *Liver Int* 33:79–85
9. Behrens CB, Langholz JH, Eiler J et al (2013) A pilot study of the characterization of hepatic tissue strain in children with cystic-fibrosis-associated liver disease (CFLD) by acoustic radiation force impulse imaging. *Pediatr Radiol* 43:552–557
10. Belei O, Sporea I, Gradinaru-Tascau O et al (2016) Comparison of three ultrasound based elastographic techniques in children and adolescents with chronic diffuse liver diseases. *Med Ultrason* 18:145–150
11. Chen B, Schreiber RA, Human DG, Potts JE, Guttman OR (2016) Assessment of liver stiffness in pediatric Fontan patients using transient elastography. *Can J Gastroenterol Hepatol* 2016:7125193

12. Cho Y, Tokuhara D, Morikawa H et al (2015) Transient elastography-based liver profiles in a hospital-based pediatric population in Japan. *PLoS One* 10:e0137239
13. de Lédinghen V, Le Bail B, Rebouissoux L et al (2007) Liver stiffness measurement in children using fibroscan: feasibility study and comparison with fibrotest, aspartate transaminase to platelets ratio index, and liver biopsy. *J Pediatr Gastroenterol Nutr* 45:443–450
14. Dhyani M, Gee MS, Misdraji J, Israel EJ, Shah U, Samir AE (2015) Feasibility study for assessing liver fibrosis in paediatric and adolescent patients using real-time shear wave elastography. *J Med Imaging Radiat Oncol* 59:687–694 quiz 751
15. Dillman JR, Heider A, Bilhartz JL et al (2015) Ultrasound shear wave speed measurements correlate with liver fibrosis in children. *Pediatr Radiol* 45:1480–1488
16. Eiler J, Kleinholdermann U, Albers D et al (2012) Standard value of ultrasound elastography using acoustic radiation force impulse imaging (ARFI) in healthy liver tissue of children and adolescents. *Ultraschall Med* 33:474–479
17. Engelman G, Gebhardt C, Wenning D et al (2012) Feasibility study and control values of transient elastography in healthy children. *Eur J Pediatr* 171:353–360
18. Fitzpatrick E, Quaglia A, Vimalasvaran S, Basso MS, Dhawan A (2013) Transient elastography is a useful noninvasive tool for the evaluation of fibrosis in paediatric chronic liver disease. *J Pediatr Gastroenterol Nutr* 56:72–76
19. Fontanilla T, Cañas T, Macia A et al (2014) Normal values of liver shear wave velocity in healthy children assessed by acoustic radiation force impulse imaging using a convex probe and a linear probe. *Ultrasound Med Biol* 40:470–477
20. Franchi-Abella S, Como L, Gonzales E et al (2016) Feasibility and diagnostic accuracy of supersonic shear-wave elastography for the assessment of liver stiffness and liver fibrosis in children: a pilot study of 96 patients. *Radiology* 278:554–562
21. Friedrich-Rust M, Koch C, Rentzsch A et al (2008) Noninvasive assessment of liver fibrosis in patients with Fontan circulation using transient elastography and biochemical fibrosis markers. *J Thorac Cardiovasc Surg* 135:560–567
22. Garcovich M, Veraldi S, Di Stasio E et al (2017) Liver stiffness in pediatric patients with fatty liver disease: diagnostic accuracy and reproducibility of shear-wave elastography. *Radiology* 283:820–827
23. Goldschmidt I, Stieghorst H, Munteanu M et al (2013) The use of transient elastography and non-invasive serum markers of fibrosis in pediatric liver transplant recipients. *Pediatr Transplant* 17:525–534
24. Goldschmidt I, Streckenbach C, Dingemann C et al (2013) Application and limitations of transient liver elastography in children. *J Pediatr Gastroenterol Nutr* 57:109–113
25. Hamidieh AA, Shazad B, Ostovaneh MR et al (2014) Noninvasive measurement of liver fibrosis using transient elastography in pediatric patients with major thalassemia who are candidates for hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant* 20:1912–1917
26. Hanquinet S, Courvoisier D, Kanavaki A, Dhoubi A, Anooshiravani M (2013) Acoustic radiation force impulse imaging-normal values of liver stiffness in healthy children. *Pediatr Radiol* 43:539–544
27. Hukkinen M, Kivisaari R, Lohi J et al (2016) Transient elastography and aspartate aminotransferase to platelet ratio predict liver injury in paediatric intestinal failure. *Liver Int* 36:361–369
28. Kummer S, Klee D, Kircheis G et al (2017) Screening for non-alcoholic fatty liver disease in children and adolescents with type 1 diabetes mellitus: a cross-sectional analysis. *Eur J Pediatr* 176:529–536
29. Lee MJ, Kim MJ, Han KH, Yoon CS (2013) Age-related changes in liver, kidney, and spleen stiffness in healthy children measured with acoustic radiation force impulse imaging. *Eur J Radiol* 82:e290–e294
30. Lewindon PJ, Balouch F, Pereira TN et al (2016) Transient liver elastography in unselected control children: impact of age and intercurrent illness. *J Paediatr Child Health* 52:637–642
31. Malbrunot-Wagner AC, Bridoux L, Nousbaum JB et al (2011) Transient elastography and portal hypertension in pediatric patients with cystic fibrosis Transient elastography and cystic fibrosis. *J Cyst Fibros* 10:338–342
32. Nobili V, Vizzutti F, Arena U et al (2008) Accuracy and reproducibility of transient elastography for the diagnosis of fibrosis in pediatric nonalcoholic steatohepatitis. *Hepatology* 48:442–448
33. Pawar SV, Zanwar VG, Choksey AS et al (2016) Most overweight and obese Indian children have nonalcoholic fatty liver disease. *Ann Hepatol* 15:853–861
34. Rubio A, Monpoux F, Huguon E et al (2009) Noninvasive procedures to evaluate liver involvement in HIV-1 vertically infected children. *J Pediatr Gastroenterol Nutr* 49:599–606
35. Sagir A, Ney D, Oh J et al (2015) Evaluation of acoustic radiation force impulse imaging (ARFI) for the determination of liver stiffness using transient elastography as a reference in children. *Ultrasound Int Open* 1:E2–E7
36. Schenk JP, Alzen G, Klingmüller V et al (2014) Measurement of real-time tissue elastography in a phantom model and comparison with transient elastography in pediatric patients with liver diseases. *Diagn Interv Radiol* 20:90–99
37. Shin HJ, Kim MJ, Kim HY, Roh YH, Lee MJ (2016) Optimal acquisition number for hepatic shear wave velocity measurements in children. *PLoS One* 11:e0168758
38. Shin NY, Kim MJ, Lee MJ et al (2014) Transient elastography and sonography for prediction of liver fibrosis in infants with biliary atresia. *J Ultrasound Med* 33:853–864
39. Trout AT, Dillman JR, Xanthakos S et al (2016) Prospective assessment of correlation between US acoustic radiation force impulse and MR elastography in a pediatric population: dispersion of US shear-wave speed measurement matters. *Radiology* 281:544–552
40. Voutilainen S, Kivisaari R, Lohi J, Jalanko H, Pakarinen MP (2016) A prospective comparison of noninvasive methods in the assessment of liver fibrosis and esophageal varices in pediatric chronic liver diseases. *J Clin Gastroenterol* 50:658–663
41. Yoon HM, Kim SY, Kim KM et al (2017) Liver stiffness measured by shear-wave elastography for evaluating intrahepatic portal hypertension in children. *J Pediatr Gastroenterol Nutr* 64:892–897
42. Zhou LY, Jiang H, Shan QY et al (2017) Liver stiffness measurements with supersonic shear wave elastography in the diagnosis of biliary atresia: a comparative study with grey-scale US. *Eur Radiol* 27:3474–3484
43. Castera L, Foucher J, Bernard PH et al (2010) Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* 51:828–835
44. Liberati A, Altman DG, Tetzlaff J et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 151:W65–W94
45. Castera L, Forns X, Alberti A (2008) Non-invasive evaluation of liver fibrosis using transient elastography. *J Hepatol* 48:835–847
46. Suh CH, Park SH (2016) Successful publication of systematic review and meta-analysis of studies evaluating diagnostic test accuracy. *Korean J Radiol* 17:5–6
47. Kim KW, Lee J, Choi SH, Huh J, Park SH (2015) Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part I. General guidance and tips. *Korean J Radiol* 16:1175–1187
48. Lee J, Kim KW, Choi SH, Huh J, Park SH (2015) Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part II. Statistical methods of meta-analysis. *Korean J Radiol* 16:1188–1196

49. Higgins JPT, Green S (eds) (2011) Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration. http://handbook.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm. Updated March 2011. Accessed 8 Jan 2017
50. Higgins JP, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. *Stat Med* 21:1539–1558
51. Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315:629–634
52. Dietrich CF, Bamber J, Berzigotti A et al (2017) EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version). *Ultraschall Med* 38:e48
53. Shiina T, Nightingale KR, Palmeri ML et al (2015) WFUMB guidelines and recommendations for clinical use of ultrasound elastography: part 1: basic principles and terminology. *Ultrasound Med Biol* 41:1126–1147
54. Yoon HM, Cho YA, Kim JR et al (2017) Real-time two-dimensional shear-wave elastography for liver stiffness in children: interobserver variation and effect of breathing technique. *Eur J Radiol* 97: 53–58
55. Ferraioli G, Tinelli C, Zicchetti M et al (2012) Reproducibility of real-time shear wave elastography in the evaluation of liver elasticity. *Eur J Radiol* 81:3102–3106