Ultrasound in Med. & Biol., Vol. 46, No. 8, pp. 1854–1864, 2020

0301-5629/$ - see front matter

https://doi.org/10.1016/j.ultrasmedbio.2020.03.025

● *Original Contribution*

LIVER ELASTOGRAPHY IN PRIMARY SCLEROSING CHOLANGITIS PATIENTS USING THREE DIFFERENT SCANNER SYSTEMS

ANDERS BATMAN MJELLE,[*,†] GURI FOSSDAL,[‡,§,¶] ODD HELGE GILJA,[*,†] and METTE VESTERHUS[†,‡,§,¶]

[*] Department of Clinical Medicine, University of Bergen, Bergen, Norway; [†] National Centre for Ultrasound in Gastroenterology, Department of Medicine, Haukeland University Hospital, Bergen, Norway; [‡] Department of Medicine, Haraldsplass Deaconess Hospital, Bergen, Norway; [§] Norwegian PSC Research Center (NoPSC), Division of Surgery, Inflammatory Medicine and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway; and [¶] Department of Clinical Science, University of Bergen, Bergen, Norway

(*Received* 16 *October* 2019; *revised* 17 *March* 2020; *in final from* 24 *March* 2020)

Abstract—The aim of the study described here was to characterize three different liver elastography methods in primary sclerosing cholangitis (PSC) patients, for the first time exploring 2-D shear wave elastography (2-D-SWE) in PSC patients and its putative advantages over point shear wave elastography (pSWE). Sixty-six adult PSC patients (51 males, 77%) underwent liver elastography: Transient elastography (TE), pSWE and 2-D-SWE were applied head-to-head after B-mode ultrasonography and blood tests. Liver stiffness measurements (LSMs) by pSWE yielded lower values than those by TE; 2-D-SWE had less steep slope but was overall not significantly different from TE. Correlation between LSMs by pSWE and TE was excellent (intraclass correlation coefficient = 0.92); correlation for 2-D-SWE with either pSWE or TE was moderate but improved with exclusion of overweight individuals. LSMs correlated with the Enhanced Liver Fibrosis test (ELF) across all scanner systems. Our study indicates that LSM by different systems is feasible in PSC patients and that 2-D-SWE tends to underestimate stiffness compared with TE. (E-mail addresses: abmjelle@gmail.com, adnj@helse-bergen.no)

Key Words: **Elastography, Liver fibrosis, Point shear wave elastography, Primary sclerosing cholangitis, Shear wave elastography, Transient elastography, Ultrasound.**

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a progressive fibroinflammatory disease affecting primarily the bile ducts, causing strictures and dilations and progressing over time through increasing stages of liver fibrosis and eventually cirrhosis. The natural history of PSC is notoriously unpredictable, with population-based studies reporting substantial variation in disease progression (Broome et al. 1996; Boonstra et al. 2013). Histologic disease stage is associated with prognosis in PSC but requires invasive biopsies and is flawed by sampling error and inter-observer variation.

Liver elastography has gained a significant role as a method for non-invasive evaluation of liver fibrosis in chronic liver diseases, enabling quantification of liver stiffness as a proxy for fibrosis with a high diagnostic accuracy compared with liver biopsy, which is considered the reference standard (Sandrin et al. 2003; Corpechot et al. 2006; Friedrich-Rust et al. 2008). Transient elastography (TE) has exhibited excellent ability to stratify between milder and severe stages of fibrosis in PSC compared with liver biopsy, clinical scoring systems and serologic markers and liver stiffness by TE was associated with clinical outcome in two independent studies (Corpechot et al. 2014; Ehlken et al. 2016a, 2016b). However, intermittent cholestasis caused by (dominant) strictures is common in PSC and has been reported to affect liver stiffness measurement (LSM) levels, thus constituting an important confounder (Millonig et al. 2008).

Liver elastography encompasses several technically different methods all based on the measurement of shear wave velocity, such as TE, point shear wave elastography (pSWE) and 2-D shear wave elastography (2-D-SWE). pSWE and 2-D-SWE allow simultaneous B-mode visualization of the liver, which may be

Address correspondence to: Anders Batman Mjelle, Department of Clinical Medicine, University of Bergen, Bergen, Norway.E-mail addresses: abmjelle@gmail.com, adnj@helse-bergen.no

particularly useful in PSC to exclude cholestasis as a confounder. Potentially, 2-D-SWE might have advantages because of visualization of an elastogram before LSM in PSC, owing to the patchy distribution of fibrosis. Previous studies in healthy controls of all ages (Mulabecirovic et al. 2018a, 2018b; Mjelle et al. 2019) and in patients with various liver diseases (Sporea et al. 2014, 2018; Piscaglia et al. 2017; Ferraioli et al. 2019; Iijima et al. 2019; Lefebvre et al. 2019) have revealed differences between methods with respect to both LSM levels and feasibility.

In PSC, data on liver elastography are scarce, particularly for 2-D-SWE and pSWE and head-to-head comparisons of methods. Thus, in this study, we aimed to perform a parallel assessment of three different scanner systems for LSM in PSC patients: TE, pSWE and 2-D-SWE.

METHODS

Patients

The study was performed at Haukeland University Hospital, Bergen, Norway, in 2017 and 2018. All patients were part of a well-characterized cohort of non-transplanted PSC patients. PSC was diagnosed according to acknowledged criteria. Data collection was performed prospectively as part of annual study visits consisting of patient history, clinical examination, blood tests and ultrasound investigation including liver elastography. Informed written consent was obtained from all participants.

B-Mode ultrasound examination

All patients were examined by B-mode ultrasound scanning of the liver and spleen before liver stiffness measurements. All examinations were performed by a single operator (M.V.) using a Philips iU22 (Philips Healthcare, Andover, MA, USA) scanner with software Version 6.3.2.2, and a C5-1 convex probe. Scores were registered for visual signs of liver fibrosis, including parenchyma heterogeneity, liver capsule regularity, liver angle, ascites, bile duct variability, sludge or gallbladder stones. Splenomegaly was defined as a spleen length $\geq$12 cm.

Liver stiffness measurements

LSM was measured using a right intercostal approach in fasting ($\geq$3 h) patients placed in a supine position, with the right hand resting under the head. All measurements were acquired in relaxed mid-ventilation breath hold with minimal probe pressure. pSWE using the Philips iU22 (ElastPQ, iU22, Philips Healthcare) was performed by a single operator (M.V.) using a convex C5-1 probe, followed head-to-head by examinations by another single operator (A.B.M.) using 2-D-SWE.GE (GE S8, GE Healthcare, Milwaukee, WI) with a C1-6 probe, and TE using Fibroscan incorporated into the GE S8 (GE Healthcare) using an M-probe, or XL-probe if the machine indicated that the M-probe was not suitable. 2-D-SWE.GE and TE results were given in kilopascals (kPa), and pSWE results in meters per second (m/s). The latter values were converted into kilopascals for comparison, using the equation kPa = 3 $(m/s)^2$. Operators were very experienced (M.V., many years of experience) or moderately experienced (A.B.M., >300 elastography measurements with 2-D-SWE.GE and pSWE); both were certified Fibroscan users. For both pSWE and 2-D-SWE.GE, a region of interest (for pSWE, a fixed region of 0.5 × 1.5 cm, and for 2-D-SWE.GE, a fixed circle with a diameter of 1 cm) was placed in a homogenous area 2−6 cm under the liver capsule, avoiding vessels and visible bile ducts. Quality criteria were applied according to the specific manufacturer's recommendations: For all systems, a valid measurement required a success rate $\geq$60%, and for 2-D-SWE.GE and TE valid measurements required an interquartile range divided by the median (IQR/M) $\leq$30%.

Enhanced liver fibrosis test

Serum samples were analyzed with the commercially available enhanced liver fibrosis (ELF) test (Siemens Medical Solutions Diagnostics Inc., Tarrytown, NY, USA), with essays performed with the Siemens ELF Test kits and an ADVIA Centaur XP analyzer (Siemens Medical Solutions Diagnostics Inc.).

Statistical analyses

For all analyses, SPSS Version 25 (IBM, Armonk, NY, USA) was used. All variables were tested for normality, and data were presented as the mean (standard deviation [SD]) or median (range) as appropriate. We applied Student's *t*-test or Mann−Whitney *U*-test as appropriate. Correlations were tested with Pearson's correlation or Spearman's rank correlation coefficient as appropriate. Degree of correlation was defined as poor (<0.40), moderate (0.40−0.69), good (0.70−0.89) or excellent ($\geq$0.9). p Values < 0.05 were considered to indicate statistical significance.

Ethical aspects

The protocol was in accordance with the Declaration of Helsinki and approved by the Regional Committee on Medical and Health Research Ethics of Western Norway (2012/2214/REK VEST).

RESULTS

We prospectively recruited and included 66 non-transplanted PSC patients (51 males, 77.3%) with a mean (SD) age of 49 y (16.3 y). Median (range) time since diagnosis was 8 y (0−37 y). Fifty-eight (78.8%) patients had inflammatory bowel disease (IBD) and 3 (4.5%) had overlapping features of autoimmune hepatitis. Generally, 64.1% of these patients had B-mode findings of liver pathology and/or bile duct pathology, while 25 (39.7%) patients had B-mode findings indicating liver fibrosis. Liver pathology was defined as coarse parenchyma, absence of a smooth liver capsule or blunt liver angle, while bile duct pathology was mainly segmental dilations or multifocal strictures or bile duct wall thickening. In 23 patients (35.9%), neither liver pathology nor bile duct variability was observed. Splenomegaly was observed in 33 (52.4%) patients. Only a few patients (n = 3) had significant hyperbilirubinemia (bilirubin >50 μmol/L). Baseline characteristics are summarized in Table 1.

Table 1. Baseline characteristics in patients with PSC undergoing elastography

Number of patients	66
Males	51 (77.3%)
Age, y, mean (SD)	49.0 (16.3)
Age at diagnosis, y, mean (SD)	37.7 (14.9)
Body mass index, kg/m^2, mean (SD)	25.7 (4.4)
Body mass index class	33 (52.4%)
<25	20 (30.3%)
25−30	13 (19.7%)
≥30	
PSC duration, y, median (range)	8 (0−37)
IBD (n, %)	52 (78.8%)
Ulcerative colitis	38 (57.6%)
Crohn's disease	9 (13.6%)
Indeterminate IBD	5 (7.6%)
Feature of autoimmune hepatitis	3 (4.5%)
Decompensated liver disease*	3 (4.5%)
Mayo risk score, median (range)	−0.37 (−1.89 to 2.94)
APRI score, median (range)	0.47 (0.12−3.36)
Fib4 score, median (range)	1.29 (0.23−8.0)
B-Mode ultrasound	
Irregular liver capsule†	22 (34.4%)
Subtle irregularity	17 (26.6%)
Moderate-severe irregularity	5 (7.8%)
Coarse liver parenchyma†	11 (16.7%)
Blunt liver angle* (missing = 1)	11 (16.7%)
Any liver pathology (missing = 3)	25 (39.7%)
Bile duct variability†	31 (47.0%)
Any liver or bile duct pathology	41 (62.1%)
Spleen length, cm, mean (SD)	12.1 (2.2)
Splenomegaly (≥12 cm) (missing = 3)	33 (52.4%)
Laboratory values, median (range)	
Alanine transaminase	47.5 (4−657)
Aspartate transaminase	46 (14−299)
γ-Glutamyl transferase	205 (15−2389)
Alkaline phosphatase	159 (36−863)
Bilirubin	11 (4−99)
Thrombocytes	232 (68−618)
Albumin	45 (31−51)

APRI = aspartate transaminase-to-platelet ratio index SD = standard deviation; IBD = inflammatory bowel disease; PSC = primary sclerosing cholangitis.

* Decompensating event before baseline (ascites, variceal bleeding, encephalopathy or liver synthesis failure).

† Missing = 2.

LSMs by three scanner systems

Feasibility was good to excellent for all systems, with valid results for TE in 89.4%, for pSWE in 93.9% and for 2-D-SWE.GE in 71.2%. Median (range) LSM was 7.1 kPa (3.5−61.4 kPa) for TE, 4.9 kPa (2.6−64.9 kPa) for pSWE and 6.4 kPa (4.2−40.1 kPa) for 2-D-SWE.GE. Valid results for all three systems were available for 42 patients (63.6%), with LSM values of 6.4 (3.5−32.7), 4.7 (2.7−25.8) and 6.4 (4.5−34.9) for TE, pSWE and 2-D-SWE.GE, respectively. Median LSM was significantly lower by pSWE than by TE ($p < 0.001$), whereas there were no significant overall differences between 2-D-SWE.GE and either pSWE or TE. Intersystem differences were however not linear: 2-D-SWE.GE was significantly *higher* than TE for low-average LSM values, and significantly *lower* for middle-range and high-average LSM values (*i.e.,* LSM by 2-D-SWE.GE exhibited a less steep slope compared with TE; Fig. 1). Similarly, 2-D-SWE.GE yielded significantly higher values than pSWE for low to middle-range LSM values.

LSM values exhibited a good correlation between systems (ρ = 0.70, 0.72, 0.65 for pSWE vs. TE, TE vs. 2-D-SWE.GE and pSWE vs. 2-D-SWE.GE, respectively; $p < 0.001$). Correlation improved when patients with body mass indexes (BMIs) ≥30 were excluded (ρ = 0.81, 0.78 and 0.76 for pSWE vs. TE, TE vs. 2-D-SWE.GE and pSWE vs. 2-D-SWE.GE, respectively). The intraclass correlation coefficient (ICC) was excellent for pSWE versus TE (ICC = 0.91, $p < 0.001$), while moderate for pSWE versus 2-D-SWE.GE and for TE versus 2-D-SWE.GE (ICC = 0.49, $p = 0.013$ and 0.43, $p = 0.035$, respectively). This discrepancy was largely caused by BMI. When analyzed for normal-weight individuals only (n = 31), the ICC values improved to 0.93, 0.92 and 0.81, respectively. By excluding only individuals with a BMI ≥30, ICC values were 0.91, 0.6 and 0.43, respectively. In trying to establish a threshold BMI, similar high ICC values were kept with a cutoff of BMI <28 kg/m^2 for pSWE versus TE (n = 45, ICC = 0.90, $p < 0.001$) and 27 kg/m^2 for 2-D-SWE.GE versus TE (n = 33; ICC = 0.81, $p < 0.001$).

Factors associated with LSM differences between scanner systems

Difference in LSM (ΔLSM) between pSWE and TE was associated with BMI (ρ = 0.53, $p < 0.001$) (Fig. 2) in linear regression, with higher ΔLSM in the obese. When patients were classified as either obese (BMI ≥30)

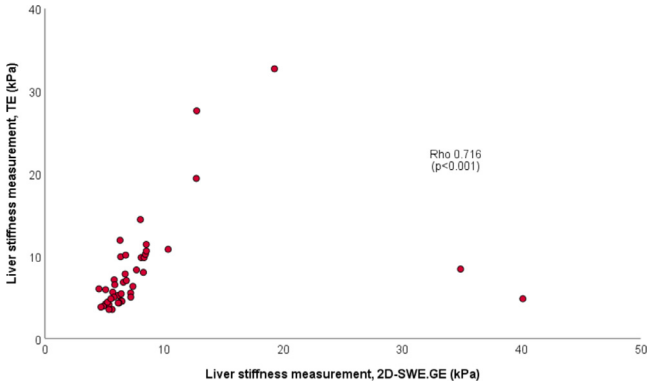

Fig. 1. Liver stiffness measurements (LSMs) for 2-D shear wave elastography (2-D-SWE.GE) plotted against those for transient elastography (TE). LSMs by 2-D-SWE.GE were higher than LSMs by TE in patients with a relatively low average LSM value, while TE values were higher than 2-D-SWE.GE values in patients with a higher average LSM value. A good correlation was observed (ρ = 0.716), but two outliers exhibited a major discrepancy. Exclusion of the two outliers increased ρ to 0.785. Patients representing these two outliers had quite average body mass indexes (27.5 and 28.1, respectively) and normal laboratory values. One had B-mode signs of steatosis, and the other had a pathologic liver capsule, but otherwise normal B-mode findings. Both had point shear wave elastography (pSWE) LSM measurements mimicking TE values, although one of them were deemed invalid because of a success rate just under 60%.

or not obese, the ΔLSM between pSWE and TE was significantly higher in the obese ($p = 0.002$). Use of the XL probe for TE in 8 PSC patients was not associated with increased ΔLSM between pSWE and TE. ΔLSM was not associated with any laboratory value, fibrosis marker or B-mode finding, nor to LSM levels.

For ΔLSM for pSWE versus 2-D-SWE.GE or 2-D-SWE.GE versus TE, there was no difference between the obese and non-obese ($p > 0.3$). The ΔLSM between pSWE and 2-D-SWE.GE was shown in linear regression to be affected mainly by the pSWE LSM value ($p < 0.001$), but otherwise no single factor. ΔLSM for 2-D-SWE.GE and TE was significantly associated with LSM values, mainly by TE ($p < 0.001$), with serum ALP ($p = 0.009$) and GT ($p = 0.02$) and BMI ($p = 0.047$).

Comparison of 2-D-SWE.GE with either TE or pSWE revealed that there were some outliers with highly deviant 2-D-SWE.GE LSM values, particularly in the higher LSM range (Fig. 1). This was further illustrated with Bland–Altman plots (Fig. 3). We did not reveal any common characteristic explaining this deviance, and it is thus possible that these were operator dependent.

LSMs and suggested prognostic markers (ELF test, Mayo risk score and spleen length)

LSMs correlated significantly with liver enzymes and serum-based fibrosis scores (Table 2). In general, TE and pSWE exhibited higher correlations with all laboratory values compared with 2-D-SWE.GE (Table 2). LSMs by pSWE, TE and 2-D-SWE.GE all correlated significantly (ρ 0.57, 0.59 and 0.40, $p \leq 0.009$) with the

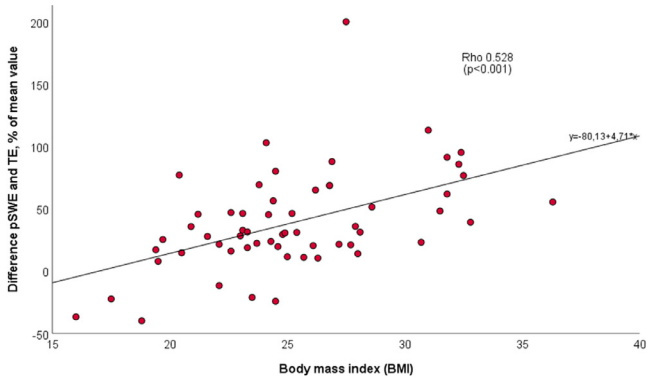

Fig. 2. Scatterplot of difference in liver stiffness values between point shear wave elastography and transient elastography in percent (y-axis) and body mass index (BMI, x-axis), ρ = 0.528, $p < 0.001$. The intersystem difference increases with increasing BMI. Looking at the systems separately, there was no increasing LSM with increasing BMI.

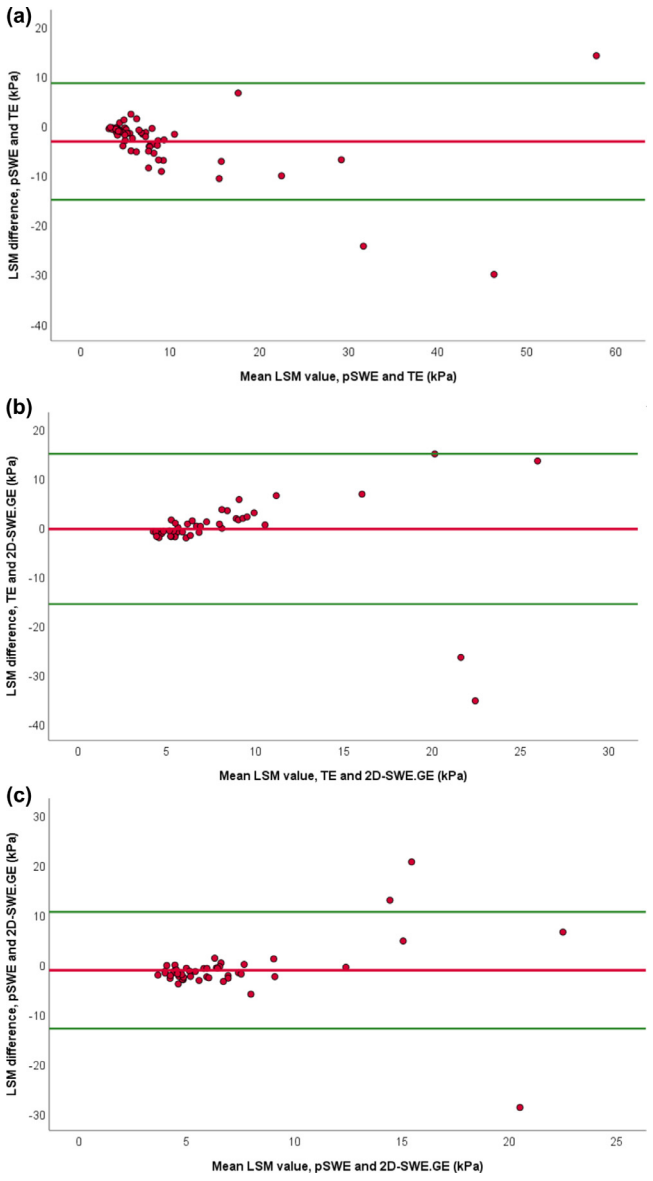

Fig. 3. Bland–Altman plots comparing scanner systems, with average liver stiffness measurements (LSMs) on the x-axis and intersystem differences on the y-axis, for (a) point shear wave elastography (pSWE) and transient elastography (TE); (b) TE and 2-D shear wave elastography (2-D-SWE.GE); and (c) pSWE and 2-D-SWE.GE.

Table 2. Correlation between liver stiffness measurements and laboratory values including fibrosis markers*

Laboratory value	ρ (*p* value)		
	TE (Fibroscan)	pSWE (Philips)	2-D-SWE.GE (GE)
Aspartate transaminase	0.71 (<0.001)	0.60 (<0.001)	0.50 (<0.001)
Alkaline phosphatase	0.69 (<0.001)	0.58 (<0.001)	0.50 (<0.001)
Enhanced liver fibrosis test	0.59 (<0.001)	0.57 (<0.001)	0.40 (0.009)
Mayo risk score	0.57 (<0.001)	0.67 (<0.001)	0.45 (0.001)
Fibrosis 4 (Fib4) score	0.48 (<0.001)	0.59 (<0.001)	0.47 (0.001)
γ-Glutamyl transferase	0.59 (<0.001)	0.41 (0.001)	0.39 (0.006)
Bilirubin	0.49 (<0.001)	0.47 (<0.001)	0.35 (0.015)
Alanine transaminase	0.53 (<0.001)	0.36 (0.003)	0.36 (0.01)
Albumin	−0.23 (0.081)	−0.46 (<0.001)	−0.18 (0.23)
IgG4	0.008 (0.955)	0.17 (0.194)	0.173 (0.257)

TE = transient elastography; pSWE = point shear wave elastography; 2-D-SWE.GE = 2-D shear wave elastography.
* Sorted after average ρ value between all three scanner systems.

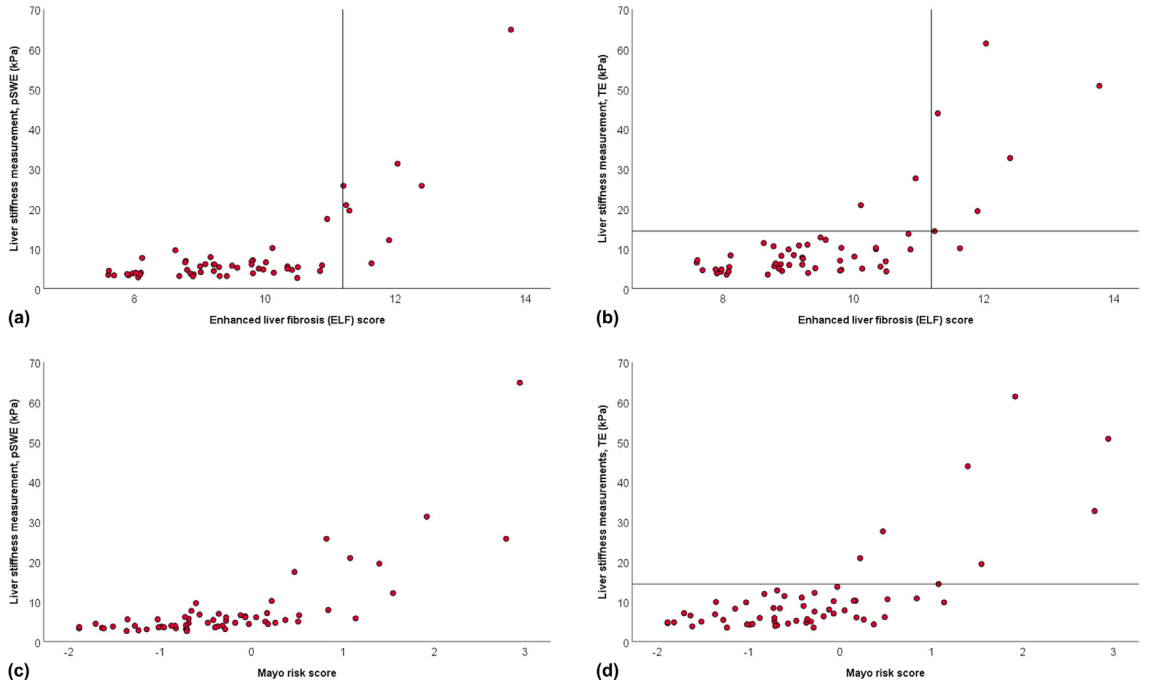

Fig. 4. Correlation between between liver stiffness and scores of fibrosis (enhanced liver fibrosis [ELF] score) or prognosis (Mayo risk score). For ELF scores, a previously published cutoff value of 11.2 is noted (vertical line), while for transient elastography (TE), a published cutoff value for cirrhosis of 14.4 is noted (horizontal line). (a) ELF score versus point shear wave elastography (pSWE). (b) ELF score versus TE. (c) Mayo risk score versus pSWE. (d) Mayo risk score versus TE.

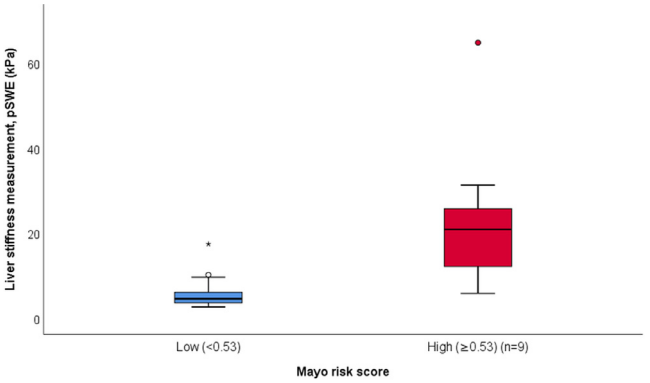

Fig. 5. Boxplot revealing a vast difference in liver stiffness measurements between patients with a high Mayo risk score and a low Mayo risk score (median values: 20.9 kPa vs. 4.5 kPa, $p < 0.001$).

enhanced liver fibrosis (ELF) test, a well-validated serum biomarker panel based on three direct markers of fibrosis that has been found to be strongly associated with clinical outcome in PSC (Vesterhus et al. 2015; de Vries et al. 2017). The suggested cutoff value for ELF of 11.2 discriminated well between high and low LSMs (Fig. 7), exhibiting an abrupt rise in LSM values beyond the cutoff value (Fig. 4a, 4b). LSMs by all methods were correlated with the Mayo risk score (Table 2). There seemed to be a cutoff value of 0.5, after which a rapid increase in LSM values was seen (Fig. 4c, 4d), and LSMs were elevated in high-risk compared with low-risk Mayo risk score groups (Fig. 5).

There was no LSM difference between patients with and without splenomegaly (spleen length ≥12 cm). However, at a cutoff of 13 cm, as proposed by a study on a large healthy patient panel (Chow et al. 2016), patients with splenomegaly had significantly higher LSM values by pSWE (4.7 vs. 6.6 kPa, $p = 0.019$) and TE (6.2 vs. 8.9 kPa, $p = 0.034$), but not by 2-D-SWE.GE.

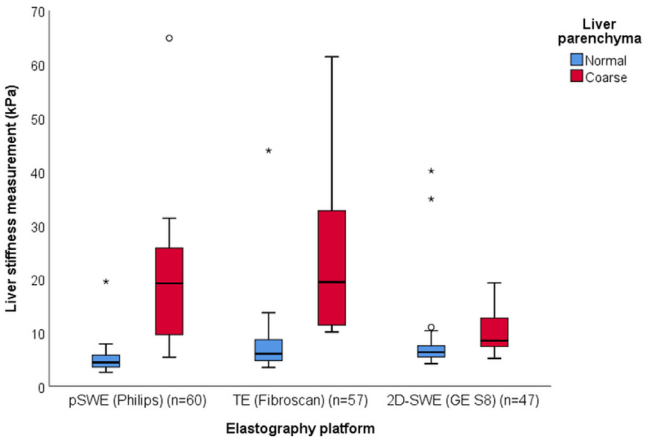

Fig. 6. Liver stiffness measurement (LSMs) for all methods, grouped by liver parenchyma on B-mode ultrasound (normal or coarse liver tissue). LSM values were significantly higher by all scanner systems in patients with coarse liver tissue, although substantially less significant by 2-D-SWE.GE ($p = 0.028$ vs. $p < 0.001$ and $p < 0.001$). TE = transient elastography; pSWE = point shear wave elastography; 2-D-SWE.GE = 2-D shear wave elastography.

LSMs and associations with clinical and B-mode characteristics in PSC patients

LSMs did not differ between men and women, nor between obese and non-obese patients, for any scanner system. LSMs were significantly higher for all systems in patients with either coarse liver parenchyma or an irregular liver capsule (Table 3, Fig. 6) (TE and pSWE: $p < 0.001$, 2-D-SWE.GE: $p = 0.028$).

Table 3. Comparison of liver stiffness measurements in patients with visible pathology or normal findings on B-mode ultrasound scanning*

	TE (kPa)		pSWE(kPa)		2-D-SWE.GE (kPa)	
	Normal vs. pathologic	*p* Value	Normal vs. pathologic	*p* Value	Normal vs. pathologic	*p* Value
Liver parenchyma	6.0 vs. 19.4	<0.001	4.4 vs. 19.2	<0.001	6.3 vs. 8.5	0.028
Liver capsule	5.9 vs. 10.1	0.008	4.0 vs. 7.0	<0.001	6.2 vs. 7.7	0.009
Liver angle	6.1 vs. 16.6	0.005	4.6 vs. 12.1	0.019	6.4 vs. 7.2	0.485
Any liver pathology	5.9 vs. 9.2	0.005	4.0 vs. 6.3	0.001	6.2 vs. 7.5	0.055
Bile duct variability	6.0 vs. 7.1	0.5	4.4 vs. 5.5	0.086	6.3 vs. 6.6	0.434
Splenomegaly	6.5 vs. 8.0	0.252	4.7 vs. 5.4	0.442	6.2 vs. 7.2	0.115

TE = transient elastography; pSWE = point shear wave elastography; 2-D-SWE.GE = 2-D shear wave elastography.
* All values are medians.

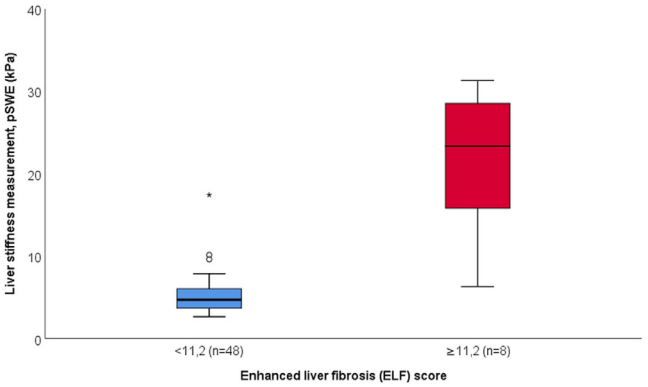

Fig. 7. Boxplot of liver stiffness measurement (LSM) by point shear wave elastography (pSWE) using a cutoff enhanced liver fibrosis (ELF) score of 11.2, discriminating high-risk (11.2) from low-risk patients. LSM values were significantly higher for patients with a high ELF score.

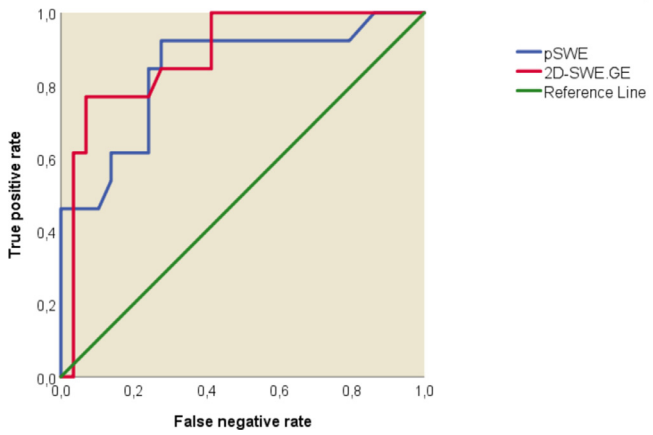

Fig. 8. Receiver operating characteristic curve of liver stiffness measurements by point shear wave elastography (pSWE) and 2-D shear wave elastography (2-D-SWE.GE) for the diagnosis of advanced fibrosis (F3−F4), using published cutoff values for transient elastography in a primary sclerosing cholangitis cohort.

Discriminative ability of pSWE and 2-D-SWE.GE to identify significant fibrosis and advanced disease

By applying published cutoff values for fibrosis staging in PSC using TE (Corpechot et al. 2014), we identified 50.8% as F0, 10.2% as F1, 1.7% as F2, 23.7% as F3 and 13.6% as F4. The ability of pSWE as well as 2-D-SWE.GE to discriminate between mild and advanced (F3−F4) disease as defined by TE was good (area under the receiver operating characteristic curve [AUROC] = 0.85 for both pSWE and 2-D-SWE.GE), with optimal cutoff values as decided by Youden's index of 4.9 kPa (sensitivity 90.9%, specificity 72.5%) and 7.8 kPa (sensitivity 76.9%, specificity 88.2%) for pSWE and 2-D-SWE.GE, respectively (Fig. 8).

DISCUSSION

To the best of our knowledge, this is the first study to explore the use of 2-D-SWE.GE in PSC with different ultrasound elastography scanner systems in a head-to-head fashion in an exclusive PSC cohort. Previous studies have described the use of either TE (Corpechot et al. 2014; Ehlken et al. 2016a, 2016b; Krawczyk et al. 2017) or pSWE (Mjelle et al. 2016; Goertz et al. 2019). Intersystem differences are known (Mjelle et al. 2016; Dietrich et al. 2017; Mulabecirovic et al. 2018a, 2018b) and may result from the different system technologies, but it is unknown how these are affected by the patchy fibrosis distribution in complex cholestatic diseases such as PSC, which is histologically different from viral hepatitis and metabolic liver diseases.

We found good to excellent feasibility for all three elastography systems in PSC patients. Previous studies in patients with liver diseases are discrepant; some describe a similar lower feasibility for 2-D-SWE.GE versus TE (Cassinotto et al. 2014; Staugaard et al. 2016), while others report the opposite (Bota et al. 2015; Cassinotto et al. 2015). LSM values in PSC patients were significantly lower by pSWE than by TE, while measurements with 2-D-SWE.GE were not different from those by either TE or pSWE. This is in line with previous results, describing similar intersystem differences in healthy cohorts (Mulabecirovic et al. 2018a, 2018b; Mjelle et al. 2019) and in various liver diseases (Rizzo et al. 2011; Sporea et al. 2014; Belei et al. 2016; Thiele et al. 2016), as well as phantoms (Mulabecirovic et al. 2016).

We obtained overall lower LSM values by pSWE than by TE in PSC patients, with a steeper slope for TE compared with either pSWE or 2-D-SWE.GE (*e.g.,* 2-D-SWE.GE yielded higher LSM values compared with TE for low mean LSM values, but lower values in the higher LSM range; Fig. 1), confirming findings suggested by previous reports (Cassinotto et al. 2014; Thiele et al. 2016; Mulabecirovic et al. 2018a, 2018b; Mjelle et al. 2019). Such differences between elastography methods regarding the slope of the LSM curve are important to acknowledge in clinical follow-up of patients.

We found good correlations for LSM values between scanner systems (ρ between 0.65 and 0.72). The ICC for pSWE versus TE was excellent (0.91). The ICC for 2-D-SWE.GE versus pSWE or TE was only moderate (0.49 and 0.43, respectively), but improved markedly when excluding overweight individuals with a BMI >27−28 kg/m^2, agreeing with reports of increased unreliable LSM measurements in patients with a BMI >27.7 kg/m^2 (Bota et al. 2014).

Exploring the intersystem difference for LSM values (ΔLSM), we found that these correlated with a few select parameters, which differed between each intersystem comparison. ΔLSM between pSWE and TE was associated only with BMI ($p < 0.001$), while ΔLSM for pSWE and 2-D-SWE.GE was influenced by the pSWE LSM value ($p <$

0.001). ΔLSM between 2-D-SWE.GE and TE was significantly associated with individual LSM values (in particular TE) ($p < 0.001$), with serum alkaline phosphatase and γ-glutamyl transferase and ELF values and with BMI. Earlier publications report conflicting results, with some describing increased failure rate or changing LSM values with increasing BMI (Guzman-Aroca et al. 2011; Popescu et al. 2013; Cassinotto et al. 2014; Liao et al. 2015), while some report that BMI exert no effect on LSM values (Takahashi et al. 2010; Rizzo et al. 2011; Son et al. 2012; Huang et al. 2014; Mulabecirovic et al. 2018a, 2018b). Some of these discrepancies may well be the result of different study designs, as most studies include no or few obese (BMI $\geq$30 kg/m^2) patients, with some studies even setting the BMI cutoff well into the normal range before testing for differences, the lowest at 22 kg/m^2 (Liao et al. 2015). Taken together, our results suggest that LSM values in overweight patients with PSC may be subject to increased uncertainty. Further studies are warranted to tease out the effects of high BMI on LSM with the various elastography scanner systems.

Limits of agreement analysis revealed a substantial dispersion in LSM values between all machines, mainly restricted to average LSM values >10 kPa, as illustrated in Bland–Altman plots (Fig. 3). Our results indicated that 2-D-SWE.GE gave a more pronounced dispersion in intersystem differences. Patients with the highest ΔLSM between TE and pSWE often had invalid results for 2-D-SWE.GE. This could suggest that the ICC values with 2-D-SWE.GE measurements were falsely *elevated* compared with those with pSWE and TE, as 2-D-SWE.GE had mainly valid measurements for patients in which LSM was easily measured and where the different scanner systems agreed.

TE has been reported to correlate well with liver biopsy in the evaluation of liver fibrosis in PSC patients (Corpechot et al. 2014). We therefore tested the performance of pSWE and 2-D-SWE.GE using TE as a gold standard, with published cutoff values for different fibrosis stages. pSWE and 2-D-SWE.GE both performed well in the identification of advanced fibrosis (F3–F4) with similar AUROC values of 0.85, with corresponding optimal cutoffs of 4.9 and 7.8 kPa as decided by Youden's index.

We found that ELF correlates rather well with LSMs by pSWE and TE and, more importantly, that there is a clear linear relationship between the parameters, with an abrupt rise in LSM after the suggested cutoff for ELF of 11.2. A similar relationship was not observed for 2-D-SWE.GE, although there was a significant correlation.

There was no difference in LSM between those with and without splenomegaly defined as spleen length $\geq$12 cm, while LSM was elevated in PSC patients with spleen length $\geq$13 cm for pSWE as well as TE. Previous studies have reported that spleen length $\geq$12 cm is associated with poorer prognosis in PSC patients (Ehlken et al. 2016a, 2016b). However, spleen size depends on body length and weight, as well as sex, and it has been reported that one in four healthy males has a spleen length $\geq$12 cm (Chow et al. 2016). Our findings may indicate that the cutoff values for spleen length as a prognostic factor might benefit from body size adjustment or vary between countries, and perhaps change in spleen size over time may be a better parameter with respect to disease progression (Jung et al. 2019).

Limitations of the study

The lack of liver biopsies is the main limitation of this study. Although liver biopsies would have strengthened our results, this invasive procedure is not clinically indicated in PSC, is associated with adverse events and hence is not acceptable from an ethical perspective. Furthermore, LSMs were not necessarily made at the exact same spot in the liver with all three scanner systems in any given patient; in particular, pSWE measurements were performed by a different operator than TE and 2-D-SWE.GE measurements. Stricter system recommendations for TE and 2-D-SWE.GE, requiring IQR/M% $\leq$30 for valid results, compared with the Philips iU22 (lacking such recommendations), may have affected results, in particular the feasibility. Considering the patchy distribution of PSC, it is possible that LSM measurements throughout the entire liver would have affected our results. However, given the results indicating closer ICCs for TE and pSWE (performed by two observers) than for TE and 2-D-SWE.GE (performed by a single observer), we do not believe this biased our results.

CONCLUSIONS

We have for the first time described the use of 2-D-SWE.GE in an exclusive PSC patient cohort, with a head-to-head comparison using three different elastography methods. We found good feasibility and moderate to excellent correlations for 2-D-SWE.GE, pSWE and TE, respectively. LSM levels differed between scanner systems. The ICC was excellent for pSWE versus TE. For 2-D-SWE.GE compared with TE or pSWE, the ICC was moderate, but improved significantly when overweight patients were excluded. LSM was correlated with the ELF test. Our results further suggest that a spleen length cutoff of 13 cm may be more appropriate as a prognostic marker in PSC than the recommended 12 cm.

Further research is warranted to clarify the effects of BMI on LSM for the various elastography scanner systems and the factors causing diverging LSM values between scanner systems in PSC patients.

Acknowledgments—The authors thank the manufacturers of the different scanner systems, Philips and GE Healthcare, for the opportunity to use their elastography systems free of charge. The companies mentioned had no influence on the design or performance of the study. We also thank the Norwegian PSC Research Center. The work is part of the PhD project of A.B.M., funded by the University of Bergen.

Conflict of interest disclosure—OHG did consultancy for GE and Samsung in 2017.

REFERENCES

Belei O, Sporea I, Gradinaru-Tascau O, Olariu L, Popescu A, Simedrea I, Marginean O. Comparison of three ultrasound based elastographic techniques in children and adolescents with chronic diffuse liver diseases. Med Ultrason 2016;18:145–150.

Boonstra K, Weersma RK, van Erpecum KJ, Rauws EA, Spanier BW, Poen AC, van Nieuwkerk KM, Drenth JP, Witteman BJ, Tuynman HA, Naber AH, Kingma PJ, van Buuren HR, van Hoek B, Vleggaar FP, van Geloven N, Beuers U, Ponsioen CY, Epi PSG. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. Hepatology 2013;58:2045–2055.

Bota S, Sporea I, Sirli R, Popescu A, Danila M, Jurchis A, Gradinaru-Tascau O. Factors associated with the impossibility to obtain reliable liver stiffness measurements by means of acoustic radiation force impulse (ARFI) elastography—Analysis of a cohort of 1,031 subjects. Eur J Radiol 2014;83:268–272.

Bota S, Paternostro R, Etschmaier A, Schwarzer R, Salzl P, Mandorfer M, Kienbacher C, Ferlitsch M, Reiberger T, Trauner M, Peck-Radosavljevic M, Ferlitsch A. Performance of 2-D shear wave elastography in liver fibrosis assessment compared with serologic tests and transient elastography in clinical routine. Ultrasound Med Biol 2015;41:2340–2349.

Broome U, Olsson R, Loof L, Bodemar G, Hultcrantz R, Danielsson A, Prytz H, Sandberg-Gertzen H, Wallerstedt S, Lindberg G. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. Gut 1996;38:610–615.

Cassinotto C, Lapuyade B, Mouries A, Hiriart JB, Vergniol J, Gaye D, Castain C, Le Bail B, Chermak F, Foucher J, Laurent F, Montaudon M, De Ledinghen V. Non-invasive assessment of liver fibrosis with impulse elastography: comparison of Supersonic Shear Imaging with ARFI and FibroScan. J Hepatol 2014;61:550–557.

Cassinotto C, Charrie A, Mouries A, Lapuyade B, Hiriart JB, Vergniol J, Gaye D, Hocquelet A, Charbonnier M, Foucher J, Laurent F, Chermak F, Montaudon M, de Ledinghen V. Liver and spleen elastography using supersonic shear imaging for the non-invasive diagnosis of cirrhosis severity and oesophageal varices. Dig Liver Dis 2015;47:695–701.

Chow KU, Luxembourg B, Seifried E, Bonig H. Spleen size is significantly influenced by body height and sex: Establishment of normal values for spleen size at US with a cohort of 1200 healthy individuals. Radiology 2016;279:306–313.

Corpechot C, El Naggar A, Poujol-Robert A, Ziol M, Wendum D, Chazouilleres O, de Ledinghen V, Dhumeaux D, Marcellin P, Beaugrand M, Poupon R. Assessment of biliary fibrosis by transient elastography in patients with PBC and PSC. Hepatology 2006;43:1118–1124.

Corpechot C, Gaouar F, El Naggar A, Kemgang A, Wendum D, Poupon R, Carrat F, Chazouilleres O. Baseline values and changes in liver stiffness measured by transient elastography are associated with severity of fibrosis and outcomes of patients with primary sclerosing cholangitis. Gastroenterology 2014;146:970–979 quiz e15−e16..

de Vries EM, Farkkila M, Milkiewicz P, Hov JR, Eksteen B, Thorburn D, Chazouilleres O, Pares A, Nygard S, Gilja OH, Wunsch E, Invernizzi P, Carbone M, Bernuzzi F, Boberg KM, Rosjo H, Rosenberg W, Beuers UH, Ponsioen CY, Karlsen TH, Vesterhus M. Enhanced liver fibrosis test predicts transplant-free survival in primary sclerosing cholangitis, a multi-centre study. Liver Int 2017;37:1554–1561.

Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, Cosgrove D, Ferraioli G, Friedrich-Rust M, Gilja OH, Goertz RS, Karlas T, de Knegt R, de Ledinghen V, Piscaglia F, Procopet B, Saftoiu A, Sidhu PS, Sporea I, Thiele M. EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version). Ultraschall Med 2017;38:e48.

Ehlken H, Wroblewski R, Corpechot C, Arrive L, Lezius S, Hartl J, Denzer UW, Lohse AW, Chazouilleres O, Schramm C. Spleen size for the prediction of clinical outcome in patients with primary sclerosing cholangitis. Gut 2016;65:1230–1232 a.

Ehlken H, Wroblewski R, Corpechot C, Arrive L, Rieger T, Hartl J, Lezius S, Hubener P, Schulze K, Zenouzi R, Sebode M, Peiseler M, Denzer UW, Quaas A, Weiler-Normann C, Lohse AW, Chazouilleres O, Schramm C. Validation of transient elastography and comparison with spleen length measurement for staging of fibrosis and clinical prognosis in primary sclerosing cholangitis. PLoS One 2016;11:e0164224. b.

Ferraioli G, De Silvestri A, Lissandrin R, Maiocchi L, Tinelli C, Filice C, Barr RG. Evaluation of inter-system variability in liver stiffness measurements. Ultraschall Med 2019;40:64–75.

Friedrich-Rust M, Ong MF, Martens S, Sarrazin C, Bojunga J, Zeuzem S, Herrmann E. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology 2008;134:960–974.

Goertz RS, Gassmann L, Strobel D, Wildner D, Schellhaas B, Neurath MF, Pfeifer L. Acoustic radiation force impulse (ARFI) elastography in autoimmune and cholestatic liver diseases. Ann Hepatol 2019;18:23–29.

Guzman-Aroca F, Reus M, Berna-Serna JD, Serrano L, Serrano C, Gilabert A, Cepero A. Reproducibility of shear wave velocity measurements by acoustic radiation force impulse imaging of the liver: a study in healthy volunteers. J Ultrasound Med 2011;30:975–979.

Huang Z, Zheng J, Zeng J, Wang X, Wu T, Zheng R. Normal liver stiffness in healthy adults assessed by real-time shear wave elastography and factors that influence this method. Ultrasound Med Biol 2014;40:2549–2555.

Iijima H, Tada T, Kumada T, Kobayashi N, Yoshida M, Aoki T, Nishimura T, Nakano C, Ishii A, Takashima T, Sakai Y, Aizawa N, Nishikawa H, Ikeda N, Iwata Y, Enomoto H, Ide YH, Hirota S, Fujimoto J, Nishiguchi S. Comparison of liver stiffness assessment by transient elastography and shear wave elastography using six ultrasound devices. Hepatol Res 2019;49:676–686.

Jung F, Cazzagon N, Vettorazzi E, Corpechot C, Chazouilleres O, Arrive L, Lohse AW, Schramm C, Ehlken H. Rate of spleen length progression is a marker of outcome in patients with primary sclerosing cholangitis. Clin Gastroenterol Hepatol 2019;17:2613–2615.

Krawczyk M, Ligocka J, Ligocki M, Raszeja-Wyszomirska J, Milkiewicz M, Szparecki G, Ilczuk T, Gornicka B, Zieniewicz K, Krawczyk M, Lammert F, Milkiewicz P. Does transient elastography correlate with liver fibrosis in patients with PSC? Laennec score-based analysis of explanted livers. Scand J Gastroenterol 2017;52:1407–1412.

Lefebvre T, Wartelle-Bladou C, Wong P, Sebastiani G, Giard JM, Castel H, Murphy-Lavallee J, Olivie D, Ilinca A, Sylvestre MP, Gilbert G, Gao ZH, Nguyen BN, Cloutier G, Tang A. Prospective comparison of transient, point shear wave, and magnetic resonance elastography for staging liver fibrosis. Eur Radiol 2019;29:6477–6488.

Liao LY, Kuo KL, Chiang HS, Lin CZ, Lin YP, Lin CL. Acoustic radiation force impulse elastography of the liver in healthy patients: test location, reference range and influence of gender and body mass index. Ultrasound Med Biol 2015;41:698–704.

Millonig G, Reimann FM, Friedrich S, Fonouni H, Mehrabi A, Buchler MW, Seitz HK, Mueller S. Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. Hepatology 2008;48:1718–1723.

Mjelle AB, Mulabecirovic A, Hausken T, Havre RF, Gilja OH, Vesterhus M. Ultrasound and point shear wave elastography in livers of patients with primary sclerosing cholangitis. Ultrasound Med Biol 2016;42:2146–2155.

Mjelle AB, Mulabecirovic A, Havre RF, Rosendahl K, Juliusson PB, Olafsdottir E, Gilja OH, Vesterhus M. Normal liver stiffness values in children: A comparison of three different elastography methods. J Pediatr Gastroenterol Nutr 2019;68:706–712.

Mulabecirovic A, Vesterhus M, Gilja OH, Havre RF. In vitro comparison of five different elastography systems for clinical applications, using strain and shear wave technology. Ultrasound Med Biol 2016;42:2572–2588.

Mulabecirovic A, Mjelle AB, Gilja OH, Vesterhus M, Havre RF. Liver elasticity in healthy individuals by two novel shear-wave

elastography systems—Comparison by age, gender, BMI and number of measurements. PLoS One 2018;13:e0203486. a.

Mulabecirovic A, Mjelle AB, Gilja OH, Vesterhus M, Havre RF. Repeatability of shear wave elastography in liver fibrosis phantoms—Evaluation of five different systems. PLoS One 2018;13:e0189671. b.

Piscaglia F, Salvatorea V, Mulazzani L, Cantisani V, Colecchia A, Di Donato R, Felicani C, Ferrarini A, Gamal N, Grasso V, Marasco G, Mazzotta E, Ravaioli F, Ruggieri G, Serio I, Nkamgho JFS, Serra C, Festi D, Schiavone C, Bolondi L. Differences in liver stiffness values obtained with new ultrasound elastography machines and Fibroscan: A comparative study. Digest Liver Dis 2017;49:802–808.

Popescu A, Bota S, Sporea I, Sirli R, Danila M, Racean S, Suseanu D, Gradinaru O, Ivascu Siegfried C. The influence of food intake on liver stiffness values assessed by acoustic radiation force impulse elastography—Preliminary results. Ultrasound Med Biol 2013;39:579–584.

Rizzo L, Calvaruso V, Cacopardo B, Alessi N, Attanasio M, Petta S, Fatuzzo F, Montineri A, Mazzola A, L'Abbate L, Nunnari G, Bronte F, Di Marco V, Craxi A, Camma C. Comparison of transient elastography and acoustic radiation force impulse for non-invasive staging of liver fibrosis in patients with chronic hepatitis C. Am J Gastroenterol 2011;106:2112–2120.

Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, Mal F, Christidis C, Ziol M, Poulet B, Kazemi F, Beaugrand M, Palau R. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 2003;29:1705–1713.

Son CY, Kim SU, Han WK, Choi GH, Park H, Yang SC, Choi JS, Park JY, Kim DY, Ahn SH, Chon CY, Han KH. Normal liver elasticity values using acoustic radiation force impulse imaging: a prospective study in healthy living liver and kidney donors. J Gastroenterol Hepatol 2012;27:130–136.

Sporea I, Bota S, Gradinaru-Tascau O, Sirli R, Popescu A. Comparative study between two point shear wave elastographic techniques: Acoustic radiation force impulse (ARFI) elastography and ElastPQ. Med Ultrason 2014;16:309–314.

Sporea I, Mare R, Lupusoru R, Popescu A, Danila M, Bende F, Sirli R. Comparative study between four ultrasound shear wave elastographic methods for liver fibrosis assessment. Med Ultrason 2018;20:265–271.

Staugaard B, Christensen PB, Mossner B, Hansen JF, Madsen BS, Soholm J, Krag A, Thiele M. Feasibility of transient elastography versus real-time two-dimensional shear wave elastography in difficult-to-scan patients. Scand J Gastroenterol 2016;51:1354–1359.

Takahashi H, Ono N, Eguchi Y, Eguchi T, Kitajima Y, Kawaguchi Y, Nakashita S, Ozaki I, Mizuta T, Toda S, Kudo S, Miyoshi A, Miyazaki K, Fujimoto K. Evaluation of acoustic radiation force impulse elastography for fibrosis staging of chronic liver disease: a pilot study. Liver Int 2010;30:538–545.

Thiele M, Detlefsen S, Sevelsted Moller L, Madsen BS, Fuglsang Hansen J, Fialla AD, Trebicka J, Krag A. Transient and 2-dimensional shear-wave elastography provide comparable assessment of alcoholic liver fibrosis and cirrhosis. Gastroenterology 2016;150:123–133.

Vesterhus M, Hov JR, Holm A, Schrumpf E, Nygard S, Godang K, Andersen IM, Naess S, Thorburn D, Saffioti F, Vatn M, Gilja OH, Lund-Johansen F, Syversveen T, Brabrand K, Pares A, Ponsioen CY, Pinzani M, Farkkila M, Moum B, Ueland T, Rosjo H, Rosenberg W, Boberg KM, Karlsen TH. Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. Hepatology 2015;62:188–197.