Validation of liver elastography in patients with primary sclerosing cholangitis and healthy individuals. Normal values and correlation to fibrosis parameters.

Anders Batman Mjelle

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2021

UNIVERSITY OF BERGEN

Validation of liver elastography in patients with primary sclerosing cholangitis and healthy individuals. Normal values and correlation to fibrosis parameters.

Anders Batman Mjelle

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 08.01.2021

Year: 2021

Title: Validation of liver elastography in patients with primary sclerosing cholangitis and healthy individuals. Normal values and correlation to fibrosis parameters.

Name: Anders Batman Mjelle

Print: Skipnes Kommunikasjon / University of Bergen

"So, a graduate student, huh? How come you guys can go to the moon but you can't make my shoes smell good?"

Homer J. Simpson

Contents

SCIENTIFIC ENVIRONMENT

The project began as a cooperation with the National Center for Ultrasound in Gastroenterology (NCUG) and Medical Department, both at Haukeland University Hospital in Bergen. This is where I met professor Gilja and my main supervisor, Mette Vesterhus, resulting in Paper I. The studies were financed by the University of Bergen from 2017, as part of the Ph.D. program.

The National Center for Ultrasound in Gastroenterology (NCUG) was founded in 2001, established to perform teaching and research on gastroenterological ultrasound, focusing on standard B-mode ultrasound and more advanced methods such as elastography, contrast-enhanced ultrasound, gastrointestinal motility, and endoscopic ultrasound. NCUG is a European Learning Centre, acknowledged by the European Federation of Ultrasound Societies in Medicine and Biology (EFSUMB).

We have also been dependent on the collaboration with the Norwegian PSC Research Center (NoPSC), established in 2008 at Oslo University Hospital, Rikshospitalet. NoPSC is internationally acknowledged for spearheading research related to basic and clinical aspects of primary sclerosing cholangitis and is involved in a vast amount of both national and international projects.

ACKNOWLEDGMENTS

First of all, I want to thank all participants in our studies: the fantastic children of all ages, their parents and guardians, the patients, and the healthy adults. Without their invaluable participation, there would be no research and no thesis.

I have to thank the University of Bergen for letting me spend these years producing the research on which this thesis is based. The University has been a very supportive and flexible employer, and I am grateful for being a part of this great institution. The administration at the Department of Clinical Medicine has been fantastic, always answering my endless line of questions.

Without my supervisors, there would be no thesis. I am in deep debt to my great supervisors: professor Gilja, the ultrasound guru and head of the National Centre for Ultrasound in Gastroenterology; Edda Olafsdottir, who has taught me the field of pediatric gastroenterology; and my main supervisor, Mette Vesterhus, who has been there constantly throughout the entire period, always supporting, never resting, giving me valuable feedback within (and way outside) regular working hours. I could not have wished for a better supervisor, and because of you, this whole process became rather easy.

To my co-culprit (and when it comes to making graphs, my gifted instructor), Anesa Mulabecirovic: thank you for the scientific discussions and friendship.

I want to thank the rest of the gastroenterological milieu at the Depart of Medicine: Roald Flesland Havre, Trond Engjom, Roy Cato Solheim, and everyone else, in particular all the fantastic nurses who are always there to help me.

To all my colleagues and friends at the Department of Pediatric and Adolescent Medicine: thank you for all your support during these years (many of you also helped me find study participants). I particularly need to thank my colleagues in pediatric gastroenterology, who have had to work harder in the years I have been gone, never complaining, always supportive. The same goes to my colleagues working with emergency medicine simulation training and to the Department's leadership, particularly Ansgar Berg, my supervisor during my training in pediatrics, and whom I still regard as one of my supervisors in life.

Finally, I have to thank my family. My parents, siblings, and my grandmother ("Ho Mor"), whom all live too far away, and Jim, who sadly passed away decades too soon. And of course, my wife Karin (lskdg!) and our fantastic children: Selma, Sindre, and Sebastian, who too often have to be patient with their ever working parents. The best thing (by far) with these years of doing my Ph.D. has been more time with my family. (Our fourth child should probably be mentioned, the Portuguese water dog Professor Frank Zappa av Målvatnet, who is always there, though often not doing what I ask of him.)

ABBREVIATIONS

2D-SWE – Two-dimensional shear wave elastography

ALP – Alkaline phosphatase

APRI – Aspartate aminotransferase-to-platelet-ratio index

ARFI – Acoustic radiation force impulse

AST – Aspartate aminotransferase

AUC – Area under the curve

BMI – Body mass index

B-mode – Brightness mode

CI – Confidence interval

CT – Computerized tomography

EFSUMB – European Federation of Societies for Ultrasound in Medicine and Biology

ELF – Enhanced liver fibrosis test

FIB-4 – Fibrosis-4

IBD – Inflammatory bowel disease

ICC – Intraclass correlation coefficient

IQR – Interquartile range

IQR/M – Interquartile range divided by the median

kPa – Kilopascals

LS – Liver stiffness

LSM – Liver stiffness measurement

MRI – Magnetic resonance imaging

M/S – Meters per second

PPV – Positive predictive value

PSC – Primary sclerosing cholangitis

pSWE – Point shear wave elastography

ROI – Region of interest

SD – Standard deviation

SWE – Shear wave elastography

SWS – Shear wave speed

TE – Transient elastography

ULN – Upper limit of normal

WFUMB – World Federation for Ultrasound in Medicine and Biology

SUMMARY OF THE THESIS

Liver elastography applies ultrasound-based measurements of liver stiffness (LS), used as a quantification of liver fibrosis. Correlation between liver stiffness measurements (LSM) and histological stages of liver fibrosis was already established; however, technological and methodological differences between ultrasound equipment and biological differences related to age or liver disease etiology, may influence the interpretation of LSM. Correct assessment of disease relies on a precise definition of normality, but reference values and inter-system differences were lacking for several elastography systems. Data were particularly scarce for children and patients with primary sclerosing cholangitis (PSC). In this rare and complicated liver disease, non-invasive surveillance of fibrosis development is recommended as part of follow-up.

Thus, we aimed to establish reference values for LSM in children and adults using several elastography systems and to compare principally different methods such as transient elastography (TE), point shear wave elastography (pSWE), and two-dimensional shear wave elastography (2D-SWE) in a head-to-head setup. We furthermore investigated, for the first time, the feasibility of pSWE and 2D-SWE in PSC patients.

In studies II and III, we included 343 healthy individuals aged 4-70 years. Applying two or three elastography systems head-to-head in every participant, we defined age-specific reference values for each system.

Reference values were determined for different age groups: 4-7 years, 8-11 years, 12–14 years, 15–17 years, and 20–70 years. We found a gender difference in subjects 12–70 years, with no similar difference in young children.
LSM was higher in adolescents and adults compared to younger children.
Correlation between different observers was good.

In studies I and IV, a cohort of adult PSC patients was followed, collecting clinical data and performing LSM and blood tests. Paper I describes pSWE in PSC for the

first time, comparing assessments of both liver lobes and concluding that left liver LSM is unreliable; hence, subsequent measurements forming Paper IV were applied in the right liver lobe only. All PSC patients were examined by pSWE at every visit. For Paper IV, all patients were examined head-to-head by both pSWE, 2D-SWE, and TE: all three systems were feasible in PSC patients, and all were highly correlated with other indications of liver fibrosis (B-mode findings, liver biochemistry, fibrosis scores, and prognostic scores).

In conclusion, the results demonstrate that elastography systems representing three principally different methods are feasible and perform well in healthy subjects and PSC patients. We have established reference values for healthy children and adults, with head-to-head inter-system comparisons and descriptions of interobserver differences. Ultrasound elastography of the liver should be adopted broadly in the medical environment; in screening, diagnostics, and clinical patient follow-up of both children and adults.

LIST OF PUBLICATIONS

I. 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 Sep; 42(9):2146-55.* PMID: 27262519

II. Mjelle AB, Mulabecirovic A, Havre RF, Rosendahl K, Juliusson PB, Olafsdottir E, Gilja OH, Vesterhus M. Normal liver stiffness values in children: A comparison between three different elastography methods. *J Pediatr Gastroenterol Nutr. 2019 May;68(5):706-712.* PMID: 30889132.

III. 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 Sep 14;13(9):e0203486.* PMID: 30216377.

IV. Mjelle AB, Fossdal G, Gilja OH, Vesterhus M. Liver elastography in primary sclerosing cholangitis patients using three different scanner systems. *Ultrasound Med Biol. 2020 Aug;46(8):1854-1864.* PMID: 32507342

1. INTRODUCTION

1.1 PREFACE

Liver elastography is well established as a non-invasive method for liver fibrosis assessment, but opportunities and limitations remain to be fully explored, partly due to an ever-increasing jungle of different systems and platforms. For every system, there is a need for investigations in both healthy individuals and chronic liver patients to establish reference and cutoff values where fibrosis or cirrhosis can be suspected. Comparably, suppose a producer wants to introduce a new way of measuring serum sodium values. In that case, this will need proof of performance using reference standard testing *before* being applied in clinical practice – this is not necessarily the case with liver elastography.

In addition to establishing reference values, studies in healthy individuals are warranted to evaluate the effect of various factors and confounders, such as gender, age, or body mass index (BMI), as the impact of such factors may be obscured by unpredictable diseases with fluctuating or elevated LSM values. Published studies have concluded differently regarding the effect of gender and age on LSM.

Liver elastography has been shown to correlate well with liver fibrosis evaluation by liver biopsy, with an overweight of investigations performed in adults with infectious hepatitis [1-3]. Data are scarce for rare diseases such as PSC, and the few published studies prior to the present thesis were all limited to TE [4, 5], demonstrating that LSM by TE is associated with clinical outcome and indicating elastography as a sorely needed prognostic tool in this severe, but hitherto unpredictable disease.

Serum markers are shown to correlate with the outcome. Specifically, the Enhanced liver fibrosis (ELF) test has demonstrated a strong and independent correlation with prognosis in PSC patients [6, 7]. While LSM is indeed a robust

surrogate marker of liver fibrosis, it may not outperform all serum-based tests under all conditions. Some studies have demonstrated that the combination of LSM and fibrosis markers performs better than LSM alone [8-10]. In PSC, dominant stenosis causing cholestasis is a frequent finding and may influence LSM, without being indicative of true liver fibrosis [11]. Thus, comparing each elastography platform to the ELF test as the leading serum-based liver fibrosis test in PSC is of high interest.

Non-invasive evaluation of liver fibrosis encompasses serum based tests and scores directly or indirectly assessing fibrogenesis, as well as imaging-based methods assessing LS, a physical property of the liver. Acknowledging that each strategy carries its advantages and disadvantages, both showing utility [12], we compared LSM results to serum markers and fibrosis scores. These approaches can be used in a step-up manner: several publications advocate using simple serum biomarkers in primary care, with non-invasive tests, including elastography as part of the subsequent specialist evaluation [13-17].

Literature search was ended August 26th, 2020.

1.2 ULTRASOUND

1.2.1 Basic concept of ultrasound

Sound is a wave of mechanical energy, caused by vibration of molecules within a medium [18]. Humans can hear sound with frequencies between 20 and 20.000 Hz [19]. Ultrasound does not differ from sound waves, except for the fact that the frequency is higher than what is audible for man (>20.000 Hz), thus the word *ultra*sound (ultra: from Latin, meaning 'beyond' [20]). Ultrasound used in medical applications is commonly in the range of 1–20 MHz (MHz = 10^6 Hz).

Diagnostic ultrasound uses ultrasound waves to generate an image, providing a valuable source of clinical information. A B-mode (brightness mode) image (Fig. 1) is constructed from echoes generated by the reflection of ultrasound waves from boundaries between different tissues and from scattering from subtle irregularities within tissues [21, 22].

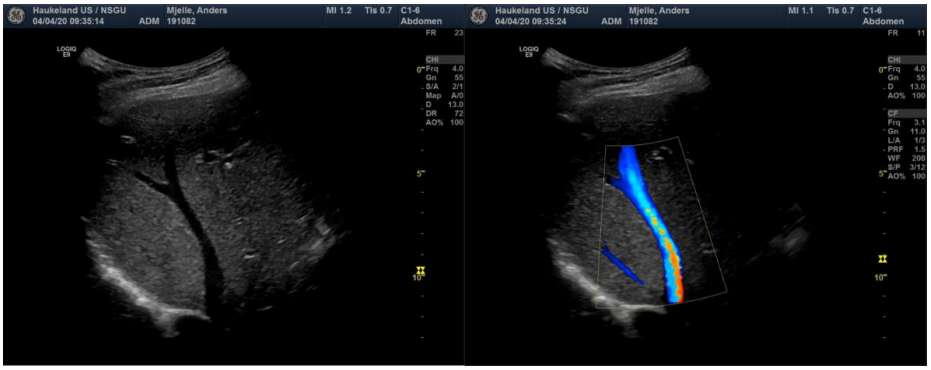


Fig. 1. B-mode ultrasound image without (left) and with (right) color flow, highlighting blood flow. *Images: Batman*

The brightness of each point in the created image is related to the returning echo's strength. The same principle is applied when echo sounding is used to determine the distance to the bottom of the sea: an ultrasound wave is sent from the boat towards the seabed, which reflects the wave and transmits it back to the boat, where the signal is received and interpreted (Fig. 2). If the speed of sound through water is known, the distance to the seabed can easily be calculated once the time taken is measured.

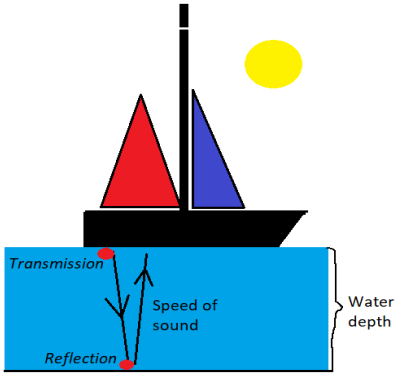


Fig. 2. Echo sounding. *Image: Batman*

The human anatomy is obviously more complex than in this example, with multiple interfaces and tissues with different characteristics. This leads to a tremendous amount of information, with echoes returning from various depths and with different amplitudes or strength of the echo, creating a multifaceted picture. The return time will define the depth and thus, where the signal is placed in the B-mode image, while the echo's strength will determine the intensity of the specific points.

Ultrasound systems used in medicine typically consist of a transmitter, a transducer, and a screen for image display. The transmitter generates electric energy and activates the transducer, causing it to vibrate and create an acoustic pulse transmitted through tissues [19, 22]. The transducer holds ceramic crystal elements, referred to as piezoelectric crystals. This is where the conversion of electric energy to pulse waves, and back again to electrical energy, takes place

(Fig. 3). Piezoelectric materials generate voltages when an external force imposes compression or stretching [21, 22]. The transducer receives the reflected echo, and the wave is then converted back to an electric signal.

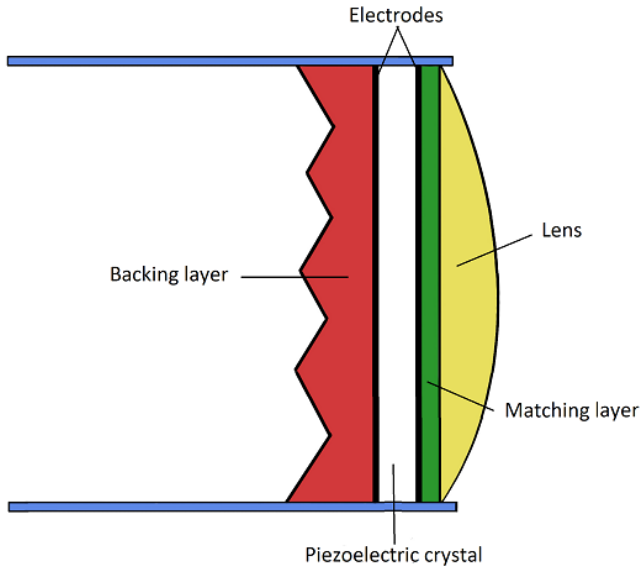


Fig. 3. Transverse section of an ultrasound transducer. *Image: Selma Sundt Mjelle*.

1.2.2 Development of ultrasound technology

The first experiments on ultrasound demonstrating the probable existence of ultrasound were performed in bats by Lazzaro Spallanzani (1729-1799), an Italian priest and physiologist [23]. He found that when the bat's mouth was covered, it would crash into objects in the dark, concluding that the bat used its ears for "seeing" and measuring depths. The concept of piezoelectricity, necessary for ultrasound, where electric charge accumulates in solid materials in response to mechanical stress, was first demonstrated in 1880 by the Curie brothers [24]. Devices using echolocation were patented in 1912, with the first sonar (*so*und *na*vigation and *r*anging) being built in 1914. With the fear of German submarines, World War I pushed technology forward, ending in the first

construction of an underwater sound generator, considered to be the prototype of modern ultrasound devices [25]. Ultrasound in medical diagnostics started with the attempt of locating brain tumors and cerebral ventricles measuring the transmission of ultrasound beams through the head, published in 1942 [23]. The next decade, in 1958, saw the Lancet publication describing differentiating cystic and solid abdominal masses [26], a publication considered a hallmark paper in ultrasound development. In the latter part of the 1960s, pulsed Doppler technology was developed, allowing investigation of blood flow in the heart [27]. In the 1970s, several developments were described, including continuous wave, spectral wave Doppler, and color Doppler [28]. In 1986, three-dimensional imaging was performed on a fetus [28]. The overall quality of imaging gradually improved and is still improving. The last 10–20 years have seen exponential improvement in both image quality and new platforms, with handheld devices, including tablets and probes designed for use with smartphones, applied in all fields of medicine.

1.2.3 Clinical use of ultrasound

Ultrasound is widely used in health care globally, in radiology only second to conventional x-ray imaging. The 2018–2019 report from the UK National Health Service showed that ultrasound was used 10.3 million times during 12 months in the UK – more than all computer tomography (CT) and magnetic resonance imaging (MRI) scans combined (9.7 million) [29].

One of the major advantages of diagnostic ultrasound compared to conventional x-ray imaging is the dynamic imaging in real-time. It allows the investigation of rapidly moving organs, such as the heart or the blood. It is especially relevant for application in children, performing well even in those too small to lay still or hold their breath while lacking the ionizing radiation, which is particularly harmful in growing children. Furthermore, children are often thin and easy to examine, with a short distance from the skin to the organs of interest. Ultrasound may easily be performed bedside, which is a significant advantage

compared to MRI or CT and can thus be performed even when patient transportation is difficult or unwanted, e.g., in an intensive care setting.

1.3 ELASTOGRAPHY

1.3.1 The concept of elastography

Elastography is essentially an upgraded version of palpation, the ancient diagnostic technique, described more than 3'500 years ago [30]. Palpation gives the clinician the opportunity to locate organs and tumors and assess its' properties, e.g., size, form, position, and stiffness. A typical example is lymph nodes, which typically will be harder when growing due to malignancy than in the event of an infectious etiology. A malignant tumor will feel stiffer than a benign tumor, and scar tissue/fibrosis will feel stiffer than healthy organ tissue. Standard ultrasound may solve some of this issue, analyzing the size, position, and form of the organ or lesion. Still, with little correlation between elasticity and echogenicity, there was an obvious need for a new modality. Elastography, applied in both ultrasound and MRI technology, allows the clinician to evaluate what is beyond palpation, yielding a more objective description of the tissue's mechanical properties. Different forms of ultrasound elastography exist (see 1.3.3): commonly, a pulse, either ultrasound-induced or mechanically induced, is transmitted by the transducer, causing waves through the adjacent tissue, which subsequently give rise to slow traveling, perpendicular *shear waves* (Fig. 4). Since waves will travel faster in a stiff medium compared to a soft medium, a high shear wave speed (SWS) indicates a stiff tissue. Measurements are either given directly as SWS in meters per second (m/s) or converted to Young's modulus (see 3.3 DEFINITIONS) value using the equation $kPa = 3\rho(m/s)^2$. As this equation requires some assumptions, SWS (using m/s) is considered more accurate. However, most publications report values in kPa [31].

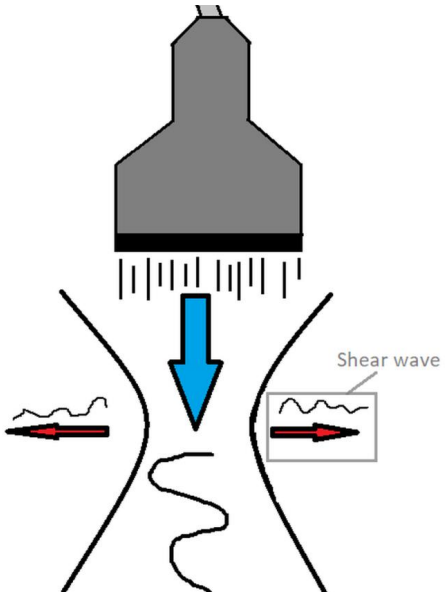


Fig. 4. Shear waves are created perpendicularly to the acoustic push pulse. The shear wave speed (SWS) will vary depending on the tissue stiffness, traveling faster in stiffer tissues. A high SWS thus suggests the presence of liver fibrosis. *Image: Batman*

1.3.2 The development of elastography

The field of elastography starts with *strain*, a term used in physical sciences, describing how an object is displaced when subjected to an external force. It is intuitive that applying the same amount of force to a soft and a hard object, will cause different deformation between objects; but, how to describe this? The concept of imaging tissue strain was developed in the 1970s and early 1980s [32]. In the early 1990s, different research groups expanded on this, describing tissue stiffness imaging, either as a quasistatic pressure (producing an elastogram, a depiction of the estimated tissue strain [33]) or vibrator-induced excitation, measuring the tissue motion with a color Doppler instrument [34, 35], eventually resulting in lesion detection [36]. Doppler ultrasound was also used to estimate strain applied in Strain Rate Imaging, providing relative strain in contracting tissue [37-39] or stiffness in tumors [40-42]. Studies on the

propagation of shear waves in phantoms and beef muscle were also performed [43, 44]. In 1999, transient elastography was first described, demonstrating biases during shear wave propagation, with the formation of different wave types [45, 46]. These different wave types could be identified using transient excitation, facilitating measurement of shear elasticity and shear viscosity. The method was termed transient elastography (TE). The same year featured a paper describing a system setup with a piston built on a vibrator, with an imaging system producing up to 1'000–2'000 frames of shear wave propagation images per second [47]. From these images, both wavelength and SWS were estimated. The same research group continued to develop the system further, and in 2003 they reported on the use of TE using the Fibroscan system in patients with hepatitis C [1]. At the beginning of the current century, acoustic radiation force was applied for tissue stiffness evaluation. This allowed elastography coupled with standard B-mode ultrasonographic imaging [48, 49], based on a technique described a decade earlier [50]. The last decade has seen an explosion in technique development, with a wide variety of systems [51].

1.3.3 Different elastography techniques

Elastography systems can be divided into either strain imaging, shear wave speed *measurement,* or shear wave *imaging*.

In **strain imaging**, the system uses echo signals before and during external compression, comparing these signals. This is then used to calculate how the tissue is displaced [52]. An elastogram is produced, with different colors depicting the degree of strain, indicating tissue stiffness. The method is qualitative and does not give a measurement value. Fig. 5 shows an example of strain imaging: with B-mode imaging side-by-side, it is easy to appreciate how regions with different degrees of stiffness are impossible to single out using B-mode imaging. Thus, the mechanical properties of the various tissues are emphasized, irrespective of their echogenicity. Strain imaging is particularly

useful when screening for focal lesions, which often will have elastographic properties differing from healthy parenchyma. Correspondingly, it will be less useful when searching for subtle progression of (average) organ stiffness during clinical follow-up.

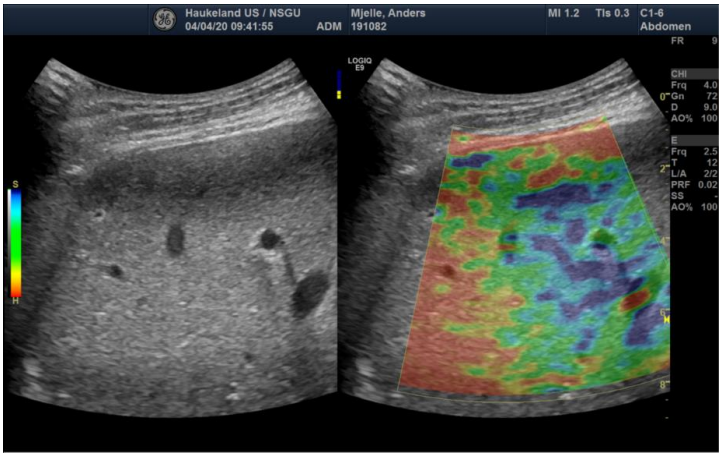


Fig. 5. Strain elastography of a healthy liver. *Image: Batman*

A dynamic variant of strain imaging is acoustic radiation force impulse (ARFI) imaging, where the same transducer both monitors tissue displacement and generates the original push pulse.

In **point shear wave elastography (pSWE)**, which is available in many ultrasound systems and sometimes referred to as ARFI quantification, an acoustic radiation force impulse described above produces perpendicular shear waves measured in a region of interest (ROI). The speed given is the average speed within the boundaries of the ROI (Fig. 6, Fig. 7). Conventional B-mode imaging is used, with no visualization of LS as in strain elastography or 2D-SWE. Thus, the operator cannot choose a measurement site based on assumed stiffness but uses the B-mode image to select a homogenous area free of vessels and visible biliary tracts.

Measurements are given in m/s but can be converted to Young's modulus, displaying values in kPa. Guidelines recommend using the median value of 10 acquisitions, similar to TE.

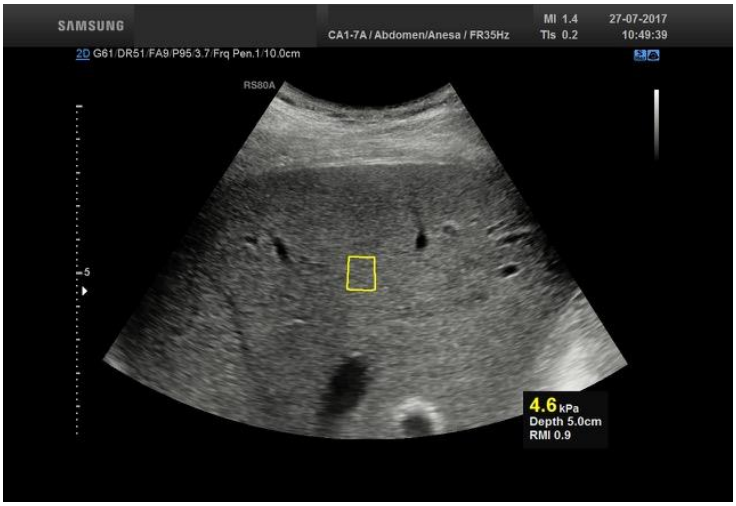


Fig. 6. Point shear wave elastography (pSWE) by Samsung RS80A.
Image: Anesa Mulabecirovic

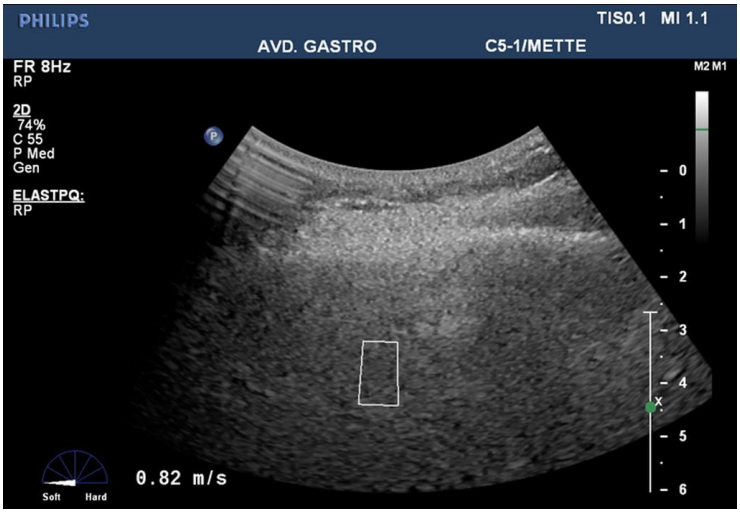


Fig. 7. Point shear wave elastography (pSWE) by Philips iU22. *Image: Mette Vesterhus*

Two-dimensional shear wave elastography (2D-SWE) is, similar to pSWE, incorporated into high-end ultrasound systems. Acoustic radiation force creates tissue displacement at all positions along the pulse axis [51]. A high frame rate allows following the shear waves in 2D, creating a vast and fine-meshed map of information, depicted as a color elastogram (Fig. 8). The operator determines the placement of the color elastogram, before placing a circular ROI in the most homogenous part of the elastogram. Having chosen the ROI site, an LSM value is generated. Guidelines recommend using the median value of at least three acquisitions, though this is subject to an ongoing debate.

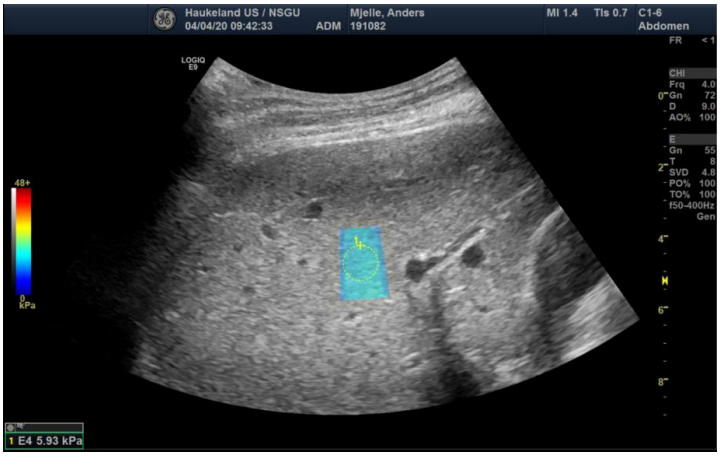


Fig. 8. Two-dimensional shear wave elastography (2D-SWE) by GE Logiq E9. The color elastogram is seen as a light blue rectangle, with the yellow stapled ROI in its center. *Image: Batman*

Transient elastography (TE) is the most studied version of ultrasound elastography systems, launched as early as 2003. It has a system-specific transducer that applies a mechanical push at the skin level, causing an acoustic wave into the adjacent tissue. The ultrasound system tracks the arising shear waves, and a median value of 10 acquisitions is reported in kPa. The often-used criterion of IQR/median ≤30%, is derived from the Fibroscan system, where it

was first applied. TE does not produce a B-mode image, making the operator unable to evaluate signs of liver disease or select a homogenous measurement site free of vessels or lesions. Moreover, TE does not function in the presence of ascites, which is easily visualized during a conventional ultrasound examination.

TE transducers come in three different sizes: S, M, and XL. M is most commonly used, while S is developed for small children, and XL for the obese (with a skin-to-capsule distance of ≥ 25 mm).

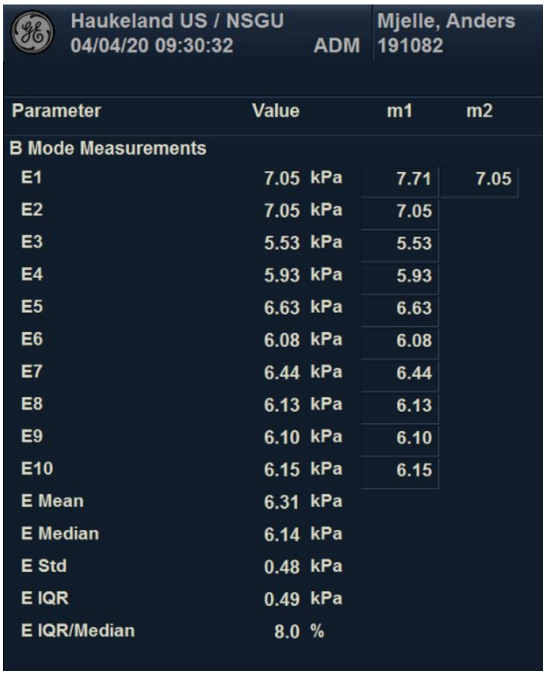
1.3.4 Clinical application of elastography

Liver elastography represents an estimation of LS, used to assess both the current level and the progression of liver fibrosis. There is a progressive replacement of healthy liver tissue with scar tissue in most chronic liver diseases, making the liver stiffer and less elastic. This is a continuous process, and surveillance of fibrosis development is a crucial part of clinical follow-up of patients with chronic liver disease. Inflammation plays a vital role in this process, sometimes contributing to the increased LSM, thus constituting a possible confounder in fibrosis evaluation. Inflammation causes increased LSM along two separate paths: an abrupt, often fully reversible LSM increase due to the ongoing inflammation process *per se*, and a chronic, less reversible LSM increase due to inflammation-induced scar tissue development. It is thus essential to look for ongoing acute inflammation when evaluating LSM values: significantly elevated liver transaminases (as an indication of inflammation; >5x of upper normal limit (ULN) is commonly used as a cutoff value) may indicate a falsely elevated LSM. Similarly, reduced LSM (e.g., when treating hepatitis C patients) can signify either reduced ongoing inflammation or actual regression of fibrosis.

The widely accepted criterion standard of liver fibrosis evaluation, the liver biopsy, has inherited limitations, such as sampling variability and its non-linear fibrosis grading [53, 54], with cirrhosis being missed by as much as 10–30% [55-57]. It is also an invasive procedure with the potential of clinical complications.

Furthermore, it necessitates general anesthesia in children, which should be avoided unless absolutely necessary, as it has been shown to affect nerve cells and learning and memory abilities [58, 59].

In all qualitative, dynamic elastography methods, a predetermined number of acquisitions are acquired: EFSUMB recommends different amounts for different methods, but always recommends reporting the median value. For TE, "10 measurements should be obtained"; for pSWE "at least 10 measurements" and for 2D-SWE "a minimum of three" [31]. A datasheet is used to show all acquisitions, with an automatic calculation of the median value (Fig. 9). The same guidelines recommend an interquartile range/median (IQR/M) ≤30% for valid measurements (See 3.3 DEFINITIONS) for TE and pSWE. EFSUMB states that there is no agreement on the use of quality criteria for 2D-SWE, but mentions that the IQR/M criterium is used in several studies on 2D-SWE and that the producer has recommended its use for 2D-SWE by GE Logiq E9 [31].



Parameter	Value	m1	m2
B Mode Measurements			
E1	7.05 kPa	7.71	7.05
E2	7.05 kPa	7.05	
E3	5.53 kPa	5.53	
E4	5.93 kPa	5.93	
E5	6.63 kPa	6.63	
E6	6.08 kPa	6.08	
E7	6.44 kPa	6.44	
E8	6.13 kPa	6.13	
E9	6.10 kPa	6.10	
E10	6.15 kPa	6.15	
E Mean	6.31 kPa		
E Median	6.14 kPa		
E Std	0.48 kPa		
E IQR	0.49 kPa		
E IQR/Median	8.0 %		

Fig. 9. Datasheet on GE Logiq E9 from scanning of a human liver. *Image: Batman*

1.4 THE LIVER

1.4.1 Liver physiology

The liver is an intra-abdominal organ, weighing as much as 1.5 kg in adults, in mature newborns between 0.1 and 0.2 kg and in our smallest premature infants treated at our institution, unless aberrant fetal growth, 0.03 kg [60, 61]. In pregnancies with decreased fetal growth, the liver volume will be reduced [62-65], while in some conditions, it can be increased, e.g., in pregnancies complicated by trisomy 21 [66, 67] or maternal diabetes [68].

The liver bud appears in the middle of the 3rd week of gestation, and by week 10, the liver constitutes 10% of the body weight, primarily due to a vast amount of hematopoietic cells, producing red and white blood cells [69]. This activity decreases the last eight weeks, and at term, the liver makes up 5% of body weight. Bile is produced from week 12. The neonatal liver has similar functions as in adults but is initially hampered by immaturity, leaving the neonate particularly susceptible to hypoglycemia, hyperbilirubinemia, and impaired drug metabolism [70]. The two latter conditions influence one another, as hyperbilirubinemia will cause bilirubin occupying albumin binding sites, yielding increased available drug concentration, which increases the chances of toxic effects. The liver's enzymatic ability increases quickly, and drug dosage increments are often needed after only a few days or weeks of life. For some drugs, we see an increased metabolism during childhood, necessitating higher dosages per kilogram than adults, and pediatricians have to be skilled in dosing medications according to both body weight and age group.

The liver has a wide variety of physiological tasks vital for survival. It is part of the digestive system, synthesizing proteins and detoxifying metabolites, and is our largest and most important metabolic organ, sometimes termed the body's major biochemical factory [71, 72]. Amongst other, the liver A) processes carbohydrates, proteins, and lipids after absorption of nutrition from the intestine; B) synthesizes plasma proteins necessary for blood clotting and

transport of hormones and cholesterol; C) stores glycogen, fat, iron, copper and vitamins; D) activates vitamin D (together with the kidney), and E) secretes bile/bile salts.

The liver is part of the biliary system, alongside the gallbladder and biliary ducts. Bile is an aqueous alkaline fluid containing bilirubin, cholesterol, and bile salts (derivates of cholesterol). Bile flows through the biliary system and into the duodenum, where it promotes the absorption of fat. This action depends on two abilities: bile salts' detergent action, converting large fat droplets into a lipid emulsion, making it available for the pancreatic lipase; and its role in facilitating fat absorption through forming micelles, a practical vehicle for transporting water-insoluble substances such as free fatty acids and fat-soluble vitamins [71].

1.4.2 Liver fibrosis

To put it simply, liver fibrosis is hepatic scar tissue, representing the healing process after the damage has occurred. It is the inevitable result of almost all chronic liver diseases, depending on the etiology and certain host factors. Fibrosis is characterized by an extracellular matrix (ECM) increase, where the dynamic balance between formation and degradation is skewed, leading to increased ECM and fibrosis load. While the matrix amount is similar across different etiologies, different diseases show different distribution characteristics throughout the organ [73]. Comparing PSC with viral hepatitis, we find that the former is characterized by a concentric onion-skin like scar around the bile duct, causing degeneration and bile duct stricture, while the latter by fibrosis deposition initially in the portal tracts, in time with periportal and bridging fibrosis (fibrous septa between lobules) [74]. Fibrosis commonly occurs due to chronic inflammation induced by oxidative stress, autoimmune reactions, ischemia, trauma, or as a spontaneous process [75]. The development is a rather slow process, requiring months and sometimes years in the context of ongoing inflammation or damage (although exceptions exist) and is thus absent after acute injury, even in fulminant hepatitis [73]. Cirrhosis is considered a late

product of fibrosis development. Still, fibrosis and cirrhosis are both parts of a dynamic process with some degree of reversibility, an ability which lessens as the disease progresses. Cirrhosis represents a distortion of liver architecture, and in cases of advanced cirrhosis, it is considered irreversible, with liver transplantation as the only viable treatment option. It is not established precisely when the condition becomes irreversible [76, 77]. With this change in architecture, portal hypertension follows, contributing to increased LSM, complicating interpretation of these measurements.

Although advanced cirrhosis is not very available to treatment, milder liver fibrosis can be substantially reversible in cases of treatable underlying causes, as demonstrated in several studies. In a large study on hepatitis B patients with liver fibrosis or cirrhosis, 87% showed histological improvement, and, more impressively, 74% of cirrhosis patients had no cirrhosis on follow-up biopsy (≥1 unit score decrease) [78].

Antifibrotic treatment is a major area of research in hepatology, aiming to develop liver-specific medications promoting resorption of fibrous tissue, without affecting the healthy parenchyma [79]. Animal studies have shown that degrading hepatic fibrous tissue is possible (156, 158), and in humans, medications such as obeticholic acid have been shown to reduce fibrosis in non-alcoholic steatohepatitis (NASH) [80]. Obeticholic acid is also applied in primary biliary cholangitis (PBC), hitherto showing improved serum liver tests, but high-quality randomized controlled studies are lacking [81]. Recently the same medication was shown to reduce alkalic phosphatase in PSC as well [82].

Several histologic scoring systems exist, and liver fibrosis is often characterized using a scale of F0–F4: F0 = no fibrosis; F1 = portal fibrosis without septa; F2 = few septa; F3 = numerous septa without cirrhosis; and F4 = cirrhosis. For PSC, three different scoring systems are applied: Ishak, Nakamura (developed for primary biliary cholangitis, PBC), and Ludwig and Batts, all showing histological staging of PSC to be associated with transplant-free survival [83-85].

1.5 PRIMARY SCLEROSING CHOLANGITIS

1.5.1 Epidemiology and demographics

The prevalence of PSC in northern Europe and the US is approximately 1 in 10'000, while much lower in southern Europe or Asia, with an estimated prevalence of 0.02–0.1 in 10'000 [86, 87]. The incidence lies between 0.4 – 2.0/100'000/year [88], and it has been speculated that it is rising, although some attribute this *partly* to better diagnostics [89-92]. The rising incidence is part of a general uprise in inflammatory disease, for which there are several suggested mechanisms [85, 93], but certainty is lacking.

PSC patients typically present in their 30s or 40s. The disorder is rare in children and adolescents, with a calculated incidence of 0.2/100'000 person-years [94, 95], but may be underestimated since the cholangiographic anatomy of liver disease remains to be well described [96]. Furthermore, the diagnosis of PSC may be delayed as liver pathology is rare in children, and thus often not considered early on, and the diagnosis is often seen in conjunction with autoimmune hepatitis, with the initial focus on the latter.

Studies find that 2/3 have concomitant inflammatory bowel disease (IBD) in Europe and North America, compared to only 1/3 in Japan [85, 97]. The frequency of IBD is notoriously difficult to determine as the two entities do not necessarily present themselves together: IBD can appear after liver transplantation for PSC, while PSC can be diagnosed after colectomy for ulcerative colitis [98-101]. Thus, the removal of one entity does not offer the patient full protection against the other one. One in four may have other autoimmune diseases [102]. Male patients are overrepresented, but studies suggest that this mainly reflects that women tend to have less symptoms [103, 104].

1.5.2 Pathogenesis

In summary, PSC is a mix of inflammation, fibrosis, and cholestasis, with a multifactorial and poorly understood pathogenesis, hindering the development of an effective therapy [85]. A typical description is a pathologically activated immune system with genetic and environmental factors, linked to the following:

- autoimmunity
- bile homeostasis
- leakage of gut bacteria or bacterial products into the enterohepatic circulation

The most striking pathogenic features are seemingly autoimmune, even though immunosuppressants do not alleviate symptoms or affect disease progression: prednisolone, budesonide, colchicine, penicillamine, azathioprine, ciclosporin, methotrexate, mycophenolate, and anti-tumour necrosis factors monoclonal antibodies have all been investigated without success [105, 106].
Genetic factors are involved, with increased risk in siblings of PSC or IBD patients [107], and more than 20 genes have been identified [108-116].

All these possible explanations are flawed by inseparable elements of contradictions: an autoimmune disease affected by no known anti-inflammatory medications; a condition with cholestasis and bile toxicity, though no effect of ursodeoxycholic acid treatment; a disease with inflammation affected by gut bacterial products leaking from the intestine, but with no halt in disease progression following removal of the intestine.

No matter the specific etiology, PSC is characterized by progressive inflammation affecting the biliary tree, with biliary strictures and fibrosis. The fibrosis is described as concentric layers surrounding the cholangiocytes lining the bile ducts, often referred to as “onion skin” scarring or periductal fibrosis. This finding is considered a hallmark finding of PSC, but may not be visible in the early stages of the disease [117]. Histopathology may support the diagnosis but is considered obsolete as part of diagnostic practice, rarely offering relevant

contributions if performed [118]. It should, however, be part of the diagnostic armamentarium for atypical or difficult cases.

1.5.3 Symptoms, diagnostics, and follow-up

The development of PSC is typically insidious, rarely symptomatic in the early stages. Symptoms develop over time and include right upper quadrant pain, pruritus, jaundice, fever, fatigue, and weight loss (11,15). Suspicion is often raised in the following settings: incidental finding of abnormal liver biochemistry; during diagnostics of patients with newly diagnosed cholestasis, cholangitis or IBD; (in more severe cases) liver failure, including variceal bleeding or ascites; malignancy (cholangiocarcinoma) [119].

The preferred method for diagnosing PSC, is magnetic resonance cholangiopancreatography (MRCP), with endoscopic retrograde cholangiopancreatography (ERCP) typically reserved for cases with biliary strictures requiring intervention or tissue sampling [119]. Liver biopsy is rarely recommended, mainly used when suspecting overlap variants with autoimmune hepatitis, when small duct PSC is suspected, or when there is doubt regarding the diagnosis. Since as much as 2/3 of PSC patients have IBD, colonoscopy is part of the diagnostics, even in asymptomatic patients.

Although no definitive treatment exists besides liver transplantation in cases of malignancy or end-stage liver disease, a life long follow-up, preferably in a center participating in clinical trials, is recommended [119]. However, the latter part of this advice is not always followed, including in Norway. Follow-up includes screening for liver failure, fibrosis development, and hepatobiliary and colorectal malignancy, and risk assessment regarding poor nutrition, weight loss, vitamin deficiencies, and osteoporosis. Annual ultrasound investigation of the gallbladder, to search for polyps or other pathology, is recommended in all patients. Colonoscopy is performed regularly in patients with and without an IBD diagnosis, annually/biannually or every 5 years, respectively [119].

The mean time from diagnosis to death or liver transplantation is highly variable, ranging from 10 to 22 years in different studies [89, 92, 120-122].

1.5.4 Non-invasive risk stratification

The first PSC specific prognostic model was published by the Mayo Clinic in 1989 [123]. The revised Mayo risk score is currently the most widely used worldwide: an advanced formula based on age, variceal bleeding, and specific blood tests (bilirubin, albumin and aspartate aminotransferase (AST)), although not validated at the individual level [124]. The Enhanced Liver Fibrosis (ELF) test is a specific liver fibrosis test developed to assess fibrosis in chronic liver diseases, recently validated in several independent PSC cohorts, where it has been shown to predict transplant-free survival [6, 7]. However, these tests have yet to be proven valuable in deciding treatment and further investigations in individual patients.

LSM is in some centers part of routine follow-up. Although not part of the recent British recommendations on PSC [119], it is recommended in international guidelines to evaluate liver disease severity [125]. LSM using transient elastography (TE) has been independently validated in PSC, proven to correlate well with the histological degree of fibrosis (in particular when severe) and to be associated with clinical outcome both at baseline and over time [4, 5, 126]. Diagnostic accuracy by TE for severe fibrosis and cirrhosis in PSC was shown to be superior to prognostic scores such as aspartate aminotransferase/platelet ratio index (APRI), Fibrosis-4 score (FIB-4) and the revised Mayo risk score, and found to increase with time [4]. Due to this, TE has been named one of the preferred candidates for surrogate endpoints by the International PSC Study Group [127]. In this thesis, we describe - for the first time - the use of pSWE and 2D-SWE in PSC patients (Papers I & IV).

2. AIMS OF THE THESIS

Liver elastography data for healthy individuals and PSC patients are scarce. The ultimate aim was **A)** to evaluate elastography in children and **B)** to establish LSM as a safe and effective tool for disease monitoring in PSC.
Our specific aims were:

1. To explore liver elastography in PSC patients
    a. *To describe the use of pSWE in PSC, including feasibility and LSM levels* ***(Paper I)***
    b. *To compare feasibility of pSWE between liver lobes* ***(Paper I)***
    c. *To describe the correlation between LSM and other signs of liver fibrosis: biochemistry, risk scores, B-mode findings* ***(Papers I & IV)***
    d. *Assuming that 2D-SWE, with its color elastogram, would offer advantages in a PSC setting (due to several PSC specific factors: scattered fibrosis, dominant stenosis with cholestasis, and transitory cholangitis), to compare 2D-SWE, pSWE, and TE head-to-head in a PSC cohort* ***(Paper IV)***
    e. *Describe the change of LSM over time in a PSC cohort* ***(data remains to be published)***
2. To explore liver elastography in healthy participants
    a. *To establish reference values for LSM in healthy children and adults (aged 4-70 years) using pSWE, 2D-SWE, and TE* ***(Papers II & III)***
    b. *To compare the different elastography systems head-to-head: LSM levels and feasibility* ***(Papers II & III)***
    c. *To evaluate any effect on LSM by age, gender, or body mass/composition, or by different observers* ***(Papers II & III)***

3. METHODS

3.1 STUDY DESIGN AND SELECTION OF PARTICIPANTS

3.1.1 Papers I and IV

Invited patients were adult non-transplanted PSC patients in western Norway, prospectively included 2013-2014 (Study I) and 2017-2018 (Study IV). All examinations were performed at Haukeland University Hospital, Bergen. For both studies, we used the same cohort, but the number of individuals grew between intervals: for study I, we included 55 PSC patients (38 males, 69.1%; mean age 46.4 years), compared to 66 PSC patients in study IV (51 males, 77.3%; mean age 49 years). In study I, 24 healthy controls were included.

The medical history was taken, and a B-mode ultrasound examination was performed. Patient records were searched for clinical information, including complications of liver disease (ascites, encephalopathy, esophageal bleeding). Blood was sampled, and standard biochemical analyses were performed as part of every visit. APRI and FIB-4 scores were calculated at both intervals. The ELF test was analyzed for the second interval (Study IV). For study I, LSM was performed using pSWE (Philips iU22); for study IV, we used three different systems: pSWE, 2D-SWE (GE S8), and TE (Fibroscan). Two observers investigated all patients (pSWE: MV; 2D-SWE and TE: ABM).

3.1.2 Paper II

A total of 246 healthy children aged 4–17 years were recruited from September 2017 through January 2018, see flowchart (Fig. 10). All subjects were healthy and were divided into predefined age groups: 4–7; 8–11; 12–14 and 15–17 years. All examinations were performed at Haukeland University Hospital by a single pediatrician (ABM), who recorded the medical history and performed a clinical examination including height, weight, and waist circumference. B-mode

scanning was performed before LSM, including scanning for signs of liver steatosis, comparing attenuation of the liver and right kidney (Fig. 11). Three subjects were excluded due to either steatosis and/or splenomegaly (n=2) or lack of fasting prior to the examination (n=1), leaving 243 for final analyses (108 boys, 44.4%). In all subjects, pSWE (Samsung RS80A with Prestige) was performed, followed directly by 2D-SWE (GE Logiq E9).

In a subset of children aged 8–17 years (n=87), TE (Fibroscan) was performed following pSWE and 2D-SWE. In another subset from all age groups (n=50), two observers performed pSWE and 2D-SWE measurements for interobserver analysis. Blood tests were not performed.

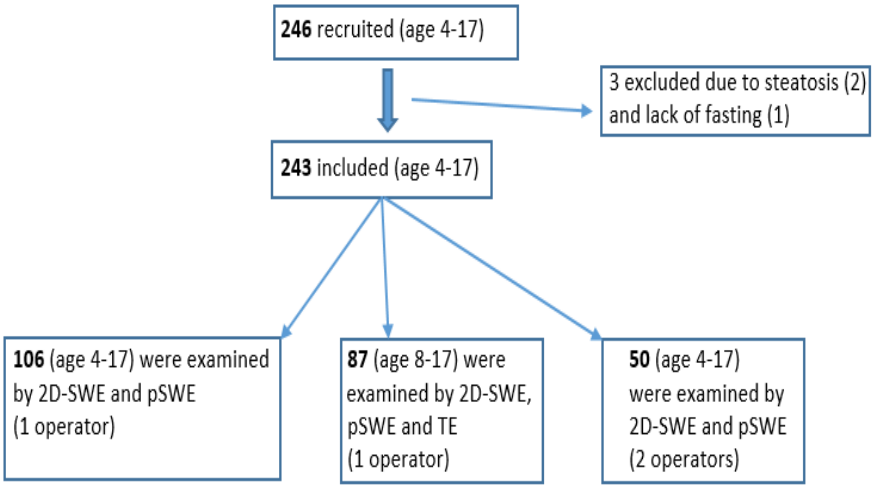


Fig. 10. Flowchart, study II.

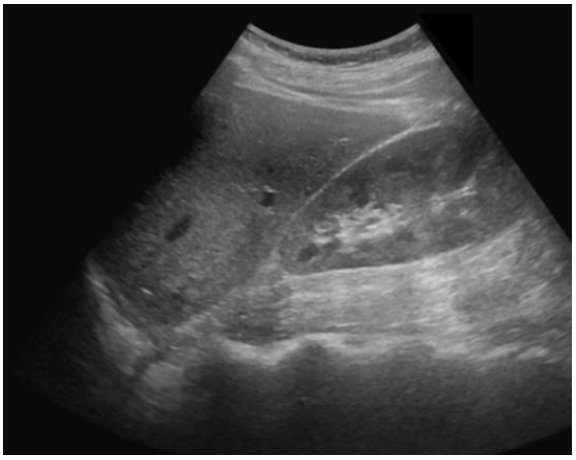


Fig. 11. B-mode image showing liver and kidney tissue, for comparison of attenuation. *Image: Batman*

3.1.3 Paper III

One hundred healthy adults were recruited and divided into five age groups (20-30; 31–40; 41–50; 51–60 and 61–70 years) with ten men and ten women. LSM was performed using three systems: pSWE (Samsung RS80A with Prestige), 2D-SWE (GE Logiq S8), and TE (Fibroscan). A B-mode ultrasound examination and blood tests were performed in all subjects. Height and weight were recorded, and participants were divided into normal (18–25 kg/m^2) and high BMI (25.0–30 kg/m^2). A flowchart for study III is shown in Fig. 12.

For interobserver analysis, pSWE and 2D-SWE were performed by two observers in a subset of 24 subjects,

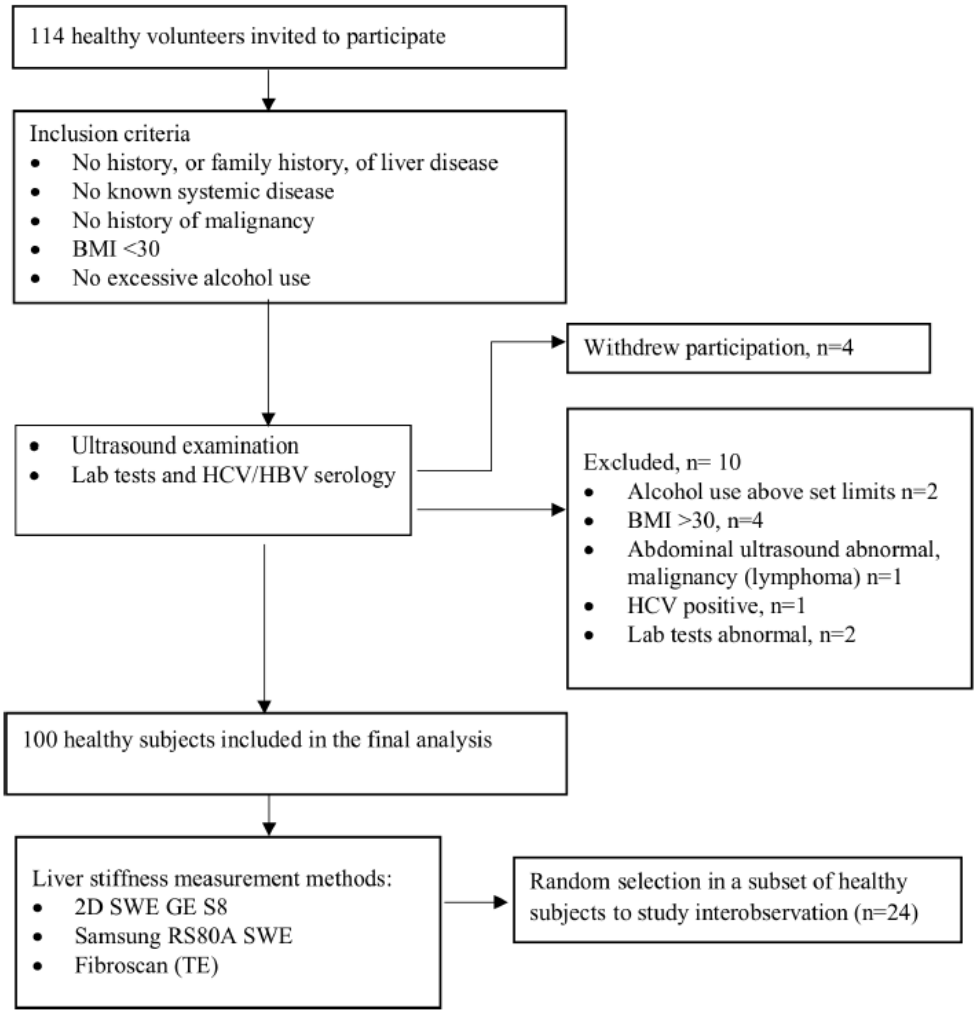


Fig. 12. Flowchart, study III. *Reprinted with permission from Paper III [128].*

3.2 ULTRASOUND EQUIPMENT

In *Study I*: Philips iU22 (ElastPQ, Philips Healthcare, Andover, MA, USA), software version 6.3.2.2, with a convex C5-1 probe.

In *Study II*: Samsung RS80A with Prestige (Samsung Medison Co, Ltd, Seoul, Korea), with a CA1-7A probe; GE Logiq E9 (GE Healthcare, Milwaukee, Wisconsin, USA), with a C1-6 probe and Fibroscan (Echosens, Paris, France; incorporated into a GE Logiq S8), using an M-probe.

In *Study III*: Samsung RS80A with Prestige (Samsung Medison Co, Ltd, Seoul, Korea), with a CA1-7A probe; GE Logiq S8, with a C1-6 probe and Fibroscan (Echosens, Paris, France; incorporated into a GE Logiq S8), using an M-probe.

In *Study IV*: Philips iU22 (ElastPQ, Philips Healthcare, Andover, MA, USA), software version 6.3.2.2, with a convex C5-1 probe; GE Logiq S8, with a C1-6 probe and Fibroscan (Echosens, Paris, France; incorporated into a GE Logiq S8), using an M or XL probe.

3.3 DEFINITIONS

The *body mass index* (BMI) is a unit defining the proportion between height and weight, used as a predictor of obesity. It is defined as the body weight in kilograms divided by the height in meters squared: kg/m^2. Although a validated predictor, it is must be converted into z-scores in children, as a BMI of 18 is considered overweight in a five-year-old child, while underweight in an adult. Furthermore, an increased BMI will not tell you if it is due to fat or muscle tissue.

The *median* is used to combat extreme values, easily affecting the mean value, pulling it away from the true value. When using the median value, 50% of all values are below the median and 50% above. The median is thus sometimes referred to as the 50^{th} percentile.

The *IQR/M* is the interquartile range divided by the median, given in percent. Just as the median is the 50th percentile, we commonly use the 25th percentile and the 75th percentile, which are the values with 25% and 75% of values directly below, respectively. Since 25% is a quartile, the 25th percentile is the lower quartile, and the 75th percentile is the upper quartile. The values between the 25th and 75th percentile are the middle 50% of the data points. This is the interquartile range (IQR) and is a measure of the dispersion of the central data points (Fig. 13). This must be seen in relationship with the median value, and it is thus divided by the median. When creating a boxplot, the lower limit of the box is the lower quartile, while the upper limit is the upper quartile (and the end of the lines represents the lowest and highest observation in the dataset).

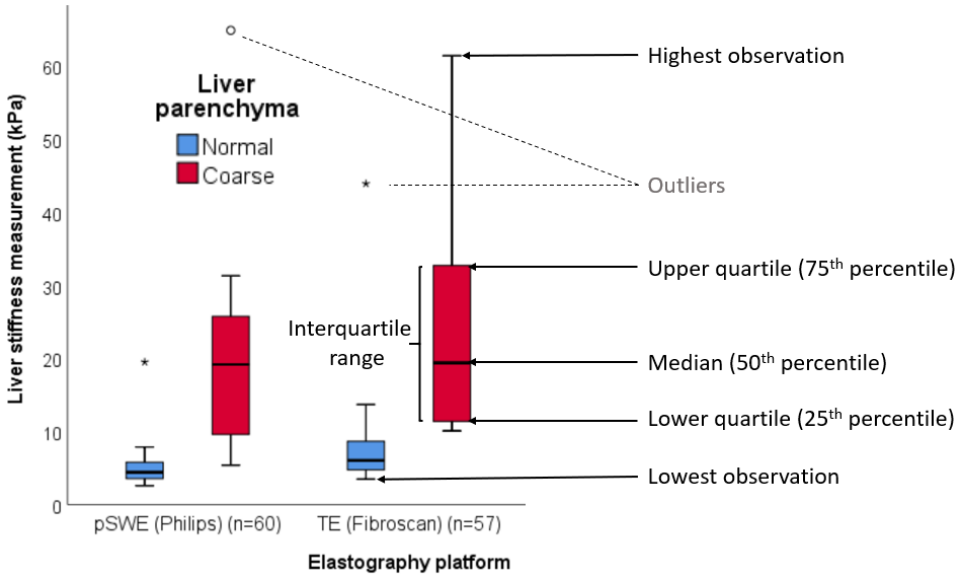


Fig. 13. Example of a boxplot.

Young's modulus is a term used in physics and is often mentioned in papers on elastography. It is merely a description of the relationship between *stress* and *strain*.

Imagine the liver (black triangle, Fig. 14) with a given length, L_0. When a force is applied as shown, the length will increase to L_n. Stress is the force applied (relative to the size of the object: in this case, the liver), while strain is the relative deformation (L_n minus L_0). In this figure, the horses exert a particular force, which is the stress, and the difference in length between the red line (L_n) and the black line (L_0), is the strain. Young's modulus thus describes the object's properties subjected to this force, how it withstands stress, fighting against deformation. In the figure, we see a relatively big difference, with the L_n approximately 2x L_0. The higher this difference becomes with a given stress, the lower Young's modulus will be.

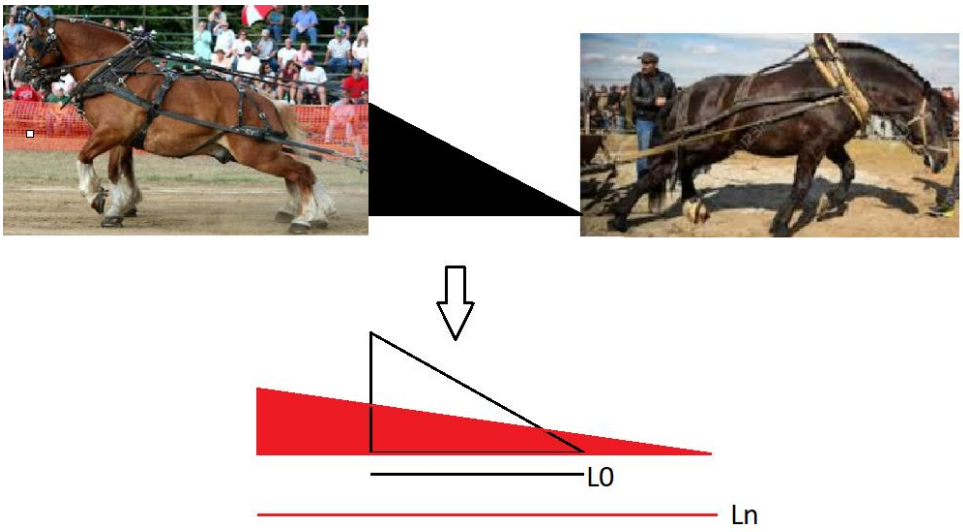


Fig. 14. Young's modulus is defined as stress/strain. Stress is the force (the pull of the horses) relative to liver's size (black triangle), while strain is the change in length. The liver deforms substantially (red triangle), doubling its length (L_n-L_0), indicating a low Young's modulus (i.e., a soft liver). *Image: Batman.*

3.4 PERFORMING THE STUDIES

For all systems, the LSM was defined as the median of 10 acquisitions. LSM was performed following B-mode ultrasound scanning, selecting a homogenous area of liver tissue free of visible vessels or biliary ducts. In study II, smooth logistics was emphasized to ensure optimal cooperation from the participating children: children were placed on the examination table and moved between ultrasound machines. As always, when examining children, there was an emphasis on keeping a pleasant atmosphere, communicating with both the children and their legal guardians, explaining everything thoroughly. Children rarely welcome surprises in a hospital setting.

Reliable elastographic measurements were defined according to the specific manufacturer's recommendations: a success ratio ≥60% (i.e., a minimum of 6 valid acquisitions for every 10 attempts) for all systems and an IQR/M ≤30% for all systems except Philips iU22.

3.5 STATISTICAL METHODS

Statistical analyses for all studies were performed using SPSS (SPSS Inc, Armonk, NY), versions 22, 24 or 25, and MedCalc version 12.7.0.0 (MedCalc Software, Mariakerke, Belgium). The candidate performed all analyses in study I, II, and IV.

All variables were tested for normality using the Shapiro Wilk test and the Q-Q-plot. Data were presented as either mean (standard deviation, SD) or median (range). Student's t-test or the Mann-Whitney U-test were applied as appropriate. We used the Pearson or Spearman's rank correlation coefficient as appropriate. P-values less than 0.05 were considered statistically significant.

The limits of agreement method with Bland-Altman plots was used to test for inter- and intraobserver agreement. The intraclass correlation coefficient (ICC) was used to assess interobserver reliability (studies II and III) or inter-system correlation (study IV).

3.6 ETHICAL CONSIDERATIONS

Oral and written information was given to all participants. Participation was always voluntary, irrespective of age. Written consent was obtained from all participants and legal guardians when the participants were too young to give legal consent (<16 years). Protocols were in accordance with the Declaration of Helsinki and approved by the Regional Committee on Medical and Health Research of Western Norway.

For PSC patients, studies were part of a routine clinical follow-up of chronic liver patients, with little risk of adverse events caused by the study, but would cause an extra visit for some patients.

There was a risk of incidental findings for the healthy participants, which could raise concern in the participants and their parents and guardians and possibly lead to unnecessary investigations and treatment. This was discussed prior to the study start and was part of the information to the individual participants and their parents.

The risk of such incidental findings is much smaller in children compared to adults. This was part of the evaluation when we deemed the study to have more possible benefits than disadvantages. There were no serious incidental findings – but liver steatosis, splenomegaly, gall stones, and gall bladder polyps were found in a few select subjects and followed up according to clinical guidelines.

4. SUMMARY OF RESULTS

4.1 PAPER I

Paper I is the first report of liver pSWE in adult PSC patients. We recruited 55 non-transplant PSC patients (38 males, 17 females; mean age 46 years) and 24 matched controls. B-mode ultrasonography and blood sampling for biochemical analysis were performed on the same day. SWV was measured using Philips iU22 in both the right and left liver lobe and the spleen. In 16 healthy controls, intra- and interobserver variation were investigated, with right liver lobe LSM performed twice by one operator or once by two independent operators, respectively.

Overall, we demonstrated that pSWE was feasible in PSC patients. Intra- and interobserver variability were low. Valid LSM in the right liver lobe was achieved in all patients and controls. Left liver lobe SWV showed a low success rate (valid LSM in 66% of patients) and did not differentiate patients and controls. PSC patients were found to have a higher median SWV in the right liver, compared to healthy controls (1.26 vs. 1.09 m/s), with the optimal cutoff as decided by Youden's index of 1.24 m/s (~4.6 kPa), yielding a sensitivity and specificity of 56.4 and 95.8%, respectively. Right liver LSM correlated with indirect fibrosis scores (APRI and FIB-4), but not significantly with the Mayo risk score (rho 0.296, p=0.06). Right liver lobe SWV was significantly higher in patients with B-mode signs of liver fibrosis (n=21, 38%), coarse liver parenchyma, irregular liver capsule, and periductal fibrosis (p-values 0.002, 0.001 and 0.049, respectively).

PSC patients without visible B-mode signs of liver fibrosis had an average LSM similar to healthy controls, but 12 patients without B-mode signs of fibrosis, expressed an LSM above our proposed cutoff value for increased liver stiffness, >1.24 m/s. The presence of splenomegaly (n=19, 35%) or bile duct dilatation did not impose a significant effect on LSM (p=0.11 and 0.61, respectively).

4.2 PAPER II

Paper II was the first report comparing three principally different elastography systems investigated head-to-head in children and established reference values for LSM in children for 2D-SWE, pSWE, and TE. A total of 243 children aged 4–17 years were included. All children were examined clinically, including height, weight, and waist circumference. Following a B-mode scan, pSWE and 2D-SWE were performed in all subjects, while TE was performed in a subset (n=87) aged 8–17. In 50 children, pSWE and 2D-SWE were performed by two independent operators.

Feasibility was shown to be high across all systems: 242/243 (99.6%) for 2D-SWE, 238/243 (97.9%) for pSWE, and 83/87 (95.4%) for TE. TE feasibility was significantly lower than 2D-SWE, but not different from pSWE. For the entire cohort, the average LSM value was 3.3, 4.1, and 4.1 kPa for 2D-SWE, pSWE, and TE, respectively.

LSM by 2D-SWE was lower than pSWE and TE, with no difference between the two latter systems. The inter-system difference between pSWE and 2D-SWE was not linear: in individuals with a low average LSM, pSWE was higher than 2D-SWE, while the opposite was true when average LSM was high, reflecting a steeper LSM value slope for 2D-SWE (Fig. 15).

Interobserver reliability analysis did not show any significant difference between observers for any system, with ICC of 0.83–0.84. Individual LSM differences between observers were rarely clinically relevant, with only 1 in 50 exceeding 1.6 kPa for 2D-SWE and none for pSWE.

Liver stiffness increased with age, reaching adult values in adolescence. We found a gender difference in adolescence, with boys having stiffer livers than girls, similar to some adult studies.

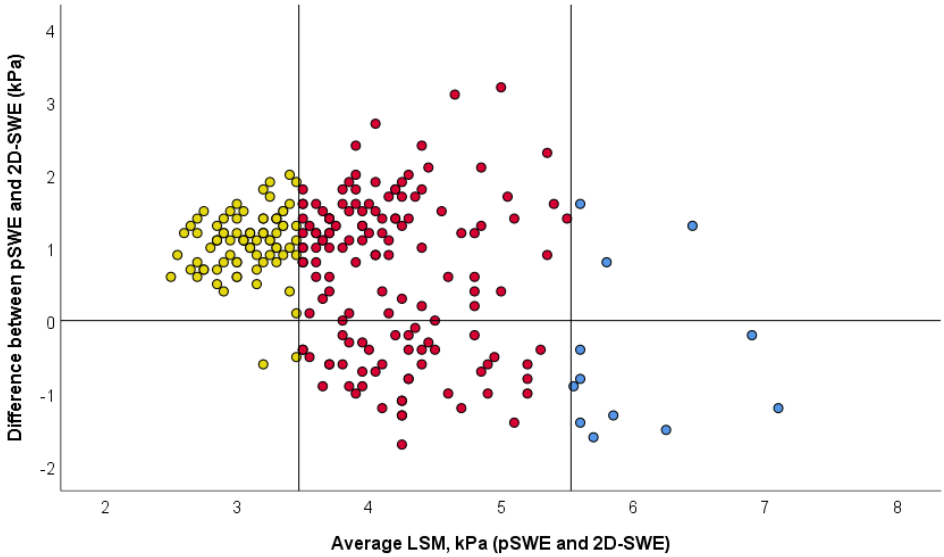


Fig. 15. Scatter plot showing the difference between pSWE and 2D-SWE against average LSM. Dots above the horizontal line (0-line), indicate a higher pSWE than 2D-SWE; while dots below the line mean that 2D-SWE is higher than pSWE. Dividing average LSM into <3.5 kPa, 3.5–5.5 kPa and >5.5 kPa (vertical lines), the number (%) of subjects with a pSWE LSM above the 2D-SWE LSM, was 80/82 (97.6%), 96/143 (67.1%) and 3/12 (25%), respectively.

4.3 PAPER III

The aim of study III was to establish reference values for LSM in healthy adults, using two new pSWE and 2D-SWE systems: Samsung RS80A and GE Logiq S8. TE was used as reference. One hundred healthy adults aged 20–70 were included, ten subjects for each ten year age period. B-mode ultrasound was performed before LSM, and blood was drawn for biochemical analyses in all subjects. Interobserver agreement was evaluated in 24 participants.

All methods showed perfect feasibility (100%). Average LSM was 4.5, 4.1, and 4.2 kPa for 2D-SWE, pSWE, and TE, respectively. LSM by 2D-SWE was significantly higher than pSWE and TE, while there was no difference between the two latter systems. There was no difference between average LSM between operators for pSWE, but a significant difference was found for 2D-SWE (4.5 vs. 5.1 kPa). However, the interoperator reliability was deemed good for both systems, with rho values 0.74 and 0.65 (p-values <0.001) for 2D-SWE and PSWE, respectively. Males had significantly higher LSM compared to females for TE and 2D-SWE. A similar difference failed to reach significance for pSWE (p=0.06).

There was no significant difference when comparing median LSM values using 5 or 10 acquisitions with pSWE and 2D-SWE.

4.4 PAPER IV

In Paper IV, we aimed to explore feasibility and LSM levels of three different elastography systems in PSC patients. Sixty-six patients were evaluated with B-mode ultrasonography, blood tests, and LSM by pSWE, 2D-SWE, and TE. Feasibility was 93.9%, 71.2, and 89.4% for pSWE, 2D-SWE, and TE, respectively, with corresponding median LSM values of 4.9, 7.1, and 6.4 kPa, respectively. LSM values correlated well across systems, with rho values 0.65–0.72, p-values <0.001. ICC was excellent for pSWE vs. TE (rho 0.91), while moderate for pSWE vs. 2D-SWE, and TE vs. 2D-SWE (rho values 0.49 and 0.43, respectively). Including normal weight patients only, the latter ICC values increased to 0.92 and 0.81, respectively.

Inter-system differences were not linear. 2D-SWE displayed higher LSM than TE when average LSM was low, while *lower* than TE when average LSM was high.

LSM values correlated significantly with liver enzymes and serum-based fibrosis scores, including the ELF test. The suggested ELF cutoff value of 11.2 discriminated well between high and low LSM. Similarly, LSM rose steeply when the Mayo risk score exceeded a value of 0.5, and LSM was higher in patients with B-mode signs of fibrosis.

There was no difference in LSM between those with splenomegaly and normal spleen length. However, applying a cutoff of 13 cm, we found a significant relationship between splenomegaly and LSM, and we argue that a universal cutoff value of 12 cm, regardless of gender or body size, should be reevaluated.

5. DISCUSSION

5.1 METHODOLOGICAL CONSIDERATIONS

This thesis delves into liver elastography, with liver stiffness measurements of both healthy individuals and chronic liver disease patients. Both European (EFSUMB) and global (WFUMB) guidelines have explained basic principles and technology, and given recommendations on the use of the different elastography technologies in practice [31, 51, 52, 129]. Its use has been advocated by the international liver disease community [125]. These guidelines are the basis of our practice both clinically and in this thesis.

Chronic liver disease leads to fibrosis, replacing soft liver tissue with stiff fibrous tissue. The degree of stiffness corresponds with the fibrosis stage, which corresponds with disease severity. An easy, non-invasive way of measuring stiffness represents a great leap forward in follow-up of chronic liver diseases.

New ultrasound models and systems are continuously being developed, necessitating new investigations and inter-system comparisons. LSM differs due to many differences: elastography methods; underlying algorithms and assumptions for calculation; probes/transducers; system users (interoperator variability); intraoperator variability; choice of measurement site/scattered fibrosis; other underlying biases, e.g., suboptimal measuring technique, non-fasting state, recent physical exertion, inflammation, cholestasis or right-sided heart failure. Some factors represent inherent sources of bias existing irrespective of system chosen and should be searched prior to LSM.

Every system, including different models and probes, needs to be evaluated before application in a clinical setting to establish reference values and values corresponding to stages of fibrosis. This thesis set out to establish reference values for healthy individuals, particularly children, with inter-system comparison, and to explore different elastography systems in PSC.

Reference values are typically determined by selecting healthy individuals using stringent methodology, collecting samples based on a predefined protocol, evaluating data distribution and defining the middle 95% as normal, while continually searching for errors and/or outliers [130]. These reference values do not necessarily correspond to medical decision limits, which may vary based on the need for test sensitivity or specificity in a given clinical setting.

When establishing a quantitative method, we have to consider both accuracy and precision [131]. A liver subject for evaluation will always have a *true* LS. The true value can never be established with absolute certainty; thus, we aim to establish an approximation of the true value. The true LS will have a mean value for the entire liver, while different regions may have different LS. Measuring LS through multiple acquisitions, results will concentrate around a mean value, with a variation that can be either low or high, thus indicating the quality of the analysis. In LSM, the variation is described using IQR/M – a high IQR/M indicates a less reliable measurement (See 3.3 DEFINITIONS).

In our case, accuracy would be how close the LSM (median value) is to the true LS, while precision describes the degree of variation. Fig. 16 illustrates the concept of accuracy and precision.

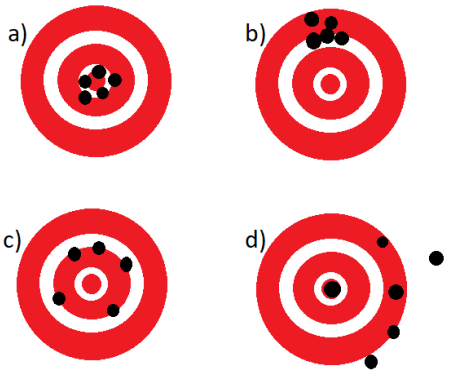


Fig. 16a-d. a) High precision, high accuracy; b) High precision, low accuracy; c) Low precision, high accuracy; d) Low precision, low accuracy. *Image: Batman*

A good test does not need both high precision and high accuracy, as long as the bias is known: in Fig. 16b, the measured value is not close to the true value, but this constitutes no problem if the difference is known.

Imagine that we have a patient with a true LS of 1.0 m/s and two different elastography systems. System A always shows measured LSM twice the true value, meaning a value of 2.0 m/s. System B yields different values: sometimes 0.8 m/s, sometimes 1.3 m/s, but always within 50% of the true value. Thus, the latter system yields a more accurate LSM, always closer to the true value, but it is impossible to know if the LSM at hand is below or above the true value. Knowing the systematic difference for system A, this system would, without doubt, be preferred. If this difference is unknown, one will choose system B, which gives a value closer to the true value.

It is a problem that clinicians often do not consider these biases in quantitative measurements, whether it measures serum lab values or LSM. The measured value will always be an approximation, and there will always be some variability. The value approximation, or the accuracy, will typically be handled by the producer based on internal control and research such as in this thesis, while the clinician should focus on the precision level: does a higher value during follow-up represent an actual rise, or is it just an expression of random variation? These factors are often subject to more control in laboratory medicine than in elastography, and there is obviously a need for more investigations.
In laboratory medicine, the internal control is strict, with calibration on a regular basis: at regular intervals, a test with a known value is analyzed to check for accuracy and precision. A similar system for elastography would mean checking accuracy and precision regularly using liver fibrosis phantoms with known stiffness values.

Investigating the accuracy of LSM is notoriously difficult since the true LS is unknown. The criterion standard is the liver biopsy, but this method is flawed by sampling errors and interobserver discrepancy [132, 133], meaning it does not

tell the full tale. Furthermore, it is an invasive procedure with serious complications seen in approximately 1% [134, 135], and necessitates general anesthesia in specific patient groups.

When investigating normal LSM in healthy individuals (Papers II & III), we are quite confident that all or close to all subjects were free of significant liver fibrosis, although this is impossible to prove. Performing liver biopsies would strengthen our belief's foundation but would also be unethical in this setting. Liver biopsies in our PSC patients (Papers I & IV) would be have been of interest as these patients often have liver fibrosis but would be unethical since there was no clinical benefit for the patients. In Papers II-IV, we did what is often done when lacking the possibility of liver histology: we used TE as a proxy criterion standard. TE has been proven to correlate with histological fibrosis score and is the most widely studied method. However, it is not evident that TE correlates better with histological fibrosis than pSWE or 2D-SWE, as pSWE, 2D-SWE, and TE show similar results when compared to biopsy-proven liver fibrosis levels [136-138].

5.2 LIVER ELASTOGRAHY IN HEALTHY INDIVIDUALS

A total of 343 healthy individuals were included: 243 children aged 4–17 years and 100 adults aged 20–70 years. One operator examined all children, another operator examined all adults, and the two operators were each others' interobservers. All subjects were examined with B-mode ultrasound scanning of the liver, kidneys, and spleen prior to LSM. Blood tests were performed in all adult participants to rule out liver disease but were deemed unethical in the children. This was compensated by a thorough anamnesis and clinical examination. All subjects underwent liver elastography using pSWE and 2D-SWE. TE was performed in all adults and a subset of children aged 8–17 years (n=87). A total of 81 subjects had valid measurements for all three systems.

		Mean LSM (normal range, mean±1.96 SD) (lowest observation), kPa				
		4–7 y	**8–11 y**	**12–14 y**	**15–17 y**	**Adults**
pSWE	Males	**3.9** (2.8–5.0)	**4.2** (2.7–5.6) (3.1*)	**4.9** (3.3–6.6) (3.4*)	**4,3** (2.6–6.0) (3.2*)	**4,2** (2.8–5.6)
	Females			**4.4** (2.9–5.8) (3.1*)	**4.2** (2.4–6.0) (2.9*)	**3.9** (2.1–5.7)
2D-SWE	Males	**2.9** (1.8–4.0) (2.0*)	**3.4** (1,4–5.5) (2.1*)	**4.1** (1.4–6.9) (2.5*)	**4.4** (2.4–6.5) (2.9*)	**4.7** (3.3–6.1)
	Females			**3.5*** (1.5–5.5) (2.0*)	**3.7** (1.8–5.6) (2.0*)	**4.3** (2.9–5.7)
TE	Males	–	**3,8** (1.9–5.6) (2.4*)	**4.6** (3.0–6.3) (3.5*)	**4.8** (2.8–6.7) (3.2*)	**4.5** (2.5–6.5)
	Females			**4.0** (2.6–5.4)	**4.2** (2.4–6.1)	**3.9** (1.7 – 6.1)

* lowest observation; given when higher than the lower end of calculated normal range

Table 1.

Liver stiffness measurement (LSM) values for all age groups and elastography systems.

For pSWE, mean values across age groups were 3.9–4.9 kPa, increasing significantly from 4–7 years to 12–14 years. LSM for age group 15–17 years was significantly lower than for 12–14 years (p=0.002). We have no reason to believe that LS decreases between these two age groups, supported by the fact that this finding only applied to pSWE, and not 2D-SWE and TE. However, LSM 15-17 years was similar to LSM in adults, and the finding of higher LSM 12-14 years compared to LSM 15-17 years and in adults, was made by two independent observers.

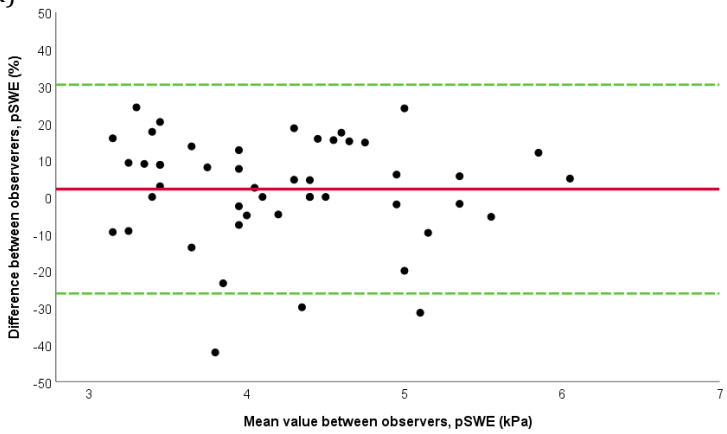
The upper limit of normal (ULN) ranged from 5.0 kPa in the youngest children to 5.7 kPa in female adults, with the highest ULN in boys aged 12–14 years (6.6 kPa). The feasibility was excellent: 97.9% in children and 100% in adults. Interobservation analysis revealed no difference between observers (medians

4.10 vs. 4.15 kPa) and an ICC of 0.83 (95% confidence interval (CI) 0.7–0.91), with no systematic differences. We found that 4/48 (8.3%) showed a discrepancy >1.0 kPa, with the highest value 1.6 kPa. Correspondingly, 7/48 (15%) showed a discrepancy >20% of the mean LSM between observers, with the highest value 42.1%. All values were still within the normal range – meaning that the discrepancy did not move healthy patients into a category where fibrosis ought to be suspected. A high discrepancy in percentage poses no clinical implications if the absolute LSM value is low, but is problematic in patients with a high LSM.

Bland-Altman plots visualize discrepancies, searching for any systematic bias. We can assume that in a stiff liver, there is an increased risk of a big difference in absolute values (since LSM values are higher) between two observations. However, in our normal material, there are only soft livers, albeit with variations in softness. The *relative* differences are not necessarily different in soft and stiff livers. However, due to system differences, e.g., software technology and underlying algorithms, measurements may increase more in some systems compared to others. E.g., when a tissue's true stiffness increases with a factor of 2x, one system may display an increase of 1.7x, another system 2.4x. A further increase may be linear or non-linear, e.g., another increase of 2x can result in LSMs of 3.4x and 4.8x or 3.4x (linear) and 6.7x (non-linear), respectively.

In Paper II, we used Bland-Altman plots with the difference portrayed in percent, but it is equally attractive with absolute LSM values for clinicians. Fig. 17a-b shows identical measurements displayed either with differences in % (a) and differences in absolute LSM values (kPa) (b).

a)



b)

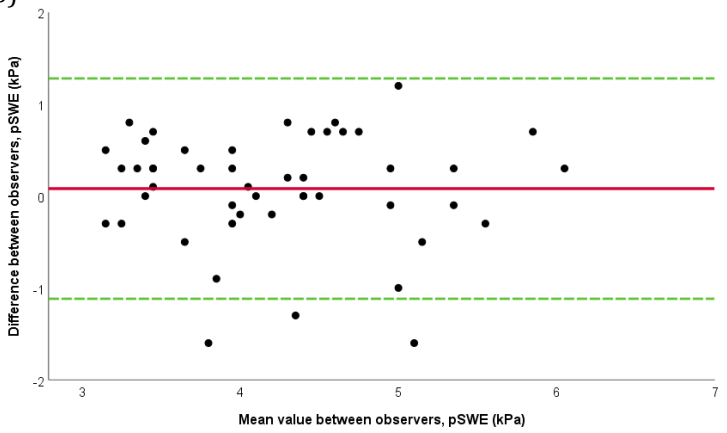


Fig. 17a-b. Bland-Altman plots with the average value between the two observers on the x-axis, and the difference between observer values, either in percent (a) or in absolute values (b), on the y-axis. The figures are created to look for systematic bias: is there any tendency of deviation or difference between observers depending on the average LSM? *Fig. 17a reprinted from Paper II with permission [139].*

In adults, the ICC for pSWE was 0.85, similar to the findings in children. Pearson's correlation coefficient was 0.74 for adults and 0.71 for children. We found 3/24 (12.5%) cases with an interobserver difference >1.0 kPa when using 10 acquisitions, compared to 1/24 (4.2%) when using 15 acquisitions. There was no observer bias for pSWE in children or adults. In adults, we found no significant difference in LSM when applying 5 or 10 acquisitions when measuring: 4.1 vs. 4.1 kPa (p=0.08).

For 2D-SWE, mean values across age groups ranged from 2.9 to 4.7 kPa. LSM rose steadily with age until age 12–17 years. The correlation between age and LSM was stronger for 2D-SWE compared to pSWE, with Pearson's correlation coefficient of 0.42 compared to 0.15 (p-values <0.001). ULN ranged from 4.0 kPa in the youngest children to 5.7 kPa in female adults, with the highest in boys aged 12–14 years (6.9 kPa). The feasibility was excellent: 99.6% in children and 100% in adults. Interobservation analysis found identical mean values between observers, with an ICC of 0.84 (95% CI 0.71–0.91), and no systematic bias. We found that 5/50 (10%) showed a discrepancy >1.0 kPa, with the highest value 2.2 kPa. Correspondingly, 12/50 (24%) showed a discrepancy >20% of the mean LSM between observers, with the highest interobserver difference of 51.3%; however, none of these discrepancies in LSM resulted in discrepant classification regarding the stage of fibrosis. In the subject with the highest percentage value, the LSM by observer A was 4.4 kPa, compared to 2.6 kPa by observer B. Performing a similar study in children with known liver fibrosis, with identical ultrasound systems and observers, would be of interest.

In the 24 adults, there was a significant difference between observers (4.5 vs. 5.1 kPa, p=0.009). The ICC, however, was good, with a value of 0.78. Our suggested explanation of this interobserver discrepancy is that 2D-SWE is particularly user-dependent and that there is a longer learning curve for this method compared to pSWE. However, this discrepancy was not present in the interobserver analysis of children, which took place only two to three months later. Furthermore, in a liver phantom study conducted one year earlier, the same two operators had no LSM discrepancy for 2D-SWE [140]. Thus, we cannot conclude on the reason for this discrepancy with any certainty.

An absolute interobserver difference >1.0 kPa was present in 2/24 (8.3%) when performing 10 acquisitions and 1/24 (4.2%) when performing 15 acquisitions. There was no significant difference in LSM applying 5 or 10 acquisitions when measuring in adults: 4.4 vs. 4.5 kPa (p=0.05).

For TE, mean values across all age groups were between 3.8 and 4.8 kPa, with values increasing significantly from 8–11 years to 12–14 years, and no significant difference between 12–14 and 15–17 years. The ULN ranged from 5.6 kPa in children aged 8–11 years to 6.5 kPa in male adults, with the highest ULN in boys aged 15–17 years (6.7 kPa). The feasibility was excellent, with 95.4% in children and 100% in adults. Interobservation analyses were not performed.

The interobserver differences seen across systems are essential to notice, with 8-12% deviating more than 1.0 kPa in healthy children and adults, for observations made only 5 minutes apart, with different observers. A recent study using TE in healthy children found 25% with an interobserver difference >1 kPa [141] for examinations >24 hours apart. Studies have shown that intraobserver ICC falls from 0.9 to 0.65 when comparing observations performed the same day or a few days apart [142, 143]. Interestingly, we find a reduced amount of clinically relevant interobserver differences when using 15 compared to the standard 10 acquisitions, though the number of participants (n=24) precludes firm conclusions. At the same time, we find no significant LSM difference using 5 acquisitions, compared to the standard of 10, for the individual observer.

Typically, studies report correlations and p-values, but this does not necessarily indicate the frequency of large individual discrepancies or whether these differences have clinical implications. Studies may report a very high overall ICC, while simultaneously reporting individual variability leading to patients moving between fibrosis groups F0-1 and F2-4, in one study 17% [144]. Studies report the ICC, but do not mention individual differences, whether for TE [145-147], pSWE [148-156] or 2D-SWE [142, 143, 157]. Some report agreements for fibrosis stages, ranging from 75-92% [145, 146, 155, 156]. In some cases, it is possible to find individual differences in tables or figures; however, even in studies with individual differences >10 kPa in healthy participants, this may not

be specifically mentioned in the manuscript [154]. An expected difference between expert investigators of 0.5 kPa has been suggested [142].

A near-perfect correlation using consistency instead of absolute agreement could still mean significant differences, no matter how low the p-value. In clinical practice, we need to know if there is a risk of values fluctuating >1-2 kPa, not whether the rho value of correlation is 0.92 or 0.97. Current guidelines highlight reproducibility, but only report the degree of correlation, and not the percentage of clinically significant differences in individuals [31, 129].

For all methods, the feasibility was high, ranging from 95.4–100%. High feasibility is described in most studies, with values commonly above 90–95%. A lower feasibility is described in very young children: one study reporting 66.7% in those <2 years, compared to an overall feasibility of 90% [158]. Feasibility is affected by several factors: mainly obesity, but also inexperienced operators, narrow intercostal spaces, ascites, high age, and female gender [159, 160]. In a clinical setting with the same subjects, the feasibility would probably be higher: in cases with an IQR/M above 30%, the operator can choose to erase all acquisitions and repeat a second time or carry on with acquisitions, choosing a total of 15 or 20 acquisitions, increasing the probability of a valid result. In a research setting demanding stringency, we strictly followed our protocol, but it would be interesting to perform studies with a protocol describing such actions in cases with invalid LSM.

There was a significant gender difference in adolescents aged 12–17 years across all three systems, in adults for 2D-SWE and TE, with males showing a higher LSM. Fig. 18 depicts this difference for 2D-SWE. A similar difference was noted for pSWE in adults without reaching significance (p=0.06). Adult studies have reported a gender difference, for both 2D-SWE [128, 161, 162], pSWE [151] and TE [128, 163, 164], but results are conflicting. European guidelines [31] mention gender only when describing pSWE, reporting no difference, while global guidelines describe higher LSM in males for TE [129]. All studies

describing a gender difference find that men have increased LS, never the opposite. Based on our findings, we believe that a true gender difference is likely.

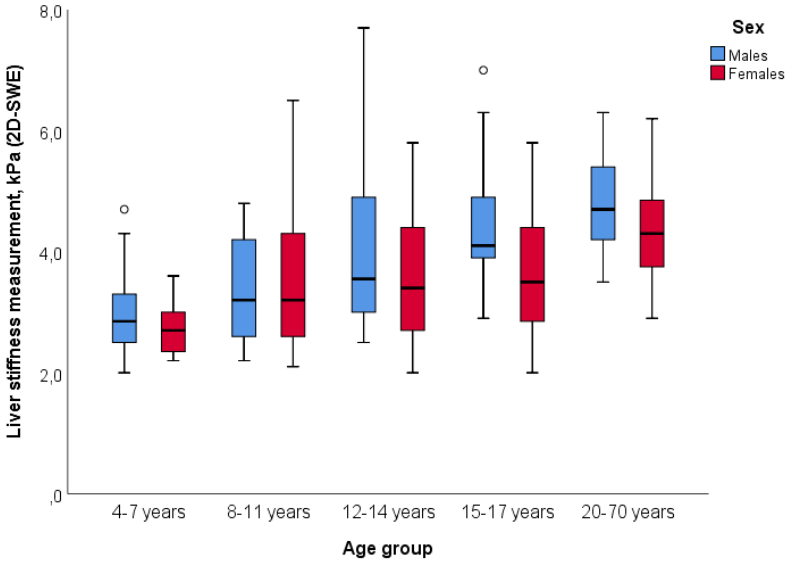


Fig. 18. Liver stiffness measurements by 2D-SWE for all age groups, with males in blue and females in red.

There is no definite explanation of why liver stiffness should be increased in men. However, we know that ovarian hormones, including estrogen, affect collagen synthesis, which is the basis for liver fibrosis development [165]. It has been shown in animal studies that the response to antifibrotic treatment is lower in females [166]. Likewise, estradiol treatment has been shown to reduce liver transaminase levels and suppress hepatic collagen content, with a negative relationship between the percentage of affected liver tissue and serum estradiol levels [167]. Female gender is regarded as protective against fibrosis, and menopause and late menarche have been shown to correlate with increased severity of liver fibrosis in untreated chronic hepatitis B infections [168]. Muscle stiffness, as measured by elastography, is higher during menstruation than during ovulation, indicating that fluctuating sex hormones affect tissue stiffness

[169]. The same has not been found for breast tissue [170]. As the described gender difference in LSM is observational and somewhat inconsistent, a study on LSM in women could help answer this question.

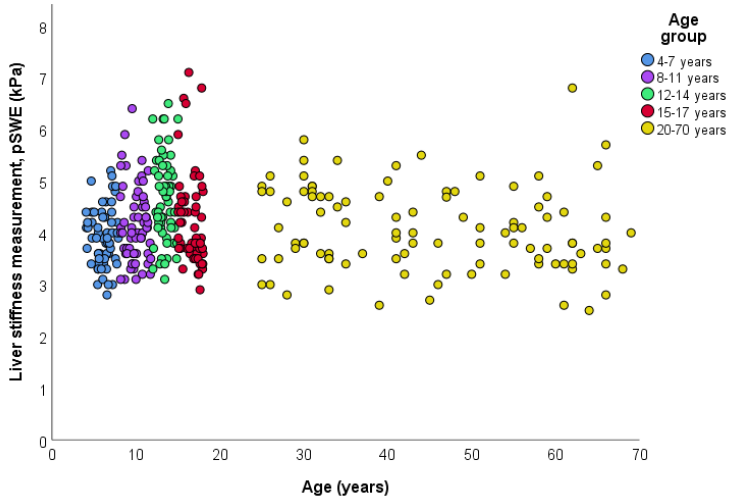
The finding of increasing LS through childhood was consistent across all systems in our study, with no difference between younger and older adults in the range 20 to 70 years. Previous studies have produced conflicting results in children, with several studies describing the influence of age [171-179], while some report no correlation [158, 180-183]. The largest pediatric study to date was recently published, describing a significant increase with age and a higher LSM in males [179]. One study concludes on no influence by age, while reporting increasing LSM in subjects above 10 years [184]. All these studies have been performed with different approaches, with different age groups and only one elastography system. Some studies describe higher LSM in the youngest children: one study in children less than one month [174], another in children <6 years [175]. Several studies find that children above 12 years have higher LSM than those under [171, 172, 176, 177], one describing a U-shaped curve with higher LSM in ages 0–2 and 12–18 compared to those 2–11 years [176]. There are only three large studies with no significant correlation between age and LSM [158, 181, 182], and all report a tendency of higher LSM in older children and adolescents. One study found a mean LSM of 4.6 kPa in children >6 years, compared to 4.1–4.2 kPa in younger children [158], another with mean LSM 1.14 m/s (equals 3.9 kPa) (95% CI 1.1–1.7 m/s) in children 10–17 years, compared to 1.08 m/s (equals 3.5 kPa) (95% CI 1.05–1.13 m/s) in those <10 years [182]. It is particularly difficult to compare studies since they define age groups differently. It would interesting if all studies used similar definitions or at least published scatter plots with individual values across all ages, as in Fig. 19.

Fig. 19 demonstrates several aspects regarding age influence on LSM values. In our studies, age seems to exert a more substantial influence on LSM by 2D-SWE than pSWE. We see a rising LSM from early childhood to adolescence, influenced

mainly by a raise of the top values more than a raising of the lower bar (the lower limit of normality is not much changed with age, while the upper limit is). For both systems, there are seemingly more outliers for adolescents than for adults. When comparing 12–14 and 15–17 years for pSWE, we see that the highest outliers are in the oldest age group. However, it is still evident that LSM 12–14 is higher than LSM 15–17 years, where values seem to stabilize below 5 kPa, similar to adult LSM. The same tendency is not seen for 2D-SWE.

The tendency of outliers in children and adolescents may be the result of several underlying factors. With more participants (243 vs. 100), there will be more outliers. It is also possible that sex hormone levels during puberty and adolescence contribute. We were prepared for more outliers in young children since cooperation can be difficult, but there was no such tendency. However, it was noted in the preparations to the adult study, that some, in particular lean and fit adults, had surprisingly high values. These values lowered significantly when participants were instructed carefully regarding mid-expiratory breath-hold. Participating children were instructed in mid-expiratory breath-hold, but no pre-study LSM was performed, and there was no new investigation in cases with a high LSM. Reinvestigation would possibly lower the outlier rate.

a)



b)

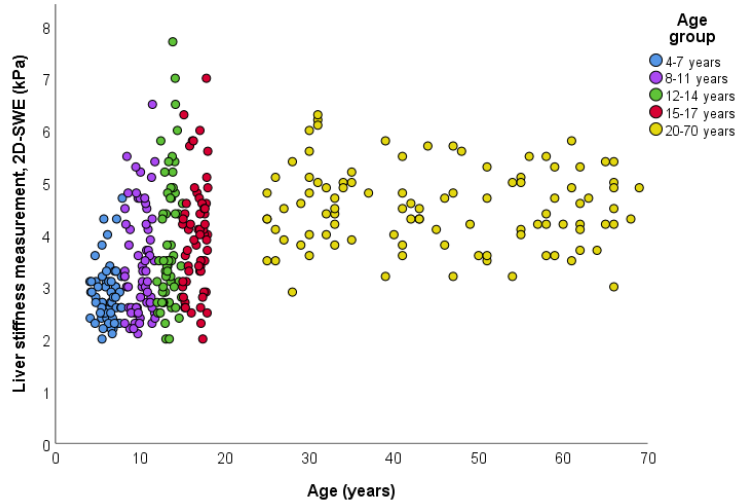


Fig. 19. Liver stiffness measurements according to age for (a) pSWE and (b) 2D-SWE. Of note, 2D-SWE was performed by two systems from the same producer: GE Logiq E9 in children and GE S8 in adults. pSWE is performed using the same machine across all ages. Two different observers performed LSM in children and adults, respectively.

We describe increasing LSM with age for all systems, from group 1 (4-7 years) to group 2 (8-11 years) to group 3 (12-14 years), where LS seemingly reaches a plateau after age 12. LSM in group 4 (15-17 years) was similar to LSM in group 3 for 2D-SWE and TE but to our surprise, significantly *lower* than group 3 for pSWE, for which we have no good explanation.

When comparing LSM by pSWE for 15–17 and 12–14 years, we used median values since data were not normally distributed: 3.9 kPa and 4.6 kPa, respectively, yielding a difference of 0.7 kPa. Some high outliers were noted in the older age group, which hindered normal distribution. Using mean values, the difference was reduced to 0.4 kPa. However, no matter how this is approached, LSM for the older adolescents remains lower. Thus, there is no question whether pSWE LSM is reduced in the oldest adolescents, but a question if this represents a true LS difference. Why we find this LSM difference, remains to be answered.

To explore this inconsistent finding of reduced LSM for pSWE, we performed post hoc analyses of LSM by observer B. The same phenomenon appeared (Fig. 20), but the statistical strength was substantially lower due to fewer participants. Mean LSM was 4.0, 4.5, 4.5 and 3.8 for age groups 4–7, 8–11, 12–14 and 15–17 years, respectively: LSM group 4 was significantly lower than group 2 and 3. It is possible to assume that the boxplots (For observer B: Fig. 20a, for observer A: Fig. 20b) reveal more information that the inter-group t-tests: we see a rising LSM with age from 4–7 to 15–17 years for 2D-SWE, while only a rise between 4–7 and 12–14 years for pSWE, with a subsequent fall in LSM for those 15–17 years. It is easy to assume that such a discrepancy is operator dependent, but this is less likely given similar findings in both observers. Furthermore, pSWE LSM for adolescents aged 15-17 years, and adults 20-70 are identical. There seems to be some factor during early adolescence for which pSWE is particularly sensitive to, be it a biological factor or user-subject factors, e.g., a different ability to perform breath-hold. We have no indications that 12-year-olds behaved differently or took instructions differently than 17-year-olds, but this cannot be ruled out.

It is crucial to keep in mind that the actual LS when measuring was close to constant in our studies, as the inter-system investigation was performed in the same subjects with only minutes apart. The differences are thus system and/or user-dependent. Hypothesizing that LSM increases during childhood, it is impossible that both pSWE and 2D-SWE yield perfect measurements of the true liver stiffness: the former with a 0% increase in LSM from 4–7 years to adulthood in females, compared to nearly 50% in the latter. It is the *measurement values* that increase. It is possible that LSM reaches a peak value during puberty before slightly decreasing before entering adulthood and that some systems may be more sensitive to such changes during this period, but we can only speculate.

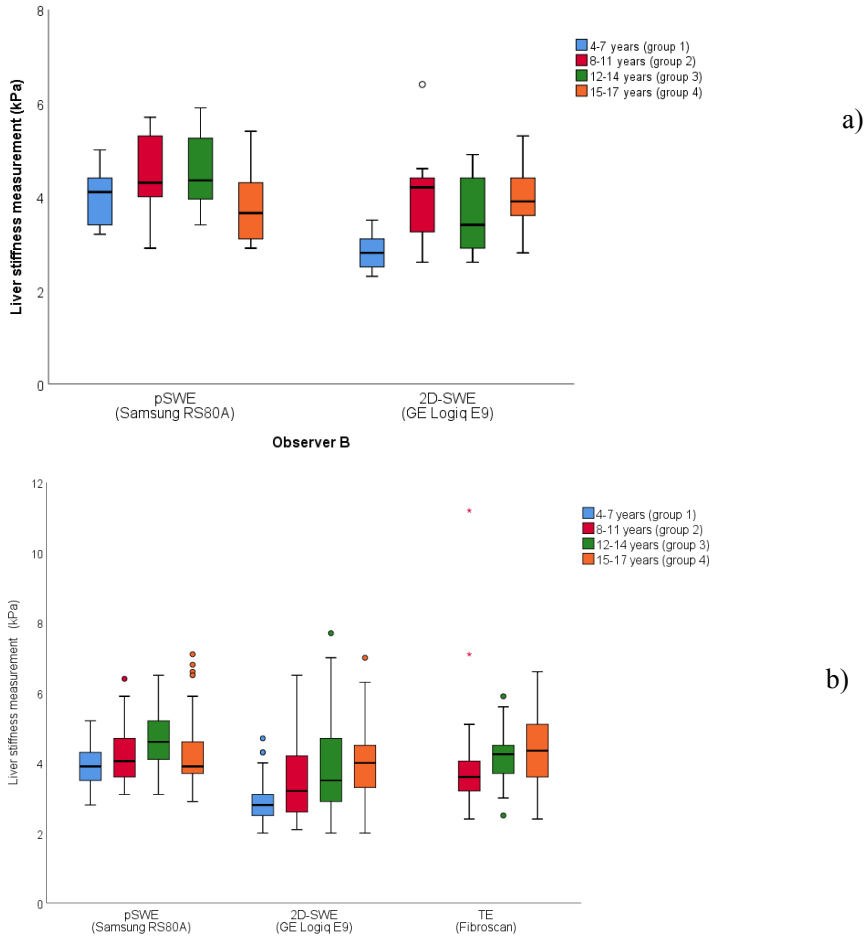


Fig. 20a-b. Liver stiffness measurements by age groups and elastography systems for a) observer A and b) observer B. *Fig. 20b reprinted from Paper II with permission [139].*

Figures 22–24 depict boxplots across all age groups for the different systems. For 2D-SWE, we see LSM from 4–7 years steadily rising into adulthood (Fig. 21), while for pSWE, we see a decline in LSM for age 15–17, stabilizing in adulthood (Fig. 22). For TE, we see a rise from 8–11 to 12–14 years, with no difference between the three older groups, including adults (Fig. 23). It should be mentioned that for 2D-SWE, we used one system for adults and another system for children and adolescents, while systems were identical for pSWE and TE across all ages.

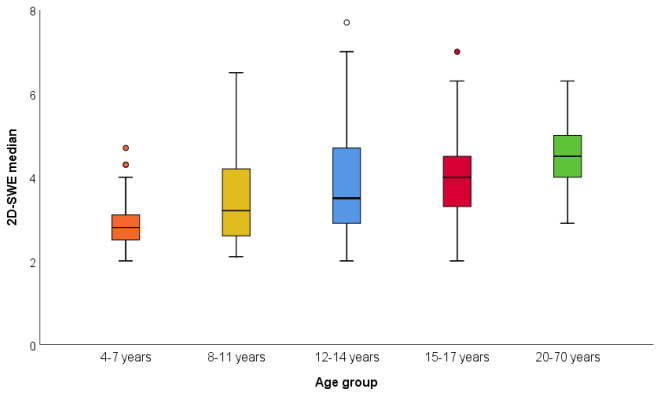


Fig. 21. Liver stiffness measurements by two-dimensional shear wave elastography (2D-SWE) for all age groups.

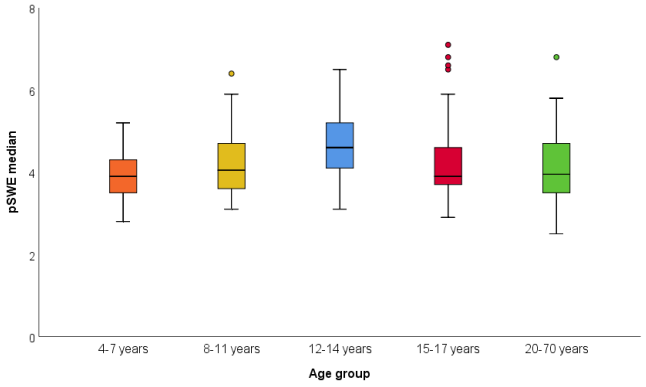


Fig. 22. Liver stiffness measurements by point shear wave elastography (pSWE) for all age groups.

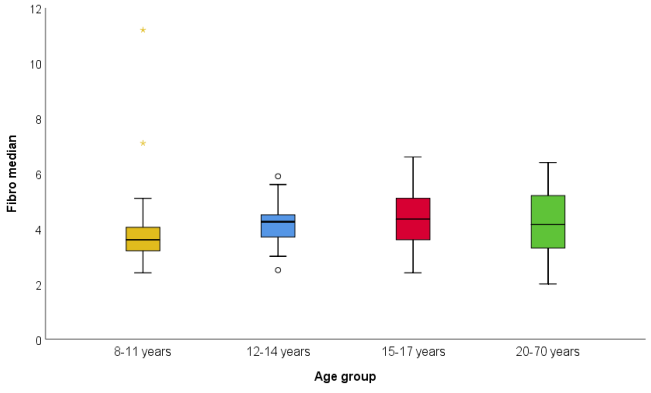


Fig. 23. Liver stiffness measurements by transient elastography (TE) for all age groups.

The influence of age on liver stiffness, although commonly observed, cannot presently be explained, and we have not been able to find any specific hypotheses on this subject. Similar findings are described in other organs, including the pancreas [185] and spleen [186]. A possible explanation can be similar to the cause of a gender difference, as discussed above. The body of the developing child and adolescent goes through massive changes well beyond a mere gain in weight and height, and influences of different hormones are present during the entire childhood, not only a part of adolescence and puberty.

5.3 LIVER ELASTOGRAPHY IN PSC

In Paper I, we describe how pSWE is feasible in PSC patients, with successful measurements in 100% using a right intercostal approach, compared to 93.9% in Paper IV (see Discussion section of Paper IV regarding feasibility and quality criteria). Left liver lobe LSM showed only moderate feasibility (66%) and (more importantly) poor diagnostic value. The 55 PSC patients were compared to 24 healthy controls, demonstrating a significant difference in LSM, with an optimal cutoff of 1.24 m/s (~4.6 kPa) as decided by Youden's index. In Paper IV, we describe a similar cutoff of 4.9 kPa, but this is calculated using TE values as gold standard. Our cutoff value of 1.24 m/s is identical with an earlier published cutoff for the same system, yielding a sensitivity of 75% and a specificity of 90% [187]. We report a lower sensitivity of 56%, but with a higher specificity of 96%: only 1 in 24 healthy controls had an LSM above cutoff. Using our cutoff value, we can be quite confident of disease if above this value (high positive predictive value, PPV), but the corresponding negative predictive value was low.

The LSM difference between patients and controls is readily appreciated using a boxplot (See Fig. 3, Paper I [188]).

The area under the curve (AUC) value was 0.775, which is considered good. However, the curve displays a problem (Fig. 24): We will catch 50% of PSC patients without false positive results, but the test shows poor qualities when

trying to find more patients, rapidly including healthy controls.
If we want to find 90% of patients, 75% of healthy controls will follow, which is clearly unwanted.

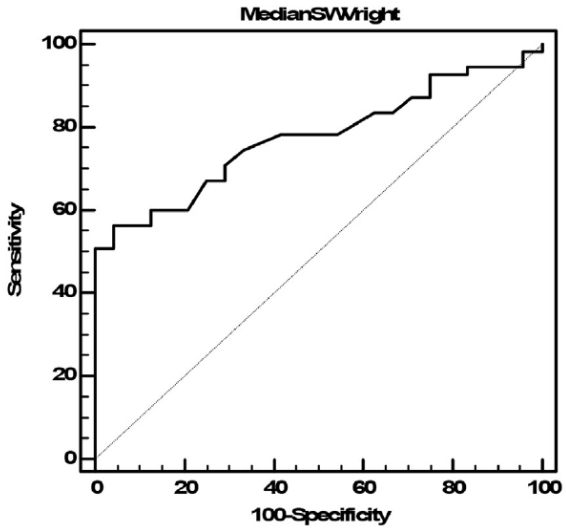


Fig. 24. Area under the curve (AUC), displaying sensitivity and specificity of LSM, where sensitivity is the percentage of PSC patients found by the test; and 100-specificity is the percentage of healthy controls with a "positive" test, falsely indicating that they have PSC. *Reprinted from Paper I with permission [188].*

Thus, if LSM was above 1.24 m/s, PSC was deemed very probable in our mixed cohort consisting of both patients and controls. Two clarifications are needed having said this:

1) LSM is not a diagnostic test, but as chronic liver disease leads to a higher liver stiffness, increased LSM will heighten the probability of liver disease;

2) when applying a cutoff value to separate patients and the healthy, the test's predictive values will depend on the disease prevalence. With a sensitivity of 56.4% and a specificity of 95.8%, our cutoff value will cause us to label 56.4% of patients and 4.2% of healthy controls as 'positive.' With 55 patients and 24 controls, we will have 31 patients and one healthy, yielding a PPV of 31/32 = 0.97. Thus, in our cohort, chances of disease will be 97%, given a value >1.24

m/s. Screening the general public, where the prevalence is very low, the same cutoff value would cause a much lower PPV: our 55 PSC patients would be matched by 550'000 non-PSC individuals. Using the same test criteria, we would still label 31 PSC patients as 'positive,' but now compared to 23'100 non-PSC individuals with a positive test, lowering PPV from 0.97 to 0.0013. If implementing a screening program, the cutoff value would have to be set much higher, lowering sensitivity and increasing specificity. Prevalence and pretest probability is important to consider before implementing such testing [189]. Our example ignores the fact that there any many other liver diseases and other confounding factors which may increase LSM, and that there may be PSC patients without a diagnosis in the general public.

In those with a lower LSM, other tests are warranted to separate PSC patients from the healthy. Analyzing PSC patients from Paper IV, 41% had an LSM <1.24 m/s, and 49% an LSM <1.28 m/s. No PSC patients with an LSM below these values had an elevated ELF test ≥11.2 or an increased Mayo risk score (≥0.5).

Analyzing PSC patients with an elevated pSWE LSM, we find that 21/31 (67.7%), displayed normal liver parenchyma at baseline (Paper I), with nearly identical figures during follow-up (Paper IV). All participants displaying coarse parenchyma had an LSM indicating fibrosis, both for pSWE and TE. B-mode findings will appear well into the disease course and are thus not a sensitive marker for subtle progression of fibrosis. The literature reports sensitivity and interobserver agreement as low as 41 and 35%, respectively [190-192]. It has been shown that this can be improved using artificial intelligence [193], but this is not currently available in clinical practice.

Splenomegaly is not part of the mentioned B-mode signs of liver fibrosis, but may indicate portal hypertension secondary to the liver disease and is easily quantified by B-mode ultrasound. A spleen length ≥12 cm has been reported to be negatively associated with prognosis in PSC patients [5, 194]. We could not replicate this finding, but setting the cutoff at 13 cm, we found a difference with

significantly higher LSM in those with splenomegaly. Spleen size depends on body size and weight, and we propose that cutoff values should consider body size and that a *change* in spleen size may be an even better prognostic marker.

There was a highly significant relationship between LSM and serum markers at baseline and in follow-up, with APRI scores showing the highest correlation to LSM in both investigations, in Paper IV equal with Mayo risk score correlation (Table 2). All correlations increased between the two time points. A possible contributing factor to this rise is increased fibrosis over time in the cohort, as tests are more precise in cases of more advanced fibrosis.

Another factor may be missing data: for APRI and Mayo risk score there was missing data in 25% at baseline, compared to only 1.5% in follow-up. It is also possible that an increasing number of participants contributed, and that the operator technique improved over the years, yielding more precise measurements.

	Baseline (Paper I)		Follow-up (Paper IV)	
	pSWE LSM	**P-value**	**pSWE LSM**	**P-value**
ELF	–	–	0.57	<0.001
APRI	0.49	0.001	0.67	<0.001
Mayo RS	0.30	0.06	0.67	<0.001
FIB-4	0.37	0.017	0.59	<0.001
AST	0.21	0.187	0.60	<0.001
Albumin	–0.36	0.011	–0.46	<0.001
Thrombocytes	–0.32	0.022	–0.40	0.001
ALP	0.25	0.085	0.58	<0.001

Table 2. Correlation between liver stiffness measurements and laboratory values and fibrosis/prognostic scores.

The ELF test was only investigated in follow-up; thus, we cannot provide any data on change over time. A strong and independent correlation with prognosis in PSC patients has previously been reported for ELF [6, 7].

While LSM is indeed a robust surrogate marker of liver fibrosis, it may not outperform all serum-based tests under all conditions. Some studies have

demonstrated that the combination of LSM and fibrosis markers performs better than LSM alone [8-10]. In PSC, a dominant stenosis causing cholestasis is a frequent finding and may influence LSM, without being indicative of true liver fibrosis [11]. Thus, comparing each elastography platform to the ELF test as the leading serum-based liver fibrosis test in PSC is of high interest.

PSC is a complex disease with many factors other than fibrosis affecting disease progression and risk of death or transplantation. Even a simple blood test such as alkaline phosphatase (ALP) is strongly linked with outcome [195]. Several factors other than traits of the disease, may negatively influence the ability to perform successful LSM, necessitating the use of serum markers [12]. A recent study demonstrated the excellent properties of liver elastography while arguing that serum markers still have a role due to both availability and cases of failure to obtain a valid LSM [196]. Furthermore, serum-based tests are not operator dependent [7]. Several publications advocate using simple serum biomarkers in primary care, with non-invasive tests including elastography as part of the subsequent specialist evaluation [13-17].

When trying to separate PSC patients with or without advanced fibrosis, it is paramount to evaluate the puzzle pieces at our disposal. The perfect test does not exist, and LSM remains only a part of the puzzle. Together with clinical information, biochemistry, serum fibrosis markers and prognostic markers, the LSM value, and perhaps especially a tendency of changing LSM over time, has its place in the diagnostic armamentarium in chronic liver patients.

The interquartile range divided by the median, given in percent (IQR/M, see 3.3 DEFINITIONS), is a measurement of dispersion: when measuring, are values close to each other, or is there a substantial variability? A high IQR/M means that measurements vary, representing uncertainty. When TE (Fibroscan) was introduced in 2003, the producer recommended two reliability criteria: the IQR/M, which should be 30% or less, and success rate (the ratio between valid and the total number of acquisitions), which should be 60% or above [31]. The

latter has since been removed [197]. The same authors introduced the concept of dividing IQR/M into "very reliable," "reliable" and "poorly reliable," for IQR/M values of ≤10%, 10–30%, and >30%, respectively. This finding has not been externally validated but is still adopted by the producer. The concept of IQR/M ≤30% has been adopted in pSWE and 2D-SWE systems, even though it is specified that evidence is lacking [31]. Guidelines recommend using this reliability criterion but do not specify whether this is independent of the measuring unit. In Paper II, we describe how IQR/M is different for identical measurements given in either kPa or m/s, with IQR/M twice as high in kPa as in m/s (Fig. 25). Thus, when using m/s, an IQR/M of 30% equals 60% with measurements in kPa; still, publications have been made where IQR/M ≤30% is applied even in the setting of measurements in m/s [198].

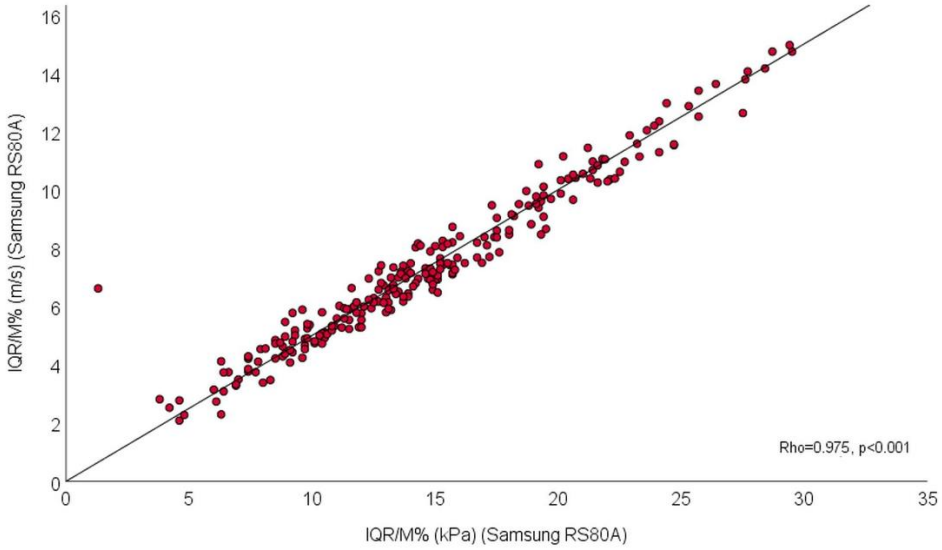


Fig. 25. Scatter plot of interquartile range / median (IQR/M) in % for kPa (x-axis) and m/s (y-axis). We see a near-perfect linear correlation, with IQR/M in kPa twice the IQR/M in m/s. *Reprinted from Paper II with permission [139].*

In Paper IV, we used three systems for measuring LSM in adult PSC patients: 2D-SWE and TE, both recommending IQR/M ≤30% as a reliability criterion, and pSWE from a producer not recommending this criterion. In our paper, we noted

that the feasibility was 93.9% for pSWE, compared to 89.4 and 71.2% for TE and 2D-SWE, respectively. We speculated that the difference in reliability criteria was part of this difference, even if inter-system differences have been demonstrated, with higher feasibility for TE compared to 2D-SWE in difficult-to-scan patients [199].

Post hoc calculation of the IQR/M for pSWE reveals that applying this criterion for all systems, would have dramatically changed the outcome: only 45/66 (68.2%) patients had an IQR/M ≤30%. Since measurements with pSWE were performed using m/s and not kPa, the real feasibility would be even lower: only 11/66 (16.7%) had an IQR/M ≤15% (Table 2). Thus, most of the accepted and valid measurements by pSWE would be deemed invalid if all systems applied the same criteria. This finding should raise suspicion regarding *how* we use IQR/M as a reliability criterion, as we found that pSWE seemingly performed as good or better than the other systems, even when accepting measurements with a very high IQR/M.

	Requirements for valid liver stiffness measurement			
	Producer recommendations	**SR >60 only**	**IQR/M ≤30 + SR>60**	**IQR/M_{kPa} ≤30 + SR>60**
pSWE	93.9% (4 SR)	93.9% (4 SR)	65.2% (4 SR, 19 IQR/M)	13.6 % (4 SR, 53 IQR/M)
2D-SWE	71.2% (4 SR, 15 IQR/M)	93.9% (4 SR)	71.2% (4 SR, 15 IQR/M)	71.2% (4 SR, 15 IQR/M)
TE	87.9% (6 SR, 2 IQR/M)	90.9% (6 SR)	87.9% (6 SR, 2 IQR/M)	87.9% (6 SR, 2 IQR/M)

* Ignoring the success rate requirement, 2D-SWE would still have four patients with invalid results (not able to obtain LSM), pSWE would still have four patients (2 patients: unable to obtain invalid results, the final two had an IQRM >30 as well), for TE 5 had no valid acquisitions, but the last one would have yielded a valid result.

Table 2. Feasibility based on different recommendations for point shear wave elastography (pSWE), two-dimensional shear wave elastography (2D-SWE), and transient elastography (TE). The different recommendations are combinations of either interquartile range divided by the median in percent (IQR/M) or success rate (SR).

For TE, IQR/M is a key factor for discrepancy in fibrosis staging by TE and histology [200, 201]. Although this link is not found in all investigations [202], it is recognized by EFSUMB guidelines [31]. These studies, comparing histological fibrosis and TE, found that an IQR/M <21% was the best cutoff for a reliable LSM, but with acceptable reliability for IQR/M 21–50% in patients where LSM was below 9.5 kPa. This is consistent with what we see in clinical practice, that a low LSM probably reflects a truly soft liver, even in the presence of a high IQR/M, while a high LSM «needs» a lower IQR/M to be trusted. Thus, both the real LS and the different systems play a role, affecting the IQR/M. In Paper IV, where we discussed the vast difference in feasibility, we also describe that the different systems revealed differences in LS measurements. Applying the mentioned cutoff of 9.5 kPa, we find that 57 patients were below this limit when using pSWE, compared to 48 and 41 by 2D-SWE and TE, respectively. Our findings are consistent with the mentioned study, with acceptable measurements, even with IQR/M up to 50%, or even above.
Median (IQR) for the entire cohort was 4.8 (2.8), 6.8 (4.5), and 6.9 (5.6) for pSWE, 2D-SWE, and TE, respectively.

The inter-system differences are challenging to characterize in a few words. Differences were not linear, but subject to a more complex relationship, as described in healthy individuals and liver elasticity phantoms. In Paper IV, we were able to show that various inter-system differences were affected by different factors, such as increasing differences between pSWE and TE levels with increasing BMI. pSWE LSM was the single factor affecting the difference between pSWE and 2D-SWE, which only reflects the non-linear relationship: when pSWE levels increased, the 2D-SWE level did not follow accordingly, which is expected based on earlier comparisons between these systems (Fig. 26) [140].

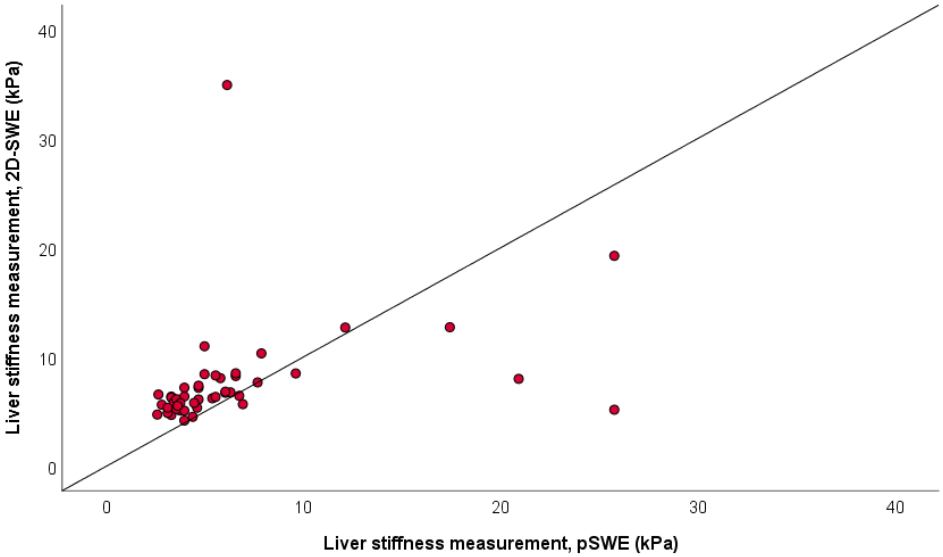


Fig. 26. Scatter plot showing liver stiffness measurement (LSM) by pSWE vs. 2D-SWE: with very high pSWE values, the 2D-SWE did not follow accordingly.

The correlation between LSM values across systems was assessed using ICC. The ICC was excellent between pSWE and TE, but only moderate, nearly poor, for 2D-SWE versus the two other systems. This was primarily caused by a high BMI, and excluding obese and moderately overweight individuals (cutoff 27-28 kg/m^2), caused the ICC to increase from 0.43-0.49 to 0.81-0.92. We also suggest that the real inter-system ICC involving 2D-SWE may have been falsely *elevated*, as 2D-SWE often yielded invalid measurements in cases with a massive LSM difference between pSWE and TE. If one system provides no value in difficult-to-scan patients with a high degree of variability, the results will necessarily seem better than what is reality. It is thus not only essential to provide readers with feasibility in numbers, but also try to explain factors causing invalid measurements. This can be compared to a study on LSM variability where one system is applied in a cohort of morbidly obese patients with cirrhosis, while another system investigates lean, healthy young adults. Highly significant differences would be expected without providing any useful information on real inter-system differences.

5.4 STRENGTHS AND LIMITATIONS

The main strength to be mentioned is a substantial amount of data acquisition, gathered using stringent methodology: 343 healthy persons were subject to ultrasound investigation and liver elastography using two different elastography systems – with 187 investigated by three different systems and 74 investigated by two observers. Every measurement consisted of a minimum of 10 acquisitions. Our pediatric material represents the first description of multiple elastography systems in a head-to-head comparison in healthy children. Care was taken to conduct everything according to protocol and to ensure that all participants were healthy.

For the PSC cohort, we investigated a substantial amount of patients (55 in study I and 66 in study IV), considering that PSC is a rare disease and with investigations performed in a sparsely populated country. In both papers, investigations were performed in a clinical setting with blood tests, clinical examination, and B-mode ultrasonography being performed, allowing a broad approach to the included patients. We explored three different systems (two of which for the first time) and evaluated interobserver differences and feasibility of LSM in both the right and left liver lobe.

The main limitation is the lack of a criterion standard. In liver elastography, where an estimation of liver fibrosis is the goal, the criterion standard, although flawed, is the liver biopsy. It would be especially welcome in the studies on PSC, where liver fibrosis was probable but the degree unknown. However, liver biopsy was considered an unethical procedure as it was not clinically indicated. Liver histology in the healthy subjects would be highly unethical. Furthermore, it would have necessitated general anesthesia in the children, and the risk of complications would be present in all age groups.

The use of TE as a gold standard is complicated. TE has been shown to correlate with liver histopathology but represents a surrogate assessment with imperfect

correlation with a flawed criterion standard. Still, it is the elastography method most validated thus far. Investigations comparing both pSWE and TE with liver histopathology, do not find any convincing evidence of these systems performing differently [155, 203, 204], and 2D-SWE has demonstrated diagnostic performance similar to TE, with a high degree of correlation with fibrosis staging by liver biopsy [204-207].

Thus, we are left with an imperfect gold standard at best performing at the same level as the very systems we are trying to evaluate.

Another way to increase the probability of a healthy liver is blood tests, which we performed in adults, but not in the children. Liver disease in children is rare in Norway, and children are less able than adults to give proper consent. It was thus judged unethical to impose blood tests on healthy children. Furthermore, negative blood tests would not rule out liver disease with absolute certainty.

Ideally, we should have included children from infancy, to create reference values for all ages, and TE should have been performed in smaller children. The latter would require a different TE probe, which we did not possess. It would also be better to have the B-mode ultrasound investigation performed by an experienced pediatric radiologist, but was practically not feasible.

Paper I has no gold standard, with no liver biopsy and no TE. Ideally, there should have been a TE examination to strengthen findings and, preferably, a liver biopsy. The study's scope was not to calibrate LSM against a gold standard but to explore pSWE in PSC and compare LSM against other signs of fibrosis. Furthermore, the ELF test was not available at the time.

Another limitation is *where* LSM was gathered. In all studies, all 10 valid acquisitions were made in the same part of the liver. It is unknown whether having a protocol attempting to establish LSM in the entire liver, would have yielded different results. This would have been especially interesting in the PSC cohort, with its inherent scattered fibrosis.

6. CONCLUSION

Liver elastography is feasible across multiple ultrasound systems, both in the healthy and in patients with chronic cholestatic liver disease, and there is a high degree of interobserver correlation. We have established LSM reference values for both children and adults, and find that LSM seems to increase with age until adolescence, stabilizing on adult levels. In children, 2D-SWE yields lower values than pSWE and TE, while 2D-SWE yields the highest values in adults. There is a gender difference, with males exhibiting higher LSM than females in adolescence and adulthood, but not in early childhood.

LSM is increased in PSC, and it is correlated with other signs indicating liver pathology: serum biomarkers; fibrosis and prognostic scores; and B-mode ultrasonographic findings, including spleen length ≥13 cm. The latter finding indicates that splenomegaly should be defined according to gender and body size.

We find that the correlation between LSM by 2D-SWE and the other systems is affected by BMI, and care should be made when investigating obese patients.
It is interesting to note that while 2D-SWE showed near-perfect feasibility in the healthy, its feasibility dropped dramatically in a cholestatic liver disease cohort. This could be due, at least partly, to the fact that 2D-SWE demands more training than pSWE and TE.

The IQR/M varies greatly depending on the measurement unit and should be used actively when interpreting LSM values.

Through demonstration of feasibility and correlation with liver fibrosis, this thesis strengthens the belief that elastography in combination with conventional ultrasound allows us as clinicians to perform a thorough and in-depth evaluation of the patient's liver in a single session, without spending much time or causing the patient any discomfort.

7. FUTURE PERSPECTIVES

Elastography is already established as a valuable tool for non-invasive liver stiffness measurement, and reference values for both adults and children have been published. Similar studies with a head-to-head comparison of different systems are missing in younger children (0-3 years). Gender difference is not fully explored, and future studies should include a thorough examination of this, including more healthy subjects of both genders, assessing puberty and menarche as part of the study. It could also be of interest to perform LSM in adult women in different parts of the menstrual cycle, to establish whether hormones affect measurements. Similarly, it would be of value to explore how LSM fluctuates in both healthy subjects and chronic liver patients, as it is unknown how much LSM can increase without representing a true worsening: such a study calls for interobserver analyses and the use of different elastography systems. It is paramount that not only correlation is reported, but also discrepancies in absolute values. Research questions should reflect clinical needs whenever possible.

The use of quality criteria such as IQR/M and SR is still under debate, necessitating more investigations. More studies with multiple systems and simultaneous liver biopsy, emphasizing quality criteria, should be performed. That pSWE seemingly outperformed 2D-SWE even in cases where the measurement should have been deemed invalid if the criteria of IQR/M≤30 had been applied, warrants future studies. Maybe a lack of such criteria could be met by an increased number of acquisitions? A protocol could include 20 acquisitions, noting the IQR/M level at the first 10, 15, and 20 acquisitions, respectively. Rather than accepting old criteria established by TE, care should be taken to find if IQR/M should be used and if so: at what level.

Although having cemented the obvious strengths of liver evaluation through elastography, there is still much to do.

8. REFERENCES

1. Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, Mal F, et al. Transient elastography: A new noninvasive method for assessment of hepatic fibrosis. Ultrasound in Medicine and Biology. 2003;29(12):1705-13. doi: DOI 10.1016/j.ultrasmedbio.2003.07.001. PubMed PMID: WOS:000187755200007.
2. Friedrich-Rust M, Ong MF, Martens S, Sarrazin C, Bojunga J, Zeuzem S, et al. Performance of transient elastography for the staging of liver fibrosis: A meta-analysis. Gastroenterology. 2008;134(4):960-74. doi: 10.1053/j.gastro.2008.01.034. PubMed PMID: WOS:000254853800016.
3. Herrmann E, de Ledinghen V, Cassinotto C, Chu WC, Leung VY, Ferraioli G, et al. Assessment of biopsy-proven liver fibrosis by two-dimensional shear wave elastography: An individual patient data-based meta-analysis. Hepatology. 2018;67(1):260-72. Epub 2017/04/04. doi: 10.1002/hep.29179. PubMed PMID: 28370257; PubMed Central PMCID: PMCPMC5765493.
4. Corpechot C, Gaouar F, El Naggar A, Kemgang A, Wendum D, Poupon R, et al. 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(4):970-U511. doi: 10.1053/j.gastro.2013.12.030. PubMed PMID: WOS:000333254500026.
5. Ehlken H, Wroblewski R, Corpechot C, Arrive L, Rieger T, Hartl J, et al. 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(10). doi: ARTN e0164224

10.1371/journal.pone.0164224. PubMed PMID: WOS:000385504100031.
6. de Vries EM, Farkkila M, Milkiewicz P, Hov JR, Eksteen B, Thorburn D, et al. Enhanced liver fibrosis test predicts transplant-free survival in primary sclerosing cholangitis, a multi-centre study. Liver International. 2017;37(10):1554-61. doi: 10.1111/liv.13402. PubMed PMID: WOS:000411512100016.
7. Vesterhus M, Hov JR, Holm A, Schrumpf E, Nygard S, Godang K, et al. Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. Hepatology. 2015;62(1):188-97. doi: 10.1002/hep.27825. PubMed PMID: WOS:000356864800025.
8. Castera L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. Gastroenterology. 2005;128(2):343-50. Epub 2005/02/03. doi: 10.1053/j.gastro.2004.11.018. PubMed PMID: 15685546.
9. Boursier J, Guillaume M, Leroy V, Irles M, Roux M, Lannes A, et al. New sequential combinations of non-invasive fibrosis tests provide an accurate diagnosis of advanced fibrosis in NAFLD. J Hepatol. 2019;71(2):389-96. Epub 2019/05/19. doi: 10.1016/j.jhep.2019.04.020. PubMed PMID: 31102719.
10. Lannerstedt H, Konopski Z, Sandvik L, Haaland T, Loberg EM, Haukeland JW. Combining transient elastography with FIB4 enhances sensitivity in detecting

advanced fibrosis of the liver. Scand J Gastroentero. 2013;48(1):93-100. doi: 10.3109/00365521.2012.746389. PubMed PMID: WOS:000312505700013.
11. Millonig G, Reimann FM, Friedrich S, Fonouni H, Mehrabi A, Buchler MW, et al. Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. Hepatology. 2008;48(5):1718-23. Epub 2008/10/07. doi: 10.1002/hep.22577. PubMed PMID: 18836992.
12. Karlas T, Dietrich A, Peter V, Wittekind C, Lichtinghagen R, Garnov N, et al. Evaluation of Transient Elastography, Acoustic Radiation Force Impulse Imaging (ARFI), and Enhanced Liver Function (ELF) Score for Detection of Fibrosis in Morbidly Obese Patients. PLoS One. 2015;10(11):e0141649. Epub 2015/11/04. doi: 10.1371/journal.pone.0141649. PubMed PMID: 26528818; PubMed Central PMCID: PMCPMC4631322.
13. Anstee QM, Lawitz EJ, Alkhouri N, Wong VWS, Romero-Gomez M, Okanoue T, et al. Noninvasive Tests Accurately Identify Advanced Fibrosis due to NASH: Baseline Data From the STELLAR Trials. Hepatology. 2019;70(5):1521-30. doi: 10.1002/hep.30842. PubMed PMID: WOS:000482168500001.
14. Castera L. Diagnosis of non-alcoholic fatty liver disease/non-alcoholic steatohepatitis: Non-invasive tests are enough. Liver International. 2018;38:67-70. doi: 10.1111/liv.13658. PubMed PMID: WOS:000424860900012.
15. Davyduke T, Abraldes JG, Tandon P, Ma MM. Impact of Pre-Screening with Fibrosis-4 Index on a Referral Pathway for Patients with Suspected NAFLD. Hepatology. 2018;68:44a-a. PubMed PMID: WOS:000446020500070.
16. Tsochatzis EA, Newsome PN. Non-alcoholic fatty liver disease and the interface between primary and secondary care. Lancet Gastroenterol. 2018;3(7):509-17. doi: Doi 10.1016/S2468-1253(18)30077-3. PubMed PMID: WOS:000436554300027.
17. Davyduke T, Tandon P, Al-Karaghouli M, Abraldes JG, Ma MM. Impact of Implementing a "FIB-4 First" Strategy on a Pathway for Patients With NAFLD Referred From Primary Care. Hepatol Commun. 2019;3(10):1322-33. doi: 10.1002/hep4.1411. PubMed PMID: WOS:000488812200005.
18. Gibbs V, Cole, D, Sassano, A. Ultrasound physics and technology: How, why and when: Churchill Livingstone Elsevier; 2009. 137 p.
19. Siegel MJ. Pediatric Sonography, 5th edition: Wolters Kluwer; 2019. 716 p.
20. Johanssen JN, M; Schreiner, E. Latinsk - norsk ordbok [Latin-norwegian dictionary]. 4th ed: J.W. Cappelen; 1998.
21. Hoskins P. Diagnostic ultrasound: Physics and equipment. R HP, Kevin M, Abigail T, editors: CRC Press; 2019. 387 p.
22. Ødegaard S GO, Gregersen H. Basic and new aspects of gastrointestinal ultrasonography: World scientific; 2005. 502 p.
23. Kane D, Grassi W, Sturrock R, Balint PV. A brief history of musculoskeletal ultrasound: 'From bats and ships to babies and hips'. Rheumatology. 2004;43(7):931-3. doi: 10.1093/rheumatology/keh004. PubMed PMID: WOS:000222270100021.
24. Curie P, Curie Jacques. Développement par compression de l'électricité polaire dans les cristaux hémièdres à faces inclinées. Bulletin de Minéralogie. 1880;3-4:90-3.

25. Hill CR. Medical Ultrasonics - Historical Review. Brit J Radiol. 1973;46(550):899-905. doi: Doi 10.1259/0007-1285-46-550-899. PubMed PMID: WOS:A1973R080300025.
26. Donald I, Macvicar J, Brown TG. Investigation of Abdominal Masses by Pulsed Ultrasound. Lancet. 1958;1(Jun7):1188-94. PubMed PMID: WOS:A1958WB50000002.
27. Baker DW. Pulsed Ultrasonic Doppler Flowmeter Biological and Engineering Applications. Ieee T Son Ultrason. 1970;Su17(1):65-&. PubMed PMID: WOS:A1970F503400085.
28. Maulik D. Doppler sonography in obstetrics and gynecology: Springer; 2005.
29. National health service England. Diagnostic imaging dataset statistical release 2019 [updated 23.03.20]. Available from: https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2019/07/Provisional-Monthly-Diagnostic-Imaging-Dataset-Statistics-2019-07-18.pdf.
30. Wilkins RH. Neurosurgical Classic .17. J Neurosurg. 1964;21(3):240-&. doi: DOI 10.3171/jns.1964.21.3.0240. PubMed PMID: WOS:A19646090B00013.
31. Dietrich CF, Bamber J, Berzigotti A. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version) (vol 38, pg e16, 2017). Ultraschall Med. 2017;38(4):E52-E. doi: 10.1055/a-0641-0076. PubMed PMID: WOS:000444461600001.
32. Wilson LS, Robinson DE. Ultrasonic Measurement of Small Displacements and Deformations of Tissue. Ultrasonic Imaging. 1982;4(1):71-82. doi: Doi 10.1016/0161-7346(82)90006-2. PubMed PMID: WOS:A1982NC10700005.
33. Ophir J, Cespedes I, Ponnekanti H, Yazdi Y, Li X. Elastography - a Quantitative Method for Imaging the Elasticity of Biological Tissues. Ultrasonic Imaging. 1991;13(2):111-34. doi: Doi 10.1016/0161-7346(91)90079-W. PubMed PMID: WOS:A1991FN13800001.
34. Lerner RM, Huang SR, Parker KJ. Sonoelasticity Images Derived from Ultrasound Signals in Mechanically Vibrated Tissues. Ultrasound in Medicine and Biology. 1990;16(3):231-9. doi: Doi 10.1016/0301-5629(90)90002-T. PubMed PMID: WOS:A1990DF12400002.
35. Parker KJ, Huang SR, Musulin RA, Lerner RM. Tissue-Response to Mechanical Vibrations for Sonoelasticity Imaging. Ultrasound in Medicine and Biology. 1990;16(3):241-6. doi: Doi 10.1016/0301-5629(90)90003-U. PubMed PMID: WOS:A1990DF12400003.
36. Parker KJ, Fu DS, Graceswki SM, Yeung F, Levinson SF. Vibration sonoelastography and the detectability of lesions. Ultrasound in Medicine and Biology. 1998;24(9):1437-47. doi: Doi 10.1016/S0301-5629(98)00123-9. PubMed PMID: WOS:000077573000019.
37. Gilja OH, Heimdal A, Hausken T, Gregersen H, Matre K, Berstad A, et al. Strain during gastric contractions can be measured using Doppler ultrasonography. Ultrasound in Medicine and Biology. 2002;28(11-12):1457-65. doi: Pii S0301-5629(02)00614-2

Doi 10.1016/S0301-5629(02)00614-2. PubMed PMID: WOS:000180041500011.

38. Ahmed AB, Gilja OH, Gregersen H, Odegaard S, Matre K. In vitro strain measurement in the porcine antrum using ultrasound Doppler strain rate imaging. Ultrasound in Medicine and Biology. 2006;32(4):513-22. doi: 10.1016/j.ultrasmedbio.2005.12.019. PubMed PMID: WOS:000237007600007.
39. Ahmed AB, Gilja OH, Hausken T, Gregersen H, Matre K. Strain measurement during antral contractions by ultrasound strain rate imaging: influence of erythromycin. Neurogastroent Motil. 2009;21(2):170-9. doi: 10.1111/j.1365-2982.2007.01043.x. PubMed PMID: WOS:000262688300010.
40. Havre RF, Leh S, Gilja OH, Odegaard S, Waage JE, Baatrup G, et al. Strain Assessment in Surgically Resected Inflammatory and Neoplastic Bowel Lesions. Ultraschall Med. 2014;35(2):149-58. doi: 10.1055/s-0032-1325535. PubMed PMID: WOS:000334168000007.
41. Havre R, Gilja OH. Elastography and strain rate imaging of the gastrointestinal tract. Eur J Radiol. 2014;83(3):438-41. Epub 2013/06/19. doi: 10.1016/j.ejrad.2013.05.018. PubMed PMID: 23769191.
42. Havre RF, Odegaard S, Gilja OH, Nesje LB. Characterization of solid focal pancreatic lesions using endoscopic ultrasonography with real-time elastography. Scand J Gastroenterol. 2014;49(6):742-51. Epub 2014/04/10. doi: 10.3109/00365521.2014.905627. PubMed PMID: 24713038.
43. Muthupillai R, Lomas DJ, Rossman PJ, Greenleaf JF, Manduca A, Ehman RL. Magnetic-Resonance Elastography by Direct Visualization of Propagating Acoustic Strain Waves. Science. 1995;269(5232):1854-7. doi: DOI 10.1126/science.7569924. PubMed PMID: WOS:A1995RX19400032.
44. Dutt V, Kinnick, RR., Greenleaf, JF. . Acoustic shear wave displacement measurement using ultrasound. IEEE Ultrasonics Symposium Proceedings. 1996;2:1185-8.
45. Catheline S, Wu F, Fink M. A solution to diffraction biases in sonoelasticity: The acoustic impulse technique. J Acoust Soc Am. 1999;105(5):2941-50. doi: Doi 10.1121/1.426907. PubMed PMID: WOS:000080310200040.
46. Catheline S, Thomas JL, Wu F, Fink MA. Diffraction field of a low frequency vibrator in soft tissues using transient elastography. Ieee T Ultrason Ferr. 1999;46(4):1013-9. doi: Doi 10.1109/58.775668. PubMed PMID: WOS:000081469000029.
47. Sandrin L, Catheline S, Tanter M, Hennequin X, Fink M. Time-resolved pulsed elastography with ultrafast ultrasonic imaging. Ultrasonic Imaging. 1999;21(4):259-72. doi: Doi 10.1177/016173469902100402. PubMed PMID: WOS:000086639700002.
48. Nightingale KR, Palmeri ML, Nightingale RW, Trahey GE. On the feasibility of remote palpation using acoustic radiation force. J Acoust Soc Am. 2001;110(1):625-34. doi: Doi 10.1121/1.1378344. PubMed PMID: WOS:000169846100057.
49. Nightingale K, Bentley R, Trahey G. Observations of tissue response to acoustic radiation force: Opportunities for imaging. Ultrasonic Imaging. 2002;24(3):129-38. doi: Doi 10.1177/016173460202400301. PubMed PMID: WOS:000179920500001.

50. Sugimoto T, Ueha, S, Itoh K. . Tissue hardness measurement using the radiation force of focused ultrasound. . IEEE Symposium on Ultrasonics. 3:1377-80.
51. Bamber J, Cosgrove D, Dietrich CF, Fromageau J, Bojunga J, Calliada F, et al. EFSUMB Guidelines and Recommendations on the Clinical Use of Ultrasound Elastography. Part 1: Basic Principles and Technology. Ultraschall Med. 2013;34(2):169-84. doi: 10.1055/s-0033-1335205. PubMed PMID: WOS:000322373400011.
52. Shiina T, Nightingale KR, Palmeri ML, Hall TJ, Bamber JC, Barr RG, et al. Wfumb Guidelines and Recommendations for Clinical Use of Ultrasound Elastography: Part 1: Basic Principles and Terminology. Ultrasound in Medicine and Biology. 2015;41(5):1126-47. doi: 10.1016/j.ultrasmedbio.2015.03.009. PubMed PMID: WOS:000352803100002.
53. Colloredo G, Guido M, Sonzogni A, Leandro G. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. J Hepatol. 2003;39(2):239-44. Epub 2003/07/23. doi: 10.1016/s0168-8278(03)00191-0. PubMed PMID: 12873821.
54. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology. 2003;38(6):1449-57. doi: 10.1016/j.hep.2003.09.022. PubMed PMID: WOS:000187032100016.
55. Pagliaro L, Rinaldi F, Craxi A, Di Piazza S, Filippazzo G, Gatto G, et al. Percutaneous blind biopsy versus laparoscopy with guided biopsy in diagnosis of cirrhosis. A prospective, randomized trial. Dig Dis Sci. 1983;28(1):39-43. Epub 1983/01/01. doi: 10.1007/bf01393359. PubMed PMID: 6217961.
56. Maharaj B, Maharaj RJ, Leary WP, Cooppan RM, Naran AD, Pirie D, et al. Sampling variability and its influence on the diagnostic yield of percutaneous needle biopsy of the liver. Lancet. 1986;1(8480):523-5. Epub 1986/03/08. doi: 10.1016/s0140-6736(86)90883-4. PubMed PMID: 2869260.
57. Poniachik J, Bernstein DE, Reddy KR, Jeffers LJ, Coelho-Little ME, Civantos F, et al. The role of laparoscopy in the diagnosis of cirrhosis. Gastrointest Endosc. 1996;43(6):568-71. Epub 1996/06/01. doi: 10.1016/s0016-5107(96)70192-x. PubMed PMID: 8781934.
58. Backeljauw B, Holland SK, Altaye M, Loepke AW. Cognition and Brain Structure Following Early Childhood Surgery With Anesthesia. Pediatrics. 2015;136(1):E1-E12. doi: 10.1542/peds.2014-3526. PubMed PMID: WOS:000357296000001.
59. Rappaport BA, Suresh S, Hertz S, Evers AS, Orser BA. Anesthetic Neurotoxicity - Clinical Implications of Animal Models. New Engl J Med. 2015;372(9):796-7. doi: 10.1056/NEJMp1414786. PubMed PMID: WOS:000349901700002.
60. Guihard-Costa AM, Menez F, Delezoide AL. Organ weights in human fetuses after formalin fixation: standards by gestational age and body weight. Pediatr Dev Pathol. 2002;5(6):559-78. Epub 2002/10/26. doi: 10.1007/s10024-002-0036-7. PubMed PMID: 12399830.
61. Szpinda M, Paruszewska-Achtel M, Wozniak A, Mila-Kierzenkowska C, Elminowska-Wenda G, Dombek M, et al. Volumetric Growth of the Liver in the Human Fetus: An Anatomical, Hydrostatic, and Statistical Study. Biomed Res Int.

2015;2015:858162. Epub 2015/09/29. doi: 10.1155/2015/858162. PubMed PMID: 26413551; PubMed Central PMCID: PMCPMC4564626.
62. Boito SM, Laudy JA, Struijk PC, Stijnen T, Wladimiroff JW. Three-dimensional US assessment of hepatic volume, head circumference, and abdominal circumference in healthy and growth-restricted fetuses. Radiology. 2002;223(3):661-5. Epub 2002/05/30. doi: 10.1148/radiol.2233010656. PubMed PMID: 12034932.
63. Boito S, Struijk PC, Ursem NT, Stijnen T, Wladimiroff JW. Umbilical venous volume flow in the normally developing and growth-restricted human fetus. Ultrasound Obstet Gynecol. 2002;19(4):344-9. Epub 2002/04/16. doi: 10.1046/j.1469-0705.2002.00671.x. PubMed PMID: 11952962.
64. Chang CH, Yu CH, Ko HC, Chen CL, Chang FM. Predicting fetal growth restriction with liver volume by three-dimensional ultrasound: efficacy evaluation. Ultrasound Med Biol. 2006;32(1):13-7. Epub 2005/12/21. doi: 10.1016/j.ultrasmedbio.2005.09.002. PubMed PMID: 16364792.
65. Kuno A, Hayashi Y, Akiyama M, Yamashiro C, Tanaka H, Yanagihara T, et al. Three-dimensional sonographic measurement of liver volume in the small-for-gestational-age fetus. J Ultrasound Med. 2002;21(4):361-6. Epub 2002/04/06. doi: 10.7863/jum.2002.21.4.361. PubMed PMID: 11934092.
66. Gielchinsky Y, Zvanca M, Minekawa R, Persico N, Nicolaides KH. Liver volume in trisomy 21 and euploid fetuses at 11 to 13 weeks. Prenat Diagn. 2011;31(1):28-32. Epub 2010/12/31. doi: 10.1002/pd.2633. PubMed PMID: 21190270.
67. Hojo S, Tsukimori K, Kitade S, Nakanami N, Hikino S, Hara T, et al. Prenatal sonographic findings and hematological abnormalities in fetuses with transient abnormal myelopoiesis with Down syndrome. Prenat Diagn. 2007;27(6):507-11. Epub 2007/03/09. doi: 10.1002/pd.1718. PubMed PMID: 17345586.
68. Boito SM, Struijk PC, Ursem NT, Stijnen T, Wladimiroff JW. Assessment of fetal liver volume and umbilical venous volume flow in pregnancies complicated by insulin-dependent diabetes mellitus. BJOG. 2003;110(11):1007-13. Epub 2003/11/01. PubMed PMID: 14592586.
69. Sadler T. Langman's Medical embryology. 14th ed: Wolters Kluwer.; 2019.
70. Grijalva J, Vakili K. Neonatal liver physiology. Semin Pediatr Surg. 2013;22(4):185-9. Epub 2013/12/18. doi: 10.1053/j.sempedsurg.2013.10.006. PubMed PMID: 24331092.
71. Sherwood L. Human physiology - From Cells to Systems: Brooks/Cole; 2001.
72. Hall. Ga. Textbook of medical physiology. 14th ed: Elsevier; 2021.
73. Friedman SL, Flier JS, Epstein F, Glickman R, Scheele G. Seminars in Medicine of the Beth-Israel-Hospital, Boston - the Cellular Basis of Hepatic-Fibrosis - Mechanisms and Treatment Strategies. New Engl J Med. 1993;328(25):1828-35. PubMed PMID: WOS:A1993LH21000008.
74. Kumar VA, AK; Aster, JC. Robbins basic pathology. 10th ed: Elsevier; 2018.
75. Wynn TA. Cellular and molecular mechanisms of fibrosis. J Pathol. 2008;214(2):199-210. doi: 10.1002/path.2277. PubMed PMID: WOS:000252597000008.

76. Bonis PAL, Friedman SL, Kaplan MM. Is liver fibrosis reversible? New Engl J Med. 2001;344(6):452-4. doi: Doi 10.1056/Nejm200102083440610. PubMed PMID: WOS:000166800600010.
77. Iwaisako K, Brenner DA, Kisseleva T. What's new in liver fibrosis? The origin of myofibroblasts in liver fibrosis. J Gastroen Hepatol. 2012;27:65-8. doi: 10.1111/j.1440-1746.2011.07002.x. PubMed PMID: WOS:000300047000013.
78. Marcellin P, Gane E, Buti M, Afdhal N, Sievert W, Jacobson IM, et al. Regression of cirrhosis during treatment with tenofovir disoproxil fumarate for chronic hepatitis B: a 5-year open-label follow-up study. Lancet. 2013;381(9865):468-75. doi: 10.1016/S0140-6736(12)61425-1. PubMed PMID: WOS:000314810900032.
79. Yoon YJ, Friedman SL, Lee YA. Antifibrotic Therapies: Where Are We Now? Semin Liver Dis. 2016;36(1):87-98. doi: 10.1055/s-0036-1571295. PubMed PMID: WOS:000369951000010.
80. Younossi ZM, Ratziu V, Loomba R, Rinella M, Anstee QM, Goodman Z, et al. Obeticholic acid for the treatment of non-alcoholic steatohepatitis: interim analysis from a multicentre, randomised, placebo-controlled phase 3 trial. Lancet. 2019;394(10215):2184-96. doi: 10.1016/S0140-6736(19)33041-7. PubMed PMID: WOS:000502546300028.
81. Li X, Liao M, Pan Q, Xie Q, Yang H, Peng Y, et al. Combination therapy of obeticholic acid and ursodeoxycholic acid in patients with primary biliary cholangitis who respond incompletely to ursodeoxycholic acid: a systematic review. Eur J Gastroenterol Hepatol. 2020. Epub 2020/07/11. doi: 10.1097/MEG.0000000000001785. PubMed PMID: 32649329.
82. Kowdley KV, Vuppalanchi R, Levy C, Floreani A, Andreone P, LaRusso NF, et al. A randomized, placebo-controlled, phase II study of obeticholic acid for primary sclerosing cholangitis. J Hepatol. 2020;73(1):94-101. Epub 2020/03/14. doi: 10.1016/j.jhep.2020.02.033. PubMed PMID: 32165251.
83. de Vries EMG, de Krijger M, Farkkila M, Arola J, Schirmacher P, Gotthardt D, et al. Validation of the Prognostic Value of Histologic Scoring Systems in Primary Sclerosing Cholangitis: An International Cohort Study. Hepatology. 2017;65(3):907-19. doi: 10.1002/hep.28963. PubMed PMID: WOS:000397301300016.
84. de Vries EMG, Verheij J, Hubscher SG, Leeflang MMG, Boonstra K, Beuers U, et al. Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis. Journal of Hepatology. 2015;63(5):1212-9. doi: 10.1016/j.jhep.2015.06.008. PubMed PMID: WOS:000363307600022.
85. Karlsen TH, Folseraas T, Thorburn D, Vesterhus M. Primary sclerosing cholangitis - a comprehensive review. Journal of Hepatology. 2017;67(6):1298-323. PubMed PMID: WOS:000415325900022.
86. Molodecky NA, Kareemi H, Parab R, Barkema HW, Quan H, Myers RP, et al. Incidence of Primary Sclerosing Cholangitis: a Systematic Review and Meta-Analysis. Hepatology. 2011;53(5):1590-9. doi: 10.1002/hep.24247. PubMed PMID: WOS:000289956100020.
87. Tanaka A, Takikawa H. Geoepidemiology of primary sclerosing cholangitis: A critical review. J Autoimmun. 2013;46:35-40. doi: 10.1016/j.jaut.2013.07.005. PubMed PMID: WOS:000327004200005.

88. Jepsen P, Gronbaek L, Vilstrup H. Worldwide Incidence of Autoimmune Liver Disease. Digest Dis. 2015;33:2-12. doi: 10.1159/000440705. PubMed PMID: WOS:000366740200002.
89. Bambha K, Kim WR, Talwalkar J, Torgerson H, Benson JT, Therneau TM, et al. Incidence, clinical spectrum, and outcomes of primary sclerosing cholangitis in a United States community. Gastroenterology. 2003;125(5):1364-9. doi: 10.1016/j.gastro.2003.07.011. PubMed PMID: WOS:000186621500013.
90. Escorsell A, Pares A, Rodes J, Solisherruzo JA, Miras M, Delamorena E. Epidemiology of Primary Sclerosing Cholangitis in Spain. Journal of Hepatology. 1994;21(5):787-91. doi: Doi 10.1016/S0168-8278(94)80240-8. PubMed PMID: WOS:A1994PV81600016.
91. Lindkvist B, de Valle MB, Gullberg B, Bjornsson E. Incidence and Prevalence of Primary Sclerosing Cholangitis in a Defined Adult Population in Sweden. Hepatology. 2010;52(2):571-7. doi: 10.1002/hep.23678. PubMed PMID: WOS:000280513900019.
92. Boonstra K, Weersma RK, van Erpecum KJ, Rauws EA, Spanier BWM, Poen AC, et al. Population-Based Epidemiology, Malignancy Risk, and Outcome of Primary Sclerosing Cholangitis. Hepatology. 2013;58(6):2045-55. doi: 10.1002/hep.26565. PubMed PMID: WOS:000327385000052.
93. Bach JF. Infections and autoimmune diseases. J Autoimmun. 2005;25:74-80. doi: 10.1016/j.jaut.2005.09.024. PubMed PMID: WOS:000234371800015.
94. Kaplan GG, Laupland KB, Butzner D, Urbanski SJ, Lee SS. The burden of large and small duct primary sclerosing cholangitis in adults and children: A population-based analysis. American Journal of Gastroenterology. 2007;102(5):1042-9. doi: 10.1111/j.1572-0241.2007.01103.x. PubMed PMID: WOS:000246186700020.
95. Card TR, Solaymani-Dodaran M, West J. Incidence and mortality of primary sclerosing cholangitis in the UK: A population-based cohort study. Journal of Hepatology. 2008;48(6):939-44. doi: 10.1016/j.jhep.2008.02.017. PubMed PMID: WOS:000256683100008.
96. Suchy FS, RJ; Balistreri, WF. Liver disease in children. 4th ed2014.
97. Tanaka A, Tazuma S, Okazaki K, Tsubouchi H, Inui K, Takikawa H. Nationwide survey for primary sclerosing cholangitis and IgG4-related sclerosing cholangitis in Japan. J Hepato-Bil-Pan Sci. 2014;21(1):43-50. doi: 10.1002/jhbp.50. PubMed PMID: WOS:000328792500009.
98. Joo M, Abreu-e-Lima P, Farraye F, Smith T, Swaroop P, Gardner L, et al. Pathologic Features of Ulcerative Colitis in Patients With Primary Sclerosing Cholangitis A Case-control Study. Am J Surg Pathol. 2009;33(6):854-62. doi: DOI 10.1097/PAS.0b013e318196d018. PubMed PMID: WOS:000266464200006.
99. Riley TR, Schoen RE, Lee RG, Rakela J. A case series of transplant recipients who despite immunosuppression developed inflammatory bowel disease. American Journal of Gastroenterology. 1997;92(2):279-82. PubMed PMID: WOS:A1997WG31200019.
100. Schrumpf E, Elgjo K, Fausa O, Gjone E, Kolmannskog F, Ritland S. Sclerosing Cholangitis in Ulcerative-Colitis. Scand J Gastroentero. 1980;15(6):689-97. doi: Doi 10.3109/00365528009181516. PubMed PMID: WOS:A1980KS74800007.

101. Befeler AS, Lissoos TW, Schiano TD, Conjeevaram H, Dasgupta KA, Millis JM, et al. Clinical course and management of inflammatory bowel disease after liver transplantation. Transplantation. 1998;65(3):393-6. doi: Doi 10.1097/00007890-199802150-00017. PubMed PMID: WOS:000072089100017.
102. Saarinen S, Olerup O, Broome U. Increased frequency of autoimmune diseases in patients with primary sclerosing cholangitis. American Journal of Gastroenterology. 2000;95(11):3195-9. PubMed PMID: WOS:000165186200031.
103. Lunder AK, Hov JR, Borthne A, Gleditsch J, Johannesen G, Tveit K, et al. Prevalence of Sclerosing Cholangitis Detected by Magnetic Resonance Cholangiography in Patients With Long-term Inflammatory Bowel Disease. Gastroenterology. 2016;151(4):660-+. doi: 10.1053/j.gastro.2016.06.021. PubMed PMID: WOS:000389548500024.
104. Bungay HK, Buchel OC, Cummings F, Travis SPL, Chapman RWG. Prevalence and determinants of primary sclerosing cholangitis in a cohort of patients with inflammatory bowel disease and normal liver function tests. Gut. 2008;57:A41-A. PubMed PMID: WOS:000253090600107.
105. Cullen SN, Chapman RW. The medical management of primary sclerosing cholangitis. Semin Liver Dis. 2006;26(1):52-61. doi: DOI 10.1055/s-2006-933563. PubMed PMID: WOS:000237981400006.
106. Goode EC, Rushbrook SM. A review of the medical treatment of primary sclerosing cholangitis in the 21st century. Ther Adv Chronic Dis. 2016;7(1):68-85. doi: 10.1177/2040622315605821. PubMed PMID: WOS:000367810700005.
107. Bergquist A, Montgomery SM, Bahmanyar S, Olsson R, Danielsson A, Lindgren S, et al. Increased risk of primary sclerosing cholangitis and ulcerative colitis in first-degree relatives of patients with primary sclerosing cholangitis. Clin Gastroenterol H. 2008;6(8):939-43. doi: 10.1016/j.cgh.2008.03.016. PubMed PMID: WOS:000258292500020.
108. Ji SG, Juran BD, Mucha S, Folseraas T, Jostins L, Melum E, et al. Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease. Nat Genet. 2017;49(2):269-73. doi: 10.1038/ng.3745. PubMed PMID: WOS:000393148600016.
109. Ellinghaus D, Folseraas T, Holm K, Ellinghaus E, Melum E, Balschun T, et al. Genome-Wide Association Analysis in Primary Sclerosing Cholangitis and Ulcerative Colitis Identifies Risk Loci at GPR35 and TCF4. Hepatology. 2013;58(3):1074-83. doi: 10.1002/hep.25977. PubMed PMID: WOS:000329284000028.
110. Ellinghaus D, Jostins L, Spain SL, Cortes A, Bethune J, Han B, et al. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet. 2016;48(5):510-+. doi: 10.1038/ng.3528. PubMed PMID: WOS:000374834100010.
111. Liu JZ, Hov JR, Folseraas T, Ellinghaus E, Rushbrook SM, Doncheva NT, et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. Nat Genet. 2013;45(6):670-+. doi: 10.1038/ng.2616. PubMed PMID: WOS:000319563900015.
112. Melum E, Franke A, Schramm C, Weismuller TJ, Gotthardt DN, Offner FA, et al. Genome-wide association analysis in primary sclerosing cholangitis identifies two

non-HLA susceptibility loci. Nat Genet. 2011;43(1):17-9. doi: 10.1038/ng.728. PubMed PMID: WOS:000285683500008.
113. Karlsen TH, Franke A, Melum E, Kaser A, Hov JR, Balschun T, et al. Genome-Wide Association Analysis in Primary Sclerosing Cholangitis. Gastroenterology. 2010;138(3):1102-11. doi: 10.1053/j.gastro.2009.11.046. PubMed PMID: WOS:000275109900040.
114. Folseraas T, Melum E, Rausch P, Juran BD, Ellinghaus E, Shiryaev A, et al. Extended analysis of a genome-wide association study in primary sclerosing cholangitis detects multiple novel risk loci. Journal of Hepatology. 2012;57(2):366-75. doi: 10.1016/j.jhep.2012.03.031. PubMed PMID: WOS:000307685300023.
115. Srivastava B, Mells GF, Cordell HJ, Muriithi A, Brown M, Ellinghaus E, et al. Fine mapping and replication of genetic risk loci in primary sclerosing cholangitis. Scand J Gastroentero. 2012;47(7):820-6. doi: 10.3109/00365521.2012.682090. PubMed PMID: WOS:000305008500010.
116. Jiang XJ, Karlsen TH. Genetics of primary sclerosing cholangitis and pathophysiological implications (vol 14, pg 279, 2017). Nat Rev Gastro Hepat. 2017;14(5). PubMed PMID: WOS:000399997400006.
117. Dyson JK, Beuers U, Jones DEJ, Lohse AW, Hudson M. Primary sclerosing cholangitis. Lancet. 2018;391(10139):2547-59. doi: 10.1016/S0140-6736(18)30300-3. PubMed PMID: WOS:000436403700029.
118. Burak KW, Angulo P, Lindor KD. Is there a role for liver biopsy in primary sclerosing cholangitis? American Journal of Gastroenterology. 2003;98(5):1155-8. doi: DOI 10.1111/j.1572-0241.2003.07401.x. PubMed PMID: WOS:000183449300032.
119. Chapman MH, Thorburn D, Hirschfield GM, Webster GGJ, Rushbrook SM, Alexander G, et al. British Society of Gastroenterology and UK-PSC guidelines for the diagnosis and management of primary sclerosing cholangitis. Gut. 2019;68(8):1356-78. doi: 10.1136/gutjnl-2018-317993. PubMed PMID: WOS:000482497600004.
120. Kingham JGC, Kochar N, Gravenor MB. Incidence, clinical patterns, and outcomes of primary sclerosing cholangitis in South Wales, United Kingdom. Gastroenterology. 2004;126(7):1929-30. doi: 10.1053/j.gastro.2004.04.052. PubMed PMID: WOS:000221888300048.
121. Tischendorf JJW, Hecker H, Kruger M, Manns MP, Meier PN. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. American Journal of Gastroenterology. 2007;102(1):107-14. doi: 10.1111/j.1572-0241.2006.00872.x. PubMed PMID: WOS:000243376500018.
122. Ponsioen CY, Vrouenraets SME, Prawirodirdjo W, Rajaram R, Rauws EAJ, Mulder CJJ, et al. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. Gut. 2002;51(4):562-6. doi: DOI 10.1136/gut.51.4.562. PubMed PMID: WOS:000178243600022.
123. Wiesner RH, Grambsch PM, Dickson ER, Ludwig J, Maccarty RL, Hunter EB, et al. Primary Sclerosing Cholangitis - Natural-History, Prognostic Factors and Survival Analysis. Hepatology. 1989;10(4):430-6. doi: DOI 10.1002/hep.1840100406. PubMed PMID: WOS:A1989AT37000005.

124. Kim WR, Therneau TM, Wiesner RH, Poterucha JJ, Benson JT, Malinchoc M, et al. A revised natural history model for primary sclerosing cholangitis. Mayo Clin Proc. 2000;75(7):688-94. PubMed PMID: WOS:000088050900004.
125. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. Journal of Hepatology. 2015;63(1):237-64. PubMed PMID: WOS:000356314400027.
126. Corpechot C, El Naggar A, Poujol-Robert A, Ziol M, Wendum D, Chazouilleres O, et al. Assessment of biliary fibrosis by transient elastography in patients with PBC and PSC. Hepatology. 2006;43(5):1118-24. doi: 10.1002/hep.21151. PubMed PMID: WOS:000237327200028.
127. Ponsioen CY, Chapman RW, Chazouilleres O, Hirschfield GM, Karlsen TH, Lohse AW, et al. Surrogate Endpoints for Clinical Trials in Primary Sclerosing Cholangitis: Review and Results From an International PSC Study Group Consensus Process. Hepatology. 2016;63(4):1357-67. doi: 10.1002/hep.28256. PubMed PMID: WOS:000374736300029.
128. 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(9):e0203486. Epub 2018/09/15. doi: 10.1371/journal.pone.0203486. PubMed PMID: 30216377; PubMed Central PMCID: PMCPMC6138384.
129. Ferraioli G, Filice C, Castera L, Choi BI, Sporea I, Wilson SR, et al. Wfumb Guidelines and Recommendations for Clinical Use of Ultrasound Elastography: Part 3: Liver. Ultrasound in Medicine and Biology. 2015;41(5):1161-79. doi: 10.1016/j.ultrasmedbio.2015.03.007. PubMed PMID: WOS:000352803100004.
130. Henny J, Vassault A, Boursier G, Vukasovic I, Brguljan PM, Lohmander M, et al. Recommendation for the review of biological reference intervals in medical laboratories. Clin Chem Lab Med. 2016;54(12):1893-900. doi: 10.1515/cclm-2016-0793. PubMed PMID: WOS:000387508700018.
131. BureauInternationalDesPoidsEtMesures. Guide to the Expression of Uncertainty in Measurement 2008 [19-05-2020]. Available from: https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf.
132. Tapper EB, Lok ASF. Use of Liver Imaging and Biopsy in Clinical Practice. New Engl J Med. 2017;377(8):756-68. doi: 10.1056/NEJMra1610570. PubMed PMID: WOS:000413090000008.
133. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD. Liver Biopsy. Hepatology. 2009;49(3):1017-44. doi: 10.1002/hep.22742. PubMed PMID: WOS:000263817000035.
134. Seeff LB, Everson GT, Morgan TR, Curto TM, Lee WM, Ghany MG, et al. Complication Rate of Percutaneous Liver Biopsies Among Persons With Advanced Chronic Liver Disease in the HALT-C Trial. Clin Gastroenterol H. 2010;8(10):877-83. doi: 10.1016/j.cgh.2010.03.025. PubMed PMID: WOS:000283042500014.
135. Boyum JH, Atwell TD, Schmit GD, Poterucha JJ, Schleck CD, Harmsen WS, et al. Incidence and Risk Factors for Adverse Events Related to Image-Guided Liver Biopsy. Mayo Clin Proc. 2016;91(3):329-35. doi: 10.1016/j.mayocp.2015.11.015. PubMed PMID: WOS:000371248900011.

136. Baldea V, Sporea I, Tudor A, Popescu A, Bende F, Sirli R. Virtual Touch Quantification using Acoustic Radiation Force Impulse Imaging Technology versus Transient Elastography for the Noninvasive Assessment of Liver Fibrosis in Patients with Chronic Hepatitis B or C using Liver Biopsy as the Gold Standard. J Gastrointestin Liver Dis. 2020;29(2):181-90. Epub 2020/06/13. doi: 10.15403/jgld-2256. PubMed PMID: 32530985.
137. Furlan A, Tublin ME, Yu L, Chopra KB, Lippello A, Behari J. Comparison of 2D Shear Wave Elastography, Transient Elastography, and MR Elastography for the Diagnosis of Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. Am J Roentgenol. 2020;214(1):W20-W6. doi: 10.2214/Ajr.19.21267. PubMed PMID: WOS:000503833500003.
138. Sporea I, Sirli R, Bota S, Popescu A, Sendroiu M, Jurchis A. Comparative Study Concerning the Value of Acoustic Radiation Force Impulse Elastography (Arfi) in Comparison with Transient Elastography (Te) for the Assessment of Liver Fibrosis in Patients with Chronic Hepatitis B and C. Ultrasound in Medicine and Biology. 2012;38(8). doi: 10.1016/j.ultrasmedbio.2012.03.011. PubMed PMID: WOS:000306107300002.
139. Mjelle AB, Mulabecirovic A, Havre RF, Rosendahl K, Juliusson PB, Olafsdottir E, et al. Normal Liver Stiffness Values in Children: A Comparison of Three Different Elastography Methods. J Pediatr Gastroenterol Nutr. 2019;68(5):706-12. Epub 2019/03/20. doi: 10.1097/MPG.0000000000002320. PubMed PMID: 30889132; PubMed Central PMCID: PMCPMC6510324.
140. 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(1). doi: ARTN e0189671

10.1371/journal.pone.0189671. PubMed PMID: WOS:000419101600029.
141. Rowland M, McGee A, Broderick A, Drumm B, Connolly L, Daly LE, et al. Repeatability of transient elastography in children. Pediatr Res. 2020. doi: 10.1038/s41390-020-0916-4. PubMed PMID: WOS:000534533400001.
142. Ferraioli G, Tinelli C, Zicchetti M, Above E, Poma G, Di Gregorio M, et al. Reproducibility of real-time shear wave elastography in the evaluation of liver elasticity. Eur J Radiol. 2012;81(11):3102-6. doi: 10.1016/j.ejrad.2012.05.030. PubMed PMID: WOS:000309554400028.
143. Hudson JM, Milot L, Parry C, Williams R, Burns PN. Inter- and Intra-Operator Reliability and Repeatability of Shear Wave Elastography in the Liver: A Study in Healthy Volunteers. Ultrasound in Medicine and Biology. 2013;39(6):950-5. doi: 10.1016/j.ultrasmedbio.2012.12.011. PubMed PMID: WOS:000319260000003.
144. Vignier N, Esmat G, Elsharkawy A, Hassany M, Bonnard P, Delarocque-Astagneau E, et al. Reproducibility of liver stiffness measurements in hepatitis C virus (HCV)-infected patients in Egypt. J Viral Hepatitis. 2011;18(7):E358-E65. doi: 10.1111/j.1365-2893.2010.01433.x. PubMed PMID: WOS:000292265400027.
145. Fraquelli M, Rigamonti C, Casazza G, Conte D, Donato MF, Ronchi G, et al. Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. Gut. 2007;56(7):968-73. doi: 10.1136/gut.2006.111302. PubMed PMID: WOS:000247205000016.

146. Boursier J, Konate A, Gorea G, Reaud S, Quemener E, Oberti F, et al. Reproducibility of Liver Stiffness Measurement by Ultrasonographic Elastometry. Clin Gastroenterol H. 2008;6(11):1263-9. doi: 10.1016/j.cgh.2008.07.006. PubMed PMID: WOS:000261210900021.
147. Afdhal NH, Bacon BR, Patel K, Lawitz EJ, Gordon SC, Nelson DR, et al. Accuracy of Fibroscan, Compared With Histology, in Analysis of Liver Fibrosis in Patients With Hepatitis B or C: A United States Multicenter Study. Clin Gastroenterol H. 2015;13(4):772-+. doi: 10.1016/j.cgh.2014.12.014. PubMed PMID: WOS:000351299300027.
148. Boursier J, Isselin G, Fouchard-Hubert I, Oberti F, Dib N, Lebigot J, et al. Acoustic radiation force impulse: a new ultrasonographic technology for the widespread noninvasive diagnosis of liver fibrosis. Eur J Gastroen Hepat. 2010;22(9):1074-84. doi: 10.1097/MEG.0b013e328339e0a1. PubMed PMID: WOS:000280604500006.
149. Guzman-Aroca F, Reus M, Berna-Serna JD, Serrano L, Serrano C, Gilabert A, et al. 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(7):975-9. Epub 2011/06/28. doi: 10.7863/jum.2011.30.7.975. PubMed PMID: 21705730.
150. Ferraioli G, Tinelli C, Lissandrin R, Zicchetti M, Dal Bello B, Filice G, et al. Point shear wave elastography method for assessing liver stiffness. World J Gastroentero. 2014;20(16):4787-96. doi: 10.3748/wjg.v20.i16.4787. PubMed PMID: WOS:000335393600034.
151. Ling WW, Lu Q, Quan JR, Ma L, Luo Y. Assessment of impact factors on shear wave based liver stiffness measurement. Eur J Radiol. 2013;82(2):335-41. doi: 10.1016/j.ejrad.2012.10.004. PubMed PMID: WOS:000313093800029.
152. Rizzo L, Calvaruso V, Cacopardo B, Alessi N, Attanasio M, Petta S, et al. 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(12):2112-20. Epub 2011/10/06. doi: 10.1038/ajg.2011.341. PubMed PMID: 21971536.
153. Takahashi H, Ono N, Eguchi Y, Eguchi T, Kitajima Y, Kawaguchi Y, et al. Evaluation of acoustic radiation force impulse elastography for fibrosis staging of chronic liver disease: a pilot study. Liver Int. 2010;30(4):538-45. Epub 2009/10/31. doi: 10.1111/j.1478-3231.2009.02130.x. PubMed PMID: 19874490.
154. D'Onofrio M, Gallotti A, Mucelli RP. Tissue Quantification With Acoustic Radiation Force Impulse Imaging: Measurement Repeatability and Normal Values in the Healthy Liver. Am J Roentgenol. 2010;195(1):132-6. doi: 10.2214/Ajr.09.3923. PubMed PMID: WOS:000278998200017.
155. Friedrich-Rust M, Wunder K, Kriener S, Sotoudeh F, Richter S, Bojunga J, et al. Liver fibrosis in viral hepatitis: noninvasive assessment with acoustic radiation force impulse imaging versus transient elastography. Radiology. 2009;252(2):595-604. Epub 2009/08/26. doi: 10.1148/radiol.2523081928. PubMed PMID: 19703889.
156. Bota S, Sporea I, Sirli R, Popescu A, Danila M, Costachescu D. Intra- and Interoperator Reproducibility of Acoustic Radiation Force Impulse (Arfi) Elastography-Preliminary Results. Ultrasound in Medicine and Biology.

2012;38(7):1103-8. doi: 10.1016/j.ultrasmedbio.2012.02.032. PubMed PMID: WOS:000304995600001.
157. Suh CH, Kim SY, Kim KW, Lim YS, Lee SJ, Lee MG, et al. Determination of Normal Hepatic Elasticity by Using Real-time Shear-wave Elastography. Radiology. 2014;271(3):895-900. doi: 10.1148/radiol.14131251. PubMed PMID: WOS:000336894600033.
158. Goldschmidt I, Streckenbach C, Dingemann C, Pfister ED, di Nanni A, Zapf A, et al. Application and Limitations of Transient Liver Elastography in Children. J Pediatr Gastr Nutr. 2013;57(1):109-13. doi: 10.1097/MPG.0b013e31829206a0. PubMed PMID: WOS:000326744500029.
159. Castera L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of Liver Stiffness Measurement: A 5-Year Prospective Study of 13,369 Examinations. Hepatology. 2010;51(3):828-35. doi: 10.1002/hep.23425. PubMed PMID: WOS:000275472300014.
160. Sirli R, Sporea I, Bota S, Jurchis A. Factors influencing reliability of liver stiffness measurements using transient elastography (M-probe)-Monocentric experience. Eur J Radiol. 2013;82(8):E313-E6. doi: 10.1016/j.ejrad.2013.03.002. PubMed PMID: WOS:000321496900002.
161. Bende F, Mulabecirovic A, Sporea I, Popescu A, Sirli R, Gilja OH, et al. Assessing Liver Stiffness by 2-D Shear Wave Elastography in a Healthy Cohort. Ultrasound in Medicine and Biology. 2018;44(2):332-41. doi: 10.1016/j.ultrasmedbio.2017.10.013. PubMed PMID: WOS:000418603600005.
162. 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(11):2549-55. Epub 2014/10/06. doi: 10.1016/j.ultrasmedbio.2014.05.008. PubMed PMID: 25282481.
163. Roulot D, Costes JL, Buyck JF, Warzocha U, Gambier N, Czernichow S, et al. Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. Gut. 2011;60(7):977-84. doi: 10.1136/gut.2010.221382. PubMed PMID: WOS:000291306900017.
164. Colombo S, Belloli L, Zaccanelli M, Badia E, Jamoletti C, Buonocore M, et al. Normal liver stiffness and its determinants in healthy blood donors. Digest Liver Dis. 2011;43(3):231-6. doi: 10.1016/j.dld.2010.07.008. PubMed PMID: WOS:000293036900011.
165. Garate-Carrillo A, Gonzalez J, Ceballos G, Ramirez-Sanchez I, Villarreal F. Sex related differences in the pathogenesis of organ fibrosis. Transl Res. 2020. Epub 2020/04/15. doi: 10.1016/j.trsl.2020.03.008. PubMed PMID: 32289256.
166. Xu JW, Gong J, Chang XM, Luo JY, Dong L, Hao ZM, et al. Estrogen reduces CCL4- induced liver fibrosis in rats. World J Gastroentero. 2002;8(5):883-7. doi: DOI 10.3748/wjg.v8.i5.883. PubMed PMID: WOS:000178464200023.
167. Liu QH, Li DG, Huang X, Zong CH, Xu QF, Lu HM. Suppressive effects of 17 beta-estradiol on hepatic fibrosis in CCl4-induced rat model. World J Gastroentero. 2004;10(9):1315-20. PubMed PMID: WOS:000208097800017.
168. Xiong M, Li JY, Yang SL, Zeng FS, Ji YL, Liu J, et al. Influence of Gender and Reproductive Factors on Liver Fibrosis in Patients With Chronic Hepatitis B Infection. Clin Transl Gastroen. 2019;10(10). doi: ARTN e00085

10.14309/ctg.0000000000000085. PubMed PMID: WOS:000498805500008.
169. Ham S, Kim S, Choi H, Lee Y, Lee H. Greater Muscle Stiffness during Contraction at Menstruation as Measured by Shear-Wave Elastography. Tohoku J Exp Med. 2020;250(4):207-13. Epub 2020/04/03. doi: 10.1620/tjem.250.207. PubMed PMID: 32238619.
170. Li X, Wang JN, Fan ZY, Kang S, Liu YJ, Zhang YX, et al. Determination of the Elasticity of Breast Tissue during the Menstrual Cycle Using Real-Time Shear Wave Elastography. Ultrasound Med Biol. 2015;41(12):3140-7. Epub 2015/09/26. doi: 10.1016/j.ultrasmedbio.2015.07.013. PubMed PMID: 26403697.
171. Engelmann G, Gebhardt C, Wenning D, Wuhl E, Hoffmann GF, Selmi B, et al. Feasibility study and control values of transient elastography in healthy children. Eur J Pediatr. 2012;171(2):353-60. doi: 10.1007/s00431-011-1558-7. PubMed PMID: WOS:000300669500020.
172. Witters P, De Boeck K, Dupont L, Proesmans M, Vermeulen F, Servaes R, et al. Non-invasive liver elastography (Fibroscan) for detection of cystic fibrosis-associated liver disease. J Cyst Fibros. 2009;8(6):392-9. doi: 10.1016/j.jcf.2009.08.001. PubMed PMID: WOS:000272897400006.
173. Rubio A, Monpoux F, Huguon E, Truchi R, Triolo V, Rosenthal-Allieri MA, et al. Noninvasive Procedures to Evaluate Liver Involvement in HIV-1 Vertically Infected Children. J Pediatr Gastr Nutr. 2009;49(5):599-606. doi: DOI 10.1097/MPG.0b013e3181a15b72. PubMed PMID: WOS:000271666500010.
174. Fontanilla T, Canas T, Macia A, Alfageme M, Junquera CG, Malalana A, et al. 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 in Medicine and Biology. 2014;40(3):470-7. doi: 10.1016/j.ultrasmedbio.2013.10.024. PubMed PMID: WOS:000330573400002.
175. Matos H, Trindade A, Noruegas MJ. Acoustic Radiation Force Impulse Imaging in Paediatric Patients: Normal Liver Values. J Pediatr Gastr Nutr. 2014;59(6):684-8. doi: 10.1097/Mpg.0000000000000539. PubMed PMID: WOS:000345853800007.
176. Galina P, Alexopoulou E, Zellos A, Grigoraki V, Siahanidou T, Kelekis NL, et al. Performance of two--dimensional ultrasound shear wave elastography: reference values of normal liver stiffness in children. Pediatr Radiol. 2019;49(1):91-8. doi: 10.1007/s00247-018-4244-3. PubMed PMID: WOS:000454837300011.
177. Tokuhara D, Cho Y, Shintaku H. Transient Elastography-Based Liver Stiffness Age-Dependently Increases in Children. Plos One. 2016;11(11). doi: ARTN e0166683

10.1371/journal.pone.0166683. PubMed PMID: WOS:000388350300089.
178. Lewindon PJ, Balouch F, Pereira TN, Puertolas-Lopez MV, Noble C, Wixey JA, et al. Transient liver elastography in unsedated control children: Impact of age and intercurrent illness. J Paediatr Child H. 2016;52(6):637-42. doi: 10.1111/jpc.13151. PubMed PMID: WOS:000379592300011.
179. Zeng F, Wang H, Li X. The defining of the reference range of liver stiffness and fat attenuation parameter for healthy Chinese children. Eur J Gastroenterol

Hepatol. 2020. Epub 2020/08/17. doi: 10.1097/MEG.0000000000001874. PubMed PMID: 32796362.
180. Menten R, Leonard A, Clapuyt P, Vincke P, Nicolae AC, Lebecque P. Transient elastography in patients with cystic fibrosis. Pediatr Radiol. 2010;40(7):1231-5. doi: 10.1007/s00247-009-1531-z. PubMed PMID: WOS:000278582300008.
181. Eiler J, Kleinholdermann U, Albers D, Dahms J, Hermann F, Behrens C, et al. Standard value of ultrasound elastography using acoustic radiation force impulse imaging (ARFI) in healthy liver tissue of children and adolescents. Ultraschall Med. 2012;33(5):474-9. doi: 10.1055/s-0032-1313145. PubMed PMID: WOS:000311589400010.
182. Hanquinet S, Rougemont AL, Courvoisier D, Rubbia-Brandt L, Mclin V, Tempia M, et al. Acoustic radiation force impulse (ARFI) elastography for the noninvasive diagnosis of liver fibrosis in children. Pediatr Radiol. 2013;43(5):545-51. doi: 10.1007/s00247-012-2595-8. PubMed PMID: WOS:000317665700005.
183. Hamidieh AA, Shazad B, Ostovaneh MR, Behfar M, Tayebi S, Malekzadeh R, et al. 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 Tr. 2014;20(12):1912-7. doi: 10.1016/j.bbmt.2014.07.025. PubMed PMID: WOS:000345606500008.
184. Lee MJ, Kim MJ, Han KH, Yoon CS. Age-related changes in liver, kidney, and spleen stiffness in healthy children measured with acoustic radiation force impulse imaging. Eur J Radiol. 2013;82(6):E290-E4. doi: 10.1016/j.ejrad.2013.01.018. PubMed PMID: WOS:000318908300007.
185. Ozturk M, Caliskan E, Bayramoglu Z, Adaletli I. Normative values of pancreas stiffness by shear wave elastography in healthy children and adolescents. J Med Ultrason. 2020. doi: 10.1007/s10396-020-01039-1. PubMed PMID: WOS:000549246600001.
186. Canas T, Fontanilla T, Miralles M, Macia A, Malalana A, Roman E. Normal values of spleen stiffness in healthy children assessed by acoustic radiation force impulse imaging (ARFI): comparison between two ultrasound transducers. Pediatr Radiol. 2015;45(9):1316-22. doi: 10.1007/s00247-015-3306-z. PubMed PMID: WOS:000359292100006.
187. 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(4):309-14. PubMed PMID: WOS:000345159600006.
188. 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 in Medicine and Biology. 2016;42(9):2146-55. doi: 10.1016/j.ultrasmedbio.2016.04.016. PubMed PMID: WOS:000388314000014.
189. Thiele M, Madsen BS, Krag A. Is Liver Stiffness Equal to Liver Fibrosis? Hepatology. 2017;65(2):749-. doi: 10.1002/hep.28791. PubMed PMID: WOS:000397300700033.
190. Vigano M, Visentin S, Aghemo A, Rumi MG, Ronchi G. US features of liver surface nodularity as a predictor of severe fibrosis in chronic hepatitis C. Radiology.

2005;234(2):641-. doi: DOI 10.1148/radiol.2342041267. PubMed PMID: WOS:000226483200050.
191. Colli A, Fraquelli M, Andreoletti M, Marino B, Zuccoli E, Conte D. Severe liver fibrosis or cirrhosis: Accuracy of US for detection - Analysis of 300 cases. Radiology. 2003;227(1):89-94. doi: 10.1148/radiol.2272020193. PubMed PMID: WOS:000181901100013.
192. Soresi M, Giannitrapani L, Cervello M, Licata A, Montalto G. Non invasive tools for the diagnosis of liver cirrhosis. World J Gastroentero. 2014;20(48):18131-50. doi: 10.3748/wjg.v20.i48.18131. PubMed PMID: WOS:000347460700006.
193. Lee JH, Joo I, Kang TW, Paik YH, Sinn DH, Ha SY, et al. Deep learning with ultrasonography: automated classification of liver fibrosis using a deep convolutional neural network. Eur Radiol. 2020;30(2):1264-73. doi: 10.1007/s00330-019-06407-1. PubMed PMID: WOS:000511977900065.
194. Ehlken H, Wroblewski R, Corpechot C, Arrive L, Lezius S, Hartl J, et al. Spleen size for the prediction of clinical outcome in patients with primary sclerosing cholangitis. Gut. 2016;65(7):1230-+. doi: 10.1136/gutjnl-2016-311452. PubMed PMID: WOS:000377170200023.
195. Goode EC, Clark AB, Mells GF, Srivastava B, Spiess K, Gelson WTH, et al. Factors Associated With Outcomes of Patients With Primary Sclerosing Cholangitis and Development and Validation of a Risk Scoring System. Hepatology. 2019;69(5):2120-35. doi: 10.1002/hep.30479. PubMed PMID: WOS:000466449400021.
196. Thiele M, Madsen BS, Hansen JF, Detlefsen S, Antonsen S, Krag A. Accuracy of the Enhanced Liver Fibrosis Test vs FibroTest, Elastography, and Indirect Markers in Detection of Advanced Fibrosis in Patients With Alcoholic Liver Disease. Gastroenterology. 2018;154(5):1369-79. Epub 2018/01/11. doi: 10.1053/j.gastro.2018.01.005. PubMed PMID: 29317276.
197. Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, Lebail B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology. 2013;57(3):1182-91. doi: 10.1002/hep.25993. PubMed PMID: WOS:000315644200038.
198. Popescu A, Bota S, Sporea I, Sirli R, Danila M, Racean S, et al. The influence of food intake on liver stiffness values assessed by acoustic radiation force impulse elastography-preliminary results. Ultrasound Med Biol. 2013;39(4):579-84. Epub 2013/02/19. doi: 10.1016/j.ultrasmedbio.2012.11.013. PubMed PMID: 23415282.
199. Staugaard B, Christensen PB, Mossner B, Hansen JF, Madsen BS, Soholm J, et al. Feasibility of transient elastography versus real-time two-dimensional shear wave elastography in difficult-to-scan patients. Scand J Gastroentero. 2016;51(11):1354-9. doi: 10.1080/00365521.2016.1193217. PubMed PMID: WOS:000385581200013.
200. Lucidarme D, Foucher J, Le Bail B, Vergniol J, Castera L, Duburque C, et al. Factors of Accuracy of Transient Elastography (Fibroscan) for the Diagnosis of Liver Fibrosis in Chronic Hepatitis C. Hepatology. 2009;49(4):1083-9. doi: 10.1002/hep.22748. PubMed PMID: WOS:000264862100005.
201. Myers RP, Crotty P, Pomier-Layrargues G, Ma M, Urbanski SJ, Elkashab M. Prevalence, risk factors and causes of discordance in fibrosis staging by transient

elastography and liver biopsy. Liver International. 2010;30(10):1471-80. doi: 10.1111/j.1478-3231.2010.02331.x. PubMed PMID: WOS:000282634100010.
202. Wong VWS, Vergniol J, Wong GLH, Foucher J, Chan HLY, Le Bail B, et al. Diagnosis of Fibrosis and Cirrhosis Using Liver Stiffness Measurement in Nonalcoholic Fatty Liver Disease. Hepatology. 2010;51(2):454-62. doi: 10.1002/hep.23312. PubMed PMID: WOS:000274131200015.
203. Piscaglia F, Salvatorea V, Mulazzani L, Cantisani V, Colecchia A, Di Donato R, et al. Differences in liver stiffness values obtained with new ultrasound elastography machines and Fibroscan: A comparative study. Digest Liver Dis. 2017;49(7):802-8. doi: 10.1016/j.dld.2017.03.001. PubMed PMID: WOS:000406221200013.
204. Cassinotto C, Boursier J, de Ledinghen V, Lebigot J, Lapuyade B, Cales P, et al. Liver stiffness in nonalcoholic fatty liver disease: A comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. Hepatology. 2016;63(6):1817-27. Epub 2015/12/15. doi: 10.1002/hep.28394. PubMed PMID: 26659452.
205. Yoo HW, Kim YS, Kim SG, Yoo JJ, Jeong SW, Jang JY, et al. Usefulness of noninvasive methods including assessment of liver stiffness by 2-dimensional shear wave elastography for predicting esophageal varices. Dig Liver Dis. 2019;51(12):1706-12. Epub 2019/07/10. doi: 10.1016/j.dld.2019.06.007. PubMed PMID: 31281068.
206. Thiele M, Detlefsen S, Moller LS, Madsen BS, Hansen JF, Fialla AD, et al. Transient and 2-Dimensional Shear-Wave Elastography Provide Comparable Assessment of Alcoholic Liver Fibrosis and Cirrhosis. Gastroenterology. 2016;150(1):123-33. doi: 10.1053/j.gastro.2015.09.040. PubMed PMID: WOS:000366832800030.
207. Abe T, Kuroda H, Fujiwara Y, Yoshida Y, Miyasaka A, Kamiyama N, et al. Accuracy of 2D shear wave elastography in the diagnosis of liver fibrosis in patients with chronic hepatitis C. J Clin Ultrasound. 2018;46(5):319-27. Epub 2018/04/07. doi: 10.1002/jcu.22592. PubMed PMID: 29624693.

9. PAPERS I, II, III, IV

Ultrasound in Med. & Biol., Vol. 42, No. 9, pp. 2146–2155, 2016

0301-5629/$ - see front matter

http://dx.doi.org/10.1016/j.ultrasmedbio.2016.04.016

● *Original Contribution*

ULTRASOUND AND POINT SHEAR WAVE ELASTOGRAPHY IN LIVERS OF PATIENTS WITH PRIMARY SCLEROSING CHOLANGITIS

ANDERS BATMAN MJELLE,* ANESA MULABECIROVIC,* TRYGVE HAUSKEN,*† ROALD FLESLAND HAVRE,* ODD HELGE GILJA,*† and METTE VESTERHUS*‡

*National Centre for Ultrasound in Gastroenterology, Department of Medicine, Haukeland University Hospital, Bergen, Norway; †Department of Clinical Medicine, University of Bergen, Bergen, Norway; and ‡Norwegian PSC Research Center, Division of Cancer Medicine, Surgery and Transplantation, Department of Transplantation Medicine, Oslo University Hospital Rikshospitalet, Oslo, Norway

(*Received* 12 *October* 2015; *revised* 21 *April* 2016; *in final form* 26 *April* 2016)

Abstract—Point shear wave elastography (pSWE) is an ultrasound-based method for non-invasive quantification of liver fibrosis. The objective of this study was to explore liver pSWE in patients with primary sclerosing cholangitis (PSC) for assessment of fibrosis. Fifty-five non-transplant patients with PSC (38 males, 17 females; mean age: 46.4 y) were included and compared with 24 matched controls. Median (range) PSC duration was 8.1 (0–33) y. Ultrasonographic scanning followed by liver stiffness measurement by pSWE was performed using a conventional ultrasound system (Philips iU22). Signs of liver fibrosis on B-mode were identified in 21 patients (38%). Splenomegaly was found in 19 patients (35%) and ascites in two patients (4%). Successful pSWE measurements were achieved in the right liver lobe of all individuals and in the left liver lobe of 36 patients (65.5%). PSC patients had significantly higher median shear wave velocity (SWV) than controls in the right liver (median [range] SWV 1.26 [0.73–2.57] m/s vs. 1.09 [0.88–1.25] m/s, $p < 0.001$). SWV measured in the left liver lobe and spleen did not differ between PSC patients and controls. Our findings indicate that PSC patients have increased median SWV, indicating more fibrosis compared with controls; however, a wide range of SWV values were obtained among PSC patients, possibly reflecting the various stages in disease development. (E-mail: mette.vesterhus@helse-bergen.no) **© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).**

Key Words: **Primary sclerosing cholangitis, Point shear wave elastography, Non-invasive, Liver fibrosis, Elastography, Ultrasound.**

INTRODUCTION

Primary sclerosing cholangitis (PSC), a chronic inflammatory disease affecting the biliary tree, leads to liver fibrosis and cirrhosis over time, with a reported median transplant-free survival time of 12–21 y (Boonstra et al. 2013; Broomé et al. 1996). Medical therapy with proven benefit is lacking, and PSC is a frequent indication for transplantation.

A major challenge in PSC is the lack of valid prognostic markers and biomarkers of disease activity (Hirschfield et al. 2013; Karlsen et al. 2014). Fibrogenesis is an important pathogenetic pathway in PSC and a target of treatment in several clinical trials. A serum marker panel of fibrosis, the enhanced liver fibrosis (ELF) test, was reported to distinguish mild from advanced disease in PSC by an area under the curve of 0.81 and to predict prognosis independently of other biomarkers, underscoring the importance of accurate liver fibrosis estimation in PSC (Vesterhus et al. 2015). However, for other etiologies of liver fibrosis, some studies indicate an improved performance of ultrasound elastography compared with ELF or an incremental value of the combination of the ELF test and liver stiffness evaluation by ultrasound elastography (Cobbold et al. 2010; Wahl et al. 2012). Hence, better methods for the diagnosis, grading and monitoring of liver fibrosis are warranted.

Ultrasound elastography is a technique measuring liver stiffness as an expression of fibrosis and has emerged as an important tool in the diagnosis and follow-up of liver fibrosis and cirrhosis, largely replacing

Address correspondence to: Mette Vesterhus, Department of Medicine, Haukeland University Hospital, Jonas Lies vei 49, N-5021 Bergen, Norway. E-mail: mette.vesterhus@helse-bergen.no

liver biopsy in hepatitis B and C (Cosgrove et al. 2013; Ferraioli et al. 2015). The status of liver biopsy as the gold standard for liver fibrosis assessment has long been challenged because of its invasiveness and risk of serious complications, as well as the substantial sampling error and inter-observer variation between pathologists (Castera and Pinzani 2010; Cholongitas et al. 2006; Thampanitchawong and Piratvisuth 1999). Liver biopsy is generally not indicated in PSC for either diagnosis or follow-up because of the patchy disease distribution and consequent sampling bias, except in cases of suspected small-duct disease or autoimmune hepatitis overlap (Chapman et al. 2010; European Association for the Study of the Liver 2009). Ultrasound elastography has the advantages of being non-invasive and repeatable and offers the possibility of investigating several regions of the liver, thus reducing sampling bias. Guidelines for the use of elastography in clinical practice have been published (Bamber et al. 2013; Cosgrove et al. 2013); however, reports on elastography in PSC are scarce (Corpechot et al. 2006; Hagstrom et al. 2012; Righi et al. 2012).

Interestingly, a recent publication reported that baseline values of transient elastography (TE), as well as the change in liver stiffness measured by TE, are associated with clinical outcome in PSC (Corpechot et al. 2014). Point shear wave elastography (pSWE) is a more recent technology than TE, with the advantage of being incorporated into high-end ultrasound equipment, allowing B-mode ultrasound guidance of elastography measurements and an integration of liver stiffness measurement with a full evaluation of the liver. Some studies of pSWE in patient populations with chronic liver disease of heterogeneous etiologies have included PSC patients in small numbers insufficient for sub-analysis (Righi et al. 2012). To our knowledge, there are no studies exploring pSWE in PSC alone. In this study, we aimed to evaluate liver stiffness in PSC patients and compare them with healthy controls using ultrasound pSWE.

METHODS

Patient population and data collection

The protocol was in accordance with the Declaration of Helsinki and approved by the Regional Committee for Health and Research Ethics in Western Norway. Patients invited to participate in the study belonged to a known cohort of non-transplanted PSC patients in western Norway. Informed written consent was obtained from each patient enrolled. PSC patients with a histologically confirmed diagnosis of autoimmune hepatitis (AIH) were classified as PSC–AIH overlap. Patients were examined, and patient records were searched for information on clinical data, including ascites, encephalopathy, esophageal varices, variceal bleeding and inflammatory bowel disease status at the time of serum extraction. On the day of ultrasound and elastography, blood was sampled and biochemical analyses were performed using the standard routine laboratory protocols, including C-reactive protein, hemoglobin, leukocytes, platelets, creatinine, total bilirubin, albumin, International Normalization Ratio, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase and γ-glutamyl transferase. The Aspartate Aminotransferase-to-Platelet Ratio Index (APRI) and Fibrosis-4 (FIB-4) scores of fibrosis were calculated using published algorithms (Sterling et al. 2006; Wai et al. 2003). Mayo risk scores were calculated using the algorithm for the revised Mayo risk score (Kim et al. 2000). Blood samples were not taken from healthy controls.

B-Mode ultrasound examination

Immediately before pSWE examination, all patients underwent B-mode ultrasound scanning of the liver and spleen. All examinations were performed by a single operator (M.V.) using a standardized scanning protocol on a Philips iU22 (Philips Healthcare, Andover, MA, USA) scanner. Scores were registered for liver capsule regularity, parenchyma heterogeneity, liver angle appearance, presence of ascites, gallbladder stones or polyps and the presence of bile duct variability or sludge. Measures were taken for liver size in a sagittal section in the medioclavicular line; gallbladder length, width and area; spleen length and width; and portal vein diameter. Splenomegaly was defined as spleen length >12 cm.

Point shear wave elastography

Liver and spleen stiffness was measured in the fasting condition by pSWE using a conventional ultrasound system (ElastPQ, iU22, Philips Healthcare) equipped with a convex probe (C5-1). For liver measurements, patients and controls were examined in the supine position with their right arm maximally abducted. A 0.5×1.5-mm region of interest was placed 2–6 cm deeper than the liver capsule in hepatic tissue, avoiding large vessels or bile ducts (Fig. 1). Right lobe measurements were made in an intercostal position, whereas left liver lobe measurements were performed in a subcostal epigastric position, with sampling from the central portion of the left liver lobe. Spleen stiffness was measured by pSWE from a left-side intercostal position. All pSWE measurements were acquired in relaxed mid-breath hold with minimal scanhead pressure being applied. A valid measurement was defined as the median value of 10 acquisitions, provided the requirement for a success rate $\geq$60% was also fulfilled. The acquisitions were performed during separate breath holds in the same general area within one segment, avoiding visible bile ducts and blood

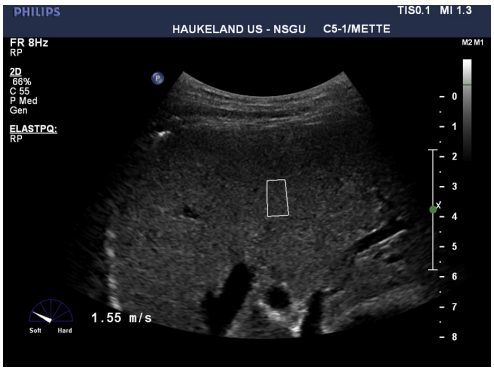


Fig. 1. Screen image of Philips iU22 in Elasto mode. B-Mode ultrasound image of a section of the right liver lobe in a patient with percutaneous sclerosing cholangitis. The *rectangular box* represents the region of interest where the elastography measurements are being performed. The region of interest has a fixed size but can be moved freely within the image down to a maximum depth of 8 cm. Measurements within 1 cm of the liver capsule or close to large vessels or bile ducts should be avoided. The measured SWV is given in the lower left corner; in this case it is 1.55 m/s, which indicates some degree of fibrosis. SWV = shear wave velocity.

vessels. Results were given as median shear wave velocity (SWV) in meters per second. All measurements were performed by a single investigator (M.V.). To evaluate the intra- and inter-observer variation, pSWE of the right liver lobe was performed twice in 16 healthy controls by the same investigator (M.V.) or two independent investigators (M.V. and A.M.), respectively, according to the protocol described previously.

Statistical analyses

Version 12.7.0.0 of SPSS 22 (IBM, Armonk, NY, USA) and MedCalc were used to perform all statistical analyses, and p values < 0.05 were considered to indicate significance. Variables were tested for normal distribution, and Student's t-test or the Mann–Whitney U-test were applied as appropriate. Data are presented as mean (SD), or as median (range) when the data were not normally distributed. Correlations between SWV and clinical parameters, biochemical scores of fibrosis or Mayo risk scores were tested by Spearman's rank order correlation coefficient (ρ). Intra-observer agreement was tested using the limits of agreement method (Bland and Altman 1999).

RESULTS

Sixty-four non-transplant PSC patients in a region of western Norway were identified and invited to participate; 55 (86%, 38 males, 17 females; mean age: 46.4 y; 95% confidence interval [CI]: 42.0–50.8) were included and compared with 24 healthy controls matched for age and gender. Median (range) PSC duration was 8.1 (0–33) y. Baseline demographic characteristics and clinical and laboratory data are summarized in Table 1. One patient with small-duct PSC and 5 patients with PSC–AIH overlap syndrome were included. In total, 3 patients had biochemical signs of clinically significant cholestasis or hepatitis as determined by a bilirubin level >30 (2 patients) or an alanine aminotransferase or aspartate aminotransferase level >5× the upper limit of normal (1 patient). There were no significant differences in distribution of age, gender or body mass index (BMI) between patients and controls. On the basis of B-mode ultrasound evaluation, signs compatible with advanced liver fibrosis, including liver capsule irregularity, periductal fibrosis and coarse liver parenchyma, were identified in 16 (29%), 5 (9%) and 12 (22%) patients, respectively, whereas 34 (62%) patients displayed no signs of fibrosis on B-mode ultrasound. Splenomegaly was found in 19 patients (35%) and ascites in 2 (4%) patients. Bile duct

Table 1. Baseline characteristics of 55 patients with PSC and 20 healthy controls undergoing ultrasound elastography

Baseline characteristics	PSC patients	Controls	p
N	55	24	
Males	37 (67%)*	11 (46%)	0.07
Age, mean (95% CI)	46 (42, 51)	43 (37, 49)	0.35
Age at diagnosis	34 (12–73)	NA	
Body mass index, mean (95% CI)	25.8 (24.8–26.8)	24.3 (22.9–25.8)†	0.09
PSC duration, y	8.1 (−0.6 to 32.8)	NA	
Inflammatory bowel disease, ever	47 (85.5%)	0	
Ulcerative colitis	32 (58.2%)	0	
Crohn's disease	8 (14.5%)	0	
Indeterminate colitis	7 (12.7%)	0	
Cholecystectomy	4 (7.3%)	0	
Mayo risk score	−0.4 (−2.1 to 1.9)	NA	
Laboratory data			
ALP, U/L	138 (25–838)	NI	
AST, U/L	45 (20–129)	NI	
ALT, U/L	49 (19–390)	NI	
Total bilirubin, μmol/L	11 (5–75)	NI	
γ-Glutamyl transpeptidase, U/L	177 (17–1576)	NI	
Albumin, g/L	46 (36–53)	NI	
INR	1.0 (0.9–1.2)	NI	
Platelets, 10^9/L	227 (60–765)	NI	

ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CI, confidence interval; INR = International Normalization Ratio; NA = not applicable; NI = not investigated; PSC = primary sclerosing cholangitis.

* Values are expressed as n (%) or median (range) unless otherwise noted.

† N = 17.

dilation was identified in 25 (46 %) patients. B-Mode ultrasound findings are summarized in Table 2.

pSWE of the right liver lobe

The intra- and inter-observer agreement for pSWE of the right liver lobe in the healthy controls was good, as evaluated by the limits of agreement method (Fig. 2). In PSC patients, valid pSWE measurements were achieved in all patients for the right liver lobe. The median success rate for the individual patients was 100% (range: 71.4%–100%). PSC patients had higher median SWV compared with the healthy controls (1.26 [0.73–2.57] and 1.09 [0.88–1.25] m/s, respectively, $p < 0.001$) (Fig. 3a). Area under the receiver operating characteristic curve (AUROC) analysis revealed fairly good discrimination for median SWV of the right liver lobe between PSC patients and controls, with an AUROC of 0.775 (95% CI: 0.67–0.86) (Fig. 4). For the discrimination of PSC patients from controls, the statistically optimal cutoff for SWV as decided by Youden's index was 1.24 m/s, with a sensitivity and specificity of 56.4 and 95.8, respectively.

Figure 5 illustrates right liver SWVs correlated with APRI and FIB-4 scores ($\rho = 0.494$, $p = 0.001$, and $\rho = 0.368$, $p = 0.017$, respectively) (Fig. 5a and b), whereas there was no significant correlation with the Mayo risk score ($\rho = 0.296$, $p = 0.06$) (Fig. 5c). No correlation was found between right liver SWVs and BMI, age or PSC duration (Table 3).

SWV in the right liver lobe was significantly higher in PSC patients with coarse liver parenchyma (median [range]: 1.88 [0.99–2.57] m/s vs. 1.22 [0.73–2.34] m/s, $p = 0.002$), irregular liver capsule (1.81 [1.11–2.57] m/s vs. 1.17 [0.73–2.43] m/s, $p = 0.001$) and periductal fibrosis (1.76 [1.27–2.09] m/s vs. 1.24 [0.73–2.57] m/s, $p = 0.049$), compared with patients with normal findings (Fig. 6a–c). Patients with none of these three visual signs of liver fibrosis had a median right liver SWV of 1.17 (0.73–2.34) m/s, compared with 1.76 (0.99–2.57) m/s among patients with minimum one B-mode sign of fibrosis ($p = 0.001$) (Fig. 6c). Right liver stiffness

Table 2. B-Mode ultrasound findings in PSC patients

Liver	
Liver size in MCL, cm	14.2 (8.0–28.5)*
Hepatomegaly (>16 cm)	10 (18.2%)
Liver capsule irregularity	16 (29.1%)
Coarse liver parenchyma	12 (21.8%)
Blunted liver angle	16 (29.1%)
Ascites	2 (3.6%)
Gallbladder and bile ducts	
Gallbladder length, cm	6.6 (2.1–10.0)
Gallbladder width, cm	3.0 (1.5–5.4)
Gallbladder wall thickness, mm	2.3 (0.3–10.4)
Gallbladder stone(s)	6 (10.9%)
Gallbladder polyp	1 (1.8%)
Cholecystectomy	4 (7.3%)
Bile duct variability	25 (45.5%)
Periductal fibrosis	5 (9.1%)
Spleen	
Spleen length, cm	12.6 (8.2–22.7)
Spleen area, cm^2	58.3 (25.4–165.7)
Splenomegaly (>13 cm)	19 (34.5%)

MCL = medioclavicular line; PSC = percutaneous sclerosing cholangitis.

* Values are expressed as n (%) or median (range).

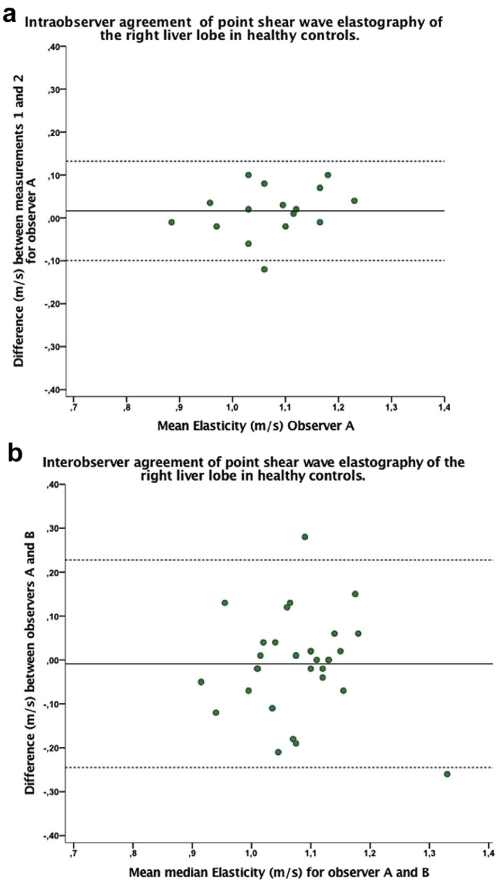


Fig. 2. Intra- and inter-observer agreement of point shear wave elastography of the right liver lobe in healthy controls. The Bland–Altman plots illustrate the (a) intra-observer and (b) inter-observer differences in liver stiffness evaluation measured by point shear wave elastography using ElastPQ (iU22, Philips) and expressed as shear wave velocity in meters per second. The *horizontal solid lines* represent the intra- or inter-observer mean ± 2 SD (limits of agreement, *dashed horizontal lines*), respectively. A valid measurement was defined as the median of 10 valid acquisitions with a success rate >60%. The measurements were performed twice by one (a) or two (b) observer(s) on the same day in the right liver lobe of healthy controls for the intra-observer agreement assessment.

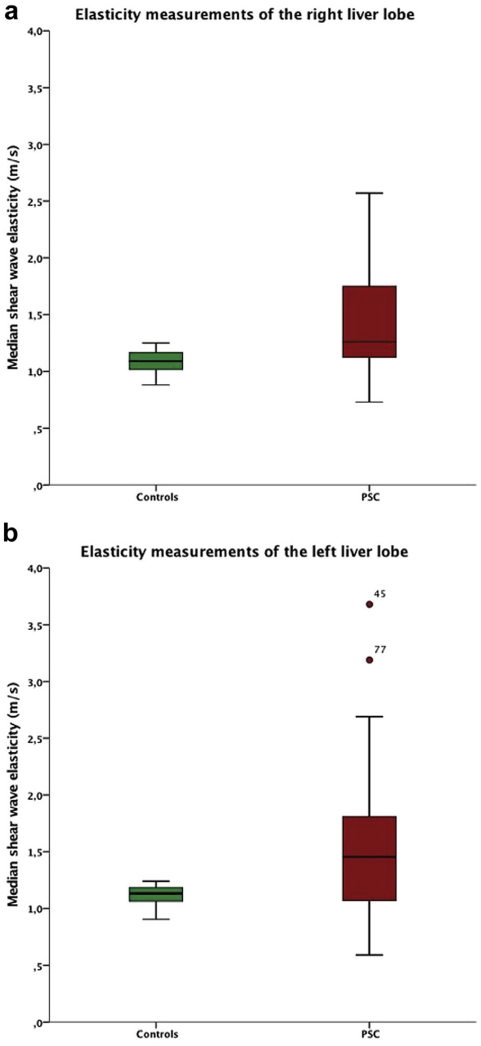


Fig. 3. Liver stiffness in PSC patients compared with controls. Liver stiffness evaluation by point shear wave elastography using iU22 (Philips) in 55 PSC patients and 24 healthy controls matched for age and gender in (a) the right liver lobe, and (b) the left liver lobe revealed increased liver stiffness in the right liver lobe of PSC patients compared with controls ($p < 0.001$). No significant difference could be found in the left liver lobe ($p = 0.11$). Liver stiffness is expressed as shear wave velocity in meters per second. PSC = primary sclerosing cholangitis.

assessed by median SWV did not differ significantly between patients with splenomegaly and patients without splenomegaly (1.42 [0.73–2.57] m/s vs. 1.24 [0.93–2.34] m/s, $p = 0.11$). Bile duct dilation was identified

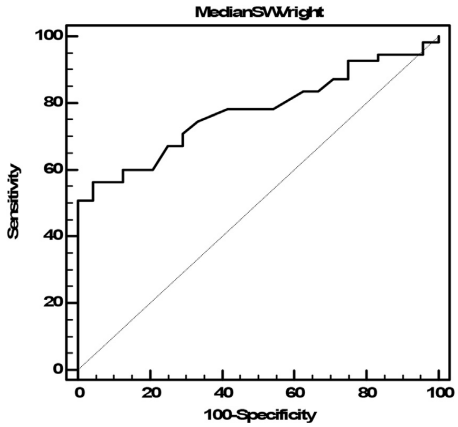


Fig. 4. Point shear wave elastography discriminates between primary sclerosing cholangitis patients and controls. Area under the receiver operating characteristic curve analysis revealed fairly good discrimination for median SWV measured by point shear wave elastography between 55 primary sclerosing cholangitis patients and 24 age- and gender-matched controls with an area under the curve of 0.775 (95% confidence interval: 0.67–0.86). The optimal cutoff for SWV as decided by Youden's index was 1.24 m/s, with a sensitivity and specificity of 56.4 and 95.8, respectively. SWV = shear wave velocity.

in 26 (47.3%) patients, but median right liver SWV did not differ between these patients and patients without bile duct dilation (1.32 [0.93-2.57] m/s vs. 1.24 [0.73-2.43] m/s, $p = 0.61$).

pSWE of the left liver lobe

In the left liver lobe, valid SWV measurements were acquired in 36 patients (66%), whereas in 19 patients (35%) the measurements were considered invalid because of too many failed acquisitions (success rate: <60%). The median success rate was 83% (46%–100%). Left liver lobe SWV did not significantly differ between PSC patients and controls (median [range] SWV: 1.46 [0.59–3.68] m/s vs. 1.13 [0.91–1.24] m/s, $p = 0.11$) (Fig. 3 b). There was no significant difference between median SWV of the right and left liver lobes in PSC patients ($p = 0.41$). Paired SWV values of the right and left liver lobes in the individual patient did not significantly correlate ($\rho = 0.233$, $p = 0.17$). Similarly, no significant correlation was found between left liver SWVs and BMI, age or PSC duration (Table 3).

pSWE of the spleen

Valid measurements were obtained in 37 PSC patients (67.3%), whereas in the remainder of the patients, measurements either were not performed (n = 16) or

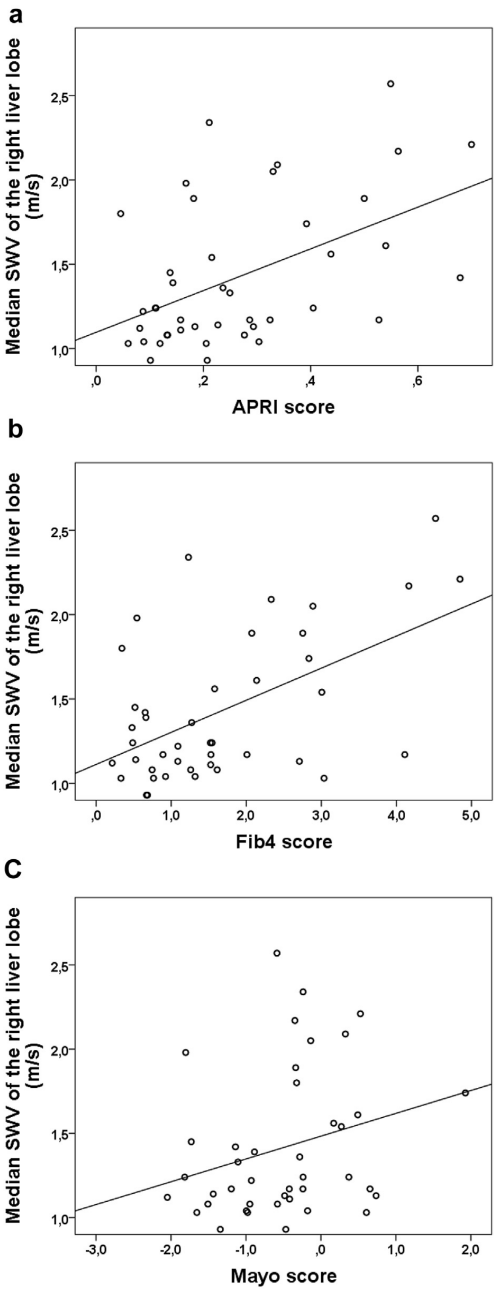


Fig. 5. Associations of liver stiffness with fibrosis scores and prognosis in percutaneous sclerosing cholangitis patients. The scatterplots with regression lines illustrate that SWV (m/s) of the right liver lobe as measured by point shear wave elastography using the iU22 (Philips) was correlated with the (a) APRI and (b) Fibrosis-4 (Fib4) scores of fibrosis, but not with (c) the Mayo risk score, a commonly used prognostic score in primary sclerosing cholangitis. SWV = shear wave velocity; APRI = Aspartate Aminotransferase-to-Platelet Ratio Index.

failed to fulfill the quality criteria (n = 2). The median success rate for the patients with valid measurements was 100% (76.9%–100%). There was no significant difference between PSC patients and controls (median SWV: 1.47 [0.79–3.13] m/s and 1.48 [1.17–1.80] m/s, respectively, $p = 0.83$). A tendency toward higher spleen SWV in patients with splenomegaly compared with patients without splenomegaly did not reach statistical significance (1.71 m/s [0.89–2.71] vs 1.39 m/s [0.79–3.13], respectively, $p = 0.05$).There was no correlation between spleen SWV and right or left liver SWV (1.47 m/s vs. 1.24 m/s, N = 37, $\rho = 0.104$, $p = 0.54$, and 1.39 m/s vs. 1.50 m/s, N = 22, $\rho = 0.331$, $p = 0.13$, respectively). One patient had variceal bleeding but did not have a high spleen SWV (median SWV: 1.55 m/s). No correlation was found between spleen SWV and BMI, age or PSC duration.

DISCUSSION

Our study indicates how liver fibrosis can be evaluated by both SWE and traditional B-mode findings during the same procedure. We found excellent intra- and inter-observer variation for pSWE using the iU22 system for the right liver lobe, in line with previous reports (Ling et al. 2013). The non-invasive evaluation of the degree of and change in liver fibrosis in PSC may be of major importance in evaluating the stage and prognosis of the disease, as indicated by recent reports (Corpechot et al. 2014, Vesterhus et al. 2015).

Liver biopsy is generally not indicated in PSC, based on the inherent sampling error resulting from the patchy distribution of fibrosis (European Association for the Study of the Liver 2009). The flaws of biopsies were illustrated in a study of whole-section scanning of 50 liver explants from patients with primary biliary cirrhosis, another disease of the biliary tree, in which only 20% of the primary biliary cirrhosis livers had a consistent histologic stage of fibrosis throughout the liver at clinically defined end-stage disease (Garrido and Hubscher 1996). Likewise, the distribution of liver fibrosis in PSC is uneven and follows the bile ducts to a large extent. In our opinion, it may therefore be preferable to use non-invasive methods assessing liver fibrosis covering larger areas of the liver in PSC. Ultrasound shear wave elastography is non-invasive and repeatable, can be integrated into a full liver examination and has been documented in viral hepatitis as a means of measuring liver fibrosis, but has not been previously explored in PSC

Table 3. Correlations of median SWV with other continuous variables in PSC patients

Clinical and laboratory variables	Median SWV right liver N	Spearman's ρ	*p*	median SWV left liver N	Spearman's ρ	*p*
Age	55	−0.004	0.974	36	−0.064	0.709
Age at diagnosis	54	−0.06	0.669	36	0.115	0.504
PSC duration	54	0.01	0.945	36	−0.179	0.295
Bilirubin	49	0.175	0.23	32	0.207	0.255
Albumin	50	−0.358	*0.011**	33	−0.222	0.215
ALP	50	0.246	0.085	33	0.335	0.057
AST	42	0.208	0.187	26	0.191	0.35
ALT	51	−0.072	0.616	34	0.222	0.207
Platelet count	51	−0.319	*0.022*	34	−0.301	0.084
Body mass index	55	−0.101	*0.463*	36	0.281	0.097
INR	47	0.325	*0.026*	31	−0.115	0.539
Mayo score	41	0.296	0.06	25	0.495	*0.012*
APRI score (AST/TRC)†	42	0.494	*0.001*	26	0.468	*0.016*

ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; APRI = AST-to-Platelet Ratio Index; INR = International Normalization Ratio; NA = not applicable; PSC = primary sclerosing cholangitis; SWV = shear wave velocity.

* *p*-Values in italic denote significance.

† See Figure 5a.

(Bota et al. 2013a, 2013b; Friedrich-Rust et al. 2012; Sporea et al. 2012a, 2012b).

Our findings suggest that PSC patients have increased median liver stiffness as expressed by SWV compared with healthy controls. The literature reveals scarce information about pSWE in PSC, but our data are in line with pSWE findings in a pilot study of patients with autoimmune liver diseases causing fibrosis, including PSC (Righi et al. 2012), and support previous findings describing TE in PSC (Corpechot et al. 2006, 2014; Hagstrom et al. 2012). In the present PSC cohort, 21 (38%) of the patients expressed B-mode signs of liver fibrosis. SWV was significantly higher in patients with B-mode signs of liver fibrosis; however, 12 patients without visual signs of liver fibrosis also had increased SWV (>1.24 m/s). Previous studies have reported good to excellent AUROCs between pSWE and histologic evaluation of fibrosis, even in autoimmune liver diseases (Friedrich-Rust et al. 2012; Righi et al. 2012). Thus, our findings could suggest an increased sensitivity in identifying fibrosis in PSC patients by adding elastography to an ultrasound liver evaluation. Furthermore, pSWE was associated with currently acknowledged signs of fibrosis, including TE in cystic fibrosis liver disease, which displays a patchy disease distribution similar to that of PSC (Behrens et al. 2013; Karlas et al. 2012; Manco et al. 2012; Monti et al. 2012). Because liver biopsy is generally not indicated in PSC, histologic correlates are lacking in the present study, but SWV correlated with serum-based scores of fibrosis, including APRI and FIB-4 scores.

Previous studies have reported that the performance of SWE and the cutoff values for significant fibrosis and cirrhosis may vary with the etiology of liver disease (Friedrich-Rust et al. 2012; Guzman-Aroca et al. 2012; Karlas et al. 2012; Sporea et al. 2012b). In the present study, the cutoff SWV value with the best statistical power to discriminate between PSC patients and healthy controls was 1.24 m/s with an area under the curve of 0.775, in line with previous findings suggesting 1.23 m/s as the statistically best cutoff between patients with chronic liver disease and controls (Sporea et al. 2014). It is an interesting characteristic of patients with PSC that this cutoff is similar to that of other liver disease populations, although it should be kept in mind that the clinically ideal cutoff value may differ depending on the aim of stratification (*e.g.,* early diagnosis of liver fibrosis or identification of a high-risk group). Longitudinal studies are needed to resolve whether pSWE can be used to follow disease progression in the individual patient for prognostic purposes.

The wider variability and lower success rate of ultrasound elastography of the left liver lobe has been debated (Karlas et al. 2011; Ling et al. 2013; Toshima et al. 2011). We were able to obtain valid liver stiffness measurements of the left liver lobe in 66% of the patients. SWVs of the left liver lobe correlated with the APRI score of fibrosis. There was a wider range of SWV measurements in the left compared with the right liver, probably caused by respiratory and cardiac movements affecting elastography measurements and suggesting reduced reliability of measurements in the left liver lobe. The lack of correlation between the two liver lobes may be due to the higher variability in SWV of the left liver lobe, and the definition of stricter quality criteria for pSWE measurements of the left liver lobe might yield better correlation for the valid measurements.

Previously published studies have indicated that cholestasis influences the accuracy of pSWE for the non-invasive evaluation of liver fibrosis (Pfeifer et al. 2014). PSC is a cholestatic disease, and this might be expected to complicate the evaluation of liver stiffness by pSWE in this patient group. However, only two patients had significantly elevated bilirubin >30 μmol/L; and although some degree of cholestasis was indicated in 26 (47%) patients by bile duct dilation on B-mode ultrasound, there was no difference in liver stiffness by pSWE in these patients compared with patients without bile duct dilation.

Several articles have reported that increased spleen stiffness alone or the spleen/liver stiffness ratio may predict high-risk esophageal varices and, thus, aid the

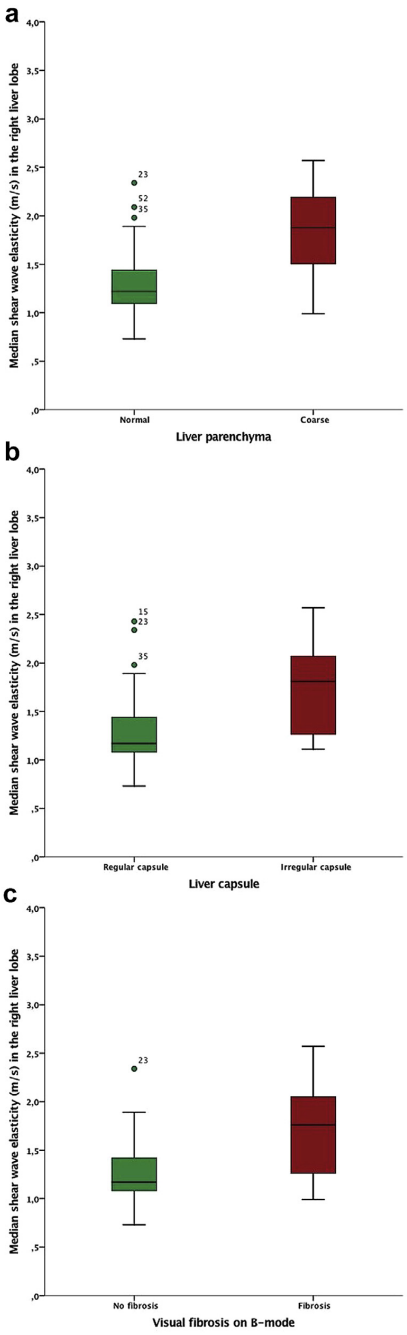


Fig. 6. Point shear wave elastography and B-mode signs of liver fibrosis. Boxplots illustrate SWV (m/s) reflecting liver stiffness as evaluated by point shear wave elastography using the iU22 (Philips) in patients with and without B-mode ultrasound signs of fibrosis, including (a) coarse liver parenchyma, (b) irregular liver capsule and (c) any of three signs of liver fibrosis (liver capsule irregularity, parenchyma coarseness, periductal fibrosis), compared with none of these.

identification of patients who should be selected to undergo gastroscopy (Berzigotti et al. 2013, 2014; Takuma et al. 2013; Sirli et al. 2015). In view of this, the spleen stiffness of PSC patients is of interest. In the present study, spleen stiffness did not significantly differ between PSC patients and controls. However, the number of patients was small, and only one variceal bleeding was observed in this cohort, precluding definitive conclusions. The association of liver and spleen stiffness with portal hypertension and esophageal varices in PSC should be further investigated in a larger cohort and, preferably, in a prospective setting.

Limitations of the study

The lack of liver biopsies in our cohort represents a limitation to the study. However, liver biopsies are not indicated in PSC, and biopsy was considered unethical for study purposes. Previous studies of chronic liver disease with a range of etiologies have repeatedly indicated excellent correlation between pSWE and histology findings. The question of prognostic value cannot be answered by the cross-sectional design of the present study, and further studies, including prospective follow-up, are warranted.

Although no standards of quality control have been agreed on for pSWE, we applied a standardized protocol and strict quality criteria as previously proposed, demanding 10 valid acquisitions with a success rate >60% to have reliable results (Bota et al. 2013a, 2013b). Measurements were made in two selected sites in the right and left liver lobes, respectively. Considering the patchy disease distribution in PSC, it is conceivable that SWV measurements throughout the entire liver would have revealed variable results within each lobe of the individual patient, and further studies should attempt to explore this.

CONCLUSIONS

We found that PSC patients have increased SWV in their liver parenchyma compared with healthy controls, indicating increased liver fibrosis. However, a wide range of SWV values were obtained for PSC patients, possibly reflecting various stages in PSC disease development. This novel method exhibited low intra- and inter-observer variation, making it suitable for further studies analyzing prospective follow-up data evaluating pSWE as a prognostic tool.

Acknowledgments—The authors thank Professor Tom H. Karlsen for kind advice regarding the constitution of the PSC cohort. The National Centre for Ultrasound in Gastroenterology, Haukeland University Hospital, is acknowledged for providing the ultrasound equipment, room and computer facilities. The study was supported by MedViz (http://medviz.uib.no/), an interdisciplinary research cluster from Haukeland University Hospital, University of Bergen and Christian Michelsen Research AS. M.V. is funded by the Norwegian PSC Research Center.

REFERENCES

Bamber J, Cosgrove D, Dietrich CF, Fromageau J, Bojunga J, Calliada F, Cantisani V, Correas JM, D'Onofrio M, Drakonaki EE, Fink M, Friedrich-Rust M, Gilja OH, Havre RF, Jenssen C, Klauser AS, Ohlinger R, Saftoiu A, Schaefer F, Sporea I, Piscaglia F. EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography: Part 1. Basic principles and technology. Ultraschall Med 2013;34:169–184.

Behrens CB, Langholz JH, Eiler J, Jenewein R, Naehrlich L, Fuchs K, Harth S, Krombach GA, Alzen GF. 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 2013;43:552–557.

Berzigotti A, Bosch J, Boyer TD. Use of noninvasive markers of portal hypertension and timing of screening endoscopy for gastroesophageal varices in patients with chronic liver disease. Hepatology 2014;59:729–731.

Berzigotti A, Seijo S, Arena U, Abraldes JG, Vizzutti F, Garcia-Pagan JC, Pinzani M, Bosch J. Elastography, spleen size, and platelet count identify portal hypertension in patients with compensated cirrhosis. Gastroenterology 2013;144:102–111.e1.

Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–160.

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, Ep-iPSCPBC Study group. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. Hepatology 2013;58:2045–2055.

Bota S, Herkner H, Sporea I, Salzl P, Sirli R, Neghina AM, Peck-Radosavljevic M. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver Int 2013a;33:1138–1147.

Bota S, Sporea I, Sirli R, Popescu A, Jurchis A. Factors which influence the accuracy of acoustic radiation force impulse (ARFI) elastography for the diagnosis of liver fibrosis in patients with chronic hepatitis C. Ultrasound Med Biol 2013b;39:407–412.

Broomé U, Olsson R, Lööf 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.

Castera L, Pinzani M. Biopsy and non-invasive methods for the diagnosis of liver fibrosis: Does it take two to tango? Gut 2010;59:861–866.

Chapman R, Fevery J, Kalloo A, Nagorney DM, Boberg KM, Shneider B, Gores GJ. Diagnosis and management of primary sclerosing cholangitis. Hepatology 2010;51:660–678.

Cholongitas E, Senzolo M, Standish R, Marelli L, Quaglia A, Patch D, Dhillon AP, Burroughs AK. A systematic review of the quality of liver biopsy specimens. Am J Clin Pathol 2006;125:710–721.

Cobbold JF, Crossey MM, Colman P, Goldin RD, Murphy PS, Patel N, Fitzpatrick J, Vennart W, Thomas HC, Cox IJ, Taylor-Robinson SD. Optimal combinations of ultrasound-based and serum markers of disease severity in patients with chronic hepatitis C. J Viral Hepat 2010;17:537–545.

Corpechot C, El Naggar A, Poujol-Robert A, Ziol M, Wendum D, Chazouillères O, de Lédinghen 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, Chazouillères 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 e915–e976.

Cosgrove D, Piscaglia F, Bamber J, Bojunga J, Correas JM, Gilja OH, Klauser AS, Sporea I, Calliada F, Cantisani V, D'Onofrio M, Drakonaki EE, Fink M, Friedrich-Rust M, Fromageau J, Havre RF, Jenssen C, Ohlinger R, Saftoiu A, Schaefer F, Dietrich CF, EFSUMB. EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography: Part 2. Clinical applications. Ultraschall Med 2013;34:238–253.

European Association for the Study of the Liver. EASL clinical practice guidelines: Management of cholestatic liver diseases. J Hepatol 2009;51:237–267.

Ferraioli G, Filice C, Castera L, Choi BI, Sporea I, Wilson SR, Cosgrove D, Dietrich CF, Amy D, Bamber JC, Barr R, Chou YH, Ding H, Farrokh A, Friedrich-Rust M, Hall TJ, Nakashima K, Nightingale KR, Palmeri ML, Schafer F, Shiina T, Suzuki S, Kudo M. WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 3. Liver. Ultrasound Med Biol 2015;41:1161–1179.

Friedrich-Rust M, Nierhoff J, Lupsor M, Sporea I, Fierbinteanu-Braticevici C, Strobel D, Takahashi H, Yoneda M, Suda T, Zeuzem S, Herrmann E. Performance of acoustic radiation force impulse imaging for the staging of liver fibrosis: A pooled meta-analysis. J Viral Hepat 2012;19:e212–e219.

Garrido MC, Hubscher SG. Accuracy of staging in primary biliary cirrhosis. J Clin Pathol 1996;49:556–559.

Guzmán-Aroca F, Frutos-Bernal MD, Bas A, Lujàn-Mompeán JA, Reus M, Berná-Serna Jde D, Parrilla P. Detection of non-alcoholic steatohepatitis in patients with morbid obesity before bariatric surgery: Preliminary evaluation with acoustic radiation force impulse imaging. Eur Radiol 2012;22:2525–2532.

Hagström H, Stål P, Stokkeland K, Bergquist A. Alcohol consumption in patients with primary sclerosing cholangitis. World J Gastroenterol 2012;18:3105–3111.

Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary sclerosing cholangitis. Lancet 2013;382:1587–1599.

Karlas T, Neuschulz M, Oltmanns A, Guttler A, Petroff D, Wirtz H, Mainz JG, Mössner J, Berg T, Tröltzsch M, Keim V, Wiegand J. Non-invasive evaluation of cystic fibrosis related liver disease in adults with ARFI, transient elastography and different fibrosis scores. Plos One 2012;7:e42139.

Karlas T, Pfrepper C, Wiegand J, Wittekind C, Neuschulz M, Mössner J, Berg T, Tröltzsch M, Keim V. Acoustic radiation force impulse imaging (ARFI) for non-invasive detection of liver fibrosis: Examination standards and evaluation of interlobe differences in healthy subjects and chronic liver disease. Scand J Gastroenterol 2011;46:1458–1467.

Karlsen TH, Vesterhus M, Boberg KM. Review article: Controversies in the management of primary biliary cirrhosis and primary sclerosing cholangitis. Aliment Pharmacol Ther 2014;39:282–301.

Kim WR, Therneau TM, Wiesner RH, Poterucha JJ, Benson JT, Malinchoc M, LaRusso NF, Lindor KD, Dickson ER. A revised natural history model for primary sclerosing cholangitis. Mayo Clin Proc 2000;75:688–694.

Ling W, Lu Q, Quan J, Ma L, Luo Y. Assessment of impact factors on shear wave based liver stiffness measurement. Eur J Radiol 2013;82:335–341.

Manco M, Zupone CL, Alghisi F, D'Andrea ML, Lucidi V, Monti L. Pilot study on the use of acoustic radiation force impulse imaging in the staging of cystic fibrosis associated liver disease. J Cyst Fibros 2012;11:427–432.

Monti L, Manco M, Lo Zupone C, Latini A, D'Andrea ML, Alghisi F, Lucidi V, Tomà P, Bonomo L. Acoustic radiation force impulse (ARFI) imaging with Virtual Touch tissue quantification in liver disease associated with cystic fibrosis in children. Radiol Med 2012;117:1408–1418.

Pfeifer L, Strobel D, Neurath MF, Wildner D. Liver stiffness assessed by acoustic radiation force impulse (ARFI) technology is considerably increased in patients with cholestasis. Ultraschall Med 2014;35: 364–367.

Righi S, Fiorini E, De Molo C, Cipriano V, Cassani F, Muratori L, Lenzi M, Morselli Labate AM, Serra C. ARFI elastography in patients with chronic autoimmune liver diseases: A preliminary study. J Ultrasound 2012;15:226–231.

Sirli R, Sporea I, Popescu A, Danila M. Ultrasound-based elastography for the diagnosis of portal hypertension in cirrhotics. World J Gastroenterol 2015;21:11542–11551.

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, Bota S, Peck-Radosavljevic M, Sirli R, Tanaka H, Iijima H, Badea R, Lupsor M, Fierbinteanu-Braticevici C, Petrisor A, Saito H, Ebinuma H, Friedrich-Rust M, Sarrazin C, Takahashi H, Ono N, Piscaglia F, Borghi A, D'Onofrio M, Gallotti A, Ferlitsch A, Popescu A, Danila M. Acoustic radiation force impulse elastography for fibrosis evaluation in patients with chronic hepatitis C: An international multicenter study. Eur J Radiol 2012a;81: 4112–4118.

Sporea I, Sirli R, Bota S, Popescu A, Sendroiu M, Jurchis A. Comparative study concerning the value of acoustic radiation force impulse elastography (ARFI) in comparison with transient elastography (TE) for the assessment of liver fibrosis in patients with chronic hepatitis B and C. Ultrasound Med Biol 2012b;38:1310–1316.

Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, S Sulkowski M, Torriani FJ, Dieterich DT, Thomas DL, Messinger D, Nelson M, APRICOT Clinical Investigators. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;43:1317–1325.

Takuma Y, Nouso K, Morimoto Y, Tomokuni J, Sahara A, Toshikuni N, Takabatake H, Shimomura H, Doi A, Sakakibara I, Matsueda K, Yamamoto H. Measurement of spleen stiffness by acoustic radiation force impulse imaging identifies cirrhotic patients with esophageal varices. Gastroenterology 2013;144:92–101.e102.

Thampanitchawong P, Piratvisuth T. Liver biopsy: Complications and risk factors. World J Gastroenterol 1999;5:301–304.

Toshima T, Shirabe K, Takeishi K, Motomura T, Mano Y, Uchiyama H, Yoshizumi T, Soejima Y, Taketomi A, Maehara Y. New method for assessing liver fibrosis based on acoustic radiation force impulse: A special reference to the difference between right and left liver. J Gastroenterol 2011;46:705–711.

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

Wahl K, Rosenberg W, Vaske B, Manns MP, Schulze-Osthoff K, Bahr MJ, Bantel H. Biopsy-controlled liver fibrosis staging using the enhanced liver fibrosis (ELF) score compared to transient elastography. PLoS One 2012;7:e51906.

Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, Lok AS. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 2003;38:518–526.

OPEN

Normal Liver Stiffness Values in Children: A Comparison of Three Different Elastography Methods

*†‡Anders B. Mjelle, †‡Anesa Mulabecirovic, ‡Roald F. Havre, †§Karen Rosendahl, *||¶Petur B. Juliusson, *Edda Olafsdottir, †‡Odd H. Gilja, and ‡||#**Mette Vesterhus

ABSTRACT

Objectives: Noninvasive tests for the evaluation of liver fibrosis are particularly helpful in children to avoid general anesthesia and potential complications of invasive tests. We aimed to establish reference values for 2 different elastography methods in a head-to-head comparison for children and adolescents 4 to 17 years, using transient elastography as common reference in a subset.

Methods: A total of 243 healthy participants aged 4 to 17 years were examined by a single observer with a full liver B-mode scan before elastography, following a minimum of 3 hours fasting. Liver stiffness measurements (LSMs) using 2-dimensional shear wave elastography (2D-SWE, GE Logiq E9) and point shear wave elastography (pSWE, Samsung RS80A with Prestige) were performed in all participants, and compared to transient elastography (TE, FibroScan) in a subset (n = 87). Interobserver agreement was evaluated in 50 children aged 4 to 17 years.

Results: Valid measurements were obtained in 242 of 243 (99.6%) subjects for 2D-SWE, 238 of 243 (97.9%) for pSWE, and in 83 of 87 (95.4%) for TE. Median liver stiffness overall was 3.3 (interquartile range [IQR] 2.7–4.3), 4.1 (IQR 3.6–4.7), and 4.1 kPa (IQR 3.5–4.6) for 2D-SWE, pSWE, and TE, respectively. Intraclass correlation coefficients between observers were 0.84 and 0.83 for 2D-SWE and pSWE, respectively. LSM values were significantly lower for 2D-SWE compared to pSWE and TE, and increased with advancing age. Higher LSM values in males were observed in adolescents.

Conclusions: All methods showed excellent feasibility. 2D-SWE showed significantly lower LSM values than pSWE and TE, and lower failure rate compared to TE. Our results further indicate an age and sex effect on LSM values.

Key Words: liver fibrosis, pediatric, shear wave elastography, transient elastography, ultrasound

(*JPGN* 2019;68: 706–712)

What Is Known

- Liver fibrosis is difficult to assess without a liver biopsy, which in children often require general anesthesia.
- Liver elastography is increasingly used as a noninvasive surrogate marker of liver fibrosis.
- Different ultrasound platforms yield different values.

What Is New

- This is the first publication of reference values for both point shear wave elastography and 2-dimensional shear wave elastography in healthy children, with comparison to transient elastography.
- Two-dimensional shear wave elastography yielded significantly lower liver stiffness measurements than point shear wave elastography and transient elastography.
- Liver stiffness measurements increased with age and were higher in male than female adolescents.

Ultrasound elastography is increasingly used to estimate liver fibrosis in children and adolescents, and has been reported in pediatric populations with mixed liver diseases (1–3) and in more homogenous populations (eg, biliary atresia) (4–8). The technique is continuously improving, and a sensitivity and specificity of 81% and 91%, respectively, was recently reported for liver fibrosis in children (2). The criterion standard for liver fibrosis evaluation is the liver biopsy, challenged due to its invasive nature, the need for general anesthesia and the potential of sampling errors and clinical complications (9). Ultrasound elastography is noninvasive and allows evaluation of the entire organ, with minimal discomfort for the patient, albeit lacking the liver biopsy's ability to assess the

Received October 31, 2018; accepted February 11, 2019.

From the *Department of Pediatric and Adolescent Medicine, Haukeland University Hospital, the †Department of Clinical Medicine, University of Bergen, the ‡National Centre for Ultrasound in Gastroenterology, the §Department of Radiology, Section of Pediatric Radiology, Haukeland University Hospital, the ||Department of Clinical Science, University of Bergen, Bergen, the ¶Department of Health Registries, Norwegian Institute of Public Health, the #Norwegian PSC Research Center, Oslo University Hospital, Oslo, and the **Department of Medicine, Haraldsplass Deaconess Hospital, Bergen, Norway.

Address correspondence and reprint requests to Anders B. Mjelle, MD, Department of Pediatric and Adolescent Medicine, Haukeland University Hospital, 5021 Bergen, Norway (e-mail: abmjelle@gmail.com).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (*www.jpgn.org*).

A.B.M., R.F.H., O.H.G., and M.V. have received speaker's fees from GE Healthcare. O.H.G. did consultancy for GE and Samsung (2017). The work is part of the PhD program for A.B.M., funded by the University of Bergen. The remaining authors report no conflicts of interest.

DOI: 10.1097/MPG.0000000000002320

etiology. Awareness of factors other than fibrosis which can influence liver stiffness is necessary (10). Transient elastography (TE) was introduced first and is established in several clinical settings. However, recent years have seen the introduction of alternative platforms allowing simultaneous B-mode imaging, including point shear wave (pSWE) and 2-dimensional shear wave elastography (2D-SWE), based on similar principles, with similar recommendations regarding use and application (10). Liver stiffness measurements (LSMs) have been demonstrated to vary between different elastography methods in both children (2) and adults (11,12); thus, normal values should be defined for each platform. There are currently no publications comparing LSM across platforms in healthy children. We aimed to establish and compare reference values for pSWE and 2D-SWE, with head-to-head comparison to TE in a subset.

METHODS

Subjects

The study was performed at Haukeland University Hospital in Bergen, Norway from September 2017 through January 2018. Participants were recruited through hospital employees, local schools, and social media. Exclusion criteria were a history of liver disease or chronic disease which could affect the liver. Informed written consent was obtained. A total of 246 children aged 4 to 17 years were recruited and grouped into 4 predefined age categories: 4 to 7; 8 to 11; 12 to 14; and 15 to 17 years. Two hundred thirty (94.7%) had Caucasian parents. The medical history was recorded, including the use of alcohol or nicotine. All were evaluated clinically by a pediatrician (A.B.M.) with >10 years' experience. Height, weight, waist circumference, and body mass index (BMI) were recorded and converted into z scores by the means of the Norwegian growth references (13,14). Weight classes were defined using International Obesity Task Force (IOTF) definitions (15). Blood tests were not performed. Participants classified as obese (n = 5) were included. Subjects with B-mode signs of steatosis or splenomegaly (n = 2), or fasting <3 hours (n = 1), were excluded, leaving 243 (108 boys, 44.4%) for further analyses.

B-mode Ultrasound Evaluation

B-mode ultrasound was performed after a standardized protocol with evaluation of the liver, gall bladder, spleen, and kidneys before elastography measurements, using Samsung RS80A with Prestige, with a convex 1 to 7 MHz probe. Skin to liver capsule distance was recorded. Examinations were conducted by a single operator (A.B.M.) with >2 years' experience in abdominal ultrasound.

Liver Stiffness Measurements

Ultrasound elastography measurements were performed in a supine position with the right arm maximally abducted, after ≥3 hours of fasting. Participants were examined with both GE Logiq E9 2D-SWE, using a C1–6 probe (GE Healthcare, Milwaukee, WI) and Samsung RS80A with Prestige pSWE, using a CA1–7A probe (Samsung Medison Co, Ltd, Seoul, Korea). SWE measurements were obtained in the right liver lobe applying minimal pressure through an intercostal space, perpendicular to the capsule, avoiding large liver vessels, bile ducts, and rib shadowing. Acquisitions were made during mid-expiratory breath hold if possible, otherwise during calm expiration. LSM values are expressed in meters per second (m/s) or kilopascals (kPa), the latter being calculated using the equation $\text{kPa} = 3(\text{ms}^{-1})^2$. We performed 10 valid acquisitions, and reported median values expressed in kPa for all systems. Every acquisition and mean, median, and interquartile range/median (IQR/M%; measure of dispersion) was recorded. A valid LSM value was defined as the median of 10 valid acquisitions with an IQR/M% ≤30%. For Samsung, values in m/s, average measurement depth and reliability measurement index, were automatically recorded, with a fixed region of interest (ROI) of 1 × 1.5 cm. For GE (2D-SWE), the ROI was a fixed circle with a diameter of 1 cm. ROIs were placed 2 to 5 cm from the liver capsule. In a subset of 50 subjects, 2 observers (A.B.M. and A.M.) both obtained data using GE and Samsung for interobserver reliability analysis. Both investigators had >2 years of experience in liver elastography. In a subset of patients ≥8 years (n = 87), TE using FibroScan (M-probe) incorporated in a GE S8 (GE Healthcare) was performed, reporting LSM results in kPa. For TE, the additional criterion of success rate ≥60% was adopted. The M-probe has been used extensively in children and adults with thorax perimeter under the recommended 75 cm, having been shown to affect feasibility only slightly (16); furthermore, we did not have access to the smaller S probe. The XL probe is known to yield lower values, and none of the subjects had skin to capsule distance ≥2.5 cm, for which an XL probe is warranted.

Controlled Attenuation Parameter

Fat deposits in hepatocytes affect ultrasound propagation, increasing the attenuation. Controlled attenuation parameter (CAP) evaluates the ultrasonic attenuation in the liver at 3.5 MHz at depth 25 to 65 mm using FibroScan, and represents a noninvasive assessment of liver steatosis (17). CAP values in dB/m were reported as the median of 10 acquisitions for all subjects evaluated by TE.

Statistical Analysis

For all analyses, SPSS version 25 (SPSS Inc, 2016, Armonk, NY) was used. Variables were tested for normality, and data were presented as mean (standard deviation [SD]) or median (range), as appropriate. When establishing age-specific reference values, mean ± 1.96 SD was used. For comparison of groups, standard paired T test, Wilcoxon signed rank test, or Pearson Chi-Square test were used as appropriate. Correlations were tested by Pearson correlation coefficient. Intraclass correlation coefficients (ICCs) were calculated to present interobserver reliability. Limits of agreement were assessed to reveal differences between platforms and observers. P values <0.05 were considered significant.

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 (2017/290/REK Vest).

RESULTS

The characteristics of the 243 subjects included are displayed in Table 1. Valid measurements were obtained in 242 of 243 (99.6%) for 2D-SWE, 238 of 243 (97.9%) for pSWE, and 83 of 87 (95.4%) for TE. TE feasibility was significantly lower than 2D-SWE feasibility ($P = 0.03$), but not different from pSWE ($P = 0.47$); nonfeasibility was most often due to wide dispersion (IQR/M% >30%) reflecting insufficient reliability. Median (range) IQR/M values were 10.1 (1.4–40), 13.9 (1.3–44), and 12.0 (3–44) for 2D-SWE, pSWE and TE, respectively. Among the excluded was an extreme outlier of 395% for pSWE, with corresponding reliability measurement index of 0.1 (the producer recommends ≥0.4).

TABLE 1. Baseline characteristics for all children and adolescents included for liver elastography in a study using ultrasound 2-dimensional shear wave elastography, point shear wave elastography, and transient elastography

	Total panel	4–7 y	8–11 y	12–14 y	15–17 y
Number	243	59	64	59	61
Males, number (%)	108 (44.4%)	31 (52.5%)	26 (40.6%)	30 (50.8%)	21 (34.4%)
Age, y, median (range)	11.7 (4.1–17.9)	6.3 (4.1–7.9)	10.0 (8.1–11.8)	13.4 (12–15.0)	17.1 (15–17.9)
Waist circumference, cm, median (range)	60.0 (45–98)	52 (45–59)	58 (50–75)	64 (51–85)	70 (60.5–98)
Weight, kg, median (range)	40.8 (13.7–105)	22.0 (13.7–32.7)	33.5 (22.2–61.7)	47.5 (28.7–80.7)	63.2 (41.6–105)
Body mass index (BMI)	17.6 (12–30.6)	15.5 (12–18.9)	17.2 (14–25.7)	17.9 (13.4–28.9)	21.5 (17.5–30.6)
Overweight or obese according to IOTF, n (%)	27 (11.1)	3 (5.1)	6 (9.4)	5 (8.5)	13 (21.3)
Mid-expiratory breath hold during measurement	195 (80.2%	11/59 (18.6%)	64/64 (100%)	59/59 (100%)	61/61 (100%)
Skin-to-capsule*, cm	1.11 (0.70–2.64)	0.9 (0.7–1.24)	1.03 (0.72–1.77)	1.17 (0.85–2.14)	1.41 (1.05–2.64)
Alcohol consumption last 72 h, n (%)	7 (2.9%)	0 (0%)	0 (0%)	0 (0%)	7 (11.5%)

IOTF = International Obesity Task Force.
*Distance from skin to liver capsule in centimeters.

Two hundred thirty-seven subjects (97.5%) showed valid results for both 2D-SWE and pSWE; 81 (93.1%) for all 3 platforms.

Median Liver Stiffness and Measurement Variability for 2-Dimensional Shear Wave Elastography, Point Shear Wave Elastography, and Transient Elastography

LSM values for 2D-SWE were significantly lower compared with pSWE or TE (median LSM 3.3, 4.1, and 4.1 kPa, respectively; $P < 0.001$), with no difference between pSWE and TE ($P = 0.65$) (Table 2). Moreover, the slope of LSM values was steeper for 2D-SWE compared to pSWE, with lower values for 2D-SWE compared to corresponding pSWE values for LSM <4 kPa by 2D-SWE, but higher values for 2D-SWE compared to corresponding pSWE values for LSM >5 kPa by 2D-SWE (Supplementary Fig. 1, Supplemental Digital Content, *http://links.lww.com/MPG/B612*). In the total panel, 2D-SWE and pSWE showed moderate correlation (rho = 0.51, $P < 0.001$).

The coefficient of variation (CV) between the 10 serial acquisitions forming a single LSM value, was low for all systems, ranging from 0.03 to 1.54, across all age groups (0.03–0.21 for pSWE, 0.03–0.24 for 2D-SWE, and 0.03–1.54 for TE). The highest CV for TE was due to a single high acquisition in 1 subject, which if excluded would have yielded a range of 0.03 to 0.29. The CV was slightly lower for 2D-SWE compared to pSWE ($P = 0.009$) and TE ($P = 0.006$), with similar values for pSWE and TE ($P = 0.12$).

Interobserver Evaluation

Interobserver reliability was evaluated in participants from all age groups with valid LSM for pSWE (n = 48) and 2D-SWE (n = 50). There were no significant differences between observers for pSWE (medians 4.10 kPa [IQR 3.6–4.9] vs 4.15 kPa [3.4–4.6]) or 2D-SWE (medians 3.55 kPa [IQR 2.8–4.3] vs 3.55 kPa [2.8–4.3]). ICCs between observers were 0.83 (95% confidence interval 0.7–0.91) and 0.84 (95% confidence interval 0.71–0.91) for pSWE and 2D-SWE, respectively, with no systematic differences between observers (Suppl. Figs. 2A and 2B, Supplemental Digital Content, *http://links.lww.com/MPG/B612*). The average difference (95% limits of agreement) was +2.1% (−26.2%–30.4%) and −0.1% (−35.9%–35.7%) for pSWE and 2D-SWE, respectively. A small number of subjects showed a discrepancy >1 kPa between observers, 4 of 48 (8.3%) and 5 of 50 (10%) for pSWE and 2D-SWE, respectively. Only 1 of 50 (2%) of subjects showed a difference >1.6 kPa using 2D-SWE, and none using pSWE.

TABLE 2. Liver stiffness measurements values by 2-dimensional shear wave elastography and point shear wave elastography for children aged 4 to 17 years

	GE Logiq E9 (2D-SWE)			Samsung RS80A (pSWE)		
	Mean value, kPa ± SD (97.5 percentile)	Range, kPa	Calculated mean value in m/s (97.5 percentile)	Mean value, kPa ± SD (97.5 percentile)	Range, kPa	Mean value, m/s ± SD (97.5 percentile)
4–7 y	2.87 ± 0.56 (3.96)	2.0–4.7	0.98 m/s (1.15)	3.93 ± 0.56 (5.03)	2.8–5.2	1.14 ± 0.08 (1.30)
8–11 y	3.45 ± 1.03 (5.47)	2.1–6.5	1.07 m/s (1.35)	4.17 ± 0.74 (5.61)	3.1–6.4	1.17 ± 0.10 (1.37)
12–14 y	3.83 ± 1.27 (6.32)	2.0–7.7	1.13 m/s (1.45)	4.65 ± 0.83 (6.27)	3.1–6.5	1.24 ± 0.11 (1.46)
15–17 y	3.96 ± 1.06 (6.03)	2.0–7.0	1.15 m/s (1.42)	4.23 ± 0.91 (6.02)	2.9–7.1	1.18 ± 0.12 (1.42)

Liver stiffness measurement (LSM) values represent the mean of LSM results from children in each age group. Individual LSM results are based on median values from 10 valid acquisitions.
2D-SWE = 2-dimensional shear wave elastography; pSWE = point shear wave elastography; SD = standard deviation.

Difference in Liver Elasticity by Age and Sex

For 2D-SWE, LSM was associated with age (rho = 0.421, $P < 0.001$) and increased steadily until 12 to 17 years, with no significant difference between the 2 older age groups (Table 2, Fig. 1).

For pSWE, LSM showed a weak association with age (rho = 0.146, $P = 0.02$) for the 3 youngest age groups (Table 2, Fig. 1). LSM age 8 to 11 was significantly lower than LSM age 12 to 14 years ($P = 0.001$), and significantly higher than LSM age 4 to 7 ($P = 0.04$), whereas LSM 15 to 17 years was significantly lower than 12 to 14 years ($P = 0.002$).

Overall, boys showed higher values compared to girls for pSWE. In subgroup analyses, sex difference was found in adolescence only: isolating ages 12 to 17 years, we found significantly higher LSM values in boys across all platforms, with mean values 4.27 versus 3.62 kPa ($P = 0.002$), 4.68 versus 4.27 kPa ($P = 0.01$), and 4.68 versus 4.13 kPa ($P = 0.02$) for 2D-SWE, pSWE, and TE, respectively (Fig. 2).

Obesity-related Characteristics, Liver Stiffness Measurement and Controlled Attenuation Parameter

BMI z scores or IOTF weight classes were not distributed evenly across age groups, and analyses and comparisons were performed within age groups. Overweight or obese groups yielded no significant differences for pSWE, when compared with the nonoverweight, but for 2D-SWE there was significantly higher LSM values for the overweight aged 8 to 11 (3.35 vs 4.45 kPa, $P = 0.01$) and 15 to 17 years (3.78 vs 4.64 kPa, $P = 0.008$), but not for ages 4 to 7 or 12 to 14 years. TE showed *lower* LSM values in the overweight aged 15 to 17 years (4.52 vs 3.4 kPa, $P = 0.03$), but not for 12 to 14 years.

The skin to capsule distance correlated with LSM for 2D-SWE (rho = 0.355, $P < 0.001$), but not with LSM for pSWE or TE ($P > 0.2$). In multiple linear regression (correcting for anthropometric measurements and age) skin to capsule distance was not independently associated with LSM. Skin to capsule distance was moderately correlated to anthropometric measures (rho = 0.516, 0.471 and 0.456 for BMI, waist, and weight z scores, respectively; all P values <0.001).

CAP values (estimating liver steatosis) were normally distributed, with an overall mean value 191.9 dB/m (SD 38.1), range 100 to 296. Six subjects displayed values above the proposed cut-off of 249 dB (18); 5 of these had normal weight. Using linear regression, CAP was associated with BMI z score ($P = 0.005$), but not with LSM by pSWE or 2D-SWE, age, or sex. CAP was shown to rise steadily already from BMI z score 0 and was significantly higher when comparing children with BMI z score ≥0 and BMI z score <0 (205 vs 180 dB, $P = 0.002$).

Quality Indicators and Associations With Body Mass Index and Skin-to-capsule Distance

IQR/M% is a frequently used quality indicator in LSM, with IQR/M% <30% commonly used as cutoff for a valid result in kPa,

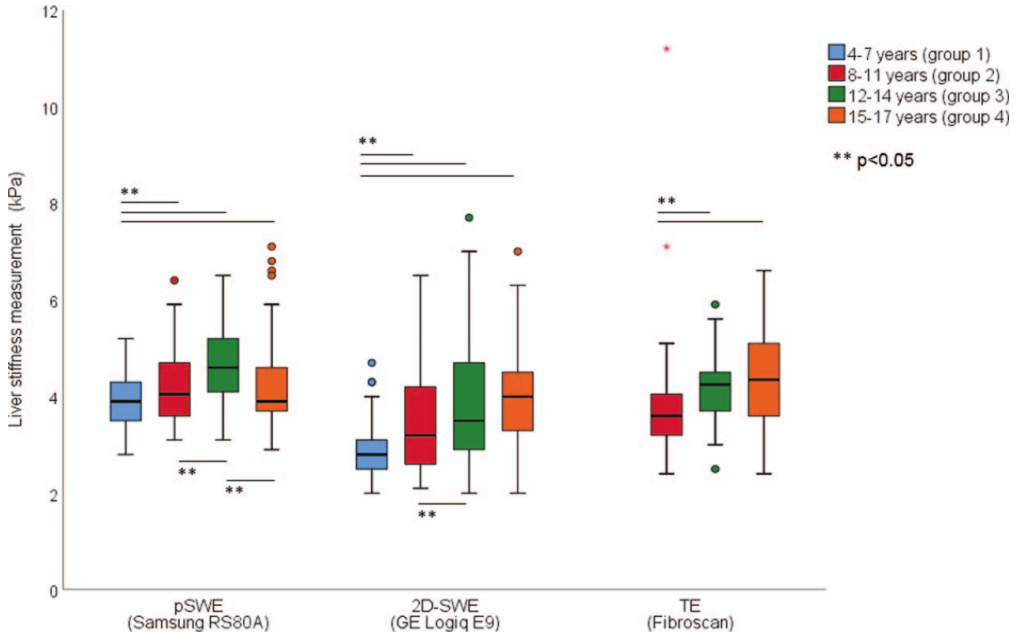


FIGURE 1. Liver stiffness measurements by age groups and elastography systems: point shear wave elastography (pSWE, all age groups, n = 238), 2-dimensional shear wave elastography (2D-SWE, all age groups, n = 243), and transient elastography (TE, age groups 2–4, n = 83). For pSWE: group 1 significantly lower liver stiffness measurements (LSM) values than groups 2 to 4 and group 3 significantly higher than group 2 and 4. For 2D-SWE: LSM values rising significantly from group 1 to group 2 and from group 2 to group 4. For TE: LSM values in group 3 and 4 were significantly higher than that in group 2.

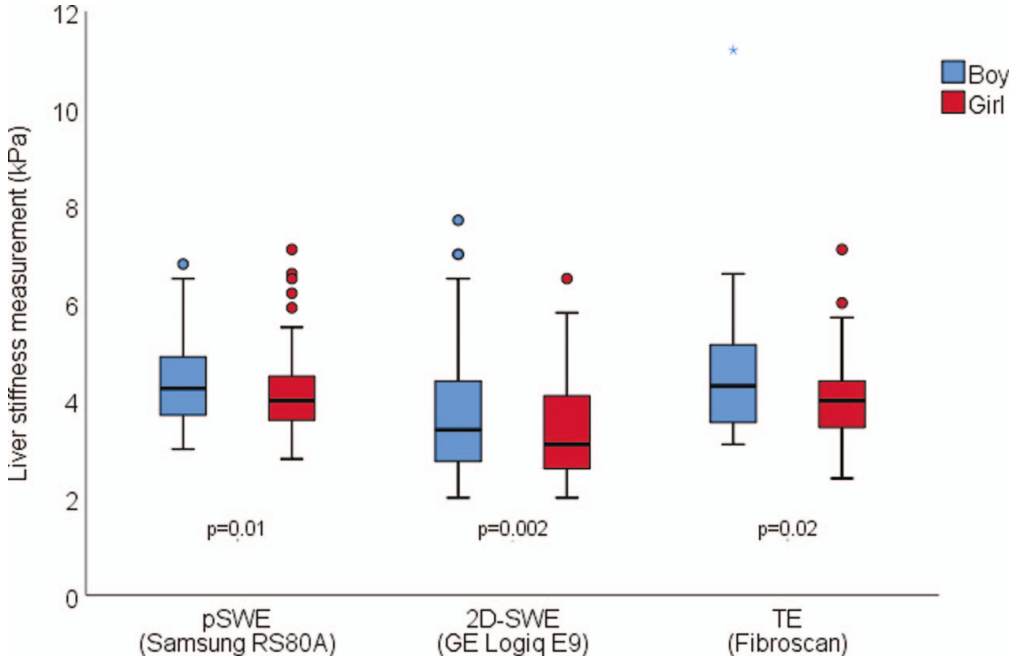


FIGURE 2. Liver stiffness measurements in boys (blue) and girls (red) aged 12 to 17 years as assessed by point shear wave elastography (pSWE), 2-dimensional shear wave elastography (2D-SWE), and transient elastography (TE). The figure shows liver stiffness as assessed by pSWE (Samsung RS80A with Prestige; n = 117), 2D-SWE (GE Logiq E9, n = 120), and TE (FibroScan, n = 60), respectively. Boxes represent the central 50% of the values, with the median value given as a horizontal line, and whiskers representing minimum and maximum, excluding outliers (small circles). Overall liver stiffness measurements (LSM) values are significantly lower in girls for all platforms: pSWE ($P = 0.01$), 2D-SWE ($P = 0.002$), and TE ($P = 0.02$).

reflecting limited spread in the values of acquisitions that collectively make up 1 LSM value. We found no correlation between IQR/M% and a high or low LSM value for pSWE and TE, and not for 2D-SWE after adjusting for age. As expected, IQR/M% values were approximately twice as high for measurements in kPa compared to m/s in the same individuals.

DISCUSSION

We report on the normal liver stiffness in healthy children aged 4 to 17 years. To our knowledge, this is the first head-to-head comparison between 2D-SWE and pSWE, with TE as common reference in a subset. Feasibility was excellent for all systems with failure rates of 0.4%, 2.1%, and 4.6% for 2D-SWE, pSWE, and TE, respectively, showing slightly but significantly superior feasibility for 2D-SWE compared to TE, in line with previous studies in adults. Failure was due to a high IQR/M >30% (n = 9) or success rate <60% (n = 1; TE). Feasibility was good using the M-probe for TE in children with a thorax perimeter <75 cm. We have demonstrated a low variation (CV) for the different platforms, and a good ICC between observers, but the average interobserver difference shown as 95% limits of agreement (−26.2%–30.4% and −35.9%–35.7% for pSWE and 2D-SWE, respectively) should be noted.

In the total panel, median liver stiffness was 3.3 kPa (1.05 m/s; range 2.0–7.7 kPa), 4.1 kPa (1.17 m/s; range 2.8–7.1 kPa), and 4.1 kPa (1.17 m/s; range 2.4–11.2 kPa) for 2D-SWE, pSWE, and TE, respectively. This is in line with previously published values for healthy children, reporting mean LSM of 4.6 kPa (19–24) and 1.12 m/s (25–31) by TE and acoustic radiation force impulse (ARFI), respectively (medians of reported values) (19–31).

Overall, we found significantly lower values using 2D-SWE compared to pSWE and TE ($P < 0.001$), whereas pSWE and TE showed similar LSM values ($P = 0.65$). This contrasts with reports in healthy adults, showing higher LSM values for 2D-SWE compared to pSWE and TE (11,12). Our results, however, indicate that the slope of the curve for LSM is steeper for 2D-SWE compared to pSWE; thus, in subjects with LSM >4 kPa the mean LSM by 2D-SWE was higher compared to LSM by pSWE (Suppl. Fig. 1, Supplemental Digital Content, *http://links.lww.com/MPG/B612*). Several publications comparing different platforms in adults (32–34) and phantoms (35,36) have shown similar relationships, but with these differences only evident outside the normal range. As higher LSMs are reported in healthy adults compared to children, the finding of higher values when using 2D-SWE compared to pSWE in adults (11), fits with our results. We can only speculate that 2D-SWE may yield higher LSM values than pSWE in children with significant liver fibrosis. We have not found other studies comparing the methods used herein. In a small study, children with heterogeneous chronic liver diseases were investigated with 2D-SWE, ARFI, and TE, but LSM values were not compared across platforms in the individual subjects (37) as in our head-to-head comparison.

Our findings strongly indicate that liver stiffness increases during childhood and adolescence. Previous studies on children investigating single elastography platforms (TE, 2D-SWE, and

ARFI, respectively) are discordant regarding the association of age with LSM; some report increasing values with age (19,21,23,25,31,38), some do not (20,22,26–28,39,40). We found higher values in boys compared to girls, during adolescence. This is in line with several adult studies, reporting higher values in males for 2D-SWE (12,41), pSWE (42), and TE (43,44); however, some describe no significant sex difference for ARFI (45) and pSWE (12). Three pediatric studies have similarly demonstrated higher LSM values in boys compared with girls (19,26,38), one of which only in older children, whereas 6 others did not find such a difference (21,22,25,27,31,39).

We aimed to describe liver stiffness in a healthy pediatric population, and therefore excluded subjects with B-mode signs of steatosis. Nevertheless, a minority of the subjects were overweight (n = 24) or obese (n = 3) as defined by IOTF. Although we found significant differences in LSM between the overweight and the normal weight in 2 age groups for 2D-SWE and 1 age group for TE, our results should be interpreted with caution due to a low number of overweight and obese in each age group. A recent publication (46) using a different pSWE method, studied the effect of BMI on LSM values, and found higher LSM values in the obese. The same study found no correlation between LSM value and BMI within the nonobese cohort, which included overweight children.

EFSUMB guidelines (10) describe the use of IQR/M% $\leq$30% as a quality indicator for pSWE and 2D-SWE, mimicking the FibroScan criterion, but do not mention that this parameter will change significantly based on the measurement unit. Comparing IQR/M% for m/s and kPa in our material, we illustrated a linear relationship, with the latter twice as high, demonstrating that the IQR/M% cut-off applies for kPa only (Suppl. Fig. 3, Supplemental Digital Content, *http://links.lww.com/MPG/B612*).

The present article does not allow conclusions regarding the superiority of any platform. In our experience, however, the availability of simultaneous B-mode imaging adds important value, allowing a full investigation in 1 session and sometimes a better assessment of factors influencing results. The learning curve for pSWE seems steeper and may thus be better for training users without extensive experience in B-mode imaging, whereas 2D-SWE may seem less prone to failed measurements in experienced users.

The study only includes healthy children and adolescents, and the findings are not automatically applicable in pediatric liver patients. Our findings and differences found have to some extent been described in adult liver patients, and this should be further investigated in children and adolescents with chronic liver diseases. Although there is little suspicion of liver diseases in our cohort, a biochemical evaluation of the participants would have strengthened our results.

CONCLUSIONS

For the first time we have described normal values for liver stiffness in healthy children using ultrasound elastography by pSWE and 2D-SWE in a head-to-head comparison, and compared to TE as reference in a subset. We demonstrated high feasibility in children for all methods with best results for the combination of ultrasound and elastography, but interobserver variability demonstrates the need of standardization of methods. Our results indicate different reference values depending on age and sex with increasing liver stiffness with age and particularly higher LSM values in boys during adolescence. This must be considered when using the methods in clinical practice. Further studies exploring elastography methods head-to-head in pediatric liver patients may further enhance our understanding of the differences between methods and the clinical utility of this tool.

Acknowledgments: The authors thank all the participants and their parents or guardians for their contribution. The authors also want to thank Samsung Medison Co, Ltd (Seoul, Korea) and GE Healthcare (Milwaukee, WI) for the opportunity to use the Samsung RS80A with Prestige and GE S8 with FibroScan, respectively, both free of charge. The companies mentioned had no influence on the design or performance of the study.

REFERENCES

1. Andersen SB, Ewertsen C, Carlsen JF, et al. Ultrasound elastography is useful for evaluation of liver fibrosis in children: a systematic review. *J Pediatr Gastroenterol Nutr* 2016;63:389–99.
2. Kim JR, Suh CH, Yoon HM, et al. The diagnostic performance of shear-wave elastography for liver fibrosis in children and adolescents: a systematic review and diagnostic meta-analysis. *Eur Radiol* 2018;28:1175–86.
3. Franchi-Abella S, Corno L, Gonzales E, et al. 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* 2016;278:554–62.
4. Rath T, Menendez KM, Kugler M, et al. TIMP-1/-2 and transient elastography allow non invasive diagnosis of cystic fibrosis associated liver disease. *Dig Liver Dis* 2012;44:780–7.
5. Monti L, Manco M, Lo Zupone C, et al. Acoustic radiation force impulse (ARFI) imaging with virtual touch tissue quantification in liver disease associated with cystic fibrosis in children. *Radiol Med* 2012;117:1408–18.
6. Chen S, Liao B, Zhong Z, et al. Supersonic shearwave elastography in the assessment of liver fibrosis for postoperative patients with biliary atresia. *Sci Rep* 2016;6:31057.
7. Tomita H, Hoshino K, Fuchimoto Y, et al. Acoustic radiation force impulse imaging for assessing graft fibrosis after pediatric living donor liver transplantation: a pilot study. *Liver Transpl* 2013;19:1202–13.
8. Pinto J, Matos H, Nobre S, et al. Comparison of acoustic radiation force impulse/serum noninvasive markers for fibrosis prediction in liver transplant. *J Pediatr Gastroenterol Nutr* 2014;58:382–6.
9. Potter C, Hogan MJ, Henry-Kendjorsky K, et al. Safety of pediatric percutaneous liver biopsy performed by interventional radiologists. *J Pediatr Gastroenterol Nutr* 2011;53:202–6.
10. Dietrich CF, Bamber J, Berzigotti A, et al. EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version). *Ultraschall Med* 2017;38:e48.
11. Mulabecirovic A, Mjelle AB, Gilja OH, et al. 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.
12. Bende F, Mulabecirovic A, Sporea I, et al. Assessing liver stiffness by 2-d shear wave elastography in a healthy cohort. *Ultrasound Med Biol* 2018;44:332–41.
13. Juliusson PB, Roelants M, Nordal E, et al. Growth references for 0-19 year-old Norwegian children for length/height, weight, body mass index and head circumference. *Ann Hum Biol* 2013;40:220–7.
14. Brannsether B, Roelants M, Bjerknes R, et al. Waist circumference and waist-to-height ratio in Norwegian children 4-18 years of age: reference values and cut-off levels. *Acta Paediatr* 2011;100:1576–82.
15. Cole TJ, Bellizzi MC, Flegal KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240–3.
16. Kim S, Kang Y, Lee MJ, et al. Points to be considered when applying FibroScan S probe in children with biliary atresia. *J Pediatr Gastroenterol Nutr* 2014;59:624–8.
17. Sasso M, Beaugrand M, De Ledinghen V, et al. Controlled attenuation parameter (CAP): a novel VCTE guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes. *Ultrasound Med Biol* 2010;36:1825–35.
18. Ferraioli G, Calcaterra V, Lissandrin R, et al. Noninvasive assessment of liver steatosis in children: the clinical value of controlled attenuation parameter. *BMC Gastroenterol* 2017;17:61.

19. Engelmann G, Gebhardt C, Wenning D, et al. Feasibility study and control values of transient elastography in healthy children. *Eur J Pediatr* 2012;171:353–60.
20. Menten R, Leonard A, Clapuyt P, et al. Transient elastography in patients with cystic fibrosis. *Pediatr Radiol* 2010;40:1231–5.
21. Witters P, De Boeck K, Dupont L, et al. Non-invasive liver elastography (Fibroscan) for detection of cystic fibrosis-associated liver disease. *J Cyst Fibros* 2009;8:392–9.
22. Goldschmidt I, Streckenbach C, Dingemann C, et al. Application and limitations of transient liver elastography in children. *J Pediatr Gastroenterol Nutr* 2013;57:109–13.
23. Rubio A, Monpoux F, Huguon E, et al. Noninvasive procedures to evaluate liver involvement in HIV-1 vertically infected children. *J Pediatr Gastroenterol Nutr* 2009;49:599–606.
24. Honsawek S, Vejchapipat P, Payungporn S, et al. Soluble receptor for advanced glycation end products and liver stiffness in postoperative biliary atresia. *Clin Biochem* 2013;46:214–8.
25. Fontanilla T, Canas T, Macia A, et al. 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* 2014;40:470–7.
26. Eiler J, Kleinholdermann U, Albers D, et al. Standard value of ultrasound elastography using acoustic radiation force impulse imaging (ARFI) in healthy liver tissue of children and adolescents. *Ultraschall Med* 2012;33:474–9.
27. Hanquinet S, Courvoisier D, Kanavaki A, et al. Acoustic radiation force impulse imaging-normal values of liver stiffness in healthy children. *Pediatr Radiol* 2013;43:539–44.
28. Lee MJ, Kim MJ, Han KH, et al. Age-related changes in liver, kidney, and spleen stiffness in healthy children measured with acoustic radiation force impulse imaging. *Eur J Radiol* 2013;82:e290–4.
29. Marginean CO, Marginean C. Elastographic assessment of liver fibrosis in children: a prospective single center experience. *Eur J Radiol* 2012;81:e870–4.
30. Noruegas MJ, Matos H, Goncalves I, et al. Acoustic radiation force impulse-imaging in the assessment of liver fibrosis in children. *Pediatr Radiol* 2012;42:201–4.
31. Matos H, Trindade A, Noruegas MJ. Acoustic radiation force impulse imaging in paediatric patients: normal liver values. *J Pediatr Gastroenterol Nutr* 2014;59:684–8.
32. Thiele M, Detlefsen S, Sevelsted Moller L, et al. Transient and 2-dimensional shear-wave elastography provide comparable assessment of alcoholic liver fibrosis and cirrhosis. *Gastroenterology* 2016;150:123–33.
33. Sporea I, Mare R, Lupusoru R, et al. Comparative study between four ultrasound shear waves elastographic methods for liver fibrosis assessment. *Med Ultrason* 2018;20:265–71.
34. Cassinotto C, Lapuyade B, Mouries A, et al. Non-invasive assessment of liver fibrosis with impulse elastography: comparison of supersonic shear imaging with ARFI and FibroScan(R). *J Hepatol* 2014;61:550–7.
35. Mulabecirovic A, Mjelle AB, Gilja OH, et al. Repeatability of shear wave elastography in liver fibrosis phantoms: evaluation of five different systems. *PLoS One* 2018;13:e0189671.
36. Mulabecirovic A, Vesterhus M, Gilja OH, et al. In vitro comparison of five different elastography systems for clinical applications, using strain and shear wave technology. *Ultrasound Med Biol* 2016;42:2572–88.
37. Belei O, Sporea I, Gradinaru-Tascau O, et al. Comparison of three ultrasound based elastographic techniques in children and adolescents with chronic diffuse liver diseases. *Med Ultrason* 2016;18:145–50.
38. Galina P, Alexopoulou E, Zellos A, et al. Performance of two-dimensional ultrasound shear wave elastography: reference values of normal liver stiffness in children. *Pediatr Radiol* 2019;49:91–8.
39. Fitzpatrick E, Quaglia A, Vimalesvaran S, et al. Transient elastography is a useful noninvasive tool for the evaluation of fibrosis in paediatric chronic liver disease. *J Pediatr Gastroenterol Nutr* 2013;56:72–6.
40. Hamidieh AA, Shazad B, Ostovaneh MR, et al. 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* 2014;20:1912–7.
41. Huang Z, Zheng J, Zeng J, et al. Normal liver stiffness in healthy adults assessed by real-time shear wave elastography and factors that influence this method. *Ultrasound Med Biol* 2014;40:2549–55.
42. Ling W, Lu Q, Quan J, et al. Assessment of impact factors on shear wave based liver stiffness measurement. *Eur J Radiol* 2013;82:335–41.
43. Roulot D, Costes JL, Buyck JF, et al. Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. *Gut* 2011;60:977–84.
44. Colombo S, Belloli L, Zaccanelli M, et al. Normal liver stiffness and its determinants in healthy blood donors. *Dig Liver Dis* 2011;43:231–6.
45. Madhok R, Tapasvi C, Prasad U, et al. Acoustic radiation force impulse imaging of the liver: measurement of the normal mean values of the shearing wave velocity in a healthy liver. *J Clin Diagn Res* 2013;7:39–42.
46. Bailey SS, Youssfi M, Patel M, et al. Shear-wave ultrasound elastography of the liver in normal-weight and obese children. *Acta Radiol* 2017;58:1511–8.

PLOS ONE

RESEARCH ARTICLE

Liver elasticity in healthy individuals by two novel shear-wave elastography systems—Comparison by age, gender, BMI and number of measurements

Anesa Mulabecirovic[1,2]*, Anders Batman Mjelle[1,2], Odd Helge Gilja[1,2], Mette Vesterhus[2,3], Roald Flesland Havre[2]

1 Department of Clinical Medicine, University of Bergen, Bergen, Norway, **2** National Centre for Ultrasound in Gastroenterology, Haukeland University Hospital, Bergen, Norway, **3** Norwegian PSC Research Center, Department of Transplantation Medicine, Division of Cancer Medicine, Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital, Oslo, Norway

* anesa.mulabecirovic@uib.no

OPEN ACCESS

Citation: Mulabecirovic A, Mjelle AB, Gilja OH, Vesterhus M, Havre RF (2018) Liver elasticity in healthy individuals by two novel shear-wave elastography systems—Comparison by age, gender, BMI and number of measurements. PLoS ONE 13(9): e0203486. https://doi.org/10.1371/journal.pone.0203486

Editor: Pavel Strnad, Medizinische Fakultat der RWTH Aachen, GERMANY

Received: May 3, 2018

Accepted: August 21, 2018

Published: September 14, 2018

Data Availability Statement: Due to ethical restrictions and local regulations on sharing a de-identified or anonymized data set, we are prohibited from sharing our data publicly. Public availability of our data would compromise participant privacy and confidentiality. Contact information for the Regional Committee for Medical and Health Research Ethics (REC West) to whom data requests may be sent is provided below. Please cite the study reference number: REK vest 2012/2214. Postal address: REK Vest,

Abstract

Objective

Establishing normal liver stiffness (LS) values in healthy livers is a prerequisite to differentiate normal from pathological LS values. Our aim was to define normal LS using two novel elastography methods head-to-head and to assess the number of measurements, variability and reproducibility.

Materials and methods

We evaluated shear wave elastography (SWE) methods integrated in Samsung RS80A and GE S8 by obtaining LS measurements (LSM) in 100 healthy subjects (20–70 years). Transient Elastography (TE) was used as reference method. Data were analyzed according to age, sex, BMI and 5 vs. 10 measurements. All subjects underwent B-mode ultrasound examination and lab tests to exclude liver pathology. Interobserver variation was evaluated in a subset (n = 24).

Results

Both methods showed excellent feasibility, measuring LS in all subjects. LSM-mean for GE S8 2D-SWE was higher compared to TE (4.5±0.8 kPa vs. 4.2±1.1, p<0.001) and Samsung RS80A (4.1±0.8 kPa, p<0.001). Both methods showed low intra- and interobserver variation. LSM-mean was significantly higher in males than females using 2D-SWE, while a similar trend for Samsung SWE did not reach significance. No method demonstrated statistical significant difference in LSM across age and BMI groups nor between LSM-mean based on 5 vs. 10 measurements.

Conclusion

LSM was performed with high reproducibility in healthy adult livers. LSM-mean was significantly higher for GE S8 2D-SWE compared to Samsung RS80A and TE in healthy livers.

Universitetet i Bergen, Det medisinske fakultet, Postboks 7804, 5020 Bergen E-mail: post@helseforskning.etikkom.no Telephone: 55 97 50 00 (switchboard at Haukeland University Hospital).

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Males had higher LSM than females. No method demonstrated statistical significant difference in LSM-mean across age- and non-obese BMI groups. Our results indicate that five LSM may be sufficient for reliable results.

Introduction

Chronic liver disease is one of the leading causes of morbidity and mortality worldwide [1, 2]. Assessment of liver fibrosis is important for chronic liver disease of various aetiologies for outcome prediction, risk stratification and selection for screening programs (e.g. endoscopy for oesophageal varices) as well as therapeutic decisions [2]. Non-invasive methods including ultrasound elastography have emerged within the past decade and are increasingly replacing liver biopsy for liver fibrosis assessment, avoiding the risks and discomforts of this invasive method. Nonetheless, ultrasound elastography encompasses several methods with important technological differences, ranging from vibration-controlled transient elastography (TE, Fibroscan) to methods based on deposition of an acoustic pulse such as point shear wave elastography (pSWE) and more recently 2D-SWE. TE has been extensively validated and is recommended for clinical use by several international guidelines, and an increasing number of studies evaluating the accuracy of various elastography methods have provided evidence for the utility of elastography imaging. However, with the expanding spectrum of ultrasound based elastography systems, it has become increasingly clear that the various technologies and platforms may yield different estimates of liver stiffness (LS) within the same liver. Hence, current guidelines acknowledge a need to establish reference values for normal liver stiffness in healthy livers for each specific equipment model in order to allow accurate diagnosis of pathological liver stiffness[3, 4].

To our knowledge, this is the first study to evaluate liver stiffness measurements (LSM) in healthy liver subjects using 2D-SWE from GE Logiq S8 (GE Healthcare, Milwaukee, Wi, USA) as well as SWE from Samsung RS80A (Samsung Medical, Seoul, Korea). Our study primarily aimed to define normal values of liver stiffness (LS) for males and females across adult age groups using these two novel platforms. We applied TE using Fibroscan integrated in the GE Logiq S8 ultrasound scanner (Echosens, Paris, France) as a reference method. Furthermore, we aimed to analyse influencing factors, such as BMI, and to assess the inter- and intraobserver variability and reproducibility, as well as to investigate the difference between obtaining five and ten consecutive liver stiffness measurements in order to calculate a representative median liver stiffness measurement (LSM).

Material and methods

Study design and subject population

The study was designed as a single-centre cross-sectional prospective study in selected healthy individuals. The protocol was in accordance with the Declaration of Helsinki for research in medicine and biology, and was approved by the Regional Committee for Medical and Health Research Ethics in Western Norway. All subjects were given oral and written information about the study and were invited to participate. Informed written consent was obtained from each subject enrolled. The study was performed in August and September 2017, at the Department of Gastroenterology, Haukeland University Hospital in Bergen, Norway.

Table 1. The characteristics of healthy subjects, by age group.

	20–30 (n = 20)	31–40 (n = 20)	41–50 (n = 20)	51–60 (n = 20)	61–70 (n = 20)
Characteristics					
Age, years$^{\perp}$ (range)	27.8 ± 2 (25–30)	34.1 ± 2.7 (31–40)	44.3 ± 3 (41–50)	55.7 ± 3 (51–60)	64.15 ± 2.5 (61–69)
Gender; Female/Male, n	10/10	10/10	10/10	10/10	10/10
BMI, kg/m$^{2\perp}$ (range)	22.5 ± 2.4 (19.4–27.2)	23.8 ± 2.3 (20.7–28.4)	24.4 ± 3 (18.1–28.7)	24.5 ± 2.3 (20–29.9)	24.7 ± 2.9 (20–29.6)
Weight group$^{\perp}$, n (%)					
18.0–25 kg/m^2	17 (17%)	15 (15%)	15 (15%)	13 (13%)	13 (13%)
25.0–30 kg/m^2	3 (3%)	5 (5%)	5 (5%)	7 (7%)	7 (7%)
Alcohol units per week$^{\perp}$(range)	4.4 ± 2.5 (0–8)	3.8 ± 2 (0–6)	3.7 ± 2.6 (1–10)	4.4 ± 2.6 (0–10)	4.2 ± 3 (0–10)
Biochemical profile					
Total bilirubin μmol/L* [<19 μmol/L] *(IQR; range)*	10.5 (8.7–12.5; 4–18)	17 (5.5–15.3; 3–39)	9.5 (7.5–11.4; 4–21)	8.5 (7.1–12.1; 4–28)	8.6 (7.4–9.7; 4–15)
AST, U/L * [15–45 U/L] *(IQR; range)*	23.5 (21.9–28.6; 17–41)	25.0 (21.6–27.7; 15–45)	22.0 (20.5–26.1; 15–40)	23.0 (21.7–25.7; 16–32)	23.5 (21.5–27.8; 13–37)
ALT, U/L * [10–70 U/L] *(IQR; range)*	17.0 (16.9–28.4; 11–55)	22.5 (18.7–27.1; 8–45)	23.0 (21.5–30.7; 15–47)	22.5 (20.1–30.7; 14–40)	24.5 (22–28.7; 14–46)
GGT, U/L* [10–115 U/L] *(IQR; range)*	18.5 (15.5–22.1; 7–42)	15.0 (13.5–19.5; 9–29)	19.0 (17.0–34.4; 5–68)	17.0 (16.7–26.0; 11–48)	22.5 (17.0–35.3; 9–96)
Serum Albumin, g/L* [39–50 g/L] *(IQR; range)*	48.0 (46.4–49.3; 41–53)	47.0 (45.6–48.2; 43–52)	46.5 (45.2–47.4; 41–51)	46.5 (45.2–49.7; 43–66)	46.0 (45.1–46.8; 43–50)
Platelet counts, 10^9 /L* [145–387 10^9 /L] *(IQR; range)*	237.5 (210.7–255.5; 152–352)	222.5 (210.6–254.8; 155–334)	256.5 (229.9–284.4; 141–400)	244.5 (221.8–273.8; 144–355)	245.5 (221.3–271.8; 165–365)
APRI score* *(IQR; range)*	0.29 (0.26–0.37; 0.17–0.56)	0.29 (0.26–0.37; 0.15–0.68)	0.26 (0.22–0.31; 0.13–0.45)	0.26 (0.24–0.34; 0.18–0.58)	0.31 (0.25–0.35; 0.10–0.55)
FIB-4 score* *(IQR; range)*	0.66 (0.59–0.77; 0.4–1.09)	0.79 (0.72–0.93; 0.42–1.41)	0.82 (0.73–1.09; 0.42–1.93)	1.12 (0.99–1.36; 0.61–2.01)	1.38 (1.05–1.56; 0.52–2.38)

*Data are presented as median ⊥Data are presented as mean ± SD. SD, Standard deviation; IQR, Interquartile range (representing upper and lower bound); Range (from minimum value to maximum value). BMI, Body Mass Index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; APRI, AST to Platelet Ratio Index; FIB4, Fibrosis-4. Reference values for our laboratory tests are given in the brackets, normal values cover both genders.

https://doi.org/10.1371/journal.pone.0203486.t001

The characteristics of healthy subjects are shown in Table 1. The subjects consisted of volunteers with various occupational backgrounds recruited amongst staff, their families and social network. Volunteers were recruited into five groups by age, with 10 males and 10 females per group: 20–30, 31–40, 41–50, 51–60 and 61–70 years (Table 1). Liver disease was ruled out as far as possible by patients' history, laboratory tests and negative viral markers. In total ten subjects were excluded, weekly alcohol use extending 10 units for males and 6 units for females (n = 2), abnormal laboratory tests (n = 3) or evidence of malignancy on ultrasound examination (n = 1). Individuals with BMI >30 kg/m^2 were excluded (n = 4). Four subjects withdrew their participation consent (Fig 1). We included 100 healthy subjects in the final analysis. A random subset of subjects (n = 24) were included for assessment of interobserver variability. For analyses regarding the effect of BMI, the subjects were divided into two groups with BMI between 18.0 and 25 kg/m^2 (n = 73) and BMI between 25 and 30 kg/m^2 (n = 27), respectively.

Laboratory analyses

On the day of ultrasound and elastography, blood was sampled and biochemical analyses were performed using standard routine laboratory protocols. The tests included C-reactive protein (CRP), haemoglobin, leukocytes, platelets, creatinine, total bilirubin, albumin, international

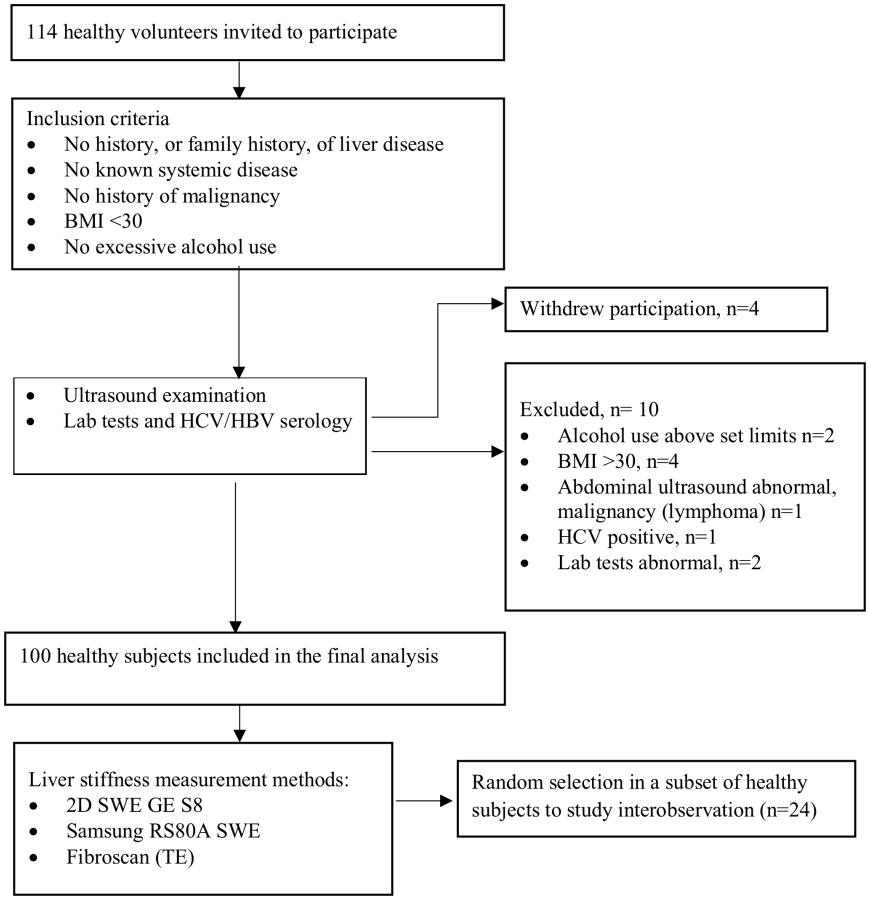


Fig 1. Method of selection of healthy subjects. Flow chart of data collection.

https://doi.org/10.1371/journal.pone.0203486.g001

normalization rate (INR), aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP) and gamma-glutamyl transferase (GGT). The laboratory analyses were performed in our hospital's laboratory, and reference values were gender specific. Three subjects had a bilirubin value outside the gender specific reference range, but none of these were excluded as the values normalized and diagnostic work-up showed no evidence of liver disease. Viral markers for hepatitis C virus (HCV) and hepatitis B virus (HBV) were also included. APRI and FIB-4 scores of fibrosis were calculated using published algorithms [5, 6].

B-mode ultrasound examination

All subjects underwent B-mode ultrasound examination of the liver, gallbladder, spleen and kidneys using a Samsung RS80A before SWE examination. All examinations were conducted after a minimum of four hours of fasting, using a standardized scanning protocol and by a

single operator (AM) with >3 years' experience in abdominal ultrasound. Small hepatic capillary haemangiomas were found in 9 subjects; none of these subjects were excluded as the lesions were confirmed by contrast enhanced ultrasound and were considered small and unlikely to influence the liver stiffness.

Elastography methods and SWE examination

Three shear wave elastography (SWE) methods were assessed in the study and are listed below in chronological order of assessment. The scanner settings were standardized for all systems. All measurements were performed by a single operator (A.M.). In order to evaluate interobserver variation, a subset of subjects (n = 24) were examined by two independent observers (A. M. and A.B.M.). Observer A (A.M.) and B (A.B.M.) had >3 and 1 years' experience in ultrasound liver scanning and elastography, respectively. The subjects were fasting (minimum 4 hours) and examined in the supine position with their right arm abducted. All SWE measurements were obtained in the right liver lobe through an intercostal space in relaxed mid-breath hold with minimal transducer pressure being applied; for Samsung RS80A and GE S8 the measurements were acquired in the right lobe about 2 cm beneath the Glisson capsule, perpendicular to the capsule, avoiding large liver vessels, bile ducts and rib shadow in B-mode. Each observer performed first 10, and then 5 separate measurements in the same area with each of the ultrasound based elastography methods. A valid LS assessment was considered as the median value and range of 10 and 5 measurements, acquired in a homogenous area (Samsung RS80A) or in a homogenous elastogram (GE S8 2D-SWE) with an interquartile range (IQR)/median <30% and a success rate (SR) ≥60%.

Samsung RS80A SWE. The Prestige ultrasound system (Samsung Medison Co. Ltd., Seoul, Korea) was applied using a CA1-7A convex array probe with a frequency of 1–7 MHz. The software version was 3.00.03.0824. The method measured the average liver elasticity within a region of interest (ROI). Within the brightness mode (B-mode) window, using default scanner settings, the ROI could be placed freely, with a fixed height of 10 mm. The width was automatically adjusted depending on the measurement depth (Fig 2). LSM was expressed in kilopascals (kPa) and meters per second (m/s).

GE Logiq S8 2D-SWE. 2D-SWE from the S8 Ultrasound scanner (GE Healthcare, Milwaukee, Wisconsin, USA), Version R4.1.2, was applied using the C1-6 convex array probe with a frequency of 1–6 MHz. Within the elastogram a circular ROI was placed, standardized to 10 mm in our study and under default scanner settings. The elastic modulus of the liver was automatically acquired by the system. The colour 2D-SWE images were captured and 2–3 elasticity frames per breath-hold (3–5 seconds) were recorded. One ROI was placed within each homogenously coloured elastogram (Fig 3). LSM was expressed in m/s and kPa.

Transient elastography (TE). Integrated in the GE Logiq S8 ultrasound scanner, TE (Fibroscan®, EchoSens, Paris, France), was applied using the M-probe with a frequency of 3.5 MHz and used according to the manufacturer's instructions. A reliable and valid measurement acquisition was defined as SR ≥60% and IQR/median <30% [7].

Statistical analysis

The statistical analysis was performed using SPSS, Version 24.0, IBM Statistics (Armon, New York, NY, USA). We used descriptive statistics for demographic, clinical and laboratory characteristics. Sample size power estimation was performed using a 2-sided comparison of two-means model. Estimating a difference in means of 4.0–4.5 kPa with a standard deviation of 0.5 kPa between the methods, 80% power and type I error of 5% yielded a sample size of 16; we compared groups consisting of 20 individuals or more. Variables were tested for normal

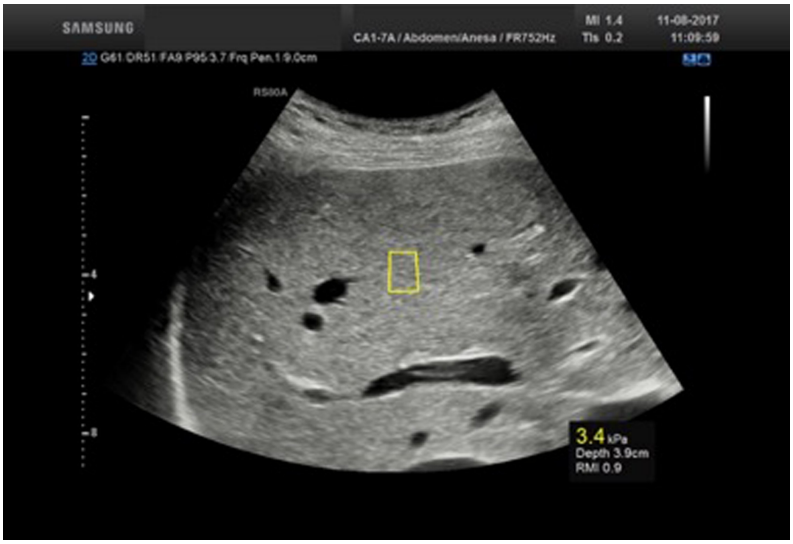


Fig 2. Samsung RS80A SWE performed on a healthy liver. The figure illustrates Samsung RS80A SWE method performed on a healthy subject. The yellow box (centre) represents the shear-wave measurement area and is expressed below the obtained elasticity measurement of 3.4 kPa.

https://doi.org/10.1371/journal.pone.0203486.g002

distribution by calculation and graphics using the Shapiro Wilk test and Q-Q Plot. Differences between numerical variables with a normal distribution were assessed with parametric tests (t-test), and those with a non-normal distribution, with nonparametric tests (Mann-Whitney). P-values of < 0.05 were considered significant. Data are presented as mean (SD) when the data were normally distributed. We calculated the coefficient of variation (CV) of the intraobserver variability. Inter-class correlation coefficients (ICC) were calculated to present the

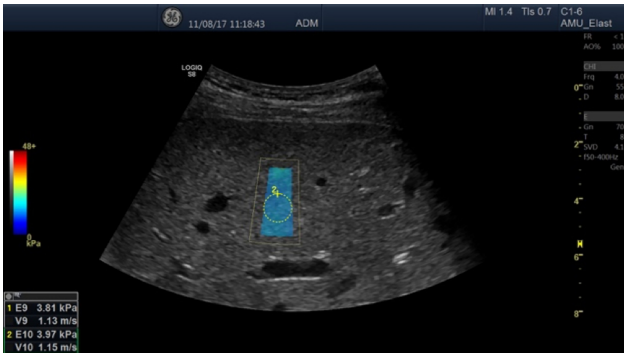


Fig 3. 2D-SWE by GE S8 performed on a healthy liver. The figure illustrates the method of 2D-SWE by GE performed on a healthy subject. The coloured box (centre) represents the elastogram, and the circle represents the ROI where the elastic modulus (LSM, liver stiffness measurement) of the liver is acquired. The blue colour indicates soft liver tissue, as semi-quantitatively presented by the colour scale to the left.

https://doi.org/10.1371/journal.pone.0203486.g003

interobserver reliability. Inter-observer agreement was classified as poor (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80) and excellent (0.81–1.00) [8]. Correlations were tested by Pearson correlation coefficient. Limits of agreement were assessed according to Bland and Altman to discover differences between individual measurements and to detect possible biases for each method [9, 10]. IQR/Median (%) was calculated for both observers individually as well as together, and for all systems [3, 11].

Results

A total of 100 healthy subjects were included. LSM was obtained by three different elastography methods (Samsung RS80A, GE S8 2D-SWE and TE). The feasibility of the methods was excellent and successful measurements were obtained in all 100 subjects by all three methods. The characteristics of the healthy subjects are shown in Table 1.

Measurement variability for the different elastography methods

The overall mean value of the median liver stiffness (LSM-mean) in 100 healthy subjects ranged from 2–6.8 kPa (Table 2).

LSM-mean by GE S8 2D-SWE was significantly higher compared to LSM-mean by TE (4.5 ± 0.8 kPa vs. 4.2 ± 1.1, respectively, $p<0.001$) and Samsung RS80A (4.1 ±0.8 kPa, $p>0.001$), whereas no significant difference was seen between Samsung RS80A and TE ($p = 0.11$) (Fig 4).

The coefficient of variation (CV) ranged from 0.03–0.28 for all systems (0.03–0.28 for Samsung RS80A SWE, 0.05–0.28 for GE S8 2D-SWE and 0.04–0.20 for TE). TE had a significantly higher CV than GE S8 2D-SWE ($p<0.001$) and Samsung RS80A ($p = 0.005$). Furthermore, between GE S8 2D-SWE and Samsung RS80A we found a small, but significant difference in CV ($p = 0.03$). Interobserver analysis was performed on 24 randomly selected subjects. No significant differences in LSM-mean between two independent observers (A.M. and A.B.M) was demonstrated for Samsung RS80A SWE (4.4 ± 0.8kPa vs. 4.4 ± 0.8 kPa, respectively, $p = 0.42$), however, we did find a significant difference between observers for GE S8 2D-SWE (4.5 ± 0.6 kPa vs. 5.1 ± 0.7 kPa, respectively, $p = 0.009$) (Fig 5).

Interoperator reliability was good for both Samsung RS80A SWE and GE S8 2D-SWE. Pearson's correlation coefficient between observers was significant for both methods ($r = 0.74$, $p<0.001$ vs. $r = 0.65$, $p<0.001$, respectively) (Fig 6).

The intraclass correlation coefficient (ICC) was good for both Samsung RS80A and GE S8 2D-SWE (ICC = 0.85 vs. ICC = 0.78, respectively). There was no indication of observer bias for either GE S8 2D-SWE or Samsung RS80A SWE as illustrated by limits of agreement

Table 2. Liver stiffness values (kPa) for the different methods.

Method	2D-SWE GE S8	Samsung RS80A SWE	Fibroscan (TE)
Mean LS, kPa	4.5	4.1	4.2
Range	2.9–6.3	2.5–6.8	2.0–6.4
SD	0.8	0.8	1.1
95% CI	4.37–4.67	3.91–4.23	4.0–4.5
CV	0.17	0.21	0.27
CV [range]	0.05–0.28	0.03–0.28	0.04–0.20

Liver stiffness (LS) values (kPa) for 2D-SWE GE, Samsung RS80A and TE. Data are presented as mean with 95% Confidence Interval (CI) and standard deviation (SD), coefficient of variation (CV) and the range of CV for the respective methods.

https://doi.org/10.1371/journal.pone.0203486.t002

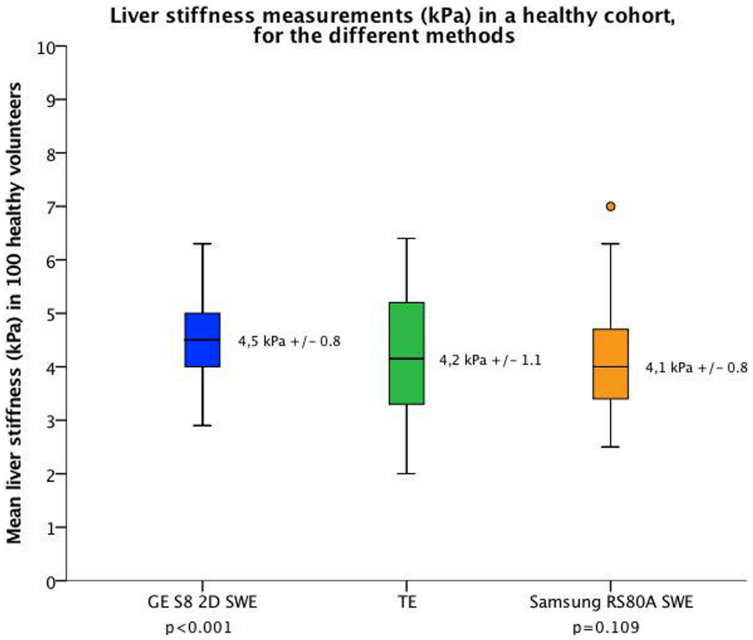


Fig 4. Liver stiffness (kPa) in a healthy cohort for the different methods. This boxplot figure displays the median and the interquartile range for LSM for each system. Whiskers represent the 90% percentile of the measured liver stiffness. The height of the box represents the variability in LSM between the healthy study subjects for each of the following three systems: blue, GE S8 2D-SWE; green, Transient Elastography (TE, Fibroscan) and orange, Samsung RS80A SWE. P-values indicate if there is a significant difference between the novel systems (Samsung RS80A or GE S8) and TE.

https://doi.org/10.1371/journal.pone.0203486.g004

analysis; however, GE S8 2D-SWE showed a trend of a slightly larger deviation of the mean than Samsung RS80A SWE (Figs 7 and 8).

Difference in liver elasticity by gender, age and BMI

LSM-mean was significantly higher in males compared to females for TE (4.5 ± 1.0 kPa vs. 3.9 ± 1.1 kPa, respectively, p = 0.006) and GE S8 2D-SWE (4.7 ± 0.7 kPa vs. 4.3 ± 0.7 kPa, respectively, p = 0.006). A similar trend for Samsung RS80A SWE did not reach significance (4.2 ± 0.7 kPa vs. 3.9 ± 0.9 kPa, respectively, p = 0.063) (Fig 9, Table 3). In a post hoc analysis of subjects consuming 5 alcohol units or less per week (n = 69) we found significant differences in LSM-mean between males (n = 33) and females (n = 36) for all systems; for GE S8 (4.8 ± 0.7 kPa vs. 4.2 ± 0.8 kPa, p = 0.003), TE (4.7 ± 1.0 kPa vs. 3.8 ± 1.1 kPa, p = 0.001) and Samsung (4.2 ± 0.7 kPa vs. 3.8 ± 0.7 kPa, p = 0.006).

None of the systems demonstrated statistical significant difference in LSM across age groups (Table 4).

LSM-mean showed no significant difference between subjects with BMI 25–30 kg/m^2 and BMI 18.0–25.0 kg/m^2 for any individual system (GE S8 2D-SWE (4.5 ± 0.8 kPa vs. 4.4 ± 0.8 kPa, respectively, p = 0.49), TE (4.3 ± 1.1 kPa vs. 4.1 ± 1.1 kPa, respectively, p = 0.36), Samsung RS80A SWE (4.1 ± 0.9 kPa vs. 3.9 ± 0.6 kPa, respectively, p = 0.28) or all systems combined (4.1 ± 0.9 kPa vs. 4.3 ± 0.9 kPa, p = 0.128) (Fig 10).

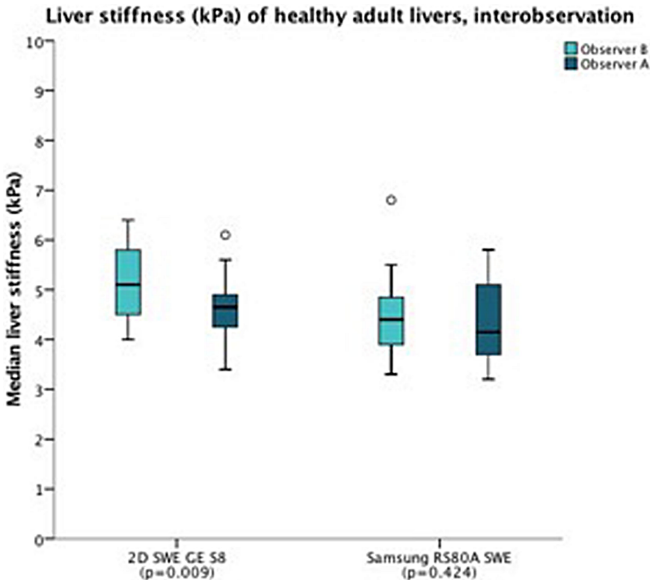


Fig 5. Liver stiffness (kPa) of healthy adult livers, interobservation. The boxplot shows interobservation between observer A (dark blue) and B (light blue). The horizontal axis represents the systems Samsung RS80A SWE and GE S8 2D-SWE and the p-value is given. For boxplot interpretation, we refer to Fig 4.

https://doi.org/10.1371/journal.pone.0203486.g005

Difference in variability and reproducibility of LSM when using 5 measurements instead of 10

There was no significant difference in LSM-mean using 5 or 10 measurements for the ultrasound based SWE methods (GE S8 2D-SWE 4.4 ± 0.66 kPa vs. 4.5 ± 0.76 kPa, respectively, p = 0.05; and Samsung RS80A SWE: 4.1 ± 0.86 kPa vs. 4.1 ± 0.81 kPa, respectively, p = 0.08) (Fig 11).

Discussion

To the best of our knowledge, this is the first study to investigate normal LSM values by two new elastography techniques (Samsung RS80A SWE and GE S8 2D-SWE) compared head-to-head and with TE as reference, in a healthy cohort. The comprehensive exclusion of liver disease as well as the direct comparison to TE as a reference standard represent strengths of our study. Data regarding normal values in liver stiffness for each of the new elastography techniques are needed to establish standardized reference bases [12], which are pivotal for clinical implementation of novel elastography systems as reliable methods for diagnostics, staging and assessment of disease progression in chronic liver diseases.

We found a mean LSM of 4.3 kPa ± 0.8 across the two methods, confirming that on average LSM for 2D-SWE GE S8 (4.5 ± 0.8 kPa) and Samsung RS80A (4.1 ±0.8 kPa) were in the same range as other elastography systems [13–15]. In this head-to-head comparison between elastography systems, 2D-SWE GE S8 demonstrated slightly higher values than both Samsung RS80A and TE, while measurements made with Samsung RS80A were not significantly different from the reference method. There was also a small, but significant difference in the

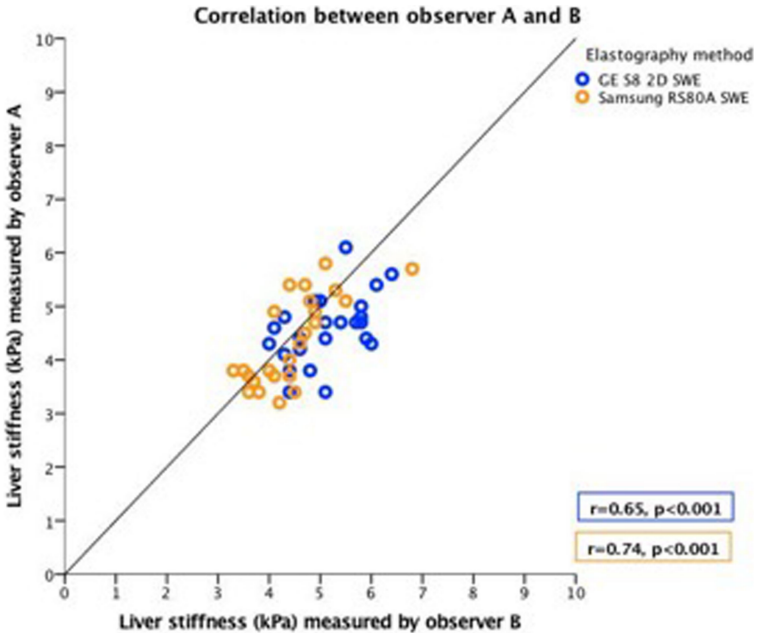


Fig 6. Correlation between observer A and B. The horizontal and vertical axes represent measurements by observer B and A, respectively. The unit measured is kilopascals (kPa). The line in the graph represents the line of unity. The Pearson correlation coefficient (r) and significance (p) for each system is given in the lower right corner. For colour representation, we refer to Fig 4.

https://doi.org/10.1371/journal.pone.0203486.g006

coefficient of variation between the two novel methods. Previous studies have shown similar results for 2D-SWE from Aixplorer [16, 17]. Our results confirm that LSM levels are significantly different depending on the method applied. We found differences both between SWE methods and TE as the reference method, and between the two different SWE systems. In clinical practice LSM greater than 6.8–7.6 kPa indicates a higher probability of significant fibrosis ($F \geq 2$) on liver biopsy; however, the EASL clinical practice guidelines state that cut-off values vary considerably and ranging 5.2–9.6 kPa for different systems. For predicting cirrhosis (F4), the optimal cut-off ranges from 11 to 15 kPa [18]. In that context, a net difference of 0.3 kPa is probably too small to represent a clinically significant difference, however, it underscores the need to compare methods also in fibrotic livers, where the differences may be more expressed, as we know that variability increases with higher liver stiffness [19].

Both methods showed good interobserver reliability and intraclass correlation. Similarly, previous studies have shown excellent interobserver agreement ranging from r = 0.80–0.97 for pSWE methods [20, 21]. However, we observed a significant difference in LS measurements between the two observers for 2D-SWE GE S8 but not for Samsung RS80A. One possible explanation for this discrepancy may be that 2D-SWE allows the examiner to place the measurement ROI within the elastogram and avoid incongruent signals, while Samsung SWE performs several automated SWE speed measurements within the elasticity measurement area without visualisation of the stiffness. The 2D-SWE method is slightly more user dependent and may acquire a longer learning curve. Previous studies on 2D-SWE measurements of liver

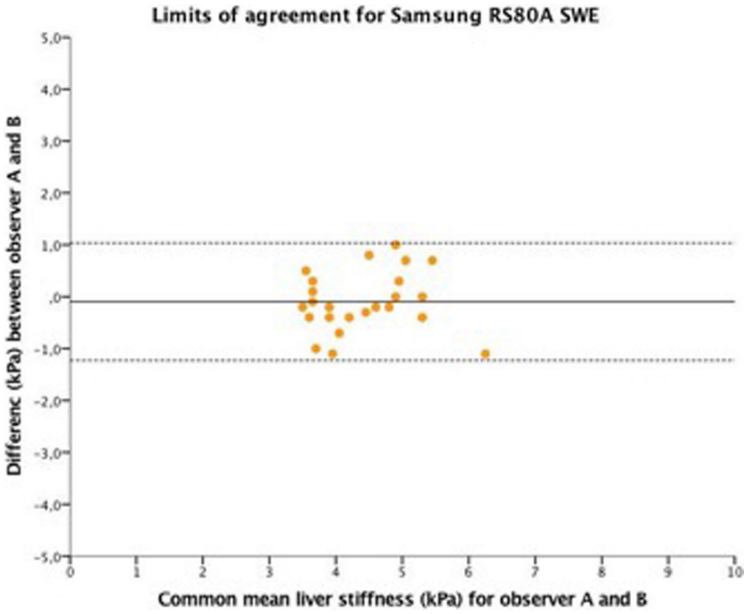


Fig 7. Limits of agreement for Samsung RS80A SWE. The figure presents the limits of agreement for Samsung RS80A. The horizontal axis represents the common mean value of all measurements in both observers for, while the vertical axis represents the difference between individual measurements and this common mean (kPa), displaying the variability of measurements. The black line within each system represents the common mean value, the dotted lines represent the 95% confidence intervals. A mean value close to 0 on the vertical axis means that the two observers apply the measurement scale without bias.

https://doi.org/10.1371/journal.pone.0203486.g007

elasticity have demonstrated a learning curve for this method, but with similar reproducibility [16, 22]. Evaluating the intra- and interobserver agreement for 2D-SWE from GE and SWE from Samsung, we demonstrated a good interobserver and better intraobserver agreement for both systems compared to the results reported for Aixplorer 2D-SWE (Figs 7 and 8).

We found a significantly higher LSM in adult male subjects for TE and GE S8 2D-SWE, whereas a similar trend for Samsung RS80A SWE did not reach significance. This is an important finding, indicating that it may be necessary to define separate cut-off values for normal liver and possibly also for levels of liver fibrosis for male and female patients. Previous studies have shown inconsistent results regarding the effect of gender on LSM [23, 24]. Using pSWE, Ling et al. demonstrated that males had 8% higher LSM than females; however, the study had more than twice as many female participants compared to male participants [14]. In contrast, using ARFI, one study found no significant difference between genders in 137 subjects [25] in line with our results for Samsung RS80A. Two studies conducting reliable LSM with TE in 1190 subjects over 45 years, and in 746 healthy subjects, found that male gender was associated with higher liver stiffness [26, 27] and our results for TE confirmed this. Using Aixplorer 2D-SWE from Supersonic Imagine, it has been suggested that males may have higher LSM than females [13]. A study of LSM in healthy children, using the same system, did not demonstrate significant difference between genders [28]. The lack of significant gender difference for LSM in healthy liver tissue for Samsung SWE in the present study may be due to different

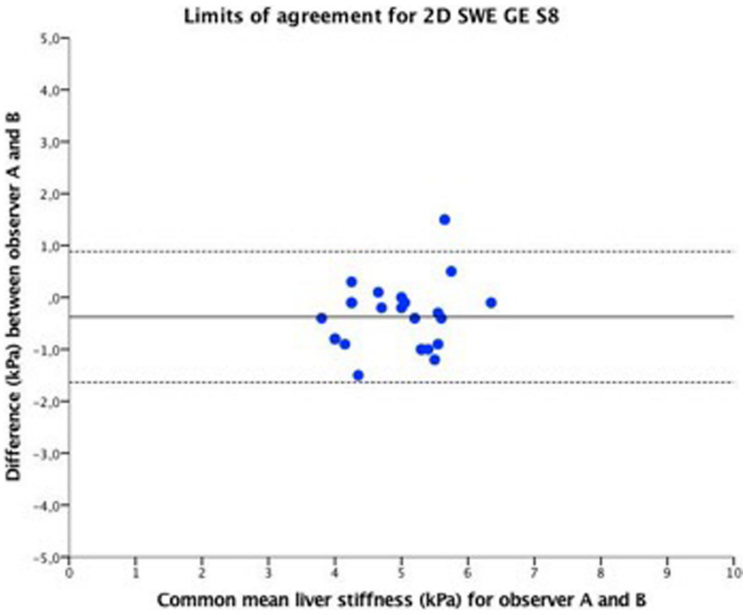


Fig 8. Limits of agreement for 2D-SWE GE S8. Limits of agreement for 2D-SWE GE S8, for legend we refer to Fig 7.

https://doi.org/10.1371/journal.pone.0203486.g008

technology and signal processing compared to the two other scanners. Despite that we found a significant difference between genders for all systems in our post hoc analysis (n = 69), the study may be underpowered considering the observed SD of 0.8 kPa for the Samsung SWE compared to our power estimation anticipating 0.5 kPa as SD. Furthermore, different hormone levels have been proposed as an explanation of LSM differences between genders, and should be investigated further in *in vivo* studies [29].

In our study, LSMs were not significantly affected by age or BMI. Multiple studies have addressed age as a variable of influencing liver stiffness in normal subjects, and the results have been inconsistent, reporting no difference across age groups [14], higher LSM in older [26] or younger [27] age. We did not demonstrate significant differences in LSM between the five age groups for any of the methods. Possibly, analyses of the effect of age on LSM is confounded by other factors such as steatosis and heart failure that are more prevalent in older populations. One study investigating GE E9 Logic 2D-SWE in healthy subjects reported an LSM-mean of 5.1 kPa ± 1.3, with higher LSM values compared to TE, similar to our findings [17]. In contrast to our results, they reported that age over 40 years was associated with higher LSM, but did not find significant difference in LSM between genders (21 males and 58 females). In the present study, we included only healthy volunteers, carefully interviewed all subjects regarding alcohol consumption, and performed full biochemical analyses and B-mode ultrasound examination of all in contrast to some other studies [12], and in our view, less strict inclusion criteria and missing data regarding liver enzymes in the other study may contribute to these differences. Higher LSM values have been reported in healthy subjects with low BMI (<18.5 kg/m^2) as well as in obese subjects compared with normal-weight subjects [30]. We did not demonstrate a difference in LSM between subjects with BMI 18.0–25.0 kg/m^2 compared to BMI 25–30 kg/m^2;

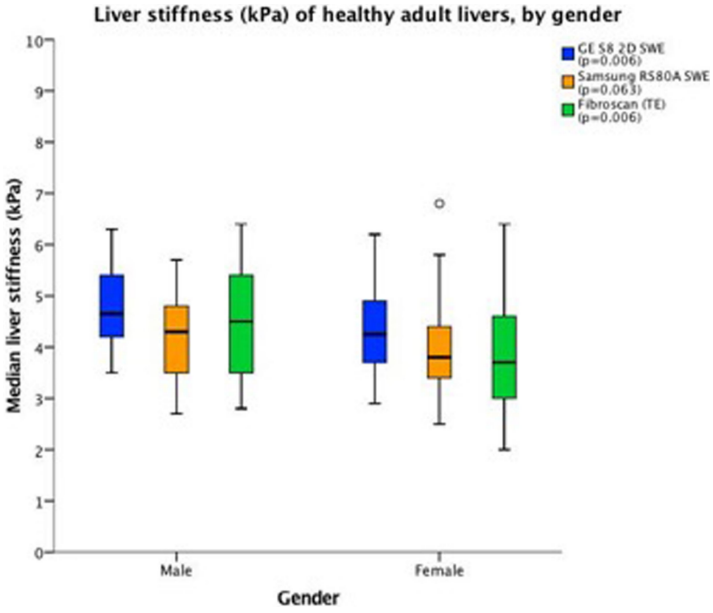


Fig 9. Liver stiffness (kPa) of healthy adult livers, by gender. The boxplot shows the liver stiffness by gender. The horizontal axis represents the gender; males and females. The colour interpretation for each system and the level of significance is given in the upper right corner. For legend interpretation, we refer to Fig 4.

https://doi.org/10.1371/journal.pone.0203486.g009

however, obese patients with BMI>30 were not included in this study. Normal values for LS in underweight and obese subjects, as well as technical feasibility of Samsung RS80A SWE and GE S8 2D-SWE, should be further investigated and established.

There is an ongoing discussion of the minimum number of measurements needed when acquiring LSM with SWE. The EFSUMB guidelines recommend at least 10 measurements for pSWE and TE, and a minimum of 3 measurements when using 2D-SWE, to obtain consistent results [31, 32]. One study reported excellent intraobserver reproducibility based on 6

Table 3. Liver stiffness values (kPa) of healthy adult livers, by gender.

Gender	Female (n = 50)	Male (n = 50)	p-value
GE S8 2D-SWE			
Mean ± SD (kPa)	4.3 ± 0.7	4.7 ± 0.7	p = 0.006
95% CI	4.1–4.5	4.5–4.9	
Samsung RS80A SWE			
Mean ± SD (kPa)	3.9 ± 0.9	4.2 ± 0.7	p = 0.063
95% CI	3.7–4.2	4.0–4.4	
Fibroscan (TE)			
Mean ± SD (kPa)	3.9 ± 1.1	4.5 ± 1.0	p = 0.006
95% CI	3.6–4.2	4.2–4.8	

Liver stiffness values (kPa) for 2D-SWE GE, Samsung RS80A and TE, by gender. Data are presented as mean ± standard deviation with 95% Confidence interval (CI).

https://doi.org/10.1371/journal.pone.0203486.t003

Table 4. Liver stiffness values (kPa) of healthy adult livers, by age group.

Age group	20–30 (n = 20)	31–40 (n = 20)	41–50 (n = 20)	51–60 (n = 20)	61–70 (n = 20)	p-value
GE S8 2D-SWE						
Mean ± SD (kPa)	4.5 ± 0.9	4.7 ± 0.8	4.5 ± 0.7	4.4 ± 0.7	4.5 ± 0.7	p = 0.843
95% CI	4.1–4.9	4.3–5.1	4.1–4.8	4.1–4.8	4.2–4.9	
Samsung RS80A SWE						
Mean ± SD (kPa)	4.3 ± 0.9	4.2 ± 0.8	4.0 ± 0.8	4.1 ± 0.6	3.9 ± 1.0	p = 0.630
95% CI	3.9–4.7	3.8–4.5	3.6–4.4	3.8–4.3	3.4–4.4	
Fibroscan (TE)						
Mean ± SD (kPa)	4.4 ± 1.1	4.3 ± 1.3	4.2 ± 1.0	4.2 ± 1.1	4.1 ± 1.1	p = 0.630
95% CI	3.9–4.9	3.7–5	3.7–4.7	3.7–4.7	3.6–4.7	

Liver stiffness values (kPa) for 2D-SWE GE, Samsung RS80A and TE, by age group. Data are presented as mean ± standard deviation with 95% Confidence interval (CI).

https://doi.org/10.1371/journal.pone.0203486.t004

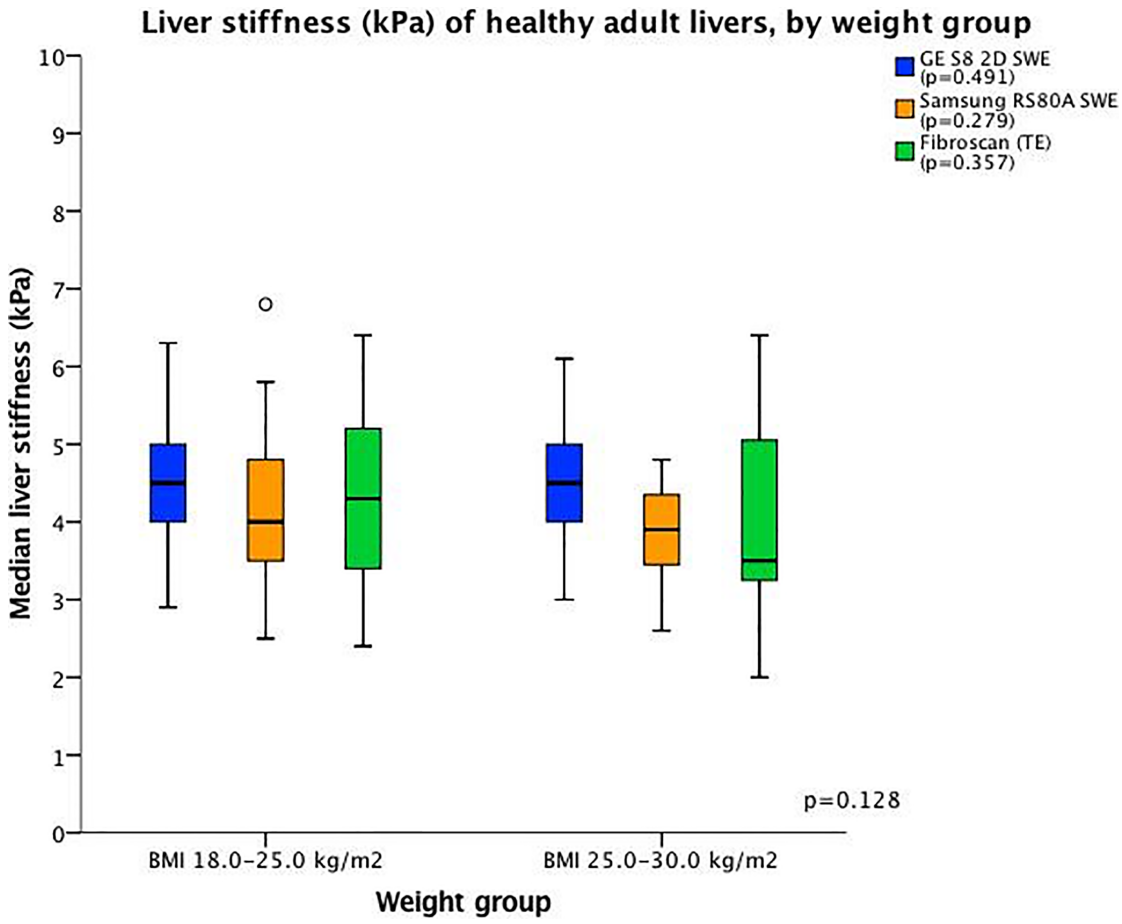


Fig 10. Liver stiffness (kPa) of healthy adult livers, by weight group. The boxplot shows the liver stiffness by weight group. The horizontal axis represents the BMI group. The colour interpretation for each system and the level of significance is given in the upper right corner. For boxplot interpretation, we refer to Fig 4.

https://doi.org/10.1371/journal.pone.0203486.g010

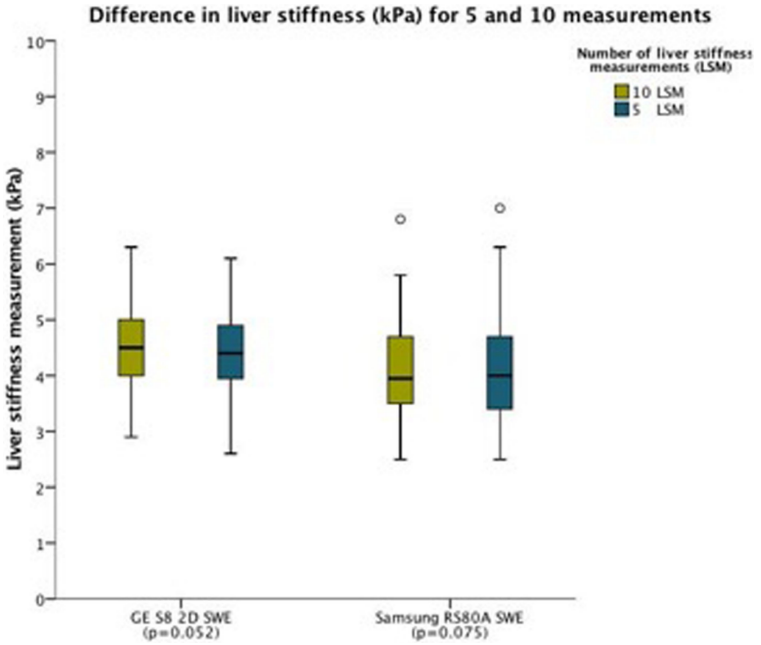


Fig 11. Difference in liver stiffness (kPa) for 5 and 10 measurements. The boxplots show difference in liver stiffness for 5 and 10 measurements. The horizontal axis represents the systems, and the vertical axis the liver stiffness measured. The colour interpretation for 5 and 10 measurements (green and blue, respectively) is given in the upper right corner. For boxplot interpretation, we refer to Fig 4.

https://doi.org/10.1371/journal.pone.0203486.g011

measurements, concluding that the optimal minimum number of measurements with 2D-SWE was 6 [15]. For pSWE (ARFI), one study concluded that 10 measurements instead of 5 should be performed to obtain a reliable estimation [33]. To the best of our knowledge there are no studies that have directly investigated the difference in mean LS between 5 and 10 separate measurements for several ultrasound SWE methods. Our results did not show significant difference in median LS for 5 versus 10 measurements. This suggests that a reliable median LSM can be obtained with fewer measurements than ten for both 2D-SWE GE S8 and Samsung RS80A in healthy livers. This is important as it indicates that adequate measurements can be made by fewer repetitions and in less time, however our results in healthy livers may not apply in patients with higher degree of liver fibrosis where measurement variability may be higher.

The main limitation of the study is the lack of liver biopsies as a reference method, which is not ethically feasible in a healthy group. Our study design included 100 healthy participants, excluding unknown liver disease by imaging, blood tests and anamnesis.

Conclusion

All methods were successfully applied in our cohort of 100 healthy subjects. The mean of median LSM for the two new elastography methods (GE S8 and Samsung RS80A) showed a slight difference. Our study shows a significantly higher liver stiffness in males compared to

females, however we found no significant difference in LS between BMI groups 18–30 kg/m^2 or between the age groups 20–70 years. Furthermore, our findings indicate that five acquisitions are sufficient to obtain a reliable LSM using Samsung RS80A or GE S8 2D-SWE in healthy subjects.

Acknowledgments

We express our gratitude to Samsung Medison and GE Healthcare for providing an ultrasound scanner and elastography software. We also thank the application specialists for valuable information of the respective elastography systems. The two companies had no role in the study design, data acquisition or analysis.

Author Contributions

Conceptualization: Anesa Mulabecirovic, Mette Vesterhus, Roald Flesland Havre.

Data curation: Anesa Mulabecirovic.

Formal analysis: Anesa Mulabecirovic.

Investigation: Anesa Mulabecirovic, Anders Batman Mjelle.

Methodology: Anesa Mulabecirovic, Mette Vesterhus, Roald Flesland Havre.

Project administration: Anesa Mulabecirovic.

Resources: Anesa Mulabecirovic, Roald Flesland Havre.

Software: Anesa Mulabecirovic.

Supervision: Anesa Mulabecirovic, Odd Helge Gilja, Mette Vesterhus, Roald Flesland Havre.

Validation: Anesa Mulabecirovic, Mette Vesterhus, Roald Flesland Havre.

Visualization: Anesa Mulabecirovic.

Writing – original draft: Anesa Mulabecirovic.

Writing – review & editing: Anesa Mulabecirovic, Anders Batman Mjelle, Odd Helge Gilja, Mette Vesterhus, Roald Flesland Havre.

References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380(9859):2095–128. https://doi.org/10.1016/S0140-6736(12)61728-0 PMID: 23245604.
2. Dultz G, Zeuzem S. Hepatitis C Virus: A European Perspective. Gastroenterol Clin North Am. 2015; 44(4):807–24. https://doi.org/10.1016/j.gtc.2015.07.008 PMID: 26600221.
3. Castera L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology. 2010; 51(3):828–35. https://doi.org/10.1002/hep.23425 PMID: 20063276.
4. Schwabl P, Bota S, Salzl P, Mandorfer M, Payer BA, Ferlitsch A, et al. New reliability criteria for transient elastography increase the number of accurate measurements for screening of cirrhosis and portal hypertension. Liver Int. 2015; 35(2):381–90. https://doi.org/10.1111/liv.12623 PMID: 24953516.
5. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology. 2003; 38(2):518–26. https://doi.org/10.1053/jhep.2003.50346 PMID: 12883497.
6. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006; 43(6):1317–25. https://doi.org/10.1002/hep.21178 PMID: 16729309.

7. Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, Lebail B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology. 2013; 57(3):1182–91. https://doi.org/10.1002/hep.25993 PMID: 22899556.
8. DG A. Practical statistics for medical research. London: Chapman & Hall; (1997).
9. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat. 2007; 17(4):571–82. https://doi.org/10.1080/10543400701329422 PMID: 17613642.
10. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999; 8(2):135–60. https://doi.org/10.1177/096228029900800204 PMID: 10501650.
11. Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, Mal F, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound in medicine & biology. 2003; 29 (12):1705–13. PMID: 14698338.
12. Dong Y, Sirli R, Ferraioli G, Sporea I, Chiorean L, Cui X, et al. Shear wave elastography of the liver—review on normal values. Z Gastroenterol. 2017; 55(2):153–66. https://doi.org/10.1055/s-0042-117226 PMID: 28192849.
13. 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 in medicine & biology. 2014; 40(11):2549–55. https://doi.org/10.1016/j.ultrasmedbio.2014.05.008 PMID: 25282481.
14. Ling W, Lu Q, Quan J, Ma L, Luo Y. Assessment of impact factors on shear wave based liver stiffness measurement. European journal of radiology. 2013; 82(2):335–41. https://doi.org/10.1016/j.ejrad.2012.10.004 PMID: 23116805.
15. Yoon JH, Lee JM, Han JK, Choi BI. Shear wave elastography for liver stiffness measurement in clinical sonographic examinations: evaluation of intraobserver reproducibility, technical failure, and unreliable stiffness measurements. J Ultrasound Med. 2014; 33(3):437–47. https://doi.org/10.7863/ultra.33.3.437 PMID: 24567455.
16. Ferraioli G, Tinelli C, Zicchetti M, Above E, Poma G, Di Gregorio M, et al. Reproducibility of real-time shear wave elastography in the evaluation of liver elasticity. Eur J Radiol. 2012; 81(11):3102–6. https://doi.org/10.1016/j.ejrad.2012.05.030 PMID: 22749107.
17. Bende F, Mulabecirovic A, Sporea I, Popescu A, Sirli R, Gilja OH, et al. Assessing Liver Stiffness by 2-D Shear Wave Elastography in a Healthy Cohort. Ultrasound in medicine & biology. 2018; 44(2):332–41. https://doi.org/10.1016/j.ultrasmedbio.2017.10.013 PMID: 29249457.
18. European Association for Study of L, Asociacion Latinoamericana para el Estudio del H. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol. 2015; 63(1):237–64. https://doi.org/10.1016/j.jhep.2015.04.006 PMID: 25911335.
19. 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(1):e0189671. https://doi.org/10.1371/journal.pone.0189671 PMID: 29293527.
20. Ferraioli G, Tinelli C, Lissandrin R, Zicchetti M, Dal Bello B, Filice G, et al. Point shear wave elastography method for assessing liver stiffness. World journal of gastroenterology: WJG. 2014; 20(16):4787–96. https://doi.org/10.3748/wjg.v20.i16.4787 PMID: 24782633; PubMed Central PMCID: PMCPMC4000517.
21. Guzman-Aroca F, Reus M, Berna-Serna JD, Serrano L, Serrano C, Gilabert A, et al. 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(7):975–9. PMID: 21705730.
22. Gradinaru-Tascau O, Sporea I, Bota S, Jurchis A, Popescu A, Popescu M, et al. Does experience play a role in the ability to perform liver stiffness measurements by means of supersonic shear imaging (SSI)? Medical ultrasonography. 2013; 15(3):180–3. PMID: 23979612.
23. Kim BK, Kim SU, Choi GH, Han WK, Park MS, Kim EH, et al. "Normal" liver stiffness values differ between men and women: a prospective study for healthy living liver and kidney donors in a native Korean population. J Gastroenterol Hepatol. 2012; 27(4):781–8. https://doi.org/10.1111/j.1440-1746.2011.06962.x PMID: 22098121.
24. Popescu A, Sporea I, Sirli R, Bota S, Focsa M, Danila M, et al. The mean values of liver stiffness assessed by Acoustic Radiation Force Impulse elastography in normal subjects. Medical ultrasonography. 2011; 13(1):33–7. PMID: 21390341.
25. Madhok R, Tapasvi C, Prasad U, Gupta AK, Aggarwal A. Acoustic radiation force impulse imaging of the liver: measurement of the normal mean values of the shearing wave velocity in a healthy liver. J Clin Diagn Res. 2013; 7(1):39–42. https://doi.org/10.7860/JCDR/2012/5070.2665 PMID: 23450092; PubMed Central PMCID: PMCPMC3576746.

26. Roulot D, Costes JL, Buyck JF, Warzocha U, Gambier N, Czernichow S, et al. Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. Gut. 2011; 60(7):977–84. https://doi.org/10.1136/gut.2010.221382 PMID: 21068129.

27. Colombo S, Belloli L, Zaccanelli M, Badia E, Jamoletti C, Buonocore M, et al. Normal liver stiffness and its determinants in healthy blood donors. Dig Liver Dis. 2011; 43(3):231–6. https://doi.org/10.1016/j.dld.2010.07.008 PMID: 20817625.

28. Franchi-Abella S, Corno L, Gonzales E, Antoni G, Fabre M, Ducot B, et al. 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. 2016; 278(2):554–62. https://doi.org/10.1148/radiol.2015142815 PMID: 26305193.

29. Yasuda M, Shimizu I, Shiba M, Ito S. Suppressive effects of estradiol on dimethylnitrosamine-induced fibrosis of the liver in rats. Hepatology. 1999; 29(3):719–27. https://doi.org/10.1002/hep.510290307 PMID: 10051473.

30. Das K, Sarkar R, Ahmed SM, Mridha AR, Mukherjee PS, Das K, et al. "Normal" liver stiffness measure (LSM) values are higher in both lean and obese individuals: a population-based study from a developing country. Hepatology. 2012; 55(2):584–93. https://doi.org/10.1002/hep.24694 PMID: 21952989.

31. Dietrich CF, Bamber J, Berzigotti A, Bota S, Cantisani V, Castera L, et al. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). Ultraschall in der Medizin. 2017; 38(4):e16–e47. https://doi.org/10.1055/s-0043-103952 PMID: 28407655

32. Sporea I, Bota S, Jurchis A, Sirli R, Gradinaru-Tascau O, Popescu A, et al. Acoustic radiation force impulse and supersonic shear imaging versus transient elastography for liver fibrosis assessment. Ultrasound in medicine & biology. 2013; 39(11):1933–41. https://doi.org/10.1016/j.ultrasmedbio.2013.05.003 PMID: 23932281.

33. Karlas T, Pfrepper C, Wiegand J, Wittekind C, Neuschulz M, Mossner J, et al. Acoustic radiation force impulse imaging (ARFI) for non-invasive detection of liver fibrosis: examination standards and evaluation of interlobe differences in healthy subjects and chronic liver disease. Scandinavian journal of gastroenterology. 2011; 46(12):1458–67. https://doi.org/10.3109/00365521.2011.610004 PMID: 21916815.

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 $\times$ 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 ($\rho = 0.70, 0.72, 0.65$ 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 ($\rho = 0.81, 0.78$ 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 ($\rho = 0.53$, $p < 0.001$) (Fig. 2) in linear regression, with higher ΔLSM in the obese. When patients were classified as either obese (BMI ≥30)

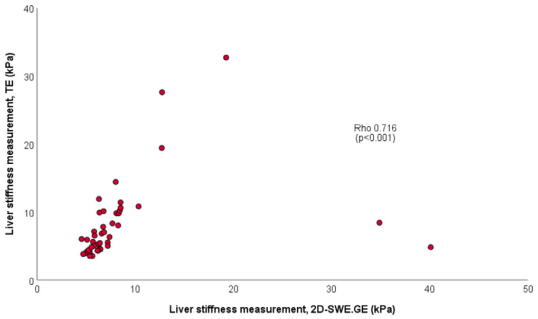


Fig. 1. Liver stiffness measurements (LSMs) for 2-D shear wave elastography (2-D-SWE.GE) plotted against those for transient elastography (TE). LSMs by 2-D-SWE.GE were higher than LSMs by TE in patients with a relatively low average LSM value, while TE values were higher than 2-D-SWE.GE values in patients 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

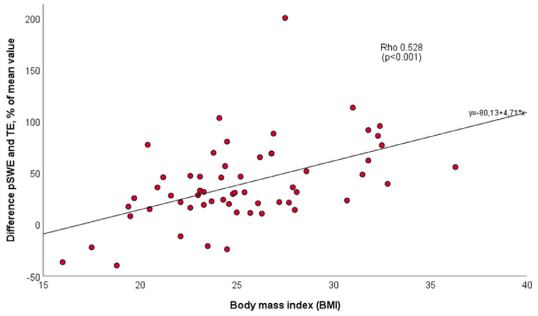


Fig. 2. Scatterplot of difference in liver stiffness values between point shear wave elastography and transient elastography in percent (y-axis) and body mass index (BMI, x-axis), ρ = 0.528, $p < 0.001$. The intersystem difference increases with increasing BMI. Looking at the systems separately, there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

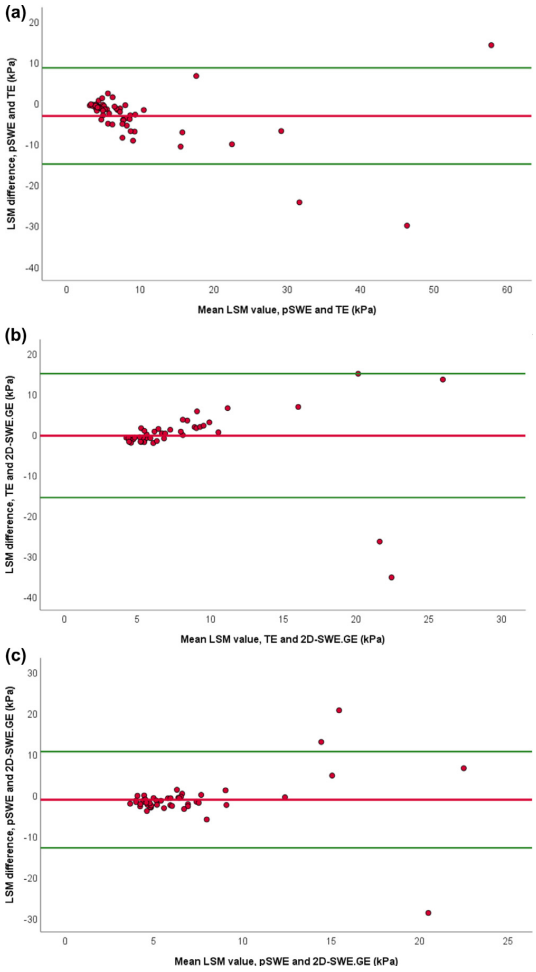


Fig. 3. Bland–Altman plots comparing scanner systems, with average liver stiffness measurements (LSMs) on the x-axis and intersystem differences on the y-axis, for (a) point shear wave elastography (pSWE) and transient elastography (TE); (b) TE and 2-D shear wave elastography (2-D-SWE.GE); and (c) pSWE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

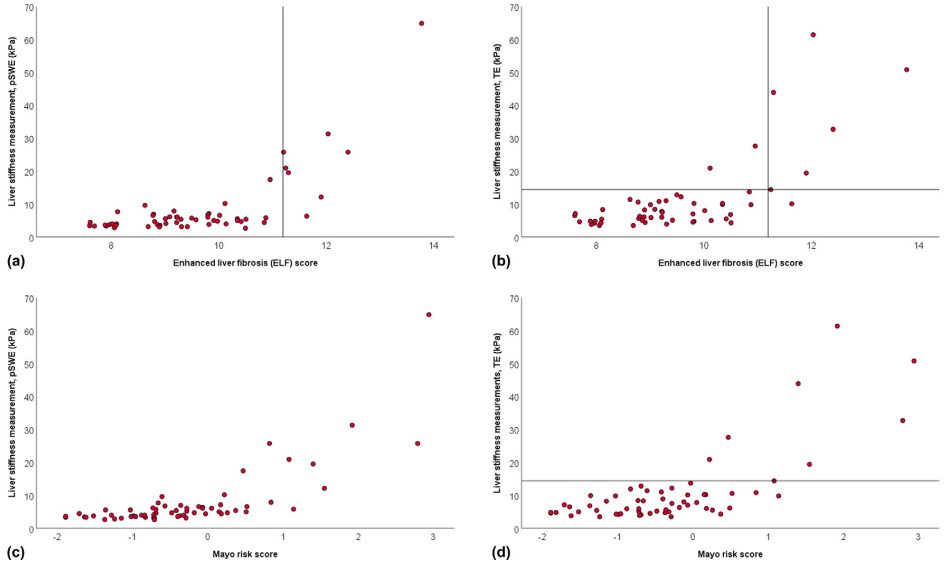


Fig. 4. Correlation between between liver stiffness and scores of fibrosis (enhanced liver fibrosis [ELF] score) or prognosis (Mayo risk score). For ELF scores, a previously published cutoff value of 11.2 is noted (vertical line), while for transient elastography (TE), a published cutoff value for cirrhosis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

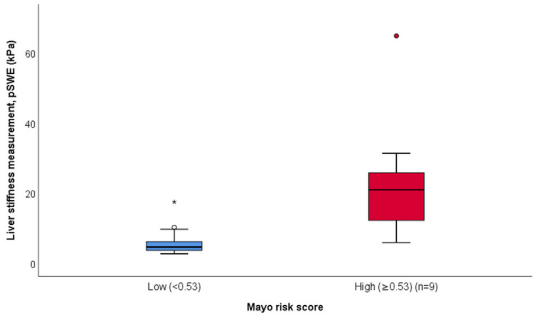


Fig. 5. Boxplot revealing a vast difference in liver stiffness measurements between patients with a high Mayo risk score and a low Mayo risk score (median values: 20.9 kPa vs. 4.5 kPa, $p < 0.001$).

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.

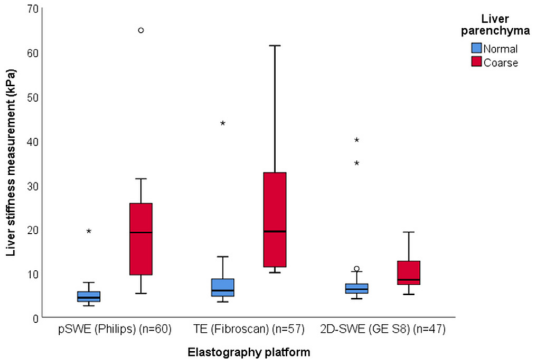


Fig. 6. Liver stiffness measurement (LSMs) for all methods, grouped by liver parenchyma on B-mode ultrasound (normal or coarse liver tissue). LSM values were significantly higher by all scanner systems in patients with coarse liver tissue, although substantially less significant by 2-D-SWE.GE ($p = 0.028$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

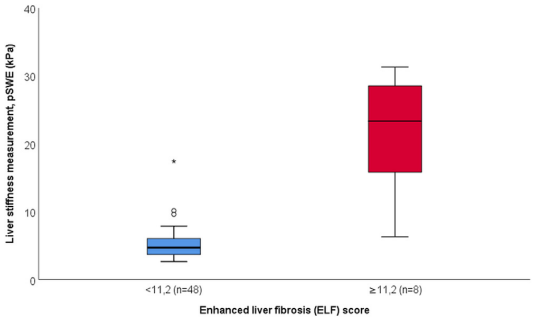


Fig. 7. Boxplot of liver stiffness measurement (LSM) by point shear wave elastography (pSWE) using a cutoff enhanced liver fibrosis (ELF) score of 11.2, discriminating high-risk (11.2) from low-risk patients. LSM values were significantly higher for patients with a high ELF score.

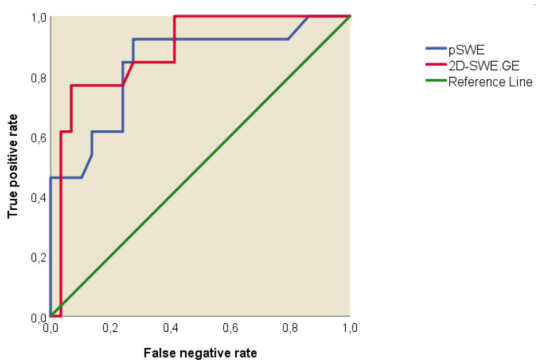


Fig. 8. Receiver operating characteristic curve of liver stiffness measurements by point shear wave elastography (pSWE) and 2-D shear wave elastography (2-D-SWE.GE) for the diagnosis of advanced fibrosis (F3−F4), using published cutoff values for transient elastography in a primary sclerosing cholangitis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

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230868072 (print)
9788230851395 (PDF)