EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version)

EFSUMB-Leitlinien und Empfehlungen zur klinischen Anwendung der Leberelastographie, Update 2017 (Langversion)

Authors
Christoph F. Dietrich1, 2, Jeffrey Bamber3, Annalisa Berzigotti4, Simona Bota5, Vito Cantisani6, Laurent Castera7, David Cosgrove8, Giovanna Ferraioli9, Mireen Friedrich-Rust10, Odd Helge Gilja11, Ruediger Stephan Goertz12, Thomas Karlas13, Robert de Knegt14, Victor de Ledinghen15, Fabio Piccaglia16, Bogdan Procopet17, Adrian Saffoliu18, Paul S. Sidhu19, Ioan Sporea20, Maja Thiele21

Affiliations
1 Department of Internal Medicine 2, Caritas Krankenhaus, Bad Mergentheim, Germany
2 Sino-German Research Center of Ultrasound in Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China
3 Joint Department of Physics and the CRUK Cancer Imaging Centre, The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust, 15 Cotswold Road, Sutton, London, SM2 5NG, UK
4 Hepatology, University Clinic for Visceral Surgery and Medicine, Inselspital, University of Bern, Bern, Switzerland
5 Department of Gastroenterology, Hepatology, Nephrology and Endocrinology, Klinikum Klagenfurt, Klagenfurt am Wörthersee, Austria
6 Department of Radiological Sciences, Policlinico Umberto I, University Sapienza, Rome, Italy
7 Department of Hepatology, Hospital Beaujon, Assistance Publique-Hôpitaux de Paris, Université Paris VII, INSERM UMR 1149, Clichy, France
8 Imaging Departments, Imperial and King’s Colleges, London, UM, UK
9 Infectious Diseases Dept, Fondazione IRCCS Policlinico San Matteo, University of Pavia, Pavia, Italy
10 Department of Internal Medicine 1, J. W.Goethe-University Hospital, Frankfurt, Germany
11 National Centre for Ultrasound in Gastroenterology, Haukeland University Hospital, Bergen and Department of Clinical Medicine, University of Bergen, Norway
12 Department of Internal Medicine 1 – Gastroenterology, Pneumology and Endocrinology, University of Erlangen-Nürnberg, Erlangen, Germany
13 Department for Internal Medicine, Division of Gastroenterology and Rheumatology, University Hospital Leipzig, Leipzig, Germany
14 Department of Gastroenterology and Hepatology, Erasmus MC University Medical Center, Rotterdam, the Netherlands
15 Non-invasive diagnosis of liver fibrosis centre, Haut-Leveque hospital, Bordeaux University Hospital, Pessac, France
16 Unit of Internal Medicine, Dpt of Medical and Surgical Sciences, University of Bologna S. Orsola-Malpighi Hospital, Bologna, Italy
17 Department of Gastroenterology, 3rd Medical Clinic, University of Medicine and Pharmacy “Iuliu Hatieganu”. Regional Institut of Gastroenterology and Hepatology “O. Fodor”, Cluj-Napoca, Romania
18 Research Center of Gastroenterology and Hepatology Craiova, University of Medicine and Pharmacy Craiova, Romania
19 King’s College London, Department of Radiology, King’s College Hospital, London
20 Department of Gastroenterology and Hepatology, University of Medicine and Pharmacy “Victor Babes”, Timisoara, Romania
21 Department of Gastroenterology and Hepatology, Odense University Hospital, University of Southern Denmark, Odense, Denmark

Key words
guidelines, liver, fibrosis, chronic liver disease, hepatitis, chronic viral hepatitis, shear wave elastography, strain elastography, alcoholic hepatitis, NAFLD

received 03.01.2017
accepted 02.02.2017

Bibliography
DOI http://dx.doi.org/10.1055/s-0043-103952
Published online: April 13, 2017 | Ultraschall in Med 2017; 38: e16–e47 © Georg Thieme Verlag KG Stuttgart · New York
ISSN 0172-4614
Basic Principles and Technology

Introduction

In these guidelines, the term elastogram refers to any image of any elastic property of tissue. This article also describes non-imaging measurements of such characteristics over a defined region. These are referred to as point elastography methods. Elastometry, on the other hand, is the measurement of an elastic characteristic of tissue, which may be obtained from either imaging or non-imaging methods.

The basic principles of elastography have not changed since they were outlined in the first part of the original EFSUMB guidelines and WFUMB guidelines on this subject [1, 2]. This paper therefore aims to provide an update of technology changes as of 2017 in this rapidly evolving field. Sufficient recapitulations are provided to allow the present paper to be understood without reference to the earlier paper, although the purpose is not to reproduce the material of the 2013 paper in detail.

Basic principles of elastography

We recommend that elastography be considered as a type of remote palpation that allows measurement and display of biomechanical properties associated with the elastic restoring forces in the tissue that act against shear deformation. This view unifies the different types of elastography and explains why they all display images with contrast to obtain the same underlying information, associated with the shear elastic modulus defined below. It is explained in detail in [1], along with different types of shear deformation. Shear deformation is generated by a force applied to a single location or broadly across the body surface. Importantly, transient shear deformations will propagate as shear waves, as discussed below.

A force may be applied by pressing on or vibrating the body surface, by naturally occurring internal physiological motion or by electronically controlling the ultrasound transducer to create focused acoustic radiation force at a controlled depth. The acoustic radiation force is: proportional to the time-averaged intensity I of the sound beam at the push location and the amplitude absorption coefficient α at the push location, inversely proportional to the speed of sound c, at the push location and increases with any reflection or scattering of the sound at the push location. In the absence of reflection or scattering, the magnitude of the radiation force acting at each point in the volume of the medium is F = 2αd / c [3]. All ultrasound beams generate such force but the force integrated over the length of a...
diagnostic pulse (< 2 µs) is too small to generate a measurable displacement. Beams designed to create a measurable displacement with acoustic radiation force use deliberately longer pulses (50 – 1000 µs) to generate micron-level tissue displacements.

All available ultrasound elastography methods (Fig. 1) employ ultrasound to measure the internal tissue shear deformations resulting from an applied force. The type of force (Fig. 1, column 2) is important. If the force varies slowly relative to the shear propagation time to the depth of interest, as is the case for probe palpation or physiological motion, it is considered quasi-static. It is difficult to acquire images that are quantitative for tissue properties using a quasi-static applied force although quantification relative to a reference tissue or phantom material may be possible. Dynamic forces permit absolute quantification of tissue property. They include impulses and continuous vibrations, and may be produced mechanically at the body surface or deep within the body using acoustic radiation force. Only impulsive dynamic forces are shown in Fig. 1 because current commercial ultrasound elastography systems still do not use a continuously vibrating force, although there are examples of this in the research reported in the ultrasound elastography literature [4–6]. This is unlike magnetic resonance elastography (MRE), where a continuous harmonic vibration induced at the body surface is the most common approach [7–9].

The signal processing within the scanner for all current commercial ultrasound elastography methods begins with the measurement of tissue displacement as a function of spatial position and time, which is done using cross-correlation tracking, Doppler, or other signal processing. The various elastography methods differ with respect to what they do with these displacement data to create an elastogram or elasticity measurement. There are three options for the property to be displayed (Fig. 1, column 4):

a) Display of displacement without further processing (Fig. 1, row 2), as in acoustic radiation force impulse (ARFI) imaging. Tissue displacement is associated with shear deformation. The greater the force, the greater the displacement, but stiff tissues are displaced less than soft tissues. Stiffness is defined as applied force divided by the generated tissue displacement. If the force (not normally known) is assumed to be the same for all image locations, an image of displacement can be thought of as an inverse relative stiffness map (it is the inverse, because the higher the Young’s modulus, the smaller the strain). Although strain itself is a quantitative measurement (units of %), image brightness is typically scaled between soft (bright) and hard (dark).

b) Display of tissue strain or strain rate (Fig. 1, row 1), calculated from the spatial gradient of displacement or velocity, respectively, as in strain elastography (SE). Displacement imaging works well when a highly localized force is applied internally at a depth, as in ARFI imaging. However, when the force is applied across a large region of the surface of the body, strain provides a better measure of deformation, as, unlike displacement, it does not decay with depth. Young’s modulus E is a true tissue property, measurable by axial compression of a tissue sample with a constant cross-sectional area along its length. It is equal to the applied stress σ divided by the measured strain ε, i.e., the well-known Hooke’s law E = σ/ε, where stress is the applied force per unit area and strain is the change in length of the sample divided by its original length. If the stress (not normally known) is assumed to be the same for all image locations, an image of strain can be thought of as an inverse relative Young’s modulus map (it is the inverse, because the higher the Young’s modulus, the smaller the strain). Although strain itself is a quantitative measurement (units of %), image brightness is typically scaled between soft (bright) and hard (dark).

c) Display of shear wave speed (Fig. 1, rows 3 – 7), calculated by using the time varying displacement data to measure the arrival time of a shear wave at various locations. All such methods are grouped under the heading shear wave elastography (SWE), and include transient elastography (TE), point shear wave elastography (pSWE) and multidimensional shear wave elastography (2D-SWE and 3D-SWE). This is only possible when the force is applied dynamically. The shear deformation generated locally and momentarily within tissue will propagate as a shear wave and, so long as it is not completely attenuated as it travels, will reach a distant location after a time determined by the shear wave speed c_s. Shear wave speed may be displayed in units of m s⁻¹. Alternatively, assuming that it does not vary with the magnitude or frequency of the applied force, or with position and direction in the tissue, c_s may be converted to either Young’s modulus E or shear modulus G, using the equations E = 3pc_s² and G = E / 3 (assuming the tissue is incompressible), where p is the density of the tissue. E and G are expressed in units of kilopascal (kPa). The choice of whether to display speed or modulus may be under user control, or such a choice may be unavoidable (sometimes determined by regulatory authorities for the region of the world in which the scanner or measurement system is manufactured or sold).

Explanation: It is typical of the MRE literature that data are provided for shear modulus G, whereas in the ultrasound elastography literature it is more common to provide values for Young’s modulus E. The reader should be wary when comparing data across published papers, as both quantities are expressed in the same units (kPa), and yet E = 3G. When publishing results of studies, it is important to state which modulus has been measured when using units of kPa, avoiding non-specific terms such as stiffness, and to say how it was calculated.

The reasons why ultrasound can be used to make elastograms

Ultrasound is the propagation of a transient density deformation. In soft tissues, it travels at speeds in the range of 1350 – 1600 m s⁻¹, whereas shear deformation travels much slower, in the range of 1 – 10 m s⁻¹ [10]. This speed difference means that ultrasound may be used to measure tissue displacements at precise phases of shear deformation. Ultrasound also suffers relative-
*Fig. 1* The types of elastography described in this paper.
ly little attenuation, allowing high-resolution (mm) measurements of displacements of a few microns at depths of many centimeters. Furthermore, the speckle in ultrasound images allows measurements of tissue displacements even when there are no resolved tissue structures to track.

The differences between elastograms and ultrasound images

The tissue contrast in all types of elastograms is potentially very high because it is related to differences in shear modulus G, which varies by more than five orders of magnitude across all soft tissues, as compared with bulk modulus K which varies by about 12%. Ultrasound is able to propagate through fluids, and in tissue the molecular composition largely determines ultrasound speed and absorption, whereas shear will not propagate in a non-viscous pure fluid and is transferred largely via the high-level tissue structure. Although ultrasound scattering occurs at the large-scale structural level, it also arises at the cellular level and does not require the structure to be connected. Shear on the other hand will not propagate across zero friction breaks in the structural matrix. For example, a dilute suspension of cells in saline will scatter ultrasound but will not support a shear deformation or a shear wave. Finally, and very importantly, unlike the situation for ultrasound propagation and scattering, the shear modulus of most tissues increases with vascular and interstitial pressure. These features may make elastography sensitive to the presence of disease in situations where conventional ultrasound imaging fails to detect it.

Technology of ultrasound elastography

The types of elastography referred to above and commercially available systems are shown in Fig. 1. Note that they may be complementary in many respects for general clinical use. For applications in the liver, methods that display shear wave speed (SWS) are the most common, followed by displacement imaging (for lesions in the liver) and finally by strain imaging.

Strain elastography (SE) and strain rate imaging (SRI): quasi-static strain imaging (Fig. 1, row 1)

Strain elastography, although it remains the most widely implemented elastography method on commercial systems, is the least used method for the liver. As motion occurs due to either palpation with the ultrasound probe or to physiological motion, images of axial displacement between sequentially acquired ultrasound echo images are calculated using either radiofrequency echo correlation tracking or Doppler processing (when Doppler processing is used the method is often called strain rate imaging). A moving window axial-gradient estimator converts the axial displacement images to strain images. Little has changed concerning the available systems, artefacts and advantages and disadvantages of strain imaging, although the trend is towards increased robustness, using more frame and pixel quality selection methods, and increased sensitivity allowing useful elastograms to be obtained with little or no palpation, relying on involuntary hand motions or physiological motion, on some systems.

Advice concerning semi-quantitative analysis of strain images using, for example, strain ratios and strain histograms, remains as described in [1].

Acoustic radiation force impulse (ARFI) imaging: imaging tissue displacement induced by radiation force (Fig. 1, row 2)

The axial displacement for locations down a given image line at and around a single pushing focus is measured by comparing the locations of tissue echoes and/or speckle for imaging pulses emitted before and immediately (<1 ms) after the push. A displacement image within a small region of interest (ROI) is then acquired by repeating the pulse sequence “image-push-image” for a sequence of A-lines in the ROI. ARFI remains the proprietary imaging technology Siemens Virtual Touch™ as described in 2013, and the advantages and disadvantages and the various image artefacts remain unchanged [1]. It is not used for assessment of diffuse liver conditions.

Shear wave elastography (SWE); methods that use shear wave speed

Transient elastography (TE): shear wave elastometry by measurement of the speed of a shear wave that has been generated using a surface impulse (Fig. 1, row 3)

Despite the name, this is not the only method that uses a transient force. All of the dynamic methods in Fig. 1 do so. This device is designed only for liver elasticity measurement and for use also by persons who are not imaging specialists. An automated movement of a piston, which is also a disc-shaped ultrasound transducer, applies a single cycle 50 Hz push to the body surface with controlled applied force. The transient shear deformation created in this way, propagates into the tissue. Its near constant speed for about 4 cm in the liver (before being rendered non-detectable due to attenuation) is measured by a straight line automatically fitted to the displacement M-mode shown on the right in Fig. 2, obtained by processing the ultrasound radio frequency A-mode echo signal versus time.

Point shear wave elastography (pSWE): shear wave elastometry at a location by measurement of the speed of a shear wave generated using acoustic radiation force (Fig. 1, row 4)

When an ARFI is applied at a depth within tissue, a shear wave is generated. Its average speed of propagation from the focal point positioned on one lateral boundary of a measurement ROI to another on the opposite lateral boundary of the ROI may be measured by detecting its time of arrival at that point, relative to that of the ARFI. As with TE, no elastograms are created. Only a regional average of SWS is measured. Ultrasound imaging is used, however, to guide placement of the ROI, and measurement may be made even when it is necessary to propagate through ascites. Other advantages and disadvantages are described in [1]. The main update since [1] is that, as listed in Fig. 1, more manufacturers are now providing this type of measurement, which was known as ARFI quantification when first introduced by Siemens.
Currently, this option is also offered by Esaote, GE, Hitachi, Philips and Samsung.

The quality of the SWS estimate is judged by the SWS estimation algorithms in pSWE systems and may warn the user if it is not adequate.

Multidimensional shear wave elastography (2D-SWE, 3D-SWE): quantitative SWE imaging (and elastometry) by measuring the speed of shear waves generated using acoustic radiation force (Fig. 1, rows 5 – 7)

The method described above for pSWE may be extended to make quantitative images of SWS in a large ROI by placing the ARFI focus (push) at multiple sequential locations and, at each, detecting the shear wave arrival time at multiple lateral locations. This creates patches of small SWS images that may be mosaicked to create a large ROI 2D-SWE image, which is displayed in color or grayscale. This may be overlaid with adjustable transparency on the B-mode image or, optionally, the SWE image and B-mode may be displayed separately, side-by-side. In addition to visual judgment of such elastograms against a color scale (calibrated in ms\(^{-1}\) or kPa), they may be quantitatively analyzed using measurement boxes. These ROIs can be placed at desired locations to obtain statistical quantities such as the mean, standard deviation, minimum and maximum values of the SWS or Young’s modulus in kPa (although the minimum is often not useful because of missing data values). It takes time to push sequentially and measure at multiple push locations. Therefore, in some implementations, a static SWE image is created (in about 700 ms on the Siemens S3000\(^{\text{TM}}\)), which may be repeated upon user command, after allowing time for the transducer to cool. In the SuperSonic Imagine Aixplorer\(^{\text{TM}}\) ‘real-time’ 2D-SWE images (up to several frames per second) are created using a combination of an axially swept push focus, which generates a push line in the SWE image ROI and a shear wave that should decay less rapidly than that from a single push focus. The echo images are generated at thousands of frames per second by using plane transmitted ultrasound waves and highly parallel receive beam-forming to allow detection of shear wave arrival times at many locations without the need for many repeated ARFI pushes.

The main update since [1] is that other manufacturers, including Toshiba, GE, Philips and Mindray, have now released 2D-SWE products that are broadly similar but differ with respect to the details of the method and sampling rate.

Toshiba, Philips and Mindray seem to have implemented a conventional ARFI pushing method similar to Siemens. Mindray employs highly parallel receive beam-forming and a zoned transmit wave to sample the travelling shear wave at more than 1000 frames per second within areas up to 4 cm wide. Philips also employs highly parallel processing, allowing free-running 2D-SWE over areas up to 5 cm high by 7 cm wide with an abdominal curved array transducer at a frame rate of 0.4 – 1.6 Hz, while the underlying B-mode runs at 20 Hz for visual guidance (Fig. 3b).

GE has implemented the “comb push” technique [11], which (as illustrated in Fig. 4) pushes along a number of ARFI lines simultaneously. As a consequence, the propagating shear waves travelling away from different push lines pass each other. To avoid confusion of detected times of arrival for waves travelling in different directions, the system analyzes the arrival time at a constant depth as a function of lateral position and uses Fourier domain filtering to separate waves travelling to the left from those travelling to the right. In combination with time-interleaved interpolation of the displacement tracking results, this allows an acceptably large 2D-SWE ROI to be formed in about 100 ms using a conventional (hardware) beamformer.

Note that direction filtering is also employed by other manufacturers, because it has the additional benefit of reducing SWS measurement artefacts that are, for example, due to shear wave reflections at shear impedance boundaries (where either the tissue density or the shear wave speed changes). It is also becoming common on the recent 2-SWE systems (e.g. Toshiba, GE, Mindray) to provide a choice between a continuously refreshed image (almost real time) and a more slowly acquired, said to be of higher quality, single frame.

All 2D-SWE systems have built-in indicators of the quality of the shear-wave speed estimate and take appropriate action to adjust the display when the quality falls too low, such as dropping out the offending pixel of the 2D-SWE image overlay to allow the
B-mode image to show through, or setting that pixel to black. Typically, this will occur as the signal-to-noise ratio deteriorates with depth, limiting the penetration capability of 2D-SWE, as in the example shown in Fig. 5.

An additional approach to quality judgement is provided by Toshiba as shown in Fig. 6 (right), which optionally shows the detected times of arrival in a ”wavefront” style of display. The concept is that if consecutive lines are parallel (even if irregular), the quality is acceptable, but it would not be acceptable if they are not parallel. In the image shown, it can be seen that time-of-arrival lines become spread further apart when the speed increases (c.f. image on the left). It appears as if the shear wave has propagated from left to right, becoming irregular in shape as it emerges on the right of the relatively stiff inclusion, where the elastogram appears to show a stiffer region in the background but the non-parallel time-of-arrival lines in the image on the right indicate that this is an artefact.

Philips provides an additional optional confidence map (Fig. 3c) that combines multiple image quality metrics, which may be displayed concurrently with the elastogram, and that guides the operator with respect to making measurements (either online or retrospectively on stored images) in areas where confidence in the SWS assessment is high.

ROIs with a side length of many centimeters may usually be set and 2D-SWE image penetration depth may be at least 8 cm, depending on the system and transducer used. Extensions to make 3D quantitative SWE images (3D-SWE) have to date been limited to the use of 3D probes that contain mechanically swept 1D transducer arrays. Such a feature is available on the SuperSonic Imagine Aixplorer™, which is able to take advantage of its real-time 2D-SWE capability to acquire a 3D-SWE image stack.

The majority of advantages, disadvantages and artefacts for SWE imaging remain as discussed in [1]. Pressure of the probe may cause superficial tissues to stiffen because of the nonlinear stress-strain character of tissue, but the resulting high shear wave speed artefact is generally not a problem for liver imaging because it only occurs near the surface.

Note that an absence of echo signal on the B-mode image does not always produce an absence of SWE signal, as should happen in a cystic structure (e.g. Fig. 6, left). The artefactual SWE signal may result from an interaction between the relatively large slice thickness typical of most SWE images and limited elevational dimensions of the cystic structure. Similar influences of the background on quantitative shear wave speed values for any small structures, whether solid or cystic, stiff or soft, can be expected.

**Penetration depth of the various technologies**

When SE relies on the application of a quasi-static surface deformation rather than internal physiological motion, the limited area of the probe (among other factors) causes the strain to decay rapidly with depth. This limits the use of SE to only the most superficial regions of the liver, although the use of a probe footprint extender may increase the depth over which strain information is useful.

For the dynamic deformations employed by the SWE methods, as with ultrasound waves, the attenuation of shear waves in tissue increases with frequency, except that the attenuation coefficient for shear waves is about 10^14 times that for ultrasound at the same frequency [13]. As a result, very low force-vibration frequencies must be used to generate shear waves for elastography. When surface mechanical sources are used to create the shear waves (as in TE), shear wave penetrations of many centimeters are needed, requiring the use of frequencies of tens of Hz. When
focused ARFI is used to create the shear waves, the shear waves do not need to reach the depth of interest by propagation from the surface and higher shear wave frequencies (hundreds of Hz) may be employed because the penetration is achieved by the ultrasound beam used to generate the shear wave. Even though the shear wave may then travel only a short distance (because of its higher frequency), another shear wave may easily be generated at a new source position using another ARFI beam.

Although the wavelength is about a thousand times shorter for shear waves than for ultrasound waves of the same frequency, the fact that the frequencies are so low for SWE means that wavelengths are in the region of many centimeters (wavelength = speed divided by frequency). Shear wave sources such as surface thumpers or focused acoustic radiation force impulses are typically therefore smaller than a wavelength, producing diverging shear waves. Nevertheless, the method described above for ultrasonically sweeping the ARFI focus faster than the SWS, which has been used to create a line-like source deep in the tissue, could be used to create a source that is to some extent focused.

kPa or m s⁻¹?

As mentioned in 1.2c) some SWE systems allow a choice for the image scale and the measurements to be reported in units of either m s⁻¹ or kPa. There is more than one reason why it is preferable to report results in units of m s⁻¹ rather than kPa. First, this is the unit of shear wave speed cₛ, which is the quantity measured by the scanner. If the user requests the results in kPa, the unit of elastic modulus, the scanner must convert the measured data to an elastic modulus, such as Young’s modulus E. This conversion typically makes use of a very simple equation, \( E = 3\rho c^2_s \). As with most simple equations, many assumptions are made and these are generally not valid. In this case, these assumptions include that:

a) The tissue density is always 1000 kg m⁻³. This may not always be true, although the error resulting from the assumption may be small.
b) There is no variation of the elastic modulus with the magnitude of applied force (i.e., either the push magnitude that generates the shear wave or any pre-stress, such as probe pressure or portal hypertension, upon which the push is superimposed). This is equivalent to assuming that the tissue’s elastic response is linear. In general, this is not true and indeed, in the future it is expected that the amount of non-linearity will be found to be an important tissue property.
c) There is no variation of the elastic modulus with the shear wave frequency. This is equivalent to assuming that the tissue is purely elastic, although in fact all tissues are viscoelastic, causing the tissue to become stiffer the faster it is strained, i.e., the higher the frequency.
d) There is no variation of the elastic modulus with direction, i.e., the tissue is mechanically isotropic, although almost all tissues are on some length scale anisotropic.
e) There are no tissue boundaries or structures nearby which may cause the wave to travel as another kind of wave (e.g., a Lamb or Rayleigh wave), or shear wave scatterers, all of which will invalidate the assumed relationship between speed and elastic modulus.
The fact that these assumptions may not be correct for many tissues makes the elastic modulus an indirect measurement. Accurate conversion of measured speed to Young’s modulus requires much more complex equations and many more measurements, and is in general impossible. Shear wave speed is the direct measurement, and the U.S. Food and Drug Administration has thus far only approved systems for sale that use numerical elastography displays in units of SWS.

Second, tissue elastic modulus values in the literature are often given for G, the shear modulus, rather than Young’s modulus. In particular, this is the convention in the MRE literature. Both moduli have units of kPa, which may mislead into thinking that data can be pooled, whereas the numerical values for the two moduli are different. Conversion using the relation $E = 3G$ is possible, under the assumption that the tissue is incompressible, but to apply this relationship authors must state which modulus has been measured.

Reporting results in terms of SWS ($\text{ms}^{-1}$) would appear to resolve these difficulties. However, it does not free the scientist from having to provide full details of all of the above experimental variables (a-e), particularly the shear wave frequency, since the SWS will vary with each of them.

Finally, although it is advisable to report SWS rather than elastic modulus, there are complexities. First, measurements of viscoelasticity are likely to be provided on ultrasound elastography systems. This is already the case for MRE, where results are often cited in the form of a complex shear modulus with real and imaginary components. This means that the situation for ultrasound elastography will have to be kept under review, although a preferred method (rather than in the future switching to an indirect measurement such as the complex shear modulus) may be to report direct measurements such as SWS and shear wave attenuation coefficient, and their frequency dependences. It is important to realize that because of the squared relationship between modulus and speed, means and standard deviations calculated for data recorded in $\text{ms}^{-1}$ and then converted to kPa will not be equal to the means and standard deviations calculated after first converting the original data values to kPa. This will be true across patient datasets and for means and standard deviations calculated over ROIs for inhomogeneous tissues. Consideration of this issue is essential when trying to pool data.

**Comparability of data between the different SWE technologies**

As already indicated, measured values of SWS will vary with a number of system factors, in particular shear wave vibration frequency and bandwidth. In addition, measurement bias may occur due to the software method employed, for example, to calculate relative shear wave arrival time and speed. Thus, although it is always a good idea to report as many of the experimental and system variables as possible, at present, corrections for the effects of these variables is not possible. Therefore, unless
a specific study is undertaken to derive an experimental factor in the subject population of interest, to correct for the biases of different equipment, data should not be pooled when acquired using more than one product. Importantly, stiffness thresholds for clinical use known for specific equipment should not be utilized for other equipment. Unless proven otherwise, this recommendation should be taken to apply not only to equipment from different manufacturers but even to different equipment from the same manufacturer or different settings for the same equipment. For example, depth-dependent biases in SWS measurement have been reported. Thus, depth setting standardization should be employed when accumulating data within a population.

Investigator education

EFSUMB is working to promote high quality in ultrasound education and sustain excellent professional standards in elastography training and practice [14]. To ensure the lowest possible intra-operator variability, EFSUMB recommends that ultrasound elastography be performed by operators that have passed competence Level 1. However, it may also be possible to train dedicated personnel to perform elastographic measurements only [15]. It is known from TE [16, 17] that the required learning period is short and this has also been observed in studies using ultrasound elastography [18, 19]. Nevertheless, it appears that experience in liver ultrasonography plays a role in the ability to perform stiffness measurements, at least by means of shear wave imaging, and leads to more reliable elastographic measurements, particularly in obese subjects [20].

Operator experience significantly influences the reliability of liver stiffness measurement (LSM) performed with TE, with an operator with less than 500 procedures representing a risk of 2.6 fold for unreliable results [21, 22]. TE can be readily learned by a nurse or a technician after approximately 100 examinations [16, 17]. 100 examinations should be considered the minimum required training, and training including > 500 examinations yields an experienced TE operator.

There is no agreement as to what constitutes an experienced operator for pSWE and 2D-SWE. Proposed definitions for 2D-SWE include having performed > 300 abdominal ultrasound scans or > 50 supervised 2D-SWE examinations [20, 23].

A major question remains: Who should be allowed to acquire and interpret elastography data? To date, there is insufficient data in the literature to answer this question adequately. However, TE acquisition can be performed by nurses or dedicated persons. Training for non-clinicians should insist on reporting standards. Interpretation should be performed by a physician with knowledge of the clinical condition of the patient. The key issue is to obtain adequate competence and training and to perform elastography in accordance with the medico-legal regulations of the relevant country. The following recommendations are in accordance with the EASL guidelines.

The basic principles and technology for elastography were developed by the academic research community before commercial translation, and it remains a heavily researched and rapidly developing field. EFSUMB recommends that users maintain an awareness of this field.

RECOMMENDATION 1
The operator must acquire appropriate knowledge and training in ultrasound elastography (LoE S, GoR C). Strong consensus (13/0/0, 100 %)

RECOMMENDATION 2
Data acquisition should be undertaken by dedicated and specially trained personnel. For pSWE and 2D-SWE, experience in B-mode ultrasound is mandatory (LoE S, GoR C). Strong consensus (13/0/0, 100 %)
Shear wave elastography (TE, pSWE and 2D-SWE), general technical comments

Introduction
The liver is an important target organ for the use of elastography; stiffness correlates with the degree of fibrosis and indirectly with portal hypertension (see liver application).

Examination procedure
Subjects should be examined in a supine position with the right arm in maximal extension. The transducer is positioned in a right intercostal space to visualize the right liver lobe in A or B mode. Artifacts and large vessels on the A-mode (TE) or B-mode image (pSWE and 2D-SWE) should be avoided. Optimal measurement quality of pSWE and 2D-SWE occurs with the ROI placed a minimum of 1 – 2 cm and a maximum of 6 cm beneath the liver capsule [24 – 27]. In the authors’ opinion, the focus should be placed at the level of the ROI.

Deep inspiration affects the measurement. A transient breath hold in a neutral position is optimal. Details of the examination techniques are documented in the specific sections. Measurements in the left liver lobe are significantly higher and show more variability than those in the right lobe [28, 29]. 2D-SWE LSMs correlate best with the stage of liver fibrosis when obtained in the upper right lobe of the liver [30]. LSM in the left liver lobe should be avoided. The time required for the measurement with TE, pSWE and 2D-SWE is usually < 5 minutes [31].

Factors influencing liver stiffness independent of liver fibrosis (confounders)
Liver stiffness does not solely reflect liver fibrosis, but can reflect many other physiological or pathological conditions. The majority of these variables have been described in studies addressing TE, but should be the same for other SWE methods. Liver stiffness is increased with hepatic inflammation (often but not exclusively shown by an elevated transaminase level) [41 – 44], obstructive cholestasis [45] and hepatic congestion [46, 47]. Also acute toxic hepatitis (phenprocoumon-induced liver failure) increases measurements [48]. For patients with falsely elevated LSMs due to alcoholic hepatitis, liver stiffness decreases following 1 – 4 weeks of abstinence [49 – 51]. Other diseases, which cause increased liver stiffness, independent of liver fibrosis include amyloidosis, lymphomas and extramedullary hemopoiesis. Presently, it is uncertain whether hepatic steatosis modulates liver stiffness [52, 53] or does not [54, 55].

Normal values
TE measurements of Young’s modulus in healthy people vary between 4.4 and 5.5 kPa (95th percentile 6.7 kPa) [60 – 63]. LSMs are generally higher in men than in women [60, 61] and may be affected by steatosis [60] but are not influenced by age [62, 63].
Transient elastography (TE)

Procedure

Transient elastography uses an ultrasound displacement M-mode and A-mode image produced by the system. The operator locates a portion of the liver at least 6 cm thick and free of large vascular structures. By pressing the acquisition button, the machine displays the median of the measured Young’s modulus in kPa, the IQR (the difference between the 75th and the 25th percentile), IQR/median (IQR/M), the value of the current measurement and, only in the old version of the system, the success rate (the ratio between valid and total number of acquisitions). The system displays a result only if the acquisition is valid, since the software automatically rejects acquisitions without correct vibration shape or a correct follow-up of the vibration propagation [1, 32, 76].

The new version of the system suggests using the M or the XL probe in adult subjects based on calculations of the skin-to-liver capsule distance. For children as well as in adults with a thoracic circumference ≤ 75 cm, the S probe is recommended, either S1 for a thoracic circumference < 45 cm or S2 for 45 – 75 cm [77].

How to measure?

Following the manufacturer’s recommendation, assessment is reliable when 10 valid readings and an IQR < 30 % of the median (IQR/M < 30 %) are obtained. The majority of studies have used these reliability criteria as well as a success rate ≥ 60 %. Areas under the receiver operator characteristics (AUROCs) curves were compared for the prediction of cirrhosis between the median values of the 3 first, the 5 first and the 10 first successful acquisitions of TE [78]. No significant loss in performance was reported if only 5 valid acquisitions were obtained. It is likely that IQR/M is the key factor for discordance between TE and histological fibrosis stage [79, 80]. In a recent large population study, new reliability criteria based on IQR/M and median LSMs were proposed [81]. According to this study, TE results should be classified as “very reliable” if IQR/M ≤ 10 %, regardless of LSM; “reliable” if 10 < IQR/M ≤ 30 %, or IQR/M > 30 % with Young’s modulus < 7.1 kPa and “poorly reliable” if IQR/M > 30 % with Young’s modulus ≥ 7.1 kPa. However, these criteria have not been externally validated. A reliable TE assessment can be achieved in over 90 % of adults, when both the M and XL probes are used as required [82 – 85]. Because the M probe takes measurements between 25 and 65 mm from the probe, to increase viability, those patients with a skin-to-liver capsule distance (SCD) of > 25 mm should be assessed with the XL probe. In practice, < 8 % of patients with a body mass index (BMI) < 30 kg/m² have an SCD > 25 mm. The rate of SCD > 25 mm increases to 50 % for patients with a BMI of 35 – 40 kg/m² [86].

Although the problem of overweight patients is not completely resolved with the XL probe [85, 87], the applicability of TE increases when both the M and XL probes are available [82 – 85, 88 – 91]. The diagnostic accuracy of the XL probe appears similar to that of the M probe but the Young’s modulus values are lower than those obtained with the M probe by a mean of 1.5 kPa (range of 0.8 – 2.3 kPa) [82 – 85, 92]. The lower values obtained with the XL probe can be explained by the fact that the validation of the XL probe in overweight patients was undertaken in comparison with the M probe. In these patients there is a high probability of SCD > 25 mm and the measurements obtained with the M probe are likely to be more sub-capsular and, thus, higher [93].
Reproducibility

The intra- and interobserver agreements are excellent, with reported intraclass correlation coefficients (ICC) above 0.90 [94–96]. The agreement decreases in overweight patients or in early stages of fibrosis [94, 95]. Although the LSM seems to be reproducible at different examination sites, the best examination site is the median axillary line on the first intercostal space under the liver percussion dullness upper limit, with the patient lying in dorsal decubitus [95].

Point shear wave elastography (pSWE)

Experience with point shear wave elastography (pSWE) has been mainly acquired with the VTQ® product, because it was the first method available, subsequently followed by ElastPQ® and, more recently, by pSWE methods from many companies.

Procedure (how to measure?)

The operator can select the depth at which liver elasticity is evaluated by placing a “measuring box” (size depending on the manufacturer) in the right liver lobe (segment V, VIII or VII), via an intercostal approach and with the transducer at 90° in relation to the liver capsule, in an area free of large vessels. Published studies [97–99] show that reliable measurements can be obtained in >90–95 % of patients. In a pSWE study using VTQ® to measure SWS [24], the best correlation with histological fibrosis was observed for measurements performed 1–2 cm and 2–3 cm beneath the liver capsule (0.675 and 0.714, respectively), but in up to 15 % of cases, measurements could not be obtained if performed 2–3 cm under the liver capsule.

How many measurements?

Most studies perform 10 valid measurements by pSWE and report the median of these values. A few studies have used only 5 [99, 100] or 6 [101] valid measurements. Another study [102] calculated the mean and standard deviation (SD) of 10 valid measurements. A high SD correlated with misclassification of fibrosis. Additionally, higher stages of fibrosis were associated with a higher SD, indirectly indicating that “more” measurements should be obtained in patients with suspected fibrosis.

Reproducibility

pSWE has excellent intra- and inter-operator reproducibility for liver elastography assessment in both healthy subjects and patients with chronic liver disease [66, 70, 72, 103–105].

Healthy subjects

In a study on SWS measurement, reproducibility was assessed by VTQ® in 50 healthy volunteers. A very good ICC (0.86) was found [66]. Age, gender and BMI did not influence the reproducibility of this method. Similar results (ICC = 0.87) were obtained [103] in 20 healthy subjects. In a cohort of 69 healthy subjects, ElastPQ® showed very good intra- (ICC = 0.96) and interobserver (ICC = 0.93) reproducibility [70].

Patients

The reproducibility of SWS assessed by VTQ® was evaluated in 61 patients with chronic hepatopathies [104]. The following cut-off values for liver fibrosis staging using SWS were applied: F2 ≥ 2.137 m/s; F3 ≥ 3.145 m/s and F4 = 1.75 m/s. There was 87 % agreement between repeated measurements for different stages of fibrosis. For the differentiation between patients with significant fibrosis (F ≥ 2 Metavir) and those with fibrosis stage F < 2, the interoperator agreement was 90 %. In another study [19] excellent agreement (ICC = 0.84) was observed in a cohort of 50 patients with various etiologies of chronic liver disease. Other authors [101] assessed the correlation of SWS values obtained with VTQ® by 2 operators in 41 patients with chronic hepatopathies. A good interoperator correlation was obtained (Spearman correlation coefficient r = 0.874, p < 0.0001) with similar results [99, 100, 105]. VTQ® reproducibility for SWS values and influencing factors were assessed in 33 and 58 patients, respectively. The overall interoperator agreement was better than the interoperator agreement: ICC 0.90 vs. ICC 0.81. For both intra- and interoperator reproducibility, the ICCs were lower in women vs. men (0.88 vs. 0.91 and 0.67 vs. 0.86, respectively), in patients with a high BMI (≥ 25 kg/m²) vs. BMI < 25 kg/m² (0.88 vs. 0.91 and 0.79 vs. 0.82, respectively), in patients with ascites vs. no ascites (0.80 vs. 0.93 and 0.78 vs. 0.84, respectively) and in non-cirrhotic vs. cirrhotic patients (0.77 vs. 0.82 and 0.70 vs. 0.83, respectively) [106]. The interobserver reproducibility of the ElastPQ® technique was assessed in 291 consecutive patients with chronic hepatitis B (CHB) who underwent partial hepatectomy or liver biopsy due to hepatocellular carcinoma (HCC). Very good reproducibility of this technique was achieved (ICC = 0.798) with a similar conclusion reached in a further study [70].
Quality criteria

When VTQ® was first introduced, no recommendations regarding the use of quality criteria for measurements were defined. One study [107] evaluated factors that influenced the correlation of SWS assessed by VTQ® with histological fibrosis in a cohort of 106 chronic hepatitis C (CHC) patients. In univariate and multivariate analysis, an IQR/M ≥ 30 % was associated with a discordance of at least 2 stages of fibrosis between SWS and histological fibrosis. The importance of IQR as a quality criterion was further demonstrated in another study [108]. Others studied the importance of standard deviation (SD) as a quality criterion [102].

10 valid SWS measurements were adopted in this study and the mean value was calculated. The authors obtained the following accuracy for predicting the presence of F≥2, F≥3 and F=4 in all cohorts of patients: 80.7 %, 86.2 % and 88.7 %, respectively. After the exclusion of patients with SD ≥ 30 %, the accuracy of VTQ® for predicting the same stages of liver fibrosis increased significantly up to 86.1 %, 91.2 % and 91.5 %, respectively. Using ElastPQ®, a recent study has suggested that an IQR/M ≤ 30 % is the most important quality criterion, whereas the number of measurements seems not to affect the performance, provided that they are at least five [109]. Thus, the compliance with quality criteria may increase the diagnostic accuracy of pSWE [102, 107]. Quality parameters have been described for other manufacturers as well [110].

2D-SWE

Almost all 2D-SWE studies for liver applications have been carried out using SSI, because other companies have only recently introduced 2D-SWE products. This description is therefore limited to the SSI system, but the principles may be applied to other 2D-SWE products.

Procedure

Obtaining an elastogram

2D-SWE evaluation should be performed in a well-visualized area of the right liver lobe, free of large vessels, liver capsule, ligaments and the gallbladder [30]. Since movement greatly influences results, the subject is asked to suspend breathing.

With 2D-SWE working in continuous and not with single shot emissions, the SWE acquisition is continued for 4–5 seconds (can be longer for other scanning systems) once a stable SWE image is obtained. The operator should aim to achieve homogeneous color filling of the SWE ROI. With SSI, the scale of the elastogram can be modified. Usually a Young’s modulus scale of up to 30 kPa is sufficient, but a higher scale of up to 150 kPa can be adopted on a case-by-case basis. The operator freezes the image (and optionally saves the clip for further post-processing) and an analysis box (Q Box, for SSI) is placed on the most homogeneous, stable elastogram for a few seconds to measure Young’s modulus (SWS, if the scanner is set to that mode). Experience seems to play a role in reliable measurement, particularly in obese patients [20]. However, there is no unified understanding of what constitutes an experienced 2D-SWE operator. Adequate ultrasound knowledge with at least 300 abdominal ultrasound investigations may be considered a minimum requirement [20]. 2D-SWE measurement of liver elasticity should be performed by experienced operators after achieving a stable, homogeneous elastogram [20, 23].

How to measure?

Analysis box size and shape

For 2D-SWE measurements, the analysis box should be set to at least 10 mm, preferably 15 mm or more. A round shape is usually chosen [57, 58]. The ROI should be placed over an isoechoic area of liver parenchyma, as seen on the grayscale image (no vessel, no nodule, no other structure), in priority in the middle line of the elastogram (avoiding positioning the Q Box on the edges of the elastogram), while also avoiding SWS artefactual areas (reverberation, noisy areas from rib shadowing).

Valid and invalid measurements

There is no agreement on objective quality criteria. Some authors suggest that a minimal Young’s modulus value of ≤ 0.2 kPa in the analyzed region is useful to identify invalid measurements as indicated by a lack of concordance with TE [111], while others use a minimal Young’s modulus < 1 kPa. Furthermore, among valid measurements an IQR/M ≤ 30 % is recommended by other studies mimicking TE reliability criteria. The latter criterion, however, is not supported by evidence. For 2D-SWE with Logiq E9 (GE), the manufacturer recommends an IQR/M value below 30 % as a quality criterion. Temporal stability of the elastogram for three seconds or more during breath hold, in combination with placement of the analysis box in a homogeneous area with complete filling results in high accuracy, high reliability and low variance of measurements with SSI [112 – 114]. The new software version of the Aixplorer® system also shows the stability index (SI) and according to the manufacturer a reliable LSM should exclude measurements with an SI < 90 %. Aplio 500 (Toshiba) provides a display of shear waves travelling within the box, allowing selection of areas not affected by artefacts for analysis. For 2D-SWE with the Philips system, a confidence map guides the operator to perform measurements in areas where the signal-to-noise ratio of the SWS assessment is high.

Data regarding the best criterion for judging a measurement as reliable are limited. Hence, following the manufacturer’s advice for acquisition is suggested in the absence of better evidence [56, 112 – 114].

How many measurements?

From 3 to 15 measurements are used in published studies, but data from several studies suggest that 3 measurements suffice to obtain consistent results for the assessment of liver fibrosis and portal hypertension, and for optimal correlation with TE [27, 58, 59, 114 – 116]. Some studies report the median, some the mean, of the available measurements. There is no convincing evidence to suggest superiority of the mean versus median of the SSI measurements. However, since the median and IQR are
Robust against non-normal distributed data, they should be preferred for reporting.

**RECOMMENDATION 13**
For 2D-SWE a minimum of three measurements should be obtained; the final result should be expressed as the median together with the interquartile range (LoE 2b, GoR B) [27, 117]. Strong consensus (18/0/0, 100 %)

**Reproducibility**

**Reproducibility in healthy subjects**
In three studies, the intraobserver reproducibility of SSI during the same session was excellent (ICC ranging from 0.92 to 0.95) [23, 73, 118]. The interobserver agreement on different days was affected by operator experience and ranged from 0.63 to 0.84 [23, 118].

**Reproducibility in patients**
The intraobserver reproducibility of 2D-SWE for liver stiffness assessment in liver fibrosis patients is excellent, with the ICC ranging from 0.90 to 0.95 in published studies [56, 57, 117]. The intra-subject reproducibility (evaluated over periods of 2 days to 4 weeks) ranges between 0.83 and 0.90 [31]. Interobserver reproducibility on the same day ranges from 0.83 [31] to 0.94 [112]. Intra- and interobserver variance may be inferior to pSWE using VTQ® [31].

**Limitations**

**Failures**
Failure – most often defined as the “inability to obtain an adequate signal for all acquisitions” – occur in up to 10 % of cases reported in the published series. Common causes of failure are: depth below 4 – 5 cm [26], poor ultrasound window, reverberations, pulsatile movement, poor breath hold, large amounts of ascites [117], intercostal wall thickness ≥ 25 mm [119], BMI ≥ 30 kg/m², histological steatosis and waist circumference ≥ 102 cm [57, 59].

**Unreliable assessments**
Unreliable measurements are randomly defined across studies and comparison is limited. Reported predictors of unreliable results with SSI are similar to those reported for TE and pSWE [56, 117]. In one study, a lower SD and larger analysis boxes independently correlated with correct classifications of severe fibrosis and cirrhosis [114].

The main factors limiting the applicability of 2D-SWE include obesity, poor acoustic window or presence of artefacts and inability of the subjects to hold their breath [26, 56, 57, 59, 117].

**Strain elastography (SE): quasi-static strain imaging**

**Procedure**
The best B-mode ultrasound image is required since real-time strain elastography (SE) depends on the quality of the B-mode images. Acquisition can be improved as follows [32]:
- Image the right liver lobe through an intercostal space with the patient supine and the right arm elevated to widen the intercostal space;
- Place the transducer lightly on the skin without moving it since intrinsic, mainly cardiac, movement is often enough to displace the tissue. However, sometimes the operator needs to induce repeated pressure movements with the probe to create the necessary stress needed for strain measurements;
- Select a region of interest in which B-mode images are free from interfering structures and artefacts;
- Obtain images displaying axial, not lateral or elevational, movement by pointing the transducer toward the heart;
- Have the patient perform a short breath hold to ensure that SE images are displayed consistently [120].

**Analysis region placement and size**
Similarly to pSWE and 2D-SWE, the analysis box should be placed at least 1 cm below the liver capsule [120, 121]. Measurement should not be performed too deep in the parenchyma as acoustic shadows, reverberation artefacts, and a lack of sufficient penetration will generate incorrectly high readings. Some researchers include the surrounding tissues, such as the subcutaneous and muscle layers [122, 123]. However, placing the ROI entirely inside the liver is key to generating uniform images to calculate the liver fibrosis (LF) index [124].

Large vessels should be avoided by using a 2.5 × 2.5 cm ROI [120, 125]. Elimination of artefacts requires attention to technique. Trialling the placement of the transducer between the ribs will lead to optimal positioning. When an examination is difficult, attempt another intercostal space, perhaps selecting one that is softer and with a thinner subcutaneous layer. Other subcutaneous structures, such as the ribs and lungs, should not be included in the image. For the analysis, frames with strain generated in the depth direction with no artefacts should be selected. Good images may be obtained at the end of diastole with electrocardiographic gating or at the largest downward wave.

**Evaluation methods**
An examiner’s experience influences the outcome of visual assessment. To overcome this, various quantitative and semi-quantitative methods have been developed to assess tissue strain more objectively. In the research literature “image pattern recognition” indices obtained by analyzing grayscale, histogram, and binarization transformations of SE images have been shown to correlate with liver fibrosis. The strain image is converted to numerical values, for example using color gradations, with blue being 0 and red being 255. Some authors have reported that mean strain...
values inversely correlated with liver stiffness and fibrosis in patients with CHC. On the other hand, the standard deviation of values of strain, the percentage of area of low strain and its complexity were positively correlated with liver stiffness and fibrosis [126).

Scores
Different scores for semi-quantitative interpretation of SE have been proposed: the German Elasticity Score [121], the Japanese Elasticity Score [126] and the LF Index [127]. The LF Index (an algorithm based of nine features evaluated directly by software) is the only one recommended by the manufacturer [32]. However, evidence is not sufficient to allow recommendation for clinical use, at least in European patients [121, 125].

Reproducibility (published literature)
The interobserver consistency among four separate locations in the liver using a semi-quantitative method was almost perfect (ICC 0.97) [128]. On the other hand, the intraobserver variability and intraobserver agreement of SE for the assessment of liver fibrosis have been criticized in several studies [121, 125], but no new data are available.

Comparison of results between systems
Introduction
Different US-based SWS technologies are available for the non-invasive assessment of liver fibrosis and the measurements produced can be slightly to moderately different between systems from different manufacturers. Even systems that use the same technique but were developed by different manufacturers can yield different values due to different and proprietary methods to measure the SWS. Moreover, the units of measurement may be meters per second if the speed of the shear wave is expressed as m/s or kPa if used to calculate the Young modulus. There is another potential difference: it is assumed that the SWS is independent of the shear wave frequency which is incorrect since higher frequencies generate shear waves that travel faster. The ‘dispersion error’ could render results from different studies incomparable.

Studies on phantoms
The Ultrasound Shear Wave Speed technical committee of the Radiological Society of North America, Quantitative Imaging Biomarker Alliance (QIBA) has quantified the differences between commercially available systems. An inter-laboratory study comparing shear wave velocities obtained with four different machines ( Fibrosan, Philips iU22, ACUSON S2000 and Aixplorer) was conducted [129]. Working on elastic phantoms, a statistically significant difference in the SWS estimates among systems and depth of measurement in the phantom was shown, whereas no statistically significant differences were found among operators using the same or equivalent systems under the same conditions [129]. Similar results were obtained using phantoms with viscoelastic properties similar to those observed in normal and fibrotic liver [130]. The measurements were performed at multiple focal depths (3.0, 4.5 and 7.0 cm). The deepest focal depth (7.0 cm) produced the greatest inter-system variability for each phantom (up to 17.7 %) as evaluated by the interquartile range. Inter-system variability was consistent across all phantoms and was not related to stiffness.

Sources of variability
Several sources of variability are detailed in published studies, including technical and patient-dependent factors that could affect comparability between systems.

Technical factors
Measurement depth
As shown by the studies on phantoms [129], the influence of depth on the estimation of elastic properties is not negligible. Furthermore, with curved transducers used in liver imaging, the angle affects the readings, with the best results being achieved when the ROI is straight ahead. Using VTQ®, it has been shown that the results with the lowest variability are obtained at a depth of 4–5 cm with a convex transducer (1–4 MHz; mean push pulse: 2.67 MHz) and at a depth of 2–3 cm with a linear transducer (4–9 MHz; mean push pulse: 4 MHz) [131]. The acoustic push pulse is progressively attenuated as it traverses the tissue. Attenuation is higher in a stiffer liver. Thus, measurements are more variable in cirrhotic patients [25].

Frequency of the transducer
In a prospective study on 89 patients with CHC, pSWE (based on VTQ®) was performed using both available transducers (4C1 and 9L4) [132]. The linear transducer gave higher values (1.91 ± 0.87 m/s vs. 1.70 ± 0.67 m/s). However, the results were correlated to each other (r = 0.70). Using the same method in a phantom and a series of eight volunteers, it was found that the convex transducer showed values that were significantly higher than those obtained with the linear transducer [131].

Position of the transducer
The highest intra- and interobserver agreement was obtained for the measurements performed through the intercostal space rather than the subcostal approach [72, 103, 133].

Operator experience
Methods using pSWE have shown excellent interobserver agreement, with concordance ranging from 0.80 to 0.97 for measurements performed via an intercostal approach, and independent
of operator experience, suggesting that operators require only a short period of training to perform reliable LSMs [66, 70, 103, 106]. Using 2D-SWE (SSI) an expert operator had higher reproducibility of measurements over time than a novice operator [23]. It is suggested that at least 50 supervised 2D-SWE measurements should be performed by a novice operator in order to obtain consistent measurements.

It was also demonstrated that a novice was able to achieve reliable results with TE after a single learning session. However, the success rate progressively increased with the experience of the operator [16]. On the contrary, it was reported that operator experience increases the applicability and diagnostic accuracy of TE [92]. Echosens, the manufacturer of the FibroScan device, recommends that TE be performed by an experienced operator (> 100 examinations).

Using liver biopsy as the reference standard, a recent study has shown that ElastPQ® matched TE for accuracy after the operator had performed at least 130 examinations [134].

Equipment

Proprietary elastographic technologies generally give different estimates of the SWS within the same liver. This translates into the need to define threshold values for fibrotic stages for each specific equipment model.

Patient-dependent factors

The most important patient-dependent factors for variability between different elastographic methods are: position (supine position has been used in most published studies), respiration phase (deep inspiration affects the measurement; it has been suggested that a breath hold for a few seconds during quiet breathing may lead to the best results) [25, 29, 135], non-fasting state [36–38], fibrosis stage (lower stages of liver fibrosis seem to reduce TE reproducibility [94, 136, 137] or the presence of steatosis [52].

Obesity limits the applicability of elastography methods (TE XL probe, pSWE) and impairs measurement accuracy [138]. The theoretical possibility that values are different rather than identical in different parts of the liver, even in the same region, may also contribute to the variability between different equipment reported in many studies. Unfortunately, there is no indication in most studies as to whether the same sampling position and the same intercostal space were predetermined and utilized.

Clinical studies

Diagnostic accuracy

The AUROCs of ElastPQ® and VTQ® for predicting the presence of liver fibrosis were similar (0.85 and 0.82, p = 0.48) [71]. Other authors compared 2D-SWE (SSI), VTQ®, and TE in 349 consecutive patients with chronic liver diseases [57]. They found that 2D-SWE, TE and VTQ® correlated significantly with the histological fibrosis score (r = 0.79; r = 0.70; r = 0.64 respectively, all p < 0.0001). The AUROCs of 2D-SWE, TE, and VTQ® were 0.89, 0.86, and 0.84, respectively, for the diagnosis of mild fibrosis; 0.88, 0.84, and 0.81, respectively, for the diagnosis of significant fibrosis; 0.93, 0.87, and 0.89, respectively, for the diagnosis of severe fibrosis; and 0.93, 0.90, and 0.90, respectively, for the diagnosis of cirrhosis. 2D-SWE had a higher accuracy than TE for the diagnosis of severe fibrosis (≥F3) (p = 0.0016), and a higher accuracy than VTQ® for the diagnosis of significant fibrosis (≥F2) (p = 0.0003). No significant difference was observed for the diagnosis of mild fibrosis and cirrhosis. Similar comparison was made in patients with NAFLD [57]: 2D-SWE had a higher accuracy than VTQ® for diagnoses of significant fibrosis (≥F2) (p = 0.004).

In a study on 132 patients with chronic liver disease, 2D-SWE, pSWE (VTQ®) and TE were compared. The reference methods were liver biopsy for the assessment of liver fibrosis (n = 101) and magnetic resonance imaging/computed tomography for the diagnosis of liver cirrhosis (n = 31). No significant difference in diagnostic accuracy was found between the three elastography methods for the diagnosis of significant and advanced fibrosis and liver cirrhosis in the ‘per protocol’ and ‘intention to diagnose’ cohorts [57].

Several studies have compared the results obtained with B-mode-based SWE techniques (pSWE, 2D-SWE) to those of TE and have found that the accuracy of these techniques is similar to that of TE [59, 139, 140].

Some comparative studies show similar accuracy for different elastographic systems. Larger prospective studies are necessary to find if there are differences in accuracy between each different system for liver stiffness evaluation.

RECOMMENDATION 15

The results with the lowest variability in comparing different pSWE or 2D-SWE systems were obtained at a depth of 4–5 cm from the transducers (with convex transducers) (LoE 4, GoR C) [131]. Accordingly, this location is recommended if it is technically suitable. Broad consensus (17/0/1, 94 %)

Liver diseases

Introduction

The assessment of fibrosis in chronic liver diseases is pivotal for prognosis and guiding management, including whether to commence antiviral treatment. Liver biopsy is considered the “gold standard” for fibrosis assessment and stage classification and can also grade necro-inflammatory activity. However, liver biopsy is limited by its invasiveness, with potential severe complications in up to 1 % of cases, sampling error, since the specimen represents roughly only 1/50 000 of the liver volume, and the inter- and intraobserver variability at microscopic evaluation. Therefore, noninvasive methods for liver fibrosis assessment including ultrasound elastographic methods have been an intense field of research [1].

Due to the differences among elastography methods outlined in previous sections, cut-off values for fibrosis are system-specific and cannot be equated across machines. Currently, TE is the most
widely studied method outside of radiology practice, but a considerable amount of evidence regarding other methods has been published. The results obtained by pSWE (mainly VTO and 2D-SWE (mainly SSI) show similar accuracy as compared to TE for the assessment of liver stiffness, with a potentially higher applicability [59, 112].

It is becoming increasingly clear that the best cut-off values of the different elastography techniques used to evaluate the presence and severity of liver fibrosis depend upon the etiology of the underlying liver disease, and upon the prevalence of the condition under study in the target population. Although different cut-off values have been proposed for cirrhosis according to etiology (ranging, for instance, from a Young’s modulus of 11 kPa in CHB to 22.6 kPa in alcoholic liver disease), it should be kept in mind that these cut-off values have been defined in a single population using ROC curves in order to maximize sensitivity and specificity – and have not been applied to a validation cohort. Differences between cut-offs may be simply related to differences in cirrhosis prevalence and severity in the studied populations, known as the spectrum bias. Therefore, elastography values should be interpreted by a liver specialist aware of the clinical aspects of the liver disease to be assessed and aware of the peculiarities of elastography in general and each elastography technique in particular.

Besides increasing evidence regarding liver stiffness measured by different techniques to detect liver fibrosis and cirrhosis (described in detail in the following paragraphs), other new applications of elastography are being tested in the field of liver diseases. They include spleen stiffness assessment for portal hypertension and the evaluation of the stiffness of focal liver lesions to differentiate between benign and malignant nodules. These applications appear promising, but remain under development and cannot yet be recommended in clinical practice.

Clinical needs

The ranges for intermediate fibrosis stages (F2–F3) are quite narrow, in the order of a Young’s modulus of 2–3 kPa (over a total range spanning 2 to 75 kPa with the TE), so that small differences in outputs could shift the assessment of patients from one stage to another [141]. However, in “the real life” situation, attention should be focused on the patient and what is appropriate from a clinical point of view. Following the availability of novel antiviral agents, the European Association for the Study of the Liver (EASL) together with the Asociación Latinoamericana para el Estudio del Hígado (ALEH) have produced guidelines for the clinical use of noninvasive tests for the evaluation of liver disease severity and prognosis [76]. These guidelines have outlined that the two clinically relevant endpoints in patients with viral hepatitis are the detection of significant fibrosis (F≥2) and the detection of cirrhosis (F=4), and the most important endpoint is the detection of cirrhosis, because it guides treatment (A1 recommendation). In patients with NAFLD and with chronic liver diseases of other etiologies, the detection of cirrhosis is also the most important clinical endpoint (A1 recommendation).

Clinical decision-making before elastography

Elastography is designed as a noninvasive tool to identify and stage liver fibrosis, and it should be used in an appropriate manner. Specifically, it should be interpreted along with a diagnostic workup aimed at identifying the cause of liver disease. This is important since the interpretation of liver elastography values depends on the specific clinical scenario, the prevalence of disease in the population under investigation, the current patient’s comorbidities and the etiology of the liver disease.

Liver stiffness evaluation is not part of the screening process and initial evaluation. The main reason for this is that a rational stepwise approach to the diagnosis and staging of liver disease should follow the same rationale used for any other disease in order to optimize the use of resources, limit costs and minimize the risk of iatrogenic injury. The backbone of any workup is anamnesis, physical examination and laboratory tests, followed by the use of noninvasive diagnostic methods, ideally from less to more expensive, and following a balanced view of the cost and accuracy of each for the specific clinical question. The use of invasive methods, which always imply a certain risk for patients, should be limited when the previous phases did not provide sufficient information. In addition, despite elastography methods being increasingly used and available in reference centers, they are not yet universally available and require experience. Therefore, elastography is part of a clinical decision-making process.

Anamnesis, clinical examination and abdominal ultrasound examination are recommended prior to elastographic liver evaluation. This strategy detects the presence of ascites, tumoral masses in the liver, obstruction of the biliary tree, and signs of right heart dysfunction. Direct liver elastographic evaluation reduces the time of examination, decreases the cost and can be performed by trained technicians (TE).

Not recommending elastography as a first-line screening approach for liver disease allows an appropriate diagnostic workup in patients without significant fibrosis at elastography. These patients may have a liver disease prone to progression, despite a normal elastography measurement.

Screening

According to the WHO principles for screening for disease [142], a screening test should offer high validity, reliability and acceptance in the screened population, in addition to a positive balance between yield and costs and available follow-up services in the case of a positive test. Thus, SWE may be well suited as a screening method for asymptomatic liver disease in low prevalence populations. However, there is insufficient evidence to evaluate the value of SWE in the general population, in high-risk populations with intravenous drug use, diabetes mellitus, obesity and/ or alcohol abuse or in migrants from areas with a high prevalence of HCV and HBV [61, 143–145].

TE as a screening procedure in a general population > 45 years has been evaluated in two large studies, with a total of 4700 apparently healthy subjects [61, 143]. One study found that the presence of diabetes mellitus and obesity was associated with a Young’s modulus above 8 kPa. However, neither of the studies...
systematically investigated the causes of elevated liver stiffness in participants with a positive result.

**Initial evaluation of chronic liver disease (prerequisites for SWE)**

**Current complaints and past medical history**

The initial evaluation of a patient with abnormal liver biochemical and/or function tests includes knowledge (anamnesis) of current complaints, obtaining a past medical history to identify potential risk factors for liver disease and performing a physical examination to look for clues to the etiology and for signs of chronic liver disease. Elastography should not be performed without the knowledge of current complaints, past medical history and basic laboratory testing.

**Conventional ultrasound**

Conventional ultrasound examination could be part of the initial evaluation [146], especially if the equipment is located within the clinic. Conventional B-mode ultrasound can be used not only to detect the classic features of established cirrhosis but also to determine the presence of biliary dilation indicating extrahepatic cholestasis and of signs of cardiac failure or hepatic vein disease, which are conditions limiting the accuracy of elastography. B-mode US screening for focal liver lesions, splenic size, venous collaterals, ascites and other signs of portal hypertension can also be performed. Ascites precludes the use of TE elastography. Knowledge of conventional B-mode ultrasound findings is mandatory before strain and SWE elastographic examination, or at least affects interpretation of the elastography results. Doppler ultrasound is a basic tool for the optimization of B-mode ultrasound and the exclusion of vascular liver disease and is helpful for determining complications of chronic liver disease, e.g., portal hypertension.

In conclusion, in patients with liver disease of unclear etiology, elastography should not be interpreted without a complete ultrasound assessment of the upper abdomen (B-mode and Doppler ultrasound).

**Chronic hepatitis C (CHC)**

**Fibrosis staging**

**Transient elastography**

In patients with CHC, TE can differentiate absent and mild fibrosis from significant fibrosis and cirrhosis, but is not accurate enough to distinguish between separate stages of fibrosis (F1–F4) [147–149]. A Young’s modulus greater than 6.8–7.6 kPa indicates a high probability of significant fibrosis (F ≥ 2) on biopsy. However, optimal cut-off values vary considerably depending on fibrosis prevalence and may range between 5.2 kPa–9.5 kPa as indicated in the clinical practice guidelines of the EASL [76]. Accordingly, the optimal cut-off values for predicting cirrhosis (F = 4) range between 11 and 15 kPa [76]. Therefore, the local fibrosis prevalence and the diagnostic aim (sensitive screening vs. secure exclusion strategies) must be considered when adapting cut-off values for clinical use. It should be emphasized that TE gives the best diagnostic performance in the context of cirrhosis diagnosis, and for this purpose it is better at ruling out than at ruling in cirrhosis.

In CHC patients with HIV co-infection, TE can be used with similar diagnostic accuracy as compared to HCV infection alone for fibrosis and cirrhosis detection [150]. TE can also be helpful in liver transplant recipients for the staging of recurrent fibrosis and cirrhosis [151].

The use of TE for the diagnosis of cirrhosis and the estimation of fibrosis severity in CHC has been endorsed in the recommendations for the management of viral hepatitis by the EASL and the ALEH, ideally in combination with an alternative and unrelated noninvasive approach such as laboratory tests/serum markers of fibrosis [152]. In the case of TE failure or inconclusive noninvasive test results, biopsy is still recommended when fibrosis staging is relevant for clinical decisions, although a preliminary attempt with an alternative SWE method could be considered.

---

**Point shear wave elastography (pSWE)**

As with TE, VTQ® SWS quantification has been studied extensively in patients with CHC. Cut-offs of 1.21–1.34 m/s predict significant fibrosis (F ≥ 2) (AUROC 0.85–0.89), while VTQ® cut-offs between 1.55 and 2 m/s (AUROC 0.89–0.93) predict cirrhosis [104, 153]. The diagnostic performance of VTQ® is comparable to TE [154] with high accuracy for predicting significant (F ≥ 2, AUROC 0.87) and severe fibrosis (F ≥ 3, AUROC 0.91) as well as cirrhosis (AUROC 0.93) [155]. However, discordance (> one fibrosis stage) between VTQ® and histology occurred in > 30% in a study including 106 patients infected with HCV. The discordance was associated with female gender and a high interquartile range (IQR/M ≥ 30%) of Young’s modulus [107]. Therefore, pSWE results require cautious interpretation. Evidence regarding ElastPQ® is limited [70, 156]. The results in the pilot study [70] for fibrosis staging are similar to those reported for VTQ®, but more data are needed. Fibrosis biomarkers may help to clarify indistinct cases [157].

---

**2D shear wave elastography (2D-SWE)**

2D-SWE using SSI was reported in several studies in patients with CHC [139, 158]. Diagnostic accuracy was high for the detection of
significant and advanced fibrosis and cirrhosis (AUC > 0.90). In these studies the diagnostic performance for 2D-SWE was better than for TE [139, 159] and serum fibrosis markers (FIB-4 index, APRI and Forns’ index) [159]. In one study including 102 obese CHC patients, 2D-SWE had excellent diagnostic accuracy for the detection of severe fibrosis and cirrhosis (AUROC > 0.90 for both) [91].

**RECOMMENDATION 18**

2D-SWE as demonstrated with SSI can be used as a first-line assessment for the severity of liver fibrosis in patients with chronic viral hepatitis C. It performs best with regard to the ruling out of cirrhosis (LoE 1b, GoR A) [139, 158, 159]. Broad consensus (17/0/1, 94 %)

**Strain Elastography**

Among the manufacturers, only Hitachi SE has been evaluated for the assessment of fibrosis in liver disease and in viral hepatitis. Different scores for semi-quantitative interpretation of SE in the liver have been proposed [1]. While the method produced promising results for the prediction of liver fibrosis in Asian patients, its performance was inferior to TE in European patients [121, 125]. New data on this technique are lacking.

**Prediction of hepatic complications**

Growing evidence supports the use of TE for risk stratification and the prediction of clinical endpoints. Liver stiffness predicts 5-year mortality with better accuracy than histological fibrosis staging (METAVIR) for HCV mono-infected (TE > 9.5 kPa) [160] and HIV co-infected patients (TE > 9.0 kPa) [161]. Furthermore, an elevated Young’s modulus indicates an increased risk of HCC development (TE > 10 kPa) [162], hepatic decompensation and variceal bleeding [163]. Especially for patients with an established diagnosis of cirrhosis, TE can be used for grouping patients into different risk classes [162]. The combination of liver stiffness (Young’s modulus of > 14 kPa) with platelet count (< 141 × 10³/μL) and response to antiviral treatment may increase the predictive value of TE for HCC development [164]. In patients with compensated liver cirrhosis, liver stiffness helps to identify patients with portal hypertension (see section recommendations on portal hypertension). Data regarding firm endpoint prediction using pSWE and 2D-SWE are lacking and no evidence-based recommendation can be specified. Two groups evaluated the predictive value of TE in transplant patients with a recurrence of HCV infection [165, 166]. A Young’s modulus > 8.7 kPa at 12 months after liver transplantation was associated with a significantly reduced five-year graft and cumulative patient survival [165].

**Role of elastography in the setting of anti-HCV treatments**

**Role of SWE to identify patients to be treated**

In the absence of universal access to direct-acting antiviral agents (DAA) as a consequence of high cost, different countries have implemented strategies to prioritize patients for treatment. TE is used as the first-line investigation for the prioritization of HCV patients for DAA (e.g. Young’s modulus values ≥ 7.1 kPa are considered equivalent to fibrosis ≥F2, and values ≥ 9.5 kPa are considered equivalent to fibrosis ≥F3). In countries using interferon-based strategies, TE can help to identify patients with cirrhosis who have a lower likelihood of achieving a sustained virological response (SVR: undetectable HCV-RNA, 24 weeks after completion of antiviral therapy).

**Role of SWE during treatment (monitoring)**

Limited evidence in the setting of interferon-based therapy postulated that increasing liver stiffness as measured by TE during the treatment period might indicate a reduced possibility of achieving a sustained virological response [167, 168]. Data in patients undergoing IFN-free antiviral therapies suggest that liver stiffness rapidly declines during treatment, even in patients with advanced fibrosis and cirrhosis. This decline appears to reflect the reduction in liver inflammation, restoration of liver function and the decrease in portal pressure, like an effect of HCV eradication [169, 170].

**Role of SWE after treatment (monitoring in follow-up)**

Data regarding the usefulness of liver stiffness monitoring during antiviral therapy are scant. In the largest prospective study published (n = 91), a significant liver stiffness decrease was observed during therapy with peg-interferon and ribavirin; the decrease in liver stiffness continued after treatment only in patients who achieved SVR [171]. In the era of DAAs, it is important to remark that after successful HCV eradication, the use of pre-treatment cut-off values can impair the accuracy of TE [172, 173], and it might lead to erroneous conclusions if the SVR status is not carefully taken into account [170]. Several studies have evaluated the use of VTQ® for monitoring IFN-based antiviral therapy in HCV patients: SWE decrease or increase reflects response or no response to treatment, respectively [174 – 176]. For IFN-free antiviral therapies, no data are available.

The monitoring of cirrhotic patients after SVR will become the new standard in the era of DAAs. Although it is tempting to use SWE in this setting to observe the dynamics of liver stiffness over time, no recommendation can be made at this stage on cut-offs and the time interval to identify cirrhosis regression.

**RECOMMENDATION 19**

SWE is not recommended to monitor fibrosis changes during anti-HCV treatment (LoE 3, GoR D) [172, 173]. Strong consensus (18/0/0, 100 %)

**RECOMMENDATION 20**

LSM changes after successful anti-HCV treatment should not affect the management strategy (e.g. surveillance for HCC occurrence in patients at risk) (LoE 3, GoR D) [76]. Broad consensus (16/0/1, 94 %)
Chronic Hepatitis B (CHB)

Introduction
A large amount of evidence regarding elastography in CHB is available. There are >50 published studies. The majority of them use TE but also validate the elastographic methods of pSWE and 2D-SWE.

Fibrosis staging
The most important goal of noninvasive diagnostic tools is diagnosis of compensated cirrhosis that would benefit from treatment regardless of the transaminase level [177].

Transient elastography
Transient elastography is the most validated elastographic method for staging CHB and has similar accuracy in this clinical scenario compared to CHC [76]. Three meta-analyses confirmed the good performance of TE in CHB staging [178 – 180]. Despite LSMS showing a substantial overlap among adjacent stages of fibrosis (particularly at lower fibrosis stages), LSM may effectively identify patients with ≥F2 and F4. Recent publications confirmed previous evidence, suggesting that the AUROCs for ≥F2 vary between 0.80 and 0.90 [181, 182] with Young’s modulus cut-off values between 6.6 kPa and 8.8 kPa [182, 183]. Regarding the identification of cirrhosis (F4), recent data confirm previous evidence, with AUROCs ranging between 0.81 and 0.97 [182] and cut-off values between 9.4 and 13.4 kPa [184, 185]. A recent meta-analysis suggested that a value > 11.7 kPa should raise suspicion of cirrhosis [178]. It has been suggested that LSM cut-offs should be adapted to transaminase levels [185] since transaminase levels tend to influence LSM in CHB, and hepatitis flares are often observed in CHB. However, recent studies showed that ALT-adapted cut-offs do not influence TE diagnostic performance [186] and that the only variable associated with overestimation of F4 stage in CHB is moderate/severe necro-inflammatory activity without any direct correlation with transaminase level [187]. Interestingly, a Young’s modulus of < 5 kPa in patients with normal ALT and low serum HBV DNA levels (< 2000 IU/ml) characterize inactive HBV carriers [188, 189]. TE can be used to rule out significant fibrosis and cirrhosis in HBV inactive carriers, which is the best indication for TE in HBV.

Point shear wave elastography (pSWE)
VTQ® has the advantage of a lower failure rate and has a similar diagnostic performance as TE. The discriminative ability for staging fibrosis in CHB is good, with AUROCs for ≥F2 and F4 of 0.76 – 0.91 [184, 190] and 0.72 – 0.97 [155, 184], respectively. These findings were confirmed in a meta-analysis that included patients with several etiologies of liver disease. The analysis of data of patients with CHB showed an AUROC for ≥F2 of 0.88 and the best cut-off was 1.35 m/s and the AUROC for F4 was 0.93 and the best cut-off was 1.87 m/s [191, 192]. There is limited data about other pSWE methods. ElastPQ® has been used for staging CHB patients in four studies, with good performance for staging liver fibrosis. Further validation is required [106, 192 – 194].

2D shear wave elastography (2D-SWE)
Recently, 2D-SWE (SSI) was tested in patients with CHB and proved to have a lower failure rate than TE and at least similar performance for fibrosis staging [195]. The AUROC for ≥F2 varies between 0.85 and 0.91 [58, 196], with Young’s modulus cut-offs between 7.1 and 8.0 kPa [195, 196], while the AUROC for F4 varies between 0.92 and 0.98 [195, 197], with optimal cut-offs between 10.1 and 11.7 kPa [58, 195]. The best indications of 2D-SWE in HBV are inactive carriers to rule out significant fibrosis and cirrhosis diagnosis.

Monitoring (evaluation of) response to treatment
Under nucleoside/nucleotide analogs, liver stiffness measured by TE significantly decreases regardless of the baseline values of ALT [198, 199]. The only factors associated with decline of LSM are higher baseline LSM and HBV DNA levels [198]. The diagnostic accuracy and thresholds of liver stiffness using TE may differ in untreated and treated patients with chronic hepatitis B and C. This aspect should be taken into account when interpreting the results of elastography. Moreover, even if liver stiffness declines with antiviral treatment, it is unknown if this reflects disease regression and LSM changes should not affect management.
Dietrich CF et al. EFSUMB Guidelines and...
Ultraschall in Med 2017; 38: e16–e47

**RECOMMENDATION 25**
LSM changes under HBV treatment should not affect the management strategy (e.g. surveillance for HCC occurrence in patients at risk) (LoE 2b, GoR B) [198, 199]. Strong consensus (16/0/0, 100 %)

Prognostic relevance
Baseline LSM by TE has modest prognostic relevance with an AUROC between 0.70 and 0.73 for liver-related events [200, 201], which may be increased by adding spleen diameter and platelet count as a prediction model called LSPS (= LSM × spleen diameter/platelet count) up to a level of 0.83 [202]. TE is a good prognostic marker for HCC development that may occur without cirrhosis. A Young’s modulus of > 8 kPa might be a value indicating the need to start screening for HCC, even if a complete virological response was achieved [203, 204]. The risk of HCC is even higher if the Young’s modulus by TE is > 12 – 13 kPa, which also implies an increased risk of decompensation [181, 203 – 205].

Non-alcoholic fatty liver disease (NAFLD)
In NAFLD patients, noninvasive markers should aim at the following: a) identify the risk of NAFLD/NASH among individuals with metabolic syndrome; b) identify those with a worse prognosis; c) monitor disease progression; d) predict response to treatment.

Fibrosis staging
Transient elastography
A particular problem in NAFLD patients is the failure rate of elastography. Different studies report increased failure rates owing to increased body mass index (BMI > 30 kg/m²) or waist circumference, which may interfere with the transmission of the push impulses and the tracking ultrasound, thus preventing correct estimation of liver stiffness. A second issue in NAFLD patients is that two probes are available (M and XL probes), with the latter producing lower stiffness values than the M probe in overweight subjects. Comparing the diagnostic performance of the two probes (M and XL probes), the XL probe obtains more reliable results than the M probe in obese patients. The main limiting factors for the XL probe are a skin-to-liver capsule distance > 3.4 cm and extreme obesity (BMI > 40 kg/m²). The third issue is that F3-F4 fibrosis is the most important prognostic factor for liver-related outcomes and mortality in NAFLD patients.

TE performance is better for cirrhosis than for significant fibrosis [206, 207]. TE has a higher rate of false-positive than false-negative results and a higher negative predictive value (NPV) than positive predictive value (PPV). Therefore, the ability to diagnose bridging fibrosis or cirrhosis is insufficient for clinical decision-making [208, 209]. A systematic review of TE in patients with NAFLD involved 9 studies and 1047 patients [210]. TE was excellent in diagnosing F3 fibrosis (85 % sensitivity, 82 % specificity) and cirrhosis (92 % sensitivity, 92 % specificity), but had only moderate accuracy for F2 fibrosis (79 % sensitivity, 75 % specificity).

With the M probe, patients with steatosis > 66 % at liver biopsy had higher LSM values, which led to higher false-positive LSM results [52]. Thus, in obese patients with a high degree of steatosis, TE (using the M probe) may be less accurate in diagnosing severe fibrosis in NAFLD, and additional evaluation may be warranted to avoid overestimation of fibrosis. However, additional studies on the effects of steatosis on LSM measured with an XL probe are needed. The XL probe produces lower stiffness values than the M probe. Different cut-offs should be used [85]. With the M probe, at a Young’s modulus cut-off value of 7.9 kPa, the sensitivity, specificity, PPV and NPV for F3 or greater disease are 91 %, 75 %, 52 %, and 97 %, respectively [55].

In NAFLD patients, the best cut-off for F3 or greater disease is 7.2 kPa. With this cut-off, the NPV to exclude F3 or greater disease is 89 % (95 % CI 84 – 95 %). Cut-off values of 5.7 kPa and 9.3 kPa have 90 % sensitivity and specificity to rule out and rule in F3 disease, respectively [55].

**RECOMMENDATION 26**
TE can be used to exclude cirrhosis in NAFLD patients (LoE 2a, GoR B) [76, 210]. Broad consensus (13/0/3, 81 %)

Point shear wave elastography (pSWE)
A systematic review of 7 studies for a total of 723 patients who underwent SWS measurements with VTQ® technique to evaluate the diagnostic efficacy of pSWE in patients with NAFLD was recently published [211]. The summary sensitivity was 80.2 % for detecting significant fibrosis, which is not an appropriate endpoint.

2D shear wave elastography (2D-SWE)
There are only two studies that evaluate the performance of 2D-SWE (SSI) [57, 212]. The results are too limited to make recommendations.

Comparison of different elastographic methods for NAFLD
One study recently compared TE (using the M probe), pSWE (VTQ®) and 2D-SWE (SSI) in 291 patients with NAFLD enrolled in two different hospitals [57]. All methods showed AUROCs ≥ 0.84 for severe fibrosis and cirrhosis and had a similar performance for the diagnosis of these endpoints. The diagnostic performance of 2D-SWE was superior to that of VTQ® for the diagnosis of significant fibrosis.

Follow-up of patients
Monitoring of the progression of fibrosis is also necessary in the follow-up of these patients. Patients who achieved a ≥ 5 % weight loss at the 6-month follow-up showed a decrease in LSM by TE, independent of the changes in aminotransferase levels [208]. No data are available for pSWE and 2D-SWE.
Prediction of liver-related complications

A recent study supports the use of TE for risk stratification and prediction of clinical endpoints [213]. For pSWE and 2D-SWE data on this aspect are lacking and no evidence-based recommendation can be given.

Alcoholic liver disease (ALD)

Fibrosis staging

Transient elastography in patients with prior or current chronic alcohol abuse can distinguish absent and mild fibrosis (F0 – 1) from severe fibrosis and cirrhosis, but similar to other etiologies, there is no evidence to suggest that it can differentiate absent and mild fibrosis from significant fibrosis. [214] Additionally, in the eight published single etiology studies on TE for staging liver fibrosis, there is no consensus regarding optimal Young’s modulus cut-off values for significant fibrosis (≥F2), severe fibrosis (≥F3) or cirrhosis (≥F4) [49, 56, 215 – 220]. The optimal cut-off values range from 7.8 [216] to 9.6 [56] kPa for significant fibrosis, from 8.0 [49] to 17.0 [218] kPa for severe fibrosis and from 12.5 [49] to 22.7 [215] kPa for cirrhosis. The considerable discrepancy between cut-off values in individual studies is likely a consequence of differences in fibrosis stage prevalence with overrepresentation of cirrhotic patients, place of recruitment, and whether patients with alcoholic hepatitis or decompensated disease were excluded.

TE is more suited to rule out than rule in cirrhosis. At a Young’s modulus of 12.5 kPa, TE may rule out cirrhosis with a negative likelihood ratio of 0.07 if the disease prevalence is 50 % or lower. [214].

**RECOMMENDATION 27**

TE can be used to exclude cirrhosis in patients with alcoholic liver disease, provided that acute alcoholic hepatitis is not present (LoE 2b, GoR B) [56, 214, 219, 220]. Strong consensus (15/0/0, 100 %)

Prognostication of alcoholic liver cirrhosis

There is scant evidence to suggest a role for ultrasound elastography for determining the prognosis and for the monitoring of patients with alcoholic liver disease [221], or for TE to predict esophageal varices [222] and the hepatic venous pressure gradient in patients with alcoholic cirrhosis [221].

Timing of liver stiffness measurements with regard to alcohol abstinence

In patients undergoing alcohol detoxification, 0.5 to 4 weeks of abstinence causes a clinically significant decrease in TE [49 – 51, 223]. However, the decrease is associated with a normalization of transaminases, bilirubin, alkaline phosphatase and/or gamma-glutamyltransferase. It is, therefore, unclear whether alcohol alone or alcohol-induced hepatitis and cholangiocytosis damage cause the increase in liver stiffness. [220]. One study suggests that TE is accurate for staging in patients with ongoing alcohol abuse but normal gamma-glutamyltransferase [56], while another study suggests that AUC for the diagnosis of cirrhosis in alcoholic patients diminishes when AST is above 100 – 150 U/L [220].

Screening the general population or high-risk groups in primary care

Ultrasound elastography for systematic screening of high-risk populations in primary care for alcoholic liver disease was performed in one study. TE was offered to primary care patients with an AST:ALT ratio ≥ 0.8 [224]. However, this study did not include biopsy confirmation in patients with elevated TE. In a diagnostic study, TE had excellent diagnostic accuracy for significant fibrosis and cirrhosis in a subgroup of 71 patients recruited from primary alcohol rehabilitation centers [56].

With a prevalence of cirrhosis of 2 – 4 % in a background population of at-risk individuals [225], the PPV of TE should be considered low regardless of cut-off values.

Point shear wave elastography (pSWE) and 2D-SWE

There is only one study to support the use of 2D-SWE (SSI) for assessing alcoholic liver fibrosis [56]. There are three small studies on the use of pSWE [226 – 228], two of which report diagnostic accuracies and test probabilities. The results are consistent regarding diagnostic accuracy, which suggests that VTQ® may be used to rule out severe fibrosis and cirrhosis. The results regarding cut-off values are, however, inconsistent. Therefore, there is insufficient evidence to make recommendations for using VTQ® to distinguish absent and mild fibrosis (F0 – 1) from significant or severe fibrosis and cirrhosis. Thus, there is still insufficient evidence to evaluate the role of pSWE or 2D-SWE in alcoholic liver disease.

Cholestatic liver disease and autoimmune hepatitis (AIH)

Risk stratification is a major need in patients with chronic cholestatic diseases in order to allow personalized management and the selection of candidates for clinical trials of new drugs. Studies of liver stiffness as a surrogate of liver fibrosis and prognosis focusing on cholestatic liver disease (primary biliary cholangitis-PBC, primary sclerosing cholangitis-PSC) or AIH are scarce. Most available data focus on TE [229 – 231].

Transient elastography

Transient elastography is currently considered one of the best surrogates of fibrosis in PBC. High baseline or increasing values over time indicate a worse outcome in this population [191, 229]. Liver stiffness was investigated in 73 patients with PSC, regularly undergoing clinical and elastographic follow-up [232]. LSMs were able to differentiate severe vs. non-severe fibrosis with a high discriminative accuracy for cirrhosis (AUROC 0.88). There was high reproducibility between two operators. Higher baseline LSM and an increase of LSM over time were associated with adverse outcome such as death, liver transplantation, ascites, hepatic encephalopathy, gastrointestinal bleeding, or HCC [232]. Dilatation of the intrahepatic biliary system due to a dominant stricture should
be excluded in PSC before interpreting the LSMs. Cholestasis increases liver stiffness independent of liver fibrosis.

Data regarding the pediatric population with biliary atresia suggest that liver (and spleen) elastography could be a valuable tool to predict outcomes before surgery, and might be used after the Kasai operation to monitor liver disease and portal hypertension [233].

Due to the limited evidence, no recommendation can be given.

Point shear wave elastography (pSWE)

VTQ® was initially performed in 9 patients with AIH, PBC and PSC having higher shear wave velocities than healthy volunteers [234]. In total two studies dealt with VTQ® in AIH and primary biliary cholangitis. In 15 patients with treated AIH, VTQ® could differentiate between the absence of fibrosis and significant fibrosis [235]. SWS assessed by VTQ® showed good diagnostic accuracy for detecting cirrhosis (AUROC 0.91) in 61 patients with primary biliary cholangitis [236].

2D shear wave elastography (2D-SWE)

Data regarding 2D-SWE in AIH, PBC and PSC are not available. Liver stiffness assessment helped in differentiating between biliary atresia and neonatal hepatitis in one study [237].

Due to the paucity of data, no recommendation can be given.

Portal hypertension

Transient elastography

In patients with compensated advanced chronic liver disease/cirrhosis LSM correlates with the hepatic venous pressure gradient (HVPG). Even though the correlation between the two does not allow for an accurate estimate of the exact HVPG value (range: 0.59–0.70), the discriminative ability of liver stiffness for the presence of clinically significant portal hypertension (CSPH, defined as HVPG ≥ 10 mmHg, threshold for the appearance of complications) is very high, with a summary AUROC of 0.93 in a recent meta-analysis [238]. However, it should be emphasized that most of the patients included in the studies concerning HVPG had viral or alcoholic cirrhosis, and evidence regarding other etiologies is limited. In viral cirrhosis, Young’s modulus values of > 20–25 kPa are highly specific for CSPH, and values of > 21 kPa predict the onset of a first clinical decompensation with an accuracy similar to that of HVPG > 10 mmHg [239].

RECOMMENDATION 28

LSM with TE is useful to identify patients with a high likelihood of having clinically significant portal hypertension (HVPG ≥ 10 mmHg) (LoE 2b, GoR B) [238, 239]. Strong consensus (15/0/0, 100 %)

The accuracy of LSM in predicting the presence and size of gastroesophageal varices has been the subject of several studies. Despite the fact that it is currently the best single noninvasive predictor in this field with summary AUROCs of 0.84 for esophageal varices (EV) and 0.78 for large EV in a recent meta-analysis [238], the cut-offs vary widely among the studies, and the accuracy is not sufficient to replace endoscopy. The accuracy of LSM for the diagnosis of CSPH and varices improves if it is combined with platelet count and spleen size [240, 241]. Recent data indicates that if a combination of a Young’s modulus value of < 20 kPa and a platelet count of > 150 G/L is used, varices needing treatment can be ruled out with a high accuracy (< 5 % of patients missed) [242, 243], and endoscopy can be safely avoided [244].

Point shear wave elastography (pSWE)

Point SWE (VTQ®) has been used in three studies addressing the diagnosis of CSPH [245–247], and showed excellent applicability and very good diagnostic accuracy (AUROCs 0.82–0.90). VTQ® has been used in a few studies addressing the diagnosis and severity of esophageal varices. SWS was higher in patients with esophageal varices of any size, and was even higher in patients with large varices [245, 246]. However, reliable cut-offs are not available yet. No strong recommendation regarding the cut-offs to be used can be made due to the limited evidence.

2D shear wave elastography (2D-SWE)

2D-SWE (SSI) has been tested for the diagnosis of CSPH in 4 studies and a further small series [27, 248–251]. The accuracy of the method was reliable in all of the published studies (AUROC 0.80–0.92).

Two studies performed a head-to-head comparison between LSM by TE and 2D-SWE [27, 248]. TE was less applicable, and both techniques showed similar accuracy for the diagnosis of CSPH.

LSM by 2D-SWE is higher in patients with esophageal varices of any size and is further increased in patients with large varices. However, reliable cut-offs are not available yet. No strong recommendation regarding the cut-offs for 2D-SWE can be given, and further evidence is needed.

RECOMMENDATION 29

Liver stiffness using TE combined with platelet count is useful to rule out varices requiring treatment (LoE 2b, GoR B) [244]. Although preliminary results are encouraging, there is insufficient evidence to recommend pSWE and 2D-SWE in this setting. Broad consensus (13/0/1, 93 %)

Reimbursement

Social health care systems and noninvasive liver fibrosis

The data on the cost-effectiveness of liver elastography are limited to TE, are mainly derived from UK and Canadian health care systems and depend on local health care economic conditions. Although TE is generally accepted as a noninvasive alternative for the staging of liver fibrosis, which helps to avoid a liver biopsy [252], it does not represent the most effective approach
with respect to resources in common clinical scenarios [253]. TE is considered cost-effective for annual cirrhosis screening in patients infected with HCV [254], with the most cost-effective treatment scenario for patients infected with HCV not depending on the fibrosis category ("treating all") [253]. However, in countries that cannot afford the "treating all" scenario, TE is the accepted method to identify patients deserving access to treatment based on the greater severity of fibrosis, making it competitive in cost-effectiveness in comparison to biopsy to screen large populations of patients. Inaccurate noninvasive fibrosis staging may limit the cost-saving value of elastography [255]. Weighing the diagnostic limitations and costs of elastography against laboratory-based fibrosis scores, widely available and cheap scoring systems (e.g. AST-to-platelet ratio index) are the recommended method in regions with limited health care resources [256]. Considering the emerging prevalence of non-alcoholic fatty liver disease (NAFLD) in industrialized countries, TE can represent a cost-effective approach for the screening of patients at risk for disease progression and complications [257, 258]. However, a formal study assessing hard outcomes such as survival, hospitalization rates, etc., has not been performed recently and is unlikely to be performed given the vast number of patients required.

Current reimbursement in European countries

Reimbursement is often limited to conventional ultrasound, which includes shear wave-based liver stiffness assessment. For example, no reimbursement is available within the compulsory health insurance for TE (Fibroscan®) in Germany. In private practice, 50 – 150 € is usually charged for a Fibroscan® examination. In many other European countries, no specific reimbursement code is established. Where ultrasound elastography is reimbursed (e.g. some regions in Italy), it is accepted under the code of liver ultrasonography. However, the same exam cannot be reimbursed twice on the same day. Therefore, conventional ultrasonography and ultrasound elastography of the liver cannot be reimbursed together at the same appointment, thus limiting the cost-effectiveness of the technique. In Romania, TE is reimbursed in the daily hospitalization system with the possibility for future reimbursement of pSWE and 2D-SWE. In private practice, the cost of TE is between 50 and 120 €.

Outlook: what should be done?

One future possibility in liver elastography is to compare the accuracy of different technologies (TE, pSWE and 2D-SWE) in a large cohort of patients with division of the etiologies of liver diseases. Probably, as indicated by earlier comparative studies, we will have some results showing the superiority of new technologies in comparison with established ones (point or 2D-SWE vs TE). Robust cut-off values must be established for each of the different systems and diseases. The combination of elastographic methods using ultrasound with biological tests may improve the accuracy of the evaluation of the patients.

Conflict of interest

Some authors declare conflicts of interest, which are available from the publisher.

Acknowledgement

The authors thank the EFSUMB secretary Lynne Rudd for her never-ending support of the EFSUMB. We also thank the following companies for funding a consensus meeting of the authors held in London in July 2016, at which we agreed the recommendations made in this paper: bk/Ultrasound, Echosens, Esatec SPA, GE Healthcare, Hitachi Medical Systems, Philips Healthcare, Shenzhen Mindray Bio-medical Electronics Co., Ltd, Siemens Healthineers, SuperSonic and Toshiba Medical. Representatives of these companies were in attendance at this meeting, to assist with product technical information, but did not take part in forming the manuscript or the recommendations.

References


[56] 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 – 133


[206] Petta S, Vanni E, Bugianesi E et al. The combination of liver stiffness measurement and NAFLD fibrosis score improves the noninvasive diagnostic accuracy for severe liver fibrosis in patients with nonalcoholic fatty liver disease. Liver Int 2015; 35: 1566 – 1573


[232] Corpechot C, Gaujar F, El Naggar A 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: 970 – 979 quiz e915–e916


[243] Ding NS, Nguyen T, Iser DM et al. Liver stiffness plus platelet count can be used to exclude high-risk esophageal varices. Liver Int 2016; 36: 240 – 245


