Digestive and Liver Disease 49 (2017) 802–808

Contents lists available at ScienceDirect

Digestive and Liver Disease

journal homepage: www.elsevier.com/locate/dld

Liver, Pancreas and Biliary Tract

Differences in liver stiffness values obtained with new ultrasound elastography machines and Fibroscan: A comparative study

Fabio Piscaglia[a,*], Veronica Salvatore[a], Lorenzo Mulazzani[a], Vito Cantisani[b], Antonio Colecchia[a], Roberto Di Donato[a], Cristina Felicani[c], Alessia Ferrarini[a], Nesrine Gamal[a], Valentina Grasso[c], Giovanni Marasco[a], Elena Mazzotta[c], Federico Ravaioli[a], Giacomo Ruggieri[d], Ilaria Serio[a], Joules Fabrice Sitouok Nkamgho[a], Carla Serra[c], Davide Festi[a], Cosima Schiavone[e], Luigi Bolondi[a]

[a] Department of Medical and Surgical Sciences, University of Bologna, Italy
[b] Department of Radiological Sciences, Policlinico Umberto I, University La Sapienza, Rome, Italy
[c] Unit of Internal Medicine and Organ Insufficiency, S. Orsola-Malpighi Hospital, Bologna, Italy
[d] Division of Infectious Disease, G. d'Annunzio University, Chieti, Italy
[e] Unit of Internistic Ultrasound, Department of Medicine and Science of Aging, G. d'Annunzio University, Chieti, Italy

ARTICLE INFO

Article history:
Received 26 December 2016
Received in revised form 17 February 2017
Accepted 2 March 2017
Available online 10 March 2017

Keywords:
Cirrhosis
Fibroscan
Liver fibrosis
Shear wave elastography
Transient elastography

ABSTRACT

Background and aims: Whether Fibroscan thresholds can be immediately adopted for none, some or all other shear wave elastography techniques has not been tested. The aim of the present study was to test the concordance of the findings obtained from 7 of the most recent ultrasound elastography machines with respect to Fibroscan.
Methods: Sixteen hepatitis C virus-related patients with fibrosis ≥2 and having reliable results at Fibroscan were investigated in two intercostal spaces using 7 different elastography machines. Coefficients of both precision (an index of data dispersion) and accuracy (an index of bias correction factors expressing different magnitudes of changes in comparison to the reference) were calculated.
Results: Median stiffness values differed among the different machines as did coefficients of both precision (range 0.54–0.72) and accuracy (range 0.28–0.87). When the average of the measurements of two intercostal spaces was considered, coefficients of precision significantly increased with all machines (range 0.72–0.90) whereas of accuracy improved more scatteredly and by a smaller degree (range 0.40–0.99).
Conclusions: The present results showed only moderate concordance of the majority of elastography machines with the Fibroscan results, preventing the possibility of the immediate universal adoption of Fibroscan thresholds for defining liver fibrosis staging for all new machines.

1. Introduction

Fibroscan (Echosens), introduced in 2003 [1], was the first machine which used transient elastography to measure liver stiffness, thereby non-invasively predicting the degree of liver fibrosis. Its results, measured in kiloPascal (kPa), have become a sort of immediate reference for many hepatologists worldwide for categorizing patients with chronic liver disease.

In 2008, a new modality was introduced, called Acoustic Radiation Force Impulse (ARFI) quantification, classified by the European federation of societies for ultrasound in medicine and biology (EFSUMB) [2] as point shear wave elastography (pSWE) since the speed of the shear wave was measured in a small region [3]. Notably, for the first time, quantitative elastography was embedded in a conventional ultrasound machine. In 2012, a real time SWE technique for liver stiffness quantification, implemented in Supersonic Imagine by Aixplorer, also arrived on the market [4,5], providing bidimensional elastography information with a completely new technical solution (hence the term real time 2D SWE) [2,6].

Subsequently, nearly all ultrasound manufacturers have come to implement some SWE modalities for liver stiffness measurement in their most recent machines, which have therefore become rapidly available in real life practice. However, the lack of sufficient sup-

* Corresponding author at: Unit of Internal Medicine, Hospital S. Orsola-Malpighi, via Albertoni 15, 40138, Bologna, Italy. Fax: +39 051 636 2725.
E-mail address: fabio.piscaglia@unibo.it (F. Piscaglia).

http://dx.doi.org/10.1016/j.dld.2017.03.001
1590-8658/

porting evidence for the latter machines, partly due to their very recent release and partly due to the substantial decrease in the number of liver biopsies in current clinical practice [7], has raised some relevant questions, the main one being the applicability of the stiffness values adopted for staging chronic liver diseases with Fibroscan to the other SWE technologies.

The comparison of large series of patients undergoing liver stiffness measurement with two or three different techniques cannot correctly answer this question since individual differences are often not reported and could be due either to the different capacities of the machines [8–10] or to the choice of different intercostal spaces when using the various modalities, since the fibrotic involvement of chronic liver disease is heterogeneous [11].

The aim of the present study was to compare liver stiffness findings, acquired with the vast majority of the techniques available on the market as of 2016, with the results of Fibroscan in a series of patients with hepatitis C virus (HCV)-related liver disease, to test whether the values obtained with Fibroscan are identically produced by other machines and to verify the technical parameters influencing the results.

2. Patients and methods

2.1. Patients

The number of patients was planned so as to be able to practically investigate them within two days in which all the scheduled machines were present together in the same location with their expert operators and to keep each patient in the research facility for no more than half a day on a single occasion (see the study protocol below) in keeping with the Ethics Committee request. The maximum number of patients per day was calculated to be 11, leading to a maximum of 22 study patients.

The patients enrolled were selected from those attending the outpatient liver clinics of S. Orsola-Malpighi University Hospital. The inclusion criteria were: age 18–75, HCV-related liver disease, METAVIR [12] score ≥2 (before possible treatment with antivirals) either estimated by Fibroscan or histologically confirmed. The patients were selected by the clinicians in charge, who also recorded the clinical and laboratory data. The exclusion criteria were: the presence of hepatocellular carcinoma, orthotopic liver transplantation, biliary obstruction or compromised performance status. The protocol was approved by the Ethics Committee of S. Orsola-Malpighi hospital (id: 44/2016/U/SPER) and conformed to the ethical guidelines of the 1975 Declaration of Helsinki (6th revision, 2008); written informed consent was obtained from each patient.

Achieving reliable results with Fibroscan on the study day (Success Rate (SR) ≥60% and interquartile range (IQR)/median <0.30) [13] was considered to be an additional inclusion criterion since Fibroscan was held to be the reference standard for this study and, hence, only totally reliable results were acceptable. Out of the 21 patients enrolled, 5 had to be excluded due to the fact that reliable results were not obtained, leaving 16 patients as the study group.

2.2. Ultrasound elastography machines

In addition to Transient Elastography with Fibroscan, the following 7 ultrasound machines were utilized (all transducers and software versions are reported in round brackets, elastography class in squared brackets). Please note that not all machines or software were the most recently released by each manufacturer because the study was based on machines already available in the hospital.

- Aixplorer Supersonic (convex probe XC6-1, software version 10.0.0.1815) [real time 2D SWE].
- Esaote MyLab Twice (convex probe CA451, software version EVO 13.0 release 12.11) [pSWE].
- GE Logiq E9 (convex probe C1-6, software version R5 revision 1.0) [2D single shot SWE].
- Hitachi Arietta V70 version 3.0.1 (convex probe EUP-C532, software version 00-3.0.1) [pSWE].
- Philips iU22 ELASTPQ (convex probe C1-5, software version 6.3.2.2) [pSWE].
- Samsung RS80A Ugeo (convex probe CA1-7, software version 2.00.03.0629) [pSWE].
- Siemens, Acuson S2000 version VD10A (convex probe 6C1 HD) [pSWE].

Each machine was placed in a separate space from all others so that each operator was not aware of the results of the other machines during data acquisition. The operators selected were expert for their specific machine with ≥200 exams already performed.

2.3. Study protocal

Patients were asked to lie in the supine position with their right arm raised above the head. First, the right intercostal spaces were consecutively numbered in writing on all patients by one single investigator. Both the width of the intercostal space and the distance from the skin to the liver capsule were also measured for each space.

Each ultrasound elastography station was attended by the elastography operator and by an assistant. The operator was always kept blind to all the clinical and laboratory information.

First, the Fibroscan operator chose the intercostal space apparently most suitable for the Fibroscan measurement in each patient. The number of this Reference Intercostal Space (RIS) was recorded on a card and put into an envelope kept at the station. For each station, the plan was to take two sets of ten measurements in two intercostal spaces, one of which had to correspond to the RIS. The choice of the second intercostal space (called Alternative Intercostal Space – AIS) was left to the choice of the operator. Whatever the technique, the probe was applied perpendicularly to the skin and the region of interest (ROI) was placed at a distance of 3–6 cm from the skin and more than 1 cm below the liver capsule, being careful to avoid large vessels or evidently thick fibrotic bands while patients suspended breathing.

The procedure to fulfill the above requirements was as follows: the operator identified what he/she believed to be the best investigation site and, after leaving the patient relaxed in the supine position for at least 5 min, started the evaluation with the aim of collecting 10 valid measurements obtained within a maximum of 20 attempts. The number of total attempts and the values of the successful measurements were recorded by the assistant. After completion of the first set of measurements, the operator identified what he/she believed to be the second best intercostal space for the investigation. At this point, the assistant opened the envelope (kept hidden until that point so as not to inform the operator regarding the space already chosen for Fibroscan) and confirmed if at least one of the two preferred spaces was the RIS. If this was not the case, the operator was informed of the number of the space corresponding to the RIS in order to take the second set of measurements. The investigations using Fibroscan were carried out in the RIS and in the second best space according to the operator's choice. Non valid measurements were automatically rejected by the pSWE machines whereas the operator did not carry out the quantification analysis when the signal in the ROI was

absent or only minimally and/or scatteredly present in cases of 2D SWE.

After all patients had completed the first round of measurements, a rotation took place so that each patient moved to the next station separated from all the others. The rotations took place until all patients had undergone examination with all 7 machines and with Fibroscan, including some pauses for rest.

3. Statistical analysis

The values are expressed in kPa for all the techniques. Since the results using the Siemens machine are provided only as shear wave speed in m/s, they were converted to kPa (the unit of measurement of Young's modulus) according to the formula: kPa = 3(speed in m/s) [2].

Descriptive statistics were carried out for the demographic, anthropometric, clinical and laboratory parameters of the patients. Data are expressed as medians and ranges.

The SRs with different machines were compared using the Fisher's exact test.

The SWE techniques were compared with Fibroscan adopting the tests (namely, the Pearson coefficient of precision and accuracy as an index of the bias correction factor) included in Lin's Concordance Correlation Coefficient (CCC) [14]. The CCC consists of the product of the Pearson coefficient, an index of precision, and a bias correction factor which approximates how far the best fit line deviates from the 45° line, an index of accuracy. In keeping with the aims of the present analysis, precision and accuracy were reported separately.

Simply stated, the Pearson coefficient of precision tends to express how values are scattered from the reference line. Lower scattering indicates greater precision (Fig. 1a). Instead, the highest accuracy indicates that the regression line is close to 45° and that it ideally intercepts the crossing of the two axes at zero. Conversely, in the case of low accuracy, the regression line is far from 45° (in terms of either intercept and/or slope) and, thus, changes in the study method values tend to be disproportionate (inconsistently higher or lower) to the reference method values (Fig. 1b). On the other hand, in Fig. 1a, it can be observed that the same mean value could be produced with either perfect precision (if all data are aligned along the reference line) or with low precision (scattered dots). Thus, reporting data simply as mean (or median) values does not give enough information in this setting. A test method could be very precise but not accurate (different slope or intercept), indicating that the overall mean/median value of the test and reference methods could be nearly identical, but the test method could not be directly used as if it were the reference method.

We would therefore like to draw the attention to the fact that for a new method to potentially adopt the Fibroscan thresholds for staging without obtaining its own thresholds with specifically investigated patients populations with histological reference, a very high accuracy would be a requisite. Low accuracy does not indicate a poor or modest performance (suggested instead by low precision), but only that specific thresholds must be identified for that machine, and that Fibroscan thresholds for fibrosis staging cannot be directly adopted.

Differences between the two groups were tested using the Mann–Whitney U-test given the limited number of subjects. The Spearman test was used for correlations. All the analyses were carried out using Medcalc Software bvba, Ostend, Belgium, version 14.8.1.

Table 1
Patients characteristics.

	Median or absolute count	Range
Male/female	10/6	
Body mass index	26.0	20.1–31.8
Thoracic circumference (cm)	92.5	80–104
AST (U/L)	24	13–92
ALT (U/L)	19	5–102
GGT (U/L)	40	11–85
Alkaline phosphatase (U/L)	81	39–273
Total bilirubin (mg/dL)	0.70	0.26–2.13
Conjugated bilirubin (mg/dL)	0.22	0.05–0.51
INR	1.15	1.01–1.36
Albumin (g/dL)	4.1	3.2–4.7
Creatinine (mg/dL)	0.79	0.54–2.10
Platelet count (number/mcL)	128	33–451
Liver elastography measurements		
-Fibroscan (kPa)	13.3	7.3–37.4
-Aixplorer Supersonic (kPa)	12.4	5.3–25.2
-Esaote Twice (kPa)	11.2	5.9–33.7
-GE E9 (kPa)	9.4	4.4–15.6
-Hitachi Arietta V70 (kPa)	8.7	4.7–18
-Philips iu22 (kPa)	6.6	3.7–14.5
-Samsung RS80A Ugeo (kPa)	9.4	4.6–21.7
-Siemens Acuson S2000 (kPa)	11.0	2.5–23.1

Data are reported as medians and ranges or as absolute counts. AST: Aspartate transaminase, ALT: alanine transaminase, GGT: gamma gluthamyl synthetase, kPa: kilopascal, INR: International Normalized ratio.
All measurements were taken in the same reference intercostal space with the different machines. More detailed information regarding the machines, tranducers and working software are reported in the text of the manuscript.

4. Results

A total of 16 patients were included in the analysis, the characteristics of whom are reported in Table 1. Twelve of the 16 patients had achieved sustained virological response (SVR) to direct antivirals whereas 4 patients were untreated.

Considering that each of the 16 patients underwent investigation with 7 different operators using 7 different ultrasound machines, there were 112 occasions in which the best intercostal space to perform the elastography investigation was chosen. The best intercostal space chosen coincided with the best space independently selected for Fibroscan in 53 of 112 occasions (47.3%).

The 5 patients excluded from the analysis because of unreliable Fibroscan results had a slightly, although not significantly higher, median Body Mass Index (BMI) than patients with reliable findings (BMI 24.3, range 22.1–27.6 and 26.0, range 20.1–31.8, respectively in included and excluded patients; p = not significant (n.s.)).

4.1. Comparison of liver elastography techniques with Fibroscan.

The median stiffness results of the various machines differed, by a greater or lesser extent, from Fibroscan (Table 1). Briefly, real time 2D SWE with Supersonic Imagine showed average values similar to Fibroscan whereas all pSWE techniques and single impulse 2D SWE by GE showed (at different degrees) lower values. In this series of patients, GE, Philips and Hitachi did not record any value higher than 18 kPa (Table 1).

The comparison of the various elastography methods with Fibroscan showed different correlations in terms of both precision and accuracy (Table 2; Figs 2 and 3). This indicated that the magnitude of stiffness captured by the various elastography methods might significantly differ from that obtained with Fibroscan in the same patient, even in the same intercostal space.

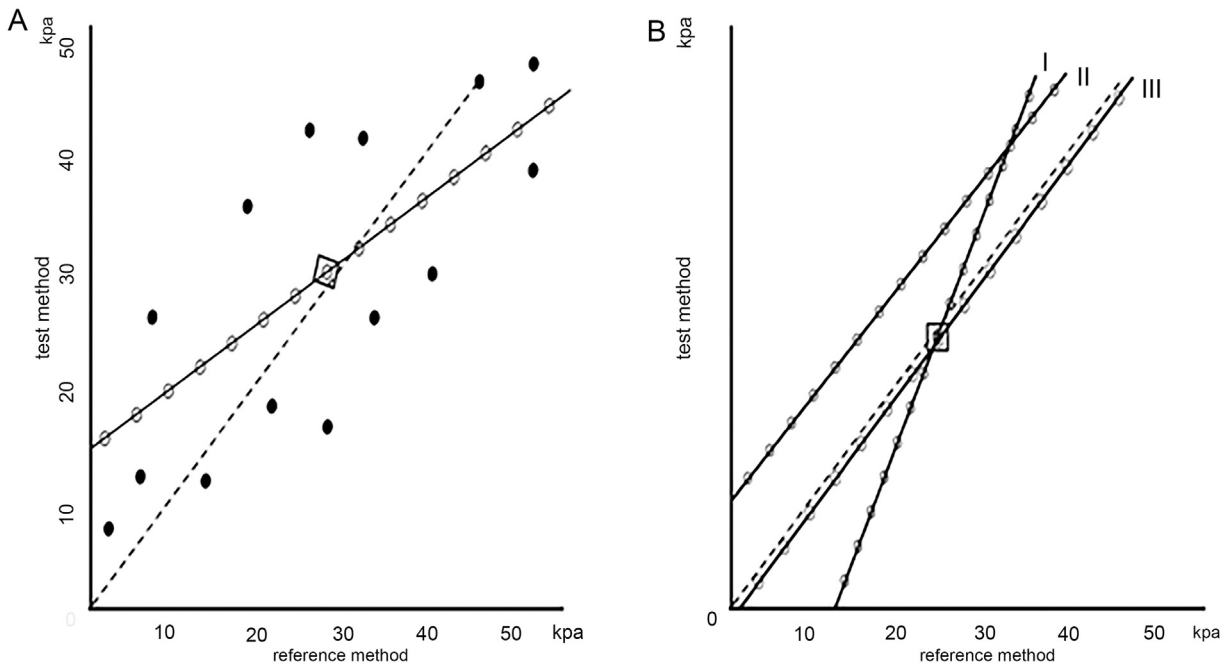

Fig. 1. Demo plots to better illustrate Lin's analysis.
a: Same accuracy but different precision: two different samples (circles and dots) are plotted on the graph; the regression line (continuous line) and the median (square) are exactly the same for the two samples, but the two sets of samples differ overtly when considering the scattering of data from the regression line. As a consequence, notwithstanding the same accuracy (indicated by the regression line), Lin's analysis is able to demonstrate a different concordance of the two sample sets by including the precision (indicated by the Pearson's coefficient) which would be higher in the circles than in the dots.
b: Same precision but different accuracy: three different sets of data are represented with circles; each set of data is perfectly aligned to the respective regression line (I, II and III); thus, Pearson's analysis (an index of precision) would give the same result in the three samples in spite of the different slopes and intercepts of the regression lines. Conversely, accuracy would be very high for set III, and lower for sets I and II. The box indicates the median value, which is identical in the I and the III lines, indicating that the median value is not sufficient to confirm that two techniques produce the same data.

Table 2
Concordance analysis of the various SWE techniques versus Fibroscan.

Ultrasound Elasto machine	Median (kPa)	Precision coefficient	Accuracy coefficient	Median (kPa)	Precision coefficient	Accuracy coefficient
	Measurements taken in the reference intercostal space			Measurements taken in two intercostal spaces		
Aixplorer Supersonic	12.4	0.727	0.871	11.8	0.900	0.992
Esaote Twice	11.2	0.542	0.842	10.4	0.782	0.843
GE E9	9.4	0.540	0.467	9.5	0.717	0.492
Hitachi Arietta V70	8.7	0.569	0.464	9.0	0.720	0.497
Philips iu22	6.6	0.637	0.284	6.8	0.754	0.405
Samsung RS80A Ugeo	9.4	0.725	0.641	8.7	0.798	0.731
Siemens Acuson S2000	11.0	0.652	0.834	9.4	0.729	0.911
	Patients with skin-capsule distance <2 cm (n = 9)			Patients with skin-capsule distance ≥2 cm (n = 7)		
Aixplorer Supersonic	11.5	0.912	0.935	13.4	0.667	0.936
Esaote Twice	12	0.962	0.991	8.25	−0.567	0.386
GE E9	9.5	0.703	0.652	9.2	0.563	0.365
Hitachi Arietta V70	10.2	0.947	0.595	7.4	0.437	0.420
Philips iu22	6.8	0.943	0.426	6.4	0.386	0.142
Samsung RS80A Ugeo	10.3	0.949	0.730	8.5	0.601	0.748
Siemens Acuson S2000	9.4	0.899	0.907	14.1	0.472	0.879

Liver stiffness was sampled in the reference intercostal space (10 measurements) and in two intercostal spaces (reference intercostal space and alternative intercostal space; 20 measurements). The precision results of all US SWE techniques were statistically correlated with the Fibroscan results (all $p < 0.05$ using 10 measurements and $p < 0.002$ using 20 measurements).
In the second part of the table, patients were subgrouped according to a skin to liver capsule distance of 2 cm.

4.2. The role of the number of measurements on liver stiffness measurement concordance

In consideration of the fact that the size of the tissue sampled varied considerably between the different techniques, that the choice of where to put the ROI or where to carry out the sampling with Fibroscan was left up to each individual operator and that the liver fibrotic involvement was expected to be heterogeneous, it was decided to carry out an additional analysis.

First, the correlation between stiffness measurements taken in the RIS or AIS was tested. The results showed good, but not perfect, reproducibility (Table 3). Interestingly, the precision for Fibroscan was suboptimal ($r = 0.70$), indicating that an intrinsic variability in stiffness assessment in different liver sites exists also with the reference method, at least in our HCV-treated patients. Conversely, as expected, the accuracy was very high for each technique when comparing the results in two different intercostal spaces (almost all >0.90).

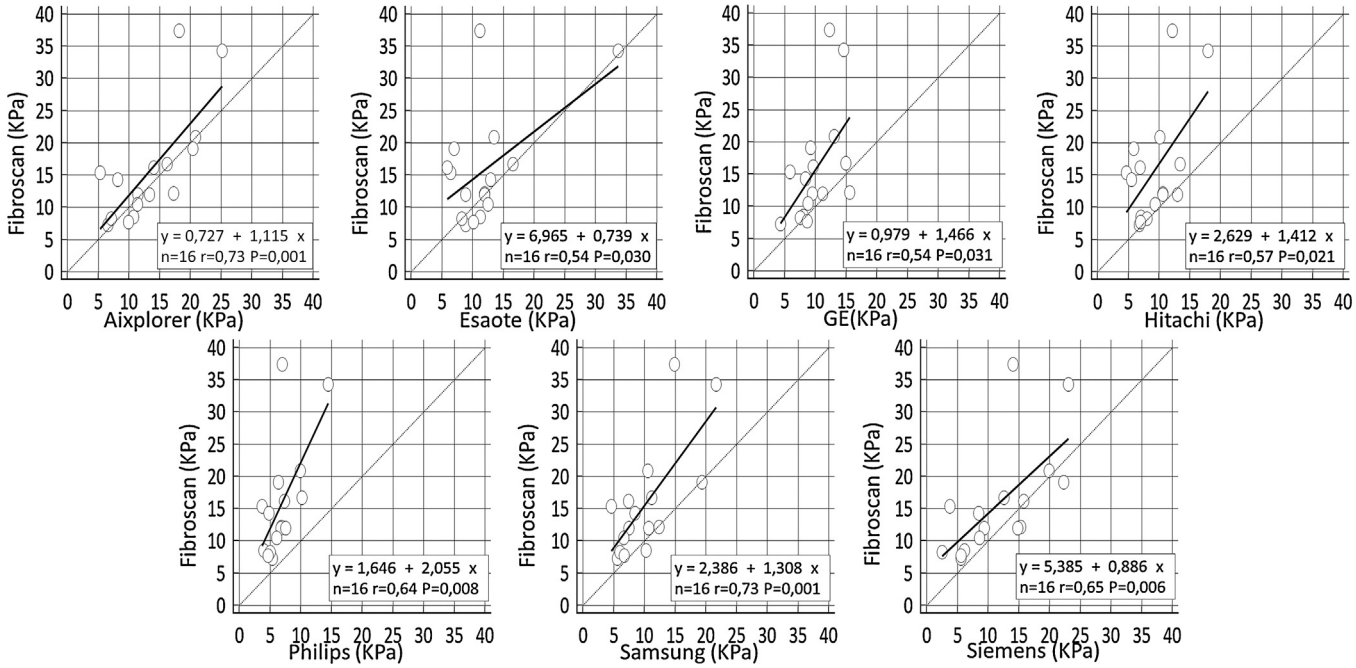

Fig. 2. Comparison of each SWE technique versus Fibroscan in determining liver stiffness sampled in the reference intercostal space.
Liver stiffness sampled with each SWE machine (x axis) is plotted against the respective Fibroscan values (y axis). The regression line (thick) and the reference line (y = x; dashed) are drawn on each plot. Reports of the regression equation ($y = ax + by$), sample size (n), Pearson's coefficient (r) and p value (P) for each comparison are reported in the box for each diagram.

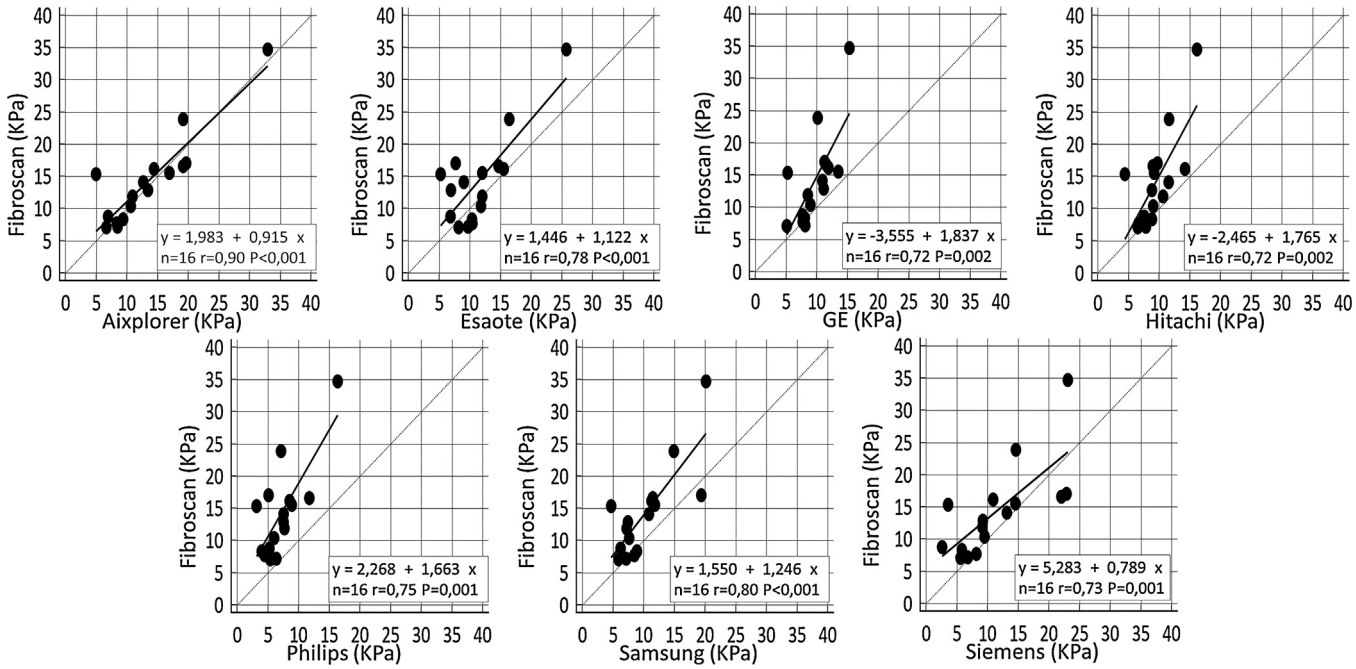

Fig. 3. Comparison of each SWE technique versus Fibroscan in determining liver stiffness from two intercostal spaces.
Scatter plots obtained considering liver stiffness as the median of all 20 samplings achieved in the reference (10 measurements) and alternative (10 measurements) intercostal spaces. The results of each SWE machine (x axis) are plotted against the respective Fibroscan values (y axis). The regression line (thick) and the reference line (y = x; dashed) are drawn on each plot. The regression equation ($y = ax + by$), sample size (n), Pearson's coefficient (r) and p value (P) are reported in the box included with each comparison.

To take into account heterogeneity of liver disease involvement and hence variability in sampling the 10 measurements in RIS were averaged with the 10 in AIS. Not necessarily were the AISs the same for all patients and machines, according to the protocol, but a more comprehensive averaging of the stiffness measurement of the right liver lobe would be expected. Indeed, a notable increase in concordance between all elastography technologies and the Fibroscan results (which was also averaged over 20 measurements) became apparent (Table 2). All the devices had good precision and some also had concurrent good accuracy, thus confirming a statistically significant correlation with Fibroscan for all ultrasound SWE methods.

Table 3
Concordance analysis for liver stiffness measurement taken in two different intercostal spaces and success rate.

Ultrasound Elasto machine	Concordance analysis		Success rate					
	Precision	Accuracy	All patients	Patients with Fibroscan < 10 kPa (n = 6)		Patients with Fibroscan ≥ 10 kPa (n = 10)		p
			S/A	S/A	SR	S/A	SR	
Fibroscan	0.730	0.960	391/507	115/143	0.804	276/364	0.758	0.710
Aixplorer Supersonic	0.920	0.962	420/425	120/120	1.000	300/305	0.984	0.939
Esaote Twice	0.666	0.911	354/686	113/174	0.649	241/512	0.471	0.028
GE E9	0.717	0.984	420/420	120/120	1.000	300/300	1.000	1.000
Hitachi Arietta V70	0.652	0.972	349/585	100/144	0.694	249/441	0.565	0.191
Philips iu22	0.911	0.861	420/519	120/142	0.845	300/377	0.796	0.715
Samsung RS80A Ugeo	0.935	0.987	312/711	106/173	0.613	206/538	0.383	0.002
Siemens Acuson S2000	0.894	0.994	418/454	120/122	0.984	298/332	0.898	0.546

Concordance analysis for liver stiffness measurement taken in two different intercostal spaces with the same machine by the same operator and Success Rate of the different SWE elastography techniques, either global or subgrouped according to the presence of severe fibrosis with Fibroscan (>10 kPa) based on the median of all 20 measurements (10 in the reference intercostal space and 10 in the alternative intercostal space). S/A: Successful/Attempted, SR: success rate.

4.3. The role of sampling depth on liver stiffness measurement concordance.

In an attempt to further examine possible causes for the suboptimal concordance among the various machines, our group attempted to account for the fact that the Fibroscan focuses ultrasound waves at 2.5 cm, where it starts measuring, and that, in patients who have a longer skin to capsule distance, the beginning of the measurements may fall in proximity of the liver capsule. Moreover, the majority of elastography methods use a push impulse technique whose signal weakens in depth due to acoustic attenuation, making the measurements theoretically less reliable. Concordance analysis showed good results for a skin to capsule distance ≤20 mm (9 patients) but poor results with a distance ≥20 mm (7 patients) (Table 2). The Fibroscan values of the 2 distance subgroups did not differ (12.0 kPa, range 7.3–34.3 and 15.4 kPa, range 8.3 and 37.4, respectively; $p > 0.05$) nor did the BMI (24.5, range 20.1–27.2 and 26.0, range 23.1–31.8, respectively; $p > 0.05$).

4.4. Technical success rate of US elastography measurement with the various machines

The SR was overtly better for bidimensional SWE than for pSWE and Fibroscan (Table 3). Since the deposition of fibrotic scars makes the liver more heterogeneous, potentially hampering successful sampling of pSWE, the influence of stiffness values was tested. In fact, for almost all pSWE systems, and for Fibroscan as well, the SRs were substantially poorer in livers showing a stiffness ≥10 kPa at Fibroscan. Statistically significant differences, with poorer SR, were evident only for Esaote ($p = 0.028$) and Samsung ($p = 0.002$).

No statistically significant differences emerged for any parameter, excluding patients not yet undergoing antiviral therapy.

5. Discussion

The results of the present study showed that the correlation of different SWE techniques with Fibroscan is variable in terms of both precision and accuracy. Although such findings do not allow making any judgment regarding the quality of each individual technology, they warn against the immediate adoption of the thresholds established for Fibroscan for diagnosing significant fibrosis or cirrhosis [13] for the majority of the other SWE methods currently on the market. Moreover, even a class effect was not demonstrated, as different results were provided by some of the methods categorized as pSWE compared to other pSWE methods. Consequently, dedicated studies for each proprietary technology having histology as reference standard are warranted to make a more specific classification of fibrotic stages based on stiffness thresholds.

In line with our findings, a recent study comparing Philips ElastoPQ technology with Fibroscan having histology as a reference standard, showed an acceptable precision of Philips ElastoPQ ($r = 0.67$), but the thresholds were lower than those of Fibroscan for the same fibrosis stages [15]. In fact, in that case series of 186 patients, only 4 patients showed ElastoPQ values >20 kPa as compared with 15 patients who underwent Fibroscan.

The need for new dedicated studies for each technology, which should theoretically have histology as reference standard, is not a trivial issue as the number of biopsies has substantially decreased in recent years for various reasons, making the investigation of new elastography methods with histology as a reference standard more and more complex or nearly impossible or misleading (as the case series may become composed of atypical clinical cases) [7,16]. This situation poses a challenge to scientific societies to identify possible reliable indicators of the fibrotic stage to be adopted as a reference standard for new studies as an alternative to histology.

It was decided to carry out the present study on real patients and only on patients with significant fibrosis rather than on phantoms or healthy subjects since previous technical studies had demonstrated that the reproducibility of various machines tended to worsen in cases of stiff phantom targets as compared to softer ones [17]. Moreover, previous studies had always found a greater dispersion of data and lower reproducibility in fibrotic rather than in healthy livers [4,18,19]. This suggests that the real challenge is to accurately estimate stiffness in fibrotic patients rather than confirm normality in healthy livers. Hence, our study attempted to target the real clinical challenge. Our study confirmed that the success rate tends to decrease at higher stiffness values, indicating that this is the population to be addressed for technological verification of performance.

Our data showed that a different intercostal space from that adopted for Fibroscan would have been selected in almost half the cases. The implication is that suboptimal comparability between different elastography methods or with biopsy may not only reflect different technological results, but potentially also the choice of different intercostal spaces, a topic which has hardly ever been mentioned in the literature to date.

The current study has limitations. First of all, this study reflects the technological standpoint as of 2016 and, in a rapidly evolving field such as elastography, new advances may improve the results of one or the other technology in terms of comparability with Fibroscan. Technology involves not only specific hardware (machines and transducers), but also the adopted software ver-

sions. This is relevant information with which the user must be acquainted when performing elastography investigations.

The number of patients in our study was quite limited. This was mainly due to the need of obtaining the availability of the largest number of machines at the same time in the same place, requiring a significant number of expert operators and the consent of the patients to be investigated for many hours in one day with several machines. The conflict between the elevated number of patients and the elevated number of machines to be tested appeared to have no practical solution; we favored the highest number of machines (Fibroscan plus 7 machines, where Fibroscan plus 2 machines is the maximum reported so far). In addition, our study was not aimed at establishing any threshold for fibrosis assessment, but simply at verifying the correspondence with the Fibroscan values, thus making a lower number of patients acceptable.

Obtaining a fresh, recent bioptic specimen for histological analysis in a consecutive series of patients would have greatly improved our data but, to date, on practical grounds, this is almost impossible in western countries since only selected patients undergo liver biopsy in the era of Direct acting antiviral agents for HCV and the majority of biopsies are performed on HCV patients today only due to the presence of concurrent possible cofactors. Thus, enrolling only patients with a recent liver biopsy would have prevented any possibility of carrying out such a study with many different machines all with expert operators.

Another potential limitation is the fact that the majority of our patients had already achieved an SVR. Clearly, this condition would have been an insurmountable hurdle if stiffness thresholds had to be defined but, again, for the simple purpose of comparability with Fibroscan, no major biases were expected. In fact, the majority of patients with significant fibrosis and HCV infection have today already been treated and new untreated patients with significant fibrosis are seen only occasionally, precluding the possibility of gathering untreated patients with significant fibrosis on one single day.

In conclusion, our study showed that the new ultrasound elastography machines tend to have moderate to high precision in measuring liver stiffness in HCV-related patients in comparison to Fibroscan which indicates that they have good potential for fibrosis assessment. However, the accuracy was not high enough, at least for the majority of them, to adopt the thresholds already established with Fibroscan for fibrotic stage prediction and each machine must provide its own data; this appears to be a very important piece of information for the hepatology community today when many new elastography machines have just become available.

Conflict of interest

Fabio Piscaglia received Speaker bureau honoraria from Bracco and Research Contract with Esaote. Vito Cantisani received Speaker bureau honoraria from Bracco, Samsung and Toshiba.

Funding

None.

References

[1] Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 2003;29:1705–13.
[2] Bamber J, Cosgrove D, Dietrich CF, et al. EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. Part 1: basic principles and technology. Ultraschall Med 2013;34:169–84.
[3] Friedrich-Rust M, Wunder K, Kriener S, et al. Liver fibrosis in viral hepatitis: noninvasive assessment with Acoustic Radiation Force Impulse imaging versus transient elastography. Radiology 2009;252:595–604.
[4] Bavu E, Gennisson JL, Couade M, et al. Noninvasive in vivo liver fibrosis evaluation using supersonic shear imaging: a clinical study on 113 hepatitis C virus patients. Ultrasound Med Biol 2011;37:1361–73.
[5] Ferraioli G, Tinelli C, Dal Bello B, et al. Accuracy of real-time shear wave elastography for assessing liver fibrosis in chronic hepatitis C: a pilot study. Hepatology 2012;56:2125–33.
[6] Cosgrove D, Piscaglia F, Bamber J, et al. EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. Part 2: clinical applications. Ultraschall Med 2013;34:238–53.
[7] Piscaglia F, Salvatore V, Mulazzani L, et al. Ultrasound shear wave elastography for liver disease. A critical appraisal of the many actors on the stage. Ultraschall Med 2016;37:1–5.
[8] Cassinotto C, Lapuyade B, Mouries A, et al. Non-invasive assessment of liver fibrosis with impulse elastography: comparison of Supersonic Shear Imaging with ARFI and FibroScan(R). J Hepatol 2014;61:550–7.
[9] Gerber L, Kasper D, Fitting D, et al. Assessment of liver fibrosis with 2-D shear wave elastography in comparison to transient elastography and Acoustic Radiation Force Impulse imaging in patients with chronic liver disease. Ultrasound Med Biol 2015;41:2350–9.
[10] Cassinotto C, Boursier J, de Ledinghen V, et al. Liver stiffness in nonalcoholic fatty liver disease: a comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. Hepatology 2016;63:1817–27.
[11] Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology 2003;38:1449–57.
[12] Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology 1996;24:289–93.
[13] EASL-ALEH Clinical Practice Guidelines: non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol 2015;63:237–64.
[14] Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989;45:255–68.
[15] Fraquelli M, Baccarin A, Casazza G, et al. Liver stiffness measurement reliability and main determinants of point shear-wave elastography in patients with chronic liver disease. Aliment Pharmacol Ther 2016;44:356–65.
[16] Sporea I, Popescu A, Gheorghe L, et al. "Quo vadis" liver biopsy? A multi-centre Romanian study regarding the number of liver biopsies performed for chronic viral hepatitis. J Gastrointest Liver Dis 2012;21:326.
[17] Hall TJ, Milkowski A, Garra B, et al. RSNA/QIBA: shear wave speed as a biomarker for liver fibrosis staging. Paper Presented at the IEEE International Ultrasonics Symposium, IUS 2013:397–400.
[18] Sporea I, Bota S, Peck-Radosavljevic M, et al. Acoustic Radiation Force Impulse elastography for fibrosis evaluation in patients with chronic hepatitis C: an international multicenter study. Eur J Radiol 2012;81:4112–8.
[19] Piscaglia F, Salvatore V, Di Donato R, et al. Accuracy of VirtualTouch Acoustic Radiation Force Impulse (ARFI) imaging for the diagnosis of cirrhosis during liver ultrasonography. Ultraschall Med 2011;32:167–75.