



Magnitude dependent discordance in liver stiffness measurements using elastography point quantification with transient elastography as the reference test

Ivica Grgurevic¹ · Nermin Salkic² · Tonci Bozin¹ · Sanda Mustapic¹ · Vladimir Matic¹ · Ivo Dumic-Cule³ · Ida Tjesic Drinkovic¹ · Tomislav Bokun¹

Received: 5 July 2018 / Revised: 25 September 2018 / Accepted: 12 October 2018 / Published online: 28 November 2018

© European Society of Radiology 2018

Abstract

Objectives To investigate diagnostic performance of point shear wave elastography by elastography point quantification (ElastPQ) for non-invasive assessment of liver fibrosis in patients with chronic liver diseases (CLD).

Methods Liver stiffness measurement (LSM) by transient elastography (TE) and ElastPQ was performed in patients with CLD and healthy volunteers. The stage of liver fibrosis was defined by TE which served as the reference. We compared two methods by using correlation, area under the receiver operating characteristics curve (AUC) analysis, Bland and Altman plot and Passing-Bablok regression.

Results A total of 185 subjects (20 healthy volunteers and 165 patients with CLD (128 non-alcoholic fatty liver disease), 83 (44.9%) females, median age 53 years, BMI 27.3 kg/m²) were evaluated. There were 24.3%, 13.5% and 11.4% patients in $\geq F2$, $\geq F3$ and F4 stage, respectively. The best performing cutoff LSM values by ElastPQ were 5.5 kPa for $F \geq 2$ (AUC = 0.96), 8.1 kPa for $F \geq 3$ (AUC = 0.98) and 9.9 kPa for F4 (AUC = 0.98). Mean (SD) difference between TE and ElastPQ measurements was 0.98 (3.27) kPa (95% CI 0.51–1.45, range 4.99–21.60 kPa). Two methods correlated significantly ($r = 0.86$; $p < 0.001$), yet Bland and Altman plot demonstrated difference between measurements, especially with TE values > 10 kPa. Passing and Bablok regression analysis yielded significant constant and proportional difference between ElastPQ and TE.

Conclusion ElastPQ is reliable method for assessment of liver fibrosis but LSM values are not interchangeable with TE, especially above 10 kPa. Diagnostic performance of ElastPQ for sub-classification of patients with compensated advanced chronic liver disease should therefore be furtherly investigated.

Key Points

- *ElastPQ appears to be reliable method for assessment of liver fibrosis, with data presented here mostly applicable to NAFLD.*
- *LSM values produced by TE and ElastPQ are NOT interchangeable—in values < 10 kPa, they are similar, but in values > 10 kPa, they appear to be increasingly and significantly different.*
- *Diagnostic performance of ElastPQ for sub-classification of patients with compensated advanced chronic liver disease should be furtherly investigated.*

Keywords Liver fibrosis · Diagnostic performance · Liver stiffness measurement · Non-alcoholic fatty liver disease

✉ Ivica Grgurevic
ivica.grgurevic@zg.htnet.hr

¹ Department of Gastroenterology, Hepatology and Clinical Nutrition; Department of Medicine, University Hospital Dubrava, University of Zagreb School of Medicine and Faculty of Pharmacy and Biochemistry, Avenija Gojka Suska 6, 10000 Zagreb, Croatia

² Department of Gastroenterology and Hepatology, University Clinical Center Tuzla, Tuzla, Bosnia and Herzegovina

³ Department of Diagnostic and Interventional Radiology, University Hospital Dubrava, University of Zagreb School of Medicine, Zagreb, Croatia

Abbreviations

AUC	Area under the receiver operating characteristics curve
CLD	Chronic liver diseases
ElastPQ	Elastography point quantification
IQR	Interquartile range
LSM	Liver stiffness measurement
NAFLD	Non-alcoholic fatty liver disease
NPV	Negative predictive value
PPV	Positive predictive value
pSWE	Point shear wave elastography
SD	Standard deviation
TE	Transient elastography
US	Ultrasound

Introduction

In recent years, there has been a trend towards the non-invasive evaluation of the stage of liver fibrosis with the goal of providing this important parameter in a less cumbersome and less invasive way as compared to liver biopsy [1, 2]. This paradigm shift is, at least, the partial result of numerous reports dealing with liver biopsy complications, as well as sampling and inter-observer variability in data interpretation [3, 4].

Physical non-invasive tests measure liver stiffness (LS) as a surrogate for liver fibrosis. Ultrasound (US) methods that use elastography for liver stiffness measurement (LSM) include transient elastography (TE), point shear wave elastography (pSWE), two-dimensional shear wave elastography (2D-SWE) and MR elastography [5–7]. Due to the accumulated scientific evidence from the numerous studies performed so far, TE (developed by Echosens) is considered to be the non-invasive standard for the measurement of LS [2]. The calculation of the elastic shear wave propagation speed through different tissues is the principle behind TE [5, 8, 9]. However, the inability to acquire valid measurements in the setting of ascites and also, to a certain extent, in obesity, is one of several drawbacks [10, 11].

Point SWE utilises uses acoustic radiation force (ARFI) to generate ultrasonic pressure waves that are transmitted through the body/liver where a portion of its energy is used to induce shear waves that travel in a perpendicular direction to the plane of the excitation impulse and the shear wave velocity is subsequently measured by pulsed-Doppler using the same probe.

As a recently introduced representative of pSWE, elastography point quantification (ElastPQ[®]), developed by Philips Healthcare, has not been fully evaluated in clinical studies [12–14]. The main advantage of pSWE is its integration in the US machine and its simultaneous operation with B-mode scanning. This enables the selection of the region of

interest (ROI) and also its operation, despite the presence of ascites and obesity. Parameters, such as the cutoff values for different fibrosis stages, the predictive value and the influence of the quality criteria, have undergone considerable investigation in recent years [14–18]. The aim of this study was to investigate the diagnostic performance of ElastPQ for the non-invasive assessment of LS in patients with chronic liver disease (CLD) and healthy patients, using TE as the reference test.

Patients and methods

This was a single-centre, cross-sectional study with prospectively enrolled patients in a tertiary-care hospital setting. During a 4-month period (February–June 2017), outpatients with CLD referred for liver ultrasound examination in the ultrasound unit of the Department of Gastroenterology were considered as candidates for this study. On each day of the week, the first two patients with a referring diagnosis of CLD were included in the study, provided that a successful LSM had been performed using TE, which served as a reference method for staging liver fibrosis. In cases where the LSM had failed, the next patient with CLD was analysed and so on, until successful LSMs had been accomplished for two patients each day.

Patients had to be over 18 years of age with previously diagnosed CLD and with available laboratory results performed within a 3-month period. The diagnosis of non-alcoholic fatty liver disease (NAFLD) relied on the ultrasonographically confirmed presence of fatty liver in at least two US examinations 6 months apart, with/without elevated liver function tests if excessive alcohol consumption and the use of drugs with known steatogenic potential had been excluded. The following biochemical parameters were documented: bilirubin, aspartate transaminase (AST), alanine transaminase (ALT), gamma-glutamyltransferase (GGT), alkaline phosphatase (AP) and platelet count (Plt). From these data, the FIB4 score was calculated for each patient according to the previously published formula [19]. The exclusion criteria were overt cholestasis with dilatation of the intrahepatic bile ducts, congestive liver failure, liver transaminases greater than $5 \times$ the upper limit of the normal (ULN) value, the presence of ascites and the failure to perform a reliable LSM using TE [2]. The patients that were eligible according to the inclusion/exclusion criteria underwent LSM using ElastPQ during the same visit. We evaluated the diagnostic accuracy of the LSM obtained by ElastPQ against the LSM obtained using TE as the reference standard. As a reference, we used three points of clinical interest, according to established TE cutoff points for liver fibrosis stage: ≥ 7 kPa for $F \geq 2$; ≥ 9.5 kPa for $F \geq 3$ and ≥ 12 kPa for $F = 4$ [7].

We also included a group of healthy volunteers in order to check the LSM values obtained using ElastPQ in healthy livers. This group of participants was recruited from the subjects who came to the unit for an annual preventive check-up. These participants were considered eligible if they had no history of liver disease and their liver function tests and liver US were normal. All of the participants signed an informed consent and the local ethics committee approved the protocol of the study.

Elastography point quantification

Elastography point quantification was performed by three experienced physicians (each had performed at least 100 LSMs using ElastPQ prior to the start of this study) using the Epiq7 ultrasound system (Philips Healthcare,) with a convex transducer C5-1 (1–5 MHz). The subjects fasted for at least 3 h prior to sessions. The right liver lobe was targeted through the intercostal space with the subjects lying in a dorsal decubitus position with the right arm in maximal abduction. The US probe was lubricated with gel to improve ultrasonic wave transmission into the liver. The skin to liver capsule distance (SCD) was measured for each participant. With the help of the real-time B-mode image, a vessel-free area of at least 1.5 cm below Glisson's capsule, was selected. During measurements, patients were instructed to hold their breath in a neutral position while the operator pressed a button that launched the measurement acquisition. At least 10 valid measurements, expressed in kPa, were repeated for each patient. The median value was considered reliable only if the interquartile range/median (IQR/M) was < 30%.

Transient elastography

Three independent operators with at least 2 years of experience performed the TE (each performed > 200 LSM using TE). We used the FibroScan Touch 502 device (EchoSens). Since both the TE and the ElastPQ were performed in the same session, the conditions were the same. All sessions were performed in the same examination room, the subjects were fasting and the right liver lobe was targeted through the intercostal space with the subjects lying in a dorsal decubitus position with the right arm in maximal abduction. Only examinations with 10 valid measurements (success rate $\geq 60\%$) and an interquartile range/median (IQR/M) of < 30% for values greater than 7.1 kPa were considered reliable. The IQR < 30% criterion was not mandatory for patients with a median LSM < 7.1 kPa [20]. The LSM was considered failed when no numerical value could be obtained. We intended to use an M probe in patients with $SCD \leq 25$ mm or $BMI \leq 30$ kg/m². However, if the Fibroscan device signalled that the XL probe should be used instead, or in

cases of a high number of failed LSMs using the M probe (> 40%), we switched to the XL probe.

Statistical analysis

Statistical analysis was performed using SPSS software, version 24.0 (SPSS Inc.), and MedCalc for Windows, version 12.0 (MedCalc Software). The patient characteristics are given as the mean \pm SD, as appropriate. Student's *t* test for independent measurements was used for comparison of the means. The diagnostic accuracy of each non-invasive model was evaluated by calculating the areas under the receiver operating characteristics (ROC) curve (AUC).

In order to assess the diagnostic reliability and interchangeability of ElastPQ against TE, which served as the reference standard, we used a set of statistical tests for the purpose of method comparison. We graphically inspected the two methods using a Bland-Altman plot and then proceeded with Passing-Bablok regression analysis in order to quantify the existence of differences between the measurements produced by the two methods and to assess for the presence of the constant and proportional difference between them.

A two-sided *p* value of < 0.05 was considered to be significant for all statistical tests.

Results

A successful TE LSM, used as the reference standard, was obtained in 201 subjects (181 patients who were eligible on the basis of the inclusion/exclusion criteria and 20 healthy volunteers). The majority of the TE LSM measurements were performed using an XL probe (126; 68.1%). These 201 subjects underwent LSM using ElastPQ during the same session, which produced reliable results in 185/201 (92%) and an unreliable LSM in 16/201 (7.9%), whereas no failure of LSM was observed in this cohort. Subjects with successful measurements had, on average, lower mean values ($p < 0.01$) of BMI and skin to capsule distances (SCD) in comparison with those with unreliable measurements (BMI 27.25 vs. 30.68 kg/m² and SCD 1.81 vs. 2.06 cm, respectively).

The final analysis included a total of 185 patients with a successful LSM using both methods, with 102 (55.1%) males and 83 (44.9%) females. The mean age (SD) of the patients was 53 (14) years, ranging from 18 to 82 years. The baseline characteristics of the patients are presented in Table 1.

The liver stiffness measurement values obtained using ElastPQ had AUC = 0.955 (95% CI = 0.914–0.980; $p < 0.001$) for diagnosing fibrosis stage $F \geq 2$, with the best performing threshold point of 5.5 kPa. The sensitivity to this cutoff was 97.8% (95% CI = 88.2–99.6) and the specificity was 84.3% (95% CI = 77.2–89.9) with a positive predictive

Table 1 Baseline characteristics of the studied cohort of patients

Parameter	Value (number of patients 185)
Age, years (median; interquartile range)	53 (43–65)
Males (<i>N</i> ; %)	102; 55.1%
Body mass index, kg/m ² (mean, SD)	27.25 (5.12)
TBIL, μmol/L (mean, SD)	17.16 (19.16)
AST, U/L (mean, SD)	37.5 (27.8)
ALT, U/L (mean, SD)	43.8 (36.9)
ALP, U/L (mean, SD)	108.4 (131.2)
GGT, U/L (mean, SD)	112.4 (216.4)
PLT × 10 ⁹ /L (mean, SD)	245.9 (68.2)
FIB-4 (mean, SD)	1.54 (1.62)
Liver disease aetiology (<i>N</i> , %)	
NAFLD	128 (69.2)
Alcohol	12 (6.5)
HBV/HCV	12 (6.5)
Autoimmune	9 (4.9)
Haemochromathosis	4 (2.2)
Healthy	20 (10.8)
Skin to liver capsule distance, cm (mean, SD)	1.81 (0.40)
Controlled attenuation parameter, dB/m (CAP; mean, SD)	273 (69)
Median LS value measured by TE, kPa (mean, SD; median, IQ range)	7.04 (5.79); 5.2 (4.0–6.8)
Median LS value measured by ElastPQ, kPa (mean, SD; median, IQ range)	6.06 (3.60); 4.9 (4.1–6.4)
Stage of liver fibrosis according to TE LSM values (<i>N</i> , %)	
< 7 kPa	140 (75.7)
≥ 7 kPa for F ≥ 2	45 (24.3); 20 (10.8) in F2 stage
≥ 9.5 kPa for F ≥ 3	25 (13.5); 4 (2.2) in F3 stage
≥ 12 kPa for F = 4	21 (11.4)

value (PPV) of 66.7% and a negative predictive value (NPV) of 99.2%.

To diagnose the F ≥ 3 fibrosis stage, the LSM values obtained using ElastPQ had AUC = 0.983 (95% CI = 0.952–0.996; $p < 0.001$), with the best performing threshold point of 8.1 kPa. This cutoff had a 92% sensitivity (95% CI = 73.9–98.8) and the specificity was 96.25% (95% CI = 92.0–98.6) with a PPV of 79.3% and an NPV of 98.7%.

To diagnose cirrhosis (F = 4), the LSM values obtained using ElastPQ had AUC = 0.982 (95% CI = 0.921–1.000; $p < 0.001$), with the best performing threshold point of 9.88 kPa. The sensitivity to this cutoff was 90.5% (95% CI = 69.6–98.5) and the specificity was 98.2% (95% CI = 94.7–99.6) with a PPV of 86.4% and an NPV of 98.8%.

A graphic representation of all three AUCs is depicted in Fig. 1.

We performed a subgroup analysis in patients with NAFLD ($N = 128$; 69.2%). The AUC for F ≥ 2 ($N = 30$) was 0.961 (95% CI = 0.932–0.999; $p < 0.0001$), with the best performing cutoff value 5.5 kPa (sensitivity 93.9%; specificity 83.2%). Due to spectrum bias (F0–1 was present in 98, F2 in

15, F3 in four, and F4 in 11 patients, respectively), it was not possible to calculate reliable LSM cutoff values for F ≥ 3 and F4 in the NAFLD subgroup of patients. This warrants further study.

We analysed the correlation between the LSM values obtained using TE and those obtained with ElastPQ. There was a medium to strong significant correlation between the methods, with $\rho = 0.72$ ($p < 0.0001$; 95% CI 0.64–0.78). However, a pronounced dissipation of the LSM values was observed above the TE LSM threshold of around 10 kPa, as depicted in the scatter diagrams in Fig. 2.

The mean (SD) difference between the TE and the ElastPQ values was 0.98 (3.27) kPa, with 95% CI 0.51–1.45 and ranging from -4.99 to 21.60 kPa.

We also calculated the relative difference (in %) in LSM values between the methods ((TE/ElastPQ-1) × 100%). Although the overall mean relative difference between the TE and the ElastPQ values in the whole cohort was 3.7% (SD 26.9), there was a significant difference between the values in the TE LSM subgroups (see Table 2). In the subgroup of patients with TE LSM values ≤ 5 kPa, on average,

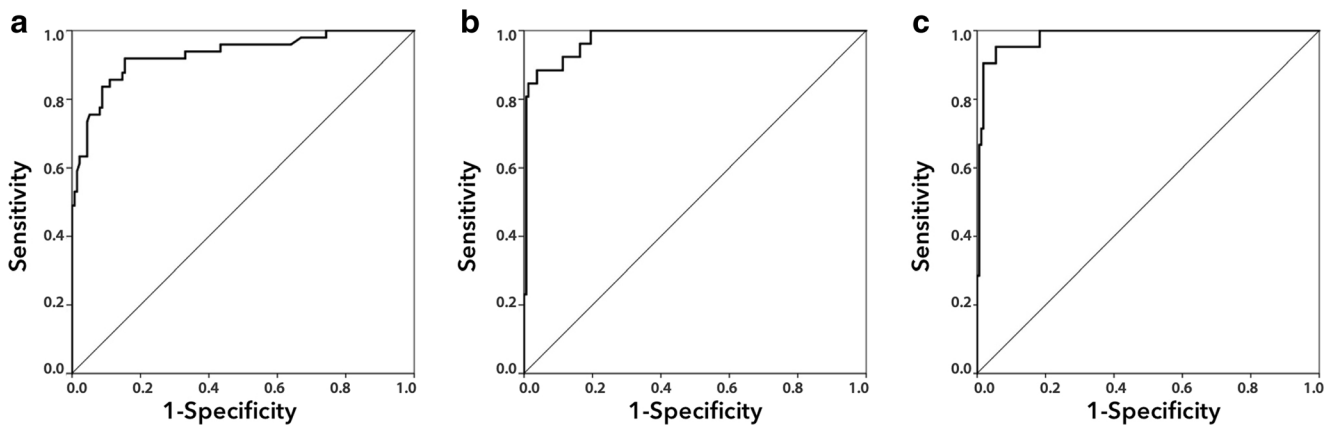


Fig. 1 Areas under the receiver operating curves for diagnostic accuracy of liver stiffness measurement (LSM) obtained by ElastPQ with transient elastography (TE) serving as the reference method. (a) $F \geq 2$ (LSM by $TE \geq 7$ kPa). (b) $F \geq 3$ (LSM by $TE \geq 9.5$ kPa). (c) $F = 4$ (LSM by $TE \geq 12$ kPa)

the measurements of liver stiffness using the ElastPQ revealed 8.4% higher values in favour of ElastPQ. However, in the other subgroups of patients, the magnitude of the difference increased progressively, with the average highest difference of 37.5% in favour of TE, as observed in the subgroup of patients with TE values > 15 kPa.

In order to evaluate the diagnostic reliability and interchangeability of the LSM obtained using ElastPQ, we compared this against the established reference standard (TE) using statistical methods for the comparison of the diagnostic methods.

We initially used a Bland and Altman plot to visually compare the two methods (see Fig. 3). The dotted lines represent the limits of agreement between ElastPQ and TE, which are defined as the mean difference ± 1.96 SD of the differences Δ .

If these limits do not exceed the maximum allowed difference between methods Δ (the differences within the mean ± 1.96 SD are not *clinically* important), the two methods are considered to be in agreement and may be used interchangeably. In our case, the two methods are mostly in agreement with values of less than 10 kPa. Yet, with values over 10 kPa, there is a clear dispersion of the measurements, with a tendency of ElastPQ to produce lower values than TE, although, clearly, this effect is dependent on the magnitude of the measurements.

The Passing-Bablok regression analysis yielded a regression formula of $y = 1.06 + 0.75x$ (Fig. 4) [21]. The intercept of 1.06 had a 95% CI of 0.55–1.46, which indicated a small but statistically significant constant difference (CI does not contain the value of 0). The slope of 0.75 had a 95% CI of 0.66–

Fig. 2 Scatter diagram of LSM values obtained by TE and ElastPQ. Diagram depicts dissipation of values, especially with $TE > 10$ kPa

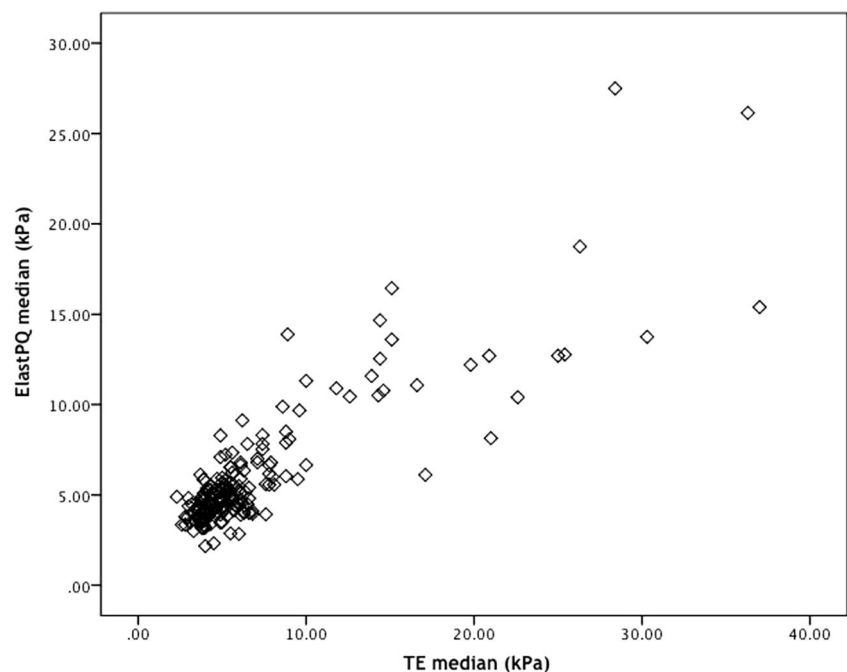


Table 2 Relative difference between TE and ElastPQ liver stiffness values. Relative difference was calculated as $(TE / ElastPQ - 1 \times 100\%)$ for each patient

TE LSM value	≤ 5 kPa	5–10 kPa	10–15 kPa	> 15 kPa
Relative difference between TE and ElastPQ liver stiffness values; mean (95% CI)	-8.4% (-13.6 to 3.1)	10.5% (5.3 to 15.6)	15.0% (5.7 to 24.4)	37.5% (25.3 to 49.8)

0.85 (CI does not contain the value of 1), which corresponded with the significant proportional difference between ElastPQ and TE. According to the cusum test ($p > 0.10$), there was no significant deviation from linearity.

We also created a residuals diagram (Fig. 5) based on the Passing-Bablok regression equation. The TE LSM values (used as a reference point) are on the *x*-axis, while the residuals of difference between the ElastPQ values and the values calculated from the regression equation (under perfect conditions, each difference would be zero, in other words, these methods would produce identical results) are shown on the *y*-axis. As observed, although there was a fair grouping of values with measurements of < 10 kPa, there was an obvious disagreement in cases with measurements of > 10 kPa.

Discussion

The results of this study demonstrate the excellent diagnostic performance of ElastPQ as the representative of the pSWE methods for the non-invasive staging of liver fibrosis in patients with CLD when TE was used as a reference method.

The areas under the receiver operating characteristics curve for three clinically relevant points of interest (significant fibrosis, advanced fibrosis and cirrhosis) were in the range of 0.95 to 0.99.

ElastPQ is a relatively new US technique that is based on the elastography method for the quantitative assessment of liver fibrosis via measurements of liver stiffness. The studies published so far have revealed inconsistent results both in terms of the diagnostic performance of ElastPQ and the calculated cutoff values for differentiating between the stages of liver fibrosis (Table 3). The use of different reference methods for liver fibrosis assessment (i.e., TE or liver biopsy) and the mixture of the aetiologies of the CLD studied, added to the heterogeneity of the results. Although our results were obtained from a cohort of mixed aetiology CLD, they mostly reflect the performance of ElastPQ in NAFLD since the majority of the patients (almost 70%) had NAFLD. This is a finding that has not previously been reported.

In the studies published to date, there has been a clear trend pointing to the high reliability of the diagnostic performance of ElastPQ, although there remain some unresolved issues that need to be analysed in greater detail. First, from the results

Fig. 3 Bland and Altman plot comparing transient elastography (TE) and elastography point quantification (ElastPQ) measurements of liver stiffness. Most of the values < 10 kPa remain within ± 1.96 SD boundaries; values > 10 kPa show prominently lower values of ElastPQ when compared with TE

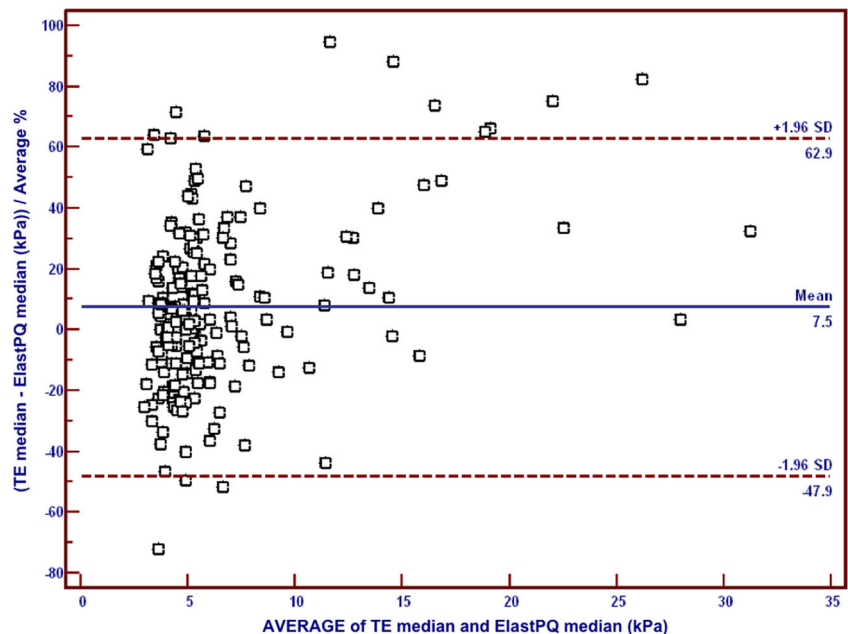
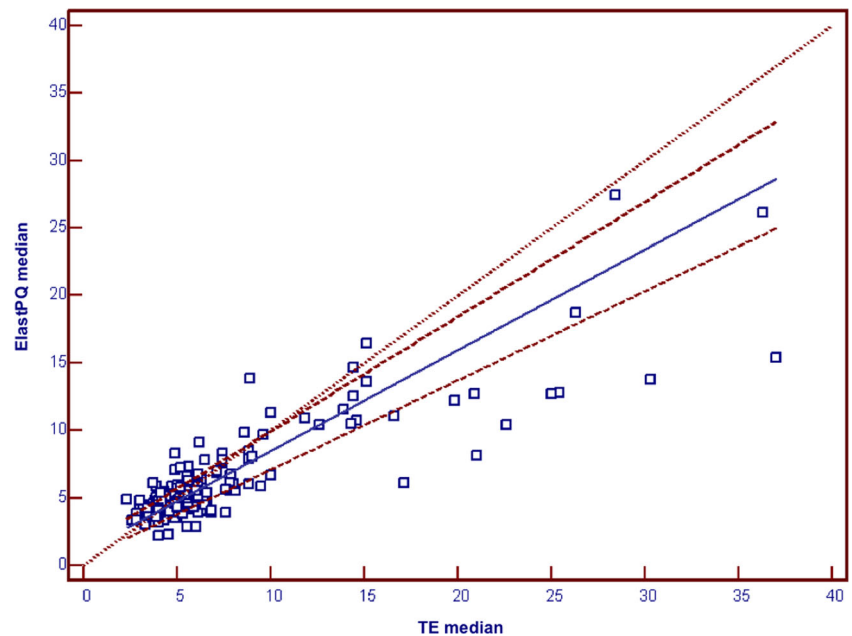


Fig. 4 Scatter diagram of transient elastography (TE) and elastography point quantification (ElastPQ) values. Passing and Bablok regression line (blue), red dashed line represents confidence interval for regression line. Red dotted line represents reference line where both methods would be in perfect correlation



published by Ferraioli et al (and confirmed in the papers that followed), it became evident that the LSM values measured using ElastPQ were lower compared to the TE measure when undertaken in typical clinical and scientific scenarios that recruited patients with compensated CLD without overt portal hypertension or liver decompensation (this is because this subset of patients were candidates for liver biopsy) [13]. These patients usually have an LSM in the range of 5–15 kPa, as measured using TE. Within this range, the LSM measured using ElastPQ is, on average, 1 kPa lower, as also demonstrated in our study. However, in patients with more advanced liver disease and a stiffer liver, the difference

between TE and ElastPQ becomes progressively divergent (Fig. 2). Although this 1 kPa approximation might be considered acceptable for LSM values within the range of 5–10 kPa, the lack of linearity and the increasing dissipation above this threshold do not allow for any meaningful correction to make the LSM values interchangeable with TE.

This observation calls for further investigation of ElastPQ, specifically in patients with compensated advanced chronic liver disease (defined by a TE LSM ≥ 10 kPa), in order to test its ability for risk stratification in this group of patients (presence of clinically significant portal hypertension, large oesophageal varices and prognostication).

Fig. 5 Residuals diagram of transient elastography (TE) and elastography point quantification (ElastPQ) comparison, based on Passing and Bablok regression equation ($F(x)$). TE values are entered on x -axis, while on y -axis are residuals of difference between ElastPQ values and values calculated from regression equation (in perfect condition, each difference would be zero, in other words, both methods produce identical results)

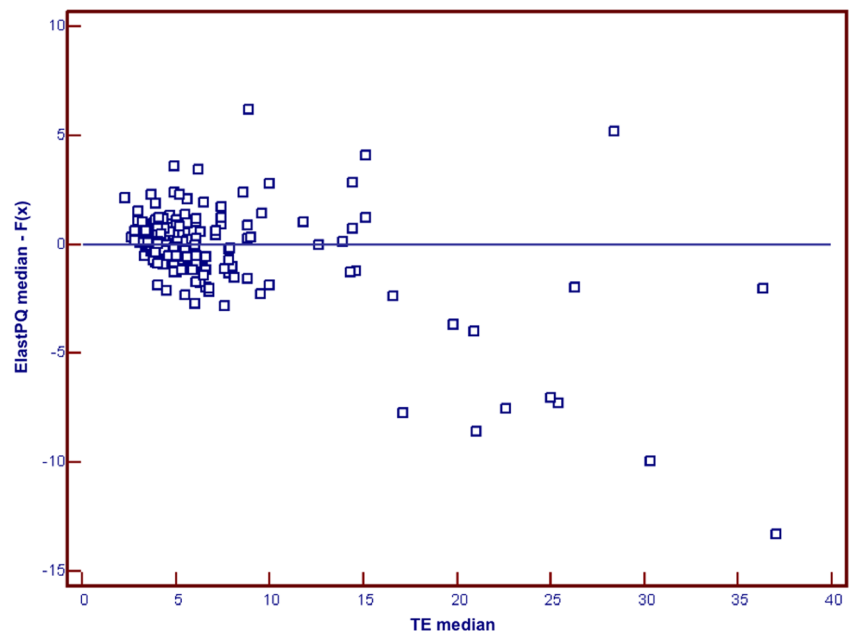


Table 3 Published cutoff values and the diagnostic performance of elastography point quantification for non-invasive staging of liver fibrosis

Author, year	Number of patients	Aetiology	Reference method	Fibrosis stage $F \geq 2$		Fibrosis stage $F = 4$	
				Cutoff (kPa)	AUC (95% CI)	Cutoff (kPa)	AUC (95% CI)
Ferraioli et al 2014 [13]	134	HCV (76.1%), HBV (20.9%)	Histology	5.7	0.80 (0.71–0.87)	7.2	0.95 (0.89–0.99)
Ferraioli et al 2016 [15]	278	HCV (100%)	TE	6.43	Not reported	11.34	Not reported
Ferraioli et al 2016 [16]	186	HCV (56%), mixed (44%)	TE	7.6	0.77 (0.70–0.83)	10.4	0.88 (0.78–0.96)
Mare et al 2017 [17]	228	HCV (74%), HBV (26%)	TE	7.2	0.94 (0.92–0.97)	8.9	0.97 (0.95–0.99)
Lee et al 2017 [18]	106	HBV (78%), HCV (14.6%)*	Histology	4.26	0.86 (0.8–0.93)	4.86	0.85 (0.77–0.94)
Grgurevic, this study	185	NAFLD (69%), mixed (20%), healthy (11%)	TE	5.5	0.95 (0.91–0.98)	9.8	0.98 (0.92–1)

HCV hepatitis C virus, HBV hepatitis B virus, NAFLD non-alcoholic fatty liver disease, TE transient elastography

*Aetiology reported in 41/101 patients

The differences in the LSM values between the elastographic devices may be related to the specific technologies used to generate and track the shear waves. This has not only been observed between TE and ElastPQ but also between other elastography methods, as demonstrated by Piscaglia et al [22]. These authors also reported that a different intercostal space for the SWE from the one adopted for the Fibroscan was selected in almost half of the cases, which might have additionally influenced the final LSM result. Further, we must note the fact that TE measures a larger volume of liver tissue than pSWE; hence, it is probable that pSWE measures fibrosis in a local area of liver tissue that may have a different stiffness to that of the surrounding area. When using TE, such local differences are muted by averaging the stiffness across a larger sampling volume.

We need to note the limitations of our study. We used TE as a reference standard. Although liver biopsy has numerous shortcomings, it is still the de facto considered as the “gold standard”. Yet, TE has been established as a reliable surrogate and, in terms of comparing the interchangeability of the methods, a lack of the “gold standard” liver biopsy is not methodologically relevant. The studied cohort comprised different aetiologies of CLD but the majority were NAFLD patients, thus, the results mostly reflect this specific aetiology. However, there was a significant spectrum bias in terms of the fibrosis stage with only one quarter of patients presenting with significant fibrosis. Nevertheless, this reflects the prevalence of significant fibrosis in the real world outpatient population and, therefore, this enhances the clinical usability of our results. The number of analysed patients was moderate and further analyses with higher numbers of patients with a single aetiology CLD are still needed. The majority of the LSMs in this study were obtained using the XL probe, even in the patients with $SCD < 25$ mm, and it may be argued that the cutoff values are not the same for the M and the XL probe. However, in our experience, 25 mm is probably too high a

SCD to use an M probe because even an 18–20 mm SCD results in a high number of failed or unreliable LSMs. In these cases, the XL probe almost always produces a reliable LSM. This observation is supported by some other authors who have reported the better accuracy of the XL as compared to the M probe in patients with $SCD \geq 17.5$ mm [23]. Although the LSMs obtained using the XL probe tend to be 1–2 kPa lower compared to the M probe, it has recently been reported that when used in appropriate patients the LSM result can be interpreted using the same diagnostic cutoffs for both probes [24, 25].

Conclusions

ElastPQ appears to be a reliable method for the assessment of liver fibrosis, with the data presented here as being mostly applicable to NAFLD, which has not previously been evaluated. Yet, there is a need for more studies with larger samples of patients with a single aetiology of liver disease in order to establish valid thresholds for the fibrosis stages in a specific aetiology. The LSM values produced by TE and ElastPQ are NOT interchangeable—in values < 10 kPa, they are similar, but in values > 10 kPa, they appear to be increasingly and significantly different. The diagnostic performance of ElastPQ for a sub-classification of patients with compensated advanced chronic liver disease should, therefore, be further investigated.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Ivica Grgurevic, Department of Gastroenterology, Hepatology and Clinical

Nutrition; Department of Medicine, University Hospital Dubrava, University of Zagreb School of Medicine, Zagreb, Croatia.

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

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Methodology

- Prospective
- Cross sectional study
- Performed at one institution

References

1. Ferraioli G, Filice C, Castera L et al (2015) WFUMB guidelines and recommendations for clinical use of ultrasound elastography: part 3: liver. *Ultrasound Med Biol* 41:1161–1179
2. European Association for Study of Liver, Asociacion Latinoamericana para el Estudio del Hígado (2015) EASL-ALEH clinical practice guidelines: non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol* 63:237–264
3. Bedossa P, Dargère D, Paradis V (2003) Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 38:1449–1457
4. Piccinino F, Sagnelli E, Pasquale G, Giusti G (1986) Complications following percutaneous liver biopsy: a multicentre retrospective study on 68 276 biopsies. *J Hepatol* 2:165–173
5. Sandrin L, Fourquet B, Hasquenoph JM et al (2003) Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Biol* 29:1705–1713
6. Bamber J, Cosgrove D, Dietrich C et al (2013) EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. Part 1: basic principles and technology. *Ultraschall Med* 34:169–184
7. Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK (2011) Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. *J Hepatol* 54:650–659
8. Castera L (2015) Noninvasive assessment of liver fibrosis. *Dig Dis* 33:498–503
9. de Lédinghen V, Vergniol J, Barthe C et al (2013) Non-invasive tests for fibrosis and liver stiffness predict 5-year survival of patients chronically infected with hepatitis B virus. *Aliment Pharmacol Ther* 37:979–988
10. Cosgrove D, Piscaglia F, Bamber J et al (2013) EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. Part 2: clinical applications. *Ultraschall Med* 34: 238–253
11. Şirli R, Sporea I, Deleanu A et al (2014) Comparison between the M and XL probes for liver fibrosis assessment by transient elastography. *Med Ultrason* 16:119–122
12. Yoo H, Lee JM, Yoon JH, Lee DH, Chang W, Han JK (2016) Prospective comparison of liver stiffness measurements between two point shear wave elastography methods: virtual touch quantification and elastography point quantification. *Korean J Radiol* 17: 750–757
13. Ferraioli G, Tinelli C, Lissandrin R et al (2014) Ultrasound point shear wave elastography assessment of liver and spleen stiffness: effect of training on repeatability of measurements. *Eur Radiol* 24: 1283–1289
14. Sporea I, Bota S, Grădinaru-Taşcău O, Şirli R, Popescu A (2014) Comparative study between two point shear wave elastographic techniques: acoustic radiation force impulse (ARFI) elastography and ElastPQ. *Med Ultrason* 16:309–314
15. Ferraioli G, Maiocchi L, Lissandrin R, Tinelli C, De Silvestri A, Filice C (2016) Accuracy of the ElastPQ® technique for the assessment of liver fibrosis in patients with chronic hepatitis C: a “real life” single center study. *J Gastrointest Liver Dis* 25:331–335
16. Fraquelli M, Baccarin A, Casazza G et al (2016) Liver stiffness measurement reliability and main determinants of point shear-wave elastography in patients with chronic liver disease. *Aliment Pharmacol Ther* 44:356–365
17. Mare R, Sporea I, Lupuşoru R et al (2017) The value of ElastPQ for the evaluation of liver stiffness in patients with B and C chronic hepatopathies. *Ultrasonics* 77:144–151
18. Lee JE, Shin KS, Cho JS et al (2017) Non-invasive assessment of liver fibrosis with ElastPQ: comparison with transient elastography and serologic fibrosis marker tests, and correlation with liver pathology results. *Ultrasound Med Biol* 43:2515–2521
19. Sterling RK, Lissen E, Clumeck N et al (2006) Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43:1317–1325
20. Dietrich CF, Bamber J, Berzigotti A et al (2017) EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version). *Ultraschall Med* 38:e16–e47
21. Passing H, Bablok W (1983) A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, part I. *J Clin Chem Clin Biochem* 21:709–720
22. Piscaglia F, Salvatore V, Mulazzani L et al (2017) Differences in liver stiffness values obtained with new ultrasound elastography machines and Fibroscan: a comparative study. *Dig Liver Dis* 49: 802–808
23. Kumagai E, Korenaga K, Korenaga M et al (2016) Appropriate use of virtual touch quantification and FibroScan M and XL probes according to the skin capsular distance. *J Gastroenterol* 51:496–505
24. Myers RP, Pomier-Layrargues G, Kirsch R et al (2012) Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. *Hepatology* 55:199–208
25. Boursier J, Hiriart JB, Lannes A et al (2017) Liver stiffness measurement with Fibroscan: Use the right probe in the right condition. *Hepatology* 66(S1):346A