Thieme

Understanding confounding factors allows for accurate interpretation of liver stiffness measurements by ElastQ, a novel 2 D shear wave elastography technique

Das Verständnis von verzerrenden Faktoren ermöglicht die akkurate Interpretation von Lebersteifigkeitsmessungen mittels ElastQ, einer neuen 2-D-Scherwellen-Elastographie-Technik

Authors
David JM Bauer[1], Annalisa De Silvestri[2], Laura Maiocchi[3], Ambra Raimondi[4], Ruxandra Mare[5], Mattias Mandorfer[1], Ioan Sporea[5], Theresa Müllner-Bucsics[1], Giovanna Ferraioli[6], Thomas Reiberger[7]

Affiliations
1 Division of Gastroenterology and Hepatology, Medical University of Vienna, Vienna, Austria
2 Clinical Epidemiology and Biometric Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
3 Clinical Sciences and Infectious Diseases Department, Fondazione IRCCS Policlinico San Matteo, Medical School University of Pavia, Pavia, Italy
4 Clinical Sciences and Infectious Diseases Department, Fondazione I.R.C.C.S. Policlinico San Matteo, Pavia, Italy, Pavia, Italy
5 Department of Internal Medicine II, Division of Gastroenterology & Hepatology, Center for Advanced Research in Gastroenterology and Hepatology, "Victor Babes" University of Medicine and Pharmacy Timisoara, Timisoara, Romania
6 Infectious Diseases, Fondazione IRCCS Policlinico S. Matteo, Medical School, University of Pavia, Pavia, Italy
7 Internal Medicine III, Div. of Gastroenterology & Hepatology, Medical University of Vienna, Vienna, Austria

Keywords
2 D-SWE, ElastQ, liver stiffness measurement, liver elastography, two-dimensional shear wave elastography

received 18.7.2023
accepted after revision 14.5.2024
published online 2024

Bibliography
Ultraschall in Med
DOI 10.1055/a-2329-2801
ISSN 0172-4614

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence
Prof. Thomas Reiberger
Internal Medicine III, Div. of Gastroenterology & Hepatology,
Medical University of Vienna,
Waehringer Guertel 18–20, A-1090 Vienna, Austria
thomas.reiberger@meduniwien.ac.at

Supplementary Material is available at https://doi.org/10.1055/a-2329-2801.

ABSTRACT

Purpose Liver stiffness measurement (LSM) using vibration-controlled transient elastography (VCTE) or two-dimensional shear wave elastography (2 D-SWE) is recommended to assess the risk of liver fibrosis and advanced chronic liver disease. Even though both techniques measure liver stiffness, their numerical results often diverge. Confounders and reliability criteria for 2 D-SWE have not been systematically investigated.

Materials and Methods We prospectively recruited participants with paired LSM by VCTE and the novel 2 D-SWE technique ElastQ (Philips) in three European tertiary centers. The following parameters were recorded: sex, age, body mass index (BMI), etiology, laboratory markers of liver damage and function, as well as cholestasis, LSM by VCTE and controlled

attenuation parameter (CAP), interquartile range (IQR)/median for VCTE-LSM and ElastQ-LSM, and the skin-to-liver capsule distance.

Results We included 840 participants: 447 (53.2 %) males; median age 57.0 [IQR:19.0] years; median BMI 25.4 [6.0] kg/m^2; median VCTE-LSM 7.25 [9.2] kPa; median ElastQ-LSM 6.7 [5.4] kPa. On uni- and multivariable modeling (adjusted for LSM), we found that the discrepancy increased with liver stiffness and markers of disease severity. Skin-to-liver capsule distance and BMI affected VCTE-LSM more compared to ElastQ-LSM and significantly increased the discordance between the two measurements.

Conclusion The discrepancy of ElastQ-LSM to VCTE-LSM increases with liver stiffness and disease severity. BMI and skin-to-liver capsule distance increase the discrepancy between VCTE- and ElastQ-LSM but affect ElastQ-LSM less. The quality criterion IQR/median ≤ 30 % indicates reliable ElastQ-LSM.

ZUSAMMENFASSUNG

Hintergrund Trotz gleicher Messgröße (Lebersteifigkeit) zeigen VCTE und die neue 2 D-SWE-Technik ElastQ oft divergente Ergebnisse. Die verzerrenden Faktoren und Zuverlässigkeitskriterien für 2 D-SWE wurden bislang nicht systematisch untersucht.

Materialen und Methoden Wir rekrutierten prospektiv Teilnehmer mit gepaarten LSM-Messungen mittels VCTE und der neuen 2 D-SWE-Technik ElastQ (Philips) in 3 europäischen Tertiärzentren. Die folgenden Parameter wurden erfasst: Geschlecht, Alter, Body-Mass-Index (BMI), Ätiologie, Labormarker für Leberschäden und -funktion sowie Cholestase, LSM durch VCTE und der kontrollierte Dämpfungsparameter (CAP), Interquartilsabstand (IQR)/Median für VCTE-LSM und ElastQ-LSM, sowie der Abstand von der Haut zur Leberkapsel.

Ergebnisse Wir schlossen 840 Teilnehmer ein: 447 (53,2 %) Männer; mittleres Alter 57,0 [IQR: 19,0] Jahre; mittlerer BMI 25,4 [1] 99] kg/m^2; mittlere VCTE-LSM 7,25 [9],[2] kPa; mittlere ElastQ-LSM 6,7 [5],[4] kPa. In univariater und multivariater Modellierung (angepasst an LSM) stellten wir fest, dass die Diskrepanz mit der Lebersteifigkeit und den Markern für die Schwere der Erkrankung zunimmt. Der Abstand von der Haut zur Leberkapsel und der BMI beeinflussten VCTE-LSM stärker als ElastQ-LSM und erhöhten die Diskrepanz zwischen den beiden Messungen signifikant.

Schlussfolgerung Die Diskrepanz von ElastQ-LSM zu VCTE-LSM nimmt mit der Lebersteifigkeit und der Schwere der Erkrankung zu. BMI und der Abstand von der Haut zur Leberkapsel erhöhen die Diskrepanz zwischen VCTE- und ElastQ-LSM, beeinflussen jedoch ElastQ-LSM weniger. Das Qualitätskriterium IQR/Median ≤ 30 % zeigt eine zuverlässige ElastQ-LSM an.

Introduction

Ultrasound-based shear wave elastography (SWE) methods such as vibration-controlled transient elastography (VCTE) [2], point shear wave elastography (pSWE) [3], and two-dimensional shear wave elastography (2 D-SWE) [4] are used to assess fibrosis risk in people with suspected or diagnosed chronic liver disease. Elastography techniques assess the same biomechanical property of the liver tissue, namely its stiffness. However, various methods and devices may yield different numerical values for liver stiffness measurements (LSMs) [1, 3, 5, 6]. Therefore, when assessing fibrosis risk in patients, different technologies require specific cut-off values. However, these differences tend to be relatively low in the range of liver stiffness used for the assessment of the risk of liver fibrosis or compensated advanced chronic liver disease (cACLD). This minimal variation allows for the use of general rules, such as the "rule-of-four", suggested by the Society of Radiologists in Ultrasound for the acoustic radiation force impulse (ARFI) imaging techniques of different manufacturers [7]. Additionally, it has been shown that high values obtained by LSM using ARFI techniques are predictive of decompensation and death even above the range diagnostic for cACLD [8].

Interestingly, even though there is frequently speculation about factors influencing the difference between LSMs obtained with different technologies and devices [3, 4, 9, 10, 11], systematic exploration of those factors is sparse, and especially information on the interactions between multiple confounders is lacking.

Guidelines for ultrasound-based SWE techniques suggest 5–10 (depending on the technique) LSMs to assess liver stiffness, using the interquartile range (IQR) to median (IQR/med) ratio to depict measurement variability [7]. The IQR/med ratio is commonly used as a quality criterion for the reliability of LSMs. The cutoff of IQR/med ≤ 30 % has been applied to 2 D-SWE without large-scale validation [7]. However, there is now some data justifying its use in some pSWE techniques [12]. Therefore, our study aims to systematically explore factors that impact the LSM differences between 2 D-SWE and VCTE, including the impact of variability between single measurements contributing to the final median liver stiffness value as expressed by the IQR/med ratio, the depth of measurement, biomarkers of hepatic steatosis, blood parameters related to hepatic inflammation, cholestasis, and hepatic function.

Patients and methods

To investigate factors that could explain the larger discrepancy between LSMs obtained with VCTE and 2 D-SWE, we analyzed a range of biometric, clinical, and laboratory data from individuals who underwent VCTE- and ElastQ-2 D-SWE LSM on the same day as part of a prospective European multinational comparative study. Our study population was a subset of a larger cohort that we previously examined. We included all patients who presented at the respective centers for liver elastography, encompassing both individuals with suspected liver conditions and those with prior liver disease diagnoses. In that prior research, we established

the correlation between SWE methods and defined specific cutoffs for the 2 D-SWE technique ElastQ, based on fibrosis risk groups determined by VCTE [4]. The subset analyzed in this study was selected for the availability of specific variables of interest (see below). It is crucial to note that the primary aim of our previous studies differed from the focus of this investigation. While the earlier research centered on establishing correlations and cutoffs, the current study specifically aims to assess potential confounders affecting LSM results obtained by both VCTE and ElastQ technologies. Details regarding the participating study centers and the laboratory parameter analysis methodologies employed in this study are thoroughly documented in the supplementary methods section. The shear wave elastography measurements were conducted in strict accordance with established best practices and current guidelines, as referenced in current guidelines [7]. For a comprehensive and detailed account of these measurement procedures, readers are directed to the supplementary materials. For a subgroup of patients, the fibrosis grade from liver biopsy within ± 1.5 years of the elastography examination was available.

Statistical Analysis

A description of variable summarization, grouping, statistical testing to compare groups, and the editing and statistical software used can be found in the supplementary materials.

We used a group of univariate and multivariate linear models to explore the interaction of VCTE- and ElastQ-LSMs with body mass index (BMI) and the skin-to-liver capsule distance.

To create measures of numerical difference/discrepancy between VCTE and ElastQ, we calculated (1) the absolute difference (= VCTE LSM [kPa] − ElastQ LSM [kPa]), (2) the relative difference based on VCTE-LSM

$$\left(= \frac{\text{absolute difference [kPa]}}{\text{(VCTE LSM [kPa])}}\right)$$

, and (3) the ElastQ-IQR/med

$$\left(= \frac{\text{3rd Quartile of ElastQ LSM} - \text{1st Quartile of ElastQ LSM}}{\text{(median ElastQ LSM [kPa])}}\right)$$

.

(1) is robust with respect to the direction of the difference and thus would show factors increasing the discrepancy randomly (i. e., in any direction) more strongly, while being affected by the absolute values of the VCTE- and ElastQ-LSM, so that higher LSMs, leading to a higher absolute difference, are valued more. (2) corrects for the level of liver stiffness. There is, however, an expected overlap with (1), because some LSM pairs with a positive difference are valued in the same way as in (1). (3) is aimed at exploring factors that increase the variability between single ElastQ measurements constituting the median ElastQ-LSM, which is relevant for clinical use. In the next step, the impact of various biometric measurements, laboratory parameters, and scores on the measures of difference mentioned above was evaluated in two ways: First, the characteristics of 10 % highest (top) results with regards to (1) absolute difference, (2) relative difference and (3) ElastQ-IQR/med were compared to the lowest 10 % (bottom) in each of the given measures.

Two sets of linear models assessed the impact of potential confounders and their interactions. The first set used the difference (absolute or relative) between VCTE- and ElastQ-LSM as dependent variables, with potential confounders as independent variables. The second set included VCTE-LSM as an independent variable to assess the confounders' impact alongside liver stiffness. The multivariate models used a forward selection process, $p < 0.05$ as the inclusion threshold [13]. Pre-selected independent variables included sex, age, BMI, alanine aminotransferase (ALT), aspartate aminotransferase (AST), platelets (PLT), gamma-glutamyl transferase (GGT), alkaline phosphatase (AP), albumin, international normalized ratio (INR), bilirubin, skin-to-liver capsule distance, controlled attenuation parameter (CAP), and CAP-IQR +/– VCTE-LSM. The ethical approval statement can be found in the supplementary materials.

Results

Population Characteristics

We included 875 individuals with suspected or previously diagnosed liver disease from three European tertiary care centers. Elastography assessments encountered technical failures in a small number of cases: ElastQ in 24 cases (2.7 %) and VCTE in five cases (0.6 %), with both methods failing in two instances (0.2 %). Reliable VCTE-LSMs, defined as VCTE-LSM ≤ 7.1 [14] or VCTE-IQR/med ≤ 30 % [15], were obtained from 840 (96.0 %) participants with various liver disease etiologies (▶ **Table 1**) (median age 57.0 years, BMI 14.5–45.9 kg/m^2, 53.2 % male). 816 (93.3 %) of the participants had reliable VCTE as well as reliable ElastQ, characterized by an ElastQ-LSM IQR/med of ≤ 30 % (**Supplementary Table ST-1**). Liver biopsy data was available for 78 participants (**Supplementary Table ST-2**).

Analysis of the Discrepancy between VCTE-LSM and 2 D-SWE-ElastQ-LSM

We have previously demonstrated that VCTE-LSM and 2 D-SWE-ElastQ-LSM show a significant correlation, and both are of prognostic value [4]. This is illustrated in ▶ **Fig. 1** of the present paper. ▶ **Fig. 2** and ▶ **Fig. 3**, using a Bland-Altmann-Leh analysis and a modified Bland-Altmann-Leh analysis, respectively, reveal a steady rise in the absolute discrepancy between VCTE-LSM and the 2 D-SWE-ElastQ-LSM, corresponding with rising average LSM values.

The Impact of BMI and Skin-to-liver Capsule Distance on the Correlation of VCTE and ElastQ

The correlation between BMI and the skin-to-liver capsule distance was moderate (Pearson $R = 0.58$, $p < 0.001$; Spearman $\rho = 0.63$, $p < 0.001$) and decreased with each step to a higher BMI quartile and BMI strata (**Supplementary Tables ST-3 and ST-4**).

Univariate and multivariate modeling showed that a higher BMI was significantly associated with an increase in VCTE-LSM, but not in ElastQ-LSM. Details on the statistical models that were used are described in the supplementary materials and patients section.

▶ **Table 1** Characteristics of the study population by research center. Child-Pugh score and MELD score only available for center 2. CAP measurements not available for center 1.

	Center 1 (N = 255)	Center 2 (N = 298)	Center 3 (N = 287)	Overall (N = 840)	P-value
Sex, N (%)					
▪ F	165 (64.7 %)	97 (32.6 %)	131 (45.6 %)	393 (46.8 %)	< 0.001
▪ M	90 (35.3 %)	201 (67.4 %)	156 (54.4 %)	447 (53.2 %)	
Age, years					
▪ Median [IQR]	60.0 [16.0]	52.0 [20.8]	58.0 [21.0]	57.0 [19.0]	< 0.001
BMI, kg/m²					
▪ Median [IQR]	27.1 [6.6]	25.0 [6.2]	24.6 [4.988]	25.4 [6.0]	< 0.001
Etiology, N (%)					
▪ ALD	19 (7.5 %)	60 (20.1 %)	5 (1.7 %)	84 (10.0 %)	< 0.001
▪ HBV	43 (16.9 %)	18 (6.0 %)	44 (15.3 %)	105 (12.5 %)	
▪ HCV	127 (49.8 %)	126 (42.3 %)	171 (59.6 %)	424 (50.5 %)	
▪ MASLD	55 (21.6 %)	37 (12.4 %)	41 (14.3 %)	133 (15.8 %)	
▪ Cholestatic/AIH	2 (0.8 %)	21 (7.0 %)	7 (2.4 %)	30 (3.6 %)	
▪ Other/cryptogenic	9 (3.5 %)	36 (12.1 %)	18 (6.3 %)	63 (7.5 %)	
AST, U/L					
▪ Median [IQR]	43 [28]	36 [34]	25.0 [18.0]	33 [28]	< 0.001
ALT, U/L					
▪ Median [IQR]	45 [35]	30 [29]	27.0 [25.0]	32 [32]	< 0.001
GGT, U/L					
▪ Median [IQR]	67 [90]	59 [114]	276.5 [50.0]	46 [91]	< 0.001
ALP, U/L					
▪ Median [IQR]	75 [46]	81 [48]	73 [31] 1.52 [1.21]	78 [46]	0.277
Platelet, G/L					
▪ Median [IQR]	190 [88]	164 [125]	192 [75] 73.0 [30.5]	180 [101]	< 0.001
Albumin, g/dL					
▪ Median [IQR]	37.0 [11.0]	41.6 [9.3]	44.0 [4.0] 192 [74.5]	42.5 [8.2]	< 0.001
INR					
▪ Median [IQR]	1.07 [0.16]	1.10 [0.30]	1.00 [0.09] 44.0 [4.00]	1.10 [0.30]	< 0.001
Bilirubin, mg/dL					
▪ Median [IQR]	0.98 [0.76]	0.69 [0.76]	0.75 [0.42] 1.00 [0.0900]	0.73 [0.65]	0.177
FIB-4, points					
▪ Median [IQR]	1.90 [1.80]	2.26 [3.60]	1.52 [1.21] 184 [54.5]	1.81 [2.26]	< 0.001
MELD Score, points					
▪ Median [IQR]	7 [3]	8.00 [4.00]	7 [3]	8.00 [4.00]	< 0.001

▶ **Table 1** (Continuation)

	Center 1 (N = 255)	Center 2 (N = 298)	Center 3 (N = 287)	Overall (N = 840)	P-value
Child-Pugh score category, N (%)					
▪ A	NA	227 (76.2 %)	NA	227 (26.9 %)	–
▪ B	NA	40 (13.4 %)	NA	40 (4.7 %)	
▪ C	NA	6 (2.0 %)	NA	6 (0.7 %)	
VCTE probe, N (%)					
▪ M	150 (58.8 %)	259 (86.9 %)	267 (93.0 %)	676 (80.5 %)	< 0.001
▪ XL	105 (41.2 %)	37 (12.4 %)	20 (7.0 %)	162 (19.3 %)	
VCTE-LSM, median kPa					
▪ Median [IQR]	8.0 [7.4]	11.8 [24.8]	6.0 [3.3]	7.3 [9.2]	< 0.001
VCTE-LSM IQR/ med, %					
▪ Median [IQR]	13.0 [8.7]	13.0 [10.0]	13.0 [10.0]	13.0 [9.4]	0.991
CAP, dB/m					
▪ Median [IQR]	281 [86.0]	227 [95.8]	243 [73.0]	244 [86.0]	< 0.001
CAP-IQR					
▪ Median [IQR]	NA	33 [23]	37.0 [25.0]	35 [24]	0.0429
ElastQ-LSM, median kPa					
▪ Median [IQR]	6.5 [5.6]	7.2 [6.7]	6.4 [3.2]	6.7 [5.4]	< 0.001
ElastQ-IQR/Med, %					
▪ Median [IQR]	10.2 [12.9]	9.7 [10.3]	9.876 [10.1]	10.0 [11.1]	0.556

ALD: alcoholic liver disease; AIH: autoimmune hepatitis; N: number; F: female; HBV: hepatitis B virus; HCV: hepatitis C virus; M: male; IQR: interquartile range; BMI: body mass index; AST: aspartate aminotransferase; ALT: alanine transaminase; GGT: gamma-glutamyl transferase; FIB-4: fibrosis 4-index; ALP: alkaline phosphatase; INR: international normalized ratio; VCTE: vibration-controlled transient elastography; LSM: liver stiffness measurement; MASLD: metabolic dysfunction-associated steatotic liver disease, CAP: controlled attenuation parameter; med: median; NA: not available.

Incorporating BMI and ElastQ-LSM as independent variables in our model highlighted a more pronounced effect of BMI on VCTE-LSM. Similarly, a stronger correlation between VCTE-LSM and ElastQ-LSM was evident in individuals with shorter skin-to-liver capsule distances, particularly < 1.5 cm. This correlation intensified in models accounting for ElastQ-LSM (for additional details see **Supplementary Tables ST-5 and ST-6**).

Impact of ElastQ IQR/median on the Correlation of ElastQ and VCTE

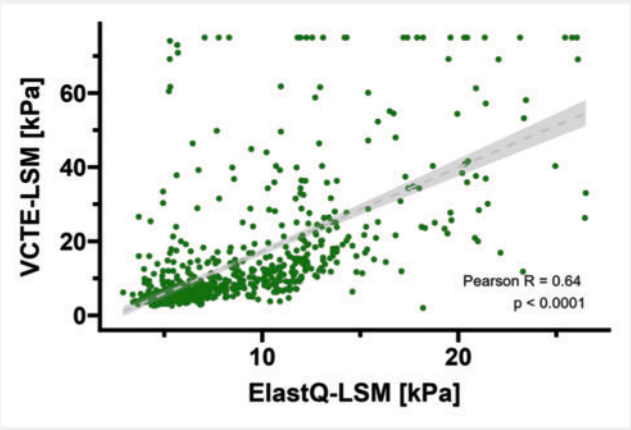
In our study, we observed a significant decrease in the correlation between ElastQ and VCTE as the interquartile range to median ratio (IQR/med) cut-offs became more lenient. The correlation peaked at $R = 0.70/\rho = 0.75$ when the ElastQ IQR/med was ≤ 10 %. In contrast, the correlation was significantly weaker at $R = 0.45/\rho = 0.41$ for ElastQ IQR/med > 30 %.

The strongest correlation was seen when combining reliable VCTE measurements and 2 D-SWE-ElastQ measurements with an IQR/med ≤ 10 %. As the IQR/med criterion for 2 D-SWE-ElastQ was relaxed, particularly beyond 30 %, this correlation progressively diminished. Accordingly, the area under the receiver operator characteristics (AUC) for the detection of VCTE-LSM > 10kPa, also decreased as the ElastQ IQR/med criterion was relaxed. The same held true for the detection of fibrosis grade ≥ F3 on liver biopsy in the subpopulation where liver biopsy was available (▶ **Table 2**, **Figure SF-1**).

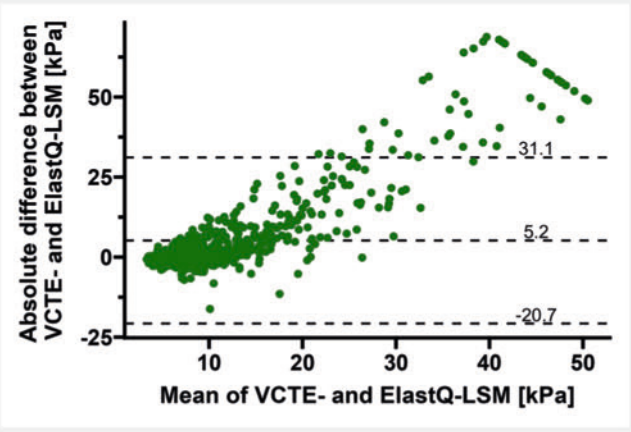
Multivariable Analysis of Factors Impacting the Correlation of VCTE and ElastQ

To explore the interaction of different factors with LSM, we used several multivariable models as described in the methods. As liver stiffness increased, so did the gap between VCTE-LSM and ElastQ-LSM. This trend was associated with indicators of advanced liver disease. Importantly, this association applied even when analyzing the absolute difference between VCTE-LSM and ElastQ-LSM, a metric unaffected by the direction of the discrepancy. More details are given in **Supplementary Tables ST-5 and ST-6** and the supplementary results.

Additionally, our analyses revealed that a higher INR, higher AP and GGT, and higher BMI were associated with a larger relative difference between VCTE and ElastQ (calculated as (VCTE – ElastQ)/VCTE). Conversely, higher albumin levels were associated with a smaller relative difference. The specific model parameters are detailed in **Supplementary Table ST-7**.



▶ **Fig. 1** Scatterplot illustrating liver stiffness measurements (LSMs) as determined by ElastQ and VCTE. The plot includes a dotted grey regression line and its 95 % confidence interval. Pearson's R value and the associated p-value are detailed in the lower right corner.

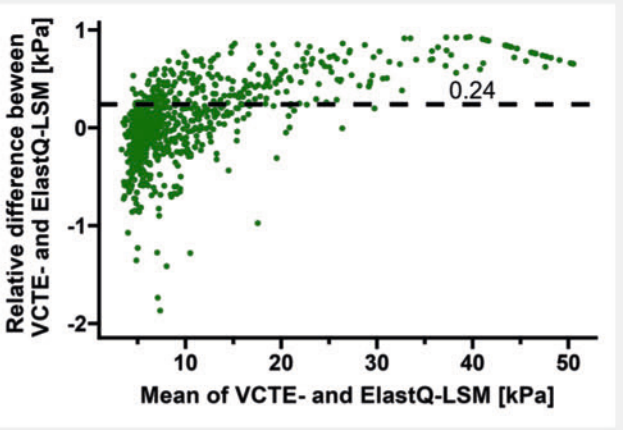


▶ **Fig. 2** Bland-Altman-Leh plot comparing LSMs from VCTE and 2 D-SWE ElastQ. The horizontal axis represents the mean values of VCTE- and 2 D-SWE ElastQ-LSM, while the vertical axis shows the absolute differences between these two measurements. The plot includes the median absolute difference (indicated by the middle dotted line, with the value specified) and the upper and lower limits of agreement (represented by the upper and lower dotted lines, with respective values provided).

Multivariable Analysis of Factors Impacting ElastQ-LSM

In our analysis aimed at determining factors that influence ElastQ-LSM, we employed VCTE-LSM as a surrogate for LSM. We then constructed a statistical model with ElastQ-LSM as the dependent variable, including VCTE-LSM and other variables as independent variables. This model revealed that higher age, GGT, and bilirubin were significantly associated with a higher ElastQ-LSM. These associations persisted even after adjusting for VCTE-LSM. Conversely, a higher PLT, greater skin-to-liver capsule distance, and ElastQ-IQR/med were significantly linked to lower ElastQ-LSM values.

The adjusted R^2 value of our model was 49.1 %, indicating that the variables included in the model accounted for approximately



▶ **Fig. 3** Modified Bland-Altman-Leh plot for LSMs from VCTE and 2 D-SWE ElastQ. The horizontal axis displays the mean values of VCTE and 2 D-SWE ElastQ-LSM, while the vertical axis presents the relative difference between these measurements (as defined in the methods section). The plot features the median absolute difference, marked by a middle dotted line, with the corresponding value provided for reference.

half of the observed variance in ElastQ-LSM. This suggests a substantial explanatory power of the model regarding ElastQ-LSM variability, but also that additional variables could explain at least part of the remaining variability. For detailed information on the model outcome parameters, please refer to **Supplementary Table ST-8**.

The results of a univariate analysis comparing the top 10 % vs. the bottom 10 % with respect to the discrepancy between VCTE and ElastQ (by different measures of discrepancy) are detailed in the supplementary results and **Supplementary Tables ST-9, ST-10, and ST-11**).

Discussion

Previous research has documented an increase in the discrepancy between LSMs obtained with ARFI-based techniques and those obtained with VCTE [3, 4, 7]. This difference is proportional to the liver stiffness. We verified this effect was present in our study cohort, underlining not only that both SWE techniques depict liver stiffness *per se*, but also that high LSM-SWE values reflect the severity of liver disease.

Factors like the size of the region of interest [10], BMI [16], and skin-to-liver capsule distance [9] are known to affect the accuracy of LSM for fibrosis evaluation. Yet, the extent, direction, and interplay of these effects are largely unexplored, primarily due to a reliance on univariate analyses in the existing literature. Therefore, we used multivariable analyses in this prospective cohort to explore factors causing significant differences between VCTE and ElastQ measurements.

We confirmed that BMI and skin-to-liver capsule distance negatively influence the agreement between VCTE and 2 D-SWE [16], finding that VCTE-LSM is more affected by these factors than ElastQ-LSM. Markers of impaired liver function and cholestasis, which correlate with advanced liver fibrosis and higher LSM, also

▶ **Table 2** Pearson correlation between ElastQ and all VCTE or all liver biopsy fibrosis grades across different IQR/med strata. P-value between correlation coefficients of the highest and the lowest stratum calculated via Fisher-Z test, and between AUC via bootstrap (N = 200). No subjects with liver biopsy with IQR/med > 30 available.

ElastQ IQR/med Strata	≤ 10 %	10–20 %	≤ 20 %	≤ 30 %	> 30	all	P-value
VCTE							
Pearson R (95 %-CI)	0.70 (0.65–0.74)	0.58 (0.50–0.66)	0.66 (0.61–0.70)	0.64 (0.60–0.68)	0.45 (0.06–0.72)	0.64 (0.59–0.67)	**0.016**
Spearman Rho (95 %-CI)	0.75 (0.69–0.80)	0.72 (0.63–0.78)	0.74 (0.69–0.78)	0.71 (0.67–0.75)	0.41 (0.00–0.72)	0.70 (0.66–0.74)	**0.002**
AUC (95 %-CI) for ACLD (VCTE > 10 kPa)	0.91 (0.88–0.95)	0.88 (0.83–0.93)	0.90 (0.88–0.93)	0.89 (0.87–0.92)	0.65 (0.42–0.88)	0.88 (0.86–0.91)	**0.035**
N (% of all N = 840)	42 (49.9 %)	277 (32.8 %)	698 (82.7 %)	812 (96.2 %)	24 (2.8 %)	844 (100 %)	–
Liver Biopsy							
Pearson R (95 %-CI)	0.42 (0.11–0.65)	0.44 (0.08–0.70)	0.40 (0.18–0.59)	0.34 (0.13–0.53)	–	0.34 (0.13–0.53)	–
Spearman Rho (95 %-CI)	0.42 (0.08–0.66)	0.44 (0.04–0.69)	0.45 (0.20–0.62)	0.36 (0.14–0.54)	–	0.36 (0.14–0.54)	–
AUC (95 %-CI) for ACLD (≥ F3)	0.83 (0.69–0.97)	0.73 (0.53, 0.94)	0.79 (0.67, 0.90)	0.73 (0.61, 0.85)	–	0.73 (0.61, 0.85)	–
N (% of all with liver biopsy N = 78)	38 (48 %)	27 (34.6 %)	65 (83.3 %)	78 (100 %)	–	78 (100 %)	–

95 %-CI: 95 %-confidence interval; ACLD: advanced chronic liver disease; AUC: area under the receiver operator characteristics curve; IQR: interquartile range; med: median; N: number.

drove the disparity between VCTE and ElastQ. Two of these multivariate models included VCTE, as a surrogate marker of liver fibrosis severity, and laboratory-based markers of liver disease severity, correcting for these important aspects.

BMI and skin-to-liver capsule distance are commonly associated with hepatic steatosis, another factor previously described to affect the discrepancy between VCTE and other 2 D-SWE techniques [17, 18]. While BMI or skin-to-liver capsule distance was associated with an increased discrepancy between VCTE-LSM and ElastQ, the VCTE-CAP – a biomarker of hepatic steatosis – was not. This suggests that the interference of the BMI and the skin-to-liver capsule distance is mainly due to the measurement depth and not hepatic steatosis.

The guideline recommendation [19] of using IQR/med with a cutoff of ≤ 30 % as a quality measure in 2 D-SWE techniques was borrowed from VCTE and pSWE, where it was validated [12]. Its applicability to all ultrasound-derived SWE techniques is uncertain. For example, the application to VCTE-CAP led to conflicting results [20]. Our study saw a significant drop in the correlation between VCTE and ElastQ at > 30 % IQR/med. Additionally, we observed a higher AUC for the diagnosis of VCTE-defined ACLD and liver biopsy-defined ACLD below ElastQ-IQR/med ≤ 30 %, providing additional evidence that IQR/med with a cutoff of ≤ 30 % might be a useful quality criterion for ElastQ and potentially other 2 D-SWE techniques.

Finally, we built a forward-selected multivariable model to explore the impact of candidate confounders that influence ElastQ-LSM independently of the level of underlying liver fibrosis, as measured by VCTE. A higher patient age and indirect biomarkers of cholestasis (higher GGT), portal hypertension (lower PLT), and liver dysfunction (higher bilirubin) were associated with higher ElastQ-LSMs. Interestingly, indirect biomarkers of liver inflammation such as AST and ALT were neither significant in univariate analysis nor included in this model, underlining that they might affect VCTE more than ElastQ. Thus, we believe that further investigations to identify further potential confounders are needed.

One of the main strengths of our study is the large prospective cohort of subjects from three European centers as well as the inclusion of patients with various etiologies of liver disease. Moreover, besides approaches targeted at known or suspected sources of discrepancy between VCTE and ElastQ, we used relatively unbiased statistical approaches, such as multivariable models, which allowed for the evaluation of multiple confounders and the description of relative effect sizes.

Our study has limitations, including having been conducted at tertiary care centers, potentially skewing results toward more severely ill patients, and possibly not being representative of the general population. Importantly, liver SWE is recommended only for populations at risk for liver fibrosis and not as a general screening tool [21].

Another limitation is the limited availability of liver biopsy data. Liver biopsies were performed only in patients for whom it was clinically indicated, thus introducing a potential selection and spectrum bias. This limitation restricts our ability to definitively confirm several aspects of our findings: 1) the independence of

the observed effects from liver fibrosis severity; 2) the precise assessment of hepatic steatosis, as grading steatosis via liver biopsy is superior to the controlled attenuation parameter (CAP); 3) which of the two techniques, VCTE or ElastQ, is more significantly influenced by potential confounders; and 4) the definitive validation of using an IQR/med ≤ 30 % as a quality criterion.

Future research could assess the impact of a progressively relaxed IQR/med quality criterion versus biopsy-defined fibrosis stages in a larger cohort to minimize risk of bias. Our study used diverse analytics but couldn't analyze operator experience effects, as all were experienced. VCTE-LSM, age, PLT, bilirubin, GGT, skin-liver distance, and ElastQ IQR/med only accounted for half of the variance in ElastQ-LSM, thus indicating other unknown factors.

Our study, primarily comprised of chronic hepatitis C patients, might have limited applicability for other liver disease etiologies. Variability in disease etiology and severity across centers could also affect results. Nonetheless, it contributes to understanding discrepancies between VCTE and 2 D-SWE techniques like ElastQ.

In conclusion, our study has identified several variables that drive the discrepancy between VCTE and ElastQ. We could describe the relative magnitude and direction of the confounding by indirect markers of liver dysfunction, such as higher INR, lower albumin, and higher bilirubin, of portal hypertension, such as lower PLT, of cholestasis, such as higher GGT and AP, and less so measures of ongoing cellular liver damage (AST, ALT) (largest standardized effect sizes for INR, PLT, and BMI). Only including cases with ALT < 5x the upper limit of normal, we could not explore the impact of higher values. Indirect indicators of measurement depth, such as BMI and skin-to-liver capsule distance but not CAP, significantly influence VCTE and ElastQ concordance, with VCTE being more affected. Our findings also support using IQR/med ≤ 30 % as a reliability criterion for 2 D-SWE techniques.

Funding Information

This research was supported by Philips Healthcare

Acknowledgement

We extend our heartfelt thanks to our colleagues for their significant contributions to patient scanning and data collection: Benedikt Simbrunner, MD; David Chromy, MD; Mathias Jachs, MD; PD Albert Stättermayer, MD; Prof. Matthias Pinter, MD; Rafael Paternostro, MD; Theresa Binter, MD; Tobias Meischl, MD (alphabetical order). Their invaluable efforts were crucial to the successful realization of this study.

Conflict of Interest

DB served as a speaker and/or consultant and/or advisory board member for AbbVie and Siemens, received travel support from AbbVie and Gilead, and received grant support from Gilead and Siemens.
ADS, LM, AR, RM: nothing to declare.
MM served as a speaker and/or consultant and/or advisory board member for AbbVie, Bristol-Myers Squibb, Gilead, Collective Acumen, and W. L. Gore & Associates and received travel support from AbbVie, Bristol-Myers Squibb, and Gilead.
IS served as speaker for AbbVie, BMS, Gilead, Janssen, Echosens, and Philips; and received advisory board fees from AbbVie, Merck; Siemens, Canon; and received research support from Philips.
GF served as a speaker for Canon Medical Systems, Fujifilm Medical Systems, Mindray Medical Systems, Philips Medical Systems, and Siemens Healthineers; served as advisory board member for Philips Medical Systems, and her institution received grant/research support from Canon Medical Systems, Esaote S.p.A., Fujifilm Medical Systems, Mindray Medical Systems, Siemens Healthineers. She receives royalties from Elsevier Publisher.
TR served as a speaker and/or consultant and/or advisory board member for AbbVie, Bayer, Boehringer Ingelheim, Gilead, Intercept, MSD, Siemens, and W. L. Gore & Associates and received grants/research support from AbbVie, Boehringer Ingelheim, Gilead, MSD, Philips, and W. L. Gore & Associates as well as travel support from Boehringer Ingelheim and Gilead.

References

[1] Foncea CG, Popescu A, Lupusoru R et al. Comparative study between pSWE and 2 D-SWE techniques integrated in the same ultrasound machine, with Transient Elastography as the reference method. Med Ultrason 2020; 22: 13–19. doi:10.11152/mu-2179

[2] de Franchis R, Bosch J, Garcia-Tsao G et al. Baveno VII – Renewing consensus in portal hypertension. J Hepatol 2022; 76: 959–974. doi:10.1016/j.jhep.2021.12.022

[3] Bauer DJ, Matic V, Mare R et al. Point Shear Wave Elastography by ElastPQ for Fibrosis Screening in Patients with NAFLD: A Prospective, Multicenter Comparison to Vibration-Controlled Elastography. Ultraschall in Med 2022; 44: 169–178. doi:10.1055/a-1724-7289

[4] Bauer DJM, Silvestri IA, Mare R et al. Two-dimensional shear wave elastography (ElastQ) accurately rules out liver fibrosis and rules in advanced chronic liver disease across liver disease etiologies: a prospective multicenter study. Ultrasonography 2023; 42: 544–554. doi:10.14366/usg.23069

[5] Ferraioli G, De Silvestri A, Lissandrin R et al. Evaluation of Inter-System Variability in Liver Stiffness Measurements. Ultraschall in Med 2019; 40: 64–75. doi:10.1055/s-0043-124184

[6] Sporea I, Bende F, Popescu A et al. Are there different cut-off values for staging liver fibrosis using 2 D-SWE implemented on different systems from the same manufacturer? Med Ultrason 2020; 22: 7–12. doi:10.11152/mu-2225

[7] Barr RG, Wilson SR, Rubens D et al. Update to the Society of Radiologists in Ultrasound Liver Elastography Consensus Statement. Radiology 2020; 296: 263–274. doi:10.1148/radiol.2020192437

[8] Trebicka J, Gu W, de Ledinghen V et al. Two-dimensional shear wave elastography predicts survival in advanced chronic liver disease. Gut 2022; 71: 402–414. doi:10.1136/gutjnl-2020-323419

[9] Uchikawa S, Kawaoka T, Fujino H et al. The effect of the skin-liver capsule distance on the accuracy of ultrasound diagnosis for liver steatosis and fibrosis. J Med Ultrason (2001) 2022; 49: 443–450. doi:10.1007/s10396-022-01210-w

[10] Wang H, Zheng P, Wang X et al. Effect of Q-Box size on liver stiffness measurement by two-dimensional shear wave elastography. J Clin Ultrasound 2021; 49: 978–983. doi:10.1002/jcu.23075

[11] Grgurevic I, Salkic N, Bozin T et al. Magnitude dependent discordance in liver stiffness measurements using elastography point quantification with transient elastography as the reference test. Eur Radiol 2019; 29: 2448–2456. doi:10.1007/s00330-018-5831-2

[12] Ferraioli G, De Silvestri A, Reiberger T et al. Adherence to quality criteria improves concordance between transient elastography and ElastPQ for liver stiffness assessment-A multicenter retrospective study. Dig Liver Dis 2018; 50: 1056–1061. doi:10.1016/j.dld.2018.03.033

[13] Aravind H. Tools for Building OLS Regression Models. In: 0.5.3 edn. Aravind, Hebbali 2020

[14] Schwabl P, Bota S, Salzl P et al. New reliability criteria for transient elastography increase the number of accurate measurements for screening of cirrhosis and portal hypertension. Liver Int 2015; 35: 381–390. doi:10.1111/liv.12623

[15] EASL. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol 2015; 63: 237–264. doi:10.1016/j.jhep.2015.04.006

[16] Petta S, Di Marco V, Camma C et al. Reliability of liver stiffness measurement in non-alcoholic fatty liver disease: the effects of body mass index. Aliment Pharmacol Ther 2011; 33: 1350–1360. doi:10.1111/j.1365-2036.2011.04668.x

[17] Kumada T, Toyoda H, Yasuda S et al. Liver Stiffness Measurements by 2 D Shear-Wave Elastography: Effect of Steatosis on Fibrosis Evaluation. AJR Am J Roentgenol 2022; 219: 604–612. doi:10.2214/AJR.22.27656

[18] Huang Z, Zhou J, Lu X et al. How does liver steatosis affect diagnostic performance of 2 D-SWE.SSI: assessment from aspects of steatosis degree and pathological types. Eur Radiol 2021; 31: 3207–3215. doi:10.1007/s00330-020-07288-5

[19] Dietrich CF, Bamber J, Berzigotti A et al. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). Ultraschall in Med 2017; 38: e48. doi:10.1055/a-0641-0076

[20] Eddowes PJ, Sasso M, Allison M et al. Accuracy of FibroScan Controlled Attenuation Parameter and Liver Stiffness Measurement in Assessing Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2019; 156: 1717–1730. doi:10.1053/j.gastro.2019.01.042

[21] EASL. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis – 2021 update. J Hepatol 2021; 75: 659–689. doi:10.1016/j.jhep.2021.05.025