Applicability and variability of liver stiffness measurements according to probe position

Patrick Ingiliz, Kim Pav Chhay, Mona Munteanu, Pascal Lebray, Yen Ngo, Dominique Roulot, Yves Benhamou, Dominique Thabut, Vlad Ratziu, Thierry Poynard

Abstract

AIM: To investigate the liver stiffness measurement (LSM) applicability and variability with reference to three probe positions according to the region of liver biopsy.

METHODS: The applicability for LSM was defined as at least 10 valid measurements with a success rate greater than 60% and an interquartile range/median LSM < 30%. The LSM variability compared the inter-position concordance and the concordance with FibroTest.

RESULTS: Four hundred and forty two consecutive patients were included. The applicability of the anterior position (81%) was significantly higher than that of the reference (69%) and lower positions (68%), (both \( P = 0.0001 \)). There was a significant difference (0.5 kPa, 95% CI 0.13-0.89; \( P < 0.0001 \)) between mean LSM estimated at the reference position (9.3 kPa) vs the anterior position (8.8 kPa). Discordance between positions was associated with thoracic fold (\( P = 0.008 \)). The discordance rate between the reference position result and FibroTest was higher when the 7.1 kPa cutoff was used to define advanced fibrosis instead of 8.8 kPa (33.6% vs 23.5%, \( P = 0.03 \)).

CONCLUSION: The anterior position of the probe should be the first choice for LSM using Fibroscan, as it has a higher applicability without higher variability compared to the usual liver biopsy position.

Key words: Fibroscan; Fibrotest; Liver fibrosis; Variability; Concordance

INTRODUCTION

A major clinical challenge is to find the best method to evaluate and to manage the increasing numbers of patients with chronic liver disease\(^1-4\). Liver biopsy, due to its risks and limitations, is no longer considered mandatory as the first-line indicator of liver injury, and several markers have been developed as non-invasive alternatives\(^1-4\).

The assessment of liver fibrosis by non-invasive techniques such as biomarkers, [FibroTest\(^8\) (FT)]\(^9\) and liver stiffness measurement (LSM) by Fibroscan\(^6,7\), is now widely performed in countries where these techniques are available and approved\(^6,9\). It is therefore essential to identify factors associated with a variability of the results of these techniques to reduce the risk of false positives or false negatives. There are no published procedures for the most accurate position of the probe in LSM. In almost all
The described method is copied from the original description by Sandrin et al.[19]: “Because liver biopsies are performed on the right lobe of the liver, so were the elasticity measurements. During the acquisition, patients were lying on their backs with their right arms behind their heads. The physician first proceeded to a sonographic examination to localize the best ultrasonic imaging window between the rib bones. Regions with large vessels were avoided and a minimal liver parenchyma thickness of 6 cm was sought”.

A total of 468 consecutive patients were pre-included, the described method is copied from the original description by Sandrin et al.[19]: “Because liver biopsies are performed on the right lobe of the liver, so were the elasticity measurements. During the acquisition, patients were lying on their backs with their right arms behind their heads. The physician first proceeded to a sonographic examination to localize the best ultrasonic imaging window between the rib bones. Regions with large vessels were avoided and a minimal liver parenchyma thickness of 6 cm was sought”.

Several studies have examined the variability possibly associated with different positions in the rather vaguely defined area called “the liver biopsy zone”. The variability associated with position could be part of the interobserver effect. Only two published studies have estimated the interobserver effect: Sandrin et al.[19] studied 10 patients involving 3 operators (standardized CV 3.3%) and Coco et al.[20] compared 2 operators in 40 patients using correlation coefficients (0.92) and paired t-tests. Tanne et al.[21] also observed a significant discordance (25%) between predicted fibrosis stages according to three different positions of the probe and suggested using three different positions, to reduce the “sampling error”.

We previously compared 9 different positions in the right lobe in 35 healthy subjects with the same operator and observed a very significant variability[18]. Three positions were therefore selected according to their applicability: the reference position, an anterior position and a lower position.

The aims of this study were to compare the applicability of these three positions, their inter-position concordance, and their concordance alone and relative to FT, a reference biomarker of fibrosis.

**MATERIALS AND METHODS**

Consecutive patients with chronic liver disease seen in the Hepatology Department of the Pitié-Salpêtrière Hospital in Paris, France were pre-included to undergo LSM and FT. Patients were not included if they did not accept the protocol, or if the quality requirements for FT were not achieved. All patients gave informed consent for the use of data and serum for research purposes in this non-interventional clinical study, which was approved by the local institutional review board. The study protocol was in accordance with the ethical guidelines of the Declaration of Helsinki.

**Biochemical markers**

FibroTest and ActiTest (Biopredictive, Paris, France) were performed according to published recommendations[5,19,28]. The operating algorithm was defined area called “the liver biopsy zone”. The variability associated with position could be part of the interobserver effect. Only two published studies have estimated the interobserver effect: Sandrin et al.[19] studied 10 patients involving 3 operators (standardized CV 3.3%) and Coco et al.[20] compared 2 operators in 40 patients using correlation coefficients (0.92) and paired t-tests. Tanne et al.[21] also observed a significant discordance (25%) between predicted fibrosis stages according to three different positions of the probe and suggested using three different positions, to reduce the “sampling error”.

We previously compared 9 different positions in the right lobe in 35 healthy subjects with the same operator and observed a very significant variability[18]. Three positions were therefore selected according to their applicability: the reference position, an anterior position and a lower position.

The aims of this study were to compare the applicability of these three positions, their inter-position concordance, and their concordance alone and relative to FT, a reference biomarker of fibrosis.

**Liver stiffness measurements**

LSM was performed with the non-invasive method of transient elastography (FibroScan, Echosens, Paris, France). The stiffness results were expressed in kilopascal (kPa). The technique was performed by two trained (more than 100 measurements) senior hepatologists, blinded to all other characteristics, and according to the manufacturers’ recommendations. During the acquisition, patients were lying on their backs with their right arms behind their heads. The physician first proceeded to a sonographic examination to localize the best ultrasonic imaging window between the rib bones. Regions with large vessels were avoided and a minimal liver parenchyma thickness of 6 cm was sought.

The reference position was the region usually recommended for biopsy located at the intersection between the xiphoid line and the median axillary line, where the operator would have performed the biopsy. The second position (lower position) was a more posterior position 2-3 cm in the next intercostal space on the same xiphoid line as the reference position and the third position (anterior position) was an anterior position 2-3 cm ahead of the reference position in the same intercostal space.

Two of the most commonly recommended cutoffs for advanced fibrosis (F2, F3 and F4 in METAVIR staging)[7,9,12,13] were used: 7.1 kPa[8] and 8.8 kPa[10].

**Applicability**

The applicability for LSM was defined as: a success rate greater than 60% (SR60)[7,9,12,13] at least 10 valid liver stiffness measurements (V10)[7,9,12,13] and an interquartile range/median LSM < 30% (IQR30)[7,9,14-19].

The applicability for FT was defined as: a security algorithm profile excluding Gilbert’s disease, hemolysis, acute inflammation profiles and extreme values of FT components, leading to a change of at least 0.30 in the FT result if the median value of each component was used[7,8].

**Statistical analysis**

The two main endpoints were the applicability rate and the discordance rate with FT, for the two new positions vs the reference position. Because of the number of statistical comparisons for these two endpoints for two positions, a P value lower than 0.01 has been taken for a significant difference.

The strength of concordance between each LSM, or their combinations, and FT was assessed using three methods, the kappa reliability test (K) for 2 fibrosis stages (advanced vs non-advanced fibrosis), the Spearman rank correlation coefficient (R), and the intraclass coefficient of correlation (ICC)[22].

Applicabilities were compared using Chi square and Fisher’s exact tests, quantitative variables were compared using Mann-Whitney test, Wilcoxon signed rank test for paired comparisons, and multivariate analysis using logistic regression analysis. All comparisons were performed separately with subpopulations of operator 1 and 2, as well as with the population of patients with all positions applicable and populations with at least one position applicable. Analyses were performed with NCSS software (Kaysville, Utah, USA)[23].

**RESULTS**

A total of 468 consecutive patients were pre-included
between April and September 2007. Twenty-six patients were not included and 442 patients were included (Table 1). There was no difference between included and non-included patient characteristics.

**Applicability**
The applicability of LSM according to position is described in Table 2. The applicability of the anterior position (81%) was significantly higher than that of the reference (69%) and lower positions (68%), (both P = 0.0001). These differences in applicability were mainly due to an IQR30 obtained more often with the anterior position than with the reference and lower positions for operator 1 (82% vs 67%; P = 0.0001; 82% vs 73%; P = 0.0004) and for operator 2 (91% vs 76%; P = 0.36). Presumed fibrosis stage (META VIR scoring system) using 7.7 kPa for F2, 8.8 kPa for F3 and 14.5 kPa for F4; *Presumed prevalence of non-advanced fibrosis was lower using the reference position than the anterior position (P = 0.04).

**Liver stiffness measurements between positions**
Among 268 patients with both anterior and reference positions applicable, the mean LSM estimated at the reference position [9.0 kPa (0.5)] was significantly higher in comparison to the anterior position [8.5 kPa (0.5); P < 0.0001].

There was no significant difference between LSM measured at the reference in comparison to the lower position [9.5 kPa (0.5) vs 9.3 kPa (0.5); n = 253, P = 0.36].

**Presumed prevalence of fibrosis**
Among 268 patients with both applicable anterior and reference positions, using a 7.1 kPa cutoff, 121/268 (45%) of patients had advanced fibrosis using the reference position vs 102/268 (38%) using the anterior position (P = 0.10). Using an 8.8 kPa cutoff, 73/268 (27%) of patients had advanced fibrosis using the reference position vs 58/268 (24%) using the anterior position (P = 0.40).

When the prevalence of presumed fibrosis stages was compared according to the probe position of all applicable patients, prevalence of non-advanced fibrosis (7.1 kPa cutoff) was lower using the reference position (55%) than the anterior position (62%, P = 0.04) (Table 2).
Using an 8.8 kPa cutoff, there was no difference between the prevalence of non-advanced fibrosis using the reference position (226/306, 74%) than the anterior position (275/357, 77%; P = 0.34).

**Concordance between positions**

The discordance rates and strength of concordance between LSM assessed in three positions are detailed in Table 3. The discordance rate between the anterior and the lower probe positions was higher (17.3%) when the 7.1 kPa cutoff was used to define advanced fibrosis, instead of 8.8 kPa (10.2%; P = 0.04), and for the anterior vs the reference position (17.9% vs 11.3%; P = 0.06). There was no significant operator effect.

The factors significantly associated with discordance between the reference and the anterior positions were thoracic fold (P = 0.0008) thickness and non-alcoholic fatty liver disease (NAFLD) as the cause of liver disease (P = 0.008) (Table 4). BMI (P = 0.02), abdominal (P = 0.03) and waist circumference (0.047), and SteatoTest (P = 0.04) were not significantly associated when protected for multiple statistical comparisons (Table 4). In multivariate analysis, only thoracic fold was significantly associated with position discordance (regression coefficient \(\beta = 0.07; 95\% \text{ CI } 0.02-0.13; \ P = 0.01\). Same results were observed in the population with three positions applicable.

**Concordance with FT**

Discordance rates and strength of concordance between LSM assessed in three positions and FT are detailed in Table 5. There were no significant differences between the discordance rates and the strength of concordance between the three probe positions at a sufficient P value protected for multiple testing.

The discordance rates between probe positions and FT were higher when the 7.1 kPa cutoff was used to define advanced fibrosis instead of 8.8 kPa for the reference position (33.6% vs 23.5%, P = 0.03) in the 196 patients with all 3 positions applicable and also among the 306 patients with only the reference position applicable (34.9% vs 26.8%, P = 0.03).

The mean of the 3 positions (a total of 30 LSM), did not increase the strength of concordance with FT.

**DISCUSSION**

This study provides an improved assessment of the variability of LSM due to the position of the probe in the right liver lobe. We confirmed the preliminary results we had observed in 35 healthy subjects, in whom 9 different positions had been assessed.

The diagnostic value of LSM and FT has been validated in the most common chronic liver diseases and FT has shown at least a similar prognostic value as liver biopsy (which is also an imperfect gold-standard) in patients with chronic hepatitis C and D. We demonstrated previously that the strength of concordance between LSM and FT could be used to identify LSM...
The results strongly suggest that the reference position for LSM has two weaknesses in comparison with a more anterior position: a significantly lower applicability and a possible higher variability for the diagnosis of advanced fibrosis using the 7.1 kPa cutoff. The third position analyzed at a lower level compared to the reference position had no advantage either in terms of applicability or in strength of concordance with FT.

The mean LSM was significantly lower (0.5 kPa) at the anterior position of the reference position. This difference was also clinically significant. When using the anterior position instead of the reference position, 7% of patients changed status from advanced fibrosis to non-advanced fibrosis when a cutoff of 7.1 kPa was chosen. The difference of 0.5 kPa is particularly clinically relevant in the zone of 7 to 9 kPa for the risk of a false negative/positive diagnosis of advanced fibrosis; it is less relevant for the diagnosis of cirrhosis as LSM cutoffs are usually recommended at a 12.5 kPa or 14 kPa cutoff with a range to 75 kPa. From these data it is possible to say that the reference position using 7.7 kPa cutoff for F2 and 8.8 kPa for F3 increases the risk of false positive conclusions in comparison with 8.8 and 12.5 kPa cutoffs, respectively.

Several anthropometric factors were associated with discordance between the reference and the anterior positions but the most significant factor was the thoracic skin fold thickness. More studies must now be conducted to better understand the role of these anthropometric factors both with regard to the applicability and to the variability of LSM.

Improved knowledge of LSM variability is also important for the definition of normal values of LSM. In contrast to FT, very few studies have assessed the normal range of LSM with biopsy without fibrosis (F0 in the METAVIR scoring system). Roulot et al. proposed the 95% percentile of a healthy non-obese population as the upper normal limit; 7.8 for females and 8.0 kPa for males. If these definitions of normal range are widely validated, several anthropometric factors both with regard to the applicability and to the variability of LSM.

Table 5  Strength of concordance between LSM and FibroTest (FT) according to positions

<table>
<thead>
<tr>
<th>Position (No. of patients)</th>
<th>Method assessing Discordance rate (%)</th>
<th>Kappa mean (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative concordance</td>
<td>Spearman mean (95% CI)</td>
<td>Intra class coefficient mean (95% CI)</td>
</tr>
<tr>
<td>All positions applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference (196)</td>
<td>0.46 (0.34-0.56)</td>
<td>0.55 (0.33-0.67)</td>
</tr>
<tr>
<td>Anterior (196)</td>
<td>0.46 (0.34-0.56)</td>
<td>0.56 (0.34-0.68)</td>
</tr>
<tr>
<td>Lower (196)</td>
<td>0.40 (0.27-0.51)</td>
<td>0.50 (0.38-0.62)</td>
</tr>
<tr>
<td>Mean of positions (196)</td>
<td>0.47 (0.35-0.57)</td>
<td>0.56 (0.34-0.68)</td>
</tr>
<tr>
<td>At least one position applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference (306)</td>
<td>0.44 (0.35-0.53)</td>
<td>0.51 (0.39-0.63)</td>
</tr>
<tr>
<td>Anterior (357)</td>
<td>0.46 (0.38-0.54)</td>
<td>0.54 (0.32-0.66)</td>
</tr>
<tr>
<td>Lower (302)</td>
<td>0.39 (0.29-0.49)</td>
<td>0.50 (0.38-0.62)</td>
</tr>
<tr>
<td>Two classes concordance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutoff 7.1 kPa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All positions applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference (196)</td>
<td>66 (33.6)</td>
<td>0.30 (0.07)</td>
</tr>
<tr>
<td>Anterior (196)</td>
<td>61 (31.1)</td>
<td>0.32 (0.07)</td>
</tr>
<tr>
<td>Lower (196)</td>
<td>71 (36.2)</td>
<td>0.24 (0.07)</td>
</tr>
<tr>
<td>Mean of positions (196)</td>
<td>67 (34.2)</td>
<td>0.28 (0.07)</td>
</tr>
<tr>
<td>At least one position applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference (306)</td>
<td>107 (34.9)</td>
<td>0.28 (0.06)</td>
</tr>
<tr>
<td>Anterior (357)</td>
<td>112 (31.6)</td>
<td>0.33 (0.05)</td>
</tr>
<tr>
<td>Lower (302)</td>
<td>109 (36.1)</td>
<td>0.24 (0.06)</td>
</tr>
<tr>
<td>Cutoff 8.8 kPa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All positions applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference (196)</td>
<td>46 (23.5)</td>
<td>0.45 (0.07)</td>
</tr>
<tr>
<td>Anterior (196)</td>
<td>54 (27.6)</td>
<td>0.33 (0.07)</td>
</tr>
<tr>
<td>Lower (196)</td>
<td>56 (28.6)</td>
<td>0.34 (0.07)</td>
</tr>
<tr>
<td>Mean of positions (196)</td>
<td>52 (26.5)</td>
<td>0.37 (0.07)</td>
</tr>
<tr>
<td>At least one position applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference (306)</td>
<td>82 (26.8)</td>
<td>0.38 (0.06)</td>
</tr>
<tr>
<td>Anterior (357)</td>
<td>108 (30.2)</td>
<td>0.30 (0.05)</td>
</tr>
<tr>
<td>Lower (302)</td>
<td>84 (27.8)</td>
<td>0.34 (0.06)</td>
</tr>
</tbody>
</table>

$^{a,b}$P = 0.03 between 7.1 and 8.8 kPa.
therefore the two operators measured LSM in different patients. There was a difference between operators for the applicability rate of the reference position due to a lower IQR30 percentage. This was not a systematic operator effect, as this lower IQR30 percentage was not observed for the anterior or the lower positions. Operator 1 had twice as many NAFLD patients [50/329 (15.2%)] as operator 2 [9/113 (8%) \( P = 0.05 \)], which could explain the greater variability of LSM and lower applicability in comparison to operator 2. As with other authors\(^7,12,13,14\), we previously observed that the non-applicability and the variability of LSM at the reference position were higher in patients with NAFLD vs non-NAFLD patients.

We acknowledge that the number of comparisons increased the risk of false positives. The comparison between the strength of concordance anterior position-FT and the strength of concordance reference position-FT did not reach a high statistical significance (\( P < 0.01 \)). However, all the comparisons indicated the same direction and at least a lower concordance of the anterior position with FT in comparison with the reference position can be excluded.

In conclusion, our results suggest that the anterior position of the probe, 2-3 cm ahead of the usual position of liver biopsy, should be the first choice for LSM using FibroScan for liver fibrosis estimates. Compared with the reference position, the anterior position improved the applicability of FibroScan without decreasing its concordance with FibroTest.

**REFERENCES**

2. Afzhal NH. Diagnosing fibrosis in hepatitis C: is the pendulum swinging from biopsy to blood tests? *Hepatology* 2003; 37: 972-974
9. La Haute Autorité de Santé (HAS) in France - The HAS recommendations for the management of the chronic hepatitis C using non-invasive biomarkers. Available from: URL: http://www.has-sante.fr/portail/display.jsp?id_c=476486
14. Coco B, Oliveri F, Maina AM, Ciccorossi P, Sacco R, Colombatto P, Bonino F, Brunetto MR. Transient...
elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. J Viral Hepat 2007; 14: 360-369


22 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-174


