ARFI cut-off values and significance of standard deviation for liver fibrosis staging in patients with chronic liver disease

Ruediger S. Goertz,* Joerg Sturm,* Lukas Pfeifer,* Dane Wildner,* David L. Wachter,** Markus F. Neurath,* Deike Strobel*

* Department of Internal Medicine 1, University of Erlangen, Ulmenweg 18, 91054 Erlangen, Germany.
** Institute of Pathology, University of Erlangen, Universitätstrasse 22, 91054 Erlangen, Germany.

ABSTRACT

Background. Acoustic radiation force impulse (ARFI) elastometry quantifies hepatic stiffness, and thus degree of fibrosis, non-invasively. Our aim was to analyse the diagnostic accuracy of ARFI cut-off values, and the significance of a defined limit of standard deviation (SD) as a potential quality parameter for liver fibrosis staging in patients with chronic liver diseases (CLD). Material and methods. 153 patients with CLD (various aetiologies) undergoing liver biopsy, and an additional 25 patients with known liver cirrhosis, were investigated. ARFI measurements were performed in the right hepatic lobe, and correlated with the histopathological Ludwig fibrosis score (inclusion criteria: at least 6 portal tracts). The diagnostic accuracy of cut-off values was analysed with respect to an SD limit of 30% of the mean ARFI value. Results. The mean ARFI elastometry showed 1.95 ± 0.87 m/s (range 0.79-4.40) in 178 patients (80 female, 98 male, mean age: 52 years). The cut-offs were 1.25 m/s for F ≥ 2, 1.72 m/s for F ≥ 3 and 1.75 m/s for F = 4, and the corresponding AUROC 80.7%, 86.2% and 88.7%, respectively. Exclusion of 31 patients (17.4%) with an SD higher than 30% of the mean ARFI improved the diagnostic accuracy: The AUROC for F ≥ 2, F ≥ 3 and F = 4 were 86.1%, 91.2% and 91.5%, respectively. Conclusion. The diagnostic accuracy of ARFI can be improved by applying a maximum SD of 30% of the mean ARFI as a quality parameter - which however leads to an exclusion of a relevant number of patients. ARFI results with a high SD should be interpreted with caution.


INTRODUCTION

Chronic hepatic injury may create hepatic fibrosis, and finally cirrhosis, leading to serious morbidity or mortality. Currently, liver biopsy is proposed as the standard procedure for staging fibrosis in patients with chronic liver disease (CLD). However, this procedure is invasive and has the potential to generate severe complications occurring in up to 0.4% of the cases.1 Prompted by these limitations of liver biopsy, the prediction of fibrosis in CLD via non-invasive methods has emerged as an important field of research. Extracellular matrix remodelling proteins and also plasmatic indices (such as AST/platelet ratio index = APRI or Fibrotest) failed to differentiate between the various degrees of liver fibrosis.2,3 Doppler measurements of the hepatic vessels are neither a valid surrogate marker of liver fibrosis nor a useful method for estimating the degree of liver fibrosis.4

Recently, elastographic ultrasound techniques have become evermore popular. The well-evaluated technique of transient elastography reveals a good correlation of liver stiffness with advanced hepatic fibrosis,5,6 but is limited in patients with severe obesity and ascites. The technology of acoustic radiation force impulse (ARFI) elastometry is a novel ultrasonic method that provides information about the local mechanical properties of tissue in vivo.7 ARFI elastometry can be applied as a real-time examination, since it is incorporated in a conventional ultrasound machine. A number of studies have explored hepatic shear wave velocities in comparison with various degrees of hepatic fibrosis.8-11
An initial meta-analysis of ARFI attempted to define cut-off values for hepatic fibrosis, but only 9 out of a wealth of manuscripts could be included because of a lack of a standardised measurement acquisition method, and poor biopsy quality. Moreover, an ARFI quality parameter is not yet in place, since measurements by transient elastography should be within an interquartile range of less than 30%. In the literature, a similar quality criterion has been proposed for ARFI elastometry, but has not yet been critically evaluated.

The aims of the present study were first to assess the performance of ARFI elastometry of the liver so as to define cut-off values for relevant hepatic fibrosis stages, and secondly to evaluate the impact of a maximum standard deviation (SD) of 30% of the mean on ARFI accuracy in a cohort of patients with CLD providing adequate hepatic specimens (portal tracts ≥ 6).

**MATERIAL AND METHODS**

Between December 2008 and September 2012, patients with CLD undergoing liver biopsy providing an adequate hepatic specimen (portal tract count ≥ 6) were included in the study. An additional 25 patients with clinically and sonographically confirmed hepatic cirrhosis were added later. Informed consent was obtained, and the project approved by the local ethics committee. Demographic data (age, gender) were obtained from all patients, as well as laboratory data (aspartate aminotransferase (AST) and platelet count) to assess the APRI score (AST/platelet ratio).

**Histological grading and staging**

The biopsy specimen of the liver was obtained with a core tissue needle having an external diameter of 1.3 mm and a sample notch of 2.2 cm in two passes of the needle. The specimens were fixed in formalin (10%) and then sent to the pathology department. All specimens were analysed by an experienced pathologist who was blinded to the patients’ elastometry results. Histological features of liver fibrosis were interpreted in accordance with the Ludwig score classification F0-F4. F0 indicates absence of fibrosis, F1 minimal portal fibrosis, F2 moderate periportal fibrosis with no architectural distortion, F3 substantial septal fibrosis with architectural distortion, F4 cirrhosis.

**ARFI measurements**

ARFI measurements were performed at 4 MHz with a curved array transducer (4C1) of the ultrasound system Acuson S2000 (Siemens Medical Solutions, Erlangen, Germany) in the Virtual touch tissue quantification mode. During real-time B-mode imaging a region of interest (10 x 5 mm) was selected in the liver parenchyma such as to avoid large vessels and an average of seven measurements were conducted in the right hepatic lobe using an intercostal approach during relaxed breath-holding. The mean and SD was documented. All biopsied patients were in the fasting state.

The ARFI technique uses short-duration acoustic radiation force pulses to generate localised displacements in the tissue. These displacements result in shear-wave propagation away from the region of excitation and are tracked using ultrasonic correlation-based methods. The tissue response to these acoustic pulses can be monitored both spatially and temporally. Displacement magnitude is inversely proportional, and shear-wave velocity direct proportional, to local tissue stiffness. A single transducer is used both to apply localised radiation forces within tissue for short time periods, and to track the resulting tissue displacements. Shear-wave propagation velocity is proportional to the square root of tissue elasticity. Results are expressed in metres per second (m/s).

**Statistical analysis**

Statistical analysis was performed using the Statistical Package for the Social Sciences (version 19.0.0.1, IBM SPSS statistics, New York, USA). Clinical and laboratory characteristics of the patients, as well as ARFI measurements, were expressed as means ± SD or as medians (with the range). Relationships between variables were examined using the Spearman’s correlation coefficient (r). In addition to the overall cohort, two subgroups were also analysed: subgroup A including patients with a 30% maximum SD of the individual mean ARFI value, and subgroup B including patients with an SD > 30% of the individual mean ARFI value. To compare the medians the Mann-Whitney-U-Test was used.

The results were presented as boxplots with the median as a thick line through each box representing the interquartile range within which 50% of values are located. Error bars mark minimum and maximum values (range). Small circles or stars mark outliers. We assessed the diagnostic perfor-
performance of hepatic ARFI elastometry by calculating the areas under the receiver operating characteristic (AUROC) curves. ARFI cut-off values for the relevant stages of liver fibrosis ($F \geq 2$, $F \geq 3$, $F \geq 4$) were optimised using the Youden’s index. All reported $p$ values are two-sided. A $p < 0.05$ indicated a significant correlation or difference.

RESULTS

A total of 178 patients (80 females, 98 males, mean age 52 years) met the inclusion criteria. The main aetiologies of CLD were hepatitis C ($n = 45$, 25.3%), hepatitis B ($n = 24$, 13.5%), auto-immune such as AIH, PBC, PSC ($n = 24$, 13.5%) and alcoholic ($n = 23$, 12.9%). The mean body mass index (BMI) was $26.2 \pm 4.8$ kg/m$^2$ (range: 16.0-43.2) (21 data unavailable) and the mean platelet count 223.6 ± 97 x $10^9$/L (range: 53-635 x $10^9$/L) and histologically, the mean specimen length 27.47 ± 9.7 mm (range: 10-54 mm) (7 data unavailable) (Table 1).

The mean ARFI values were $1.95 \pm 0.87$ m/s (range: 0.79-4.40 m/s) and correlated well with the histological classification of the fibrosis ($r = 0.670$, $p < 0.001$) (Figure 1A) as also with patient age ($r = 0.291$, $p < 0.001$), AST levels ($r = 0.151$, $p = 0.046$), APRI score ($r = 0.318$, $p < 0.001$), but negatively with the platelet count ($r = -0.397$, $p < 0.001$). The degree of fibrosis correlated positively with patient age ($r = 0.267$, $p < 0.001$), level of AST ($r = 0.308$, $p < 0.001$), and the APRI score ($r = 0.439$, $p < 0.001$), and negatively with the platelet count ($r = -0.472$, $p < 0.001$). The assessment accuracy of different degrees of fibrosis by means of ARFI is expressed in

| Table 1. Data (median with range) of the patients overall and of two subgroups A + B comprising only those with a standard deviation (SD) ≤ 30% or > 30% of the mean ARFI value. |
|-----------------|-----------------|-----------------|
|                 | All             | Subgroup A      | Subgroup B      |
|                 | (SD ≤ 30%)      | (SD > 30%)      |
| f/m             | 80/98           | 64/83           | 16/15           |
| Age (y)         | 52 (19-78)      | 50 (19-78)      | 54 (33-73)      |
| AST (U/l)       | 52 (15-2414)    | 55 (15-2414)    | 40 (17-321)     |
| APRI            | 0.54 (0.1-45.68)| 0.62 (0.18-45.68)| 0.39 (0.1-4.41) |
| Portal tracts   | 10 (6-26)       | 10 (6-26)       | 10 (6-17)       |
| F0              | 17 (9.6%)       | 15 (10.2%)      | 2 (6.5%)        |
| F1              | 43 (24.2%)      | 30 (20.4%)      | 13 (41.9%)      |
| F2              | 38 (21.3%)      | 33 (22.5%)      | 5 (16.1%)       |
| F3              | 25 (14.0%)      | 19 (12.9%)      | 6 (19.4%)       |
| F4              | 55 (30.9%)      | 50 (34.0%)      | 5 (16.1%)       |
| Total           | 178             | 147             | 31              |

![Figure 1. ARFI elastometry and fibrosis stage of A. All patients ($r = 0.670$) and B. Subgroup A - only patients with a standard deviation of maximum 30% ($r = 0.748$). Horizontals represent cut-offs for $F \geq 2$ and $F = 4$.](image-url)

938

the AUROC analyses. The AUROC values for relevant degrees of hepatic fibrosis range from 80.7 to 91.5% (Table 2).

Out of the total 178 patients, 31 (17.4%) had an SD of more than 30% of the individual mean ARFI (subgroup B). Almost half the patients in subgroup B had fibrosis grade F0 or F1. These patients of subgroup B (mean BMI 26.9 ± 4.5) had a mean individual SD of 0.88 with a mean individual SD of 0.90 for F ≤ 2 and 0.84 for F ≥ 3. The ARFI elastometry returned particularly high values at 1.98 ± 0.89 m/s for F = 1, and 2.35 ± 0.80 m/s for F = 2 (Table 3). Neither ARFI shear wave velocity (r = 0.338, p = 0.063) nor the intra-individual SD (r = 0.125; p = 0.503) correlated with hepatic fibrosis.

Analysis of the remaining 147 patients (subgroup A) with a 30% maximum SD revealed a mean BMI of 26.1 ± 4.9 and a mean ARFI value of 1.92 ± 0.89 m/s (range: 0.79-4.40 m/s). ARFI values correlated significantly with SD (r = 0.751, p < 0.001) and hepatic fibrosis (r = 0.748, p < 0.001). Subgroup A showed a mean individual SD of 0.17 for F ≤ 2 and of 0.40 for F ≥ 3. Subgroup B showed non-significantly lower levels of AST (p = 0.056), but significantly lower levels of the APRI score (p = 0.006) in comparison with subgroup A. There were no differences in patient age and number of portal tracts.

Comparative ARFI data of all patients and subgroup A including values for cut-off, sensitivity and specificity for relevant fibrosis stages, are shown head to head to the analysis in table 2. The diagnostic accuracy of ARFI for the assessment of a significant liver fibrosis (F ≥ 2) was improved in subgroup A, which included only patients with an SD ≤ 30% of the individual mean ARFI. In addition, individual absolute SD values seemed to correlate with liver fibrosis to a higher degree in subgroup A (r = 0.592, p < 0.001) than in patients as a whole (r = 0.370, p < 0.001).

Exclusion of 65 patients (36.5%) with an SD above 20% of the mean ARFI led to a higher correlation coefficient (r = 0.790) between ARFI values and hepatic fibrosis; and even higher AUROC figures of 88.8%, 93.1 and 95.0% for F ≥ 2, F ≥ 3 and F = 4 fibrosis, respectively. With regard to a 40% maximum SD, only 18 (10.1%) patients had to be excluded, resulting in AUROC figures of 85.6, 89.2 and 90.6%, respectively. Including only patients with at least 11 portal tracts (n = 93), the correlation between hepatic ARFI and fibrosis was r = 0.674 and p < 0.001 and the AUROC for F ≥ 2, F ≥ 3 and F = 4 fibrosis were 79.4, 85.6 and 89.7%, respectively.

### Table 2. Diagnostic accuracy and optimal cut-offs for the relevant fibrosis stages.

<table>
<thead>
<tr>
<th>ARFI</th>
<th>AUROC (%) [95% confidence interval]</th>
<th>Cut-off (m/s)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F ≥ 2</td>
<td>80.7 [73.8-87.7]</td>
<td>1.25</td>
<td>89.8</td>
<td>61.7</td>
<td>82.2</td>
<td>75.5</td>
</tr>
<tr>
<td>F ≥ 3</td>
<td>86.2 [80.8-91.6]</td>
<td>1.72</td>
<td>82.5</td>
<td>79.6</td>
<td>76.7</td>
<td>84.8</td>
</tr>
<tr>
<td>F = 4</td>
<td>88.7 [83.9-93.5]</td>
<td>1.75</td>
<td>94.5</td>
<td>73.2</td>
<td>61.2</td>
<td>96.8</td>
</tr>
</tbody>
</table>

Only patients with a standard deviation ≤ 30% of the individual mean ARFI (subgroup A)

| F ≥ 2 | 86.1 [79.8-92.4] | 1.48 | 72.5 | 88.9 | 93.7 | 58.8 |
| F ≥ 3 | 91.2 [86.6-95.8] | 1.72 | 84.1 | 88.5 | 86.8 | 86.3 |
| F = 4 | 91.5 [87.0-95.9] | 1.76 | 94.0 | 80.4 | 71.2 | 96.3 |

### Table 3. Hepatic ARFI elastometry and individual standard deviation (SD) (m/s) for different degrees of hepatic fibrosis in all patients and the two subgroups A + B.

<table>
<thead>
<tr>
<th>All</th>
<th>Subgroup A (SD ≤ 30%)</th>
<th>Subgroup B (SD &gt;30%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 178</td>
<td>n = 147</td>
<td>n = 31</td>
</tr>
<tr>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>F0</td>
<td>1.26 ± 0.47</td>
<td>0.19</td>
</tr>
<tr>
<td>F1</td>
<td>1.47 ± 0.66</td>
<td>0.40</td>
</tr>
<tr>
<td>F2</td>
<td>1.59 ± 0.55</td>
<td>0.28</td>
</tr>
<tr>
<td>F3</td>
<td>1.99 ± 0.76</td>
<td>0.39</td>
</tr>
<tr>
<td>F4</td>
<td>2.76 ± 0.73</td>
<td>0.50</td>
</tr>
</tbody>
</table>
DISCUSSION

In CLD, prognosis and treatment are related to the stage of hepatic fibrosis. To evaluate the degree of fibrosis, liver biopsy is regarded as the gold standard, providing additional information on inflammation, steatosis and aetiology. However, its use is limited by its invasiveness, complications, inter- and intra-observer variability and patient acceptability. For the non-invasive assessment of hepatic fibrosis, liver elasticity evaluation by transient elastography or ARFI elastometry has been, and is increasingly being, used in patients with CLD. Studies mostly including patients with chronic viral hepatitis have revealed good correlation between histological degree of advanced liver fibrosis and cirrhosis, and between the two different methods (ARFI elastometry and transient elastography). Recently, liver elastometry has been included in the German guidelines for diagnosis and treatment of chronic viral hepatitis B or C, as a potential non-invasive test for excluding cirrhosis. To date, information on the diagnostic value of ARFI in the assessment of liver fibrosis in non-viral liver disease is still limited.

As in the case of transient elastography, the investigation of ARFI needs quality criteria for the measurements, and a good gold standard. Although histology is limited by potential sampling errors, it is accepted as the reference, and should include at least 6 portal tracts and a specimen length of 15 mm.

Recent AALSD guidelines recommended at least 11 portal tracts and a specimen length of 20 mm (or better 25 mm). These quality parameters, however, were regularly not achieved and would demand more needle passages with the higher risk of complications. In our data series, the quality of hepatic specimens were equivalent or even superior to present ARFI studies and the results of patients with at least 11 portal tracts were comparable to the overall group.

To obtain robust data concerning the classification of hepatic fibrosis, a meta-analysis defining cut-off values of ARFI elastometry has recently been published. Interestingly, only nine out of 349 ARFI studies met the inclusion criteria, and in the selected studies important information defining histological characteristics had to meet different requirements (minimum length of biopsy ≥ 18 mm, ≥ 15 mm or ≥ 10 mm) or was lacking. In this meta-analysis the AUROC of the different degree of fibrosis presented (87% for F ≥ 2; 91% for F ≥ 3, 93% for F = 4) were equal to those of our subgroup with a 30% maximum SD, whereas the AUROC of all patients in our study were marginally inferior. The ARFI elastometry of the 518 meta-analysis patients reveals higher specificities as compared with those of all our own patients. These differences might be associated with the considerably higher proportion of viral hepatitis in the meta-analysis (83.2 vs. 38.8% in our data). For ARFI elastometry has been evaluated mainly in viral hepatitis and much less in other aetiologies. Nevertheless, our data of hepatic ARFI elastometry show competitive results of its potential in a cohort with various CLD.

The interquartile range and SD are measures of statistical dispersion: the interquartile range—excluding 50% of the measurements, excluding extreme values—is the difference between the upper and lower quartiles. The SD shows how much variation there is from the mean (i.e. mean ± SD includes 68.27% of the measurements) of the entire range of values. Such technical quality parameters as an interquartile range interval < 30% and a success rate of ≥ 60% were once proposed for ARFI elastometry, as is the case for transient elastography. In the study of Rifai, et al., only 10% of 121 ARFI measurements of the right lobe were considered to be unrepresentative, since these technical parameters were not met. In the remaining study group (cirrhotic and non-cirrhotic patients together with healthy controls; aetiologically 50% chronic viral hepatitis) ARFI values correlated significantly with histological degree of hepatic fibrosis (r = 0.54, p < 0.001). The analysis of Bota, et al. showed a very strong ARFI correlation with fibrosis (r = 0.722, p < 0.001) —after exclusion of 45 (9.7%) of the baseline 460 patients and healthy controls—on applying the technical quality parameters (an interquartile range interval < 30% and a success rate of ≥ 60%). In that study, the excluded patients had a significantly higher BMI, and showed no correlation between ARFI and histological fibrosis. In a following report, a discordance of at least two stages of fibrosis between the ARFI results and liver histology was often associated with an interquartile range ≥ 30% in addition to female sex. Popescu, et al. employed ARFI elastometry to investigate only subjects with no known liver pathology. Most of the patients excluded on account of the technical quality parameters (7.4%) had incomprehensibly high ARFI values. However, in none of these four studies was an analysis performed after varying the technical quality parameters.

Our data underscore the importance of establishing a quality parameter. In contrast to the technical
quality parameters (interquartile range interval < 30%, success rate of ≥ 60% mentioned, we analysed ARFI performance applying an SD of maximum 30% of the mean ARFI value as a simple, post-examination calculable parameter and since the SD is based on the entire range of shear wave velocity measurements. Including only patients with an SD maximum of 30% led to higher accuracy parameters of ARFI elastometry for the classification of hepatic fibrosis. Remarkably, 13 (41.9%) of 31 excluded patients (subgroup B) showed mild portal fibrosis F1. Apparently, the initially scattered nature of the minimal portal fibrosis led to scattering of the ARFI measurements. In these excluded patients, ARFI elastometry showed high shear wave velocities in F1/2, which would lead to a relevant overestimation of the degree of fibrosis (Table 3). High SD might reflect the sometimes inhomogeneous morphology in liver fibrosis. Subgroup B showed no difference in BMI, and failed to show statistically significant lower levels of AST as compared to the patients in subgroup A. Low levels of transaminases are considered to be associated with lower ARFI shear wave velocities.

The ARFI values and SDs of healthy livers are generally lower, namely 1.16 ± 0.14 m/s in children and 1.19 ± 0.18 m/s in older volunteers, resulting in an SD clearly below 20% of the mean. In the investigation of clinically healthy subjects, only 1 (1.2%) out of 82 had to be excluded because quality criteria were not met. In our study the individual SD correlated with the degree of fibrosis and the individual mean ARFI values.

A certain weakness of our study is the fact that the study participants could not be included consecutively over the whole of the study period, which might have led to a certain selection bias. Another possible shortcoming is that an analysis of the impact of interquartile range and success rate on the accuracy of the ARFI measurement was not done—since it was not initially intended—although it has since been shown to be important. In terms of the aetiologies of CLD and the frequencies of different stages of fibrosis, our patient group is in fact rather heterogeneous.

On the basis of our findings we are of the opinion that the level of SD should be taken into account when interpreting hepatic ARFI elastometry of CLD patients. ARFI with high SD values might be falsely interpreted as pronounced fibrosis (overstaging). But a major strength of non-invasive assessment by ARFI remains untouched: low shear wave velocities permit reliable exclusion of hepatic cirrhosis. Individual mean ARFI with high SD should be regarded to be of doubtful value. Including only patients with a maximum of 30% SD leads to higher AUROC and higher correlation coefficients of hepatic ARFI elastometry for staging fibrosis.

In conclusion, ARFI elastometry correlates well with the degree of hepatic fibrosis. The application of a 30% maximum of SD represents a quality parameter for ARFI elastometric measurements. Thus, in patients with an SD higher than 30% of the mean, ARFI elastometry is often unreliable and should be interpreted with caution.

**ABBREVIATIONS**

- **AIH**: autoimmune hepatitis.
- **APRI**: AST/platelet ratio index.
- **ARFI**: acoustic radiation force impulse.
- **AST**: aspartate transaminase.
- **AUROC**: area under the receiver operating characteristic.
- **BMI**: body mass index.
- **CLD**: chronic liver disease.
- **PBC**: primary biliary cirrhosis.
- **PSC**: primary sclerosing cholangitis.
- **SD**: standard deviation.

**ACKNOWLEDGMENTS/GRANT**

Thanks are due to the ELAN (Erlanger Leistungsbezogene Anschubfinanzierung und Nachwuchsförderung) foundation for the financial support of the project.

No other financial support.

**REFERENCES**


