Original Article

Skin-Liver Distance and Interquartile Range-Median Ratio as Determinants of Interoperator Concordance in Acoustic Radiation Force Impulse Imaging

Shu Su[1]*, Wayland Wang[1], David Nadebaum[2], Amanda Nicoll[2], Siddharth Sood[2], Alexandra Gorelik[3,4], Jeffrey Lai[1], Robert Gibson[1,4]

[1]Department of Radiology, Royal Melbourne Hospital, Victoria, Australia, [2]Department of Gastroenterology, Royal Melbourne Hospital, Victoria, Australia, [3]Melbourne Epicentre, Royal Melbourne Hospital, Victoria, Australia, [4]Department of Radiology, University of Melbourne, Victoria, Australia

Abstract

Context and Aims: The accuracy of acoustic radiation force impulse (ARFI) ultrasound compared to liver biopsy is higher when there is concordance between F-scores of two or more operators. We hypothesized that when the first operator interquartile range/median-velocity ratio (IMR) is <0.3 and skin-liver distance (SLD) is <2.5 cm, there is greater interoperator concordance and a second operator is not necessary. **Subjects and Methods:** Two-operator ARFI ultrasound measurements (F-score, SLD, and IMR) were recorded for 927 consecutive patients. Chi-squared testing compared interoperator concordance for SLD <2.5 cm versus SLD ≥2.5 cm and IMR <0.3 versus IMR ≥0.3 when SLD <2.5 cm, in each of the F-score groups of 0/1, 2, 3, and 4. **Results:** Statistically significant differences were demonstrated between SLD <2.5 cm and SLD ≥2.5 cm groups for F-scores 0/1 or 4 ($P = 0.005$) and F-scores 2 or 3 ($P < 0.001$). Concordance, when SLD measured <2.5 cm, was more than 85% for all F-score groups. In the SLD <2.5 cm group, concordance fell below 85% when IMR ≥0.3, for all F-scores except F2. Specifically, P values comparing IMR <0.3 and IMR ≥0.3 in the various first operator F-score groups were $P = 0.040$ for F0/F1, $P = 0.580$ for F2, $P = 0.342$ for F3, and $P < 0.001$ for F4. **Conclusions:** ARFI measurements from one operator can be considered acceptable when SLD <2.5 cm and IMR <0.3. Otherwise, adding a second operator can improve confidence in the result.

Keywords: Elasticity imaging techniques, fibrosis, liver, obesity, ultrasonography

Introduction

Staging of liver fibrosis is important for the prognosis and management of chronic liver disease.[1] Historically, this has relied on the gold standard of histological staging by liver biopsy, which carries a small but significant morbidity.[2]

Acoustic radiation force impulse (ARFI) imaging is a form of ultrasound elastography, which allows for a noninvasive method of assessing liver fibrosis. It is provided by conventional B-mode ultrasonography (Acuson S2000; SIEMENS Medical Solutions) and involves the transmission of ultrasonic pulses to measure a quantitative shear-wave velocity (m/s). The stiffer a region of liver, the greater the shear-wave speed.[3] Stiffness in turn corresponds to liver fibrosis, although reliability of elastography can be variably affected by obesity, ascites, necroinflammation, narrow intercostal spaces, and operator inexperience.[4-6]

Received: 16-12-2018 **Revised:** 23-01-2019 **Accepted:** 30-01-2019 **Available Online:** 17-05-2019

Access this article online

Quick Response Code:

Website: www.jmuonline.org

DOI: 10.4103/JMU.JMU_124_18

Liver stiffness assessed by shear-wave velocity has been found to significantly correlate with liver fibrosis in chronic liver diseases. Friedrich-Rust *et al*. reported that the area under the receiver operating characteristic curve for the accuracy of ARFI elastography was 0.87, 0.91, and 0.93 for the diagnosis of moderate fibrosis ($F \geq 2$), severe fibrosis ($F \geq 3$), and cirrhosis, respectively. Their cutoff value for liver cirrhosis was 1.80 m/s.[7] Other reports have identified variable cutoff values for cirrhosis, ranging from 1.6 to 1.95 m/s, although this can be explained by population heterogeneity.[8-10]

Our study focuses on the factors which significantly affect the accuracy of ARFI. In general, accuracy when compared

Address for correspondence: Dr. Shu Su, Department of Radiology, Royal Melbourne Hospital, 300 Grattan Street, Parkville, Victoria, 3050. Australia. E-mail: shusuemail@gmail.com

How to cite this article: Su S, Wang W, Nadebaum D, Nicoll A, Sood S, Gorelik A, *et al*. Skin-liver distance and interquartile range-median ratio as determinants of interoperator concordance in acoustic radiation force impulse imaging. J Med Ultrasound 2019;27:177-80.

to liver biopsy is higher when the interquartile range-median velocity ratio (IMR) is <0.3, and the skin-liver distance (SLD) is <2.5 cm.[11-13] It is a routine protocol at our institution to use at least two operators as Nadebaum *et al*. found that accuracy when compared to liver biopsy improves when there is interoperator concordance between two or more operators.[11] A previous study also found that variability in 2D-shear-wave elastography measurements was higher if the patient's SLD was higher and that this correlated with a higher IMR.[14] In addition, interoperator concordance is higher when fibrosis grade (F-score) is 0/1 or 4.[15] Hence, we hypothesize that when the first operator ARFI measurement has an F-score of 0/1 or 4, average SLD <2.5 cm, or IMR <0.3, a second operator measurement is not necessary given the high likelihood of interoperator concordance.

Subjects and Methods

The study was approved by the Melbourne Health Human Research Ethics Committee (approval number QA2015154). ARFI point shear-wave elastography velocity measurements using the Siemens Acuson S2000 ultrasound machine were recorded for consecutive patients. The 4C1 convex transducer was used, and measurements were obtained from the right hepatic lobe through an intercostal approach. The region of liver measured was selected to be away from portal tracts and hepatic veins. Patients were instructed to suspend their respiration during the acquisition of the shear-wave velocities. The region of interest for the measurement of shear-wave velocity was selected to be approximately 4–5-cm deep to the skin surface. Size of the region of interest is set by the manufacturers and unchanged for all measurements. As is our routine, two sets of 10 valid ARFI measurements were obtained by two operators, with each operator blinded to the previous operator's measurements. Operators were sonographers who had been trained in the use of ARFI measurements using the Siemens Acuson S2000 system.

Liver fibrosis was staged using shear-wave velocity cutoffs, adapted from Friedrich-Rust *et al*.:[7]

- <1.35 – absent or mild fibrosis (F0 or F1)
- 1.35–1.55 – significant fibrosis (F2)
- 1.55–1.80 – severe fibrosis (F3)
- >1.80 – cirrhosis (F4).

If the median velocities of the two operators are within the same or adjacent fibrosis stage, the result is accepted as being adequate for clinical management decisions. If not, a third operator is used in an attempt to determine the more reliable result. All measurements are obtained from the right lobe of the liver.

The F-score, SLD, and IMR of the first operator were recorded. The F-score of the second operator was then recorded, and concordance of the results from these two operators was calculated. Interoperator concordance was defined as F-scores in the same or adjacent grades. For the purposes of this study, the third operator F-score was not required.

Data analysis

Chi-squared or Fisher's exact tests were performed comparing interoperator concordance in the following groups: SLD <2.5 cm versus SLD ≥2.5 cm when the first operator F-score was 0/1 or 4; SLD <2.5 cm versus SLD ≥2.5 cm when the first operator F-score was 2 or 3; and IMR <0.3 versus IMR ≥0.3 when SLD <2.5 cm, in each of the F-score groups of 0/1, 2, 3, and 4. Given the multiple comparisons made, differences were considered to be statistically significant when $P < 0.006$, after applying the Bonferroni correction. All analyses were performed using SPSS Statistics (version 17.0, Polar Engineering and Consulting).

Results

Nine hundred and twenty-seven consecutive patients from August 2008 to December 2014 were included, with an average age of 50.8 (σ =14.6) years. Four hundred and fifty-four were female and 473 were male. One hundred and eighty cases had clinically diagnosed cirrhosis, 541 did not, and 206 were unknown. There also were multiple potential contributory etiologies for chronic liver disease. These were hepatitis B (240), nonalcoholic fatty liver disease (251), hepatitis C (201), alcohol (108), drug-induced liver injury (47), autoimmune hepatitis (42), hemochromatosis (32), primary biliary cirrhosis (33), cryptogenic (22), primary sclerosing cholangitis (20), cardiac failure (17), Wilson's disease (5), porphyria (10), and other (28). Note that, the sum of these is more than the total number of patients because some patients had multiple etiologies for chronic liver disease.

The spread over the fibrosis groups was as follows: $n = 457$ for F0/F1, $n = 122$ for F2, $n = 74$ for F3, and $n = 274$ for F4. The average SLD was 2.17 cm (σ =0.58) and average IMR was 0.23 (σ =0.19). Interoperator concordance was present in 808 of these patients, of which 370 patients were in F0/F1, 46 in F2, 19 in F3, and 207 in F4.

Statistically significant differences (i.e., $P < 0.006$ after applying the Bonferroni correction) were demonstrated between SLD <2.5 cm and SLD ≥2.5 cm groups when F-score was 0/1 or 4 ($P = 0.005$), and when F-score was 2 or 3 ($P < 0.001$). Interoperator concordance, when SLD <2.5 cm, was more than 85% for all F-score groups, in contrast to the SLD ≥2.5 cm group [Table 1].

In the SLD <2.5 cm group, interoperator concordance fell below 85% when IMR was ≥0.3, for all F-scores except F2. *P* values comparing IMR <0.3 and IMR ≥0.3 were <0.05 for F-scores of F4. There was no significant difference between IMR <0.3 and IMR ≥0.3 groups when F score was 0/1, 2, or 3 [Table 2].

Discussion

In clinical practice, it can be difficult to be confident of the fibrosis stage determined by ARFI if either SLD or IMR is high. To combat this at our institution, it is routine to use two

Table 1: Chi-squared tests for interoperator concordance comparing fibrosis grade and skin-liver distance

F-score	Concordance		*P*
	SLD <2.5 cm (%)	SLD ≥2.5 cm (%)	
F0/F1 OR F4	496/551 (90.0)	148/180 (82.2)	0.005
F2 OR F3	149/167 (89.2)	15/29 (51.7)	<0.001

Table 2: Chi-squared tests for interoperator concordance comparing fibrosis grade and interquartile range-median ratio when skin-liver distance <2.5 cm

F-score	Concordance		*P*
	IMR <0.3 (%)	IMR ≥0.3 (%)	
F0/F1	332/353 (94.1)	33/39 (84.6)	0.040*,†
F2	95/103 (92.2)	9/10 (90.0)	0.580*
F3	38/44 (86.4)	7/10 (70.0)	0.342*
F4	97/108 (89.8)	34/51 (66.7)	<0.001

*Fisher's exact test was used due to low Chi-squared test power, †*P*=0.006 is considered statistically significant. IMR: Interquartile range-median velocity ratio

operators to improve our confidence; and potentially, a third operator if the first two operators are discordant. However, applying this for every patient is time- and labor-intensive, and this study is helpful in identifying those patients who may not need a second ARFI operator.

Skin-liver distance

Our study has demonstrated that interoperator concordance is significantly higher when SLD <2.5 cm. Previous studies have demonstrated that higher interoperator concordance gives better correlation of ARFI fibrosis grade to liver biopsy grade.[11] Thus, by this association, SLD <2.5 cm is linked to improved accuracy of ARFI readings.

The Chi-squared tests comparing higher SLDs were not performed as correlation with biopsy reduces when SLD ≥2.5 cm.[11] Note that, we separated the F-scores in this manner because previous studies had shown that F0/F1 and F4 groups had higher interoperator concordance compared to F2 and F3 groups.[15] However, we found that interoperator concordance was also higher when SLD ≥2.5 cm in the F2 and F3 groups, and thus separating the patients who do not require a second operator based on first operator fibrosis grade is not necessary.

While it is conceivable that signals from deeper regions of interest may be weaker due to beam attenuation, previous studies have demonstrated that SLD is a better independent determinant of ARFI reliability. Total ARFI depth had shown positive correlations with IMR, percentage deviation between operators, and the deviation of ARFI velocities from the biopsy F-score reference range. In multiple regression analyses, however, these associations were not independent of SLD. Thus, the effects seen are likely due to SLD rather than the traversing liver.[16] This stronger association between SLD and ARFI reliability provides further justification to its adoption as a reliability metric in preference to total measurement depth.

Interquartile range-median ratio

When comparing IMR in the SLD <2.5 cm group, significant differences in interoperator concordance between the IMR <0.3 and IMR ≥0.3 groups in the F4 stage of fibrosis were observed. This was expected since we know that IMR <0.3 correlates with improved accuracy of ARFI compared to liver biopsy. High interoperator concordance is also associated with improved ARFI accuracy compared to liver biopsy.[11,12] This supports the use of a single ARFI operator when IMR <0.3 and SLD <2.5 cm for liver fibrosis stage F4.

The lack of a significant difference in the F0/F1 group was due to the strict $P < 0.006$ we selected to account for the multiple comparisons that were made. The fact that there was no difference between the IMR <0.3 and IMR ≥0.3 groups in the F2 and F3 stages of fibrosis was likely due to the high interoperator concordance in the IMR ≥0.3 groups in these stages, and thus, we would expect the difference in interoperator concordance to be more apparent had the cutoff we used to distinguish IMR groups been higher. Thus, it would be reasonable to accept a single ARFI operator reading for these groups as well, given the high absolute interoperator concordance.

Strengths and limitations

A major strength of this study is that a large consecutive cohort of patient's data was able to be collected. All of the ARFI readings were also performed on the same ultrasound machine, with the same standard technique to allow for very close replication of the method of data collection.

However, the study was limited by the relatively smaller number of patients in the F2 and F3 groups, which necessitated the use of Fisher's exact tests instead of Chi-squared tests due to the reduced power. Another limitation was that it was not possible to provide biopsy correlation for every patient. This does not detract from the value of the study though, as the aim is to establish identifiable features that allow more confidence in the result of the ARFI, in particular, when the addition of a second operator can add diagnostic confidence in the result. All means of noninvasive assessment of liver fibrosis suffer from accuracy limitations; so, any factors that can improve the reliability of an approach, in this case ARFI, is worth establishing.

Clinical decision flow chart

To help maintain accuracy while reducing demands on workforce, we propose a clinical decision flow chart based on our results, whereby if the first operator ARFI reading is obtained with a SLD <2.5 cm and IMR <0.3, the result can be considered acceptable because the interoperator concordance is likely to be high. If either one of these criteria is not met, a second operator ARFI reading should be sought. A third operator reading may also be required if the first and second operators do not concur [Figure 1]. If we apply the flow chart to our study

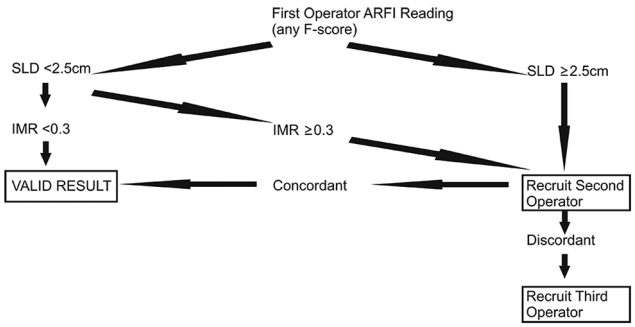

Figure 1: Clinical decision flow chart for potential acoustic radiation force impulse reading interoperator concordance based on skin-liver distance and interquartile range-median ratio

cohort, 209 would initially need a second operator given their first operator SLD ≥2.5 cm. Then, a further 110 patients with SLD <2.5 cm but IMR >0.3 would also require a second operator. Thus, instead of 927 patients requiring two-operator ARFI measurements, only 319 patients would need two operators.

Using this algorithm, resources and time can be saved in cases where a second operator ARFI measurement is not required, and confidence in the ARFI results can be improved in cases where a second (or third) operator is recommended.

Acknowledgments

All persons who have made substantial contributions to the work reported in the manuscript (e.g., technical help, writing and editing assistance, general support), but who do not meet the criteria for authorship, are named in the Acknowledgments and have given us their written permission to be named. If we have not included an Acknowledgments in our manuscript, then that indicates that we have not received substantial contributions from non-authors.

Financial support and sponsorship

Nil.

Conflicts of interest

Amanda Nicoll – Bayer speaker, MSD research support, BMS speaker.

References

1. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology 1996;24:289-93.
2. Piccinino F, Sagnelli E, Pasquale G, Giusti G. Complications following percutaneous liver biopsy. A multicentre retrospective study on 68,276 biopsies. J Hepatol 1986;2:165-73.
3. Jaffer OS, Lung PF, Bosanac D, Patel VM, Ryan SM, Heneghan MA, *et al.* Acoustic radiation force impulse quantification: Repeatability of measurements in selected liver segments and influence of age, body mass index and liver capsule-to-box distance. Br J Radiol 2012;85:e858-63.
4. Cassinotto C, Lapuyade B, Aït-Ali A, Vergniol J, Gaye D, Foucher J, *et al.* Liver fibrosis: Noninvasive assessment with acoustic radiation force impulse elastography – Comparison with FibroScan M and XL probes and FibroTest in patients with chronic liver disease. Radiology 2013;269:283-92.
5. Castéra L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, *et al.* Pitfalls of liver stiffness measurement: A 5-year prospective study of 13,369 examinations. Hepatology 2010;51:828-35.
6. Yoon KT, Lim SM, Park JY, Kim DY, Ahn SH, Han KH, *et al.* Liver stiffness measurement using acoustic radiation force impulse (ARFI) elastography and effect of necroinflammation. Dig Dis Sci 2012;57:1682-91.
7. Friedrich-Rust M, Wunder K, Kriener S, Sotoudeh F, Richter S, Bojunga J, *et al.* Liver fibrosis in viral hepatitis: Noninvasive assessment with acoustic radiation force impulse imaging versus transient elastography. Radiology 2009;252:595-604.
8. Sporea I, Sirli R, Popescu A, Danilă M. Acoustic radiation force impulse (ARFI) – A new modality for the evaluation of liver fibrosis. Med Ultrason 2010;12:26-31.
9. Takahashi H, Ono N, Eguchi Y, Eguchi T, Kitajima Y, Kawaguchi Y, *et al.* Evaluation of acoustic radiation force impulse elastography for fibrosis staging of chronic liver disease: A pilot study. Liver Int 2010;30:538-45.
10. Yoneda M, Suzuki K, Kato S, Fujita K, Nozaki Y, Hosono K, *et al.* Nonalcoholic fatty liver disease: US-based acoustic radiation force impulse elastography. Radiology 2010;256:640-7.
11. Nadebaum D, Gibson R, Howell J, Halliday J, Christie M, Gorelik A, *et al.* Inter-operator concordance in acoustic radiation force impulse imaging (arfi); a practical role for multiple operators to improve diagnostic performance. J Gastroenterol Hepatol Hepatol Clin Abstr 2014;29:68-101.
12. Bota S, Sporea I, Sirli R, Popescu A, Dănilă M, Sendroiu M, *et al.* Factors that influence the correlation of acoustic radiation force impulse (ARFI), elastography with liver fibrosis. Med Ultrason 2011;13:135-40.
13. Frulio N, Trillaud H, Perez P, Asselineau J, Vandenhende M, Hessamfar M, *et al.* Acoustic radiation force impulse (ARFI) and transient elastography (TE) for evaluation of liver fibrosis in HIV-HCV co-infected patients. BMC Infect Dis 2014;14:405.
14. Nadebaum DP, Nicoll AJ, Sood S, Gorelik A, Gibson RN. Variability of liver shear wave measurements using a new ultrasound elastographic technique. J Ultrasound Med 2018;37:647-56.
15. Wang W, Nadebaum D, Gorelik A, Nicoll A, Sood S, Gibson R. Acoustic Radiation Force Impulse (ARFI) Imaging – Can we Predict When a Second Operator is Needed? Perth: Abdominal Radiology Group of Australia and New Zealand; 2016.
16. Nadebaum D. Shear Wave Elastography in the Assessment of Liver Fibrosis. Masters Research Thesis. The University of Melbourne; 2016.