A comparison of Transient Elastography (TE) with Acoustic Radiation Force Impulse Elastography (ARFI) for the assessment of liver health in patients with Chronic Hepatitis C; baseline results from the TRACER study.

Abstract:

Background

Liver stiffness measurements (LSMs) can be used to assess liver fibrosis and can be acquired by transient elastography using FibroScan® and with Acoustic Radiation Force Impulse (ARFI) imaging. The study aimed to establish LSM scores using FibroScan® and ARFI in a chronic hepatitis C (CHC) cohort and to explore the correlation and agreement between the scores and the factors influencing agreement.

Methods

Patients had LSMs acquired with FibroScan® (right lobe of liver) and ARFI (right and left lobe of liver). We used Spearman's correlation to explore the relationship between FibroScan® and ARFI scores. A Bland-Altman plot was used to evaluate bias between the mean percentage differences of FibroScan® and ARFI scores. Univariable and multivariable analyses were used to assess how factors such as BMI, age and gender influenced the agreement between LSMs.

Results

Bland-Altman showed the average (95% CI) percentage difference between FibroScan® and ARFI scores was 27.5% (17.8, 37.2), P<0.001. There was a negative correlation between the average and percentage difference of the FibroScan® and ARFI scores [r(95% CI) = -0.41(-0.57, -0.21), P<0.001], thus showing that percentage difference gets smaller for greater FibroScan® and ARFI scores. BMI was the biggest influencing factor on differences between FibroScan® and ARFI (r=0.12[0.01, 0.23], P=0.05). ARFI scores at segment 5/8 and the left lobe showed good correlation [r(95% CI)=0.83(0.75, 0.89), P<0.001).

Conclusion

FibroScan® and ARFI had similar predictive values for the assessment of liver stiffness in patients with CHC infection however the level of agreement varied across lower and higher scores.

Introduction

In patients with chronic hepatitis C (CHC), the prognosis, management and follow-up of liver disease is closely linked to the degree of liver fibrosis[1]. Liver biopsy is considered as the "gold-standard" method when diagnosing liver tissue damage and histological assessment is based on semi-quantitative scoring systems (eg METAVIR)[2]. However liver biopsy is invasive and can only examine about 1/50,000th of the liver at one time with a 0.01-0.1% risk of mortality[3]. The emergence of non-invasive tests to evaluate liver health have resulted in a sharp decline in the need for liver biopsies[4, 5]. Research shows that by determining liver stiffness measurements (LSM) at the right lobe of liver by noninvasive shear wave elastography methods such as transient elastography (TE), point shear wave elastography (pSWE) and multidimensional SWE (2D/3D-SWE), LSM may be used by physicians as a prognostic marker among patients with CHC[6] . In 2014, De Robertis et al found that transient elastography (TE) and acoustic radiation force impulse imaging (ARFI) were the most widely used noninvasive methods for staging liver fibrosis[7]. TE performed using a FibroScan® (FS) machine is the most widely used method of liver elastography[8, 9] but liver visibility is limited while acquiring the measurement, so De Robertis et al[7] state that as ARFI is incorporated into standard ultrasound units, it could provide a higher reproducibility and successful measurement rates than TE. ARFI can also be used to determine liver stiffness at the left lobe of liver which may be useful for patient management; it has been proven that liver biopsy fragments obtained in the same session from the left and right lobes may have different stages of fibrosis in almost half of the cases[10].

This study aimed to establish FibroScan® and ARFI scores in a CHC cohort before commencing direct acting antiviral therapy for the treatment of CHC. FibroScan® and ARFI scores were acquired at the right lobe of the liver and further ARFI measurements were acquired at the left lobe of liver. We aimed to explore the relationship and agreement between FibroScan® and ARFI scores and determine factors influencing agreement.

Methods

The Chronic Hepatitis C Treatment Radiographic and Clinical Outcomes (TRACER) Study was a prospective, observational cohort study which recruited over a 32-month period from March 2017 to November 2019. The study took place at two clinical sites, the Mater Misericordiae University Hospital and St. Vincent's University Hospital, both of which are large tertiary referral hospitals located in Dublin, Ireland. The study aimed to investigate longitudinal liver stiffness measurements and clinical outcomes associated with treatment of CHC infection with directly acting antiviral agents. Inclusion criteria were eligibility for direct acting antiviral (DAA) therapy on basis of non-invasive biomarkers, age >18 years, ability to give informed consent and exclusion criteria were subjects incarcerated at time of treatment initiation. The study required six visits over a two-year period and this paper describes findings from the baseline visit. The study was approved by the hospitals' Institutional Review Board, in accordance with the Helsinki Declaration of 1975 and patients gave their informed, written consent.

LSM were acquired through FibroScan® and ARFI and patients were asked to present fasting. An M probe was used for all FibroScan® readings as there was no XL probe available. Additional data collected included sex, race, body mass index (BMI) and hepatitis C genotype. The study coordinator was a sonographer with over ten years' experience imaging the liver and had received training in acquiring stiffness measurements with both ARFI and TE and the study coordinator acquired all measurements on all patients. During TE acquisition, measurements were acquired with the patient supine, the right arm raised over the head and the transducer placed perpendicularly between the ribs adjacent to the right lobe of the liver in the middle of the liver parenchyma, away from the liver border. The median value of 10 successful acquisitions, expressed in kiloPascals (kPa), was recorded as representative of LSM. As per manufacturer guidelines, we considered 10 successful acquisitions with a success rate of at least 60% (the ratio of the number of successful measurements to the total number of acquisitions) and with an interquartile range (IQR) lower than 30% of the median value as accurate measurements.

Conventional B-mode ultrasound and ARFI elastography were performed using a Siemens Acuson S2000 HELX with a 6C1HD transducer which had ARFI capability. The patients were examined in supine position with the right arm extended above their head and ARFI elastography was performed through the intercostal spaces, 2-3 cm below and perpendicular to the liver capsule at segments 5/8 of the right lobe of the liver, avoiding large vessels and intrahepatic ducts, to ensure reliable and consistent measurements[11, 12]. ARFI measurements were also acquired at the left lobe at a depth of approximately 1.5-2cm from the liver surface with a region of interest (ROI) angle as close to 0° as possible, with the patient supine. The sonographer also noted the skin to liver capsular distance (SCD) at both the left and right lobes as high BMI or a thick sono-opaque adipose layer can attenuate shear waves and result in unmeasurable values (shown as x.xx ms^{-1}) or values with high data set variability[13]. ARFI measurements were acquired 10 times at both sites for each patient and the median value was recorded, with measurements considered complete when 6-10 measurements were achieved at each site without unmeasurable values. On completion of the ultrasound examination the machine generated a report which tabulated the shear velocity (Vs) in ms^{-1} and depth measurements(cm) associated with each site. The report calculated mean, median, standard deviation, and interquartile range (IQR) for all measurements. Siemens suggest using a median value of 10 measurements with an interquartile range/median(IQR/MED) ≤0.3 to ensure a reliable liver assessment[12]. However this can lead to an exclusion of a relevant number of ARFI scores[14] and hence those with an IQR >0.3 were included in the study.

Interpretating stiffness scores

The METAVIR scoring system incorporates five stages of liver fibrosis (F0: no scarring (no fibrosis); F1: minimal scarring; F2: scarring has occurred and extends outside the liver area (significant fibrosis); F3: fibrosis spreading and forming bridges with other fibrotic liver areas (severe fibrosis); F4: cirrhosis or advanced scarring) and is widely used to semi-quantitatively assess the histological of chronic hepatitis C[2]. Echosens

(FibroScan® manufacturers) provide a scoring card guide, which was used to group results measured in kPa into the hepatitis C specific scores, which have been shown to correspond to the METAVIR scale of F01-F4, with cut off points of F0/F1= <8.5kPa, F2= 8.5-12.5kPa, F3= >12.5-14.5kPa and F4= >14.5kPa[15].

Siemens suggest the following figures for interpreting ARFI scores; 0-1.34ms^{-1} as normal, 1.34-2.2ms^{-1} as moderate stiffness and >2.2ms^{-1} as high stiffness however we wished to stratify results into fibrosis scores similar to those provided by METAVIR scoring, so required more specific figures. A meta-analysis study published in 2013 looked at the efficiency of ARFI imaging for the staging of liver fibrosis and found that ARFI provided good diagnostic performance for assessing significant/severe hepatic fibrosis and showed excellent diagnostic accuracy cirrhosis staging[16]. Nierhof et al (2013) determined the following interpretations from their meta-analysis; No liver scarring or mild liver scarring = < 1.35ms^{-1}, significant fibrosis = 1.35-1.6ms^{-1}, severe fibrosis = 1.61 - 1.87ms^{-1} and liver cirrhosis = >1.87ms^{-1}. These figures, unlike FibroScan® scores, are not disease specific. The same figures were used for interpreting stiffness measurements at the right and left lobes of liver.

FibroScan® estimates liver stiffness by calculating Young's modulus, where $Y=3\rho c^2$ where Y=Young's Modulus (kPa), ρ=mass density of liver tissue and c=Shear Wave Velocity(ms^{-1}). The FibroScan® machine automatically performs this calculation and expresses the result in kPa[17]. For both TE and ARFI, the density of the liver tissue is assumed to be a constant of approximately 1000 kg/m^3 [17]. To explore how TE scores compared to ARFI scores, ARFI scores were converted to kPa using Young's modulus equation. We did not group the converted ARFI scores into equivalent METAVIR scores in order to minimize error. The frequency at which the FibroScan and ARFI scores were acquired were predetermined by the respective machines, with FibroScan® measurements acquired at a central frequency of 3.5MHz[17] and ARFI measurements were acquired at a fixed transmit frequency of 2.67 MHz[18].

Blood results

Aminotransferase-platelet ratio index (APRI) and Fibrosis-4 Score (FIB-4) scores have been validated for chronic hepatitis C and show satisfactory sensitivity and specificity, particularly in advanced fibrosis and cirrhosis[19-21]. These predictive formulae of liver fibrosis were used as a non-invasive method to measure liver fibrosis. They are defined as follows: APRI= AST(/ULN) x 100/platelet($x10^9$/L)

FIB4= age x AST/(platelet count [$x\ 10^9$/L] x $ALT^{1/2}$)

Statistical analysis

The data obtained from our patients were collected in a Microsoft Access file, and statistical analysis was performed using the SAS program. Patient characteristics were described using number and percentage or median and interquartile range for categorical and continuous variables, respectively. Spearman's rank correlation (r) was used to evaluate the relationship between ARFI and TE scores after ARFI scores were converted into kPa. r was interpreted as follows; 0.90 to 1.00 (−0.90 to −1.00), very high positive (negative) correlation; 0.70 to 0.90 (−0.70 to −0.90), high positive (negative) correlation; 0.50 to 0.70 (−0.50 to −0.70), moderate positive (negative) correlation; 0.30 to 0.50 (−0.30 to −0.50), low positive (negative) correlation; 0.00 to 0.30 (0.00 to −0.30) negligible correlation. We used the Bland-Altman method to assess the agreement between the FibroScan® and ARFI scores, also stratified by FibroScan® score. In order to investigate factors associated with the strength of agreement between the two elastography methods, univariate linear regression was used with the variables of interest (one at the time) as independent variables and the absolute value of the difference between the two scores after log-transformation as outcome. A multivariate regression model was also fitted with factors showing a significant association in univariate analyses. Spearman rank correlation was used to assess the correlations between ARFI scores acquired at the right and left lobes of liver. For all tests, a P value <0.05 indicated statistically significant findings.

Results

Patient features

88 patients were recruited in total and 24(27.3%) were female. The median age of study subjects was 44 years and 85(96.6%) were Caucasian. 35 (39.8%) were of normal weight with 31(35.2%) overweight and 22(25%) had a BMI in the obese category. The most common genotype was 1a (47(53.4%)), followed by genotype 3 (34(38.6%)) and other genotypes made up a minority (7(7.9%) of study subjects. 20(22.7%) study subjects had stigmata of chronic liver disease and median FIB-4 scores and APRI scores were 1.8(1.1, 3.7) and 0.8(0.4, 2.1). Two (2.3%) subjects also had human immunodeficiency virus (HIV) infection.

Table 1. Patient characteristics at baseline visit of TRACER study

Characteristic	Description	All (n=88) n(%) or median (IQR)
Sex	Female	24 (27.3%)
Race	Caucasian	85 (96.6%)
Smoker	Current Smoker	65 (73.9%)

	Ex-Smoker/Never smoked	23 (26.1%)
Stigmata of chronic liver disease	No	68 (77.3%)
BMI (kg/m²)		26.4 (23.4, 29.9)
Obesity class by BMI (kg/m²)	Normal weight	35 (39.8%)
	Overweight	31 (35.2%)
	Class I obesity	14 (15.9%)
	Class II/III obesity	8 (9.1%)
Raised blood pressure	Yes	13 (14.8%)
Genotype	1a	47 (53.4%)
	3	34 (38.6%)
	Other	7 (7.9%)
FibroScan® score	kPa	9.6 (6.55, 21.55)
	Unobtainable/unreliable	5 (5.7%)
	<8.5kPa	34 (38.6%)
	8.5-12.5kPa	16 (18.2%)
	>12.5-14.5kPa	4 (4.5%)
	>14.5kPa	29 (33%)
ARFI score at segment 5/8	ms^{-1}	1.48 (1.17, 2.76)
	< 1.35ms^{-1}	32 (36.3%)
	1.35-1.6ms^{-1}	13 (14.8%)
	1.61 - 1.87ms^{-1}	5 (5.7%)
	>1.87ms^{-1}	38 (43.2%)
ARFI score at left lobe		1.91 (1.36, 2.75)
AUDIT alcohol score		6 (2, 10)
FIB 4 score		1.8 (1.1, 3.7)
APRI score		0.8 (0.4, 2.1)

Liver stiffness measurements

FibroScan® obtained a reliable score in 83(94.3%) subjects. The median FibroScan® score was 9.6kPa (6.55, 21.55). Using FibroScan®, 34(38.6%) participants scored <8.5kPa (suggestive of F0/F1 disease), 16(18.2%) scored 8.5-12.5kPa (suggestive of F2 disease), 4(4.5%) scored >12.5-14.5kPa (suggestive of F3 disease) and 29(33%)

scored >14.5kPa (suggestive of F4 disease). 12 subjects had a skin to liver surface distance >2.5cm and we recorded a FibroScan® score on 7/12 of these subjects.

ARFI measurements at the right lobe were acquired on all subjects with a median score of 1.48ms^{-1}(1.17, 2.76). Of note 11/12 of those whose skin to liver surface distance was >2.5cm scored >1.87ms^{-1} on ARFI.12/88 subjects had an IQR >0.3, and of these, 5 subjects had no FibroScan® recorded but scored >1.87ms^{-1} on ARFI. ARFI measurements acquired at segment 5/8 of the liver found 32(36.3%) subjects scored <1.35ms^{-1} (indicating no scarring or mild liver scarring), 13(14.8%) scored 1.35-1.6ms^{-1} (indicating significant fibrosis), 5(5.7%) scored 1.61-1.87ms^{-1} (indicating severe fibrosis) and 58(43.2%) scored >1.87ms^{1} (indicating liver cirrhosis). ARFI determined the median stiffness at the left lobe of liver to be 1.91ms^{-1}(1.36, 2.75).

Liver stiffness evaluation

Overall, when all scores were considered, there was good correlation between the FibroScan® and ARFI scores with a Spearman's coefficient (95% CI) of 0.87(0.80, 0.91), P<0.001. However, the relationship was weaker at higher readings of FibroScan® and ARFI scores, with ARFI recording higher stiffness scores than FibroScan® at higher readings. In order to explore the relationship between the two methods at higher and lower scores, subjects were grouped according to FibroScan® scores as follows <8.5kPa(n=34), 8.5-14.5kPa(n=19) and >14.5kPa(n=30). The results are displayed in table 2 below.

Table 2. Liver stiffness correlation between different score groups. There was a moderate relationship at lower and higher scores and a weak relationship between FibroScan® and ARFI when subjects scored between 8.5-14.5kPa

Group	r	95% CI	P-value
<8.5kPa	0.56	0.28, 0.76	<0.001
8.5-14.5kPa	0.29	-0.18, 0.66	0.22
>14.5kPa	0.57	0.26, 0.77	<0.001

**r interpretation: 0.90 to 1.00, very high positive correlation; 0.70 to 0.90, high positive correlation; 0.50 to 0.70; moderate positive correlation; 0 .30 to 0.50, low positive correlation; 0.00 to 0.30 negligible correlation*

A Bland-Altman plot was used to evaluate bias between the mean percentage differences of the FibroScan® and ARFI scores. The plot showed that the average (95% CI) percentage difference between the two scores was 27.5% (17.8, 37.2), P<0.001, indicating that FibroScan® scores were, on average, greater than ARFI scores. Lower and upper limits of agreement (LOA) (95% CI) were -61.3% (-77.6, -45.0) and 116.3% (100.0, 132.6). There was a negative correlation between the average and the percentage difference of the FibroScan® and ARFI scores [Spearman's r (95% CI) = -0.41(-0.57, -0.21), P<0.001] thus showing that percentage difference gets smaller for greater FibroScan® and ARFI scores.

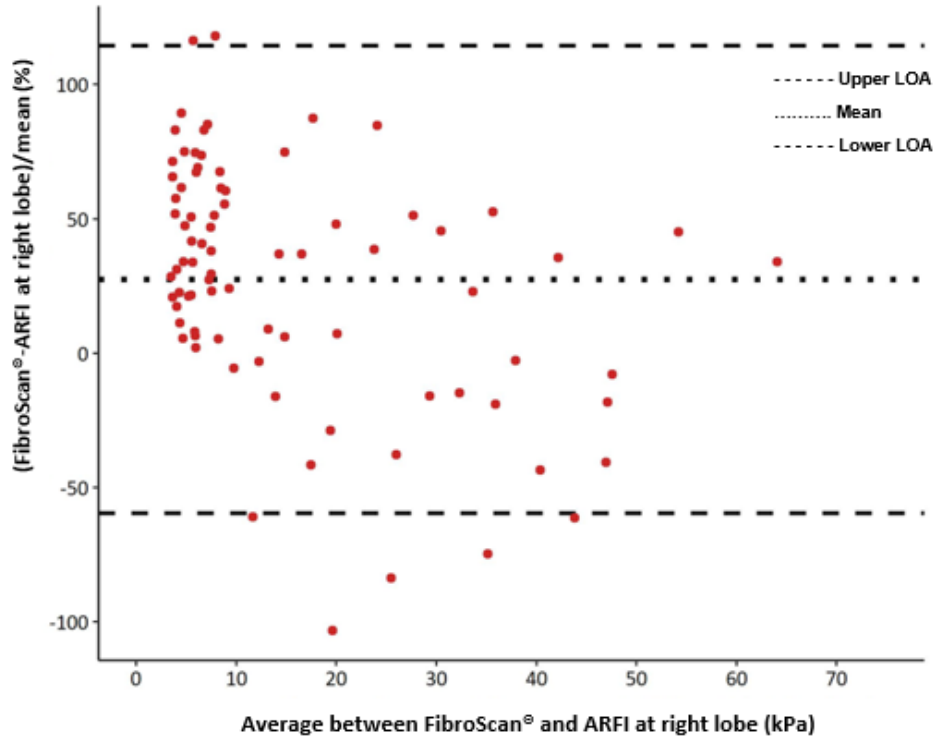

Figure 1: Bland-Altman plot of the percentage difference (y-axis) and average (x-axis) of FibroScan® and ARFI scores. The plot shows that the percentage difference gets smaller for larger FibroScan® and ARFI scores.

A Bland-Altman plot stratified by FibroScan® score groups of <8.5kPa, 8.5-14.5kPa and >14.5kPa can be seen in Figure 2. The average difference between subjects with a FibroScan® score of <8.5kPa was 1.73(1.03, 2.44), P<0.001 with a lower limit of agreement of -2.29(-3.45, -1.13) and an upper limit of 5.76(4.60, 6.91). The average difference between subjects with a FibroScan® score of 8.5-14.5kPa was 2.17(-0.96, 5.29), P=0.16 with a lower limit of agreement of -10.8(-15.8, -5.82) and an upper limit of agreement of 15.14(10.15, 20.13). For subjects that scored >14.5kPa the average difference was 1.23(-4.12, 6.57), P=0.64 with a lower limit of agreement of -27.39(-36.15, -18.64) and an upper limit of agreement of 29.85(21.09, 38.60).