



Toward acquisition protocol standardization for estimating liver fat content using ultrasound attenuation coefficient imaging

ULTRASONOGRAPHY

Giovanna Ferraioli¹, Ambra Raimondi^{1,2}, Annalisa De Silvestri³, Carlo Filice^{1,2}, Richard G. Barr^{4,5}

¹Department of Clinical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, Italy; ²Clinical Sciences and Infectious Diseases Department, Fondazione IRCCS Policlinico S. Matteo, Pavia, Italy; ³Clinical Epidemiology and Biometric Unit, Fondazione IRCCS Policlinico S. Matteo, Pavia, Italy; ⁴Department of Radiology, Northeastern Ohio Medical University, Rootstown, OH, USA; ⁵Southwoods Imaging, Youngstown, OH, USA

Purpose: This study's primary aim was to assess factors affecting ultrasound attenuation coefficient (AC) measurement repeatability using the Canon ultrasound (US) system. The secondary aim was to evaluate whether similar results were obtained with other vendors' AC algorithms.

Methods: This prospective study was performed at two centers from February to November 2022. AC was obtained using two US systems (Aplio i800 of Canon Medical Systems and Arietta 850 of Fujifilm). An algorithm combining AC and the backscatter coefficient was also used (Sequoia US System, Siemens Healthineers). To evaluate inter-observer concordance, AC was obtained by two expert operators using different transducer positions with regions of interest (ROIs) varying in terms of depth and size. Intra-observer concordance was evaluated on measurements performed intercostally, subcostally, and in the left liver lobe. Lin's concordance correlation coefficient was used.

Results: Thirty-four participants (mean age, 49.4±15.1 years; 18 females) were studied. AC values progressively decreased with depth. The measurements in intercostal spaces on best-quality US images using a 3-cm ROI with its upper edge 2 cm below the liver capsule during breath-hold showed the highest intra-observer and inter-observer concordance (0.92 [95% confidence interval, 0.88 to 0.95] and 0.89 [0.82 to 0.96], respectively). Measurements in the left lobe showed the lowest intra-observer and inter-observer concordance (0.67 [0.43 to 0.90] and 0.58 [0.12 to 1.00], respectively). Intercostal space measurements also had the highest repeatability for the other two ultrasound systems.

Conclusion: AC values obtained in intercostal spaces on best-quality images using a 3-cm ROI placed with its top 2 cm below the liver capsule were highly repeatable.

Keywords: Liver steatosis; Fat quantification; Attenuation coefficient; Fatty liver; Ultrasound; Non-alcoholic fatty liver disease; Repeatability studies

Key points: Attenuation coefficient values obtained in the right intercostal spaces were highly repeatable. Measurements in the left liver lobe are not recommended because their repeatability was only moderate, which would limit comparisons of follow-up measurements. The highest repeatability of measurement with the attenuation imaging algorithm was obtained on best-quality images, positioning the upper edge of the region of interest 2 cm below the liver capsule, avoiding reverberation artifacts, and with a 3-cm region of interest.

ORIGINAL ARTICLE

<https://doi.org/10.14366/usg.23014>
eISSN: 2288-5943
Ultrasonography 2023;42:446-456

Received: February 7, 2023
Revised: March 15, 2023
Accepted: March 27, 2023

Correspondence to:
Richard G. Barr, MD, PhD, Southwoods Imaging, 7623 Market Street, Youngstown, OH 44406, USA

Tel. +1-330-965-5100
Fax. +1-330-965-5109
E-mail: rgbarr@zoominternet.net

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2023 Korean Society of Ultrasound in Medicine (KSUM)



How to cite this article:
Ferraioli G, Raimondi A, De Silvestri A, Filice C, Barr RG. Toward acquisition protocol standardization for estimating liver fat content using ultrasound attenuation coefficient imaging. Ultrasonography. 2023 Jul;42(3):446-456.

Introduction

Non-alcoholic fatty liver disease has become the leading cause of chronic liver disease worldwide. The detection and quantification of liver fat content are of great interest not only for the diagnosis of the disease, but also for follow-up and prognostication [1,2]. Of note, a study in a large population has shown that even the presence of simple steatosis is associated with a higher risk of mortality with respect to controls [3].

Algorithms for the quantification of liver fat content with ultrasound (US) systems, based on the estimation of the US beam attenuation, backscattering, or speed of sound, are currently available [1,2,4]. Attenuation coefficient (AC) algorithms estimate the attenuation of the US beam as it traverses the tissue, whereas the backscatter coefficient is a measure of the fraction of US energy returned to the transducer from the tissue. Higher values for both parameters indicate higher liver fat content. On the contrary, the speed of sound decreases in higher-density materials, and lower speeds are therefore associated with higher liver fat content [4].

As of today, AC algorithms have been the most frequently used in published studies [1,2,4]. They have shown good to excellent performance in detecting and grading liver steatosis. In some studies, they performed significantly better than the controlled attenuation parameter (Echosens, France) for grading significant steatosis or every steatosis grade [5–7]. However, little is known about the best protocol for reliable acquisition. A panel of experts of the American Institute of Ultrasound in Medicine-RSNA Quantitative Imaging Biomarkers Alliance (QIBA) Pulse-Echo Quantitative Ultrasound (PEQUS) initiative has pointed out that, even though the literature reports clinically acceptable values obtained using vendor guidelines, the depth dependence of AC measurements needs to be evaluated [2], and recommended that, until otherwise proven, the same protocol followed for liver stiffness measurement should be applied. Currently, it is not known whether angling the transducer to the liver capsule or using different locations for the AC measurement affects its value or the intra- and inter-observer repeatability. The primary aim of this study was to assess factors that may affect the repeatability of AC measurements with the attenuation imaging (ATI) algorithm available on Canon US systems. The secondary aim was to evaluate whether the repeatability of AC measurements was affected in the same manner when using algorithms commercially available on US systems from other manufacturers.

Materials and Methods

Compliance with Ethical Standards

This institutional review board–approved, Health Insurance

Portability and Accountability Act–compliant prospective study was performed at two centers, one in Italy and the other in the USA. From February 2022 to November 2022, volunteers willing to participate in the study were consecutively enrolled after signing a written informed consent form.

Patients

The inclusion criteria were age >18 years and good health. Patients with focal fatty deposition or fatty sparing were excluded.

For each participant, the biometric characteristics (age, sex, body mass index, and waist circumference) were recorded. Liver steatosis was assessed using B-mode US [8]. The skin-to-liver capsule distance in centimeters (cm) and liver stiffness in kilopascals (kPa) were measured with the Aplio i800 US system (Canon Medical Systems, Tokyo, Japan). Liver stiffness and AC measurements were obtained after 4 hours of fasting.

AC was obtained using the ATI algorithm implemented in the Aplio i-series US systems and the "improved" attenuation (iATT) algorithm implemented in the Arietta 850 US system (Fujifilm, Tokyo, Japan). They both give measurements in decibels per centimeter per megahertz (dB/cm/MHz). Moreover, the ultrasound-derived fat fraction (UDFF) algorithm, which is a combination of AC and the backscatter coefficient implemented in the Sequoia US system (Siemens Healthineers, Erlangen, Germany), was also used for the purpose of this study. The measurements are given as the percent of fat (%).

The features of the three algorithms are the following: ATI: A large field of view (length of 100 mm, upper and lower width of 45 mm and 80 mm, respectively) covering around 70% of the B-mode image was chosen. The minimum size of the region of interest (ROI) was 30×30 mm. The quality of the measurement is displayed as an R^2 value, and only the measurements with an $R^2 \geq 0.90$ were recorded, as recommended by the manufacturer (Fig. 1). The measurements were performed with an i8CX1 convex transducer. A measurement was considered a failure when the R^2 value was below 0.70, as recommended by the vendor.

Improved Attenuation

The AC was obtained together with a liver stiffness measurement; the ROI was not color-coded and the size (both width and length) was not user-adjustable. The measurement was made in a fixed area (length, 35–75 mm from the skin; upper and lower widths of 5.8 mm and 8.5 mm, respectively) (Fig. 2). The measurements were performed with a C6-1 convex transducer.

Ultrasound-Derived Fat Fraction

The ROI had a fixed size (length of 40 mm, upper and lower

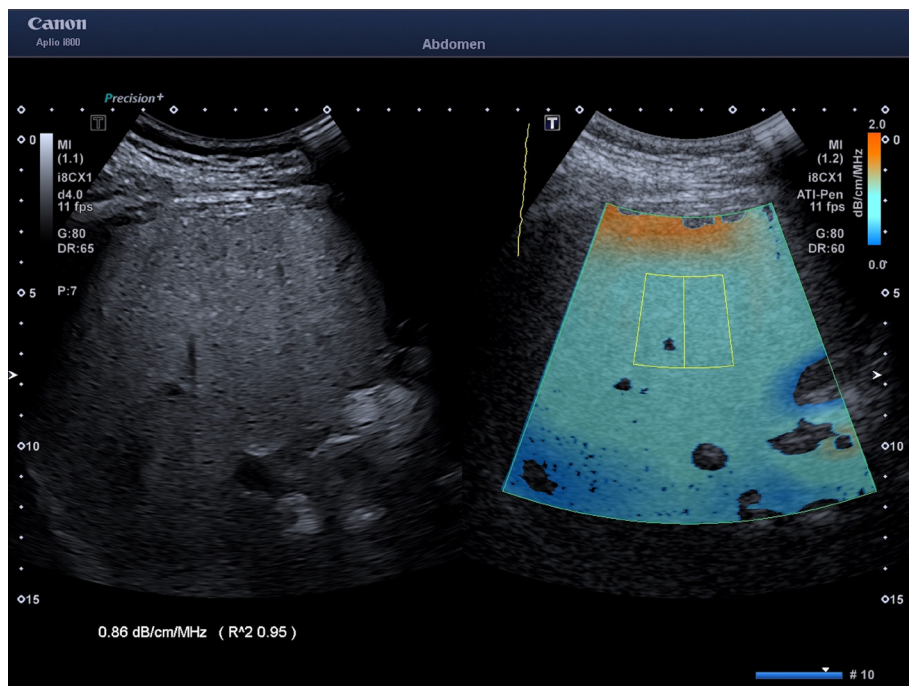


Fig. 1. Measurement performed in the "a" position using the attenuation imaging algorithm implemented on the Aplio i800-series ultrasound systems (Canon Medical Systems, Tokyo, Japan). The field of view is color-coded; therefore, artifacts generally due to reverberation under the liver capsule (dark orange area) or due to a low signal-to-noise ratio in the far field (dark blue areas) can easily be visualized and not included in the measurement box. The attenuation coefficient value is given in decibels per centimeter per megahertz (dB/cm/MHz). The algorithm provides a quality measure for fat quantification (bottom left: R^2) which should be 0.90 or above for a high-quality measurement.

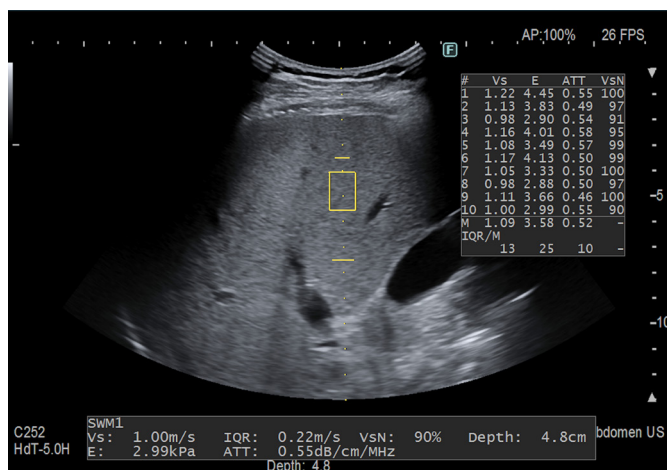


Fig. 2. Measurement performed in the "b" position using the improved attenuation (iATT) algorithm implemented on the Arietta 850 ultrasound system (Fujifilm, Tokyo, Japan). The attenuation coefficient value is given in decibels per centimeter per megahertz (dB/cm/MHz) together with liver stiffness value. If the quality of the measurement is low, the iATT and stiffness values are not shown and "Vs output failed" appears on the monitor.

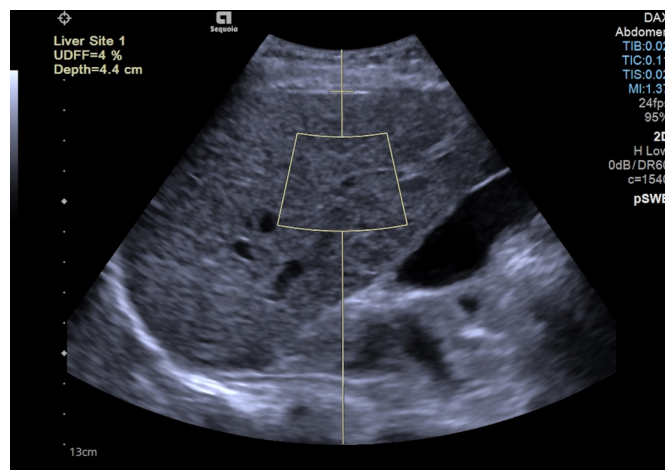


Fig. 3. Measurement performed in the "b" position using the ultrasound-derived fat fraction algorithm implemented on the Sequoia ultrasound system (Siemens Healthineers, Erlangen, Germany). The value is given as the percentage of fat (%) and it is not shown when the software judges that the quality of the acquisition is low: in that case, an "XXX" sequence instead of the percentage appears on the monitor.

widths of 30 mm and 40 mm, respectively) and the manufacturer recommended obtaining the measurement at a fixed depth (upper edge of the ROI at 15 mm below the liver capsule) (Fig. 3). The measurements were performed with the deep abdominal transducer (DAX: 1–3.5 MHz).

To assess inter-observer concordance, each patient's examinations were carried out by two expert operators at each center (G.F.

and A.R. in Italy, with 37 and 5 years of experience, respectively; and R.G.B., J.S., and B.N. in the United States, with 32, 15, and 7 years of experience, respectively), using different positions of the convex transducer or different depth/size of the ROI as follows: (a) intercostal on the best-quality image, (i.e., the one with fewer vessels, a strong B-mode signal without artifacts, and not necessarily following the protocol for stiffness acquisition), with the upper edge

of the ROI (size, 3 cm) at 2 cm below the liver capsule during a breath-hold; (b) as "a" but with the transducer always perpendicular to the liver capsule following the recommendation for liver stiffness assessment [9–11]; (c) as "a" but at a 2.5 cm depth; (d) as "a" but at a 3 cm depth; (e) as "a" but with an ROI size of 5 cm; (f) as "a" but subcostal; (g) as "a" but in the left liver lobe using a longitudinal scan; (h) each measurement in a different location with the size and depth of the ROI as "a"; and (i) as "a" but while individuals were freely breathing. For all the measurements with the ATI algorithm, care was taken not to include artifacts in the ROI.

Since for both the iATT and UDFE algorithms, the measurements were obtained using a fixed size of the ROI and at a fixed depth, positions "c," "d," and "e" were not evaluated. Moreover, since AC with the iATT algorithm is obtained together with liver stiffness, the protocol recommended for liver stiffness was followed, with the transducer always perpendicular to the liver capsule while scanning in the intercostal space [8]. UDFE measurements were obtained with the DAX transducer always perpendicular to the liver capsule because the quality of the image degrades when this transducer is angled. Therefore, for the iATT and the UDFE algorithms, position "a" was not evaluated.

Liver stiffness measurements were obtained following the recommended protocol [9–11].

The time estimated to study each volunteer was about 2 hours. Scanning was randomized to reduce bias.

The median value of five consecutive AC measurements for each patient was used for the statistical analysis.

Intra-observer concordance was assessed by repeating the measurements intercostally, subcostally, and in the left liver lobe at the end of each examination. For the ATI algorithm, the "a," "f," and "g" positions were used, whereas for the iATT and UDFE algorithms the "b," "f," and "g" positions were evaluated for intra-observer concordance. For the statistical analysis, the data obtained at both centers were merged; operator 1 was considered the expert with the highest experience at each center and operator 2 was defined as any of the other experts.

Statistical Analysis

Sample size: A sample size of 30 subjects with two observations per subject achieves more than 90% power to detect a concordance correlation coefficient (CCC) of 0.95 (considered excellent) under the alternative hypothesis when the intraclass correlation under the null hypothesis is 0.80 (considered good) using an F-test with a significance level of 0.05.

Descriptive statistics were calculated for the demographic characteristics of the study sample. The results were expressed as the mean value and standard deviation. Qualitative variables were

summarized as counts and percentages. The intra- and inter-observer repeatability was assessed with Lin's CCC [12]. The CCC combines precision and accuracy to determine how far the observed data deviate from the line of perfect concordance (i.e., the line at 45° on a square scatterplot). The CCC increases in value as a function of the nearness of the data's reduced major axis to the line of perfect concordance (the accuracy of the data) and of the tightness of the data about its reduced major axis (the precision of the data). CCC values range from 0 to +1. A CCC value of 0 indicates that most of the error originates from differences in measurements between operators. As CCC values approach 1, the measurement differences between the different operators become negligible and more consistent. The agreement was classified as poor (0 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), good (0.61 to 0.80), and excellent (0.81 to 1.00) [13]. The CCCs were reported with 95% confidence intervals (CIs). Bland-Altman analysis was also performed to calculate the mean difference and 95% limits of agreement of measurements [14].

A P-value <0.05 was considered statistically significant. All tests were two-sided. The data analysis was performed with the STATA statistical package by A.D.S. (release 17.0, 2021, Stata Corp., College Station, TX, USA).

Results

Overall, 34 participants (mean age, 49.4±15.1 years; 18 women) were studied. Twenty-four participants were studied in Italy and 10 in the United States. All 34 were studied using the ATI algorithm, while a subset of 20 participants was studied with iATT, and a subset of 18 participants with UDFE.

Table 1 reports the characteristics of the study cohort. The North American participants were older and had diabetes at a significantly higher rate. No other significant differences were observed.

A preliminary statistical analysis of the ATI data of the first 14 consecutive participants (mean age, 51.8±16.6 years; 8 males) showed that the repeatability of measurement in the "g" position was the lowest, ranging from moderate to poor (Table 2); therefore, this position was not used in the individuals that were enrolled thereafter. There was not any statistically significant difference between this sample and the remaining study cohort.

The mean values obtained with the algorithms used in this study (i.e., ATI, iATT, and UDFE) are reported in Table 2. AC values progressively decreased with depth. The mean values of AC with ATI were 0.64±0.10 dB/cm/MHz at 2 cm ("a" position), 0.62±0.10 dB/cm/MHz at 2.5 cm ("c" position), and 0.59±0.09 dB/cm/MHz at 3.0 cm ("d" position).

Table 1. Characteristics of the study cohort

Variable	Overall	Italy	USA	P-value
Sample size (%)	34	24/34 (70.6)	10/34 (29.4)	–
Female sex (%)	18/34 (52.9)	13/24 (54.2)	5/10 (50.0)	0.822
Age (year)	49.4±15.1	45.8±14.5	57.8±13.6	0.031
BMI (kg/m ²)	26.0±5.5	25.5±4.4	27.3±7.6	0.383
Waist circumference (cm)	88.1±15.3	90.9±14.1	83.5±17.9	0.272
Diabetes (%)	4/34 (11.8)	1/34 (4.2)	3/34 (30.0)	0.035
Skin-to-liver capsule distance (cm)	1.87±0.10	1.88±0.11	1.86±0.25	0.931
B-mode ultrasound steatosis (%)	11/34 (32.4)	7/34 (29.2)	4/34 (40.0)	0.540
Liver stiffness (kPa)	4.7 (1.7)	4.4 (1.4)	5.3 (2.1)	0.162
ATI sample size (%)	34	24/34 (70.6)	10/34 (29.4)	–
iATT sample size (%)	20	20/20 (100)	–	–
UDFF sample size (%)	18	15/18 (83.3)	3/18 (16.7)	–

Values are presented as number (%) or mean±SD, unless otherwise indicated.

BMI, body mass index; ATI, attenuation imaging (Canon); iATT, improved attenuation (Fujifilm); UDFF, ultrasound-derived fat fraction (Siemens); SD, standard deviation.

Table 2. Mean values of the ATI, iATT, and UDFF algorithms obtained in different positions

Variable	Mean±SD
ATI (dB/cm/MHz)	
(a) Intercostal, 2 cm ^a , best image (ROI size, 3 cm)	0.64±0.10
(b) Intercostal, 2 cm ^a , 90° (ROI size, 3 cm)	0.64±0.09
(c) Intercostal, 2.5 cm ^a (ROI size, 3 cm)	0.62±0.10
(d) Intercostal, 3.0 cm ^a (ROI size, 3 cm)	0.59±0.09
(e) Intercostal, 2 cm ^a , best image (ROI size, 5 cm)	0.60±0.09
(f) Subcostal, 2 cm ^a , best image	0.67±0.13
(g) Left lobe, 2 cm ^a , best image (n=14)	0.68±0.14
(h) Each measurement in a different location ^b	0.65±0.10
(i) Free breathing, intercostal ^b	0.65±0.12
iATT (dB/cm/MHz)	
Intercostal	0.63±0.10
Subcostal	0.62±0.09
Each measurement in a different location	0.65±0.10
Free breathing, intercostal	0.67±0.09
UDFF (%)	
Intercostal	6.6±4.2
Subcostal	9.1±7.4
Each measurement in a different location	6.6±4.2
Free breathing, intercostal	5.8±3.2

ATI, attenuation imaging; iATT, improved attenuation; UDFF, ultrasound-derived fat fraction; SD, standard deviation; ROI, region of interest.

^aThe distance refers to the upper edge of the ROI with respect to the liver capsule.

^bROI with a size of 3 cm and upper edge of the ROI at a 2 cm below the capsule, best image.

Attenuation Imaging

There were 4/14 (28.6%) failures for AC measurement in the left liver lobe and 3/34 (8.8%) failures for measurements that were obtained subcostally. No failures for measurements in the intercostal space were reported.

Table 3 reports the intra-observer and inter-observer concordance for the ATI measurements obtained using various transducer positions and ROI sizes. The "a" position, in the right intercostal space, showed the highest intra-observer and inter-observer concordance (0.92 [95% CI, 0.88 to 0.95] and 0.89 [95% CI, 0.82 to 0.96], respectively). Moreover, the limits of agreement (i.e., the interval of two standard deviations of the measurement differences on either side of the mean difference) were the narrowest for both: mean difference: 0.002 dB/cm/MHz with limits of agreement from -0.080 to 0.084 dB/cm/MHz for the intra-observer concordance, and mean difference: 0.005 dB/cm/MHz with limits of agreement from -0.085 to 0.095 dB/cm/MHz for the inter-observer concordance. These narrow ranges indicate that the measurement in this position had the highest precision. Using the "a" position as reference, the highest concordance was observed with the "b" position. Of note, when increasing the ROI size from 3 cm to 5 cm on the best-quality US image obtained in intercostal spaces, the inter-observer concordance was 0.78 (95% CI, 0.63 to 0.93), whereas in the same position but with a 3-cm ROI it was 0.89 (95% CI, 0.82 to 0.96).

Figs. 4–6 show the Bland-Altman plots of differences in measurements.

Table 3. Intra-observer and inter-observer concordance in ATI measurements

Position of the ROI ^{a)} (ROI size)	CCC (95% CI)	Pearson's r	Mean difference, dB/cm/MHz (95% limits of agreement)
Intra-observer concordance (best image, ROI at 2 cm below the liver capsule, ROI length: 3 cm)			
Intercostal	0.92 (0.88 to 0.95)	0.92	0.002 (-0.080 to 0.084)
Subcostal	0.80 (0.72 to 0.89)	0.81	-0.001 (-0.141 to 0.139)
Left lobe (n=14)	0.67 (0.43 to 0.90)	0.70	-0.001 (-0.181 to 0.183)
Inter-observer concordance			
(a) Intercostal, 2 cm, best image (3 cm)	0.89 (0.82 to 0.96)	0.89	0.005 (-0.085 to 0.095)
(b) Intercostal, 2 cm, 90° (3 cm)	0.71 (0.53 to 0.89)	0.72	0.001 (-0.142 to 0.145)
(c) Intercostal, 2.5 cm (3 cm)	0.78 (0.64 to 0.91)	0.78	-0.004 (-0.137 to 0.128)
(d) Intercostal, 3.0 cm (3 cm)	0.76 (0.62 to 0.91)	0.76	0.003 (-0.126 to 0.131)
(e) Intercostal, 2 cm, best image (5 cm)	0.78 (0.63 to 0.93)	0.78	0.004 (-0.112 to 0.120)
(f) Subcostal, 2 cm, best image	0.74 (0.57 to 0.90)	0.76	0.001 (-0.174 to 0.176)
(g) Left lobe, 2 cm, best image (n=14)	0.58 (0.12 to 1.00)	0.58	-0.008 (-0.248 to 0.232)
(h) Each measurement in a different location ^{b)}	0.69 (0.51 to 0.87)	0.70	0.005 (-0.141 to 0.151)
(i) Free breathing, intercostal ^{b)}	0.69 (0.51 to 0.87)	0.71	0.023 (-0.156 to 0.202)
Concordance assessed using the values obtained with the (a) position as reference			
(b) Intercostal, 2 cm, 90° (3 cm)	0.89 (0.84 to 0.94)	0.90	0.004 (-0.077 to 0.086)
(c) Intercostal, 2.5 cm (3 cm)	0.82 (0.74 to 0.90)	0.83	-0.020 (-0.133 to 0.093)
(d) Intercostal, 3.0 cm (3 cm)	0.61 (0.48 to 0.74)	0.70	-0.052 (-0.199 to 0.096)
(e) Intercostal, 2 cm, best image (5 cm)	0.67 (0.54 to 0.80)	0.74	-0.038 (-0.160 to 0.083)
(f) Subcostal, 2 cm, best image	0.53 (0.36 to 0.71)	0.56	0.034 (-0.174 to 0.241)
(g) Left lobe, 2 cm, best image (n=14)	0.17 (-0.26 to 0.61)	0.18	0.036 (-0.296 to 0.369)
(h) Each measurement in a different location ^{b)}	0.81 (0.72 to 0.89)	0.81	0.009 (-0.108 to 0.125)
(i) Free breathing, intercostal ^{b)}	0.81 (0.73 to 0.89)	0.82	0.002 (-0.130 to 0.134)

ATI, attenuation imaging; ROI, region of interest; CCC, concordance correlation coefficient; CI, confidence interval.

^{a)}The position refers to the upper edge of the ROI with respect to the liver capsule. ^{b)}ROI at 2 cm, best image.

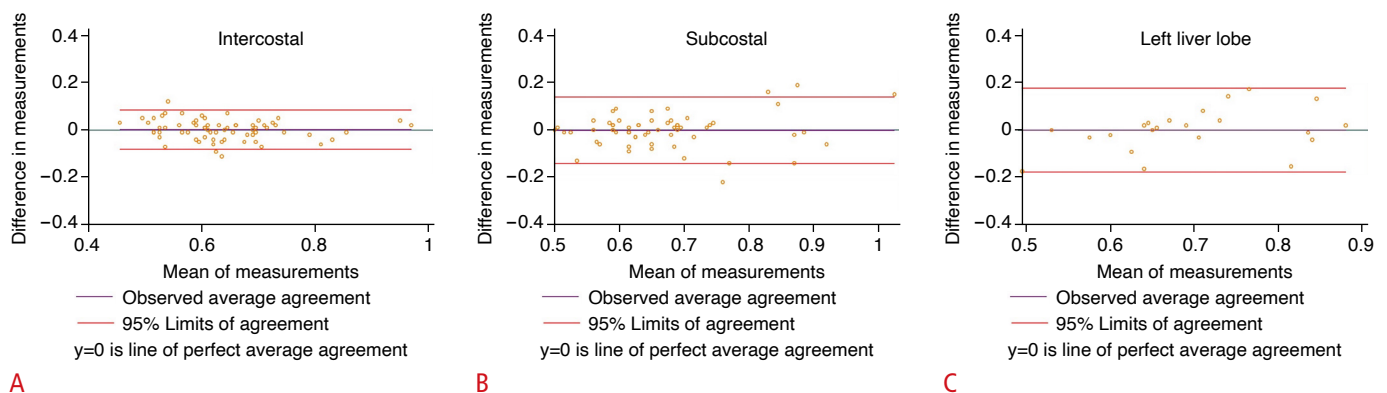


Fig. 4. Intra-observer concordance estimated with the attenuation imaging algorithm: Bland and Altman plot of differences in measurements. The green line ($y=0$) is a line of perfect average agreement. The purple line represents the mean of the difference in measurements. The red lines define the limits of agreement (mean of the difference [2 standard deviations]). Results of intercostal scanning (A), subcostal scanning (B), and left lobe scanning (C) are shown.

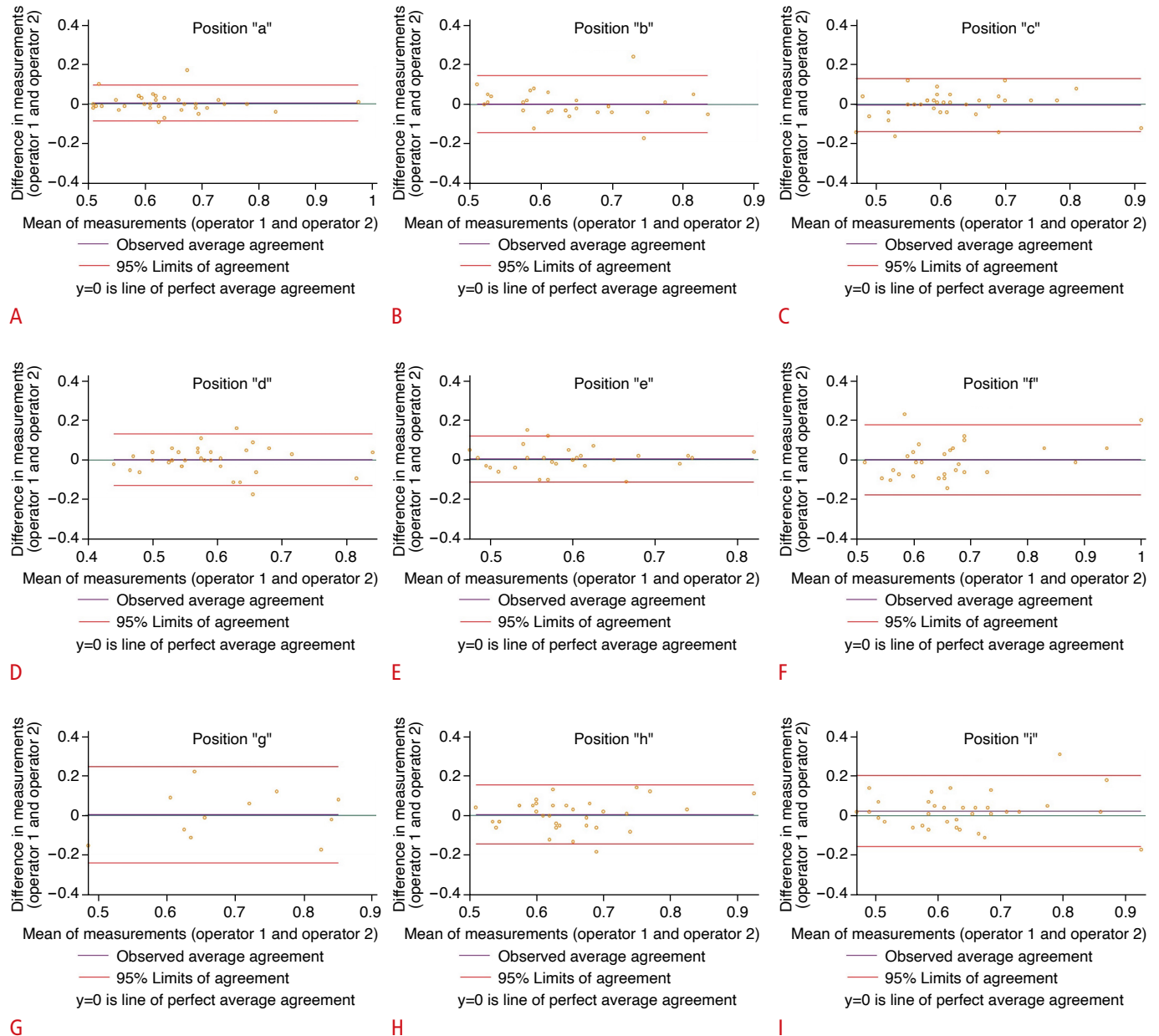


Fig. 5. Inter-observer concordance with the attenuation imaging algorithm: Bland and Altman plot of differences in measurements. The green line ($y=0$) is a line of perfect average agreement. The purple line represents the mean of the difference in measurements. The red lines define the limits of agreement (mean of the difference $[2 \text{ standard deviations}]$). Results of position "a" (A), position "b" (B), position "c" (C), position "d" (D), position "e" (E), position "f" (F), position "g" (G), position "h" (H), and position "i" (I) are shown.

Improved Attenuation

There were 3/20 (15.0%) failures for AC measurements obtained subcostally. Table 4 reports the intra-observer and inter-observer concordance for the measurements with the iATT algorithm obtained using various transducer positions. The highest intra-observer and inter-observer concordance values were observed for measurements obtained in the right intercostal space.

Ultrasound-Derived Fat Fraction

There were 2/18 (11.1%) failures for measurements obtained subcostally. Table 5 reports the intra-observer and inter-observer concordance for the measurements with the UDF algorithm obtained using various transducer positions. The highest intra-observer and inter-observer concordance values were observed for measurements acquired in the right intercostal space.

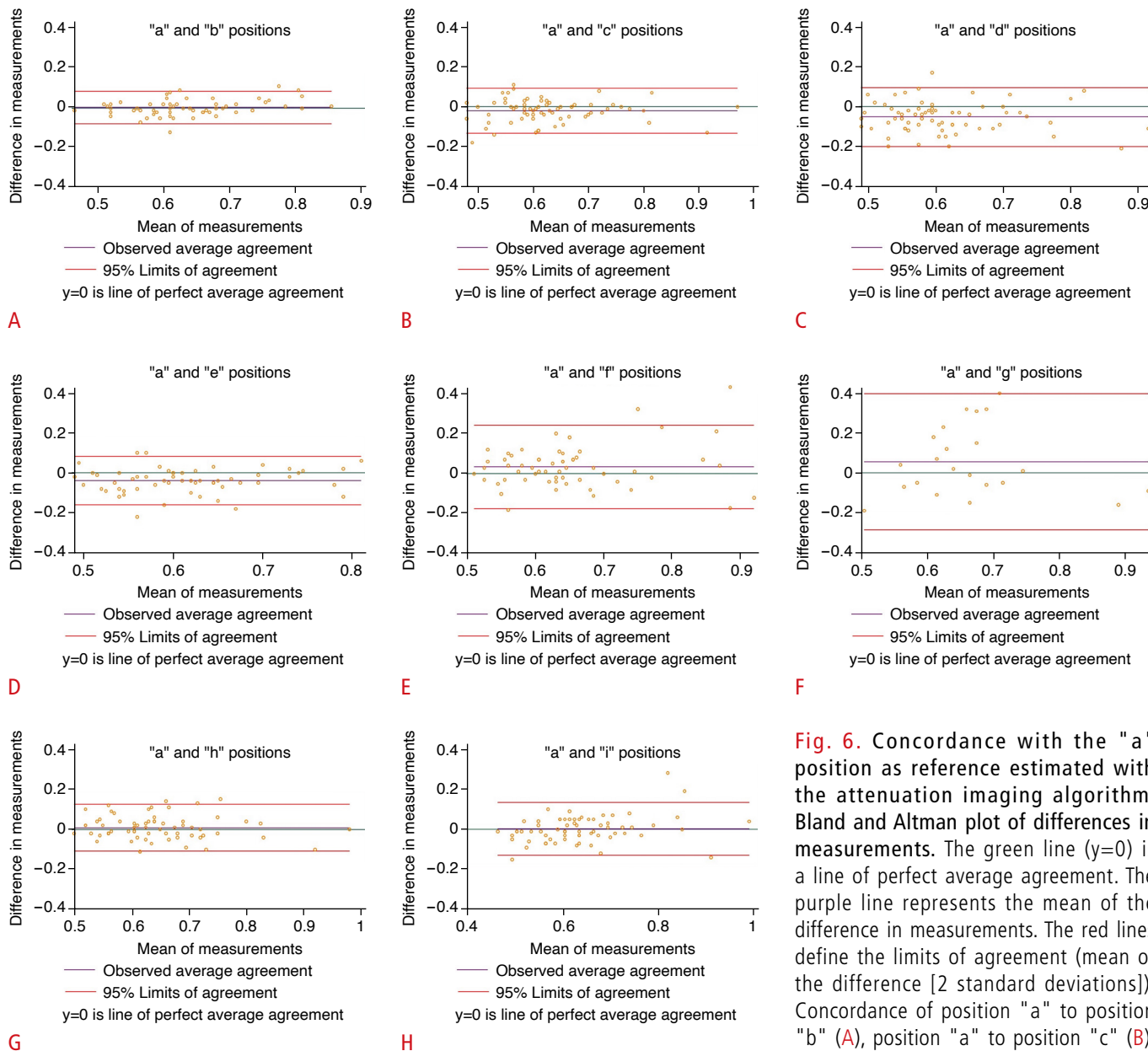


Fig. 6. Concordance with the "a" position as reference estimated with the attenuation imaging algorithm: Bland and Altman plot of differences in measurements. The green line ($y=0$) is a line of perfect average agreement. The purple line represents the mean of the difference in measurements. The red lines define the limits of agreement (mean of the difference [2 standard deviations]). Concordance of position "a" to position "b" (A), position "a" to position "c" (B), position "a" to position "d" (C), position "a" to position "e" (D), position "a" to position "f" (E), position "a" to position "g" (F), position "a" to position "h" (G), position "a" to position "i" (H) are shown.

Discussion

The results of this study show that AC measurements obtained using the right intercostal space on best-quality images showed the highest intra-observer and inter-observer repeatability across all US systems that were evaluated, and that measurements in intercostal spaces were always feasible.

It must be highlighted that increasing the ROI size from 3 cm to 5 cm on the best-quality US images obtained in intercostal spaces reduced the inter-observer concordance from excellent to good. It can be speculated that at higher depths there is a decrease in the signal-to-noise ratio that might affect the consistency of the measurements in an unpredictable manner.

A preliminary analysis of the results obtained in the first 14

Table 4. Intra-observer and inter-observer concordance in the attenuation coefficient values obtained with the improved attenuation algorithm

Position of the transducer	CCC (95% CI)	Pearson's r	Mean difference (dB/cm/MHz) (95% limits of agreement)
Intra-observer concordance			
Intercostal	0.80 (0.64 to 0.90)	0.80	0.003 (-0.133 to 0.139)
Subcostal	0.57 (0.33 to 0.81)	0.57	0.003 (-0.143 to 0.150)
Inter-observer concordance			
Intercostal	0.80 (0.63 to 0.97)	0.80	0.008 (-0.127 to 0.143)
Subcostal	0.39 (-0.04 to 0.81)	0.39	0.000 (-0.179 to 0.179)
Each measurement in a different location	0.76 (0.57 to 0.95)	0.78	0.018 (-0.119 to 0.155)
Free breathing, intercostal	0.49 (0.17 to 0.82)	0.52	0.018 (-0.191 to 0.226)

CCC, concordance correlation coefficient; CI, confidence interval.

Table 5. Intra-observer and inter-observer concordance in the measurements obtained with the ultrasound-derived fat fraction algorithm

Position of the transducer	CCC (95% CI)	Pearson's r	Mean difference, % (95% limits of agreement)
Intra-observer concordance			
Intercostal	0.91 (0.85 to 0.97)	0.91	0.188 (-5.138 to 5.513)
Subcostal	0.81 (0.69 to 0.92)	0.82	0.056 (-5.907 to 6.018)
Inter-observer concordance			
Intercostal	0.70 (0.46 to 0.94)	0.71	-0.222 (-7.194 to 6.750)
Subcostal	0.67 (0.40 to 0.95)	0.69	-0.375 (-11.011 to 10.261)
Each measurement in a different location	0.38 (-0.01 to 0.77)	0.40	0.500 (-7.404 to 8.404)
Free breathing, intercostal	0.57 (0.28 to 0.85)	0.64	-1.444 (-8.406 to 5.517)

CCC, concordance correlation coefficient; CI, confidence interval.

consecutive participants highlighted that ATI measurements performed in the left liver lobe had the lowest reproducibility and the highest rate of failure. Moreover, when the AC values obtained in the left liver lobe were compared with those obtained with the position that showed the highest repeatability of the AC values, the agreement was very poor (0.17 [95% CI, -0.26 to 0.61]). These results were confirmed through the analysis of the AC values obtained in the first four participants with the iATT algorithm (results not shown). Therefore, measurements in the left liver lobe were not performed in the participants that were enrolled thereafter in order to avoid unnecessary discomfort related to the long time needed to complete all measurements. It must be underscored that the AC values obtained in the left liver lobe were higher than those obtained subcostally or intercostally. The reason for this difference is unclear. As a speculative explanation, it might be due to artifactual areas generated by the rhythmic motion of the left liver lobe due to heartbeats, which may lead to a false increase in the AC value.

UDFF, which is one of the algorithms used in this study, is a

combination of AC and the backscatter coefficient. The results obtained using the UDF algorithm followed the same trend observed with the ATI and iATT algorithms that estimate only AC. However, it is unclear whether the repeatability of the values of the backscatter coefficient parameter included in the UDF algorithm is affected in the same manner as the AC because it was not possible to separately evaluate the two parameters. Further studies are needed to assess whether the repeatability of the backscatter coefficient is affected by different positions of the transducer. The same applies to the estimation of the speed of sound, which is another parameter available for the quantification of liver fat content that was not evaluated in the present study.

This study assessed the repeatability (i.e., the precision) of measurements and not their accuracy. It must be underscored that a measurement can be highly precise but inaccurate. However, several published studies have evaluated the accuracy of AC algorithms using either liver biopsy or magnetic resonance imaging-derived proton density fat fraction as the reference standard [1,2,4]. Those

studies generally used the protocol recommended for liver stiffness assessment; therefore, AC measurements were performed in the intercostal space, and the accuracy ranged from good to excellent. However, the AC best thresholds for detecting and grading liver fat content were different among studies [1,2,4,15]. These differences might have been due to differences in the protocols used to acquire AC values.

AC estimation can be used as a readily available non-invasive tool for diagnosing or following up patients with non-alcoholic fatty liver disease. Therefore, a standard approach that limits artifacts and that is highly reproducible for fat quantification is strongly needed. As expected, depth was found to affect the AC value and its repeatability; however, this can be controlled by standardizing the protocol for acquiring the AC values. In the present study, the influence of depth was evaluated only with the ATI algorithm. Further studies are needed to evaluate the effect of depth on AC values with different algorithms.

The major strength of this study is that the prevalence of liver steatosis in the Italian and North American samples, as evaluated by B-mode US, was 29.2% and 40.0%, respectively, with an overall prevalence of 34.2%. This prevalence is similar to that reported in the general population in a recent large meta-analysis [14]; therefore, the findings of this study would likely be the same in a different sample. Another strength is that several positions of the transducer and different depths or sizes of the ROI were evaluated, and this helped address sources of bias that may affect the repeatability of AC measurements. Moreover, the study involved two centers located on two different continents, and the data for the statistical analysis were combined to address another source of bias that is inevitable in single-center studies.

There are some limitations to this study. First, several vendors have implemented AC algorithms in their US systems. This study used only three US systems, and in one of them, the AC was combined with the backscatter coefficient. Second, the measurements with the UDFC were made using the DAX probe, which was designed for obese individuals, even though the participants enrolled in this study were mostly normal or overweight. However, the UDFC algorithm was available only on the DAX probe at the time of the study. Third, these results might not be applicable to morbidly obese individuals. Fourth the intersystem variability could not be calculated due to the small sample size. However, this was beyond the aims of the study.

In conclusion, this study shows that AC values obtained from images in the right intercostal space were highly repeatable, whereas measurements in the left liver lobe had insufficient repeatability for follow-up studies and therefore cannot be recommended. With the ATI algorithm, the highest repeatability of the AC measurement was obtained on images of the best quality, by

positioning the upper edge of the ROI 2 cm below the liver capsule, avoiding including reverberation artifacts, and with a 3-cm region of interest. To evaluate the accuracy of the AC in detecting and grading liver steatosis, a standardized protocol is needed; otherwise, the cutoff values are likely to be inconsistent between studies.

ORCID: Giovanna Ferraioli: <https://orcid.org/0000-0002-6344-697X>; Ambra Raimondi: <https://orcid.org/0000-0003-0549-310X>; Annalisa De Silvestri: <https://orcid.org/0000-0003-3128-8441>; Carlo Filice: <https://orcid.org/0000-0002-0873-3415>; Richard G. Barr: <https://orcid.org/0000-0001-8021-0072>

Author Contributions

Conceptualization: Ferraioli G, Raimondi A, De Silvestri A, Filice C, Barr RG. Data acquisition: Ferraioli G, Raimondi A, Filice C, Barr RG. Data analysis or interpretation: Ferraioli G, Raimondi A, De Silvestri A, Barr RG. Drafting of the manuscript: Ferraioli G, Raimondi A, De Silvestri A, Filice C, Barr RG. Critical revision of the manuscript: Ferraioli G, Raimondi A, De Silvestri A, Filice C, Barr RG. Approval of the final version of the manuscript: all authors.

Conflict of Interest

Giovanna Ferraioli, Carlo Filice and Richard G. Barr do have conflicts of interest not related to this study.

Giovanna Ferraioli: speaker for Canon Medical Systems, Fujifilm Healthcare, Mindray Bio-Medical Electronics Co., Philips Healthcare, Siemens Healthineers; advisory board member of Philips Healthcare and Siemens Healthineers; royalties from Elsevier Publisher.

Carlo Filice: unrestricted research grant from Canon Medical Systems, Esaote SpA, Fujifilm Healthcare, Mindray Bio-Medical Electronics Co.

Richard G. Barr: speaker for Canon Medical systems, Philips Ultrasound, Siemen Healthineers, Mindray, Samsung Ultrasound, Hologic Ultrasound. Research grants from Philips Ultrasound, Canon Ultrasound, Canon MRI, Samsung, Siemens Healthineers, Hologic, Mindray.

Acknowledgments

The authors wish to thank Mrs. Janet Scacchetti, RDMS and Mr. Brandy Neill, RDMS for performing the readings with the ultrasound systems that were used in the USA. The authors are grateful to Mrs. Nadia Locatelli for her valuable help in carrying out the study. We thank Janet Schetti and Brandy Neill for their help in collecting data.

References

1. Ferraioli G, Berzigotti A, Barr RG, Choi BI, Cui XW, Dong Y, et al. Quantification of liver fat content with ultrasound: a WFUMB

- position paper. *Ultrasound Med Biol* 2021;47:2803-2820.
2. Ferraioli G, Kumar V, Ozturk A, Nam K, de Korte CL, Barr RG. US attenuation for liver fat quantification: an AIUM-RSNA QIBA Pulse-Echo Quantitative Ultrasound Initiative. *Radiology* 2022;302:495-506.
 3. Simon TG, Roelstraete B, Khalili H, Hagstrom H, Ludvigsson JF. Mortality in biopsy-confirmed nonalcoholic fatty liver disease: results from a nationwide cohort. *Gut* 2021;70:1375-1382.
 4. Ferraioli G, Soares Monteiro LB. Ultrasound-based techniques for the diagnosis of liver steatosis. *World J Gastroenterol* 2019;25:6053-6062.
 5. Ferraioli G, Maiocchi L, Raciti MV, Tinelli C, De Silvestri A, Nichetti M, et al. Detection of liver steatosis with a novel ultrasound-based technique: a pilot study using MRI-derived proton density fat fraction as the gold standard. *Clin Transl Gastroenterol* 2019;10:e00081.
 6. Ferraioli G, Maiocchi L, Savietto G, Tinelli C, Nichetti M, Rondanelli M, et al. Performance of the attenuation imaging technology in the detection of liver steatosis. *J Ultrasound Med* 2021;40:1325-1332.
 7. Fujiwara Y, Kuroda H, Abe T, Ishida K, Oguri T, Noguchi S, et al. The B-mode image-guided ultrasound attenuation parameter accurately detects hepatic steatosis in chronic liver disease. *Ultrasound Med Biol* 2018;44:2223-2232.
 8. Barr RG. Conventional ultrasound findings in chronic liver disease. In: Barr RG, Ferraioli G, eds. *Multiparametric ultrasound for the assessment of diffuse liver disease: a practical approach*. Philadelphia, PA: Elsevier, 2022;7-24.
 9. Barr RG, Ferraioli G, Palmeri ML, Goodman ZD, Garcia-Tsao G, Rubin J, et al. Elastography assessment of liver fibrosis: Society of Radiologists in Ultrasound Consensus Conference Statement. *Ultrasound Q* 2016;32:94-107.
 10. Barr RG, Wilson SR, Rubens D, Garcia-Tsao G, Ferraioli G. Update to the Society of Radiologists in Ultrasound Liver Elastography Consensus Statement. *Radiology* 2020;296:263-274.
 11. Ferraioli G, Wong VW, Castera L, Berzigotti A, Sporea I, Dietrich CF, et al. Liver ultrasound elastography: an update to the World Federation for Ultrasound in Medicine and Biology guidelines and recommendations. *Ultrasound Med Biol* 2018;44:2419-2440.
 12. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;45:255-268.
 13. Kramer MS, Feinstein AR. Clinical biostatistics. LIV. The biostatistics of concordance. *Clin Pharmacol Ther* 1981;29:111-123.
 14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.
 15. Cassinotto C, Jacq T, Anselme S, Ursic-Bedoya J, Blanc P, Faure S, et al. Diagnostic performance of attenuation to stage liver steatosis with MRI proton density fat fraction as reference: a prospective comparison of three US machines. *Radiology* 2022;305:353-361.