

Improved Assessment of Hepatic Steatosis in Humans Using Multi-Parametric Quantitative Ultrasound

Aiguo Han

Department of Electrical and
Computer engineering
University of Illinois at Urbana-
Champaign
Urbana, IL, USA
han51@illinois.edu

Andrew S. Boehringer

Department of Radiology
University of California, San
Diego
La Jolla, CA, USA
asboehringer@ucsd.edu

Yingzhen N. Zhang

Department of Radiology
University of California, San
Diego
La Jolla, CA, USA
ynancyzhang9@gmail.com

Vivian Montes

Department of Radiology
University of California, San
Diego
La Jolla, CA, USA
v2montes@ucsd.edu

Michael P. Andre

Department of Radiology
University of California, San
Diego
La Jolla, CA, USA
mandre@ucsd.edu

John W. Erdman, Jr.

Department of Food Science and
Human Nutrition
University of Illinois at Urbana-
Champaign
Urbana, IL, USA
jwerdman@illinois.edu

Rohit Loomba

NAFLD Research Center
Division of Gastroenterology,
Department of Medicine
University of California, San
Diego
La Jolla, CA, USA
roloomba@ucsd.edu

Claude B. Sirlin

Department of Radiology
University of California, San
Diego
La Jolla, CA, USA
csirlin@ucsd.edu

William D. O'Brien, Jr.

Department of Electrical and
Computer engineering
University of Illinois at Urbana-
Champaign
Urbana, IL, USA
wdo@uiuc.edu

Abstract—Nonalcoholic fatty liver disease (NAFLD) affects ~25% of the world population. Confounder-corrected chemical-shift-encoded MRI-derived proton density fat fraction (MRI-PDF) is an established quantitative noninvasive biomarker of hepatic steatosis but has limited availability. There is a clinical need for more practical and accessible methods to noninvasively assess hepatic steatosis. Previous work has shown that two quantitative ultrasound (QUS) biomarkers - attenuation coefficient (AC) and backscatter coefficient (BSC) - are correlated with hepatic steatosis. Examining a broad range of QUS biomarkers, this study aims to develop an improved, multi-parametric QUS-based approach to diagnose NAFLD and quantify hepatic fat, with MRI-PDF as the reference standard. 102 participants recruited from the UCSD NAFLD Research Center underwent QUS exams on the right liver lobe with an Acuson S3000 ultrasound scanner and the 4C1 and 6C1HD transducers. Seven QUS biomarkers - AC, BSC, three Lizzi-Feleppa parameters (slope, intercept, midband), and two envelope parameters (k and μ) - were derived from ultrasound radiofrequency data. Two multivariable models were developed based on QUS biomarkers: a generalized linear regression model to predict hepatic PDF using stepwise regression for biomarker selection and a regularized logistic regression model to classify NAFLD (MRI-PDF > 5%, N=78/102) versus no NAFLD (MRI-PDF ≤ 5%) using LASSO regularization for biomarker selection. Leave-one-out cross-validation was performed. The final regression model selected the midband and k -parameter. The cross-validated predicted PDF values were correlated with the

reference MRI-PDF values (Spearman $\rho = 0.82$ and Pearson's $r = 0.76$). In comparison, Pearson's r was 0.59 between AC and MRI-PDF and 0.58 between BSC and MRI-PDF. The final classifier model selected the midband, k -parameter and μ -parameter, achieving an area under the receiver operating characteristic curve (AUROC) of 0.88. In comparison, AUROC was 0.83 using AC and 0.84 using BSC. The results suggest that multi-parametric QUS can improve the quantification of hepatic steatosis and diagnosis of NAFLD.

Keywords— quantitative ultrasound, nonalcoholic fatty liver disease, liver steatosis, proton density fat fraction

I. INTRODUCTION

Hepatic steatosis, defined as the accumulation of fat droplets within hepatocytes, is the earliest and hallmark histological feature of nonalcoholic fatty liver disease (NAFLD), which affects approximately 25% of the human population. Steatosis can lead to nonalcoholic steatohepatitis (NASH), a more rapidly progressive variant of NAFLD, which can contribute to the development of fibrosis, cirrhosis, and even hepatocellular carcinoma [1,2]. Disease progression may be halted or reversed in early stages [2].

There is a need to develop noninvasive, widely available, accurate, and cost-effective methods to diagnose NAFLD and quantify liver steatosis. Currently, liver biopsy is the gold standard for NAFLD diagnosis. Confounder-corrected chemical-shift-encoded MRI-derived proton density fat fraction

(PDFF) is the leading imaging biomarker for hepatic steatosis quantification [3]. Liver biopsy is invasive and MRI-PDFF is not widely accessible.

Quantitative ultrasound (QUS), a noninvasive, cost-effective, and easily accessible method, has potential for accurate NAFLD diagnosis and liver fat quantification. Previous human studies have shown that two QUS biomarkers - attenuation coefficient (AC, dB/cm-MHz) and backscatter coefficient (BSC, 1/cm-sr) - are strongly correlated with liver fat fraction [4, 5] and are repeatable and reproducible between different transducers, sonographers and scanner platforms [6-9].

Examining a broad range of QUS biomarkers for steatosis assessment, this study aims to develop multivariable QUS approaches to diagnose NAFLD and quantify hepatic fat, using contemporaneous MRI-PDFF as the reference.

II. METHODOLOGY

A. Human Study and Participants

This is an IRB-approved, HIPAA compliant study. Written informed consent was obtained. 102 research participants were prospectively recruited from the UCSD NAFLD Research Center. Inclusion criteria were age ≥ 18 years, known/suspected NAFLD, and willingness and ability to participate. Exclusion criteria were clinical, laboratory, or histology evidence of a liver disease other than NAFLD, excessive alcohol consumption (≥ 14 [men] or ≥ 7 [women] drinks/week), and steatogenic or hepatotoxic medication use.

All participants underwent QUS liver exams and contemporaneous MRI liver exams.

B. Ultrasound Data Acquisition

Ultrasound exams were performed using an Acuson S3000 (Siemens Healthineers, Munich, Germany) scanner. Each participant was scanned with a 4C1 (1-4 MHz nominal) transducer and/or a 6C1HD transducer (1-6 MHz nominal) by one or two of six registered diagnostic medical sonographers. Each participant underwent at least one but up to four same-day scanning sessions.

During each session, a sonographer made multiple repeated data acquisitions (separated by ~ 15 seconds) in the right liver lobe using a lateral intercostal approach, while participants held their breath in shallow expiration. Prior to the first data acquisition, system settings were adjusted for each participant to optimize right hepatic lobe visualization and to identify a region of the parenchyma without major vessels. System settings remained unchanged for the subsequent acquisitions. Each acquisition consisted of an operator button press that recorded a B-mode image, system settings, and the underlying radiofrequency (RF) data. Following completion of the liver acquisitions, a calibrated reference phantom (CIRS, Inc., Norfolk, VA) with known AC and BSC was scanned without changing the system settings.

C. QUS Biomarker Computation

Seven QUS biomarkers - AC, BSC, three Lizzi-Feleppa (LF) parameters (slope, intercept, midband) [10], and two envelope parameters (k and μ) - were computed from the acquired RF data. QUS biomarkers were computed offline using a custom

MATLAB (The MathWorks) graphic user interface tool that incorporates and standardizes the routines for QUS processing. Computations were performed in a field of interest (FOI) drawn by a trained image analyst within the margins of the liver boundary. Five acquisitions per participant were used for QUS computations. For each QUS biomarker, the five measurements were averaged to yield a single value. Algorithms for individual biomarkers were summarized as follows.

AC: AC was estimated using the spectral difference reference phantom method [11]. This method utilized the difference in the spectral amplitude at increasing depths to estimate local attenuation in the liver. Assuming that the liver within a small region of interest (denoted sub-ROI, much smaller than the FOI) is homogeneous and isotropic, the attenuation coefficient in dB/cm (denoted α ; AC is used to denote the attenuation coefficient in dB/cm-MHz) of the liver can be estimated at each frequency component by using

$$\alpha_s(f) = \alpha_r(f) - \frac{\gamma(f)}{4 \times 8.686}, \quad (1)$$

where f is the frequency in MHz, the subscripts s and r represent the unknown sample (i.e., liver) and the reference phantom, respectively, and $\gamma(f)$ is the slope of the straight line that fits the natural log ratio of sample power spectral density to the reference phantom power spectral density as a function of depth. The attenuation of the phantom was known *a priori*.

To implement this algorithm computationally, the FOI was subdivided into overlapping, rectangular sub-ROIs, each of which yielded an estimate of $\alpha_s(f)$. Each individual sub-ROI was subdivided into overlapping axial sections to calculate the power spectral density at different depths through the sub-ROI, as required by the spectral difference method. The size of the sub-ROI for $\alpha_s(f)$ estimation was 12 mm \times 40 A-lines (axial \times lateral), and the axial length of the overlapping sections was 4 mm. These dimensions yield sub-ROIs that are about 20 pulse lengths axially, as well as a section length of about 7 pulse lengths. The sub-ROIs were set to overlap by 50% in the axial and lateral directions.

The power spectral density at each depth was estimated separately for the sample and the reference. Afterwards, $\alpha_s(f)$ was estimated using Eq. (1) for the sub-ROI. The $\alpha_s(f)$ estimates from all of sub-ROIs were averaged to obtain the mean $\alpha_s(f)$ versus frequency. The mean $\alpha_s(f)$ curve of the liver was fit to the power law form to provide an α value for an arbitrary frequency for attenuation compensation during BSC estimation.

The AC values at different frequency points between 2.3-3.1 MHz were averaged to yield a single AC value for an FOI.

BSC: The BSC estimates were obtained using the reference phantom method [11] to account for system and settings effects. The BSC of the liver was estimated from

$$BSC_s(z, f) = \frac{S_s(z, f)}{S_r(z, f)} BSC_r(z, f) 10^{2z[\alpha_s(f) - \alpha_r(f)]/10}, \quad (2)$$

where BSC_s and BSC_r are the BSC values of the liver and the reference phantom, respectively; S_s and S_r were the power spectra of the liver and the reference phantom, respectively; and z is the depth. The term $10^{2z[\alpha_s(f) - \alpha_r(f)]/10}$ was used to

compensate for attenuation effects; note that α_s and α_r are in dB/cm for this form of compensation. The assumptions for Eq. (2) were that the transducer surface was in contact with the human skin and the reference phantom during scanning, and that α was homogenous in the liver and the reference phantom.

To implement the BSC algorithm, the FOI was re-divided into 75%-overlapped sub-ROIs with dimensions $8.6 \text{ mm} \times 20 \text{ A-lines}$ (axial \times lateral; axial size equivalent to 15 wavelengths at 2.7 MHz). The power spectral density of each sub-ROI was estimated separately for the liver and the phantom, and the BSC of the sub-ROI were then estimated using Eq. (2). BSC estimates from all the sub-ROIs were averaged to obtain the mean BSC versus frequency within an FOI.

The BSC values at different frequency points between 2.3-3.1 MHz were averaged to yield a single BSC value for the FOI.

LF parameters: LF slope, intercept, and midband were computed through linear regression of $10\log_{10}(BSC)$ against frequency. LF slope, intercept, and midband were the linear regression slope, intercept, and regression value at the center frequency (2.7 MHz), respectively.

Envelope statistics parameters k and μ : The k and μ parameters were estimated using the method described in [12]. Briefly, the distribution of envelope amplitude of the RF data was modeled by the homodyned K distribution with the following probability density function [12]:

$$p_A(A) = A \int_0^\infty x J_0(sx) J_0(Ax) \left(1 + \frac{x^2 \sigma^2}{2\mu}\right)^{-\mu} dx, \quad (3)$$

where A is the envelope amplitude, $J_0()$ is the zeroth order Bessel function of the first kind, s^2 is the coherent signal energy, σ^2 is the diffuse signal energy, and μ is a measure of the effective number of scatterers per resolution cell. The parameter k was the ratio of the coherent to diffuse signal (s/σ) and represents the level of structure or periodicity in scatterer locations.

To estimate k and μ from the raw RF data, the FOI was divided into sub-ROIs in the same way as for BSC computation. For each sub-ROI, the envelope of RF signals was calculated. Then the signal-to-noise ratio, skewness, and kurtosis of the envelope was computed from the envelop distribution and compared with theoretical values predicted by the homodyned K distribution to yield k and μ parameter estimates as described in [12]. The k and μ parameters estimated in all sub-ROIs were averaged to yield a single k and a single μ estimate for each FOI.

D. Multivariable QUS Models

We developed two QUS-based multivariable models: **a)** a classifier to differentiate participants with (PDFF>5%) versus without (PDFF≤5%) NAFLD, and **b)** a fat fraction estimator to predict PDFF. The PDFF value for each participant was obtained by averaging PDFF values from liver segments 5-8 (right lobe). Leave-one-out cross-validation was performed to avoid overestimating model performance. Feature selection and model training were repeated for each iteration.

The classifier was built using logistic regression with lasso regularization for feature selection. The cross-validation procedure was performed as follows. Model training and testing

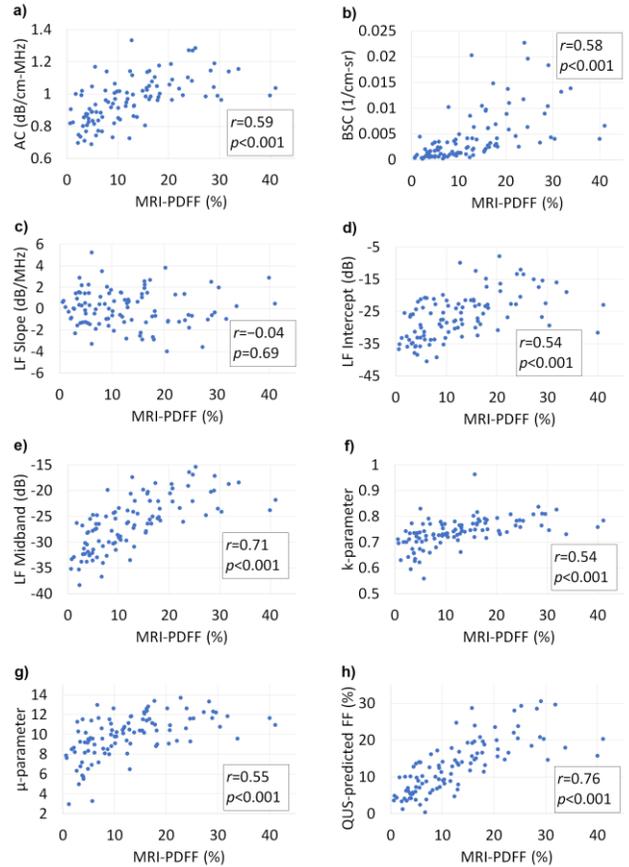


Fig. 1. (a-g) QUS biomarkers versus PDFF scatter plots and (h) multivariable fat fraction estimator predicted fat fraction (FF, leave-one-out cross-validated) versus PDFF scatter plot (h). Pearson's correlation coefficient and the associated p -value are shown for each sub-figure.

was repeated 102 times (folds), with one of the 102 participants taking turn to serve as the test sample while the remaining 101 participants serving as the training data. For instance, to predict whether Participant 1 had NAFLD, a model was built using training data from Participants 2 to 102. A logistic regression model with lasso regularization was fit to the training data (Participants 2 to 102). This fitting process was performed by calling the MATLAB function *lassoglm*, which yielded a trained classifier that took seven QUS biomarkers as the input but only used a subset of the biomarkers for prediction. The trained classifier was then applied to Participant 1 (not used for the training) for testing. This process was repeated by using each participant for testing in turn. The testing results from the 102 folds were finally pulled together for classification evaluation.

The fat fraction estimator was built using a generalized linear model with stepwise regression. The regression was performed by calling the MATLAB function *stepwiseglm*. The cross-validation procedure was performed in a similar fashion as done for the classifier.

III. RESULTS AND DISCUSSION

A. Participant Characteristics

102 participants were included. The mean (\pm standard deviation [SD]) of age was 48.5 (\pm 12.6) years for men, and 55.0

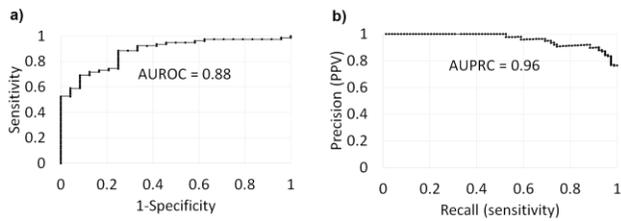


Fig. 2. Leave-one-out cross-validated a) receiver operating characteristic curve and b) precision-recall curve for diagnosing NAFLD using the multivariable classifier. AUROC = area under the receiver operating characteristic curve, AUPRC = area under the precision recall curve, PPV = positive predictive value.

(± 12.8) years for women. The mean (\pm SD) of body mass index was $31.0 (\pm 5.0)$ kg/m² for men and $31.1 (\pm 5.2)$ kg/m² for women. The mean (\pm SD) of PDFF was $11.1\% (\pm 7.9\%)$ for men and $14.5\% (\pm 9.7\%)$ for women. The PDFF ranged from 0.7 to 41.1%, with 78 of 102 participants (76.5%) having NAFLD as defined by MRI-PDFF $>5\%$. Average time duration between MRI and QUS exams was 3 days (range: 0-67 days).

B. Correlation between QUS and MRI-PDFF

Six out of the seven QUS parameters were significantly correlated with MRI-PDFF (Fig. 1a-g), with Pearson's correlation coefficients ranging from 0.54 to 0.71. The LF slope was not significantly correlated with MRI-PDFF.

C. Multivariable Fat Fraction Estimator Performance

The LF midband and k parameters were selected for the multivariable fat fraction estimator in all 102 folds. The resulting fat fraction estimator was in the form of estimated fat fraction = $a \cdot \text{midband} + b \cdot k + c \cdot \text{midband} \cdot k + d$, where a, b, c, and d are coefficients. Fat fraction values predicted by this multivariable fat fraction estimator through leave-one-out cross validation were significantly correlated with PDFF (Fig. 1g), with Spearman $\rho=0.82$ ($p < 0.001$) and Pearson's $r=0.76$ ($p < 0.001$). The Pearson's r was higher than that between any single QUS biomarker and PDFF.

There was no significant nonlinearity between predicted fat fraction and PDFF for PDFF $\leq 34\%$ following a linearity test [13]. The multivariable fat fraction estimator tended to underestimate the fat fraction for PDFF $>34\%$, suggesting a saturation effect outside the linear range, although the number of participants $> 34\%$ was very small. The cause of the saturation effect is not yet well understood and will be the subject of future studies.

D. Multivariable Classifier Performance

The LF midband, k, and μ parameters were selected for the multivariable classifier in all 102 folds. Leave-one-out cross validation of the multivariable classifier yielded an area under the receiver operating characteristic curve (AUROC) of 0.88 (95% CI: 0.82-0.94) for diagnosing NALFD and an area under the precision-recall curve (AUPRC) of 0.96. The ROC and PRC curves are shown in Fig. 2. In comparison, AUROC was 0.83 using AC and 0.84 using BSC.

IV. CONCLUSION

Multi-parametric QUS can improve the quantification of hepatic steatosis and the diagnosis of NAFLD compared with using only AC and BSC.

ACKNOWLEDGMENT

The authors thank the sonographers, Lara Callahan, Lisa Deiranieh, Elise Housman, Christopher Lucas, Susan Lynch, and Minaxi Trivedi, for their dedicated contributions and expertise.

REFERENCES

- [1] R. Loomba and A. J. Sanyal, "The global NAFLD epidemic," *Nat. Rev. Gastroenterol. Hepatol.*, vol. 10, no. 11, pp. 686–690, 2013.
- [2] S. L. Friedman, B. A. Neuschwander-Tetri, M. Rinella, and A. J. Sanyal, "Mechanisms of NAFLD development and therapeutic strategies," *Nat. Med.*, vol. 24, no. 7, pp. 908–922, 2018.
- [3] S. B. Reeder, I. Cruite, G. Hamilton, and C. B. Sirlin, "Quantitative assessment of liver fat with magnetic resonance imaging and spectroscopy," *J. Magn. Reson. Imaging*, vol. 34, no. 4, pp. 729–749, 2011.
- [4] S. C. Lin, E. Heba, T. Wolfson, B. Ang, A. Gamst, A. Han, J. W. Erdman, Jr., W. D. O'Brien, Jr., M. P. Andre, C. B. Sirlin, and R. Loomba, "Noninvasive diagnosis of nonalcoholic fatty liver disease and quantification of liver fat using a new quantitative ultrasound technique," *Clin. Gastroenterol. Hepatol.*, vol. 13, no. 7, pp. 1337–1345, 2015.
- [5] M. P. Andre, A. Han, E. Heba, J. Hooker, R. Loomba, C. B. Sirlin, J. W. Erdman, Jr., and W. D. O'Brien, Jr., "Accurate diagnosis of non-alcoholic fatty liver disease in humans via quantitative ultrasound," in 2014 IEEE Int. Ultrason. Symp. (IUS), 2014, pp. 2375–2377.
- [6] A. Han, M. P. Andre, J. W. Erdman, Jr., R. Loomba, C. B. Sirlin, and W. D. O'Brien, Jr., "Repeatability and reproducibility of a clinically based QUS phantom study and methodologies," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 64, no. 1, pp. 218–231, 2017.
- [7] A. Han, M. P. Andre, L. Deiranieh, E. Housman, J. W. Erdman, Jr., R. Loomba, C. B. Sirlin, and W. D. O'Brien, Jr., "Repeatability and reproducibility of the ultrasonic attenuation coefficient and backscatter coefficient measured in the right lobe of the liver in adults with known or suspected nonalcoholic fatty liver disease," *J. Ultrasound Med.*, vol. 37, no. 8, pp. 1913–1927, 2018.
- [8] A. Han, Y. Labyed, E. Z. Sy, A. S. Boehringer, M. P. Andre, J. W. Erdman, Jr., R. Loomba, C. B. Sirlin, and W. D. O'Brien, Jr., "Inter-sonographer reproducibility of quantitative ultrasound outcomes and shear wave speed measured in the right lobe of the liver in adults with known or suspected non-alcoholic fatty liver disease," *Eur. Radiol.*, vol. 28, no. 12, pp. 4992–5000, 2018.
- [9] A. Han, Y. N. Zhang, A. S. Boehringer, M. P. Andre, J. W. Erdman, Jr., R. Loomba, C. B. Sirlin, and W. D. O'Brien, Jr., "Inter-platform reproducibility of ultrasonic attenuation and backscatter coefficients in assessing NAFLD," *Eur. Radiol.*, vol. 29, no. 9, pp. 4699–4708, 2019.
- [10] F. Lizzi, M. Greenebaum, E. J. Feleppa, and M. Elbaum, "Theoretical framework for spectrum analysis in ultrasonic tissue characterization," *J. Acoust. Soc. Am.*, vol. 73, no. 4, pp. 1366–1373, 1983.
- [11] L. X. Yao, J. A. Zagzebski, and E. L. Madsen, "Backscatter coefficient measurements using a reference phantom to extract depth-dependent instrumentation factors," *Ultrason. Imaging*, vol. 12, no. 1, pp. 58–70, 1990.
- [12] D. P. Hruska and M. L. Oelze, "Improved parameter estimates based on the homodyned K distribution," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 56, no. 11, pp. 2471–2481, 2009.
- [13] D. L. Raunig, et al. "Quantitative imaging biomarkers: A review of statistical methods for technical performance assessment," *Stat. Methods Med. Res.*, vol. 24, no. 1, pp. 27–67, 2015.