



Quantitative ultrasound and machine learning for assessment of steatohepatitis in a rat model

An Tang^{1,2,3,4} · François Destrempe⁵ · Siavash Kazemirad⁶ · Julian Garcia-Duitama^{4,5} · Bich N. Nguyen^{7,8} · Guy Cloutier^{2,4,5}

Received: 31 July 2018 / Revised: 29 October 2018 / Accepted: 23 November 2018 / Published online: 17 December 2018
© European Society of Radiology 2018

Abstract

Objectives To develop a machine learning model based on quantitative ultrasound (QUS) parameters to improve classification of steatohepatitis with shear wave elastography in rats by using histopathology scoring as the reference standard.

Methods This study received approval from the institutional animal care committee. Sixty male Sprague-Dawley rats were either fed a standard chow or a methionine- and choline-deficient diet. Ultrasound-based radiofrequency images were recorded in vivo to generate QUS and elastography maps. Random forests classification models and a bootstrap method were used to identify the QUS parameters that improved the classification accuracy of elastography. Receiver-operating characteristic analyses were performed.

Results For classification of not steatohepatitis vs borderline or steatohepatitis, the area under the receiver-operating characteristic curve (AUC) increased from 0.63 for elastography alone to 0.72 for a model that combined elastography and QUS techniques ($p < 0.001$). For detection of liver steatosis grades 0 vs ≥ 1 , ≤ 1 vs ≥ 2 , ≤ 2 vs 3, respectively, the AUCs increased from 0.70, 0.65, and 0.69 to 0.78, 0.78, and 0.75 ($p < 0.001$). For detection of liver inflammation grades 0 vs ≥ 1 , ≤ 1 vs ≥ 2 , ≤ 2 vs 3, respectively, the AUCs increased from 0.58, 0.77, and 0.78 to 0.66, 0.84, and 0.87 ($p < 0.001$). For staging of liver fibrosis grades 0 vs ≥ 1 , ≤ 1 vs ≥ 2 , and ≤ 2 vs ≥ 3 , respectively, the AUCs increased from 0.79, 0.92, and 0.91 to 0.85, 0.98, and 0.97 ($p < 0.001$).

Conclusion QUS parameters improved the classification accuracy of steatohepatitis, liver steatosis, inflammation, and fibrosis compared to shear wave elastography alone.

Key Points

- Quantitative ultrasound and shear wave elastography improved classification accuracy of liver steatohepatitis and its histological features (liver steatosis, inflammation, and fibrosis) compared to elastography alone.
- A machine learning approach based on random forest models and incorporating local attenuation and homodyned-K tissue modeling shows promise for classification of nonalcoholic steatohepatitis.
- Further research should be performed to demonstrate the applicability of this multi-parametric QUS approach in a human cohort and to validate the combinations of parameters providing the highest classification accuracy.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00330-018-5915-z>) contains supplementary material, which is available to authorized users.

✉ An Tang
an.tang@umontreal.ca

¹ Department of Radiology, Centre hospitalier de l'Université de Montréal (CHUM), Montréal, Québec, Canada

² Department of Radiology, Radio-oncology and Nuclear Medicine, Université de Montréal and CRCHUM, 1000 rue Saint-Denis, Montréal, Québec H2X 0C2, Canada

³ Laboratory of Medical Image Analysis, Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM), Montréal, Québec, Canada

⁴ Université de Montréal, Institute of Biomedical Engineering, Montréal, Québec, Canada

⁵ Laboratory of Biorheology and Medical Ultrasonics (LBUM), Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM), Montréal, Québec, Canada

⁶ School of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran

⁷ Department of Pathology, Centre hospitalier de l'Université de Montréal (CHUM), Montréal, Québec, Canada

⁸ Department of Pathology and Cellular Biology, Université de Montréal, Montréal, Québec, Canada

Keywords Nonalcoholic steatohepatitis · Non-alcoholic fatty liver disease · Machine learning · Ultrasonography · Elasticity imaging techniques

Abbreviations

AUC	Area under the receiver-operating characteristic curve
CAP	Controlled attenuation parameter
HKD	Homodyned-K distributions
HPS	Hematoxylin phloxine saffron
IQR	Inter-quartile range
MCD	Methionine and choline deficient
NAFLD	Nonalcoholic fatty liver disease
NASH	Nonalcoholic steatohepatitis
QUS	Quantitative ultrasound
ROC	Receiver operating characteristic
ROI	Region of interest
SH	Steatohepatitis

Introduction

Nonalcoholic fatty liver disease (NAFLD) is a highly prevalent condition, found in 20 to 30% of adults in Western nations [1], which may progress to nonalcoholic steatohepatitis (NASH), an advanced form found in 3 to 5% of the general population [1]. Liver biopsy is commonly used as the reference standard to grade steatosis and inflammation, two distinctive features of steatohepatitis, and to stage fibrosis, a marker of liver disease severity [2, 3]. However, liver biopsy is invasive, associated with pain, sampling errors, and risk of bleeding [4, 5]. Hence, there is a need for a noninvasive technique for the assessment of liver steatohepatitis.

Magnetic resonance-based techniques are accurate and precise for quantification of liver fat with proton density fat fraction [6–8] and liver fibrosis with elastography [9, 10]. Recent research also suggests that advanced magnetic resonance elastography may potentially discriminate hepatic inflammation and fibrosis in the early stages of chronic liver disease [11, 12]. Despite its strengths, magnetic resonance imaging may not be practical or cost-effective for clinical screening considering the high prevalence of NAFLD [13]. Recently, quantitative ultrasound (QUS) techniques that measure controlled attenuation parameter (CAP) [14–17], local attenuation coefficient, or backscatter coefficient have been proposed for the detection or grading of liver steatosis using either magnetic resonance imaging proton density fat fraction or histology as the reference standard [18, 19]. These QUS techniques may be implemented on ultrasound scanners with an elastographic capability that can currently achieve good accuracy for noninvasive staging of liver fibrosis [20, 21]. Hence, there is a high interest in developing an ultrasound-based approach for

comprehensive assessment of liver inflammation, steatosis, and fibrosis within the same examination.

A potential strategy to characterize the livers is by analyzing the interaction of sound waves with insonified tissues to reveal microstructure properties accessible through the analysis of backscatter radiofrequency echoes [22]. Various QUS backscatter approaches for determining tissue microstructures from radiofrequency echoes have received broad interest: in particular, fitting the spectrum of radiofrequency signals to an estimated spectrum by an appropriate scattering model (spectral approach) [22], and computing 1st order statistics of radiofrequency signal echo envelope (statistical approach) [23]. We hypothesized that a machine learning model based on QUS parameters could help detect steatohepatitis by providing a cellular signature (i.e., size, density, spatial organization, and acoustic properties) of liver tissues with various histological features along the NAFLD to NASH disease continuum [2]. A recent study found that liver shear stiffness measured with ultrasound elastography shows promise as a biomarker for noninvasive diagnosis of steatohepatitis [24].

The aim of this study was to develop a machine learning model based on QUS parameters that can be used to improve the classification of steatohepatitis over shear wave elastography in rats by using histopathology scoring as the reference standard. The secondary purpose was to identify QUS parameters that provide the highest classification accuracy for steatosis, inflammation, and fibrosis.

Materials and methods

Study design and animals

This study received approval from the Institutional Animal Care Committee at the University of Montreal Hospital Research Centre. Special care was taken to follow the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines for the replacement, refinement, and reduction of animals in this study. This is an ancillary study to an experimental animal model of NASH described in detail previously [24]. To obtain a range of disease severity, we included 48 11-week-old (at the start of the study) male Sprague-Dawley rats (Charles River Canada, Saint-Constant, Quebec, Canada) fed with a methionine- and choline-deficient (MCD) diet (Dyets 518753; Research Diets) ad libitum to develop steatosis, necroinflammation, and fibrosis similar to that found in human NASH [25, 26]. These rats were divided into four groups of 12 rats each that were imaged and subsequently euthanized at 1, 4, 8, and 12 weeks, respectively.

Twelve 11-week-old healthy control rats with free access to standard food and water served as controls and were imaged and sacrificed 4 weeks after the start of the study.

The liver of animals in the experimental and control cohorts was evaluated *in vivo* by ultrasound elastography and QUS prior to euthanasia. After euthanasia, livers were explanted for *ex vivo* assessment by histopathology.

US Elastography

Shear wave ultrasound elastography measurements were performed using a research ultrasound system (model V1, Verasonics Inc.), while the rats were under anesthesia and in the supine position on an electronic heating pad (MMHP01FFC; Mansfield Medical). To generate each shear wavefront within the liver, a linear array ultrasound transducer (ATL L7-4, Philips) was used to induce three 40-V 125- μ s long radiation force pushes 4 mm apart [27]. Three reference images for the displacement field estimation were acquired before the shear wave generation. The same transducer was used to acquire plane wave radiofrequency data at a frame rate of 4 kHz [28]; image frames were reconstructed using the f-k migration algorithm [29]. Ten shear wave generation and acquisition sequences were performed for each animal. The sequence with the highest signal-to-noise ratio was chosen for post-processing.

A conventional one-dimensional normalized cross-correlation algorithm was used to obtain shear wave displacement fields within the liver from acquired radiofrequency data [30]. Shear wave (phase) velocities were estimated from the displacement fields in the wavenumber-frequency domain [31]. The regions of interest (ROIs) for all animals were selected at an approximate depth of 15 mm from the ultrasound transducer surface and with a length between 6 to 8 mm on the median lobe of the liver. Liver storage modulus (G') was estimated and averaged over the frequency range between 130 and 220 Hz, which was previously shown to provide a better separation between SH categories compared with lower frequencies [24].

Quantitative ultrasound

Similar to ultrasound elastography acquisitions, rats were under anesthesia and in a supine position on the heating pad. QUS acquisitions were performed before the elastography exam using the same Verasonics system and an ultrasound transducer, with a bandwidth of 4 to 7 MHz and a center frequency of 5 MHz. One hundred radiofrequency frames were acquired at 4 kHz, migrated [29], and Hilbert transformed to obtain the echo envelope. Received echoes were compensated for time gain attenuation, whose values were recorded during acquisition. Contours of a ROI within the median lobe or left lateral lobe of the liver were manually delineated by a fellowship-

trained radiologist on the first frame of each B-mode sequence. These contours were then propagated automatically using a segmentation algorithm [32] to consider similar ROIs on consecutive frames.

Ultrasound echo envelope statistics were modeled with homodyned-K distributions (HKD). A sliding window of 82×10 pixels², corresponding to 3 mm in both axial and lateral directions, was swept across the ROI by steps of 4×1 pixels², and estimation of HKD parameters was performed within the sliding window as in Destempes F et al [33, 34]. Thus, at each pixel of the ROI corresponded a local HKD estimation, from which could be computed local values of the mean intensity normalized by its maximal value μ_n , the reciprocal $1/\alpha$ of the scatterer-clustering parameter α , the coherent-to-diffuse signal ratio k , as well as the diffuse-to-total signal power ratio $1/(\kappa + 1)$ [34]. The mean intensity μ_n is akin to B-mode echogenicity. A decrease of $1/\alpha$ corresponds to a greater homogeneity of the scattering medium. An increase in k or κ reflects the periodic alignment of scatterers, the presence of specular reflection, or highly structured spatial organization of scatterers [34].

This image post-processing pipeline yielded four HKD parametric maps of the ROI: μ_n , $1/\alpha$, k , and $1/(\kappa + 1)$ (Supplementary Table 1). The mean and interquartile range (IQR) of each parameter were then computed on each frame, and median values over all frames were output, thus yielding eight HKD features. The local attenuation coefficient (dB/MHz/cm) within the ROI was also computed using the spectral shift algorithm [35]. Since no diffraction occurs within the image plane in images that are reconstructed from plane-wave images [36], correction for diffraction was not applied in this application of the spectral shift method. The attenuation slope was estimated with a robust fit method rather than linear regressions (c.f. eq. (4.13) in Bigelow et al [35]). The median of attenuation coefficients estimated over all the frames was considered as the last QUS feature.

Histopathology analysis

The liver specimens were stained with hematoxylin phloxine saffron, trichrome, reticulin, Sirius red, and α -smooth muscle actin stains. Histology slides of liver specimens were reviewed by an hepatopathologist. Blinded to the animal cohort, scores of the steatosis grade from 0 to 3, lobular inflammation grade from 0 to 3, hepatocellular ballooning grade from 0 to 2, and fibrosis stage from 0 to 4 were obtained according to the NASH Clinical Research Network histological scoring system [2] and used as the reference standard for *in vivo* quantitative ultrasound data. Steatohepatitis (SH) was categorized as: “not SH,” “borderline,” and “SH,” based on the NAFLD Activity Score, which is the unweighted sum of steatosis, inflammation, and ballooning grades [2]. For this study, SH was further subdivided into “SH with fibrosis stage 1 or lower” and “SH

with fibrosis stage 2 or higher” to account for the presence of fibrosis. A detailed description of the histopathological scoring system is provided in Supplementary Table 2.

Data analysis

Machine learning model A random forest classifier was used as a statistical learning model [37]. With this method, a classification is performed independently by multiple decision trees, based on input features, and the most frequent class is assigned as output decision. Feature selection was performed with random forests of 3000 trees each, on all combinations of at most three features among all chosen features. The G-mean (square root of the product of sensitivity with specificity) was computed for each tested combination, and the ten combinations with the highest G-means were selected. The G-mean is recommended in the case of imbalanced data (i.e., in which the two class groups vary much in size) [38].

Statistical analysis

Classification of NASH For each selected combination of features, the receiver-operating characteristic (ROC) curve was constructed by generating stratified samples with a proportion of the samples in one class varying from 1/40 to 39/40 by steps of 1/40. For each specified proportion, false positive and negative rates were computed according to the 0.632 + bootstrap method [39]. To avoid over-fitting, trees were restricted to a maximum number of terminal nodes, ranging from 2 to 20 by steps of 2. The combination of features with the highest AUC was then selected as the best combination, for a given classification task. The 95% confidence interval (CI) was computed based on percentiles of a sample of AUCs that was constructed with the jackknife method [40]. Jackknife samples of AUCs were also used to perform one-sided-paired Wilcoxon-signed rank tests to compare the best combination of QUS features with the elastography parameter alone.

Classification of steatosis grades, inflammation grades, and fibrosis stages The same approach was used to identify a combination of QUS parameters that improved the classification accuracy for steatosis, inflammation, and fibrosis. On each ROC curve, the point with maximal Youden’s index was computed and the corresponding sensitivity and specificity were reported. In the case of imbalanced classes, the area under the precision-recall curve (PRC-AUC) [41] was also estimated, as a better metric for comparing the performance of features’ combinations.

Correlations between steatosis grades, inflammation grades, and fibrosis stages Since it is expected that the categorical variables steatosis, inflammation, and fibrosis present some degree of association, pairwise Pearson’s chi-square test was

performed, with Holm-Bonferroni adjustment on p values, for contingency tables as a test of association between pairs of the categorical variables steatosis, inflammation, and fibrosis.

Statistical analysis was performed using R statistical software (R Foundations), with p values estimated by Monte Carlo simulations. A $p < 0.05$ was considered as significant.

Results

Effects of MCD model on histopathological findings

Compared with Sprague-Dawley control rats fed a standard chow, rats fed a MCD diet for 1, 4, 8, or 12 weeks developed steatosis, inflammation, and fibrosis. Representative stains for each cohort are shown in Fig. 1. Further histological analysis staining revealed a rapid increase in the grade of macrovesicular steatosis as early as 1 week after exposure to the MCD diet, reaching severe (grade 3) steatosis at 4 weeks and remaining at that level with a longer duration of exposure to MCD diet (Fig. 2a). Moreover, the level of inflammation increased after 1 week of exposure to the MCD diet, reaching a peak at 8 weeks and decreasing afterward (Fig. 2b). Fibrosis gradually increased after the 4-week MCD diet and developed as mild to severe fibrosis (stages 1 to 4) in the 12-week MCD diet cohort (Fig. 2c). All of the control rats were classified as not SH, whereas the classification was borderline for most of the rats (ten out of 12) of the 1-week MCD diet cohort and SH for the rest of MCD diet cohorts (except for two rats) (Fig. 2d).

Machine learning model

The storage modulus could not be estimated for one of the animals in the experimental group fed with a MCD diet for 4 weeks, as all of the ten US elastography sequences revealed a very low signal-to-noise ratio due to significant diffraction of the shear wave by the layered structure of rat liver. Therefore, analysis was performed on 59 rats (rather than 60). Table 1 provides detailed information about the combinations of G' and QUS parameters that provided the highest classification accuracy for the diagnosis of NASH and for the identification of its histopathological components (steatosis, inflammation, and fibrosis). The addition of QUS features improved all the classification tasks ($p < 0.001$). The highest improvements were for the classification of steatosis grades ≤ 1 vs. ≥ 2 (13% in AUC), and for the characterization of not steatohepatitis vs. borderline or steatohepatitis (9% in AUC). The models providing the best accuracy often included the shear elasticity modulus. B-mode images with overlaid color-coded parametric maps of the scatterer-clustering parameter, coherent-to-diffuse signal ratio, and diffuse-to-total signal power ratio are shown for five animal representatives of their respective cohorts (Fig. 3).

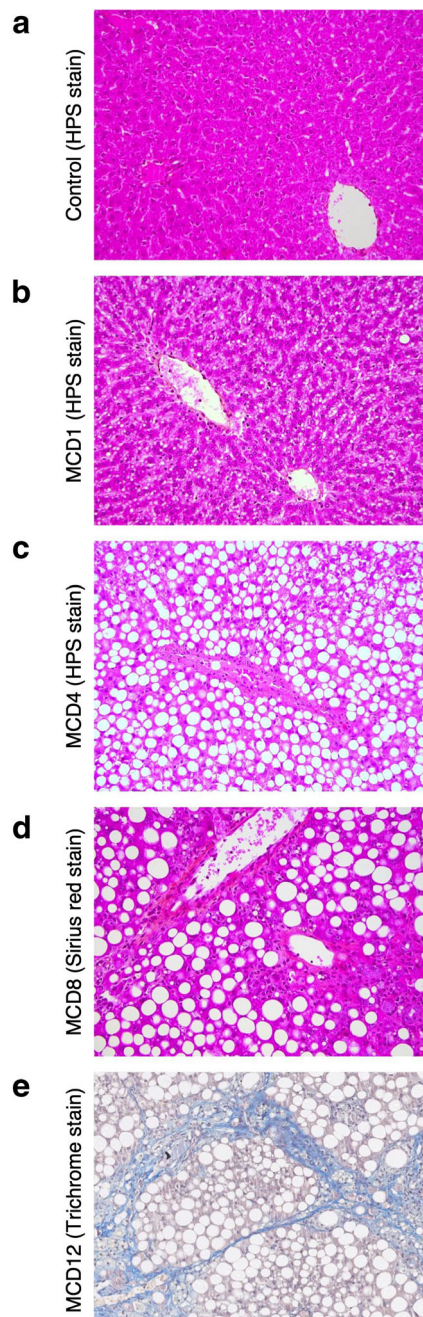


Fig. 1 Histopathology and representative stains at $\times 20$ magnification of animals from the five groups of animals tested in this study. **a** Control group fed a standard chow (hematoxylin phloxine saffron [HPS] stain). **b** Experimental group fed a methionine and choline deficient (MCD) diet for 1 week (HPS stain). **c** Four weeks (HPS stain). **d** Eight weeks (Sirius red stain). **e** Twelve weeks (trichrome stain). Steatosis can be seen after 1 week, inflammation at 8 weeks, and fibrosis at 12 weeks

Classification accuracy of NASH

Depending on the dichotomized NASH categories, AUCs were 0.63–0.92 for G' only and 0.72–0.98 for combinations of parameters (Table 1). Specificity and sensitivity corresponding to the points on ROC curves that maximize

Youden's index for each dichotomization are provided in Fig. 4a.

Classification of steatosis grades, inflammation grades, and fibrosis stages

Table 1 also provides AUC and 95% confidence intervals for G' only and the combinations of QUS parameters that provided the highest classification accuracy for dichotomized steatosis grades, inflammation grades, and fibrosis stages. For detection of liver steatosis grades 0 vs. ≥ 1 , ≤ 1 vs. ≥ 2 , ≤ 2 vs. 3, AUCs were respectively 0.70, 0.65, and 0.69 for elastography alone and 0.78, 0.78, and 0.75 for a model that combined QUS features and elastography for the latter category. For detection of liver inflammation grades 0 vs. ≥ 1 , ≤ 1 vs. ≥ 2 , ≤ 2 vs. 3, AUCs were respectively 0.58, 0.77, and 0.78 for elastography alone and 0.66, 0.84, and 0.87 for a model that combined elastography and QUS techniques. For staging of liver fibrosis grades 0 vs. ≥ 1 , ≤ 1 vs. ≥ 2 , and ≤ 2 vs. ≥ 3 , AUCs were improved from 0.79, 0.92, and 0.91 for elastography alone to 0.85, 0.98, and 0.97 for a model that combined elastography and QUS techniques. Optimal thresholds that maximize Youden's index for each steatosis, inflammation, and fibrosis dichotomizations are provided in Fig 4b–d, respectively.

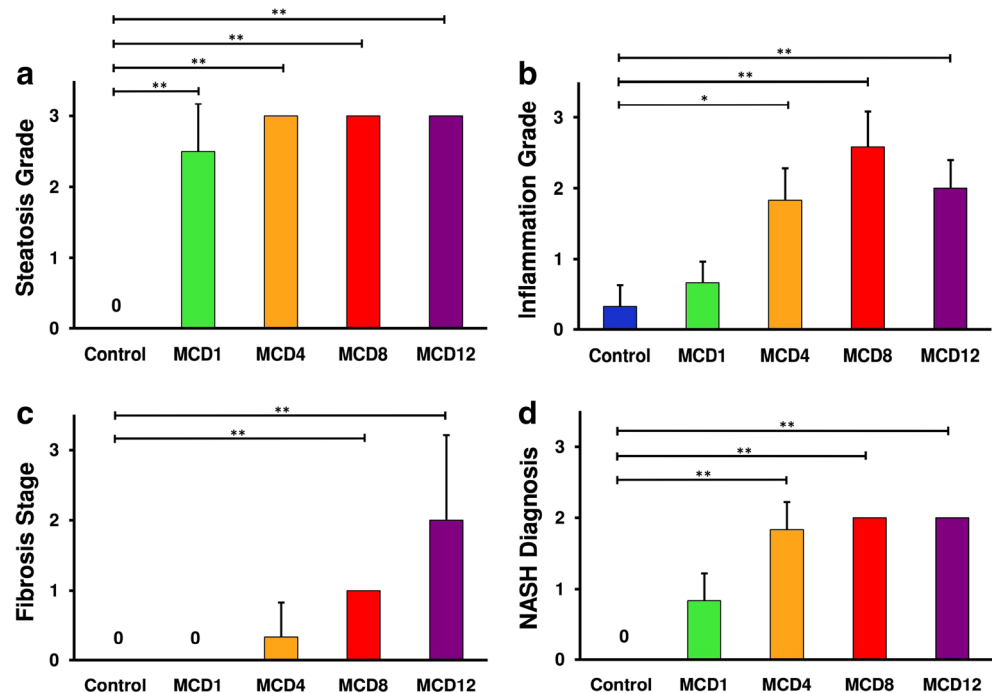
Correlations between steatosis grades, inflammation grades, and fibrosis stages

There was an association between steatosis and inflammation ($p = 0.0015$), and between inflammation and fibrosis ($p = 0.0015$), but not between steatosis and fibrosis ($p = 0.10$).

Discussion

The key findings of this study indicate that a machine learning approach adding QUS parameters (μ_n , $1/\alpha$, k , $1/(\kappa + 1)$, and local attenuation) to elastography can significantly improve the diagnosis of steatohepatitis and the classification of steatosis, inflammation, and fibrosis in an animal model of NASH. Historically, ultrasound imaging has been used to diagnose disease based on B-mode to classify tissue structures and Doppler-mode to assess vascularity. In recent years, ultrasound elastography has been very successful to image liver elasticity as a marker of liver fibrosis [42, 43]. However, steatosis, inflammation, and fibrosis may coexist in the NASH continuum and confound liver stiffness. Inflammation, through edema and the presence of inflammatory cells, may increase the internal pressure and the liver stiffness [44–47], whereas liver fat may decrease liver stiffness [24, 48]. Another strategy to

Fig. 2 Plots showing histological grading and staging. **a** Steatosis grade for the control and the MCD diet cohorts. **b** Inflammation grade for the control and the MCD diet cohorts. **c** Fibrosis stage for the control and the MCD diet cohorts. **d** NASH diagnosis category for the control and the MCD diet cohorts (category 0 = no steatohepatitis, category 1 = borderline, category 2 = steatohepatitis with fibrosis stage 1 or lower, and category 3 = steatohepatitis with fibrosis stage 2 or higher). Comparisons were limited to experimental cohorts with the control cohort (* $p < 0.01$; ** $p < 0.001$). MCD1 = 1-week MCD diet cohort, MCD4 = 4-week MCD diet cohort, MCD8 = 8-week MCD diet cohort, MCD12 = 12-week MCD diet cohort



characterize the liver tissue is to analyze its interaction with sound waves to reveal properties accessible through the analysis of backscatter radiofrequency echoes [22]. QUS imaging analyzes sub-resolution echoes produced by constructive and destructive interferences for the

purpose of characterizing the tissue microstructure. Recently, statistical machine learning approaches such as random forest classifiers [37] have been proposed to identify elastography and QUS features providing the highest classification accuracy for a given dataset [49, 50].

Table 1 Diagnostic accuracy of storage modulus (G') alone and in combination with quantitative ultrasound (QUS) parameters for diagnosis of steatohepatitis (SH) and for classification of steatosis, inflammation, and fibrosis in an animal model of NASH

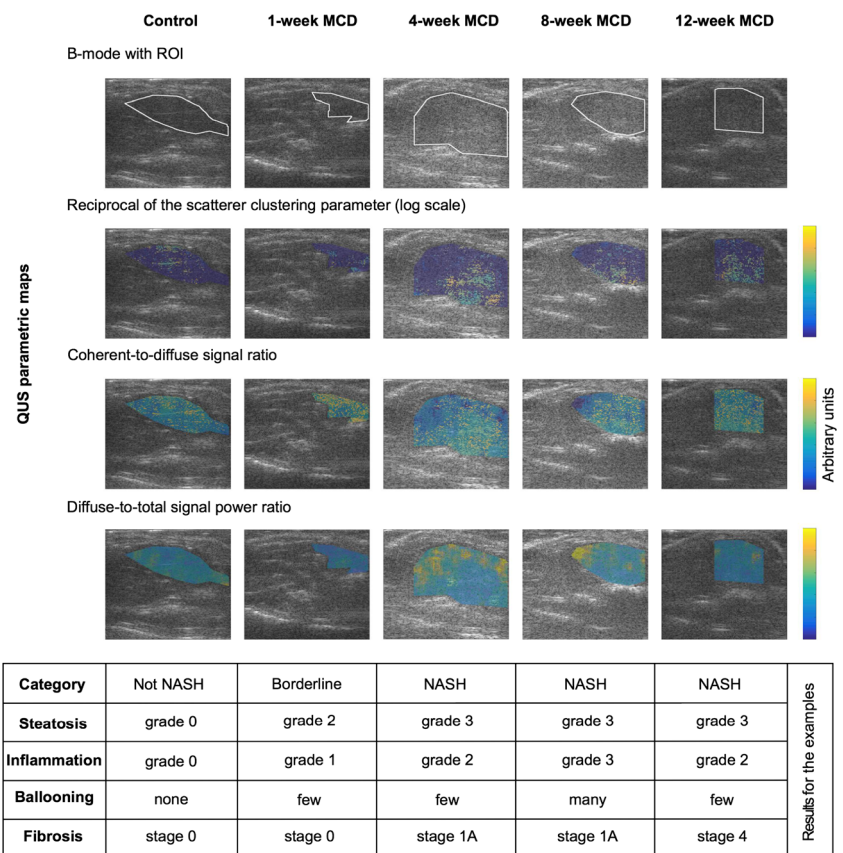
Pathological features	Groups	Size	AUC G' only	AUC multi-parameter	Parameters	p value
Diagnosis of steatohepatitis	Not SH vs. (borderline or SH)	14/45	0.66 (0.63–0.70)	0.75 (0.73–0.77)	$1/\alpha$ IQR + $1/(\kappa + 1)$ IQR + local Att.	< 0.001
	(Not SH or borderline) vs. SH	25/34	0.79 (0.78–0.81)	0.82 (0.81–0.85)	$G' + 1/\alpha$ IQR + μ_n mean	< 0.001
	(Not SH or borderline or SH with fibrosis stage ≤ 1) vs. SH with fibrosis stage ≥ 2	52/7	0.92 (0.91–0.93)	0.98 (0.97–0.98)	$G' + k$ IQR	< 0.001
Steatosis	0 vs. ≥ 1	12/47	0.73 (0.70–0.76)	0.80 (0.78–0.82)	$1/\alpha$ IQR + $1/(\kappa + 1)$ IQR + local Att.	< 0.001
	≤ 1 vs. ≥ 2	13/46	0.69 (0.65–0.73)	0.80 (0.79–0.81)	$1/\alpha$ IQR + $1/(\kappa + 1)$ IQR + local Att.	< 0.001
	≤ 2 vs. 3	17/42	0.71 (0.69–0.74)	0.78 (0.76–0.80)	$G' + k$ mean + $1/(\kappa + 1)$ IQR	< 0.001
Inflammation	0 vs. ≥ 1	12/47	0.59 (0.54–0.62)	0.69 (0.66–0.73)	$G' + 1/\alpha$ IQR + $1/\alpha$ mean	< 0.001
	≤ 1 vs. ≥ 2	31/28	0.77 (0.76–0.79)	0.84 (0.83–0.86)	$G' + 1/\alpha$ IQR + μ_n mean	< 0.001
	≤ 2 vs. 3	46/13	0.79 (0.77–0.82)	0.86 (0.84–0.87)	$G' + \mu_n$ IQR + $1/(\kappa + 1)$ mean	< 0.001
Fibrosis	0 vs. ≥ 1	31/28	0.79 (0.78–0.80)	0.85 (0.84–0.86)	$G' + k$ mean + μ_n IQR	< 0.001
	≤ 1 vs. ≥ 2	52/7	0.92 (0.91–0.93)	0.98 (0.97–0.98)	$G' + k$ IQR	< 0.001
	≤ 2 vs. ≥ 3	53/6	0.91 (0.90–0.92)	0.96 (0.95–0.97)	$G' + k$ IQR + $1/\alpha$ mean	< 0.001

Holm-Bonferroni correction was applied to p values

AUC = area under the receiver-operating characteristic curve. Numbers in parentheses are 95% confidence intervals

G' storage modulus, μ_n = mean intensity normalized by its maximal value, $1/\alpha$ reciprocal of the scatterer-clustering parameter, k coherent-to-diffuse signal ratio, $1/(\kappa + 1)$ diffuse-to-total signal power ratio, IQR inter-quartile range

Fig. 3 Representative B-mode and QUS parametric maps of five rats fed a standard chow or a methionine- and choline-deficient diet for 1, 4, 8, or 12 weeks. Top row shows B-mode images for each animal with white lines outlining segmented regions of interests. Second row shows color-coded QUS parametric map for the scatterer-clustering parameter (displayed in log-scale). Third row shows QUS parametric map for the coherent-to-diffuse signal ratio. Bottom row shows QUS parametric map for the diffuse-to-total signal power ratio. Yellow indicates higher values and dark blue lower values. Images have been cropped from their bottom third for display



Steatosis grades ≤ 2 were best staged by the local attenuation in combination with other QUS parameters based on homodyned-K modeling. Of note, combinations of QUS parameters achieved good classification accuracy: 0.78 for steatosis grades 0 vs. ≥ 1 , 0.78 for grades ≤ 1 vs. ≥ 2 , and 0.75 for ≤ 2 vs. 3 (by also considering elastography). This is consistent with recent studies that have shown that ultrasound attenuation, either measured by a controlled attenuation parameter [15–17] or local attenuation [18, 19] constitutes accurate biomarkers for the detection and quantification of steatosis. Using liver biopsy as their reference standard, de Ledinghen et al showed that the controlled attenuation parameter evaluated with transient elastography could be useful for the detection of steatosis with an AUC of 0.80 for grade 2 or higher and 0.66 for grade 3 [15]. Using histology as the reference standard, Paige et al reported that the attenuation coefficient could achieve higher accuracy than conventional ultrasound for grading liver steatosis, with an AUC of 0.79 for grade 2 or higher and 0.80 for grade 3 [19].

Inflammation was best graded by shear wave elastography in combination with QUS parameters. This is similar to recent studies that have reported mild increases in liver stiffness in the presence of inflammation, either with ultrasound-based [45, 46] or MR-based elastography [11, 47]. Investigators

showed that inflammation, either through infiltration of inflammatory cells or through edema, which increases the internal pressure of the liver [44], may increase the liver stiffness. Our results suggest that QUS, which measures microstructural changes, may provide a technique to account for these phenomena.

As expected, fibrosis could be staged with good to excellent accuracy using shear wave elastography alone, with AUCs ranging from 0.79 to 0.92. This is consistent with an abundant literature demonstrating the high fibrosis staging accuracy of liver stiffness in humans [20, 21]. Interestingly, these results suggest that the combination of additional QUS parameters further improved the classification accuracy, with AUCs ranging from 0.85 to 0.98, in the range of diagnostic performance typically observed with MR elastography [10, 51].

Taken together, these results lead us to confirm the hypothesis that QUS-based diagnosis of NASH and quantification of its individual histological components (steatosis, inflammation, and fibrosis) are achievable noninvasively within an ultrasound examination. Unlike prior studies that have focused on one or two histological components (mainly steatosis or fibrosis), we have addressed all three at once. This is important because the coexistence of these conditions may all have a confounding

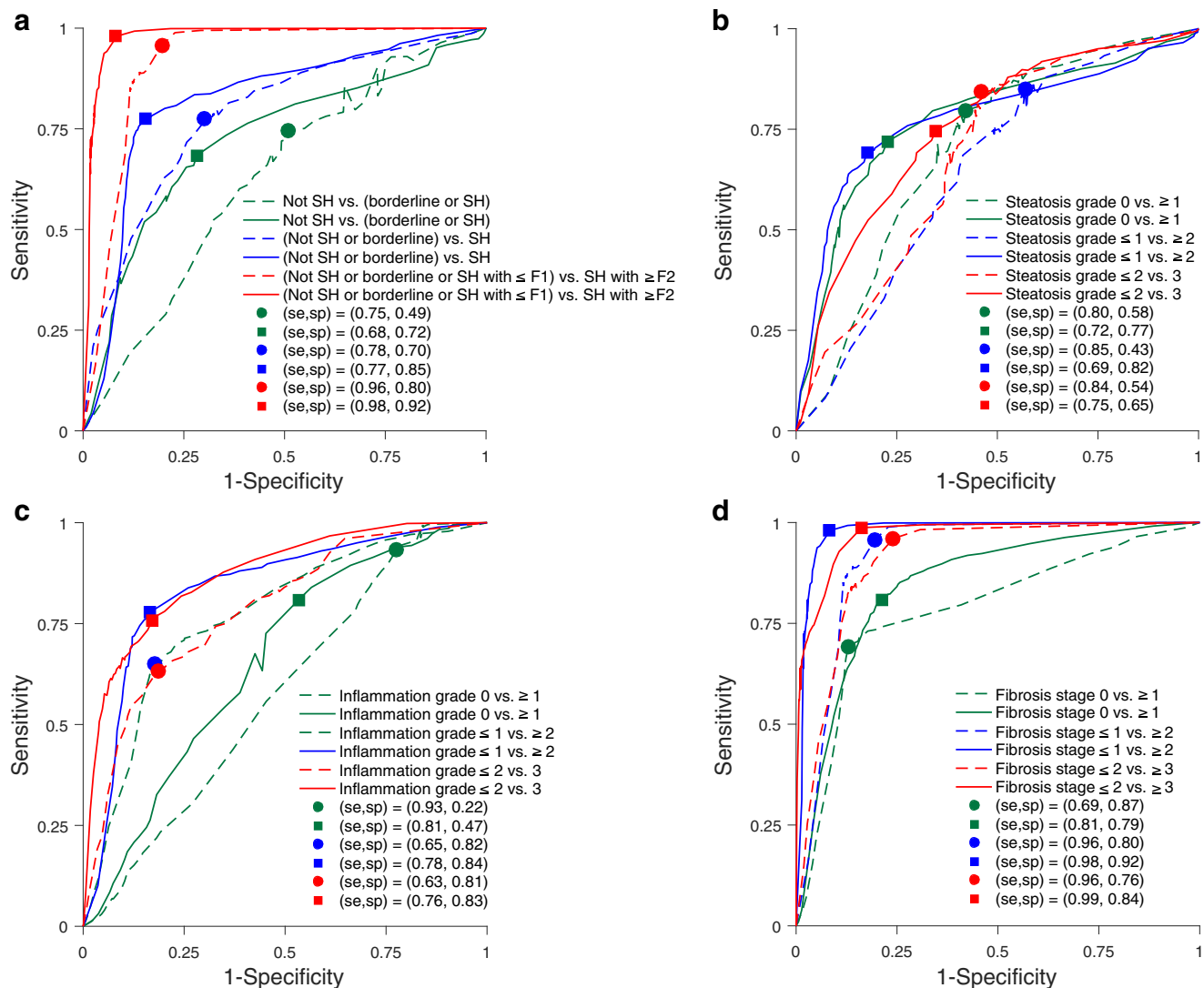


Fig. 4 ROC curves obtained for the evaluation of classification accuracy of liver steatohepatitis and its histological features. **a** Steatohepatitis categories. **b** Steatosis grade. **c** Inflammation grade. **d** Fibrosis stage. Dashed lines correspond to classification accuracy for storage modulus

(G') only and full lines to combination of QUS features. Specificity and sensitivity corresponding to points on ROC curves that maximize Youden's index (i.e., specificity + sensitivity - 1) are reported

effect on liver stiffness. Of note, our study revealed associations between steatosis and inflammation, and between inflammation and fibrosis. Hence, a multi-parametric approach may take into account the coexistence of these multiple confounders: the stiffness-lowering effect of liver steatosis overlapping with stiffness-increasing inflammation and fibrosis [24].

Our study has potential limitations. Using shear wave elastography, we only assessed the storage modulus (G'), which is related to elasticity, and did not measure the loss modulus (G''), which is related to viscosity. Although technically possible with state-of-the-art ultrasound [27], shear wave viscoelastography is difficult to achieve in an animal model due to the small size of livers but should be feasible in human livers. Future work should be performed to determine if

viscoelastography in combination with QUS parameters would further improve the classification of NASH and its histopathological components.

Another limitation was imbalanced classes for three of the dichotomic tasks. However, a comparison based on the metric AUC-PR local to the same conclusion that combining QUS parameters with elastography does improve greatly the performance of classifiers.

In summary, this animal study reveals that a random forest model based on QUS and shear wave elastography improved classification accuracy of liver steatohepatitis and its histological features (liver steatosis, inflammation, and fibrosis) compared to elastography alone. Further research should be performed to demonstrate the applicability of this multi-parametric QUS approach in a human cohort and to validate

the combinations of parameters providing the highest classification accuracy.

Acknowledgments We thank Mr. Jamal Ait Ichou for his assistance in the literature review.

Funding This study has received funding by Fonds de Recherche du Québec—Nature et Technologies (FRQNT) (PR-174387), Canadian Institutes of Health Research, Institute of Nutrition, Metabolism, and Diabetes (grant nos. 273738 and 301520), Quebec Bio-imaging Network (QBIN/RBIQ #5886), New Researcher Startup Grant from the Centre de Recherche du Centre Hospitalier de l'Université de Montréal, and by a Chercheur-Boursier Junior 2 from the Fonds de Recherche du Québec en Santé and Fondation de l'association des radiologistes du Québec (FRQS-ARQ #34939) to An Tang.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is An Tang.

Conflict of interest The authors declare that they have no conflict of interest.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Approval from the institutional animal care committee was obtained.

Ethical approval Institutional review board approval was obtained.

Study subjects or cohorts overlap The study cohorts have been previously reported in [24]. That prior article focused on a comparison between low-frequency versus high-frequency ultrasonographic shear wave elastography for the detection of steatohepatitis. In contrast, the present article investigates a machine learning model based on quantitative ultrasound parameters to improve classification of steatohepatitis compared to shear wave elastography alone.

Methodology

- Diagnostic study
- Experimental
- Performed at one institution

References

- Vernon G, Baranova A, Younossi ZM (2011) Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther* 34:274–285
- Kleiner DE, Brunt EM, Van Natta M et al (2005) Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 41:1313–1321
- Sanyal AJ, Brunt EM, Kleiner DE et al (2011) Endpoints and clinical trial design for nonalcoholic steatohepatitis. *Hepatology* 54:344–353
- Fernández-Salazar L, Velayos B, Aller R, Lozano F, Garrote JA, González JM (2011) Percutaneous liver biopsy: patients' point of view. *Scand J Gastroenterol* 46:727–731
- Myers RP, Fong A, Shaheen AA (2008) Utilization rates, complications and costs of percutaneous liver biopsy: a population-based study including 4275 biopsies. *Liver Int* 28:705–712
- Reeder SB, Cruite I, Hamilton G, Sirlin CB (2011) Quantitative assessment of liver fat with magnetic resonance imaging and spectroscopy. *J Magn Reson Imaging* 34:spcone
- Tang A, Desai A, Hamilton G et al (2015) Accuracy of MR imaging-estimated proton density fat fraction for classification of dichotomized histologic steatosis grades in nonalcoholic fatty liver disease. *Radiology* 274:416–425
- Yokoo T, Serai SD, Pirasteh A et al (2018) Linearity, bias, and precision of hepatic proton density fat fraction measurements by using MR imaging: a meta-analysis. *Radiology* 286:486–498
- Petitclerc L, Sebastiani G, Gilbert G, Cloutier G, Tang A (2017) Liver fibrosis: review of current imaging and MRI quantification techniques. *J Magn Reson Imaging* 45:1276–1295
- Singh S, Venkatesh SK, Wang Z et al (2015) Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of individual participant data. *Clin Gastroenterol Hepatol* 13:440–451.e446
- Salameh N, Larrat B, Abarca-Quinones J et al (2009) Early detection of steatohepatitis in fatty rat liver by using MR elastography. *Radiology* 253:90–97
- Yin M, Glaser KJ, Manduca A et al (2017) Distinguishing between hepatic inflammation and fibrosis with MR Elastography. *Radiology* 284:694–705
- Zhang E, Wartelle-Bladou C, Lepanto L, Lachaine J, Cloutier G, Tang A (2015) Cost-utility analysis of nonalcoholic steatohepatitis screening. *Eur Radiol* 25:3282–3294
- Park CC, Nguyen P, Hernandez C et al (2017) Magnetic resonance elastography vs transient elastography in detection of fibrosis and noninvasive measurement of steatosis in patients with biopsy-proven nonalcoholic fatty liver disease. *Gastroenterology* 152:598–607.e592
- de Lédinghen V, Wong GL, Vergniol J et al (2016) Controlled attenuation parameter for the diagnosis of steatosis in non-alcoholic fatty liver disease. *J Gastroenterol Hepatol* 31:848–855
- Chan WK, Nik Mustapha NR, Mahadeva S (2014) Controlled attenuation parameter for the detection and quantification of hepatic steatosis in nonalcoholic fatty liver disease. *J Gastroenterol Hepatol* 29:1470–1476
- Kumar M, Rastogi A, Singh T et al (2013) Controlled attenuation parameter for non-invasive assessment of hepatic steatosis: does etiology affect performance? *J Gastroenterol Hepatol* 28:1194–1201
- Lin SC, Heba E, Wolfson T et al (2015) Noninvasive diagnosis of nonalcoholic fatty liver disease and quantification of liver fat using a new quantitative ultrasound technique. *Clin Gastroenterol Hepatol* 13:1337–1345.e1336
- Paige JS, Bernstein GS, Heba E et al (2017) A pilot comparative study of quantitative ultrasound, conventional ultrasound, and MRI for predicting histology-determined steatosis grade in adult nonalcoholic fatty liver disease. *AJR Am J Roentgenol* 208:W168–W177
- Friedrich-Rust M, Nierhoff J, Lupsor M et al (2012) Performance of acoustic radiation force impulse imaging for the staging of liver fibrosis: a pooled meta-analysis. *J Viral Hepat* 19:e212–e219
- Friedrich-Rust M, Ong MF, Martens S et al (2008) Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. *Gastroenterology* 134:960–974
- Oelze ML, Mamou J (2016) Review of quantitative ultrasound: envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound. *IEEE Trans Ultrason Ferroelectr Freq Control* 63:336–351
- Mamou J, Oelze ML (2013) Quantitative ultrasound in soft tissues. Springer, Dordrecht

24. Kazemirad S, Zhang E, Nguyen BN et al (2017) Detection of steatohepatitis in a rat model by using spectroscopic shear-wave US elastography. *Radiology* 282:726–733
25. Rinella ME, Elias MS, Smolak RR, Fu T, Borensztajn J, Green RM (2008) Mechanisms of hepatic steatosis in mice fed a lipogenic methionine choline-deficient diet. *J Lipid Res* 49:1068–1076
26. Hebbard L, George J (2011) Animal models of nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol* 8:35–44
27. Kazemirad S, Bernard S, Hybois S, Tang A, Cloutier G (2016) Ultrasound shear wave viscoelastography: model-independent quantification of the complex shear modulus. *IEEE Trans Ultrason Ferroelectr Freq Control* 63:1399–1408
28. Bercoff J, Tanter M, Fink M (2004) Supersonic shear imaging: a new technique for soft tissue elasticity mapping. *IEEE Trans Ultrason Ferroelectr Freq Control* 51:396–409
29. Garcia D, Le Tarnec L, Muth S, Montagnon E, Porée J, Cloutier G (2013) Stolt's f-k migration for plane wave ultrasound imaging. *IEEE Trans Ultrason Ferroelectr Freq Control* 60:1853–1867
30. Ophir J, Céspedes I, Ponnekanti H, Yazdi Y, Li X (1991) Elastography: a quantitative method for imaging the elasticity of biological tissues. *Ultrason Imaging* 13:111–134
31. Bernal M, Nenadic I, Urban MW, Greenleaf JF (2011) Material property estimation for tubes and arteries using ultrasound radiation force and analysis of propagating modes. *J Acoust Soc Am* 129:1344–1354
32. Destrempes F, Meunier J, Giroux MF, Soulez G, Cloutier G (2011) Segmentation of plaques in sequences of ultrasonic B-mode images of carotid arteries based on motion estimation and a Bayesian model. *IEEE Trans Biomed Eng* 58:2202–2211. <https://doi.org/10.1109/TBME.2011.2127476>
33. Destrempes F, Porée J, Cloutier G (2013) Estimation method of the homodyned K-distribution based on the mean intensity and two log-moments. *SIAM J Imaging Sci* 6:1499–1530
34. Destrempes F, Franceschini E, Yu FT, Cloutier G (2016) Unifying concepts of statistical and spectral quantitative ultrasound techniques. *IEEE Trans Med Imaging* 35:488–500
35. Bigelow TA, Labyed Y (2013) Attenuation compensation and estimation. In: Mamou J, Oelze ML (eds) *Quantitative ultrasound in soft tissues*. Springer, Dordrecht, pp 71–93
36. Montaldo G, Tanter M, Bercoff J, Benech N, Fink M (2009) Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE Trans Ultrason Ferroelectr Freq Control* 56:489–506
37. Breiman L (2001) Random forests. *Mach Learn* 45:5–32
38. He H, Garcia EA (2009) Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 21:1263–1284
39. Efron B, Tibshirani R (1997) Improvements on cross-validation: the 632+ bootstrap method. *J Am Stat Assoc* 92:548–560
40. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837–845
41. Davis J, Goadrich M (2006) The relationship between precision-recall and ROC curves. *Proceedings of the 23rd international conference on machine learning*. ACM, pp 233–240
42. Parker KJ, Dooley MM, Rubens DJ (2011) Imaging the elastic properties of tissue: the 20 year perspective. *Phys Med Biol* 56:R1–R29
43. Barr RG (2014) Elastography in clinical practice. *Radiol Clin North Am* 52:1145–1162
44. Iredale JP (2007) Models of liver fibrosis: exploring the dynamic nature of inflammation and repair in a solid organ. *J Clin Invest* 117:539–548
45. Zeng X, Xu C, He D et al (2015) Influence of hepatic inflammation on FibroScan findings in diagnosing fibrosis in patients with chronic hepatitis B. *Ultrasound Med Biol* 41:1538–1544
46. Dong DR, Hao MN, Li C et al (2015) Acoustic radiation force impulse elastography, FibroScan®, Forns' index and their combination in the assessment of liver fibrosis in patients with chronic hepatitis B, and the impact of inflammatory activity and steatosis on these diagnostic methods. *Mol Med Rep* 11:4174–4182
47. Chen J, Talwalkar JA, Yin M, Glaser KJ, Sanderson SO, Ehman RL (2011) Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. *Radiology* 259:749–756
48. Yoneda M, Suzuki K, Kato S et al (2010) Nonalcoholic fatty liver disease: US-based acoustic radiation force impulse elastography. *Radiology* 256:640–647
49. Mohareri O, Ruszkowski A, Lobo J et al (2014) Multi-parametric 3D quantitative ultrasound vibro-elastography imaging for detecting palpable prostate tumors. *Med Image Comput Comput Assist Interv* 17:561–568
50. Roy-Cardinal M, Destrempes F, Soulez G, Cloutier G (2016) Homodyned K-distribution parametric maps combined with elastograms for carotid artery plaque assessment. 2016 IEEE International Ultrasonics Symposium (IUS) pp 1–4
51. Wang QB, Zhu H, Liu HL, Zhang B (2012) Performance of magnetic resonance elastography and diffusion-weighted imaging for the staging of hepatic fibrosis: a meta-analysis. *Hepatology* 56:239–247