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

An Tang, François Destrempes, Siavash Kazemirad, Julian Garcia-Duitama, Bich N. Nguyen, Guy Cloutier

Received: 31 July 2018 / Revised: 29 October 2018 / Accepted: 23 November 2018 / Published online: 17 December 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.
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-μs 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 $82 \times 10$ pixels$^2$, corresponding to 3 mm in both axial and lateral directions, was swept across the ROI by steps of $4 \times 1$ pixels$^2$, and estimation of HKD parameters was performed within the sliding window as in Destrempes 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 $\mu_n$, the reciprocal $1/\alpha$ of the scatterer-clustering parameter $\alpha$, the coherent-to-diffuse signal ratio $k$, as well as the diffuse-to-total signal power ratio $1/(\kappa + 1)$ [34]. The mean intensity $\mu_n$ 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 $\kappa$ 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: $\mu_n$, $1/\alpha$, $k$, 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 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 steatohapatitis vs. borderline or steatohapatitis (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).
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. $\geq 1$, $\leq 1$ vs. $\geq 2$, $\leq 2$ 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. $\geq 1$, $\leq 1$ vs. $\geq 2$, $\leq 2$ 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. $\geq 1$, $\leq 1$ vs. $\geq 2$, and $\leq 2$ vs. $\geq 3$, 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 ($\mu_n$, $1/\alpha$, $k$, $1/(\kappa + 1)$, 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
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

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

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, $\mu_n$ = 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.
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
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 et Fondation de l’association des radiologistes du Québec (FRQS-ARQ #34939) to An Tang.

Compliance with ethical standards

Guarantor The scientific guarantor if 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