Title

Accuracy of common proton density fat fraction thresholds for magnitude- and complex-based chemical shift-encoded MRI for assessing hepatic steatosis in patients with obesity

Journal

Abdominal Radiology, 45(3)

ISSN

2366-004X

Authors

Cunha, Guilherme Moura
Thai, Tydus T
Hamilton, Gavin
et al.

Publication Date

2020-03-01

DOI

10.1007/s00261-019-02350-3

Peer reviewed

Published in final edited form as:

Abdom Radiol (NY). 2020 March ; 45(3): 661–671. doi:10.1007/s00261-019-02350-3.

Accuracy of common proton density fat fraction thresholds for magnitude- and complex-based chemical-shift-encoded MRI for assessing hepatic steatosis in patients with obesity

Guilherme Moura Cunha, MD[1], **Tydus T. Thai, MD**[1], **Gavin Hamilton, PhD**[1], **Yesenia Covarrubias, MFS**[1], **Alex Schlein, BS**[1], **Michael S. Middleton, MD, PhD**[1], **Curtis N. Wiens, PhD**[2], **Alan McMillan, PhD**[2], **Rashmi Agni, MD**[3], **Luke M. Funk, MD, MPH**[4], **Guilherme M. Campos, MD**[5], **Santiago Horgan, MD**[6], **Garth Jacobson, MD**[6], **Tanya Wolfson, MA**[7], **Anthony Gamst, PhD**[7], **Jeffrey B. Schwimmer, MD**[8], **Scott B. Reeder, MD, PhD**[2,9,10], **Claude B. Sirlin, MD**[1]

[1]Liver Imaging Group, Radiology, University of California-San Diego, 9500 Gilman Drive, San Diego, CA, United States 92037

[2]Department of Radiology, University of Wisconsin, School of Medicine and Public Health, E3/366 Clinical Science Center, 600 Highland Avenue, Madison, WI United States 53792-3252

[3]Department of Pathology and Laboratory Medicine, University of Wisconsin School of Medicine and Public Health, 3170 UW Medical Foundation Centennial Building (MFCB), 1685 Highland Avenue, Madison, WI United States 53705-2281

[4]Surgery, University of Wisconsin, Clinical Science Center, 600 Highland Avenue, Madison, WI United States 53792-3252

[5]Virginia Commonwealth University Department of Surgery West Hospital,1200 East Broad Street 16th Floor, West Wing Box 980645, Richmond, Virginia 23298-0645

[6]Surgery, University of California-San Diego, University of California San Diego, 9500 Gilman Drive, San Diego, CA, United States 92037

[7]Computational and Applied Statistics Laboratory (CASL), San Diego Supercomputer Center, University of California-San Diego, 9500 Gilman Drive, San Diego, CA, United States 92037

[8]Pediatrics, University of California San Diego, 9500 Gilman Drive, San Diego, CA, United States 92037

[9]Medical Physics, University of Wisconsin Madison, Clinical Science Center, 600 Highland Avenue, Madison, WI United States 53792-3252

[10]Biomedical Engineering, Madison, WI, Clinical Science Center, 600 Highland Avenue, Madison, WI United States 53792-3252

Abstract

Corresponding author: Guilherme Moura Cunha, MD, Liver Imaging Group, Radiology, Altman Clinical Translational Research Institute, 9452 Medical Center Drive, Lower Level 501, La Jolla, CA 92037, **Contact info**: Telephone: (858) 246-2220; gcunha@ucsd.edu, mouracunha@hotmail.com; fax: N/A.
Institutions from which work originated: University of California at San Diego and University of Wisconsin at Madison

Purpose—MRI proton density fat fraction (PDFF) can be calculated using magnitude (MRI-M) or complex (MRI-C) MRI data. The purpose of this study was to identify, assess and compare the accuracy of common PDFF thresholds for MRI-M and MRI-C for assessing hepatic steatosis in patients with obesity, using histology as reference.

Methods—This two-center prospective study included patients undergoing MRI-C- and MRI-M-PDFF estimations within 3 days before weight-loss surgery. Liver biopsy was performed, and histology-determined steatosis grades used as reference standard. Using receiver operating characteristics (ROC) analysis on data pooled from both methods, single common thresholds for diagnosing and for differentiating none or mild (0–1) from moderate to severe steatosis (2–3) were selected as the ones achieving the highest sensitivity while providing at least 90% specificity. Selection methods were cross-validated. Performances were compared using McNemar's tests.

Results—Of 81 included patients, 54 (67%) had steatosis. The common PDFF threshold for diagnosing steatosis was 5.4%, which provided cross-validated 0.88 (95% CI, 0.77–0.95) sensitivity and 0.92 (0.75–0.99) specificity for MRI-M and 0.87 sensitivity (0.75–0.94) with 0.81 (0.61–0.93) specificity for MRI-C. The common PDFF threshold to differentiate steatosis grades 0–1 from 2–3 was 14.7%, which provided cross-validated 0.86 (95% CI, 0.59–0.98) sensitivity and 0.95 (0.87–0.99) specificity for MRI-M and 0.93 sensitivity (0.68–0.99) with 0.97(0.89–0.99) specificity for MRI-C.

Conclusion—If independently validated, diagnostic thresholds of 5.4% and 14.7% could be adopted for both techniques for detecting and for differentiating none to mild from moderate to severe steatosis, respectively, with high diagnostic accuracy.

Keywords

Liver; Magnetic resonance imaging; Non-alcoholic fat liver disease; Obesity

Introduction

The increasing prevalence of obesity has caused nonalcoholic fatty liver disease (NAFLD) to become the most common cause of chronic liver disease worldwide [1]. MRI is an accurate, reproducible and non-invasive tool to estimate liver fat in patients at risk for NAFLD [2]. Confounder-corrected chemical-shift-encoded (CSE)-MRI methods accurately quantify proton density fat fraction (PDFF), an inherent tissue property and quantitative imaging biomarker of liver fat concentration [2,3]. PDFF parametric maps showing the distribution of fat in the liver can be generated using magnitude (MRI-M) or complex (MRI-C) MRI data. Both methods minimize or correct for biases caused by T1, T2*, and the spectral complexity of fat [4–7], while MRI-C also addresses noise-related bias and the effects of eddy currents [5, 8–10]. Other differences include the range of measurable PDFF, which is up to 50% for MRI-M but up to 100% for MRI-C. Early studies on CSE-MRI accuracy were performed using MRI-M PDFF [2,11–14], while more recently, the predominant technique used in research or provided as a commercial product in the newest MRI systems is MRI-C PDFF [2, 15–18].

Due to its high accuracy, repeatability and reproducibility, MRI-PDFF is widely recognized as the leading non-invasive biomarker for liver steatosis diagnosis and quantification, but

challenges to its adoption remain [2,8,17,19,20]. For example, most studies on the accuracy of MRI-PDFF for diagnosing steatosis have investigated a single method [2,11–14,17,18,21], proposing their own diagnostic thresholds based on receiver operating characteristics (ROC) analysis applied to their population cohorts [11, 12, 18, 21, 22]. This has led to the publication of at least 8 different thresholds for diagnosing the presence of steatosis, ranging from 2.51% to 6.90% for MRI-M [11, 12, 21, 23], and from 2.94 to 5.20% for MRI-C using histology as the reference standard [18, 21, 24]. Similarly, at least 4 different thresholds ranging from 8.5% to 16.3% to distinguish none or mild from moderate to severe steatosis have been described [17,18,22,24]. The existence of multiple thresholds creates confusion and introduces inconsistency in the use of MRI-PDFF to diagnose steatosis. To reduce the inconsistency in the use of this biomarker and to further advance its standardization, a single diagnostic threshold is needed.

The purpose of this study was to identify, assess and compare the accuracy of common PDFF thresholds for MRI-M and MRI-C to diagnose the presence of steatosis, and secondarily, to differentiate none or mild (grades 0–1) from moderate to severe (grades 2–3) hepatic steatosis in patients with obesity, using contemporaneous histology as reference.

Materials and Methods

Study Design and Patients

This Health Insurance Portability and Accountability Act (HIPAA)-compliant, two-center prospective study was approved by each center's institutional review board (IRB). Inclusion criteria were adults (age ≥18 years) with obesity class II [25] or higher (baseline BMI ≥ 35kg/m^2) undergoing clinical-care weight loss surgery (WLS), either laparoscopic sleeve gastrectomy or laparoscopic Roux-en-Y gastric bypass, and willingness to participate in all research procedures, including MR examinations and intraoperative biopsy. Exclusion criteria included contraindications to MRI (e.g., claustrophobia, metallic implants), excessive alcohol consumption (>1.5 drinks per day) [26], liver disease other than NAFLD, and type 1 diabetes. Additionally, patients were excluded if one or both of the MRI methods was not acquired, or if intraoperative biopsy was not obtained.

Patients provided written informed consent and underwent multiple MR examinations as part of another study examining PDFF changes following WLS [3]. PDFF estimates for this study were derived from MR examinations that took place within 3 days prior to WLS. Patient demographics, time intervals from MRI to surgery, and reasons for exclusion were recorded.

Liver biopsy and histology

As part of the research protocol, surgeons performed wedge and core biopsies during WLS if they thought that both could be obtained safely. Otherwise, based on their judgment, they chose the safest biopsy method in individual patients. Wedge biopsies were typically about 1–4 cm^2 at the surface and 1–2 cm in depth. Core biopsies weighed 20–30 mg and were obtained using 18-gauge needles. Both types of biopsies were taken from the anterior surface of the left lobe of the liver (segments 2, 3, or 4), avoiding major blood vessels and

bile ducts. Histology slides were prepared with hematoxylin and eosin, Masson's trichome, and iron stains.

Blinded to the MRI results, two hepatopathologists (each with at least three years of experience) independently scored each biopsy specimen (wedge, core, or both) in each patient. Whenever both core and wedge biopsies were performed, the scores were averaged to yield composite scores. Based on the estimated percentage of hepatocytes containing microscopically visible fat vacuoles, hepatic steatosis was scored according to the NASH CRN system [27] as follows: grade 0, <5%; grade 1, 5–33%, grade 2, 34–66%; grade 3, >66%. Additional histology features of lobular and portal inflammation, hepatocellular ballooning, fibrosis stage and iron overload were also scored. After the independent reviews, the hepatopathologists re-reviewed all specimens of each patient in consensus, adjudicated any disagreements, and assigned a final consensus score for each histology feature. The hepatic steatosis grades for each patient were then binarized as steatosis present (grade ≥ 1) vs. absent (grade 0) and as none to mild steatosis (grades 0–1) vs. moderate to severe steatosis (grades 2–3). A NAFLD activity score was determined as the unweighted sum of the scores for steatosis (0–3), lobular inflammation (0–3), and ballooning (0–2), and represented as thresholds of ≤ 3 (low possibility of steatohepatitis), 3.5 – 4.5, and ≥ 5 (high possibility of steatohepatitis) [27]. Half grades or stages were due to averaging results from wedge and core biopsies.

Patient preparation and positioning

Patients were instructed to fast for at least four hours prior to MR examinations. At one center, patients underwent non-contrast MR examinations on a clinical 3.0T GE MRI system with an 8-element torso phased-array coil (GE Signa, EXCITE HDxt, GE Healthcare, Waukesha, WI); at this center, dielectric pads were placed over the abdomen to reduce B1 heterogeneity. At the other center, patients underwent non-contrast MR examinations on a clinical 3.0T GE 750 MRI system with a 32-element torso phased-array coil, or if the patient did not fit into the bore, on a 1.5T 450W GE wide bore system with an 8-element torso phased-array coil; dielectric pads were not placed. Coils were centered over the liver at both centers. Each MR examination lasted about 60 minutes comprising the two PDFF sequences described below in addition to other research acquisitions unrelated to this study.

MRI-PDFF sequences and parametric map reconstruction

A 2D spoiled gradient-recalled echo (SGRE) sequence was used for MRI-M while a 3D SGRE sequence was used for MRI-C. Axial images covering as much of the liver as possible within one breath hold were acquired with the parameters listed in Table 1. These parameters are similar to those used in prior MRI-M and MRI-C studies [8,11,13]. PDFF maps were reconstructed in-line on the MRI system computer from the corresponding source data as described elsewhere for MRI-M [8,28] and MRI-C [6,8]. Source images, PDFF maps, and (for MRI-C) fat and water images were transferred for off-line analysis.

MR analysis

Trained image analysts (1–3 years experience), blinded to histologic results, performed the MR analysis using the Osirix imaging software (Pixmeo, Geneva, Switzerland). Including

only liver parenchyma while excluding vessels, bile ducts, lesions, and artifacts, the analysts placed co-localized circular (1 cm radius) regions of interest (ROIs), one in each of the nine Couinaud liver segments, on the fifth echo of the source images for MRI-M and on the water images for MRI-C as these provide adequate visualization of hepatic anatomy to guide proper ROI placement. This multiple ROIs approach was used to avoid bias in case of nonuniform spatial distribution of liver fat. ROIs were placed on anatomic images rather PDFF maps to avoid feedback bias. In some patients, one or more of the nine planned ROIs could not be placed on one or both sequences, usually because the segment was not included in the acquired volume or due to imaging artifacts.

Statistical analysis

Statistical analyses were performed by a staff biostatistician under the supervision of a faculty statistician, both with more than 20 years of experience, using R software package version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria, 2018). Cohort anthropometric, laboratory, and histologic measures were summarized. Individual MRI-M and MRI-C PDFF averages from segments 2, 3, 4a, and 4b (i.e., those in correspondence to the biopsy location) were computed for each patient, using only the segmental ROIs available in both sequences. Thus, if a segmental PDFF was missing for either sequence, the unpaired segment was excluded, and averages were based only on the remaining paired segments. PDFF values from the two MR sequences were summarized and compared using a paired t-test. Bland-Altman analysis was used to evaluate agreement and differences between MRI-M and MRI-C derived PDFF values. The data from the two MR methods was pooled, and receiver operating characteristic (ROC) curves for classifying dichotomized histology-determined steatosis grades using pooled data were generated. Areas under the ROC curve (AUCs) and their significances were calculated. The threshold for diagnosing presence of hepatic steatosis (grade ≥1) and for differentiating none to mild from moderate to severe steatosis (grades 0–1 vs 2–3) were selected based on separate ROC analyses to achieve the highest possible sensitivity while providing at least 90% specificity for pooled data. Raw performance parameters (sensitivity, specificity, total accuracy, negative predictive value [NPV], and positive predictive value [PPV]) with 95% confidence intervals [CI] were calculated for each MR sequence using the common thresholds. The method of threshold selection was cross-validated, and cross-validated performance parameters (same as above) were computed for each MR sequence. Cross-validated sensitivity, specificity, and accuracy were compared using McNemar’s test.

To examine the possible confounding impact of demographic and other liver histologic features (listed in Table 2) on diagnostic accuracy, a Bayesian Information Criterion (BIC)-based stepwise logistic regression was used for each MR sequence separately, for each of the two outcomes (presence of any hepatic steatosis, and presence of moderate to severe hepatic steatosis). For each analysis the outcome was presence vs. absence of classification error, while the predictor pool included demographic and histologic features.

Results

Patients

The number of patients who satisfied the inclusion criteria was 81. A summary of cohort characteristics at the time of MRI is presented in Table 2. Ten patients lost weight between enrollment and MRI; hence, although the inclusion criterion was BMI ≥ 35 kg/m^2, it ranged below this threshold at the time of PDFF measurements. Fifty-four (67%) patients had histology-determined steatosis, while the remaining 27 (33%) had no steatosis. Forty-four had both core and wedge biopsies, twenty-eight had only wedge biopsies, and nine had only core biopsies. Among the 44 patients with both biopsies, one patient with discordant scores, the wedge biopsy had 5% steatotic hepatocytes (grade 1) and the core biopsy had 1% steatotic hepatocytes (grade 0). Since the average of steatotic hepatocytes (3%) was <5%, the final assigned steatosis grade was 0 and the outcome was binarized as steatosis absent. 66 (81%) patients had steatosis grades 0–1 and 15 (19%) patients had steatosis grades 2–3. Relatively small proportions of patients had advanced histologic features on at least one biopsy specimen as defined by bridging fibrosis or cirrhosis: 6/81 (7%); lobular inflammation grade ≥ 2: 7/81 (9%); portal inflammation grade ≥ 2: 3/81 (4%); iron grade ≥ 2: 11/81 (8%), or any ballooning 16/81 (20%).

PDFF Data

Figure 1 shows representative MRI-M and MRI-C-based PDFF images at 3.0T of a patient with histology-determined hepatic steatosis. PDFF was estimated at 3.0T in 77 (95%) of the 81 patients, and at 1.5T in the remaining four (5%). Over the entire cohort, PDFF values ranged from 1.1% to 33.5% for MRI-M and 1.2% to 32.4% for MRI-C. Figure 2 shows a boxplot of PDFF values and steatosis grades for both techniques. PDFF values were slightly but consistently lower for MRI-M than MRI-C (mean PDFF 8.9% ± 6.5 vs 9.2% ± 6.45, paired t-test p=0.0018). A Bland-Altman plot of the two methods, with the standard agreement/disagreement metrics, is presented in Figure 3.

ROC Analysis

Figure 4 shows the ROC curves for diagnosing hepatic steatosis and for differentiating grades 0–1 from 2–3 for the pooled data. The AUC was 0.956 (bootstrap p < 0.0001) for classifying presence of any steatosis and 0.965 (p<0.0001) for differentiating none to mild from moderate to severe steatosis. The best diagnostic threshold for distinguishing absence vs presence of steatosis which enforced specificity at least 90% for the pooled data was 5.4% (0.89 sensitivity, 0.92 specificity, PPV 0.96, NPV 0.82, 0.90 accuracy). The best diagnostic threshold for distinguishing none to mild from moderate to severe steatosis which enforced specificity at least 90% for the pooled data was 14.7% (0.93 sensitivity, 0.96 specificity, PPV 0.84, NPV 0.98, 0.98 accuracy).

Common PDFF Thresholds Diagnostic Performance and Comparison

Table 3 summarizes raw and cross-validated performance metrics of both sequences for diagnosing steatosis (grades ≥ 1). Cross-validated results are described below.

For MRI-M cross-validation of the threshold selection method resulted in 0.88 sensitivity (95% CI: 0.77 – 0.95), 0.92 specificity (95% CI: 0.75 – 0.99), 0.96 PPV (95% CI: 0.86 – 0.99), 0.80 NPV (95% CI: 0.62 – 0.92), and 0.90 total accuracy (95% CI: 0.81 – 0.95).

For MRI-C cross-validation of the threshold selection method resulted in 0.87 sensitivity (95% CI: 0.75 – 0.94), 0.81 specificity (95% CI: 0.61 – 0.93), 0.90 PPV 95% CI: 0.79 – 96), 0.75 NPV (95% CI: 0.56 – 0.89), and 0.85 total accuracy (95% CI: 0.75 – 0.92).

There was no significant difference between sequences on cross-validated sensitivity, specificity, or accuracy in differentiating presence vs. absence of steatosis (McNemar's test p-values: 0.999, 0.248 and 0.221, respectively).

Table 4 summarizes raw and cross-validated performance metrics of both sequences for differentiating none to mild from moderate to severe (grades 0–1 vs 2–3) steatosis. Cross-validated results are described below.

For MRI-M cross-validation of the threshold selection method resulted in 0.80 sensitivity (95% CI: 0.51 – 0.95), 0.95 specificity (95% CI: 0.87 – 0.99), 0.80 PPV (95% CI: 0.51 – 0.95), 0.95 NPV (95% CI: 0.87 – 0.99), and 0.92 total accuracy (95% CI: 0.84 – 0.97).

For MRI-C cross-validation of the threshold selection method resulted in 0.93 sensitivity (95% CI: 0.68 – 0.99), 0.97 specificity (95% CI: 0.89 – 0.99), 0.87 PPV 95% CI: 0.61 – 98), 0.98 NPV (95% CI: 0.91 – 1.00), and 0.96 total accuracy (95% CI: 0.89 – 0.99).

There was no significant difference between methods on cross-validated sensitivity, specificity, or accuracy in differentiating grades 0–1 from grades 2–3 (McNemar's test p-values: 0.480, 0.999 and 0.248, respectively).

A BIC-based, stepwise logistic regression search for confounders to classification accuracy using the common PDFF diagnostic threshold of 5.4% revealed BMI to be the sole significant predictor of classification errors for both techniques (MRI-M: odds ratio = 0.71, chi-square p-value = 0.029; MRI-C: odds ratio = 0.84, chi-square p-value = 0.045), with lower BMI associated with more frequent classification errors. Age had an inverse association with misclassification for differentiating none to mild from moderate to severe steatosis based on the 14.7% threshold, however, the association was not strong and only for MRI-M (odds ratio = 0.92, chi-square p-value = 0.054). No additional histology feature had a significant confounding effect on the relationship between PDFF and steatosis grade.

Discussion

In this study we assessed and compared the diagnostic performance of MRI-M- and MRI-C-derived PDFF using a common single threshold for detecting hepatic steatosis and for distinguishing none to mild from moderate to severe steatosis in patients with obesity using contemporaneous histology as reference standard. Our results support both MRI-M and MRI-C as equally accurate to assess hepatic steatosis in adults with obesity with a high agreement between each other. Using single common PDFF thresholds of 5.4% and 14.7% for diagnosing and for differentiating none to mild from moderate to severe steatosis,

respectively, MRI-M and MRI-C are not statistically different, achieving high cross-validated sensitivity, specificity, and accuracy.

Our study comprised adults with obesity undergoing WLS without known hepatic steatosis *a priori,* and the prevalence of steatosis found in our cohort was similar to the prevalence described in the general population with obesity [29, 30]. More important, the distribution of steatosis grades in our population may provide a more accurate PDFF threshold for diagnosing steatosis than some prior studies performed in populations known to have NAFLD and with a disproportionate predominance of high liver fat fractions [11, 12, 31]. For widespread adoption of PDFF as a non-invasive biomarker of liver steatosis, CSE-MRI sequences should provide standardized and easy to interpret results despite the complexity of each individual reconstruction algorithms. Furthermore, although the clinical relevance of differentiating mild from moderate and severe steatosis is yet to be well established, some authors have described relationships between these individual grades and clinical outcomes [32, 33]. Hence, it is important to determine the MRI-derived PDFF threshold that accurately correlate with these histological grades as they may be used in the research and clinical settings. The different and specific diagnostic thresholds proposed for each method may be burdensome for radiologists when reading cases from multiple sites and may negatively impact the utilization of such quantification MRI techniques. A single diagnostic and grading thresholds may further facilitate the application of CSE-MRI techniques in different centers and may further support both methods as equally suitable for assessing hepatic steatosis in the majority of patients at risk for NAFLD.

Previous authors have investigated the individual accuracies of both CSE-MRI techniques. Tang et al. reported in two consecutive studies an MRI-M-based PDFF threshold for diagnosing steatosis of 6.4% and 6.8%, with 84% and 86% sensitivity and 100% and 83% specificity, respectively [11, 12]. However, in these studies cohorts comprised patients with known or suspected NAFLD, which resulted in a population with a disproportionate predominance of higher grades of liver steatosis. When investigating the accuracy of MRI-C-based PDFF to diagnose hepatic steatosis, other studies have also recruited patients with known or suspected NAFLD, enriching their populations with patients with high amounts of liver fat, and potentially resulting in overly high diagnostic thresholds [22, 24]. In the later studies, PDFF thresholds to differentiate moderate or severe steatosis from mild or no hepatic steatosis were 15.03% (sensitivity of 93.0% and a specificity of 85.0%) [22] and 11.3% (sensitivity 78.9% and specificity 84.1%) [24]. When including patients with a wide range of clinical indications for liver biopsy, Kühn et al. found that an MRI-C-determined PDFF threshold of 5.1% had 86% sensitivity and 100% specificity to differentiate between no steatosis versus any steatosis [23]. Similarly, in a population composed of various indications for liver biopsy, França et al. have reported a MRI-C PDFF thresholds for the presence vs absence of steatosis of 4.8% (88.6% sensitivity and 85.1% specificity) and of 8.5% for distinguishing patients with none or mild from moderate to severe steatosis (81.3% sensitivity and 96.8% specificity) [18]. These studies show that lower thresholds may be found when accuracy assessment is performed on cohorts not overrepresented with high liver fat. Hence, it’s important that studies investigating the accuracy of non-invasive techniques for diagnosing liver steatosis are performed in cohorts that are representative of the general population at risk for this condition. A previous large-scale study with 2349

ethnically diverse participants, and a prevalence of overweight and obesity similar to the western population, defined a threshold of 5.56% fat fraction to diagnose liver steatosis using MR spectroscopy (MRS) based on population nomograms but without histologic confirmation [34].

Until now, just a few studies on the accuracy of CSE-MRI for liver fat estimations have focused on both methods in the same cohort. Hernando et al. described more accurate estimates with lower bias and standard deviations for MRI-C reconstructions compared to MRI-M models using water/fat phantoms [35]. Horng et al. compared both reconstruction methods but using MR spectroscopy as reference standard [36]. To our knowledge, only one prior study compared the diagnostic performance of both methods simultaneously, using histology as reference standard. In this study, Kang et al. reported a Youden-index-based PDFF threshold for diagnosing liver steatosis of 2.51% for MRI-M and 2.94% for MRI-C, each providing 94% sensitivity and 81% specificity [21]. That study however, similarly to a prior study by Nasr et al. [37] who found an MRS PDFF threshold of 3%, have enrolled patients with liver diseases known to result in minimal liver fat, and did not focus on an obese population in which diseases other than NAFLD were excluded. Only 32% of that study's cohort had steatosis, which is lower than commonly encountered in populations with obesity, possibly explaining the lower threshold. In our study, a common threshold of 5.4% for both methods achieved similar to higher specificity (0.81 MRI-C, 0.92 MRI-M) for diagnosing the presence of hepatic steatosis. Additionally, when applying a common threshold of 14.7% for MRI-M and MRI-C to differentiate patients with none or moderate steatosis (grades 0–1) from patients with moderate to severe steatosis (grades 2–3), both methods yield high sensitivity (80% and 93%), specificity (95% and 97%) and accuracy (92% and 96%), respectively.

Using multivariable analysis, we examined the impact of several demographic, anthropometric, and histologic variables on classification errors. Our analysis found BMI as the sole significant predictor of misclassification of presence of steatosis for both methods, with lower BMI associated with more frequent classification errors. This is an unexpected finding since we would have expected that larger habitus, not smaller, might have reduced accuracy by degrading image quality. Further research is needed to confirm and elucidate the mechanism for this finding. Unlike other authors, we did not find that hepatic fibrosis or any other histology feature confounded the performance of MRI for assessing hepatic steatosis in our cohort [18, 22, 38]. One plausible explanation is our relatively mild degrees of fibrosis and other histological abnormalities. Our findings therefore do not eliminate the possibility that severe histological abnormalities might affect the accuracy of CSE-MRI methods.

There are a few distinct contributions of our work to the existing literature that merit mentioning. In this prospective two-center study we assessed patients with obesity which introduces technical challenges from the imaging perspective, were able to acquire both sequences during the same examinations and to provide contemporaneous histology as reference standard for MRI-PDFF findings. Additionally, analyses were performed using PDFF values of ROIs placed with some degree of colocalization with the biopsy sites, which may potentially reduce statistical noise related to sampling errors. Finally, we compared the accuracy between the two CSE-MRI methods head-to-head and showed the feasibility of

using common diagnostic thresholds to detect and to differentiate none to mild from moderate to severe liver steatosis in patients with obesity.

Our study has limitations: First, we used only one MRI system manufacturer and it is conceivable that the thresholds identified in our study may not be identical in different MRI systems. Among our scanners, the choice of platforms was determined by scanner availability and patient body habitus which may introduce variability in the results. Although studies have suggested excellent reproducibility of PDFF estimation across different field strengths and MRI systems [2, 20, 28] future studies are needed to validate our results across different MRI vendors, platforms or in different study populations including children. Another limitation is the relative limited number of patients with PDFF between 5% and 6% based on one or both methods. Further studies with larger populations in that PDFF range, both with and without steatosis, may further refine or validate the diagnostic threshold for presence of steatosis (grade ≥ 1).

Conclusion

In conclusion, this study further supports both MRI-M- and MRI-C-derived PDFF as equally accurate for assessing hepatic steatosis in obese adults at risk of developing NAFLD. For simplicity, common diagnostic thresholds of 5.4% and 14.7% could be adopted for both techniques for detection and for differentiating none to mild from moderate to severe steatosis, respectively, yielding high diagnostic accuracy.

Abbreviations

CSE-MRI	Chemical-shift-encoded-MRI
MRI-M	Magnitude-based data MRI
MRI-C	Complex-based data MRI
PDFF	Proton Density Fat Fraction
WLS	Weight Loss Surgery
NAFLD	Non-Alcoholic Fatty Liver Disease

References

1. Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, George J, Bugianesi E. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nature reviews Gastroenterology & hepatology. 2018 1;15(1):11 [PubMed: 28930295]
2. Yokoo T, Serai SD, Pirasteh A, Bashir MR, Hamilton G, Hernando D, Hu HH, Hetterich H, Kühn JP, Kukuk GM, Loomba R. Linearity, bias, and precision of hepatic proton density fat fraction measurements by using MR imaging: a meta-analysis. Radiology. 2017 9 11;286(2):486–98. [PubMed: 28892458]
3. Pooler BD, Wiens CN, McMillan A, Artz NS, Schlein A, Covarrubias Y, Hooker J, Schwimmer JB, Funk LM, Campos GM, Greenberg JA, Jacobsen G, Horgan S, Wolfson T, Gamst AC, Sirlin CB, Reeder SB. Monitoring Fatty Liver Disease with MRI Following Bariatric Surgery: A Prospective, Dual-Center Study. Radiology. 2019 3;290(3):682–690. [PubMed: 30561273]

4. Bydder M, Yokoo T, Hamilton G, Middleton MS, Chavez AD, Schwimmer JB, Lavine JE, Sirlin CB. Relaxation effects in the quantification of fat using gradient echo imaging. Magnetic resonance imaging. 2008 4 1;26(3):347–59 [PubMed: 18093781]

5. Liu CY, McKenzie CA, Yu H, Brittain JH, Reeder SB. Fat quantification with IDEAL gradient echo imaging: correction of bias from T1 and noise. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2007 8;58(2):354–64.

6. Yu H, McKenzie CA, Shimakawa A, Vu AT, Brau AC, Beatty PJ, Pineda AR, Brittain JH, Reeder SB. Multiecho reconstruction for simultaneous water-fat decomposition and T2* estimation. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2007 10;26(4):1153–61.

7. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB. Multiecho water-fat separation and simultaneous R estimation with multifrequency fat spectrum modeling. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2008 11;60(5):1122–34.

8. Haufe WM, Wolfson T, Hooker CA, Hooker JC, Covarrubias Y, Schlein AN, Hamilton G, Middleton MS, Angeles JE, Hernando D, Reeder SB. Accuracy of PDFF estimation by magnitude-based and complex-based MRI in children with MR spectroscopy as a reference. Journal of Magnetic Resonance Imaging. 2017 12;46(6):1641–7. [PubMed: 28323377]

9. Yu H, Shimakawa A, Hines CD, McKenzie CA, Hamilton G, Sirlin CB, Brittain JH, Reeder SB. Combination of complex-based and magnitude-based multiecho water-fat separation for accurate quantification of fat-fraction. Magnetic resonance in medicine. 2011 7;66(1):199–206. [PubMed: 21695724]

10. Hernando D, Hines CD, Yu H, Reeder SB. Addressing phase errors in fat-water imaging using a mixed magnitude/complex fitting method. Magnetic resonance in medicine. 2012 3;67(3):638–44. [PubMed: 21713978]

11. Tang A, Tan J, Sun M, Hamilton G, Bydder M, Wolfson T, Gamst AC, Middleton M, Brunt EM, Loomba R, Lavine JE. Nonalcoholic fatty liver disease: MR imaging of liver proton density fat fraction to assess hepatic steatosis. Radiology. 2013 5;267(2):422–31 [PubMed: 23382291]

12. Tang A, Desai A, Hamilton G, Wolfson T, Gamst A, Lam J, Clark L, Hooker J, Chavez T, Ang BD, Middleton MS. Accuracy of MR imaging–estimated proton density fat fraction for classification of dichotomized histologic steatosis grades in nonalcoholic fatty liver disease. Radiology. 2014 9 22;274(2):416–25 [PubMed: 25247408]

13. Heba ER, Desai A, Zand KA, Hamilton G, Wolfson T, Schlein AN, Gamst A, Loomba R, Sirlin CB, Middleton MS. Accuracy and the effect of possible subject-based confounders of magnitude-based MRI for estimating hepatic proton density fat fraction in adults, using MR spectroscopy as reference. Journal of Magnetic Resonance Imaging. 2016 2;43(2):398–406. [PubMed: 26201284]

14. Zand KA, Shah A, Heba E, Wolfson T, Hamilton G, Lam J, Chen J, Hooker JC, Gamst AC, Middleton MS, Schwimmer JB. Accuracy of multiecho magnitude-based MRI (M-MRI) for estimation of hepatic proton density fat fraction (PDFF) in children. Journal of Magnetic Resonance Imaging. 2015 11;42(5):1223–32. [PubMed: 25847512]

15. Joshi M, Dillman JR, Singh K, Serai SD, Towbin AJ, Xanthakos S, Zhang B, Su W, Trout AT. Quantitative MRI of fatty liver disease in a large pediatric cohort: correlation between liver fat fraction, stiffness, volume, and patient-specific factors. Abdominal Radiology. 2018 5 1;43(5):1168–79. [PubMed: 28828531]

16. Armstrong T, Ly KV, Murthy S, Ghahremani S, Kim GH, Calkins KL, Wu HH. Free-breathing quantification of hepatic fat in healthy children and children with nonalcoholic fatty liver disease using a multi-echo 3-D stack-of-radial MRI technique. Pediatric radiology. 2018 5 4:1–3.

17. Middleton MS, Heba ER, Hooker CA, Bashir MR, Fowler KJ, Sandrasegaran K, Brunt EM, Kleiner DE, Doo E, Van Natta ML, Lavine JE. Agreement between magnetic resonance imaging proton density fat fraction measurements and pathologist-assigned steatosis grades of liver biopsies from adults with nonalcoholic steatohepatitis. Gastroenterology. 2017 9 1;153(3):753–61. [PubMed: 28624576]

18. França M, Alberich-Bayarri A, Martí-Bonmatí L, Oliveira P, Costa FE, Porto G, Vizcaíno JR, Gonzalez JS, Ribeiro E, Oliveira J, Miranda HP. Accurate simultaneous quantification of liver

steatosis and iron overload in diffuse liver diseases with MRI. Abdominal Radiology. 2017 5 1;42(5):1434–43 [PubMed: 28110367]

19. Reeder Scott B., Hu Houchun H., and Sirlin Claude B.. “Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration.” Journal of magnetic resonance imaging 36.5 (2012): 1011–1014. [PubMed: 22777847]
20. Artz NS, Haufe WM, Hooker CA, Hamilton G, Wolfson T, Campos GM, Gamst AC, Schwimmer JB, Sirlin CB, Reeder SB. Reproducibility of MR-based liver fat quantification across field strength: Same-day comparison between 1.5 T and 3T in obese subjects. Journal of Magnetic Resonance Imaging. 2015 9;42(3):811–7 [PubMed: 25620624]
21. Kang BK, Yu ES, Lee SS, Lee Y, Kim N, Sirlin CB, Cho EY, Yeom SK, Byun JH, Park SH, Lee MG. Hepatic fat quantification: a prospective comparison of magnetic resonance spectroscopy and analysis methods for chemical-shift gradient echo magnetic resonance imaging with histologic assessment as the reference standard. Investigative radiology. 2012 6 1;47(6):368–75. [PubMed: 22543969]
22. Idilman IS, Aniktar H, Idilman R, Kabacam G, Savas B, Elhan A, Celik A, Bahar K, Karcaaltincaba M. Hepatic steatosis: quantification by proton density fat fraction with MR imaging versus liver biopsy. Radiology. 2013 6;267(3):767–75. [PubMed: 23382293]
23. Kühn JP, Hernando D, Muñoz del Rio A, et al. Effect of multipeak spectral modeling of fat for liver iron and fat quantification: correlation of biopsy with MR imaging results. Radiology. 2012;265(1):133–142 [PubMed: 22923718]
24. Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, Fujita K, Yoneda M, Taguri M, Hyogo H, Sumida Y. Magnetic resonance imaging more accurately classifies steatosis and fibrosis in patients with nonalcoholic fatty liver disease than transient elastography. Gastroenterology. 2016 3 1;150(3):626–37 [PubMed: 26677985]
25. [accessed in December 15, 2018] https://www.cdc.gov/obesity/adult/defining.html
26. Mangus RS, Kubal CA, Fridell JA, Pena JM, Frost EM, Joseph Tector A. Alcohol abuse in deceased liver donors: impact on post-transplant outcomes. Liver International. 2015 1;35(1):171–5. [PubMed: 24512023]
27. Kleiner DE, Brunt EM, Natta MV, et al. Design and Validation of a Histological Scoring System for Nonalcoholic Fatty Liver Disease. Hepatology. 2005;41(6):1313–21 [PubMed: 15915461]
28. Kang GH, Cruite I, Shiehmorteza M, Wolfson T, Gamst AC, Hamilton G, Bydder M, Middleton MS, Sirlin CB. Reproducibility of MRI-determined proton density fat fraction across two different MR scanner platforms. Journal of magnetic resonance imaging. 2011 10 1;34(4):928–34. [PubMed: 21769986]
29. Lazo M, Hernaez R, Eberhardt MS, Bonekamp S, Kamel I, Guallar E, Koteish A, Brancati FL, Clark JM. Prevalence of nonalcoholic fatty liver disease in the United States: the Third National Health and Nutrition Examination Survey, 1988–1994. American journal of epidemiology. 2013 5 23;178(1):38–45. [PubMed: 23703888]
30. Qureshi K, Abrams GA. Prevalence of biopsy-proven non-alcoholic fatty liver disease in severely obese subjects without metabolic syndrome. Clinical obesity. 2016 4;6(2):117–23. [PubMed: 26856683]
31. Permutt Z, Le TA, Peterson MR, Seki E, Brenner DA, Sirlin C, Loomba R. Correlation between liver histology and novel magnetic resonance imaging in adult patients with non-alcoholic fatty liver disease–MRI accurately quantifies hepatic steatosis in NAFLD. Alimentary pharmacology & therapeutics. 2012 7 1;36(1):22–9. [PubMed: 22554256]
32. Chalasani N, Wilson L, Kleiner DE, Cummings OW, Brunt EM, Ünalp A, NASH Clinical Research Network. Relationship of steatosis grade and zonal location to histological features of steatohepatitis in adult patients with non-alcoholic fatty liver disease. Journal of hepatology. 2008 5 1;48(5):829–34. [PubMed: 18321606]
33. de Graaf EL, Kench J, Dilworth P, Shackel NA, Strasser SI, Joseph D, Pleass H, Crawford M, McCaughan GW, Verran DJ. Grade of deceased donor liver macrovesicular steatosis impacts graft and recipient outcomes more than the Donor Risk Index. Journal of gastroenterology and hepatology. 2012 3;27(3):540–6. [PubMed: 21777274]

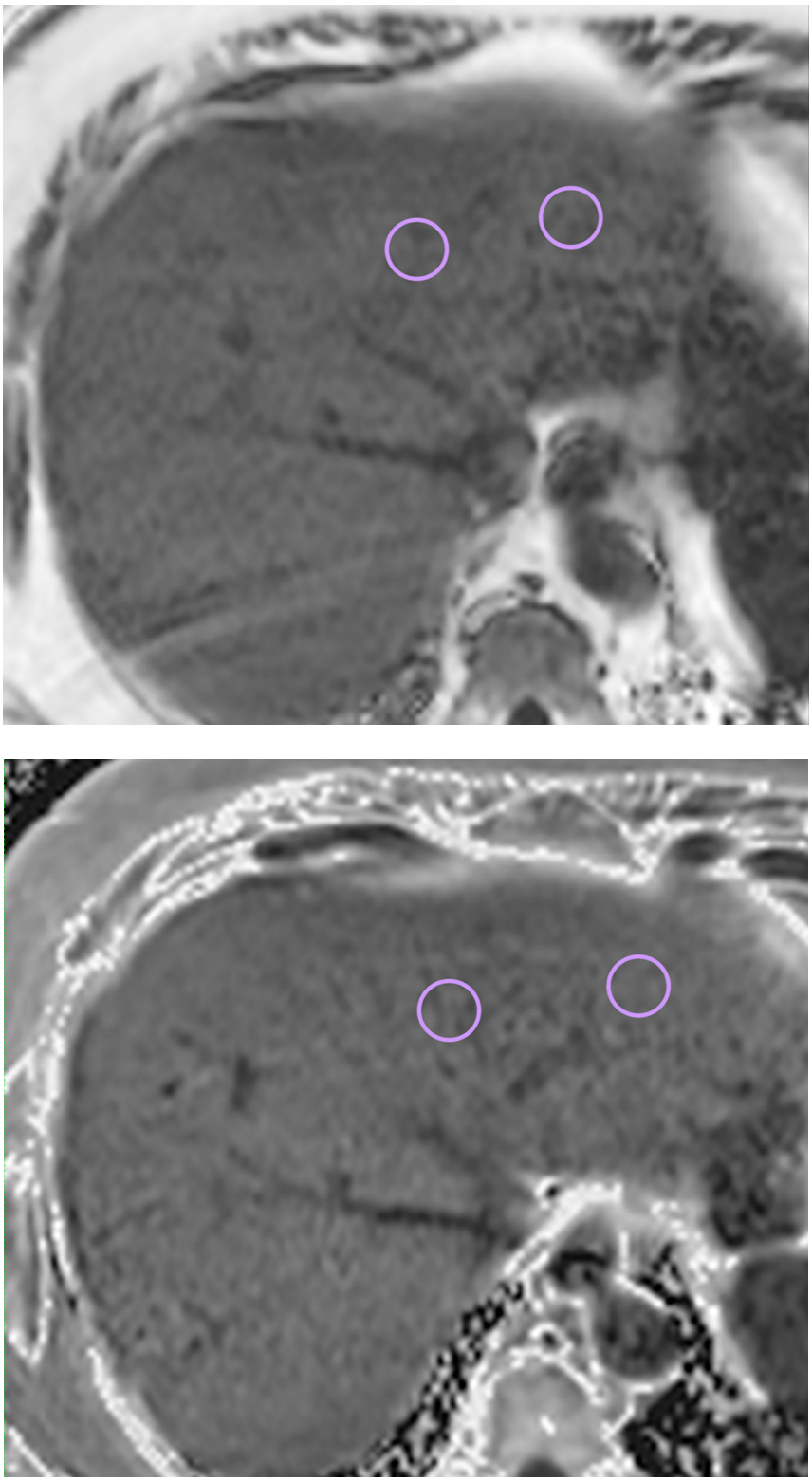
34. Szczepaniak LS, Nurenberg P, Leonard D, Browning JD, Reingold JS, Grundy S, Hobbs HH, Dobbins RL. Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. American Journal of Physiology-Endocrinology and Metabolism. 2005 2;288(2):E462–8 [PubMed: 15339742]

35. Hernando D, Liang ZP, Kellman P. Chemical shift–based water/fat separation: a comparison of signal models. Magnetic resonance in medicine. 2010 9;64(3):811–22. [PubMed: 20593375]

36. Horng Debra E., et al. “Comparison of R2* correction methods for accurate fat quantification in fatty liver.” Journal of Magnetic Resonance Imaging 37.2 (2013): 414–422. [PubMed: 23165934]

37. Nasr P, Forsgren MF, Ignatova S, Dahlström N, Cedersund G, Leinhard OD, Norén B, Ekstedt M, Lundberg P, Kechagias S. Using a 3% proton density fat fraction as a cut-off value increases sensitivity of detection of hepatic steatosis, based on results from histopathology analysis. Gastroenterology. 2017 7 1;153(1):53–5. [PubMed: 28286210]

38. Schwimmer JB, Middleton MS, Behling C, Newton KP, Awai HI, Paiz MN, Lam J, Hooker JC, Hamilton G, Fontanesi J, Sirlin CB. Magnetic resonance imaging and liver histology as biomarkers of hepatic steatosis in children with nonalcoholic fatty liver disease. Hepatology. 2015 6 1;61(6):1887–95 [PubMed: 25529941]



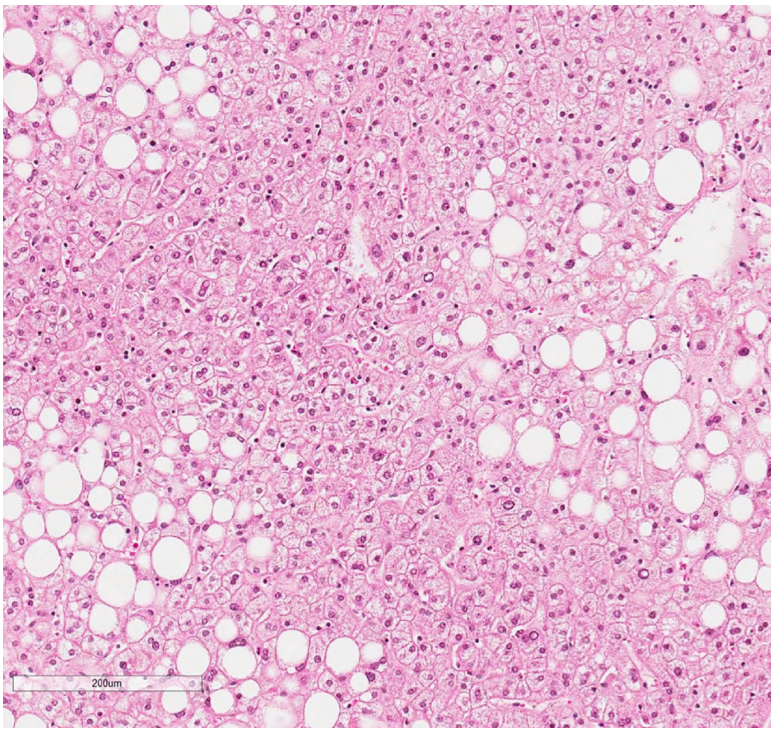


Figure 1.

(a, b) Representative MRI-C and MRI-M-based PDFF images with a one-centimeter ROIs placed over segments II and IVa. Average left lobe PDFF estimates were 11.9% and 11.2%, respectively. (c) Corresponding histology from core biopsy shows the presence of steatosis (grade 1).

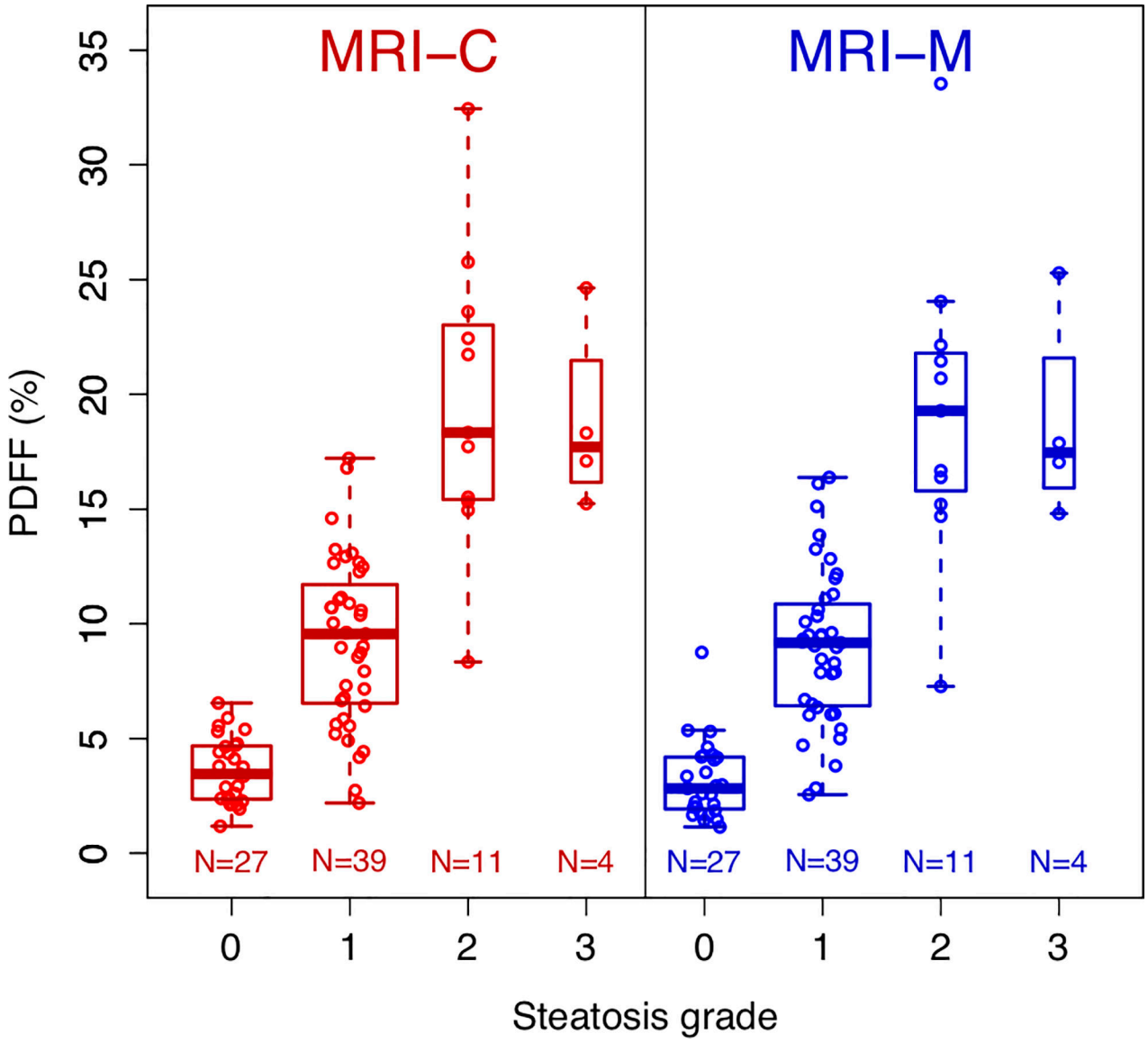


Figure 2:
Boxplot shows MRI-based PDFF values and corresponding histology-determined statosis grades for both CSE-MRI techniques.

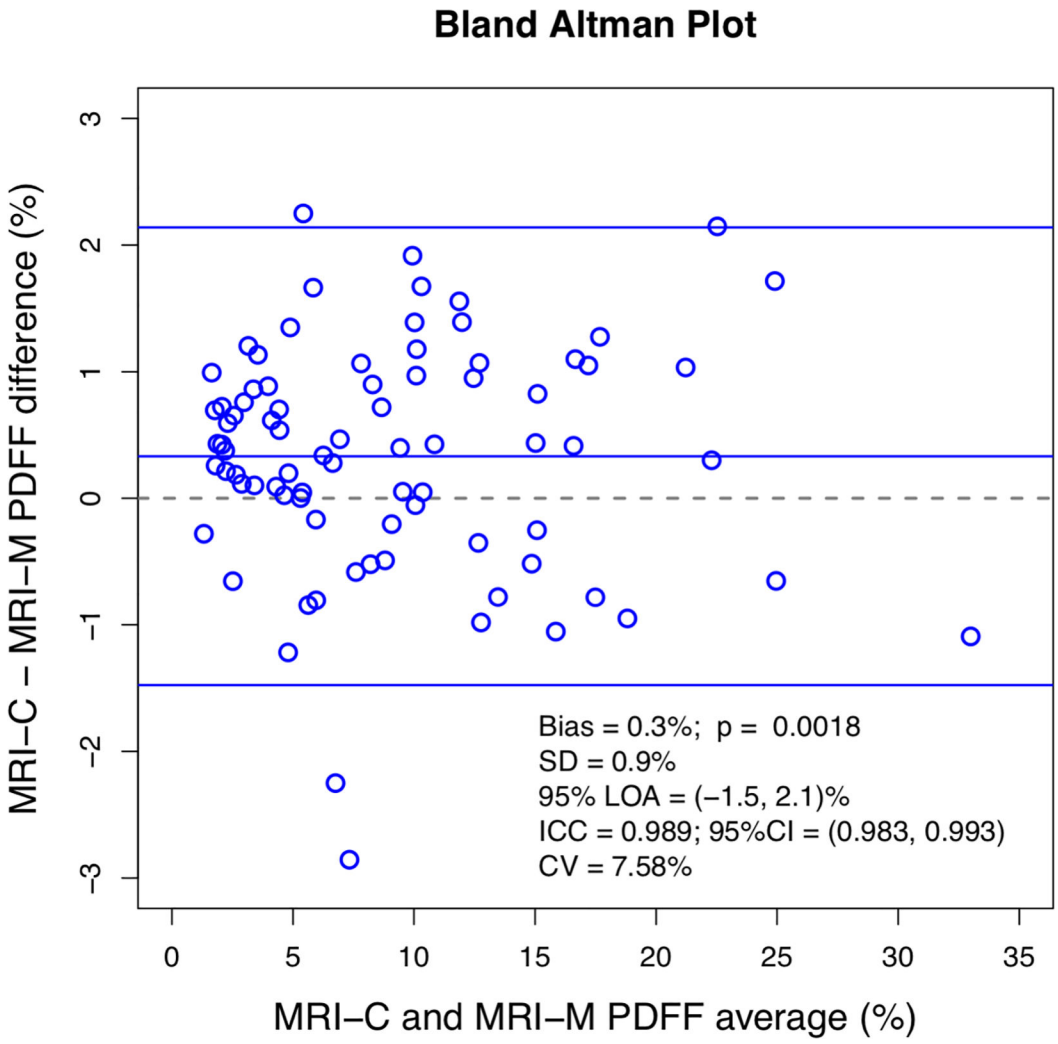


Figure 3.
Bland-Altman analysis was used to evaluate agreement and differences between MRI-M and MRI-C.

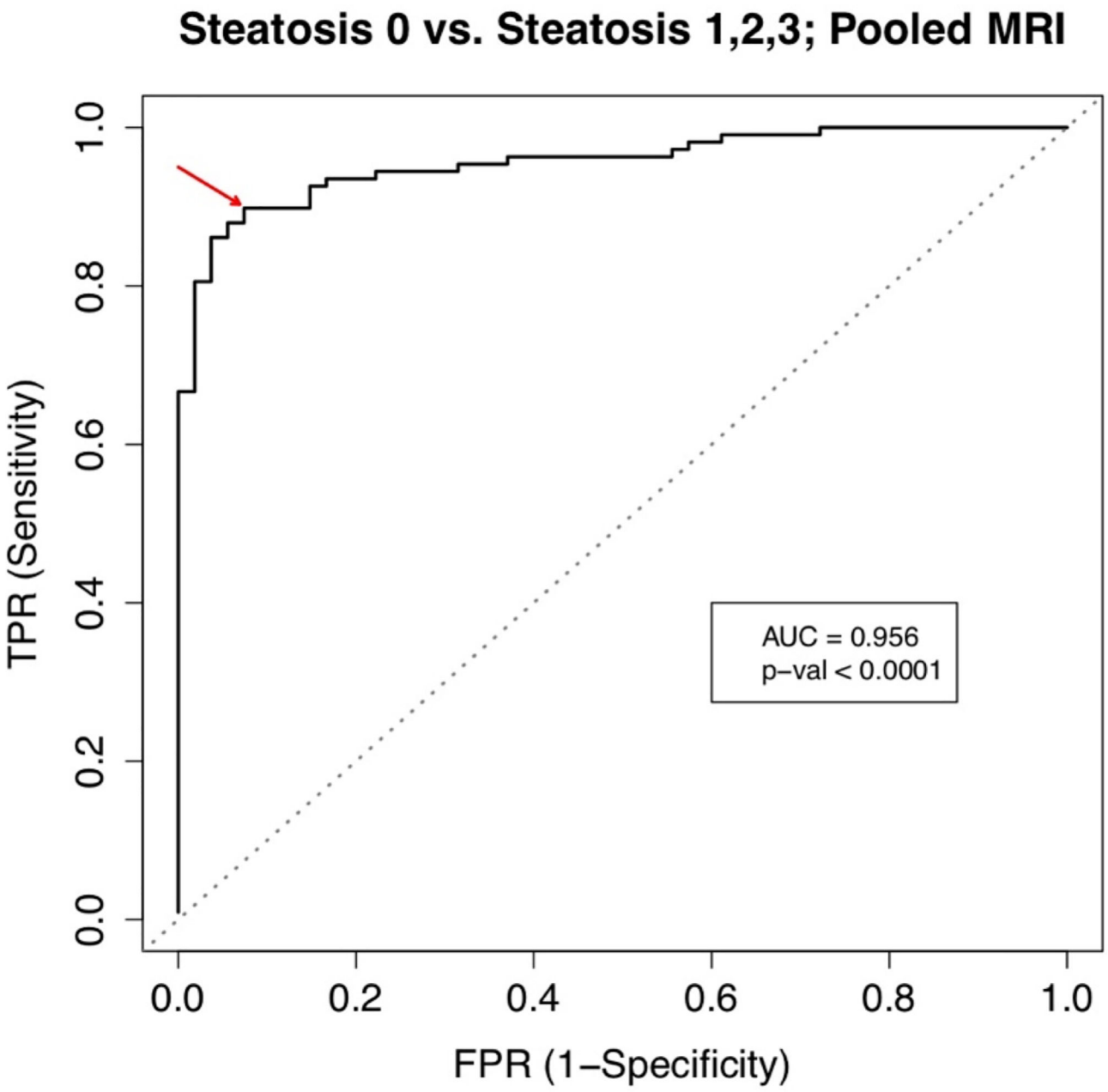


Figure 4a.
ROC curve for the diagnosis of hepatic steatosis (grades 0 vs ≥1) using pooled data. 5.4% threshold location is shown with the red arrow.

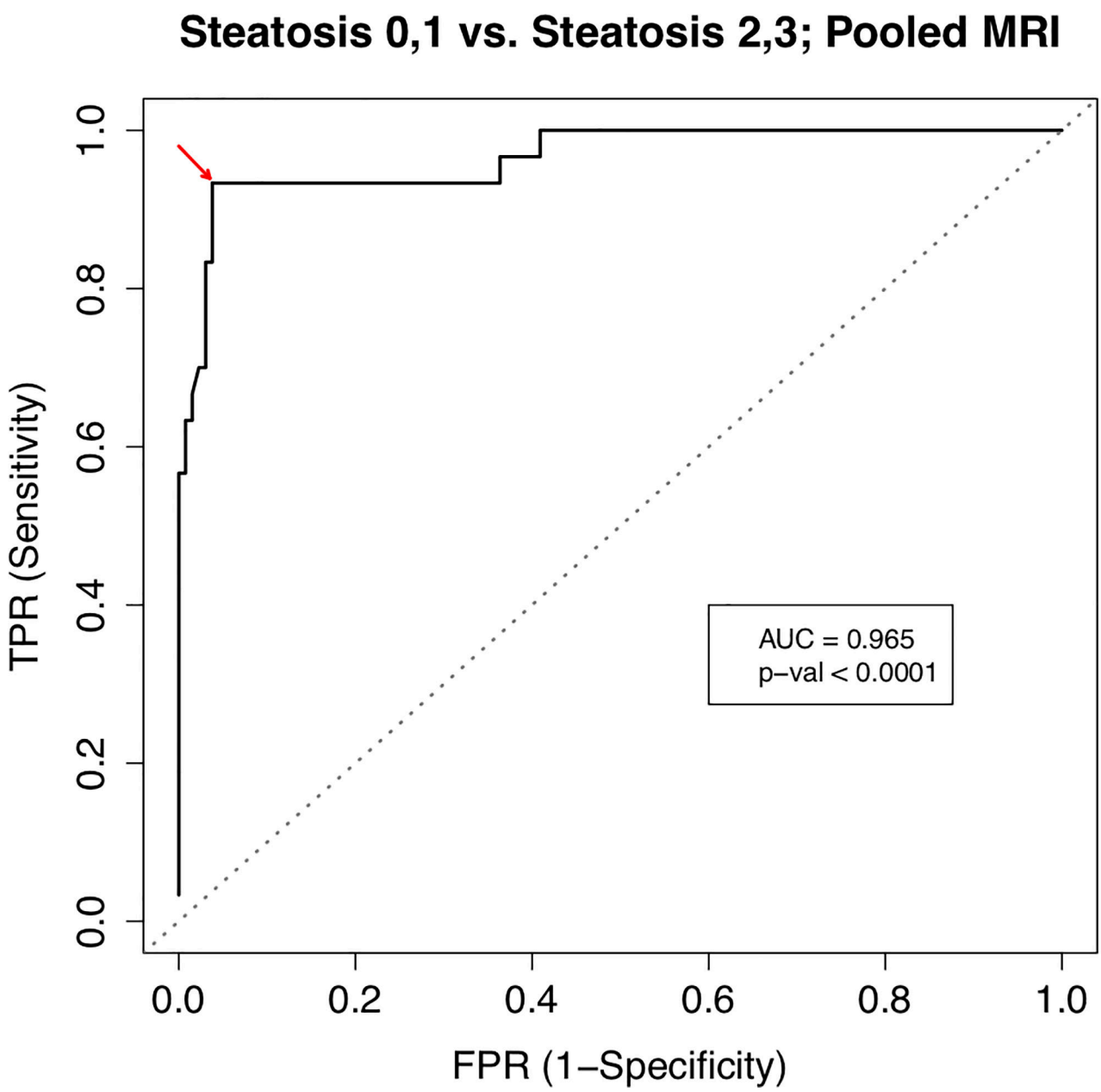


Figure 4b.
ROC curve for differentiating none to mild (grades 0–1) from moderate to severe (grades 2–3) hepatic steatosis using pooled data. 14.7% threshold location is shown with the red arrow.

Table 1:

Key acquisition parameters for magnitude-based magnetic resonance imaging (MRI-M) and complex-based magnetic resonance imaging (MRI-C) at 3.0T and 1.5T. TE = echo time; TR = repetition time.

	Field strength	TE (ms)	TR (ms)	Flip angle (°)	Bandwidth (kHz)	Slice thickness (mm)	Matrix	Number of slices	Length of single breath hold (s)
MRI-M	3.0T	1.15, 2.3, 3.45, 4.6, 5.75, 6.9	120–170	10	±125	8	192–256 X 160–256	8 – 20	12 – 34
	1.5T	2.3, 4.6, 6.9, 9.2, 11.5, 13.8	170	10	±83	8	192–256 X 160–256	8 – 20	12 – 34
MRI-C	3.0T	1.2, 2.2, 3.2, 4.2, 5.2, 6.2	8.6	3	±125	8	256 X 128	32	20
	1.5T	1.2, 3.2, 5.2, 7.2, 9.2, 11.2	13.4	5	±125	8	256 × 160	32	20

Table 2:

Cohort (n=81) demographic and liver histology characteristics.

Characteristic	Value
Sex	
Female	67 (82.7%)
Male	14 (17.3%)
Race	
Black	2 (2.5%)
Other	8 (9.9%)
White	71 (87.7%)
Ethnicity	
Hispanic/Latino	12 (14.8%)
Non-Hispanic/Latino	69 (85.2%)
Age (years)	48.2 ± 12.5 (23.7 – 70.6) *
Mean body mass index (kg/m^2)	41.7 ± 5.5 (32.5 – 56.8) *
Waist-to-Hip Ratio	0.89 ± 0.10 (0.74 – 1.19) *
Steatosis grade **	
0; <5% hepatocytes	27 (33.3%)
1; 5%-33% hepatocytes	39 (48.1%)
2; 33%-66% hepatocytes	11 (13.6%)
3; >66% hepatocytes	4 (4.9%)
Lobular inflammation ***	
0; no foci	48 (59.3%)
0.5	7 (8.6%)
1; <2 foci per 200x field	19 (23.5%)
1.5	3 (3.7%)
2; 2–4 foci per 200x field	4 (4.9%)
3; >4 foci per 200x field	0 (0%)
Portal Inflammation ***	
0; none	47 (58%)
0.5	13 (16%)
1; mild	18 (22.2%)
1.5	3 (3.7%)
2; moderate	0 (0%)
3; marked	0 (0%)
Hepatocellular ballooning ***	
0; none	65 (80.2%)
0.5	1 (1.2%)

Characteristic	Value
1; few balloon cells	10 (12.3%)
1.5	1 (1.2%)
2; many cells/prominent ballooning	4 (4.9%)
Iron grade ***	
0	44 (54%)
0.5	8 (10%)
1	13 (16%)
1.5	4 (5%)
2	5 (6%)
3	2 (3%)
N/A †	5 (6%)
Fibrosis stage ***	
0; no fibrosis	33 (41.2%)
0.5	7 (8.8%)
1; perisinusoidal or periportal	22 (27.5%)
1.5	10 (12.5%)
2; perisinusoidal and periporal	2 (2.5%)
2.5	1 (1.2%)
3; bridging fibrosis	4 (5.0%)
4; cirrhosis	1 (1.2%)
NA †	1 (1.2%)
NAS ***‡	
≤ 3	68 (84%)
3.5 – 4.5	9 (11%)
≥ 5	4 (5%)
Mean MR-determined PDFF value (%)	
MRI-M	8.9 ± 6.5 (1.1 —33.5) *
MRI-C	9.2 ± 6.5 (1.1—32.4) *

Note—

* Data are averages ± standard deviations, with ranges in parentheses.

** A 4-point ordinal steatosis score was derived from the granular steatosis score or, for patients with both wedge and core biopsies, from the average granular steatosis score.

*** Half grades/stages are due to averaging results from wedge and core biopsies.

† No data available for these patients

‡ NAS stands for the NAFLD Activity Score, which is the unweighted sum of steatosis, lobular inflammation, and hepatocellular ballooning.

Table 3:

Summary of raw and cross-validated performance metrics of the magnitude-based magnetic resonance imaging (MRI-M-) and complex-based magnetic resonance imaging (MRI-C-) derived PDFF at the common threshold for diagnosing hepatic steatosis (grades ≥1).

Threhold = 5.4%	MRI-M		MRI-C		p-value
	(raw)	(cross-validated)	(raw)	(cross-validated)	
Sensitivity	0.90 (0.79 – 0.96)	0.88 (0.77 – 0.95)	0.88 (0.77 – 0.95)	0.87 (0.75 – 0.94)	0.99
Specificity	0.96 (0.81 – 0.99)	0.92 (0.75 – 0.99)	0.85 (0.66 – 0.95)	0.81 (0.61 – 0.93)	0.24
PPV	0.98 (0.89 – 0.99)	0.96 (0.86 – 0.99)	0.92 (0.81 – 0.97)	0.90 (0.79 – 0.96)	-
NPV	0.83 (0.66 – 0.94)	0.80 (0.62 – 0.92)	0.79 (0.60 – 0.92)	0.75 (0.56 – 0.92)	-
Accuracy	0.92 (0.84 – 0.97)	0.90 (0.81 – 0.95)	0.87 (0.78 – 0.93)	0.85 (0.75 – 0.92)	0.22

Note: Numbers in parentheses are 95% confidence intervals. P-values relates to difference between performance parameters for classifying steatosis 0 vs steatosis ≥ 1. PPV and NPV differences were not computed.

Table 4:

Summary of raw and cross-validated performance metrics of the magnitude-based magnetic resonance imaging (MRI-M-) and complex-based magnetic resonance imaging (MRI-C-) derived PDFF at the common threshold for differentiating none to mild (grades 0–1) from moderate to severe (grade 2–3) steatosis.

Threshold = 14.7%	MRI-M		MRI-C		p-value
	(raw)	(cross-validated)	(raw)	(cross-validated)	
Sensitivity	0.86 (0.59 – 0.98)	0.80 (0.51 – 0.95)	0.93 (0.68 – 0.99)	0.93 (0.68 – 0.99)	0.48
Specificity	0.95 (0.87 – 0.99)	0.95 (0.87 – 0.99)	0.97 (0.89 – 0.99)	0.97 (0.89 – 0.99)	0.99
PPV	0.81 (0.54 – 0.96)	0.80 (0.51 – 0.95)	0.87 (0.61 – 0.98)	0.87 (0.61 – 0.98)	-
NPV	0.93 (0.86 – 0.98)	0.95 (0.87 – 0.99)	0.98 (0.91 – 1.00)	0.98 (0.91 – 1.00)	-
Accuracy	0.93 (0.86 – 0.98)	0.92 (084 – 0.97)	0.96 (0.89 – 0.99)	0.96 (0.89 – 0.99)	0.24

Note: Numbers in parentheses are 95% confidence intervals. P-values relates to difference between performance parameters for classifying steatosis grades 0–1 vs steatosis grades 2–3. PPV and NPV differences were not computed.