HHS Public Access

Author manuscript

J Magn Reson Imaging. Author manuscript; available in PMC 2016 November 01.

Published in final edited form as:

J Magn Reson Imaging. 2015 November ; 42(5): 1223–1232. doi:10.1002/jmri.24888.

Accuracy of Multi-echo Magnitude-based MRI (M-MRI) for Estimation of Hepatic Proton Density Fat Fraction (PDFF) in Children

Kevin A. Zand, MD[1], **Amol Shah, MD**[1], **Elhamy Heba, MD**[1], **Tanya Wolfson, MA**[2], **Gavin Hamilton, PhD**[1], **Jessica Lam, BS**[1], **Joshua Chen, BS**[1], **Jonathan C. Hooker, BS**[1], **Anthony C. Gamst, PhD**[2], **Michael S. Middleton, MD PhD**[1], **Jeffrey B. Schwimmer, MD**[1,3,4], and **Claude B. Sirlin, MD**[1]

[1]Liver Imaging Group, Department of Radiology, University of California, San Diego, CA, USA

[2]Computational and Applied Statistics Laboratory, Division of Biostatistics and Informatics, University of California, San Diego, CA, USA

[3]Division of Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of California, San Diego, CA, USA

[4]Department of Gastroenterology, Rady Children's Hospital San Diego, CA, USA

Abstract

Purpose—To assess accuracy of magnitude-based magnetic resonance imaging (M-MRI) in children to estimate hepatic proton density fat fraction (PDFF) using two to six echoes, with magnetic resonance spectroscopy (MRS)-measured PDFF as a reference standard.

Materials and Methods—This was an IRB-approved, HIPAA-compliant, single-center, cross-sectional, retrospective analysis of data collected prospectively between 2008 and 2013 in children with known or suspected non-alcoholic fatty liver disease (NAFLD). Two hundred and eighty-six children (8 – 20 [mean 14.2 ± 2.5] yrs; 182 boys) underwent same-day MRS and M-MRI. Unenhanced two-dimensional axial spoiled gradient-recalled-echo images at six echo times were obtained at 3T after a single low-flip-angle (10°) excitation with ≥ 120-ms recovery time. Hepatic PDFF was estimated using the first two, three, four, five, and all six echoes. For each number of echoes, accuracy of M-MRI to estimate PDFF was assessed by linear regression with MRS-PDFF as reference standard. Accuracy metrics were regression intercept, slope, average bias, and R^2.

Results—MRS-PDFF ranged from 0.2 – 40.4% (mean 13.1 ± 9.8%). Using three to six echoes, regression intercept, slope, and average bias were 0.46 – 0.96%, 0.99 – 1.01, and 0.57 – 0.89%, respectively. Using two echoes, these values were 2.98%, 0.97, and 2.72%, respectively. R^2 ranged 0.98 – 0.99 for all methods.

Corresponding author: Claude B. Sirlin, Tel: 619-471-0776, csirlin@ucsd.edu.
Institution address: University of California, San Diego, Department of Radiology, 200 W Arbor Dr., San Diego, CA 92103.

Conclusion—Using three to six echoes, M-MRI has high accuracy for hepatic PDFF estimation in children.

Keywords

MRI; MRS; PDFF; proton density fat fraction; NAFLD; liver steatosis

INTRODUCTION

Hepatic proton density fat fraction (PDFF) is an MR-based, standardized, reproducible biomarker of liver fat content (1). It has been applied as an endpoint in clinical trials (2–5) and its successful implementation into clinical care is an important goal. MR spectroscopy (MRS), an accepted non-invasive reference standard for PDFF measurement (6), is not widely available and samples only portions of the liver, which may introduce sampling variability. To address the limitations of MRS for clinical and research use, advanced MRI methods have been developed to estimate PDFF non-invasively across the entire liver. These MRI methods can be broadly classified as magnitude-based (M-MRI) and complex-based (C-MRI)(1,7). Both methods use low flip angle to minimize T1 weighting, acquire multiple echoes to measure and permit correction for T2* decay, and incorporate into their mathematical model the multi-frequency interference effects of protons in fat (7–10). C-MRI methods also require correction for eddy currents and noise bias; these factors have negligible effects on and usually require no correction with M-MRI methods (7, 10–11).

Several studies in adults have shown that MRI-estimated PDFF (MRI-PDFF) agrees closely with MRS-measured PDFF (MRS-PDFF), suggesting that MRI-based methods for hepatic PDFF estimation may be used in adults instead of MRS (12–17). However, independent assessment of MRI-PDFF in children is needed for two reasons. First, fatty liver disease prevalence in children is high (18–20), and so reliable methods to non-invasively assess fatty liver disease in children are needed. Ultrasound has been used to assess hepatic steatosis (21, 22), but a recent systematic review has shown that it has limited accuracy to diagnose and grade pediatric hepatic steatosis (23). Second, children are different from adults in important characteristics relevant to MRI including ability to tolerate MRI examinations, breath-holding capacity, and body habitus. These characteristics influence the feasibility, technical optimization, and quality of MRI and may affect its diagnostic performance (23, 24).

In addition to validating MRI-PDFF in children, another need is to determine the optimal number of echoes for PDFF estimation using M-MRI. Although most studies on M-MRI have evaluated a prototype sequence with six echoes, a recent study suggested that using fewer echoes might be of similar or higher accuracy (17). Using fewer echoes would reduce the duration of the echo train for any given set of MR acquisition parameters, and would permit more slices to be acquired and hence provide greater liver coverage. Accordingly, if fewer echoes provide similar or greater MRI-PDFF accuracy than the prototype, then the use of fewer echoes could be recommended for future implementation.

The purpose of this cross-sectional study was to assess the accuracy of M-MRI at 3T, over a range of numbers of echoes from two to six, for estimation of hepatic PDFF in children

using MRS-PDFF as a reference standard. We focused on M-MRI rather C-MRI because C-MRI methods are not yet widely available as they require specialized sequences; by comparison, M-MRI can be implemented using any multi-echo spoiled gradient-recalled-echo (SPGR) sequences and currently are more feasible for use in multi-center clinical trials.

MATERIALS AND METHODS

This was a single-center, cross-sectional, retrospective analysis of data collected prospectively in children at our imaging center. Inclusion criteria were children aged 8 to 20 yrs with known or suspected non-alcoholic fatty liver disease (NAFLD) who underwent research liver MR examinations at our center between August 2008 and June 2013, and had same-day MRS and M-MRI. The study was approved by an Institutional Review Board and was compliant with the Health Insurance Portability and Accountability Act. Children aged 8 to 17 years provided written informed assent with written informed consent by their parent(s)/guardian(s); subjects aged 18 to 20 years provided written informed consent. Demographic information was collected on all children. Anthropometric data was collected in a subset of children. In this study, children were considered obese for subjects with body mass index (BMI) Z-score ≥ 2. Histology data was collected in the subset of children with available liver biopsy results.

MR Examination Acquisition

All children were scanned on a single 3T scanner (GE, Signa EXCITE HDxt, GE Medical Systems, Milwaukee, WI) in supine position with an eight-channel torso phased-array coil centered over the liver. A dielectric pad was placed between the coil and the abdomen. Children were asked to fast for a minimum of four hours prior to the MR examination to reduce potential confounding effects of meals on fat quantification. No sedatives/anxiolytics were administered. Each MR examination included MRS and M-MRI fat quantification acquisitions.

MRS—MRS was performed using a previously described technique (12, 16). We selected this particular technique because it has been used as a reference standard in numerous prior studies (12–17) and, as described below, it can be acquired in a single-breathold, which facilitates registration with the breath-hold M-MRI sequence. Using three-plane localizing images, a $20 \times 20 \times 20$ mm^3 voxel was placed in the right lobe of the liver, avoiding major vessels, bile ducts, liver edges, and artifacts. To ensure a uniform spectral response, spatial saturation bands around the voxel were disabled (25). After automatic shimming, single-voxel proton MRS was performed in a 21-s breath-hold using stimulated echo acquisition mode (STEAM). Repetition time (TR) was set at 3,500 ms with the lowest allowed mixing time of 5 ms to minimize T1 weighting. Five STEAM spectra were collected at equally spaced TEs between 10 and 30 ms in a single acquisition. This range of TEs permits robust T2 correction with minimal confounding j-coupling effects (26). Spectra from the individual channels were combined using singular value decomposition (27). An anatomic image depicting the MRS voxel location was saved; as described later, this anatomic image was used to co-localize PDFF measured by MRS with PDFF estimated by M-MRI. Spectra were transferred offline for analysis.

MRI—As previously described (12, 28), unenhanced two-dimensional SPGR axial images of the liver were obtained with low flip angle (10°) and TR ≥ 120 ms to minimize T1 weighting. Six co-localized SPGR magnitude images were obtained after a single excitation at serial echo times in which protons in water and in the dominant methylene peak of triglyceride are nominally out of, or in phase (1.15, 2.3, 3.45, 4.6, 5.75, and 6.9 ms) to permit fat-water signal separation and correction for T2*. Other imaging parameters included 8 – 10 mm section thickness; 0% intersection gap; ± 142-kHz receiver bandwidth; 192 × 192 base matrix; and rectangular field of view. Complete liver coverage required one to two breath-held acquisitions for each child. Images were transferred to a picture archiving and communication system for storage and analysis.

MR Examination Analysis

MRS—MR spectra were analyzed by an MR physicist (GH; > 12 yrs experience, blinded to clinical/MRI data) using Advanced Method for Accurate, Robust, and Efficient Spectral (AMARES) fitting (29, 30). Water (4.7 ppm) and the visible fat (2.1, 1.3, 0.9 ppm) peaks were modeled as multiple Gaussian resonances and measured. T2-corrected peak areas were calculated by non-linear least squares fitting of peak areas across echo times. Obscured fat peaks (4.2 and 5.3 ppm) were estimated from the visible fat peaks (16, 31). MRS-PDFF was calculated as the integrated sum of T2-corrected fat peaks divided by the sum of T2-corrected fat and water peaks (16).

MRI—Three circular regions of interest (ROIs), each with a 1 cm radius, were manually placed by one of three image analysts (JL, JC, JH; each with > 3 mon training, blinded to clinical/MRS data) under the supervision of a research fellow and a senior author (KZ, CS, blinded to clinical/MRS data) on three adjacent slices of the fifth-echo image of the SPGR sequence, co-localized to the superior, middle, and inferior portions of the MRS voxel as depicted on the saved anatomic image mentioned earlier. The fifth echo was used for ROI placement because it consistently provides adequate anatomic delineation for this purpose (32). The ROIs then were propagated to all other echoes (first, second, third, fourth, and sixth) of that sequence. The ROI signal intensity data was analyzed with a custom-developed fitting algorithm (12, 28) that addresses the multi-frequency interference effects of protons in fat by modeling the composite fat signal as a weighted sum of six frequency signals at 0.9, 1.3, 2.1, 2.75, 4.2, and 5.3 ppm, with normalized weights of 0.088, 0.7, 0.12, 0.006, 0.039, and 0.047, respectively, and the water signal as a single frequency signal at 4.7 ppm. Using this algorithm, five PDFF estimations were computed for each acquisition using the first two, three, four, five, or six echoes (two-, three-, four-, five-, or six-echo M-MRI-PDFF, respectively). Three- to six-echo M-MRI-PDFF estimations (but not two-echo) were corrected for T2* decay, assuming mono-exponential decay (8). For illustrative purposes, PDFF maps were generated from source images by applying the algorithm pixel-by-pixel (Figure 1).

Statistical Analysis

A biostatistical analyst performed the statistical analysis under the supervision of a faculty biostatistician with statistical computing software R (R Foundation for Statistical Computing, Vienna, Austria, version 3.0.2 (2013)).

Accuracy of two- to six-echo M-MRI-PDFF was assessed using a series of five univariate linear regressions (one for each M-MRI method), modeling MRS-PDFF (reference standard) as a function of M-MRI-PDFF. As done in prior studies (16, 17), four regression parameters served as metrics of accuracy: regression intercept, slope, average bias of the fitted regression (defined as the square root of the average squared difference between the regression line and the MRS = MRI line), and R^2. Bootstrap-based 95% confidence intervals were calculated for each parameter estimate. Bland-Altman plots for MRS-PDFF and M-MRI-PDFF were generated. Multivariate models were explored to assess the confounding effects of age, sex, and BMI on M-MRI and MRS agreement. The regression parameters of two- to five-echo M-MRI were compared to the prototypical six-echo M-MRI using bootstrap-based tests. Bonferroni adjustment for multiple comparisons was applied: a pairwise comparison was considered significant at 0.05 family significance level if the individual p-value was less than 0.003. Although the major purpose of this study was to analyze agreement between various M-MRI-PDFF estimations and MRS, we also performed pairwise comparisons (using paired t-tests) of the T2* values estimated by each method, since the T2* values are estimated automatically as part of the fitting algorithm.

Subanalyses were limited to children with recorded BMIs. Using both the M-MRI method with the best agreement to MRS and the prototypical six-echo method, the accuracy of M-MRI-PDFF was assessed separately for four subgroups of children stratified by sex and presence/absence of obesity: non-obese boys, obese boys, non-obese girls, and obese girls. The same regression parameters as in the main analysis were used as metrics of accuracy. For both methods, average bias and R^2 were compared among sex/obesity subgroups using bootstrap-based tests. Bonferroni adjustment was not applied to these comparisons, as no differences were expected between the subgroups, and we sought to maximize the ability to detect unexpected differences. P-values < 0.05 were considered significant.

Additionally, correlation analyses were performed between histology steatosis grade and hepatic PDFF estimated by two- to six-echo M-MRI or MRS in the subset of children with available histology results.

RESULTS

Two hundred and eighty-six children (182 boys, 104 girls) were included in this retrospective analysis. The children ranged in age from 8 to 20 yrs (mean 14.2 ± 2.5 yrs), and in MRS-PDFF from 0.2 to 40.4% (mean 13.1 ± 9.8%). Two hundred and twenty-three children had their BMIs recorded with BMI Z-score range of −3.3 to 3.1 (mean 2.0 ± 0.8). Among those 223 subjects, 100 were obese boys, 43 non-obese boys, 36 obese girls, and 44 non-obese girls. Ten children (10 – 18 [mean 13.4 ± 2.5] yrs; 6 boys, 4 girls) had liver biopsy with available research scores of hepatic steatosis grades. Other children either had no biopsy (n=234) or had biopsies performed at other institutions for which we were not granted access. Cohort characteristics are summarized in Table 1.

Figure 2 presents the scatterplots for MRS-PDFF versus M-MRI-PDFF for each method. Over the range of three- to six-echo M-MRI, the regression intercept, slope, and average bias ranged from 0.46 to 0.96%, 0.99 to 1.01, and 0.57 to 0.89%, respectively. Two-echo M-

MRI showed regression intercept, slope, and average bias of 2.98%, 0.97, and 2.72%, respectively. R^2 ranged from 0.98 to 0.99 for all M-MRI methods (Table 2). The MRS-PDFF vs. M-MRI-PDFF regression parameters for all echoes and the statistical significance of pairwise comparisons for two-, three-, four-, and five-echo M-MRI versus the prototypical six-echo M-MRI method are summarized in Table 2. Except for those involving R^2, all pairwise comparisons were significant. Three-, four-, and five-echo M-MRI had regression intercepts significantly closer to 0%, regression slopes significantly closer to 1, and average biases significantly smaller than six-echo M-MRI (difference in intercepts had a range of −0.50 to −0.19%, slopes: 0.004 to 0.014, average biases: −0.32 to −0.13%). All these differences indicate better agreement with MRS-PDFF of the three-, four-, and five-echo M-MRI methods than six-echo M-MRI. Two-echo M-MRI had a regression intercept significantly further from 0%, a regression slope significantly further from 1, and an average bias significantly higher than six-echo M-MRI, indicating worse agreement with MRS-PDFF (difference in intercept: 2.02%, slope: −0.021, average bias: 1.83%).

Bland-Altman plots for all M-MRI methods are shown in Figure 3. The mean of MRS-PDFF and M-MRI-PDFF differences (the Bland-Altman bias) ranged from 0.57 to 0.88% for three- to six-echo M-MRI and was 2.7% for two-echo M-MRI. Multivariate models demonstrated that sex and BMI (but not age) had a significant effect on the M-MRI and MRS relationship.

T2* values estimated by each M-MRI method are summarized in Table 3. T2* values estimated by the 3-echo M-MRI method were the highest, whereas T2* values estimated by the other methods were more closely clustered. Nevertheless, all pairwise comparisons were significant ($P<0.0001$).

Only three-, and six-echo M-MRI methods were included in the subanalyses. As demonstrated in the multivariate analyses, age was not found to be a modifier of the MRS-MRI relationship, thus the subanalyses examined sex and obesity subgroups only. For both methods, average bias was higher (at unadjusted 0.05 significance level) in obese boys than in the other sex/obesity subgroups. R^2 was lower (at unadjusted 0.05 significance level) in obese boys than in non-obese boys (both methods) and obese girls (six-echo method)(Table 4). The differences in average bias and R^2 ranged from −0.59 to 0.70%, and −0.12 to 0.11, respectively.

Figure 4 presents correlation analyses between histology steatosis grade and hepatic PDFF estimated by two- to six-echo M-MRI or MRS in 10 children with available histology results.

DISCUSSION

We performed a cross-sectional retrospective analysis of same-day M-MRI and MRS data acquired over a five-year period in 286 children with known or suspected NAFLD. The main purpose was to assess the accuracy of M-MRI at 3T, over a range of numbers of echoes from two to six, for estimation of hepatic PDFF in children using MRS-PDFF as a reference standard.

We found that, for estimations using three or more echoes, M-MRI-PDFF agreed closely with MRS-PDFF, with regression intercepts close to 0%, regression slopes and R^2 close to 1, and average bias less than 1%. Compared to the agreement between the prototypical six-echo M-MRI-PDFF and MRS-PDFF, agreement was higher using five, four, and three echoes, as evidenced by regression intercepts and average biases closer to 0 and slopes closer to 1. Although small, the examined pairwise differences were statistically significant, due to large sample size and paired design. Additionally, the agreement seemed to improve progressively with decreasing number of echoes from six to three, although formal statistical analyses were not performed for all possible pairwise comparisons. As opposed to regression intercept, slope, and average bias, R^2 was similar for three- to six-echo methods. As expected, two-echo M-MRI-PDFF showed the poorest agreement with MRS-PDFF, with lower regression slope, higher intercept, greater average bias, and lower R^2.

These results suggest that, over a range of numbers of echoes from three to six, M-MRI has high accuracy for estimation of hepatic PDFF in children. Two aspects of accuracy were assessed: systematic error (assessed by regression intercept, slope, and average bias) and random error (assessed by regression R^2). Lower systematic error is connoted by regression intercept closer to 0%, slope closer to 1, and average bias closer to 0%. Lower random error is connoted by R^2 closer to 1. Using three to six echoes, M-MRI provides low systematic and random errors. The systematic error appears to decrease progressively as the number of echoes is reduced from six to three, suggesting that three echoes may be the single most accurate method. The decrease in systematic error with reduced number of echoes suggests that the signal model is imperfect and that signals acquired at progressively greater echo times diverge from the model (17); further research is needed to identify and correct for these sources of error. Additionally, pairwise differences in T2* values estimated using three to six echoes were small but significant, helping to confirm that the signal model is imperfect. Our study did not include a gold standard for T2* estimation and hence it does not indicate which of the tested methods provides the most accurate estimate.

As opposed to systematic error, random error is not meaningfully affected by the number of echoes over the range from three to six; this is to be expected because the ROIs were co-registered across echo times and, due to the large number of pixels per ROI, random fluctuations in ROI signal intensity across echo times are small. By comparison, two-echo MRI was least accurate and systematically underestimated PDFF relative to MRS. There are two possible causes for inaccuracy with two-echo method: lack of T2* correction and fewer echoes which reduces signal-to-noise ratio (SNR) performance. Of the two, the most important is probably lack of T2* correction which is not possible with two-echo PDFF estimation methods. Neglecting T2* is known to cause systematic PDFF underestimation, even in the absence of iron overload, because T2* signal decay reduces the observed signal loss between successive OP and IP images (7). The reduced SNR performance is likely less relevant since linear regression R^2 values were not meaningfully worse than R^2 values for larger-echo methods.

By comparison, sex and BMI (but not age) were shown to have confounding effects in the multivariate analysis. M-MRI-PDFF accuracy was worse in obese boys compared to non-obese boys or obese/non-obese girls. Further research is needed to confirm the lower

accuracy in obese boys and determine its cause. The underperformance was small, however, and likely to have little relevance in most clinical and research applications.

A recent pilot study showed that M-MRI was feasible in children and suggested that six-echo M-MRI was accurate for estimation of hepatic PDFF in this age group (24). The accuracy metrics (regression intercept = 1.59, regression slope = 0.97) reported by that study closely agree to our study. Other studies have assessed the accuracy of M-MRI for estimation of hepatic PDFF in adults or mixed adult/pediatric populations. Yokoo et al reported regression intercepts and slopes close to 0% and 1, respectively, using a 1.5T Siemens scanner (14) and a 3T GE scanner (12) but subanalyses to document the performance in children were not performed. Furthermore, Tang et al showed in a mixed adult/pediatric cohort that M-MRI-PDFF was significantly correlated with hepatic histologic steatosis grade and was not confounded by age or sex (33). Our study also agrees with a previous study that suggested three- and four-echo M-MRI has higher accuracy for estimation of hepatic PDFF than six-echo M-MRI (17). None of these studies performed subanalyses to assess the accuracy in sex/obesity subgroups.

Our study helps to further validate M-MRI as a method to estimate hepatic PDFF by extending the analysis to a previously underreported population, children. Thus, based on our results, M-MRI using at least three echoes may be used for accurate estimation of hepatic PDFF in children as young as eight yrs of age. This is relevant because children may differ from adults in ways that may affect the performance of MRI, including breath-hold ability, tolerance for MR examinations, and body habitus. Importantly, our sample size was relatively large, which helps to establish the reliability of our results and allowed us to evaluate age, sex, and BMI as potential confounders. Our study also suggests that M-MRI number of echoes can be reduced from six (the prototype method) to as few as three echoes without sacrificing accuracy for PDFF estimation; in fact, three echoes may be the most accurate but further studies are needed for confirmation. Moreover, if use of fewer echoes improves accuracy, it suggests the underlying model is imperfect (as discussed above) and further research may be needed to refine the model.

Utilizing three rather six echoes may facilitate whole-liver coverage in a single breath-hold and/or allow for acquisition of images with higher spatial resolution, which may potentially be helpful for focal liver lesion characterization. A disadvantage of reducing the number of echoes is that accuracy for R2* estimation may be compromised. The number of echoes influences the noise performance of R2* quantification, and a recent study by Hernando et al suggests that the minimum number of echoes for robust R2* estimation is probably six (34). Thus, reducing the number of echoes to three may complicate the simultaneous assessment of fat and iron, which may be relevant in liver conditions in which iron overload and fat accumulation coexist. The relevance of compromised R2* estimation in pediatric NAFLD is unclear, however, as children with NAFLD rarely have significant iron overload (35). By comparison, two-echo M-MRI should not be used for PDFF estimation due to its lower accuracy unless more advanced methods are unavailable.

Despite the high accuracy achieved with multi-echo M-MRI in our study, it should be emphasized that imaging children with MRI is not always straightforward. Our research

center has technologists who are experienced in working with children, and – to accommodate each child's breath-hold capacity, body habitus, and ability to cooperate – our technologists were allowed to modify certain acquisition parameters within a reasonable range shown previously to have negligible effects on PDFF estimation (36).

Our study had some limitations. First, liver histology was not available as a reference standard. Second, perfect co-localization between the MRI-ROIs and the MRS voxel is not possible because an anatomic image is not obtained simultaneously with the MRS acquisition. Any change in body position or diaphragmatic excursion between the localizing image and the MRS acquisition will cause the co-localization to be incorrect and introduce random error in M-MRI-PDFF relative to the spectroscopic reference. Third, our study was a single-center study and all our cases were scanned with a 3T scanner. However, recent studies have shown that M-MRI-PDFF is reproducible across MR scanners from different manufacturers and at different field strengths (32, 37), so our results are likely to be generalizable to other scanners and field strengths. Fourth, we did not assess C-MRI methods. A recent study showed that a C-MRI method estimated hepatic PDFF accurately, using MRS-PDFF as reference standard, in a mixed adult/pediatric population; depending on the TR, flip angle, and number of echoes, a regression bias as small as 0.24% was observed (16). Thus, it is likely that C-MRI methods also can provide accurate PDFF estimates in children, although this has not yet been explicitly shown.

In conclusion, our study suggests that over a range of numbers of echoes from three to six, M-MRI has high accuracy for estimation of hepatic PDFF in children. This helps to further validate M-MRI as a tool to estimate hepatic PDFF by expanding the relevant population to include children. Future research directions for MRI-based PDFF estimation in children may include multicenter implementation and evaluation, comparison of M-MRI to C-MRI methods, use of other reference standards such as histology and triglyceride concentration, and assessment for quantification of longitudinal changes in hepatic fat content.

Acknowledgments

Funding: This study was funded in part by R56 DK090350 and R01 DK088831 from the NIDDK, and UL1 RR031980 from the NCRR for the Clinical and Translational Research Institute at UCSD.

References

1. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging. 2012; 36:1011–1014. [PubMed: 22777847]
2. Le TA, Chen J, Changchien C, et al. Effect of colesevelam on liver fat quantified by magnetic resonance in nonalcoholic steatohepatitis: a randomized controlled trial. Hepatology. 2012; 56:922–932. [PubMed: 22431131]
3. Bjermo H, Iggman D, Kullberg J, et al. Effects of n-6 PUFAs compared with SFAs on liver fat, lipoproteins, and inflammation in abdominal obesity: a randomized controlled trial. Am J Clin Nutr. 2012; 95:1003–1012. [PubMed: 22492369]
4. Maersk M, Belza A, Stodkilde-Jorgensen H, et al. Sucrose-sweetened beverages increase fat storage in the liver, muscle, and visceral fat depot: a 6-mo randomized intervention study. Am J Clin Nutr. 2012; 95:283–289. [PubMed: 22205311]
5. Stein EA, Dufour R, Gagne C, et al. Apolipoprotein B synthesis inhibition with mipomersen in heterozygous familial hypercholesterolemia: results of a randomized, double-blind, placebo-

controlled trial to assess efficacy and safety as add-on therapy in patients with coronary artery disease. Circulation. 2012; 126:2283–2292. [PubMed: 23060426]

6. Charatcharoenwitthaya P, Lindor KD. Role of radiologic modalities in the management of non-alcoholic steatohepatitis. Clin Liver Dis. 2007; 11:37–54. [PubMed: 17544971]
7. Reeder SB, Sirlin CB. Quantification of liver fat with magnetic resonance imaging. Magn Reson Imaging Clin N Am. 2010; 18:337–357. [PubMed: 21094444]
8. Bydder M, Yokoo T, Hamilton G, et al. Relaxation effects in the quantification of fat using gradient echo imaging. Magn Reson Imaging. 2008; 26:347–359. [PubMed: 18093781]
9. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB. Multiecho water-fat separation and simultaneous R2* estimation with multifrequency fat spectrum modeling. Magn Reson Med. 2008; 60:1122–1134. [PubMed: 18956464]
10. Liu CY, McKenzie CA, Yu H, Brittain JH, Reeder SB. Fat quantification with IDEAL gradient echo imaging: correction of bias from T(1) and noise. Magn Reson Med. 2007; 58:354–364. [PubMed: 17654578]
11. Hernando D, Hines CD, Yu H, Reeder SB. Addressing phase errors in fat-water imaging using a mixed magnitude/complex fitting method. Magn Reson Med. 2012; 67:638–644. [PubMed: 21713978]
12. Yokoo T, Shiehmorteza M, Hamilton G, et al. Estimation of hepatic proton-density fat fraction by using MR imaging at 3.0. T Radiology. 2011; 258:749–759.
13. Meisamy S, Hines CD, Hamilton G, et al. Quantification of hepatic steatosis with T1-independent, T2-corrected MR imaging with spectral modeling of fat: blinded comparison with MR spectroscopy. Radiology. 2011; 258:767–775. [PubMed: 21248233]
14. Yokoo T, Bydder M, Hamilton G, et al. Nonalcoholic fatty liver disease: diagnostic and fat-grading accuracy of low-flip-angle multiecho gradient-recalled-echo MR imaging at 1.5. T Radiology. 2009; 251:67–76.
15. Hines CD, Frydrychowicz A, Hamilton G, et al. T(1) independent, T(2) (*) corrected chemical shift based fat-water separation with multi-peak fat spectral modeling is an accurate and precise measure of hepatic steatosis. J Magn Reson Imaging. 2011; 33:873–881. [PubMed: 21448952]
16. Johnson BL, Schroeder ME, Wolfson T, et al. Effect of flip angle on the accuracy and repeatability of hepatic proton density fat fraction estimation by complex data-based, T1-independent, T2*-corrected, spectrum-modeled MRI. J Magn Reson Imaging. 2013; 39:440–447. [PubMed: 23596052]
17. Levin YS, Yokoo T, Wolfson T, et al. Effect of echo-sampling strategy on the accuracy of out-of-phase and in-phase multiecho gradient-Echo MRI hepatic fat fraction estimation. J Magn Reson Imaging. 2013; 39:567–575. [PubMed: 23720420]
18. Berardis S, Sokal E. Pediatric non-alcoholic fatty liver disease: an increasing public health issue. Eur J Pediatr. 2014; 173:131–139. [PubMed: 24068459]
19. Lindback SM, Gabbert C, Johnson BL, et al. Pediatric nonalcoholic fatty liver disease: a comprehensive review. Adv Pediatr. 2010; 57:85–140. [PubMed: 21056736]
20. Schwimmer JB, Deutsch R, Kahen T, Lavine JE, Stanley C, Behling C. Prevalence of fatty liver in children and adolescents. Pediatrics. 2006; 118:1388–1393. [PubMed: 17015527]
21. Vajro P, Lenta S, Socha P, et al. Diagnosis of nonalcoholic fatty liver disease in children and adolescents: position paper of the ESPGHAN Hepatology Committee. J Pediatr Gastroenterol Nutr. 2012; 54:700–713. [PubMed: 22395188]
22. Mishra P, Younossi ZM. Abdominal ultrasound for diagnosis of nonalcoholic fatty liver disease (NAFLD). Am J Gastroenterol. 2007; 102:2716–2717. [PubMed: 18042105]
23. Awai HI, Newton KP, Sirlin CB, Behling C, Schwimmer JB. Evidence and Recommendations for Imaging Liver Fat in Children, Based upon Systematic Review. Clin Gastroenterol Hepatol. 2014; 12:765–773. [PubMed: 24090729]
24. Achmad, E.; Yokoo, T.; Hamilton, G., et al. Assessment of the Feasibility of Using MR Imaging and Proton MR Spectroscopy to Measure Liver Fat in Children. Radiological Society of North America 97th Scientific Assembly and Annual Meeting; Chicago. December 1, 2011; (SSQ15-06) (RSNA Abstract 11016236)

25. Middleton, MS.; Hamilton, G.; Sirlin, CB. Effect of Saturation Bands on Quantitative Liver Fat Proton MR Spectroscopy. Radiological Society of North America 97th Scientific Assembly and Annual Meeting; Chicago. November 28, 2011; (SSE08-06)(RSNA Abstract 11015164)
26. Hamilton G, Middleton MS, Bydder M, et al. Effect of PRESS and STEAM sequences on magnetic resonance spectroscopic liver fat quantification. J Magn Reson Imaging. 2009; 30:145–152. [PubMed: 19557733]
27. Bydder M, Hamilton G, Yokoo T, Sirlin CB. Optimal phased-array combination for spectroscopy. Magn Reson Imaging. 2008; 26:847–850. [PubMed: 18486392]
28. Negrete LM, Middleton MS, Clark L, et al. Inter-examination precision of magnitude-based MRI for estimation of segmental hepatic proton density fat fraction in obese subjects. J Magn Reson Imaging. 2014; 39:1265–1271. [PubMed: 24136736]
29. Naressi A, Couturier C, Castang I, de Beer R, Graveron-Demilly D. Java-based graphical user interface for MRUI, a software package for quantitation of in vivo/medical magnetic resonance spectroscopy signals. Comput Biol Med. 2001; 31:269–286. [PubMed: 11334636]
30. Vanhamme L, van den Boogaart A, Van Huffel S. Improved method for accurate and efficient quantification of MRS data with use of prior knowledge. J Magn Reson. 1997; 129:35–43. [PubMed: 9405214]
31. Hamilton G, Yokoo T, Bydder M, et al. In vivo characterization of the liver fat (1)H MR spectrum. NMR Biomed. 2011; 24:784–790. [PubMed: 21834002]
32. Kang GH, Cruite I, Shiehmorteza M, et al. Reproducibility of MRI-determined proton density fat fraction across two different MR scanner platforms. J Magn Reson Imaging. 2011; 34:928–934. [PubMed: 21769986]
33. Tang A, Tan J, Sun M, et al. Nonalcoholic fatty liver disease: MR imaging of liver proton density fat fraction to assess hepatic steatosis. Radiology. 2013; 267:422–431. [PubMed: 23382291]
34. Hernando D, Kramer JH, Reeder SB. Multipeak fat-corrected complex R2* relaxometry: theory, optimization, and clinical validation. Magn Reson Med. 2013; 70:1319–1331. [PubMed: 23359327]
35. Schwimmer JB, Behling C, Newbury R, et al. Histopathology of pediatric nonalcoholic fatty liver disease. Hepatology. 2005; 42:641–649. [PubMed: 16116629]
36. Hansen KH, Schroeder ME, Hamilton G, Sirlin CB, Bydder M. Robustness of fat quantification using chemical shift imaging. Magn Reson Imaging. 2012; 30:151–157. [PubMed: 22055856]
37. Mashhood A, Railkar R, Yokoo T, et al. Reproducibility of hepatic fat fraction measurement by magnetic resonance imaging. J Magn Reson Imaging. 2013; 37:1359–1370. [PubMed: 23172799]

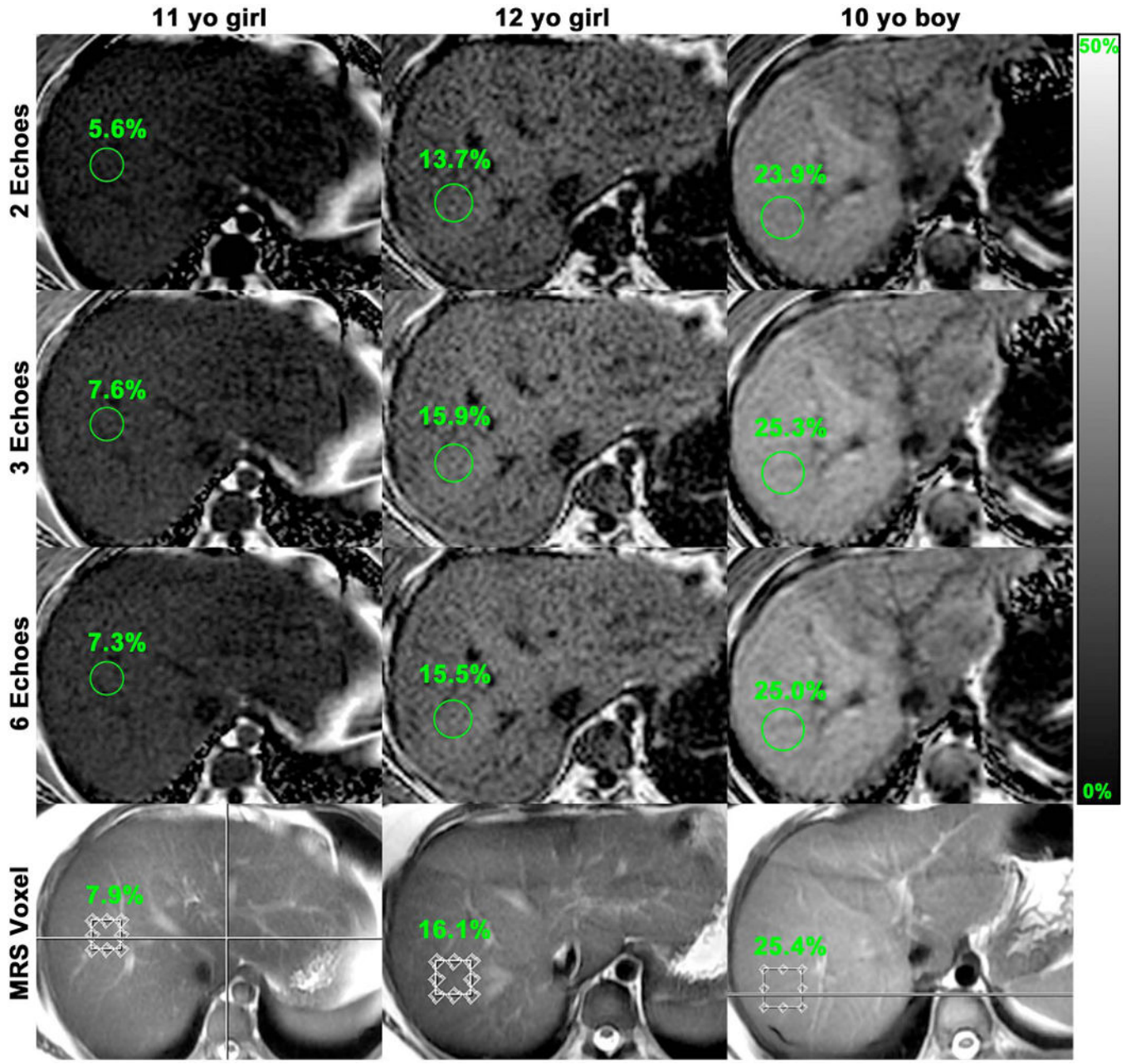

Figure 1.

Representative images from three children with MRS-PDFF values of < 10%, 10 – 20%, and > 20% to illustrate agreement of M-MRI-PDFF and MRS-PDFF at the MRS voxel location. Two-, three-, and six-echo M-MRI-PDFF parametric maps were generated from source images by applying the algorithm pixel by pixel. For each case and echo number, M-MRI-PDFF measured in a region of interest (circle) corresponding to the MRS voxel is indicated.

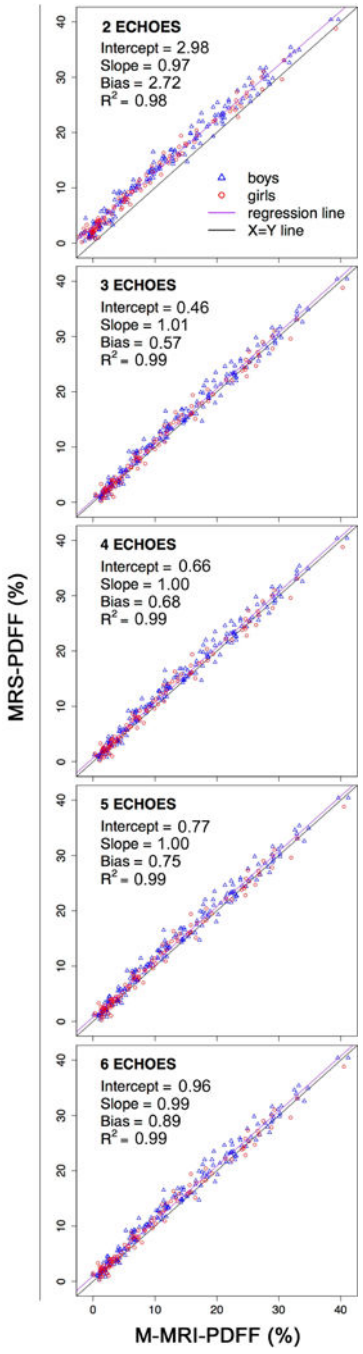

Figure 2.

Scatterplots for MRS-PDFF versus M-MRI-PDFF over a range of numbers of echoes from two to six. For each M-MRI method, regression intercept, regression slope, average bias, and R^2 are indicated.

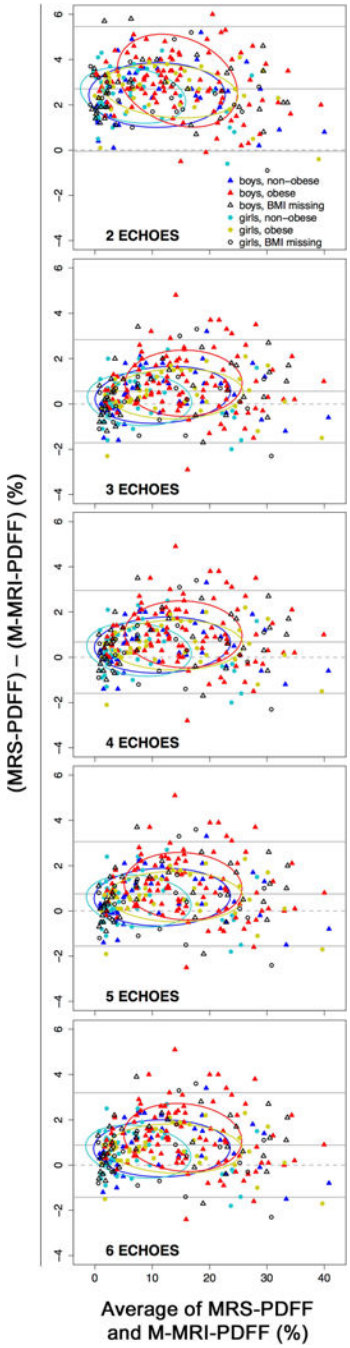

Figure 3.

Bland-Altman plots for MRS-PDFF and M-MRI-PDFF over a range of numbers of echoes from two to six.

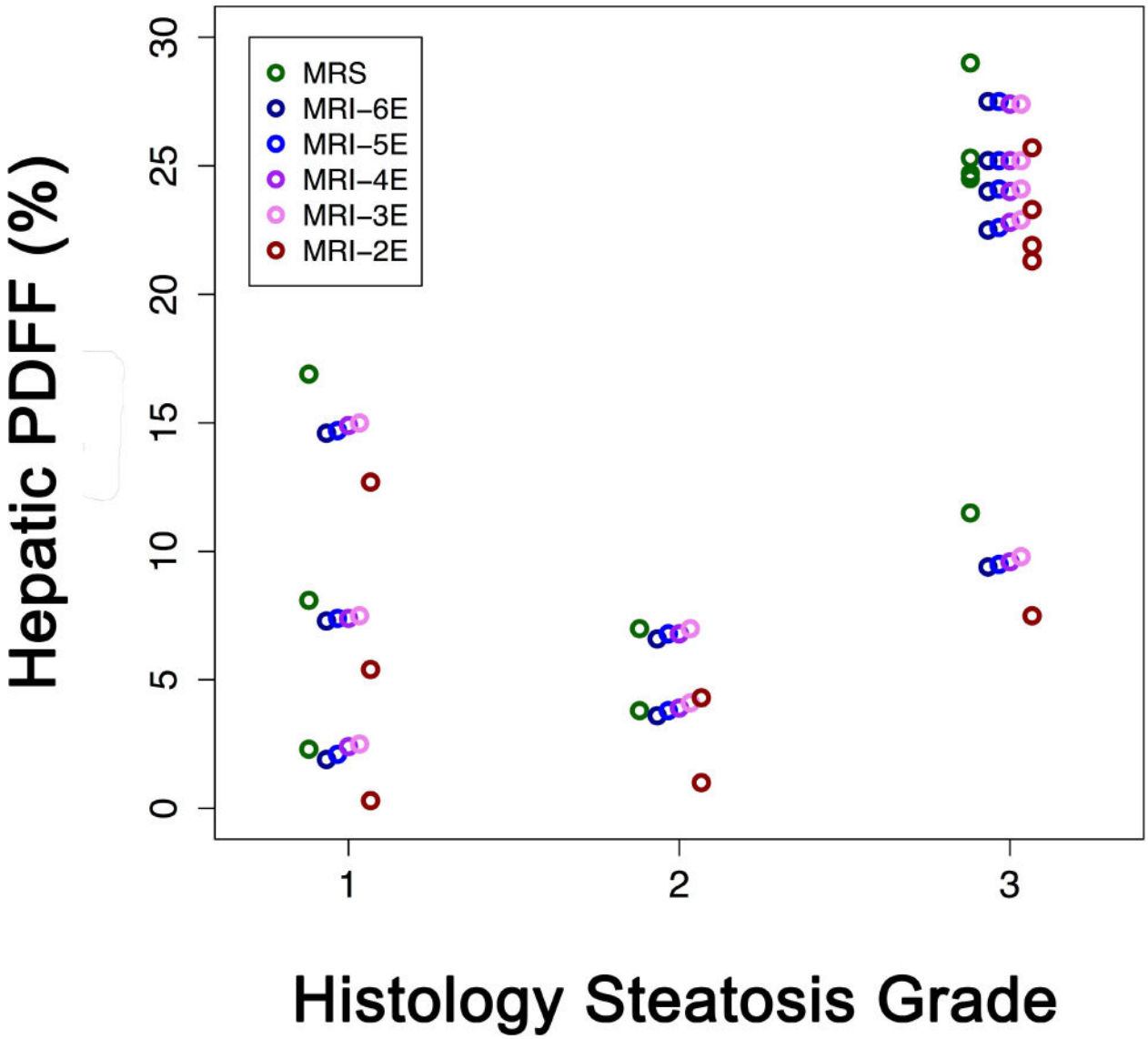

Figure 4.

Correlation analyses between histology steatosis grade and hepatic PDFF in 10 children with available liver biopsy results. The correlation coefficient was similar between histology steatosis grade and hepatic PDFF estimated by two- to six-echo M-MRI methods or MRS (Spearman's rho = 0.71, p = 0.0213).

MRS: MR Spectroscopy.

Table 1

Summary of cohort characteristics

Group	Subjects in study n (%)	Age (yrs)	MRS-PDFF (%)	Subjects with recorded BMI (n)	BMI (kg/m^2)	BMI Z-score
Total	286 (100)	14.2 ± 2.5 (8 – 20)	13.1 ± 9.8 (0.2 – 40.4)	223	31.6 ± 7.0 (14.5 – 51.1)	2.0 ± 0.8 (−3.3 – 3.1)
Boys	182 (64)	14.3 ± 2.4 (8 – 20)	14.3 ± 9.8 (0.4 – 40.4)	143	32.5 ± 7.0 (15.6 – 51.1)	2.1 ± 0.7 (−3.3 – 3.1)
Girls	104 (36)	14.1 ± 2.7 (8 – 20)	11.1 ± 9.3 (0.2 – 38.8)	80	30.0 ± 6.7 (14.5 – 49.2)	1.8 ± 0.8 (−1.1 – 2.8)

Data are means ± standard deviations, with ranges in parentheses

BMI: Body mass index

Table 2

Parameters from MRS-PDFF vs. M-MRI-PDFF regression models

M-MRI method	Intercept (%) (95% CI)	Slope (95% CI)	Average bias (%) (95% CI)	R^2 (95% CI)
6-Echo	0.96 (0.66, 1.27)	0.99 (0.97, 1.02)	0.89 (0.70, 1.09)	0.99 (0.98, 0.99)
5-Echo	0.77 (0.46, 1.09)	1.00 (0.98, 1.02)	0.75 (0.57, 0.95)	0.99 (0.98, 0.99)
vs. 6-Echo	*	*	*	NS
4-Echo	0.66 (0.36, 0.98)	1.00 (0.98, 1.03)	0.68 (0.49, 0.88)	0.99 (0.98, 0.99)
vs. 6-Echo	*	*	*	NS
3-Echo	0.46 (0.15, 0.78)	1.01 (0.99, 1.03)	0.57 (0.38, 0.78)	0.99 (0.98, 0.99)
vs. 6-Echo	*	*	*	NS
2-Echo	2.98 (2.68, 3.66)	0.97 (0.93, 1.00)	2.72 (2.51, 3.04)	0.98 (0.96, 0.99)
vs. 6-Echo	*	*	*	NS

Parameters (intercept, slope, average bias, and R^2) and in parentheses 95% bootstrap CIs for regression models of MRS-PDFF vs. M-MRI-PDFF methods.

CI: Confidence interval. * = pairwise differences in regression parameters were significant for 5-, 4-, 3-, and 2-echo methods compared to 6-echo method; NS = differences were not significant.

Table 3

T2* values estimated by different M-MRI Methods

M-MRI Method	Mean T2* ± Standard Deviation (Range) in ms
6-Echo	21.2 ± 5.0 (4.7 – 39.3)
5-Echo	22.6 ± 5.6 (4.7 – 51.9)
4-Echo	22.1 ± 5.2 (4.6 – 40.5)
3-Echo	27.4 ± 8.1 (4.8 – 71.1)

All pairwise differences in mean T2* values estimated by each M-MRI method were statistically significant.

Table 4

Sex-obesity subgroup subanalyses

Regression Metric	No. of Echoes	Boys		Girls	
		Non-obese (n=43)	Obese (n=100)	Non-obese (n=44)	Obese (n=36)
Intercept (%)	6	0.79	1.58	0.84	0.87
	3	0.28	1.06	0.35	0.43
Slope	6	1.00	0.98	0.97	0.99
	3	1.01	0.99	0.98	1.00
Average Bias (%)	6	0.74	1.24	0.65	0.73
	3	0.41	0.92	0.25	0.45
R^2	6	0.99	0.98	0.98	0.99
	3	0.99	0.98	0.98	0.99

Assessing the accuracy of six-echo (the prototype) and three-echo M-MRI-PDFF among subgroups of obese/non-obese boys and girls using four regression parameters: intercept, slope, average bias, and R^2.