HHS Public Access

Author manuscript

Hepatology. Author manuscript; available in PMC 2019 March 01.

Published in final edited form as:

Hepatology. 2018 March ; 67(3): 858–872. doi:10.1002/hep.29596.

Diagnostic Accuracy of Magnetic Resonance Imaging Hepatic Proton Density Fat Fraction in Pediatric Nonalcoholic Fatty Liver Disease

Michael S. Middleton, MD PhD[1], **Mark L. Van Natta, MHS**[2], **Elhamy R. Heba, MD**[1], **Adina Alazraki, MD**[3], **Andrew T. Trout, MD**[4], **Prakash Masand, MD**[5], **Elizabeth M. Brunt, MD**[6], **David E. Kleiner, MD PhD**[7], **Edward Doo, MD**[8], **James Tonascia, PhD**[2], **Joel E. Lavine, MD, PhD**[9], **Wei Shen, MD, MPH, MS**[9], **Gavin Hamilton, PhD**[1], **Jeffrey B. Schwimmer, MD**[10], **Claude B. Sirlin, MD**[1], and **for the NASH Clinical Research Network**

[1] -Liver Imaging Group, Department of Radiology, UCSD School of Medicine, San Diego, California [2] -Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland [3] -Emory University School of Medicine, Department of Radiology and Imaging Sciences, Atlanta, Georgia [4] -Cincinnati Children's Hospital, Department of Radiology, Cincinnati, Ohio [5] -Texas Children's Hospital, Houston, Texas [6] -Emeritus, Washington University School of Medicine, St. Louis, Missouri [7] -Laboratory of Pathology, National Cancer Institute [8] -Liver Diseases Section, Digestive Diseases Branch, National Institute of Diabetes and Digestive and Kidney Diseases [9] -Division of Pediatric Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, Columbia University Medical Center, New York, New York [10] -Division of Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of California San Diego School of Medicine, La Jolla, California; and Department of Gastroenterology, Rady Children's Hospital, San Diego, California

Abstract

We assessed the performance of magnetic resonance imaging (MRI) proton density fat fraction (PDFF) in children to stratify hepatic steatosis grade before and after treatment in the Cysteamine Bitartrate Delayed-Release for the Treatment of Nonalcoholic Fatty Liver Disease in Children (CyNCh) trial, using centrally-scored histology as reference. Participants had multi-echo 1.5T or 3T MRI on scanners from three manufacturers. Of 169 enrolled children, 110 (65%) and 83 (49%) had MRI and liver biopsy at baseline and at end-of-treatment (EOT; 52-weeks), respectively. At baseline, 17% (19/110), 28% (31/110), and 55% (60/110) of liver biopsies showed grades 1, 2, and 3 histologic steatosis; corresponding PDFF (mean ± standard deviation) values were 10.9 ± 4.1%, 18.4 ± 6.2%, and 25.7 ± 9.7%, respectively. PDFF classified grade 1 vs. 2–3 and 1–2 vs. 3 steatosis with areas under receiving operator characteristic curves (AUROCs) of 0.87 (95% confidence interval [CI]: 0.80, 0.94) and 0.79 (0.70, 0.87), respectively. PDFF cut-offs at 90%

Corresponding author contact information: Michael S. Middleton, MD PhD, Liver Imaging Group, Department of Radiology, UCSD School of Medicine, ACTRI Building, MC 8888, 9452 Medical Center Drive, La Jolla, CA 92037, phone: (858) 750-0878, fax: none, msm@ucsd.edu. mvnatta@jhu.edu, elhamyrheba@gmail.com, adina.alazraki@choa.org, andrew.trout@cchmc.org, pmmasand@texaschildrens.org, ebrunt@wustl.edu, kleinerd@mail.nih.gov, dooe@niddk.nih.gov, jtonasc1@jhu.edu, jl3553@columbia.edu, ws2003@cumc.columbia.edu, ghamilton@ucsd.edu, jschwimmer@ucsd.edu, csirlin@ucsd.edu.

specificity were 17.5% for grades 2–3 steatosis, and 23.3% for grade 3 steatosis. At EOT, 47% (39/83), 41% (34/83), and 12% (10/83) of biopsies showed improved, unchanged, and worsened steatosis grade, respectively, with corresponding PDFF (mean ± standard deviation) changes of −7.8 ± 6.3%, −1.2 ± 7.8% and 4.9 ± 5.0%, respectively. PDFF change classified steatosis grade improvement and worsening with AUROCs (95% CIs) of 0.76 (0.66, 0.87) and 0.83 (0.73, 0.92), respectively. PDFF change cut-off values at 90% specificity were −11.0% and +5.5% for improvement and worsening.

Conclusion: MRI-estimated PDFF has high diagnostic accuracy to both classify and predict histologic steatosis grade, and change in histologic steatosis grade in children with NAFLD.

Keywords

PDFF; MRI; NAFLD; cysteamine bitartrate delayed-release; CyNCh

Nonalcoholic fatty liver disease (NAFLD) is the most common pediatric chronic liver disease (1). Common co-morbidities of NAFLD include diabetes and cardiovascular disease (2–7). The long-term natural history of NAFLD is not well understood but the disease has the potential to progress to cirrhosis, hepatocellular carcinoma, and early death (8–11), and is rapidly becoming the leading cause of liver transplantation (12).

Hepatic steatosis, a key feature of NAFLD, is histologically graded on biopsy according to the proportion of hepatocytes containing fat macrovesicles on hematoxylin and eosin staining (grade 0: < 5%; grade 1: 5–33%; grade 2: 34 to 66%; and grade 3: > 66% (13). However, liver biopsy is invasive and samples only a very small portion of the liver, and thus may not be ideal for longitudinal clinical trials or for clinical monitoring and care in the early stages of the disease. Computed tomography is not well suited for use in serial-monitoring because it uses ionizing radiation and shows only modest association with liver fat content, and conventional ultrasound is of limited value because it provides only semi-quantitative estimates of liver fat content and does not permit assessment of all liver segments (14). Hepatic steatosis may also be assessed non-invasively by a transient elastography-derived controlled attenuation parameter, which is an estimate of total ultrasound attenuation (15). Although we did not perform a direct comparison of the controlled attenuation parameter and PDFF, the accuracy of MRI to estimate a fat fraction has been reported to be better than that obtained from the controlled attenuation parameter in adults, using histology as reference standard (16,17). The accuracy of the controlled attenuation parameter to assess steatosis in children is not known.

In contrast, multi-echo magnetic resonance imaging (MRI) is emerging as the non-invasive method of choice to estimate, on a continuous scale, hepatic proton density fat fraction (PDFF) (10). Single-center cross-sectional studies have shown MRI-estimated PDFF to be accurate, using magnetic resonance spectroscopy (18–22) or histology (23–27) as reference standards, and to be reproducible across field strengths (28–31) and MRI scanner manufacturers (28,31), and a recent meta-analysis has shown that MRI-estimated PDFF is reproducible across field strength and scanner manufacturer (32).

Several single-center studies have examined the diagnostic performance of hepatic PDFF to grade histologic steatosis. In a single-center study of 89 adults (33), PDFF correlated with histologic steatosis grade, and classified steatosis grade 0 vs. grade 1 at the 6.4% PDFF cutoff reported by Tang et al (23), with a sensitivity of 68%, a specificity of 98%, and an area under receiving operator characteristic curve (AUROC) of 0.82. In adults with nonalcoholic steatohepatitis (NASH) in the multi-center Farnesoid X Receptor Ligand Obeticholic Acid in NASH Treatment (FLINT) trial, PDFF correlated with histologic steatosis grade, and accurately classified baseline dichotomized steatosis grades 0–1 vs. grades 2–3, and grades 0–2 vs. grade 3 with AUROCs of 0.95 and 0.96 at cutoffs of 16.3% and 21.7%, respectively, with corresponding sensitivities of 83% and 84% at 90% specificity (34,35). In the same trial, PDFF change correlated with histologic steatosis grade change, and accurately classified steatosis grade change longitudinally (35); at end-of-treatment, PDFF change classified steatosis grade change, with cutoffs at 90% specificity of −5.1% for improvement and +5.6% for worsening, both with AUROCs of 0.81.

Assessing multi-center PDFF diagnostic performance in children is necessary to further validate PDFF as a biomarker of hepatic steatosis prior to its widespread use in clinical care or as an endpoint in pediatric clinical trials (34). Few studies assessing PDFF validation metrics have been conducted in children (36–40), and none in a multi-center setting. In the MRI Rosetta Stone Project, 174 children at a single center had hepatic MRI-estimated PDFF and histology evaluated (36). PDFF correlated with steatosis grade ($r = 0.725$, $p < 0.01$), and the overall accuracy of predicting histologic steatosis grade from PDFF was 56%. However, the relationship between change in hepatic PDFF and change in histologic steatosis grade has not been evaluated in children. Therefore, the purpose of this study was to assess cross-sectional and longitudinal diagnostic performance of hepatic PDFF to grade histologic steatosis in children with NAFLD using centrally-scored histology as the reference standard.

Patients and Methods

STUDY DESIGN

We performed a prospectively-designed study of the diagnostic performance of PDFF estimated by multi-echo MRI as part of the Cysteamine Bitartrate Delayed-Release for the Treatment of Nonalcoholic Fatty Liver Disease (NAFLD) in Children (CyNCh) trial (NCT01529268), a multi-center, randomized, double-masked, placebo-controlled, phase 2b clinical trial of treatment with either cysteamine bitartrate delayed release or placebo in children with NAFLD (41). Liver biopsy and MRI were performed at baseline and after 52 weeks of treatment allowing paired comparisons of PDFF and histologic hepatic steatosis grade, and their longitudinal changes. PDFF was a secondary endpoint, for which centrally-scored histologic hepatic steatosis grade from percutaneous biopsy served as the reference standard.

Eligibility criteria for the CyNCh trial are published elsewhere and were based on well-established biopsy criteria for pediatric NAFLD (41). Inclusion criteria for MRI were enrollment in the CyNCh trial, and willingness and ability to complete both MRI exams (the baseline exam prior to randomization and within 90 days of baseline biopsy, and the end-of-treatment (EOT) MRI exam within 120 days of EOT biopsy). Exclusion criteria for MRI

were contraindication to MRI, extreme claustrophobia, weight or girth exceeding MRI scanner capability, or any condition or circumstance that, in the opinion of the clinical trial site investigator would interfere with completion of the exam.

The CyNCh study protocol, including the MRI portion conformed to the ethical guidelines of the 1975 Declaration of Helsinki as reflected in a priori approval by the appropriate institutional review committee at each participating clinical trial site, and was in compliance with the Health Insurance Portability and Accountability Act. Children aged 8 to 17 years at enrollment were included in this study, with all children providing written informed assent with written informed consent by a parent or guardian.

All authors had access to the study data, and reviewed and approved the final manuscript.

MRI CLINICAL TRIAL SITES

Nine of the ten participating CyNCh clinical trial sites contributed MRI data to this study using 1.5T (five sites), 3T (three sites), or both 1.5 and 3T (one site) MR scanners (Table 1).

MRI CLINICAL TRIAL SITE QUALIFICATION AND QUALITY CONTROL

The NASH Clinical Research Network (CRN) Radiology Coordinating Center (RCC), in conjunction with the NASH CRN Data Coordinating Center managed the MRI portion of this study. Each individual site was approved by the RCC based on approval of technically adequate phantom or volunteer MRI data. The RCC also provided central imaging quality control.

MRI ACQUISITION AND ANALYSIS

MRI acquisition was identical to that recently reported for the MRI sub-study to the FLINT trial (35), using a non-contrast, breath-hold, gradient-recalled-echo, two-dimensional axial sequence. Images of the entire liver were obtained using a torso array coil centered over the upper abdomen. Parameters were selected to correct for or avoid confounding factors (e.g., T1 bias, T2* decay, multi-frequency interference) that could introduce fat quantification error (10,18–22) (Table 1). Images in Digital Imaging and Communications in Medicine format were transferred from clinical trial sites to the RCC.

MRI analysis was performed at the RCC in the following fashion. For each MRI, signal intensities from the 6-echo magnitude spoiled-gradient-echo source images were analyzed pixel-by-pixel using a custom MATLAB™ (The MathWorks, Natick, MA, USA) non-linear, least-squares fitting algorithm to produce a PDFF parametric map at each image level throughout the liver (18–22, 42). The algorithm computed the PDFF values of those parametric maps by assuming exponential T2* signal decay and applying a multi-peak spectral model to account for fat-fat and fat-water multi-frequency interference effects, based on the work of Hamilton et al (43). One circular 1-cm radius region of interest (ROI) was placed the on 5th echo (out-of-phase) source images in each of the nine anatomical liver segments. Those ROIs were propagated to the corresponding PDFF parametric map images, and the mean PDFF value from each evaluable parametric map-based ROI was recorded. For

each MRI, the number of ROIs that were evaluable was recorded. For MRIs that were not evaluable, reasons were recorded.

LIVER BIOPSY

In the CyNCh trial, liver biopsies were performed for clinical care within 90 days of the start of screening and no more than 120 days before randomization and demonstrated histologic evidence of NAFLD as well as a NAFLD activity score (NAS) of ≥ 4, as scored by the individual NASH CRN Pathology Committee member at each study site. Liver biopsies were performed at 52 weeks to determine the response to therapy (41).

Pathologists of the NASH CRN Pathology Committee reviewed the biopsies conjointly at a multi-head microscope as part of the CyNCh trial (41). The pathologists scored steatosis grade (proportion of hepatocytes containing fat macrovesicles: grade 0 for < 5%, grade 1 for 5 to 33%, grade 2 for 34 to 66%, and grade 3 for > 66% (13)) according to the NASH CRN histologic NAFLD scoring system that also includes scoring criteria for accompanying histologic features such as steatosis location, lobular inflammation, portal inflammation, hepatocellular ballooning, fibrosis stage, iron grade and location, and global biopsy diagnosis. The NAFLD Activity Score was calculated from unweighted sums of steatosis, lobular inflammation, and hepatocellular ballooning (13).

BLINDING

RCC analysts and other staff at the RCC were blinded to histology results, and pathologists were blinded to all clinical and imaging results. RCC analysts were blinded to treatment assignment. Pathologists knew each slide was obtained from a participant enrolled in or being considered for enrollment in a NASH CRN study, but they were blinded to all clinical information, including age, time point (baseline or EOT), or enrollment in the CyNCh MRI sub-study.

OTHER DATA

Participant demographics, laboratory, anthropomorphic measurements, and medical history were collected at each clinical trial site.

STATISTICAL ANALYSIS

Statistical analyses were done with SAS (SAS Institute 2011, Base SAS 9·3 Procedures Guide) and Stata (StataCorp 2013, Stata Statistical Software: release 13).

A single composite PDFF value was calculated for each MRI as the mean of the PDFF values for the nine anatomical liver segments. Demographic, histologic and imaging information were summarized with categorical variables expressed as numbers and percentages and continuous variables expressed by mean(± standard deviation [SD]). The proportion of participants with, and without MRIs at baseline, and for those with MRIs at baseline, the proportion of those with, and without MRIs at EOT were compared with regard to treatment group, study site, demographics, liver enzymes, lipids, metabolic factors, co-morbidities, concomitant liver medications, and histology findings.

Histologic components at baseline were linearly regressed on PDFF at baseline, and 52-week changes in histologic features were linearly regressed on 52-week changes in PDFF adjusting for baseline value of histologic feature. Analyses of follow-up data at 52 weeks were pooled across treatment groups. Beta values (mean histologic component score change per 1% increase in PDFF, 95% confidence intervals (CIs), and p-values were estimated between PDFF and histologic components (steatosis score, lobular and portal inflammation scores, hepatocellular ballooning score, and fibrosis score) at baseline and for changes from baseline to 52 weeks.

Diagnostic accuracy of PDFF to classify hepatic steatosis grade at baseline was tested for grades 1 vs. 2–3, and grades 1–2 vs. 3. Diagnostic accuracy of change in PDFF to classify change in hepatic steatosis grade from baseline to EOT was tested for reduction vs. no change/increase, and increase vs. no change/decrease. Cross-validated AUROCs using a jack-knife procedure and 95% CIs were estimated for each of these dichotomizations (44). Cut-off PDFF values were estimated using the lowest threshold value for which there was ≥ 90% specificity to distinguish between these dichotomized categories. Sensitivity, positive predictive values, and negative predictive valves were calculated along with 95% CIs fixing specificity at 90%.

Interaction between subgroups for MRI-determined steatosis vs. histologic steatosis was tested at baseline and longitudinally for the following dichotomized subgroups: age (8–12 vs. 13–18 yrs), sex, lobular inflammation score (grade 1 inflammation at baseline vs. grade 2–3 inflammation at baseline), fibrosis (no fibrosis at baseline vs. any fibrosis at baseline), scanner type (1.5T vs. 3T), and time from baseline biopsy to baseline MRI (≤ 60 days vs. > 60 days, based on the midpoint of the distribution of these times). Because of the exploratory nature of these new analyses and the increased likelihood of Type I error due to multiple comparisons, we chose the cut-off for statistical significance as 0.01.

Units of PDFF change are expressed in absolute units of change in percent PDFF. Thus, a decrease from 10 to 5% PDFF would represent a change of 5% PDFF.

Results

Of 169 children enrolled in the CyNCh trial from June 2012 to January 2014 at ten participating CyNCh clinical trial sites, all (n = 169) underwent liver biopsies at baseline, and 146 (86%) had liver biopsies at EOT. One hundred ten participants (65%) had MRI at baseline, and 85 (50%) had MRI at baseline and EOT. Of the 195 MRIs obtained for study participants, one EOT MRI was excluded because of poor image quality (phase artifact, and low signal-to-noise ratio). PDFF values were derived from all nine anatomical liver segments for 192/194 of the remaining MRIs; one baseline-only MRI had 2/9 segments that were not evaluable, and one EOT MRI had 1/9 segments that were not evaluable. Eighty-three participants (49%) had MRI and liver biopsy at both time points. Baseline MRIs occurred from 0 to 115 days after initial biopsy (mean 61 days). EOT MRIs that were paired with baseline MRIs ranged from 78 days pre-EOT biopsy to 61 days post-EOT biopsy (mean 2 days pre-EOT biopsy).

There were no statistically significant differences in study parameters among children with, and without MRI at baseline, except for differences by clinical site ($p < 0.0001$), and that children who had baseline MRIs had a lower mean fasting glucose than those who did not have baseline MRIs ($p = 0.04$). There were also no statistically significant differences in study parameters among children who had MRIs at baseline, and those with, and without MRIs at EOT, except again for differences by clinical site ($p < 0.0001$), and that children who had no EOT MRI were older ($p = 0.02$) and had lower mean alkaline phosphatase ($p = 0.008$) compared to those who had only baseline MRIs. Study parameter differences for these groups are given in Table 2.

CROSS-SECTIONAL ANALYSIS

The distribution of PDFF in the 110 children at baseline, all of whom were diagnosed with NAFLD is shown in Figure 1: PDFF mean ± SD was 21.1 ± 9.8%, and ranged from 5.3% to 46.8%.

At baseline, biopsy findings included grade 1 steatosis in 17% (19/110), grade 2 steatosis in 28% (31/110), and grade 3 steatosis in 55% (60/110) of participants. Corresponding mean ± SD PDFF values were 10.9 ± 4.1%, 18.4 ± 6.2%, and 25.7 ± 9.7%, respectively (Figure 2).

Linear regressions of PDFF values on histologic components at baseline are summarized in Table 3. PDFF was positively associated with steatosis score [0.068 mean difference in histologic steatosis score per 1% increase in PDFF; 95% CI: 0.056, 0.081; $p < 0.001$]. No associations with PDFF at baseline were found for lobular or portal inflammation scores, hepatocellular ballooning score, or fibrosis score (p-values 0.10 to 0.66), and all of the CIs for those regressions included 0.

Table 4 summarizes the diagnostic accuracy of PDFF for classifying steatosis. Using PDFF as a classifier, the AUROCs from logistic regression were 0.87 (95% CI: 0.80, 0.94) for classifying steatosis grade 1 vs. 2–3, and 0.79 (95% CI: 0.70, 0.87) for classifying steatosis grade 1–2 vs. 3. PDFF cut-off values at 90% specificity were 17.5% for grades 2–3, and 23.3% for grade 3 discrimination.

LONGITUDINAL ANALYSIS

Figure 3 shows box plots of PDFF values for each histologic hepatic steatosis grade change category (reduction, no change, and increase in steatosis grade). Mean ± SD of change in PDFF values for children with decreased histologic steatosis grade at EOT was −7.8 ± 6.3% (range: −20.1 to 1.7%) in 39 children with improvement (reduction) in steatosis grade (28 reduced one grade, eight reduced two grades, and three reduced three grades); −1.2 ± 7.8% (range: −6.8 to 12.1%) in 34 children with no change in steatosis grade; and 4.9 ± 5.0% (range: −0.5 to 12.1%) in 10 children with worsening (increase) in steatosis grade (nine increased one grade, and one increased two grades).

Linear regressions of changes in PDFF values on changes in histologic components at 52 weeks are summarized in Table 3. Change in PDFF was positively associated with change in steatosis score [0.057 mean change in histologic steatosis score per 1% increase in change of PDFF adjusted for baseline value of histologic steatosis; 95% CI: 0.034, 0.079; $p < 0.001$].

No associations with change in PDFF were found for changes in lobular or portal inflammation scores, hepatocellular ballooning score, or fibrosis score (p-values 0.40 to 0.80), and all of the CIs for those regressions included 0.

Table 4 summarizes diagnostic accuracy of change in PDFF vs. change in histologic hepatic steatosis grade from baseline to EOT. The AUROCs using PDFF change to classify histologic hepatic steatosis grade improvement and worsening, respectively, were 0.76 (95% CI: 0.66, 0.87) and 0.83 (95% CI: 0.73, 0.92). Cut-off values for PDFF change at 90% specificity were −11.0% for improvement and +5.5% for worsening hepatic steatosis grade.

TESTS OF INTERACTION BETWEEN SUBGROUPS

Table 5 summarizes subgroup analyses of regression of baseline steatosis grade on baseline PDFF, and Table 6 summarizes subgroup analyses of regression of worsening or improvement in steatosis grade on 52-week change in PDFF. None of the interaction p-values were below 0.01 (only 2 out 20 were below 0.05), although the power to detect subgroup effects in this study was low.

Discussion

In CyNCh, a multi-center, randomized controlled trial of children with known NAFLD, we found that MRI-derived PDFF values and histologic hepatic steatosis scores were associated ($p < 0.001$), that changes in MRI-derived PDFF values and changes in histologic hepatic steatosis scores were associated ($p < 0.001$), and that baseline and longitudinal PDFF values were not associated with other histologic components (p-value range 0.10 to 0.80).

Some discordance between PDFF values and histologic steatosis scores is expected since PDFF and histology do not measure the same quantity. As emphasized previously by others (14), PDFF is a quantitative marker of MRI-visible hepatic fat content, while histology scoring by grade is a semi-quantitative assessment of the proportion of microscopically-assessed steatotic hepatocytes. PDFF and histology assess liver fat on different scales ranging from 0 to 50% (rarely > 50%), and 0 to 100%, respectively. PDFF percentages are usually less than half of histologic steatosis percentages because PDFF quantifies the ratio of MRI signal from fat to the MRI signal from water, while histologic steatosis grade scores reflect the estimated percentage of hepatocytes that contain microscopically-visible fat globules. If all histologically-examined hepatocytes were half filled with fat globules, the histologic steatosis percentage would be 100% since all cells show fat globules, and the steatosis grade would be 3, but the MRI PDFF value would be about 50%. Another possible reason for discordance between PDFF and histologic measures of steatosis is that by MRI, it is possible to estimate PDFF throughout the liver, whereas biopsy samples only a small volume of a diffuse process.

Our results show that PDFF can distinguish histologic steatosis grade 1 vs. grades 2–3 with an AUROC of 0.87 (95% CI: 0.80, 0.94). This is similar to the larger single-center study of 174 children in which Schwimmer et al (36) showed that PDFF could classify histologic steatosis grades 0 and 1 with an AUROC of 0.82. Our results additionally suggest that, at 90% specificity a PDFF cut-off value of 17.5% provides 74% sensitivity for discriminating

histologic steatosis grade 1 from grades 2–3, and a value of 23.3% provides 60% sensitivity for discriminating grades 1–2 from grade 3. However, we also found that, at 90% specificity a decrease of ≥ 11.0% PDFF provided only 31% sensitivity for identifying histologic steatosis grade improvement, and an increase of ≥ 5.5% provided only 40% sensitivity for identifying histologic steatosis grade worsening.

These PDFF (17.5%, 23.3%) and change-in-PDFF (−11.0%, 5.5%) cutoffs are similar to the corresponding PDFF (16.3%, 21.7%) and change-in-PDFF (−5.1%, 5.6%) cutoffs reported in adults in the FLINT MRI sub-study (35). The associated sensitivities in this study (74% and 60% for PDFF; 31% and 40% for change in PDFF) are lower, however, than those reported for the FLINT MRI sub-study (83% and 84% for PDFF; 58% and 57% for change in PDFF). The reason for the lower sensitivity in the current study is unclear, as identical qualification procedures and imaging protocols were used. One possible explanation is that two sites changed scanners between baseline and EOT, which may have reduced the precision with which PDFF change can be measured, but this would not have impacted the performance of PDFF at baseline. Another possible explanation is that children may be less cooperative than adults. Image quality was checked as MRIs were done, but it is still possible that, although scans were deemed adequate, image quality may have been reduced compared to adults.

We found no evidence of interaction between dichotomized subgroups (age, sex, lobular inflammation, fibrosis, scanner field strength, and time between baseline biopsy and baseline MRI) for MRI-determined steatosis vs. histologic steatosis at baseline or longitudinally at or better than a conservative $p < 0.01$ cutoff significance level, although we acknowledge that the power to detect subgroup effects in this study was low. The lack of an interaction for magnetic field strength was expected based on theory, and is in agreement with other studies (32).

Results regarding the role of fibrosis in the previous literature differ from ours with regard to the the relationship between MRI-determined PDFF and histologic steatosis. We found that the correlation of PDFF with histologic steatosis was stronger for higher fibrosis stages, for histologic grade 1 vs. 2–3 (interaction p-value 0.39), and for histologic grade 1–2 vs. 3 (interaction p-value 0.03), although neither reached our conservative significance cutoff of $p < 0.01$. Idilman et al (45) found in a study of 70 adults with diagnosed NAFLD (60% with no fibrosis) that this correlation was stronger when fibrosis was absent ($r = 0.86$) than when it was present ($r = 0.60$; $p = 0.02$). Idilman et al (46) in a second, smaller study (18 adults with diagnosed NAFLD; 67% with no fibrosis) also found that this correlation was stronger when fibrosis was absent ($r = 0.83$) compared to their cohort of all cases ($r = 0.76$). In a larger single-center study, Schwimmer et al (36) (174 children, 29% with no fibrosis) reported that correlation between MRI-determined PDFF and histologic steatosis was weaker in children with higher stages of fibrosis ($r = 0.61$ for stages 2–4, $p < 0.001$) compared to those with stage 1 fibrosis ($r = 0.78$) or no fibrosis ($r = 0.76$). This difference for the two studies by Idilman et al may be due to the lower fibrosis stage distribution in our study (no fibrosis in 26% of subjects), or to NAFLD differences between children and adults. It is not clear why our results differ from those reported by Schwimmer et al (36); perhaps other factors may affect this correlation, such as differences across sites.

Results regarding the role of sex in the previous literature also differ from ours with regard to the the relationship between MRI-determined PDFF and histologic steatosis. Schwimmer et al (36) reported that this correlation was stronger in girls ($r = 0.86$) than in boys ($r = 0.70$; $p < 0.01$). We found the opposite, that these correlations were stronger in boys than in girls, although again the comparison of these subgroups in our analysis did not meet our threshold cutoff of $p < 0.01$. Perhaps this reflects under-powering of our study, although there is no physical basis to expect correlations of MRI-determined PDFF vs. histologic steatosis to be affected by sex.

Strengths of this study were the prospective longitudinal design, the well-characterized cohort of children with a racial/ethnic makeup representative of pediatric NAFLD in the United States, the availability of paired biopsies at baseline and at EOT, rigorous central scoring of histology by the NASH CRN Pathology Committee and central reading of MRIs by the NASH CRN RCC, and the utilization of a range of MRI scanner manufacturers at two field strengths and across multiple scanner types and study sites. Thus, our study results are likely to be generalizable to the entire pediatric NAFLD population in the United States, while also establishing the feasibility of MRI in multi-center pediatric trials.

While not in the design, nor a study goal, a possible limitation of this study is that none of the participants included in this MRI sub-study had histologic steatosis grade 0 at baseline, as, by definition, to be enrolled, participants all had NAFLD. The relevance of this is that PDFF cut-off values that might be used at baseline to separate participants with NAFLD, from those without NAFLD cannot be defined by this study. A detailed assessment of proposed cut-points for this purpose was reported in the MRI Rosetta Stone Project (36). Further multi-center studies in populations including participants with grade 0 steatosis are needed to continue to define associations of PDFF with hepatic steatosis in grade 0–1 range.

A more relevant limitation of this study was that not all participants enrolled in the CyNCh trial had MRIs, and that not all participants who had baseline MRIs also had EOT MRIs. Participation in the MRI sub-study was reduced in part because clinical trial sites for the sub-study were brought in after the main study started, as they completed the MRI qualification process. The only statistically significant differences between participants who did and who did not participate in the MRI sub-study were clinical site ($p < 0.0001$), and slightly lower mean fasting serum glucose (86 vs. 90 mg/dL; $p = 0.04$) among subjects who did not participate in the MRI sub-study. Comparing baseline characteristics between participants with MRI at baseline and EOT ($n = 85$) vs. those with MRI only at baseline ($n = 25$), statistically significant differences were clinical site ($p < 0.0001$), higher mean age (14.2 vs. 12.7 yrs; $p = 0.02$) and lower mean alkaline phosphatase (174 vs. 237 U/L; $p = 0.008$) in children without EOT MRI. Potentially confounding factors such as those in Table 2 were not investigated because the study was not powered to permit those investigations. While unlikely to have affected study results, there may have been a metabolic difference in these participants that is not accounted for.

In conclusion, in a well-controlled, multi-center study in children with NAFLD, PDFF and change in PDFF estimated by multi-echo MRI at sites using scanners from different manufacturers and of different field strengths showed high diagnostic accuracy (i.e.,

agreement) with histologic steatosis grade and change in histologic steatosis grade, respectively.

Acknowledgments

Financial Support

We would like to acknowledge support from the following sources:

NIDDK U01 DK061713

NIDDK U01 DK061718

NIDDK U01 DK061728

NIDDK U01 DK061730

NIDDK U01 DK061731

NIDDK U01 DK061732

NIDDK U01 DK061734

NIDDK U01 DK061737

NIDDK U01 DK061738

NCATS UL1 TR000004

NCATS UL1 TR000006

NCATS UL1 TR000040

NCATS UL1 TR000077

NCATS UL1 TR000100

NCATS UL1 TR000150

NCATS UL1 TR000423

NCATS UL1 TR000424

NCATS UL1 TR000448

NCATS UL1 TR000454

This study was supported in part by the Intramural Research Program of the NIH, National Cancer Institute.

List of Abbreviations:

MRI	magnetic resonance imaging
PDFF	proton density fat fraction
CyNCh	Cysteamine Bitartrate Delayed-Release for the Treatment of Nonalcoholic Fatty Liver Disease (NAFLD) in Children
EOT	end of treatment

SD	standard deviation
AUROC	area under receiving operating characteristic curve
CI	confidence interval
NAFLD	nonalcoholic fatty liver disease
NASH	nonalcoholic steatohepatitis
FLINT	Farnesoid X Receptor Ligand Obeticholic Acid in NASH Treatment
CRN	Clinical Research Network
RCC	Radiology Coordinating Center

REFERENCES

1). 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]

2). Schwimmer JB, Pardee PE, Lavine JE, Blumkin AK, Cook S. Cardiovascular risk factors and the metabolic syndrome in pediatric nonalcoholic fatty liver disease. Circulation 2008;118:277–283 [PubMed: 18591439]

3). Rubinstein E, Lavine JE, Schwimmer JB. Hepatic, cardiovascular, and endocrine outcomes of the histological subphenotypes of nonalcoholic fatty liver disease. Semin Liver Dis 2008;28:380–385 [PubMed: 18956294]

4). Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology 2006;44:865–873 [PubMed: 17006923]

5). Sanches PL, de Piano A, Campos RM, Carnier J, de Mello MT, Elias N, et al. Association of nonalcoholic fatty liver disease with cardiovascular risk factors in obese adolescents: the role of interdisciplinary therapy. J Clin Lipidol 2014;8:265–272 [PubMed: 24793347]

6). Kim NH, Park J, Kim SH, Kim YH, Kim DH, Cho GY, et al. Nonalcoholic fatty liver disease, metabolic syndrome and subclinical cardiovascular changes in the general population. Heart 2014;100:938–943 [PubMed: 24721975]

7). Newton KP, Hou J, Crimmins NA, Lavine JE, Barlow SE, Xanthakos SA, et al. Prevalence of prediabetes and type 2 diabetes in children with nonalcoholic fatty liver disease. JAMA Pediatr 2016 10 3;170(10):e161971. doi: 10.1001/jamapediatrics.2016.1971. Epub 2016 Oct 3. [PubMed: 27478956]

8). Feldstein AE, Charatcharoenwitthaya P, Treeprasertsuk S, Benson JT, Enders FB, Angulo P. The natural history of non-alcoholic fatty liver disease in children: a follow-up study for up to 20 years. Gut 2009;58:1538–1544 [PubMed: 19625277]

9). Alexander J, Torbenson M, Wu TT, Yeh MM. Non-alcoholic fatty liver disease contributes to hepatocarcinogenesis in non-cirrhotic liver: a clinical and pathological study. J Gastroenterol Hepatol 2013;28:848–854 [PubMed: 23302015]

10). Schwimmer JB, Middleton MS, Deutsch R, Lavine JE. A phase 2 clinical trial of metformin as a treatment for non-diabetic paediatric non-alcoholic steatohepatitis. Aliment Pharmacol Ther 2005;21:871–879 [PubMed: 15801922]

11). Dam-Larsen S, Franzmann M, Andersen IB, Christoffersen P, Jensen LB, Sørensen TI, et al. Long term prognosis of fatty liver: risk of chronic liver disease and death. Gut 2004;53:750–755 [PubMed: 15082596]

12). Townsend SA, Newsome PN. Non-alcoholic fatty liver disease in 2016. Br Med Bull 2016;119:143–156 [PubMed: 27543499]

13). Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Nonalcoholic steatohepatitis clinical research network. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–1321 [PubMed: 15915461]

14). Bonekamp S, Tang A, Mashhood A, Wolfson T, Changchien C, Middleton MS, et al. Spatial distribution of MRI-determined hepatic proton density fat fraction in adults with nonalcoholic fatty liver disease. J Magn Reson Imaging 2014;39:1525–1532 [PubMed: 24987758]

15). Sasso M, Beaugrand M, de Ledinghen V, Douvin C, Marcellin P, Poupon R, et al. Controlled attenuation parameter (CAP): a novel VCTE guided ultrasonic attenuation measurement for the evaluation of hepatic steatosis: preliminary study and validation in a cohort of patients with chronic liver disease from various causes. Ultrasound Med Biol 2010;36:1825–1835 [PubMed: 20870345]

16). Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, et al. Magnetic resonance imaging more accurately classifies steatosis and fibrosis in patients with nonalcoholic fatty liver disease than transient elastography. Gastroenterology 2016;150:626–637 [PubMed: 26677985]

17). Taouli B, Serfaty L. Magnetic resonance imaging/elastography is superior to transient elastography for detection of liver fibrosis and fat in nonalcoholic fatty liver disease (editorial). Gastroenterology 2016;150:553–556 [PubMed: 26820053]

18). Bydder M, Yokoo T, Hamilton G, Middleton MS, Chavez AD, Schwimmer JB, et al. Relaxation effects in the quantification of fat using gradient echo imaging. Magn Reson Imaging 2008;26:347–359 [PubMed: 18093781]

19). Yokoo T, Collins JM, Hanna RF, Bydder M, Middleton MS, Sirlin CB. Effects of intravenous gadolinium administration and flip angle on the assessment of liver fat signal fraction with opposed-phase and in-phase imaging. J Magn Reson Imaging 2008;28:246–251 [PubMed: 18581393]

20). Yokoo T, Bydder M, Hamilton G, Middleton MS, Gamst AC, Wolfson T, 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 [PubMed: 19221054]

21). Yokoo T, Shiehmorteza M, Hamilton G, Wolfson T, Schroeder ME, Middleton MS, et al. Estimation of hepatic proton-density fat fraction by magnetic resonance imaging at 3T. Radiology 2011;258:749–759 [PubMed: 21212366]

22). Middleton MS, Haufe W, Hooker J, Borga M, Leinhard OG, Romu T, et al. Quantifying abdominal adipose tissue and thigh muscle volume and hepatic proton density fat fraction: repeatability and accuracy of an MR imaging-based, semiautomated analysis method. Radiology 2017;283:438–449 [PubMed: 28278002]

23). Tang A, Tan J, Sun M, Hamilton G, Bydder M, Wolfson T, 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]

24). Permutt Z, Le TA, Peterson MR, Seki E, Brenner DA, Sirlin C, et al. 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. Aliment Pharmacol Ther 2012;36:22–29 [PubMed: 22554256]

25). Le TA, Chen J, Changchien C, Peterson MR, Kono Y, Patton H, 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]

26). Noureddin M, Lam J, Peterson MR, Middleton M, Hamilton G, Le TA, et al. Utility of magnetic resonance imaging versus histology for quantifying changes in liver fat in nonalcoholic fatty liver disease trials. Hepatology 2013;58:1930–1940 [PubMed: 23696515]

27). Loomba R, Sirlin CB, Ang B, Bettencourt R, Jain R, Salotti J, et al. Ezetimibe for the treatment of nonalcoholic steatohepatitis: assessment by novel magnetic resonance imaging and magnetic resonance elastography in a randomized trial (MOZART trial). Hepatology 2015;61:1239–1250 [PubMed: 25482832]

28). Kang GH, Cruite I, Shiehmorteza M, Wolfson T, Gamst AC, Hamilton G, 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]

29). Negrete LM, Middleton MS, Clark L, Wolfson T, Gamst AC, Lam J, et al. Inter-examination precision of magnitude-based magnetic resonance imaging for estimation of segmental hepatic proton density fat fraction (PDFF) in obese subjects. J Magn Reson Imaging 2013;39:1265–1271 [PubMed: 24136736]

30). Artz NS, Haufe WM, Hooker CA, Hamilton G, Wolfson T, Campos GM, et al. Reproducibility of MR-based liver fat quantification across field strength: Same-day comparison between 1.5T and 3T in obese subjects. J Magn Reson Imaging 2015;42:811–817 [PubMed: 25620624]

31). Mashhood A, Railkar R, Yokoo T, Levin Y, Clark L, Fox-Bosetti S, et al. Reproducibility of hepatic fat fraction measurement by magnetic resonance imaging. J Magn Reson Imaging 2013;37:1359–1370 [PubMed: 23172799]

32). Yokoo T, Serai SD, Pirasteh A, et al. Linearity, bias, and precision of hepatic proton-density fat fraction measurements by magnetic resonance imaging: a meta-analysis. Radiology, 2017 (accepted June 8, 2017)

33). Tang A, Desai A, Hamilton G, Wolfson T, Gamst A, Lam J, et al. Accuracy of MR imaging-estimated proton density fat fraction for classification of dichotomized histologic steatosis grades in nonalcoholic fatty liver disease. Radiology 2015;274:416–425 [PubMed: 25247408]

34). Neuschwander-Tetri BA, Loomba R, Sanyal AJ, Lavine JE, Van Natta ML, Abdelmalek MF, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. Lancet 2015;385:956–965 [PubMed: 25468160]

35). Middleton MS, Heba ER, Hooker CA, Bashir MR, Fowler KJ, Sandrasegaran K, Brunt EM, Kleiner DE, Doo E, Van Natta ML, Tonascia J, Lavine JE, Neuschwander-Tetri BA, Sanyal A, Loomba R, Sirlin CB. 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 (accepted)

36). Schwimmer JB, Middleton MS, Behling C, Newton KP, Awai HI, Paiz MN, et al. Magnetic resonance imaging and liver histology as biomarkers of hepatic steatosis in children with nonalcoholic fatty liver disease. Hepatology 2015;61:1887–1895 [PubMed: 25529941]

37). Achmad E, Yokoo T, Hamilton G, Heba ER, Hooker JC, Changchien C, et al. Feasibility of and agreement between MR imaging and spectroscopic estimation of hepatic proton density fat fraction in children with known or suspected nonalcoholic fatty liver disease. Abdom Imaging 2015;40:3084–3090 [PubMed: 26205992]

38). Tyagi A, Yeganeh O, Levin Y, Hooker JC, Hamilton GC, Wolfson T, et al. Intra- and inter-examination repeatability of magnetic resonance spectroscopy, magnitude-based MRI, and complex-based MRI for estimation of hepatic proton density fat fraction in overweight and obese children and adults. Abdom Imaging 2015;40:3070–3077 [PubMed: 26350282]

39). Zand KA, Shah A, Heba E, Wolfson T, Hamilton G, Lam J, et al. Accuracy of multiecho magnitude-based MRI (M-MRI) for estimation of hepatic proton density fat fraction (PDFF) in children. J Magn Reson Imaging 2015;42:1223–1232 [PubMed: 25847512]

40). Rehm JL, Wolfgram PM, Hernando D, Eickhoff JC, Allen DB, Reeder SB. Proton density fat-fraction is an accurate biomarker of hepatic steatosis in adolescent girls and young women. Eur Radiol 2015;25:2921–2930 [PubMed: 25916386]

41). Schwimmer JB, Lavine JE, Wilson LA, Neuschwander-Tetri BA, Xanthakos SA, Kohli R, et al. In children with nonalcoholic fatty liver disease, cysteamine bitartrate delayed release improves liver enzymes but does not reduce disease activity scores. Gastroenterology 2016;151:1141–1154 [PubMed: 27569726]

42). Manning PM, Hamilton G, Wang K, Park C, Hooker JC, Wolfson T, et al. Agreement between region-of-interest and parametric map-based hepatic proton density fat fraction estimation in adults with chronic liver disease. Abdom Radiol 2017;42:833–841

43). Hamilton G, Yokoo T, Bydder M, Cruite I, Schroeder ME, Sirlin CB, et al. In vivo characterization of the liver fat 1H magnetic resonance spectrum. NMR Biomed 2011;24:784–790 [PubMed: 21834002]

44). Lachenbruch P, Mickey M. Estimation of error rates in discriminant analysis. Technometrics 1968;10:1–11

45). Idilman IS, Aniktar H, Idilman R, et al. Hepatic steatosis: quantification by proton density fat fraction with MR imaging versus liver biopsy. Radiology 2013;267:767–775 [PubMed: 23382293]

46). Idilman IS1, Keskin O, Elhan AH, Idilman R, Karcaaltincaba M. Impact of sequential proton density fat fraction for quantification of hepatic steatosis in nonalcoholic fatty liver disease. Scand J Gastroenterol 2014;49:617–624 [PubMed: 24694249]

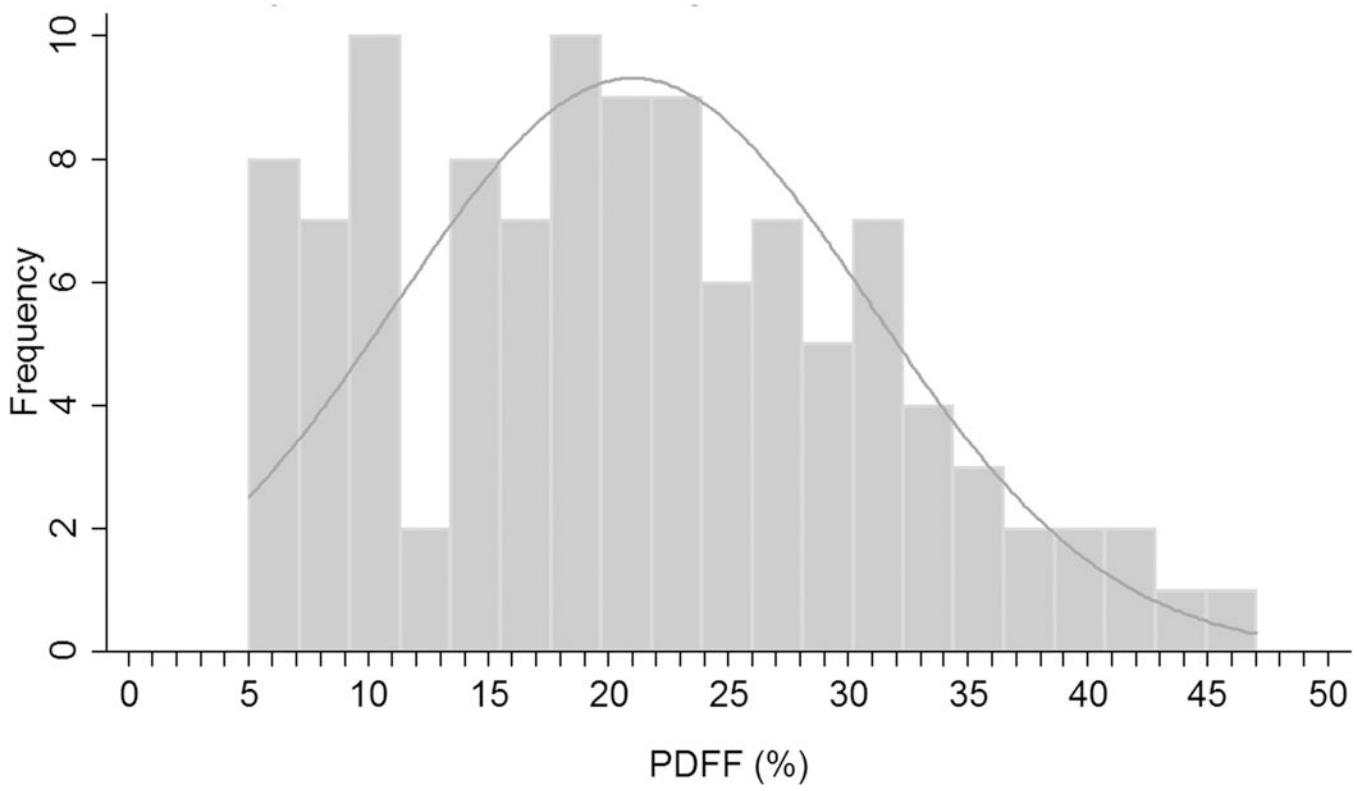

Figure 1.
PDFF distribution of study population at baseline.

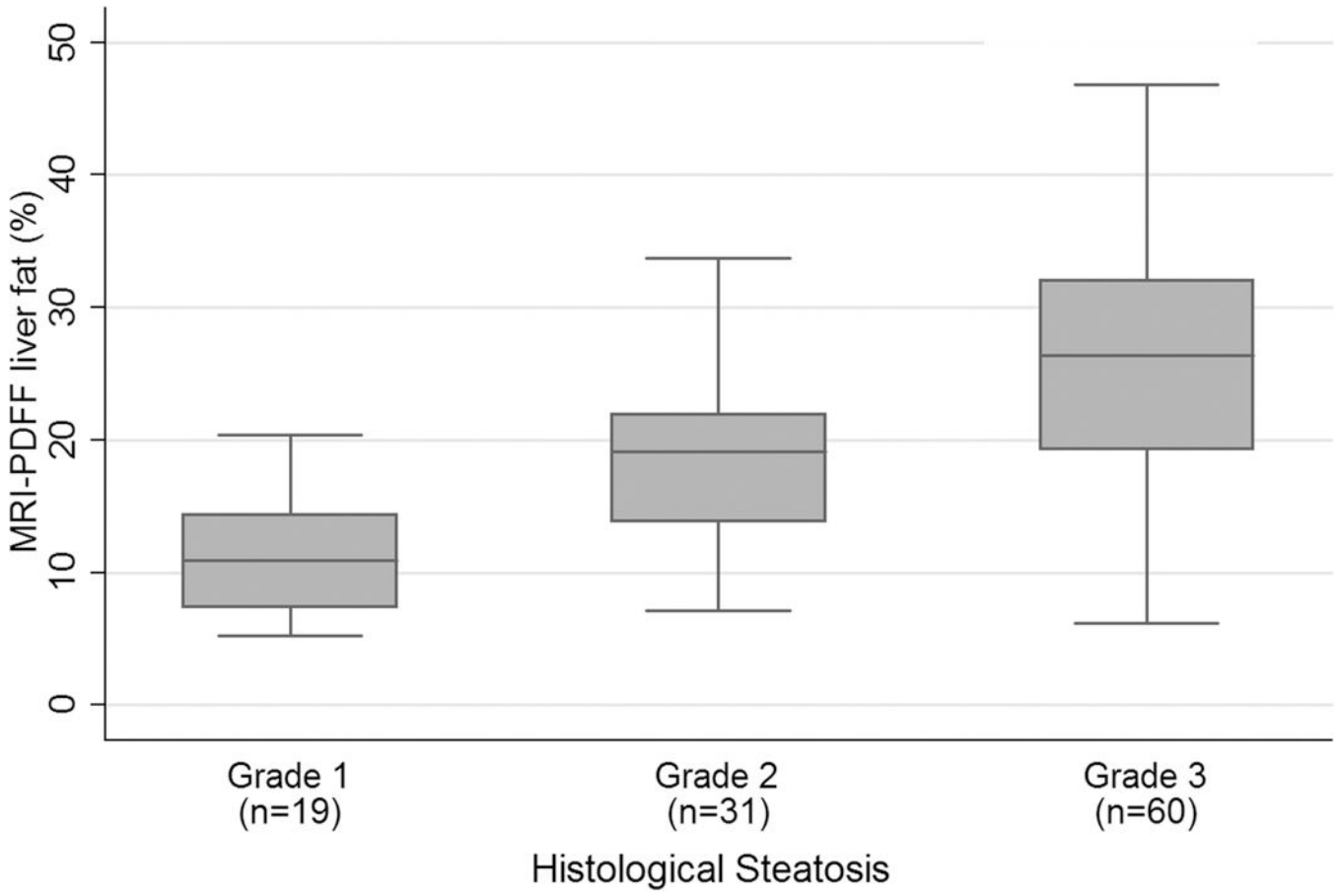

Figure 2.
Bar and whisker plot at baseline of PDFF vs. steatosis grade.

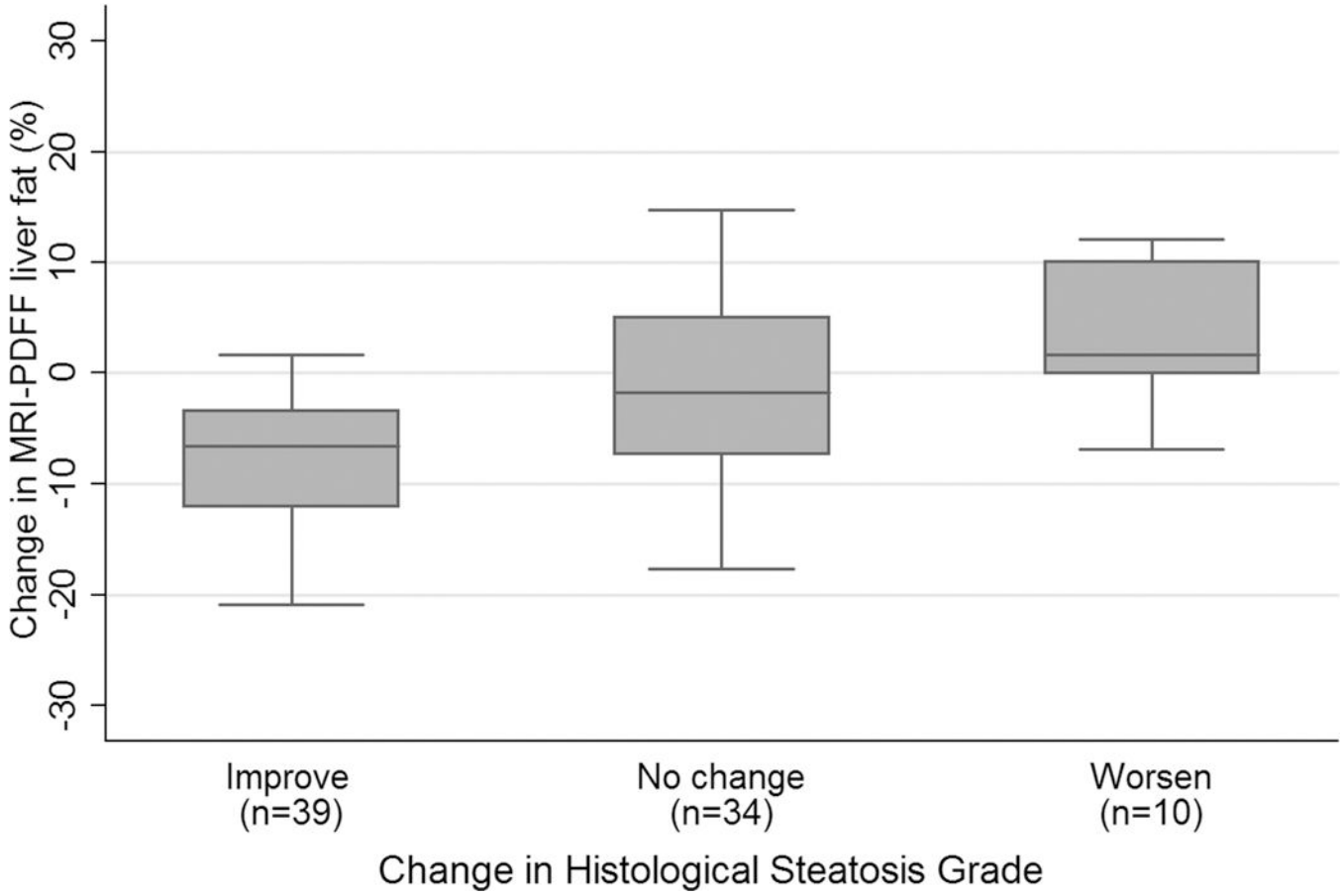

Figure 3.

Bar and whisker plot of change in PDFF vs. change in histologic hepatic steatosis grade. Mean (± SD) PDFF change values for steatosis grade reduction, no change in steatosis grade , and increase in steatosis grade were, respectively: −7.8 ± 6.3% (n = 39), −1.2 ± 7.8% (n = 34), and 4.9 ± 5.0 % (n = 10).

Table 1.

MRI scanners and techniques

Clinical trial site MR scanners	**1.5T**	**3T**
Site A		General Electric HDxt
Site B	General Electric Signa HDx, and Philips Ingenia	-
Site C	General Electric Signa HDxt	-
Site D	Siemens Aera	Siemens TrioTim
Site E	-	Siemens TrioTim
Site F	Siemens Aera	-
Site G	Siemens Avanto	-
Site H	-	General Electric Signa HDxt
Site I	Siemens Avanto	-
Site J	-	-
Parameter		
TR (ms)	≥ 120	≥ 120
First TE (ms)	2.3	1.15
Delta TE (ms)	2.3	1.15
Number of echoes	6	6
Flip angle (°)	10	10
Bandwidth (Hz/Px)	≥ 500	≥ 1,000
Slice thickness (mm)	8 or 10	8 or 10
Slice gap (mm)	0	0
Phase encoding steps	192	128
Frequency encoding steps	192	128

Notes: TR = repetition time; TE = time to echo; General Electric Healthcare (Waukesha, WI, USA); Philips Healthcare (Best, The Netherlands), Siemens Healthcare (Erlangen, Germany)

Hepatology. Author manuscript; available in PMC 2019 March 01.

Table 2.

Baseline characteristics of: a) those obtaining vs. not obtaining baseline MRI, and b) of those obtaining baseline MRI, those obtaining vs. not obtaining EOT MRI

	Baseline vs. no baseline MRI (n = 169)			**EOT vs. no EOT MRI for those who had baseline MRI (n = 110)**		
	Number participants who had baseline MRIs (n=110)	**Number participants who did not have baseline MRIs (n=59)**	**p-value**	**Number participants who had EOT MRIs (n=85)**	**Number participants who did not have EOT MRIs (n=25)**	**p-value**
Study site			< 0.0001			< 0.0001
Site A	10	13		4	6	
Site B	16	5		14	2	
Site C	17	4		13	4	
Site D	6	10		2	4	
Site E	12	1		7	5	
Site F	6	6		4	2	
Site G	2	3		2	0	
Site H	40	2		39	1	
Site I	1	3		0	1	
Site J	0	12		0	0	
Treatment group						
cysteamine bitartrate delayed release – no. (%)	53 (48.2)	35 (59.3)	0.20	39 (45.9)	14 (56.0)	0.50
Demographics						
Age (yrs)	13 ± 3	13 ± 3	0.68	12.7±2.5	14.2±2.8	0.02
Male sex - no. (%)	78 (70.9)	41 (69.5)	0.86	60 (70.6)	18 (72.0)	1.00
White race - no. (%)	64 (58.2)	41 (69.5)	0.18	45 (83.3)	19 (100.0)	0.10
Hispanic ethnicity - no. (%)	78 (70.9)	46 (78.0)	0.37	60 (70.6)	18 (72.0)	1.00
Liver enzymes						
Aspartate aminotransferase - U/L	64 ± 48	84 ± 74	0.07	64 ± 52	63 ± 32	0.86

	Baseline vs. no baseline MRI (n = 169)			EOT vs. no EOT MRI for those who had baseline MRI (n = 110)		
	Number participants who had baseline MRIs (n=110)	Number participants who did not have baseline MRIs (n=59)	p-value	Number participants who had EOT MRIs (n=85)	Number participants who did not have EOT MRIs (n=25)	p-value
Alanine aminotransferase – U/L	110 ± 85	146 ± 124	0.05	108 ± 91	119 ± 64	0.48
γ-Glutamyltransferase – U/L	48 ± 34	45 ± 25	0.46	46 ± 34	57 ± 36	0.19
Alkaline phosphatase - U/L	223 ± 106	214 ± 114	0.63	237 ± 105	174 ± 97	0.008
Total bilirubin – mg/dL	0.5 ± 0.3	0.6 ± 0.3	0.09	0.46 ± 0.33	0.58 ± 0.23	0.05
Metabolic factors						
Weight - kg	85 ± 25	85 ± 25	0.92	83 ± 26	89 ± 24	0.30
Body-mass index - kg/m^2	32 ± 6	33 ± 7	0.50	32 ± 6	33 ± 6	0.27
Waist circumference - cm	104 ± 14	102 ± 16	0.34	104 ± 15	104 ± 13	0.80
Waist to hip ratio	0.98 ± 0.06	0.97 ± 0.08	0.26	0.99 ± 0.06	0.97 ± 0.07	0.48
Fasting serum glucose - mg/dL	86 ± 12	90 ± 11	0.04	85 ± 12	89 ± 11	0.19
Insulin - umol/mL	37 ± 34	47 ± 93	0.40	35 ± 28	44 ± 51	0.40
HOMA-IR - (mg/dL * umol/mL) / 405	7.8 ± 7.7	10.2 ± 19.2	0.35	7.2 ± 5.8	9.7 ± 12.1	0.34
Hemoglobin A1c - %	5.4 ± 0.4	5.4 ± 0.4	0.47	5.5 ± 0.5	5.4 ± 0.4	0.27
Systolic blood pressure - mmHg	119 ± 11	120 ± 11	0.44	119 ± 12	119 ± 10	0.87
Diastolic blood pressure - mmHg	67 ± 9	67 ± 8	0.66	67 ± 9	69 ± 8	0.26
Comorbidities						
Hyperlipidemia - no. (%)	19 (17.3)	6 (10.2)	0.26	17 (20.0)	2 (8.0)	0.23
Hypertension - no. (%)	12 (10.9)	3 (5.1)	0.26	10 (11.8)	2 (8.)	0.73
Diabetes - no. (%)	6 (5.4)	2 (3.4)	0.71	6 (7.1)	0 (0.0)	0.33
Liver histology findings						
Definite steatohepatitis - no. (%)	30 (27.3)	17 (28.8)		22 (25.9)	8 (32.0)	0.61
Fibrosis - stage[1]	1.2 ± 1.1	1.2 ± 1.1	0.60	1.3 ± 1.0	1.2 ± 1.1	0.82
NAFLD Activity Score[2]	4.7 ± 1.4	4.7 ± 1.5	0.96	4.6 ± 1.4	4.8 ± 1.2	0.48
Hepatocellular ballooning - score	0.6 ± 0.7	0.6 ± 0.8	0.77	0.6 ± 0.8	0.6 ± 0.6	0.94

	Baseline vs. no baseline MRI (n = 169)			EOT vs. no EOT MRI for those who had baseline MRI (n = 110)		
	Number participants who had baseline MRIs (n=110)	Number participants who did not have baseline MRIs (n=59)	p-value	Number participants who had EOT MRIs (n=85)	Number participants who did not have EOT MRIs (n=25)	p-value
Steatosis - score	2.4 ± 0.8	2.4 ± 0.7	0.77	2.3 ± 0.8	2.5 ± 0.7	0.26
Lobular inflammation - score	1.7 ± 0.7	1.7 ± 0.7	0.62	1.7 ± 0.7	1.7 ± 0.6	0.99
Portal inflammation - score[3]	1.1 ± 0.5	1.0 ± 0.5	0.05	1.2 ± 0.5	1.1 ± 0.6	0.54

Notes: Plus-minus values are means ± SD; BMI = body mass index; PDFF = proton density fat fraction; SD = standard deviation.

[1]Fibrosis was assessed on a scale of 0 to 4, with higher scores indicating more severe fibrosis

[2]NAFLD Activity Score was assessed on a scale of 0 to 8, with higher scores indicating more severe disease; the components of this measure are steatosis (assessed on a scale of 0 to 3), lobular inflammation (assessed on a scale of 0 to 3), and hepatocellular ballooning (assessed on a scale of 0 to 2)

[3]Portal inflammation was assessed on a scale of 0 to 2 with higher scores indicating more severe inflammation

Table 3.

Linear regression of PDFF with histologic components at baseline, and change in PDFF with change in histologic components at 52 weeks

Histologic component	PDFF (%) at baseline (n=110)			Change in PDFF (%) at 52 weeks (n=83)		
	β*	95% CI	p-value	β†	95% CI	p-value
Steatosis (score)	0.068	0.056, 0.081	< 0.001	0.057	0.034, 0.079	< 0.001
Lobular inflammation (score)	0.012	−0.002, 0.025	0.10	0.019	0.000, 0.037	0.73
Portal inflammation (score)	0.011	−0.005, 0.018	0.24	0.005	−0.011, 0.015	0.77
Hepatocellular ballooning (score)	0.003	−0.012, 0.019	0.66	0.013	−0.010, 0.025	0.40
Fibrosis (stage)	0.013	−0.010, 0.036	0.28	0.008	−0.023, 0.030	0.80

Notes: PDFF = proton density fat fraction; CI = confidence interval

*Mean difference in histologic component per 1% increase in PDFF

†Mean change in histologic component per 1% increase in change of PDFF adjusted for baseline value of histologic component

Table 4.

Diagnostic accuracy of PDFF for classifying steatosis

Cross-sectional steatosis grade classification (n=110)	PDFF threshold (%) at 90% specificity		Sensitivity (%)	PPV (%)	NPV (%)	Correctly classified (%)	AUROC (95% CI)
1 vs. 2–3	**17.5**		**74 (67/91)**	**97 (67/69)**	**41 (17/41)**	**76 (84/110)**	**0.87 (0.80, 0.94)**
1–2 vs. 3	**23.3**		**60 (36/60)**	**88 (36/41)**	**65 (45/79)**	**74 (81/100)**	**0.79 (0.70, 0.87)**
Longitudinal steatosis grade change classification (n=83)	**Mean (SD) PDFF change (%)**	**Cutoff PDFF change at 90% specificity (%)**	**Sensitivity (%)**	**PPV (%)**	**NPV (%)**	**Correctly classified (%)**	**AUROC (95% CI)**
Decrease in steatosis grade (improvement) (47%)	−7.8 ± 6.3	−11.0	31 (12/39	75 (12/16)	60 (40/67)	63 (52/83)	0.76 (0.66, 0.87)
No change in steatosis grade (41%)	−1.2 ± 7.8	-	-	-	-		-
Increase in steatosis grade (worsening) (12%)	4.9 ± 5.0	5.5	40 (4/10)	33 (4/12)	92 (65/71)	83 (69/83)	0.83 (0.73, 0.92)

Notes: PDFF = proton density fat fraction; PPV = positive predictive value; NPV = negative predictive value; AUROC = area under receiver operating characteristic curve; CI = confidence interval; SD = standard deviation

Table 5.

Subgroup analyses of regression of baseline steatosis grade on baseline PDFF

Event	Subgroup	N	AUROC	Odds of event per 1% increase in PDFF	95% CI	p-value
Steatosis grade 2/3 vs 1						
	Age 8–12 yrs	52	0.88	1.24	1.07, 1.43	0.004
	Age 13–18 yrs	58	0.89	1.34	1.11, 1.62	0.003
	Interaction †					0.51
	Female	32	0.95	1.44	1.02, 2.05	0.04
	Male	78	0.86	1.25	1.10, 1.43	0.0008
	Interaction					0.45
	Grade 1 lobular inflammation at BL	44	0.84	1.24	1.06, 1.47	0.009
	Grade 2/3 lobular inflammation at BL	66	0.91	1.32	1.11 1.56	0.002
	Interaction					0.64
	No fibrosis at BL	29	0.78	1.23	0.97, 1.56	0.09
	Any fibrosis at BL	81	0.92	1.40	1.17, 1.66	0.0002
	Interaction					0.39
	1.5T MRI scanner	46	0.88	1.28	1.08, 1.50	0.003
	3T MRI scanner	64	0.89	1.28	1.09, 1.52	0.003
	Interaction					0.95
	BL MRI ≤ 60 days of BL biopsy	54	0.96	1.53	1.18, 1.97	0.001
	BL MRI > 60 days from BL biopsy	56	0.82	1.18	1.04, 1.35	0.01
	Interaction					0.08
Steatosis grade 3 vs 1/2						
	Age 8–12 yrs	52	0.76	1.12	1.04, 1.20	0.003

Event	Subgroup	N	AUROC	Odds of event per 1% increase in PDFF	95% CI	p-value
	Age 13–18 yrs	58	0.84	1.21	1.09, 1.33	0.0002
	Interaction					0.20
	Female	32	0.78	1.13	1.02, 1.24	0.01
	Male	78	0.82	1.18	1.09, 1.27	<0.0001
	Interaction					0.52
	Grade 1 lobular inflammation at BL	44	0.78	1.13	1.04, 1.24	0.004
	Grade 2/3 lobular inflammation at BL	66	0.82	1.16	1.07, 1.26	0.0002
	Interaction					0.68
	No fibrosis at BL	29	0.69	1.08	1.00, 1.16	0.06
	Any fibrosis at BL	81	0.85	1.22	1.12, 1.34	<0.0001
	Interaction					0.03
	1.5T MRI scanner	46	0.78	1.15	1.05, 1.26	0.003
	3T MRI scanner	64	0.81	1.15	1.07, 1.24	0.0003
	Interaction					0.99
	BL MRI ≤ 60 days of BL biopsy	54	0.79	1.16	1.06, 1.27	0.001
	BL MRI > 60 days from BL biopsy	56	0.82	1.16	1.07, 1.26	0.0004
	Interaction					0.99

[†]Test of interaction of subgroup with PDFF

Table 6.

Subgroup analyses of regression of worsening or improvement in steatosis grade on 52-week change in PDFF

Event	Subgroup	N	AUROC	Odds of event per 1% increase in change in PDFF	95% CI	p-value
Worsening of steatosis grade						
	Age 8–12 yrs	43	0.94	1.24	1.00, 1.56	0.05
	Age 13–18 yrs	40	0.77	1.16	1.01, 1.33	0.03
	Interaction					0.60
	Female	25	0.86	1.16	0.98, 1.38	0.08
	Male	58	0.86	1.22	1.05, 1.41	0.008
	Interaction					0.68
	Grade 1 lobular inflammation at BL	34	0.69	1.08	0.94, 1.22	0.27
	Grade 2/3 lobular inflammation at BL	49	0.96	1.46	1.12, 1.92	0.006
	Interaction					0.04
	No fibrosis at BL	20	NC *	NC	NC	NC
	Any fibrosis at BL	63	0.88	1.20	1.07, 1.35	0.002
	Interaction					NC
	1.5T MRI scanner	32	0.90	1.29	1.04, 1.59	0.02
	3T MRI scanner	51	0.82	1.16	1.01, 1.33	0.04
	Interaction					0.40
	BL MRI ≤ 60 days of BL biopsy	39	0.88	1.32	1.03, 1.69	0.03
	BL MRI > 60 days from BL biopsy	44	0.86	1.19	1.02, 1.39	0.02
	Interaction					0.50
Improvement in steatosis grade						

Event	Subgroup	N	AUROC	Odds of event per 1% increase in change in PDFF	95% CI	p-value
	Age 8–12 yrs	43	0.70	0.88	0.79, 0.97	0.01
	Age 13–18 yrs	40	0.80	0.83	0.73, 0.95	0.006
	Interaction					0.54
	Female	25	0.81	0.82	0.69, 0.98	0.02
	Male	58	0.76	0.86	0.78, 0.94	0.002
	Interaction					0.67
	Grade 1 lobular inflammation at BL	34	0.81	0.81	0.70, 0.94	0.005
	Grade 2/3 lobular inflammation at BL	49	0.71	0.87	0.79, 0.97	0.008
	Interaction					0.40
	No fibrosis at BL	20	0.68	0.89	0.76, 1.04	0.14
	Any fibrosis at BL	63	0.82	0.83	0.75, 0.92	0.0003
	Interaction					0.49
	1.5T MRI scanner	32	0.71	0.88	0.78, 1.00	0.05
	3T MRI scanner	51	0.82	0.81	0.72, 0.92	0.0007
	Interaction					0.34
	BL MRI ≤ 60 days of BL biopsy	39	0.77	0.86	0.77, 0.97	0.01
	BL MRI > 60 days from BL biopsy	44	0.79	0.83	0.74, 0.94	0.002
	Interaction					0.66

*Not calculable due to no events