Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Magn Reson Imaging. Author manuscript; available in PMC 2020 May 01.

Published in final edited form as:

J Magn Reson Imaging. 2019 May ; 49(5): 1456–1466. doi:10.1002/jmri.26312.

Hepatic R2* is more strongly associated with proton density fat fraction than histologic liver iron scores in patients with nonalcoholic fatty liver disease

Mustafa R. Bashir, MD[1,2], **Tanya Wolfson, MA**[3], **Anthony C. Gamst, PhD**[3], **Kathryn J. Fowler, MD**[4], **Michael Ohliger, MD, PhD**[5], **Shetal N. Shah, MD**[6], **Adina Alazraki, MD**[7], **Andrew T. Trout, MD**[8], **Cynthia Behling, MD, PhD**[9], **Daniela S. Allende, MD**[10], **Rohit Loomba, MD, MHSs**[11], **Arun Sanyal, MD**[12], **Jeffrey Schwimmer, MD**[13], **Joel E. Lavine, MD, PhD**[14], **Wei Shen, MD**[15], **James Tonascia, PhD**[16], **Mark L. Van Natta, MHS**[16], **Adrija Mamidipalli, MBBS**[17], **Jonathan Hooker**[17], **Kris V. Kowdley, MD**[18], **Michael S. Middleton, MD, PhD**[17], and **Claude B. Sirlin, MD**[17] **on behalf of the NASH Clinical Research Network (NASH CRN)**

[1]Department of Radiology, Duke University Medical Center, Durham, NC

[2]Center for Advanced Magnetic Resonance Development (CAMRD), Department of Radiology, Duke University Medical Center, Durham, NC

[3]Computational and Applied Statistics Laboratory (CASL), San Diego Supercomputing Center (SDSC), University of California-San Diego, San Diego, CA

[4]Department of Radiology, Washington University, St. Louis, MO

[5]Departments of Radiology and Biomedical Engineering, University of California-San Francisco, San Francisco, CA

[6]Section of Abdominal Imaging and Nuclear Medicine Department, Imaging Institute, Cleveland Clinic, Cleveland, OH

[7]Departments of Radiology and Pediatrics, Emory University School of Medicine/Children's Healthcare of Atlanta, Atlanta, GA

[8]Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

[9]Department of Pathology, University of California-San Diego, La Jolla, CA

[10]Department of Pathology, Cleveland Clinic, Cleveland, OH

[11]NAFLD Research Center, Division of Gastroenterology, Department of Medicine, University of California-San Diego, La Jolla, CA

[12]Division of Gastroenterology, Hepatology and Nutrition, Department of Internal Medicine, Virginia Commonwealth University, Richmond, VA

[13]Department of Pediatrics, University of California-San Diego, San Diego, CA

[14]Department of Pediatrics, Columbia College of Physicians and Surgeons, New York, NY

Corresponding author and originating institution: Mustafa R. Bashir, MD, Department of Radiology, Duke University Medical Center, DUMC 3808, Durham, NC 27710, Phone: (919) 684-7366, Fax: (919) 684-7168, mustafa.bashir@duke.edu.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[15]Division of Pediatric Gastroenterology, Hepatology and Nutrition, Department of Pediatrics and the Institute of Human Nutrition, Columbia University Medical Center, New York, NY

[16]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

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

[18]Liver Care Network and Organ Care Research, Swedish Medical Center, Seattle, WA

Abstract

Background: Liver R2* value is widely used as a measure of liver iron but may be confounded by the presence of hepatic steatosis and other covariates.

Purpose: To identify the most influential covariates for liver R2* values in patients with non-alcoholic fatty liver disease (NAFLD)

Study Type: Retrospective analysis of prospectively acquired data

Population: Baseline data from 204 subjects enrolled in NAFLD/NASH treatment trials

Field Strength: 1.5T and 3T; Chemical-shift encoded multi-echo gradient echo

Assessment: Correlation between liver proton density fat fraction and R2*; assessment for demographic, metabolic, laboratory, MRI-derived, and histological covariates of liver R2*

Statistical Tests: Pearson's and Spearman's correlations; univariate analysis; Gradient Boosting Machines multivariable machine learning method

Results: Hepatic PDFF was the most strongly correlated covariate for R2* at both 1.5T (r=0.652, p<0.0001) and at 3T (r=0.586, p<0.0001). In the GBM analysis, hepatic PDFF was the most influential covariate for hepatic R2*, with relative influences (RIs) of 61.3% at 1.5T and 47.5% at 3T; less influential covariates had RIs of up to 11.5% at 1.5T and 16.7% at 3T. Non-hepatocellular iron was weakly associated with R2* at 3T only (RI 6.7%), and hepatocellular iron was not associated with R2* at either field strength.

Data Conclusion: Hepatic PDFF is the most influential covariate for R2* at both 1.5T and 3T; non-hepatocellular iron deposition is weakly associated with liver R2* at 3T only.

Keywords

Non-alcoholic steatohepatitis; NASH; non-alcoholic fatty liver disease; NAFLD; proton density fat fraction; PDFF; R2*; hepatic steatosis

Introduction

Quantitative imaging biomarkers for characterizing non-alcoholic fatty liver disease (NAFLD) and steatohepatitis (NASH) are increasingly available and validated for multi-center applications (1–3). Two such biomarkers are proton density fat fraction (PDFF) and R2*, both of which are estimated simultaneously by multi-echo gradient echo MRI sequences (4–6). PDFF is a measure of tissue triglyceride concentration and has been validated in prior studies for the quantification of hepatic steatosis, a histopathological

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

hallmark of NAFLD (5, 7, 8). R2* has been validated as a biomarker of liver iron content in populations with iron overload states, in whom magnetic susceptibility from severe iron overload is the main driver of R2* (9, 10). Emerging evidence suggests that liver iron deposition is associated with worse histopathological features of NASH and disease progression, and a non-invasive biomarker for liver iron such as R2* may be useful in NAFLD and NASH patients (11, 12). However, R2* has not been validated as a biomarker of liver iron content in patients with lesser amounts of liver iron, such as patients with NAFLD.

To date, two retrospective single-center studies by Bashir et al and Mamidipalli et al have independently shown that liver PDFF and R2* values are positively correlated *in vivo* at 3T MRI in populations without iron overload (4, 13). Although limited by lack of histological correlation, these studies suggest that in patients without risk factors for severe iron overload (eg patients with NAFLD), R2* may not be an unconfounded measure of liver iron. Instead, R2* may be impacted by liver triglyceride content and/or other, still unidentified, metabolic or histopathologic changes, in addition to liver iron.

The Non-Alcoholic Steatohepatitis Clinical Research Network (NASH CRN) maintains a rich, well-curated database of standardized clinical, laboratory, imaging, and histopathological data collected from patients with NAFLD and NASH enrolled in recent treatment trials (2, 14–17). These data offer an opportunity to assess, in patients with NAFLD, a comprehensive set of potential clinical, laboratory, and histologic covariates for R2* that were unavailable in previous smaller, single-center investigations.

The purpose of this study was to identify and assess the most influential covariates for liver R2* values in patients with NAFLD. Our hypothesis was that, in patients with NAFLD and NASH, R2* may not be an unconfounded measure of liver iron.

Materials and Methods

This was a retrospective, cross-sectional secondary analysis of data acquired prospectively in the multicenter Farnesoid X Receptor Ligand Obeticholic Acid in NASH Treatment (FLINT) and Cysteamine Bitartrate Delayed-Release for the Treatment of NAFLD in Children (CyNCh) trials. The FLINT (NCT01265498) and CyNCh (NCT01529268) trials (termed the “parent trials”) were IRB-approved at each participating site and conducted by the NASH CRN. Subjects in the FLINT trial provided written informed consent at the time of enrollment, while for the CyNCh trial, all subjects provided written informed assent and a parent or guardian provided written informed consent. Both parent trials were multi-center, randomized, double-blind, placebo-controlled phase 2b trials for the treatment of NASH or NAFLD and incorporated MRI substudies in which subjects agreed to undergo MRI for liver PDFF estimation prior to the initiation of treatment (or placebo).

Study Population

Inclusion and exclusion criteria for the FLINT and CyNCh trials have been published (14, 15). All subjects included in the FLINT trial had a diagnosis of NASH based on established histopathological criteria for adults, and all subjects included in the CyNCh trial had a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

diagnosis of NAFLD based on established histopathological criteria for children (14, 15). For this secondary analysis, we included all 208 MRI sub-study subjects (101 from the FLINT trial and 107 from the CyNCh trial) who had baseline MRI. The population was 46% female (95/208), with a mean age of 32+/−21 years (FLINT subjects: mean age 52+/−11 years, range 22-75 years; CyNCh subjects: mean age 14+/−3 years, range 9-18 years). The mean interval between baseline MRI and liver biopsy was 56+/−24 days (range 0-115 days).

MRI Examinations and Analysis

Each subject underwent baseline MRI at one of 13 participating clinical trial sites. MRI was performed at seven sites at 1.5T, five sites at 3T, and one site at both 1.5T and 3T. The NASH CRN Radiology Coordinating Center (RCC) provided a standardized MRI acquisition protocol to each site, and all MRI examinations were performed with this protocol. Using a non-contrast, breath-held, axial 2D gradient recalled echo sequence, the entire liver was imaged using a torso phased array coil. Key pulse sequence parameters are summarized in Table 1. Image parameters were chosen to minimize confounding factors such as T1 and noise biases, while permitting quantification of fat-water signal oscillation and measurement of R2*. Magnitude images obtained at each site were transmitted in DICOM format to the RCC for central analysis including PDFF and R2* computation. The computation included corrections for the multi-peak resonance spectrum of fat using Hamilton's model (18).

PDFF and R2* values were measured by placing one circular region of interest (ROI) with at least a 1-cm radius in each of the nine Couinaud segments on the 5^{th}-echo image, avoiding major vessels, bile ducts, liver edges and artifacts. Those ROIs were then propagated onto the corresponding images for the other echo times. The mean signal intensities from those ROIs were entered into a custom MATLAB™ (The MathWorks, Natick, MA, USA) multi-peak, non-linear, least-squares fitting algorithm, and the mean PDFF and R2* values averaged across all nine ROIs were recorded; only mean values across the nine ROIs are considered in this report (5, 19, 20). Analysts were blinded to all non-MRI data at the time of image analysis.

Additional Data Collection and Analysis

Other clinical and laboratory parameters were collected locally at each site and stored by the Data Coordinating Center (DCC). Clinical data included age, gender, weight, height, body mass index (BMI), and blood pressure. Laboratory data included platelet count, glycosylated hemoglobin, liver function tests, serum triglycerides, total cholesterol, and serum insulin.

In both parent trials, biopsies were reviewed by members of the NASH CRN Pathology Committee according to the pre-established NASH CRN scoring system for steatosis, lobular inflammation, and hepatocyte ballooning, and the NASH Activity Score (NAS) was calculated as the sum of these three scores (21). Pathologists also scored the presence and severity of fibrosis (Masson's trichrome stain), hepatocellular iron deposition grade (Prussian blue stain, scale 0-4), non-hepatocellular iron deposition grade (Prussian blue stain, scale 0-2), as well as steatosis location, portal inflammation, lipid droplet size, and glycogenosis (22). Pathologists were blinded to all other data at the time of analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Statistical Analysis

All analyses were performed in R version 3.3.1 (2016. R Foundation for Statistical Computing, Vienna, Austria) by a staff statistician under the supervision of a faculty statistician (both with over 20 years' experience). Since there is no simple method for converting R2* values between field strengths, all analyses were performed separately for subjects imaged at 1.5T and those imaged at 3T. Adult and child data were pooled in each analysis, and age as well as parent trial were included as covariates in multivariable analyses.

Baseline demographic, laboratory, and histopathological features of the study population were summarized descriptively. Comparisons between subgroups were performed using Wilcoxon-Mann-Whitney tests or chi-square tests, as appropriate. Prior to statistical analysis and due to the small number of counts in some ordinal covariate categories, some covariate categories were combined, including: steatosis grade (grades 0 and 1); fibrosis stage (stages 3 and 4); lipid droplet size (mixed and predominantly small); hepatocellular iron grade (grades 1 and 2); and non-hepatocellular iron grade (grades 1 and 2).

The relationship between R2* and PDFF was examined using Pearson's correlation, since a linear relationship had been found in previous works (4, 13). This analysis was repeated for the subset of subjects with and without hepatocellular or non-hepatocellular iron at each field strength, and Spearman's correlation was used if the relationship appeared non-linear and/or numerous outliers were present (9). A permutation test, which is less sensitive to outliers than the standard Fisher's z test for independent correlations, was used to compare R2*-PDFF correlations between field strengths (23). Strength of correlation was interpreted according to the following scale: 0-0.19 = very weak; 0.2-0.39 = weak; 0.40-0.59 = moderate; 0.60-0.79 = strong; and 0.80-1.0 = very strong (24).

Then, the effect of additional factors on R2* was explored. Univariate relationships between all candidate demographic, laboratory, and histopathological covariates and R2* were examined. Features with more than two ordinal categories were assessed using Spearman's rho, as a linear relationship was no longer hypothesized. Binary features were assessed using the Wilcoxon-Mann-Whitney test. Since these univariable analyses were considered preliminary prior to the subsequent multivariable analysis, correction for multiple comparisons was not performed.

We then used a machine learning method to identify the most influential covariates for R2* at each field strength. The non-parametric method Gradient Boosting Machines (GBM) method constructs a model from an ensemble of regression trees while minimizing squared error, in our case relative to observed R2* values (25). The relative influence of any covariate in a model was determined by computing a relative influence (RI), expressed as a percentage ranging from 0% (the covariate has no influence on the outcome variable) to 100% (the covariate is the sole determinant of the outcome variable). The GBM method evaluates multiple levels of interactions between potential covariates. All demographic, laboratory and histopathological measures of potential interest were included in the candidate covariate pool for the multivariable GBM analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Results

Of the 208 subjects, 119 were imaged at 1.5T and 89 at 3T. Field strength was relatively balanced in both studies: for FLINT (adult subjects), 44 were imaged at 1.5T and 57 were imaged at 3T; for CyNCH (pediatric subjects), 45 were imaged at 1.5T and 62 were imaged at 3T. Demographic, laboratory, and histopathological features of the study population are summarized by parent trial (adult vs. pediatric) and field strength in Tables 2 and 3.

The mean liver PDFF value for all subjects was 19.8 ± 9.7%. Estimates of PDFF were similar between field strengths, but R2* values were generally higher at 3T. The mean liver R2* value was 37.2 ± 6.8 s^{-1} at 1.5T and 59.8 ± 12.9 s^{-1} at 3T, and R2* values were generally in the normal range (<60 s^{-1} at 1.5T, <126 at 3T) (26). The distributions of baseline PDFF and R2* values are summarized in Figure 1.

A minority of subjects had either type of liver iron deposition. For hepatocellular iron, 14.4% (30/208) had grade 1 deposition and 2.9% (6/208) had grade 2 deposition, and no subject had grade 3 or 4 deposition. For non-hepatocellular iron, 12.0% (25/208) of subjects had grade 1 deposition and 4.3% (9/208) had grade 2 deposition. The total number of subjects with at least grade 1 hepatocellular or non-hepatocellular iron deposition at baseline biopsy was small (n=38 total), and most of those subjects had both types (n=32). For a small number of subjects (1.9%, 4/208) iron staining had not been performed, so iron grade data were not available. These subjects were included in analyses that did not involve grade of liver iron deposition but were excluded from analyses requiring this information.

Analysis of PDFF

Scatterplots of baseline R2* and PDFF values are shown in Figure 2. The relationship between R2* and PDFF was generally linear. Pearson's correlations between R2* and PDFF were moderate to strong for subjects overall at both 1.5T ($r=0.652$, $p<0.0001$) and at 3T ($r=0.586$, $p<0.0001$). Notably, the greatest outlying values at both field strengths were in the subset of subjects with at least grade 1 liver iron deposition. For subjects with grade 0 liver iron deposition, Pearson's correlations were strong to very strong at both 1.5T ($r=0.697$, $p<0.0001$) and 3T ($r=0.835$, $p<0.0001$).

In subjects with at least grade 1 hepatocellular or non-hepatocellular iron deposition, the correlation between R2* and PDFF was moderate and significant both at 1.5T ($r=0.546$, $p=0.016$) and 3T (Spearman's $r=0.40$, $p=0.04$).

The differences between the correlations observed at 3T vs. 1.5T were not significant (permutation $p=0.45$ for all subjects, permutation $p=0.28$ for subjects without liver iron deposition, and $p=0.22$ for subjects with at least grade 1 liver iron deposition).

Univariate Analysis

The R2* values associated with various grades of hepatocellular and non-hepatocellular liver iron deposition are shown in Figure 3. R2* values for patients with grades 0 and 1 iron deposition of either type were similar at both field strengths. For grade 2 non-hepatocellular iron deposition, the median R2* value at 3T was similar to that for grades 0 and 1 (n=7). For

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

other combinations of field strength and grade 2 iron deposition, median R2* value appeared visually elevated compared to grades 0 and 1, but with very small sample sizes (n=2-3).

Results of the univariate correlation analysis of the relationship between R2* and each covariate are shown in Table 4. Using Spearman's correlation, R2* was most strongly correlated with PDFF at both field strengths (rho = 0.73-0.81, $p<0.001$), followed by histopathologic steatosis grade (rho = 0.53-0.57, $p<0.001$). At 1.5T, log-transformed ferritin (rho = 0.32, $p<0.005$) and alanine aminotransferase (rho = 0.25, $p<0.02$) were the next most strongly correlated covariates. At 3T, hepatocyte ballooning score (rho = −0.26, $p<0.01$) and log-transformed ferritin (rho = 0.23, $p<0.02$) were the next most strongly correlated covariates. At 3T, lipid droplet size was also a highly significant covariate (Wilcoxon-Mann-Whitney $p<0.001$), though a correlation coefficient could not be calculated due to the binary expression of lipid droplet size.

At 1.5T, neither hepatocellular iron grade ($p=0.13$) nor non-hepatocellular iron grade ($p=0.20$) were significantly correlated with R2*. At 3T, the correlation between R2* and both hepatocellular iron grade ($p=0.10$) and non-hepatocellular iron grade ($p=0.05$) approached but did not reach statistical significance.

Multivariable Analysis

The RIs of the covariates for the full models as well as the final optimized models of R2* at 1.5T and 3T are summarized in Supplemental Tables 1 and 2. At both field strengths, the optimal GBM models were additive (without interactions). Models with fitted interactions had higher errors.

At 1.5T, the optimal GBM model included five covariates for R2*. PDFF was the most influential covariate of R2*, with a RI of 61.3%. Serum insulin concentration, glycogenosis, Gamma-glutamyl transferase, and log-transformed ferritin concentration were weaker covariates of R2*, with RIs up to 11.5%. Neither non-hepatocellular iron nor hepatocellular iron were covariates in the optimal model at 1.5T.

At 3T, the optimal GBM model included six covariates of R2*. PDFF was the most influential covariate of R2* at 3T, with a RI of 47.5%. Log-transformed ferritin concentration, diastolic blood pressure, age, BMI, and non-hepatocellular iron were weaker covariates of R2*, with relative influences up to 16.7%. Non-hepatocellular iron was the weakest accepted covariate of R2* at 3T (RI=6.7%). Hepatocellular iron was not a covariate of R2* in the optimal model at 3T.

Discussion

In this work, we independently confirmed, at both 1.5T and 3T, the existence of a moderate to strong linear relationship between R2* and PDFF values in patients with NAFLD. Importantly, we found that in populations of adults and children with NAFLD, hepatic fat content (expressed as PDFF) was the most influential of all available covariates of hepatic R2* value, regardless of field strength. This relationship between R2* and PDFF is present regardless of whether underlying liver iron is present. The key conclusion from our study is

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that, since liver iron deposition is relatively rare and of low grade in the NAFLD/NASH population, R2* is not a useful technique for quantifying liver iron in this group of patients.

In patients with severe liver iron deposition due to transfusional hemosiderosis and other iron-loading conditions, R2* has been shown to be linearly correlated with liver iron concentration (LIC) (9, 10). MRI is used in many practices as the primary modality to monitor liver iron deposition and assess response to therapy (10, 27). However, we found that in a diverse NAFLD cohort, R2* is primarily associated with liver fat accumulation, presumably reflecting susceptibility effects from intra-hepatocellular triglyceride droplets rather than iron deposition. Although individual iron particles have much greater susceptibility effects than individual similarly-sized fat droplets, the amount of iron encountered in the NAFLD population has a narrow range from none to mild, while the range of fat content is broad. Over these ranges, liver fat has a greater impact on MRI-measured R2* values than liver iron.

In the liver, iron is often deposited within hepatocytes, non-hepatocyte depots (reticulendothelium system cells or fibrotic bands), or both. Elevated body iron stores have been linked to a worse clinical course in NAFLD, particularly for non-hepatocellular iron deposition (12, 28). We found visual trends toward higher R2* values with grade 2 hepatocellular iron at both field strengths and with grade 2 non-hepatocellular iron at 1.5T. However, grade 2 iron deposition of either type occurred rarely, and this finding may have been spurious as it was not confirmed in our formal analyses. Overall, these findings suggest that R2* may be a poor predictor of liver iron deposition in patients with NAFLD or NASH, in whom the susceptibility effects of fat predominate.

One of the reasons this study is clinically relevant is that in patients with NAFLD, liver iron deposition may portend worse clinical outcomes (11, 12). Since MRI-based measures such as R2* estimation have all but replaced liver biopsy for assessment of liver iron concentration in other populations, investigators logically may attempt to use measurements of R2* to quantify liver iron in patients with NAFLD (26, 29). However, our data show that hepatic steatosis, as quantified by PDFF, substantially confounders R2* measurements in this population. Additionally, the fact that liver iron grade has a relatively small effect on liver R2* suggests that R2* may not be a suitable method for assessing liver iron in this population, even if corrections for PDFF were implemented. Liver fibrosis has also been shown to confound R2*, though to a much lesser extent in our study (30).

Interestingly, although non-hepatocellular iron grade was weakly associated with R2* at 3T, hepatocellular iron was not associated with R2* at either field strength. The strength of susceptibility, and consequently R2* change, due to iron can vary based on the orientation of iron particles in the magnetic field, oxidation state, spin state, and other factors. Although the number of subjects with each type of iron deposition was similar, iron deposition grades are assessed visually by light microscopy, and iron particles may differ in number, size, orientation, and other factors. Such microenvironmental differences could account for the greater effect of non-hepatocellular iron deposition on R2* compared with hepatocellular iron. In addition, our study was unsuited for accurately assessing differences in effect on R2* by the two types of iron deposition, since the two types tended to occur together.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

This study has several limitations. Only a minority of subjects in the parent trials participated in the MRI substudies. Iron was graded semi-quantitatively by histochemical stain, not quantitatively. Additionally, the MRI pulse sequence utilized in this study was optimized for PDFF measurement but not necessarily R2* measurement. A sequence optimized for measuring small elevations in R2*, for example by incorporating longer echo times, might have been more precise for R2* measurements in the low range as encountered in this patient population. Additionally, biopsy locations were not recorded, so we were unable to colocalize MRI measurement locations with biopsy locations. Importantly, only a few patients had abnormal liver iron deposition, and none had grade 3 or higher deposition, however this reflects the amount of liver iron deposition typically encountered in the NAFLD population. Finally, although the subject populations studied at 1.5T and 3T were similar, the GBM modeling method produced different optimized models at 1.5T and 3T. This is not unreasonable, given the linear but different nature of the R2*-PDFF relationship and subtle distributional differences in the covariates between the two cohorts. However, with more data and the goal of refining and completely describing a multivariable predictive model for R2*, the GBM analysis could have been pursued further. As the goal of this work was to highlight the R2*-PDFF relationship we stopped at demonstrating that PDFF emerges as the most influential predictor in the presence of other predictors for both cohorts.

In conclusion, in both adults and children with NAFLD, hepatic triglyceride content as measured non-invasively by MRI-PDFF is the most influential covariate for liver R2* values at both 1.5T and 3T. In patients with NAFLD, liver R2* measurements may not provide reliable estimates of liver iron content.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Grant Support:

The Nonalcoholic Steatohepatitis Clinical Research Network (NASH CRN) is supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (grants U01DK061718, U01DK061728, U01DK061731, U01DK061732, U01DK061734, U01DK061737, U01DK061738, U01DK061730, U01DK061713). Additional support is received from the National Center for Advancing Translational Sciences (NCATS) (grants UL1TR000439, UL1TR000077, UL1TR000436, UL1TR000150, UL1TR000424, UL1TR000006, UL1TR000448, UL1TR000040, UL1TR000100, UL1TR000004, UL1TR000423, UL1TR000058, UL1TR000454).

Abbreviations

BMI –	Body mass index
CyNCh –	Cysteamine bitartrate delayed-release for the treatment of NAFLD in children
DCC –	Data coordinating center
FLINT –	Farnesoid X receptor ligand obeticholic acid in NASH treatment

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

GBM –	Gradient boosting machines
NAFLD –	Non-alcoholic fatty liver disease
NASH –	Non-alcoholic steatohepatitis
NASH CRN –	Non-alcoholic steatohepatitis clinical research network
PDFF –	Proton density fat fraction
RCC –	Radiology coordinating center

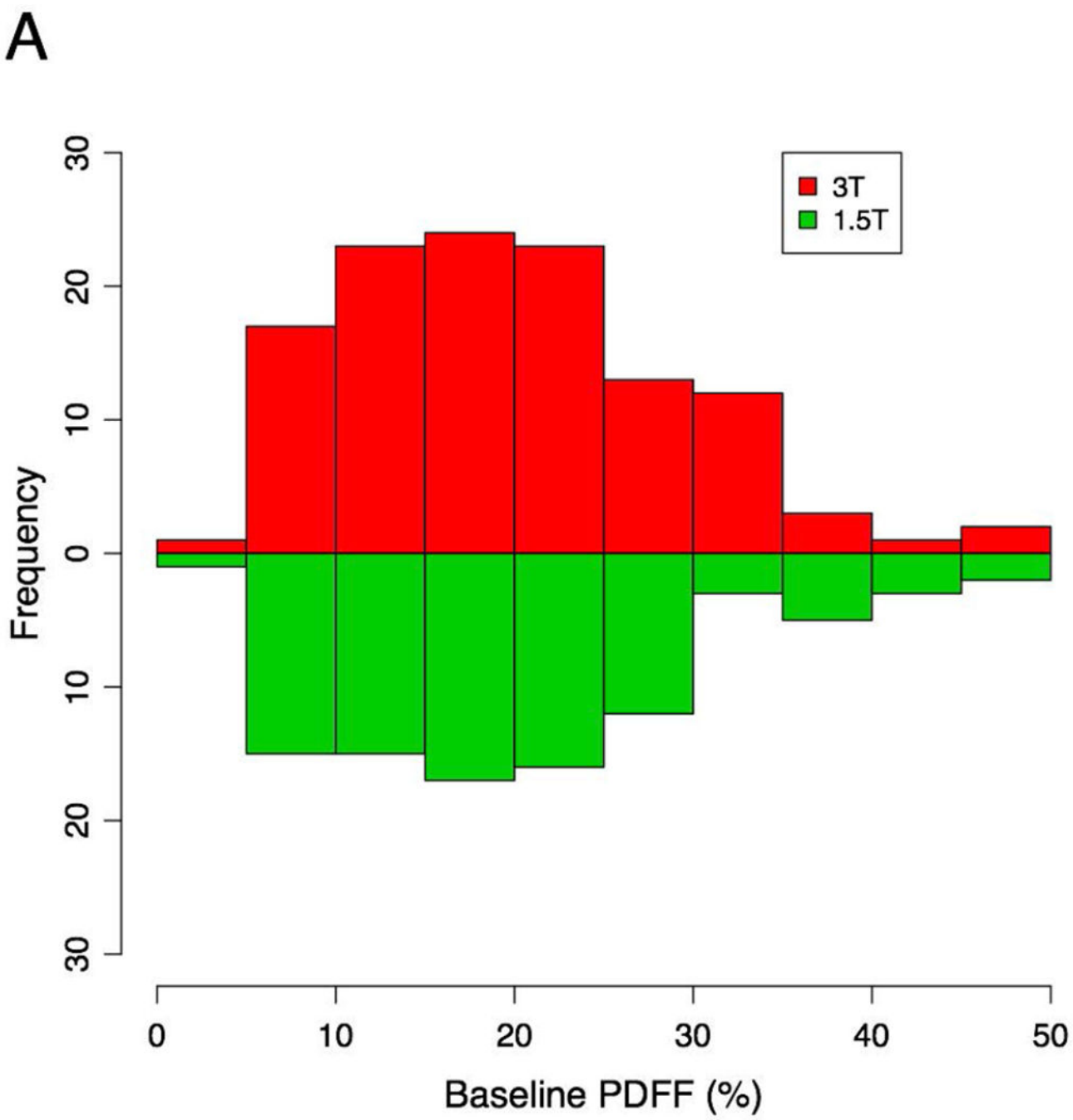
References

1. Harrison SA, Rinella ME, Abdelmalek MF, et al. NGM282 for treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet2018 3 24;391(10126):1174–85. [PubMed: 29519502]
2. Middleton MS, Heba ER, Hooker CA, et al. Agreement Between Magnetic Resonance Imaging Proton Density Fat Fraction Measurements and Pathologist-Assigned Steatosis Grades of Liver Biopsies From Adults With Nonalcoholic Steatohepatitis. Gastroenterology2017 9;153(3):753–61. [PubMed: 28624576]
3. Yokoo T, Serai SD, Pirasteh A, et al. Linearity, Bias, and Precision of Hepatic Proton Density Fat Fraction Measurements by Using MR Imaging: A Meta-Analysis. Radiology2018 2;286(2):486–98. [PubMed: 28892458]
4. Bashir MR, Zhong X, Nickel MD, et al. Quantification of hepatic steatosis with a multistep adaptive fitting MRI approach: prospective validation against MR spectroscopy. AJR Am J Roentgenol2015 2;204(2):297–306. [PubMed: 25615751]
5. Yokoo T, Shiehmorteza M, Hamilton G, et al. Estimation of hepatic proton-density fat fraction by using MR imaging at 3.0 T. Radiology2011 3;258(3):749–59. [PubMed: 21212366]
6. 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. Radiology2011 3;258(3):767–75. [PubMed: 21248233]
7. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging2012 11;36(5):1011–4. [PubMed: 22777847]
8. Zhong X, Nickel MD, Kannengiesser SA, Dale BM, Kiefer B, Bashir MR. Liver fat quantification using a multi-step adaptive fitting approach with multi-echo GRE imaging. Magn Reson Med2014 11;72(5):1353–65. [PubMed: 24323332]
9. Hankins JS, McCarville MB, Loeffler RB, et al. R2* magnetic resonance imaging of the liver in patients with iron overload. Blood2009 5 14;113(20):4853–5. [PubMed: 19264677]
10. Wood JC, Enriquez C, Ghugre N, et al. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. Blood2005 8 15;106(4):1460–5. [PubMed: 15860670]
11. Bahl M, Qayyum A, Westphalen AC, et al. Liver steatosis: investigation of opposed-phase T1-weighted liver MR signal intensity loss and visceral fat measurement as biomarkers. Radiology2008 10;249(1):160–6. [PubMed: 18796674]
12. Nelson JE, Wilson L, Brunt EM, et al. Relationship between the pattern of hepatic iron deposition and histological severity in nonalcoholic fatty liver disease. Hepatology2011 2;53(2):448–57. [PubMed: 21274866]
13. Mamidipalli A, Hamilton G, Manning P, et al. Cross-sectional correlation between hepatic R2* and proton density fat fraction (PDFF) in children with hepatic steatosis. J Magn Reson Imaging2017 5 25.
14. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. Lancet2015 3 14;385(9972):956–65. [PubMed: 25468160]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

15. Schwimmer JB, Lavine JE, Wilson LA, et al. In Children With Nonalcoholic Fatty Liver Disease, Cysteamine Bitartrate Delayed Release Improves Liver Enzymes but Does Not Reduce Disease Activity Scores. Gastroenterology2016 12;151(6):1141–54 e9. [PubMed: 27569726]
16. Bell LN, Wang J, Muralidharan S, et al. Relationship between adipose tissue insulin resistance and liver histology in nonalcoholic steatohepatitis: a pioglitazone versus vitamin E versus placebo for the treatment of nondiabetic patients with nonalcoholic steatohepatitis trial follow-up study. Hepatology2012 10;56(4):1311–8. [PubMed: 22532269]
17. Lavine JE, Schwimmer JB, Van Natta ML, et al. Effect of vitamin E or metformin for treatment of nonalcoholic fatty liver disease in children and adolescents: the TONIC randomized controlled trial. JAMA2011 4 27;305(16):1659–68. [PubMed: 21521847]
18. Hamilton G, Yokoo T, Bydder M, et al. In vivo characterization of the liver fat (1)H MR spectrum. NMR Biomed2011 8;24(7):784–90. [PubMed: 21834002]
19. Bydder M, Yokoo T, Hamilton G, et al. Relaxation effects in the quantification of fat using gradient echo imaging. Magn Reson Imaging2008 4;26(3):347–59. [PubMed: 18093781]
20. 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 Imaging2008 7;28(1):246–51. [PubMed: 18581393]
21. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA, Network NCR. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. Hepatology2011 3;53(3):810–20. [PubMed: 21319198]
22. Turlin B, Deugnier Y. Evaluation and interpretation of iron in the liver. Semin Diagn Pathol1998 11;15(4):237–45. [PubMed: 9845425]
23. Good PI. Permutation tests : a practical guide to resampling methods for testing hypotheses. 2nd ed New York: Springer; 2000.
24. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutor Quant Methods Psychol2012;8(1):23–34. [PubMed: 22833776]
25. Friedman JH. Greedy Function Approximation: A Gradient Boosting Machine. Ann Stat2001;29(5):1189–232.
26. Labranche R, Gilbert G, Cerny M, et al. Liver Iron Quantification with MR Imaging: A Primer for Radiologists. Radiographics2018 Mar-Apr;38(2):392–412. [PubMed: 29528818]
27. Yokoo T, Browning JD. Fat and iron quantification in the liver: past, present, and future. Top Magn Reson Imaging2014 4;23(2):73–94. [PubMed: 24690618]
28. Aigner E, Weiss G, Datz C. Dysregulation of iron and copper homeostasis in nonalcoholic fatty liver. World J Hepatol2015 2 27;7(2):177–88. [PubMed: 25729473]
29. Roth CG, Marzio DH, Guglielmo FF. Contributions of Magnetic Resonance Imaging to Gastroenterological Practice: MRIs for GIs. Dig Dis Sci2018 5;63(5):1102–22. [PubMed: 29549474]
30. Li J, Lin H, Liu T, et al. Quantitative susceptibility mapping (QSM) minimizes interference from cellular pathology in R2* estimation of liver iron concentration. J Magn Reson Imaging2018 3 22.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

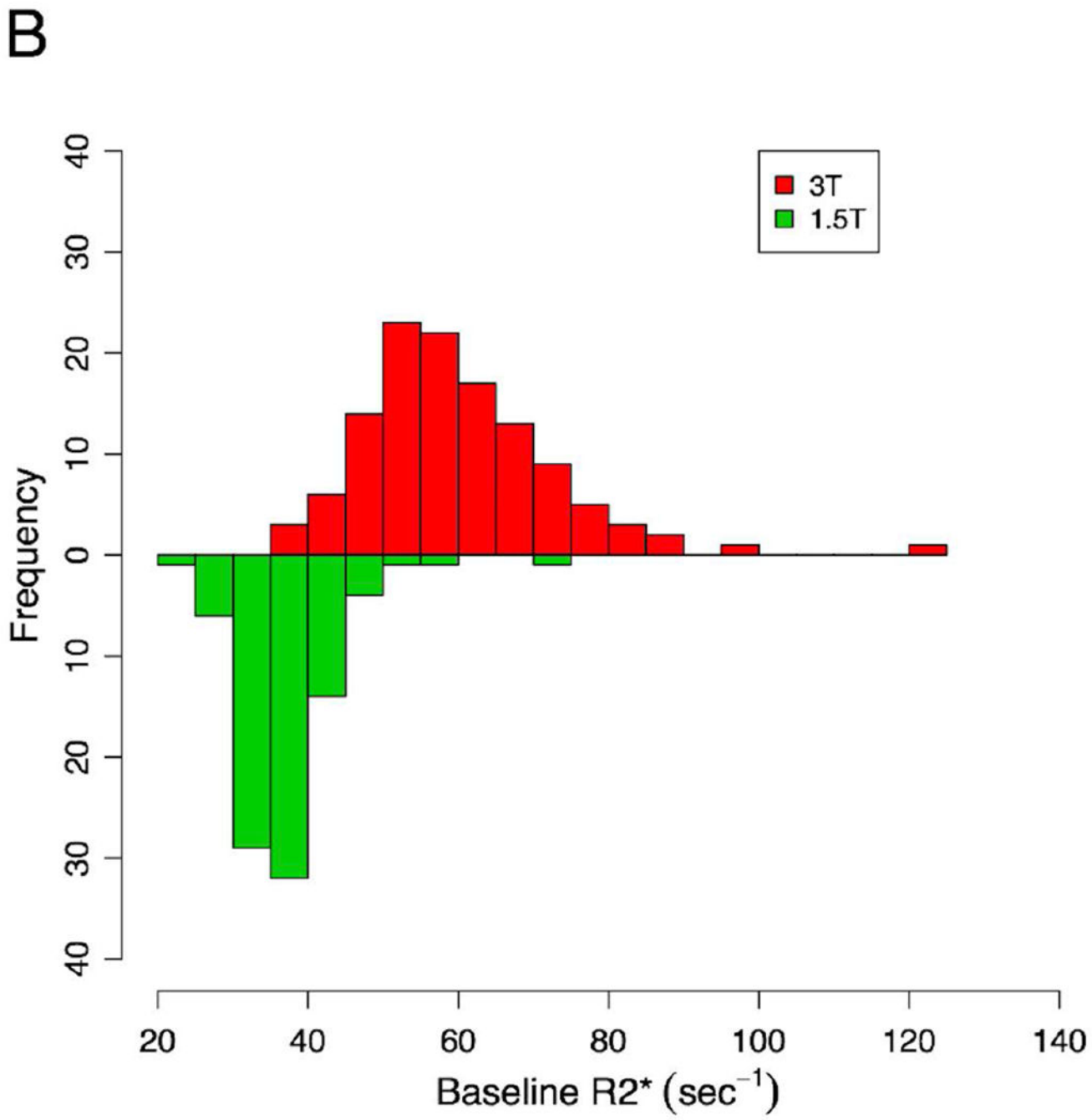


Figure 1 –.
Histograms summarizing PDFF and R2* values for subjects included in this study.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

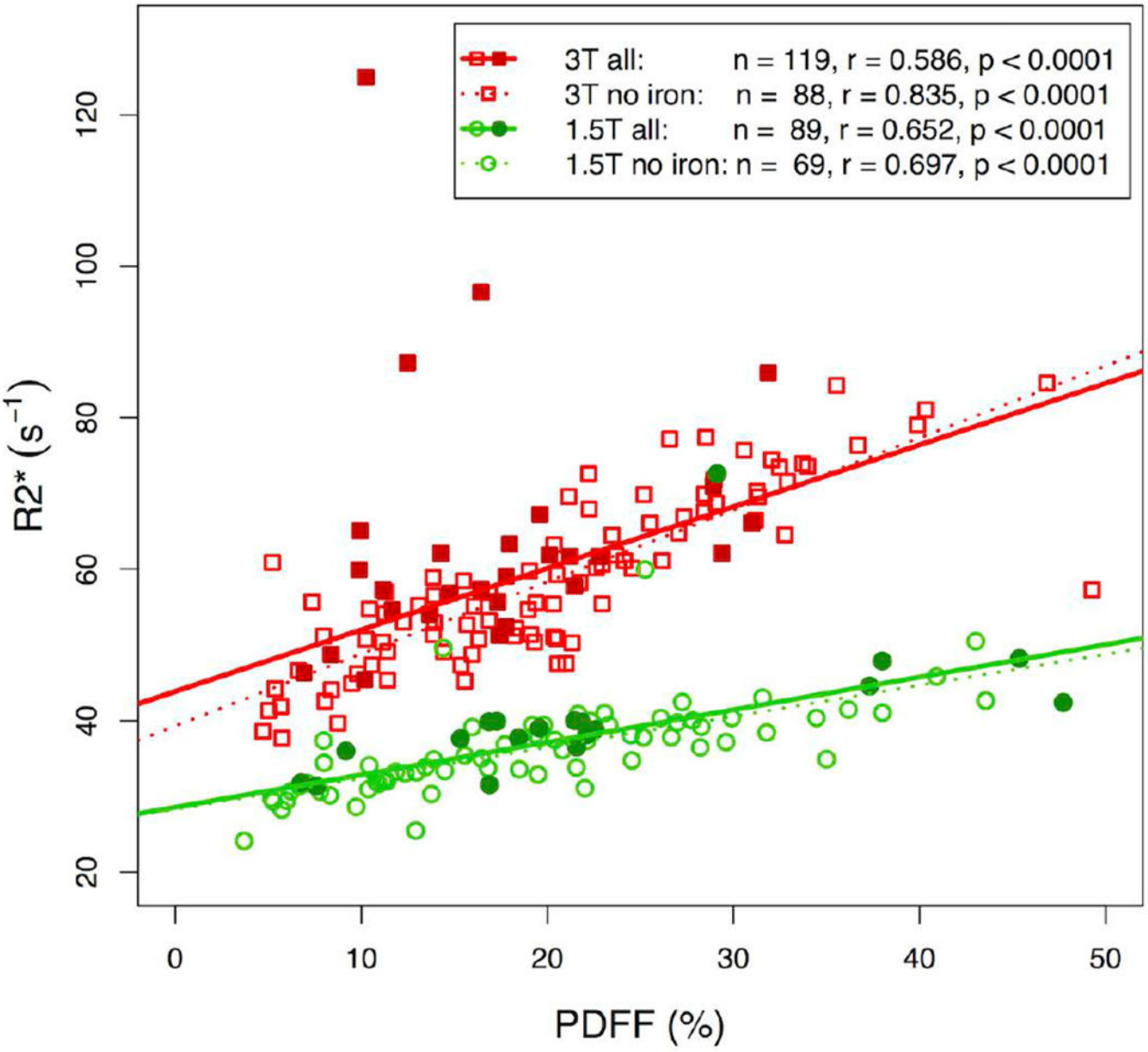
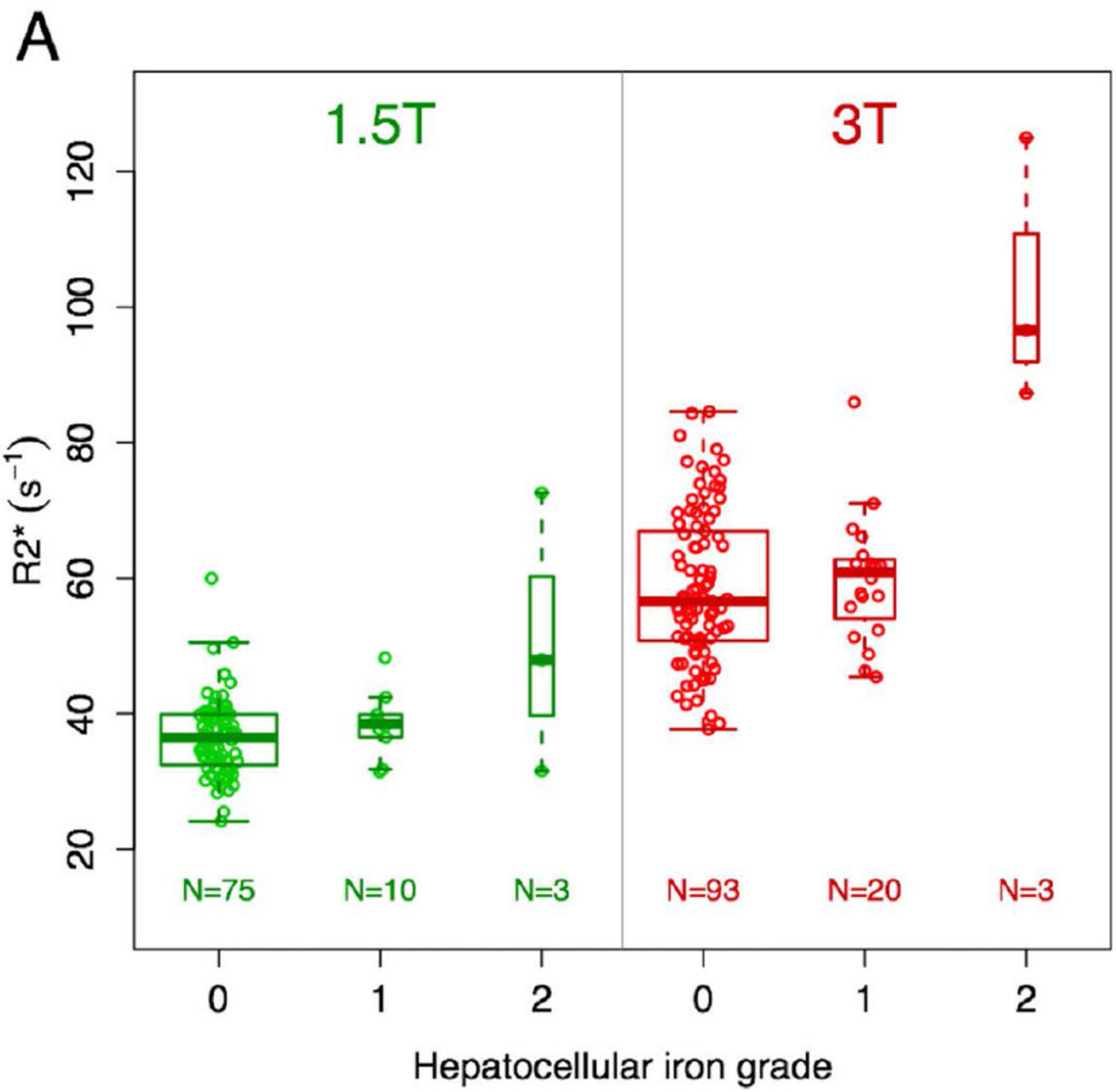


Figure 2 –.

Pearson's correlation between R2* and PDFF at baseline, shown by field strength for subjects overall ("all"), those with grade 0 hepatocellular and non-hepatocellular cell iron deposition ("no iron").

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

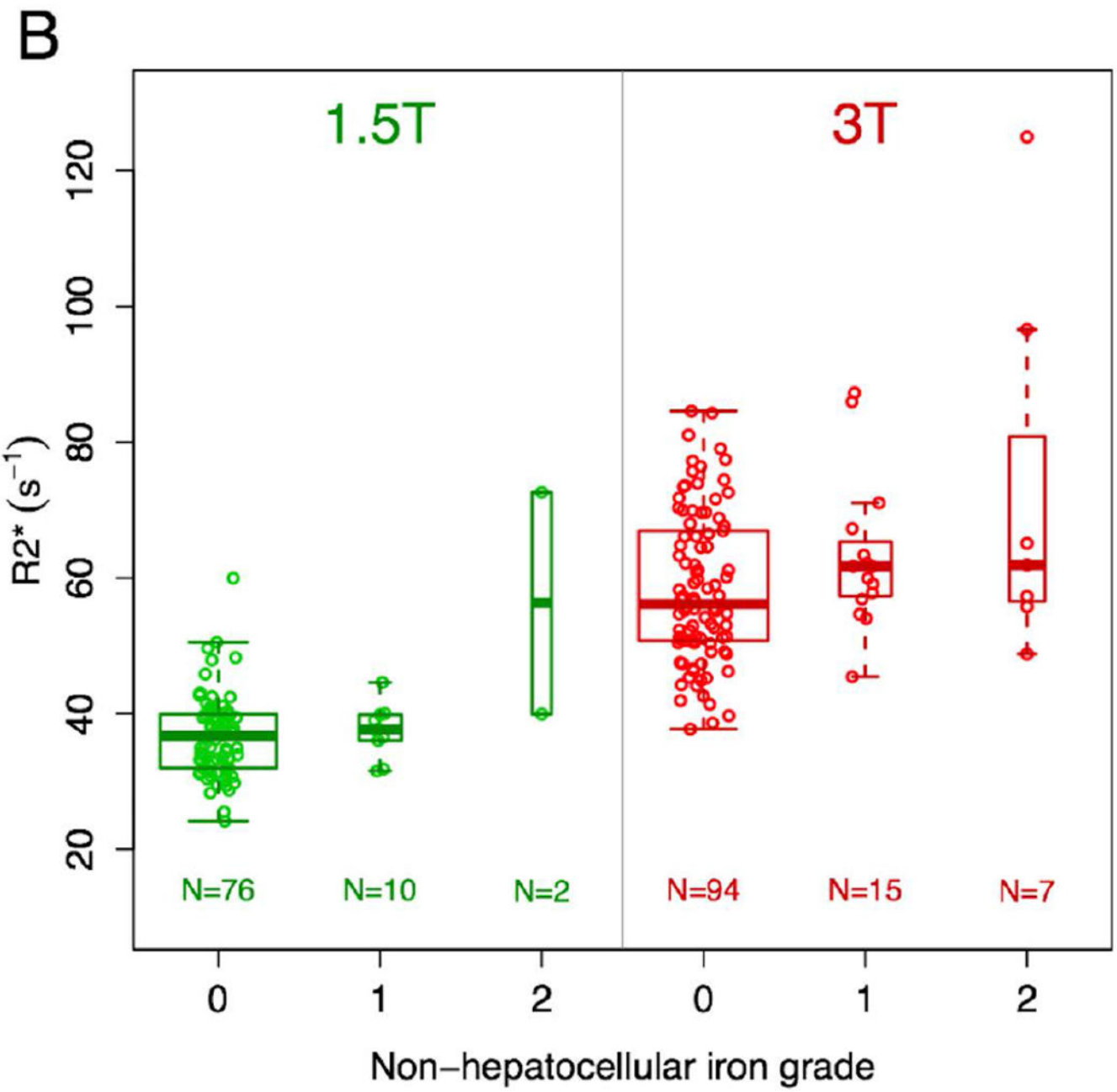


Figure 3 –.

Box and whisker plots of R2* values based on type and grade of iron deposition and magnetic field strength. There are visual trends toward higher R2* values for grade 2 iron deposition compared with grades 0 and 1 for hepatocellular iron (a) at both field strengths and non-hepatocellular iron (b) at 1.5T, though with very small samples sizes for grade 2 (n=2-3). For the largest group with grade 2 iron deposition (non-hepatocellular iron imaged at 3T, n=7), the median R2* value is similar to the median R2* values for grades 0 and 1 iron deposition.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

MRI pulse sequence parameters for the gradient recalled echo sequences used included in this study

Parameter	1.5T	3T
Repetition time (msec)	≥ 120	
Flip angle (°)	10	
First echo time (msec)	2.3	1.15
Echo spacing (msec)	2.3	1.15
Number of echoes	6	
Bandwidth (Hz/Px)	≥ 500	≥ 1,000
Slice thickness (mm)	8-10	
Slice gap (mm)	0	
Acquisition matrix	192 × 192	128 × 128
Frequency encoding steps	192	128

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2.

MRI-derived, demographic, metabolic, and laboratory factors (and summary values) assessed as covariates of liver R2* value

Covariate	1.5T (n=89)	3T (n=119)	p-values (1.5T vs. 3T)	FLINT (Adults, n=101)	CyNCh (Children, n=107)	p-values (FLINT vs. CyNCh)
MRI-Derived Measures						
Mean baseline liver PDFF (%)	20.1 ± 10.4	19.6 ± 9.3	0.84	18.5 ± 9.6	21.0 ± 9.7	<0.05
Mean baseline liver R2* (s^{-1})	37.2 ± 6.8	59.8 ± 12.9	<0.001	49.2 ± 16.0	51.0 ± 15.0	0.30
Demographic Features						
Age (years)	33.5 ± 21.9	31.7 ± 20.3	0.51	52.4 ± 11.0	13.6 ± 2.6	<0.001
Gender			0.81			<0.001
Females	47% (42/89)	45% (53/119)		62% (63/101)	30% (32/107)	
Males	53% (47/89)	55% (66/119)		38% (38/101)	70% (75/107)	
Metabolic Features						
Weight (kg)	91.5 ± 23.8	88.6 ± 21.2	0.29	95.6 ± 17.1	32.0 ± 6.2	<0.001
Body mass index (kg/m^2)	33.4 ± 6.3	32.5 ± 5.5	0.22	33.8 ± 5.1	32.0 ± 6.3	<0.05
Body mass index z-score	2.0 ± 0.5	2.0 ± 0.5	0.47	1.8 ± 0.4	2.2 ± 0.5	<0.001
Weight-to-height ratio	0.7 ± 0.1	0.7 ± 0.1	0.86	0.6 ± 0.1	0.7 ± 0.1	0.58
Systolic blood pressure (mmHg)	126.7 ± 15.1	123.6 ± 14.8	0.13	131.2 ± 15.6	118.9 ± 11.5	<0.001
Diastolic blood pressure (mmHg)	73.5 ± 10.1	71.9 ± 12.0	0.21	78.1 ± 10.8	67.3 ± 8.8	<0.001
Serum Laboratory Values						
Platelet count (×1,000 cells/mm^3)	263.3 ± 67.1	263.5 ± 67.3	0.13	131.2 ± 15.6	118.9 ± 11.5	<0.001
Glycosylated hemoglobin (%)	5.9 ± 1.0	5.9 ± 0.8	0.21	78.1 ± 10.8	67.3 ± 8.8	<0.001
Aspartate aminotransferase (U/L)	59.3 ± 40.8	59.2 ± 40.9	0.91	55.1 ± 31.8	63.1 ± 47.5	0.27
Alanine aminotransferase (U/L)	97.8 ± 71.1	89.2 ± 68.4	0.10	75.9 ± 43.1	108.9 ± 84.6	<0.002
Gamma-glutamyl transferase (U/L)	68.6 ± 108.5	59.8 ± 70.9	0.08	80.2 ± 120.6	47.9 ± 34.4	<0.02
Total bilirubin (mg/dL)	0.6 ± 0.3	0.6 ± 0.3	0.92	0.7 ± 0.3	0.5 ± 0.3	<0.001
Serum triglycerides (mg/dL)	176.6 ± 92.8	161.4 ± 80.8	0.17	180.9 ± 92.1	155.6 ± 78.8	<0.05
Total cholesterol (mg/dL)	180.9 ± 47.4	174.0 ± 41.4	0.28	193.0 ± 45.2	161.8 ± 37.2	<0.001
Fasting blood glucose (mg/dL)	98.9 ± 26.4	98.8 ± 25.8	0.57	112.0 ± 30.1	86.5 ± 12.1	<0.001
Insulin (mIU/L)	35.7 ± 46.4	31.4 ± 26.1	0.63	28.9 ± 37.2	37.4 ± 34.8	<0.001
Homeostatic Model Assessment of Insulin Resistance (HOMA-IR)	9.1 ± 13.9	7.9 ± 6.9	0.59	8.6 ± 12.7	8.2 ± 7.9	0.21

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Covariate	1.5T (n=89)	3T (n=119)	p-values (1.5T vs. 3T)	FLINT (Adults, n=101)	CyNCh (Children, n=107)	p-values (FLINT vs. CyNCh)
Log-transformed Ferritin (log ng/mL)	4.5 ± 1.0	4.7 ± 0.8	0.10	4.9 ± 1.0	4.3 ± 0.8	<0.001

For gender, data are presented as proportions (counts/total) and p-values are for the Chi-Square test. For all other measures, data are presented as mean +/− standard deviation, and p-values are for the Wilcoxon-Mann-Whitney test.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3.

Histopathological factors (and summary values) assessed as covariates of liver R2* value

Covariate	1.5T (n=89)	3T (n=119)	p-values (1.5T vs 3T)	FLINT (Adults, n=101)	CyNCh (Children, n=107)	p-values (FLINT vs CyNCH)
Steatosis grade *			0.53			<0.001
Grade 0	1% (1/89)	0% (0/119)		1% (1/101)	0% (0/107)	
Grade 1	27% (24/89)	23% (27/119)		33% (33/101)	17% (18/107)	
Grade 2	30% (27/89)	37% (44/119)		40% (40/101)	29% (31/107)	
Grade 3	42% (37/89)	40% (48/119)		27% (27/101)	54% (58/107)	
Steatosis location			0.20			<0.001
Zone 3 (acinar)	39% (35/89)	43% (51/119)		43% (43/101)	40% (43/107)	
Zone 1 (periportal)	13% (12/89)	5% (6/119)		0% (0/101)	17% (18/107)	
Azonal	18% (16/89)	19% (23/119)		32% (32/101)	7% (7/107)	
Panacinar	29% (26/89)	33% (39/119)		26% (26/101)	36% (39/107)	
Hepatocyte ballooning score			0.74			<0.001
Score 0	33% (29/89)	38% (45/119)		15% (15/101)	55% (59/107)	
Score 1	34% (30/89)	31% (37/89)		35% (35/101)	30% (32/107)	
Score 2	34% (30/89)	31% (37/89)		50% (51/101)	15% (16/107)	
Lobular inflammation score			0.070			0.35
Score 1	43% (38/89)	31% (37/119)		33% (33/101)	39% (42/107)	
Score 2	39% (35/89)	55% (66/119)		49% (49/101)	49% (52/107)	
Score 3	18% (16/89)	13% (16/119)		19% (19/101)	12% (13/107)	
Portal inflammation score			0.51			0.66
Score 0	4% (4/89)	8% (10/119)		7% (7/101)	7% (7/107)	
Score 1	73% (65/89)	68% (81/119)		67% (68/101)	73% (78/107)	
Score 2	22% (20/89)	24% (28/119)		26% (26/101)	21% (22/107)	
NAFLD activity score	4.9 ± 1.4	4.9 ± 1.4	0.75	5.1 ± 1.4	4.7 ± 1.4	<0.05
Fibrosis stage *			0.99			<0.001
Stage 0	20% (18/89)	22% (26/119)		16% (16/101)	26% (28/107)	
Stage 1	34% (30/89)	34% (40/119)		24% (24/101)	43% (46/107)	
Stage 2	19% (17/89)	19% (23/119)		26% (26/101)	13% (14/107)	
Stage 3	26% (23/89)	24% (28/119)		33% (33/101)	17% (18/107)	

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Covariate	1.5T (n=89)	3T (n=119)	p-values (1.5T vs 3T)	FLINT (Adults, n=101)	CyNCh (Children, n=107)	p-values (FLINT vs CyNCH)
Stage 4	1% (1/89)	2% (2/119)		2% (2/101)	1% (1/107)	
Lipid droplet size *			0.99			<0.01
Predominantly large droplet	70% (62/89)	69% (82/119)		78% (79/101)	61% (65/107)	
Mixed large and small droplet	28% (25/89)	26% (31/119)		17% (17/101)	36% (39/107)	
Predominantly small droplet	2% (2/89)	5% (6/119)		5% (5/101)	3% (3/107)	
Glycogenosis			0.54			0.92
Not present	55% (49/89)	49% (58/119)		52% (53/101)	50% (54/107)	
Focal, involving <50% of hepatocytes	24% (21/89)	24% (28/119)		24% (24/101)	23% (25/107)	
Diffuse, involving >= 50% of hepatocytes	21% (19/89)	28% (33/119)		24% (24/101)	26% (28/107)	
Hepatocellular iron grade *			0.45			<0.001
Grade 0	85% (75/88)	80% (93/116)		70% (68/97)	93% (100/107)	
Grade 1	12% (10/88)	17% (20/116)		24% (23/97)	7% (7/107)	
Grade 2	3% (3/88)	3% (3/116)		6% (6/97)	0% (0/107)	
Non-hepatocellular iron grade *			0.41			<0.001
Grade 0	86% (76/88)	81% (94/116)		69% (67/97)	96% (103/107)	
Grade 1	11% (10/88)	13% (15/116)		22% (21/97)	4% (4/107)	
Grade 2	2% (2/88)	6% (7/116)		9% (9/97)	0% (0/107)	

For NAFLD activity score, data are presented as mean +/− standard deviation, and p-values are for the Wilcoxon-Mann-Whitney test. For the remaining nominal and ordinal variables, data are presented as proportion (counts/total) and p-values are for the Chi-Square test.

* For the Chi-Square test, some categories were combined due to the small number of samples in each individual category, to increase statistical robustness. These included: steatosis grade (grades 0 and 1); fibrosis stage (stages 3 and 4); lipid droplet size (mixed and predominantly small); hepatocellular iron grade (grades 1 and 2); and non-hepatocellular iron grade (grades 1 and 2).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 4.

Results of the univariable analyses for covariates for R2* by field strength. For continuous and ordinal variables with more than two categories, Spearman's rho was calculated, while for binary variables, the Wilcoxon-Mann-Whitney test was utilized.

Covariate	1.5T		3T	
	Spearman's rho	p-value	Spearman's rho	p-value
Mean baseline liver PDFF	0.81	<0.001 *	0.73	<0.001
Age	−0.14	0.20	−0.09	0.35
Gender	NA	0.08	NA	0.41
Weight	−0.11	0.30	−0.12	0.21
BMI Z-score	0.12	0.27	0.04	0.64
Systolic blood pressure	0.002	0.98	0.02	0.85
Diastolic blood pressure	0.01	0.92	0.11	0.24
Platelet count	0.05	0.67	0.07	0.44
Glycosylated hemoglobin	−0.07	0.52	−0.06	0.49
Aspartate aminotransferase	0.09	0.39	0.11	0.25
Alanine aminotransferase	0.25	<0.02 *	0.17	0.07
Gamma-glutamyl transferase	0.01	0.91	0.10	0.29
Total bilirubin	0.04	0.71	−0.05	0.61
Serum triglycerides	0.07	0.49	0.15	0.10
Total cholesterol	0.06	0.57	0.14	0.15
Fasting blood glucose	0.10	0.36	−0.02	0.85
Insulin	−0.11	0.32	0.01	0.89
Log-transformed ferritin	0.32	<0.005 *	0.23	<0.02 *
Steatosis grade	0.57	<0.001 *	0.53	<0.001 *
Steatosis location (acinar)	NA	0.44	NA	0.33
Steatosis location (azonal/periportal)	NA	<0.05 *	NA	<0.05 *
Steatosis location (panacinar)	NA	<0.05 *	NA	<0.01 *
Hepatocyte ballooning score	−0.17	0.11	−0.26	<0.01 *
Lobular inflammation score	−0.18	0.09	0.11	0.23
Portal inflammation score	−0.11	0.29	0.08	0.40

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Covariate	1.5T		3T	
	Spearman's rho	p-value	Spearman's rho	p-value
NAFLD activity score	0.14	0.19	0.18	<0.05*
Fibrosis stage	−0.18	0.09	−0.05	0.57
Lipid droplet size	NA	0.57	NA	<0.001
Glycogenosis	0.02	0.88	−0.02	0.87
Hepatocellular iron grade	NA	0.13	NA	0.10
Non-hepatocellular iron grade	NA	0.20	NA	0.05

*Statistically significant, without correction for multiple comparisons