HHS Public Access

Author manuscript

AJR Am J Roentgenol. Author manuscript; available in PMC 2017 October 13.

Published in final edited form as:

AJR Am J Roentgenol. 2017 September ; 209(3): 592–603. doi:10.2214/AJR.17.17812.

Standardized Approach for ROI-Based Measurements of Proton Density Fat Fraction and R2* in the Liver

Camilo A. Campo[1], **Diego Hernando**[1,2], **Tilman Schubert**[1,3], **Candice A. Bookwalter**[1,4], **Andrew J. Van Pay**[1], and **Scott B. Reeder**[1,2,5,6,7]

[1]Department of Radiology, Wisconsin Institutes for Medical Research, University of Wisconsin-Madison, 1111 Highland Ave, Madison, WI 53705 [2]Department of Medical Physics, University of Wisconsin-Madison, Madison, WI [3]Clinic of Radiology and Nuclear Medicine, Basel University Hospital, Basel, Switzerland [4]Department of Radiology, Mayo Clinic, Rochester, MN [5]Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI [6]Department of Medicine, University of Wisconsin-Madison, Madison, WI [7]Department of Emergency Medicine, University of Wisconsin-Madison, Madison, WI

Abstract

OBJECTIVE—The purpose of this study was to evaluate the reproducibility (interreviewer agreement) and repeatability (intrareviewer agreement) of ROI sampling strategies to measure chemical shift–encoded (CSE) MRI-based liver proton density fat fraction (PDFF) and R2* (1 / T2*). A secondary purpose was to standardize ROI-based liver PDFF and R2* measurements by providing a compromise between measurement reproducibility and repeatability and time burden for image analysts.

MATERIALS AND METHODS—CSE data from two cohorts were retrospectively analyzed. Cohort A included 53 patients referred for abdominal MRI and healthy subjects recruited for a comparison study of CT and MRI. Cohort B included 37 patients with suspected liver iron overload. Three reviewers measured liver PDFF and R2* using previously reported ROI sampling strategies. Inter- and intrareviewer agreement of liver PDFF and R2* were evaluated using Bland-Altman analysis.

RESULTS—Averaging largest-fit ROIs over the nine Couinaud segments resulted in the narrowest limits of agreement (LOA) for liver PDFF and R2* measurements in both cohorts. For PDFF, interreviewer agreement had mean LOA of ± 0.8% for cohort A and ± 1.7% for cohort B. Intrareviewer agreement was ± 0.5% for cohort A and ± 0.9% for cohort B. For R2* interreviewer agreement had mean LOA of ± 3.0 s^{-1} for cohort A and ± 17.9 s^{-1} for cohort B. Intrareviewer agreement was ± 2.6 s^{-1} for cohort A and ± 14.6 s^{-1} for cohort B. This approach was the most time-burdensome, requiring a mean ± SD of 149.7 ± 8.6 s per dataset.

Address correspondence to S. B. Reeder (sreeder@wisc.edu).

Supplemental Data

Available online at www.ajronline.org.

CONCLUSION—For improved reproducibility and repeatability of liver PDFF and R2* measurements, clinicians and researchers should sample as much area of the liver as possible using multiple large ROIs.

Keywords

liver disease; MRI; proton density fat fraction; R2*; ROI

Emerging measurements of proton density fat fraction (PDFF) in the liver based on chemical shift–encoded (CSE) MRI have shown great promise as quantitative imaging biomarkers for noninvasive detection, quantitative grading, and treatment monitoring of hepatic steatosis [1–3]. PDFF-based liver fat quantification enables the noninvasive assessment of hepatic steatosis, the hallmark histologic feature of nonalcoholic fatty liver disease, which is estimated to affect 20–30% of the Western population [4–9]. Further, these techniques allow the simultaneous estimation of fat- and noise-corrected liver iron content using R2* (1 / T2*) quantification from the same acquired CSE MRI data [10–12]. R2*-based liver iron quantification enables the noninvasive assessment of diseases that result in liver iron overload, including hereditary hemochromatosis, and blood transfusion–dependent conditions such as thalassemia, sickle cell disease, and bone marrow failure [13–18].

The typical workflow for these techniques includes acquisition of CSE MRI data (typically a single 20-s breath-hold with whole-liver 3D acquisitions), reconstruction of PDFF and R2* maps from the CSE MRI data, and placement of ROIs on the PDFF and R2* maps to obtain quantitative estimates of the fat and iron content in the liver, respectively [19, 20]. Despite the growing clinical and research interest in these techniques, a standardized approach for ROI-based measurements has not been established. As a result, recent studies conducted at different sites have used a wide range of ROI sizes (from 0.785 cm^2 to the whole liver), locations (left and right lobes; anterior, posterior, and lateral segments; Couinaud segments; whole liver), and number (one to nine ROIs) for measurements of liver PDFF and R2* [6, 8, 16, 21–25]. This lack of standardization introduces an additional element of variability across sites and studies, impacting the widespread dissemination of these techniques as reproducible and repeatable quantitative imaging biomarkers. For this reason, the ROI placement approach for analysis of quantitative PDFF and R2* maps needs to be standardized.

Therefore, the primary purpose of this study was to evaluate the reproducibility (interreviewer agreement) and repeatability (intrareviewer agreement) of different ROI size, location, and number combinations for the acquisition of liver PDFF and R2* measurements. The secondary purpose of this study was to establish practical and standardized guidelines for ROI-based liver PDFF and R2* measurements that provide a compromise between measurement performance, as measured by reproducibility and repeatability, and time burden for the image analyst.

Materials and Methods

Subjects

This study retrospectively analyzed 90 liver CSE MRI datasets previously acquired for studies at the University of Wisconsin-Madison [19, 24, 26, 27]. These studies were compliant with HIPAA and approved by the institutional review board. Datasets were divided into two cohorts. Cohort A included 53 subjects (mean age, 51.9 years; age range, 23–84 years; 27 women, 26 men): 34 patients presenting with a clinical indication for an abdominal MRI [19, 24, 26] and 19 healthy subjects recruited as part of a study comparing CT and MRI [27]. To ensure that our analysis included high R2* values, cohort B included 37 patients (mean age, 43.6 years; age range, 10–78 years; 24 men, 13 women) recruited on the basis of having suspected liver iron overload.

Imaging Acquisition and Reconstruction

All imaging was performed at 1.5 T (Signa HDxt and Optima MR 450w, GE Healthcare) using an 8- or 12-channel phased-array cardiac or torso coil. Additionally, all imaging was done using an investigational version of a CSE MRI water-fat separation method. Imaging parameters included: TR range/TE_1 range, 13.5–13.7/1.2–1.3; ΔTE, 2.0 ms; number of signals averaged, 6; FOV, 35 × 35–44 × 44 cm; slice thickness, 8–10 mm; number of slices, 24–32; receiver bandwidth, ± 83 to ± 125 kHz. All image acquisitions were obtained with a low flip angle (5°) to minimize T1-related bias [28].

For subjects in cohort A, reconstruction of the PDFF and R2* maps was performed with an algorithm that provided simultaneous estimates of PDFF and R2*, incorporated spectral modeling of fat, and corrected for noise-related bias and undesired phase shifts (e.g., due to eddy currents and other sources) using a hybrid magnitude-complex fitting approach [20, 29].

For subjects in cohort B, reconstruction of PDFF and R2* maps used a complex-fitting nonlinear least-squares reconstruction algorithm [30, 31]. Complex fitting was used to maximize signal-to-noise ratio and avoid noise floor effects in the estimation of high R2* values [31]. This algorithm also included spectral modeling of fat and noise bias correction for PDFF estimation [20, 28].

Image Analysis

Because the primary purpose of this study was to evaluate the reproducibility and repeatability of different ROI size, location, and number combinations for the acquisition of PDFF and R2* measurements in the liver, our analysis used different sizes, locations, and numbers of circular ROIs that have been reported in recent studies [8, 21–23]. The ROI sizes included 1 cm^2, 4 cm^2, and the largest area that fit inside each placement designation while avoiding large vessels, bile ducts, and obvious image artifacts [8, 22, 23]. The ROI location approaches included the left and right liver lobes; the anterior, posterior, medial, and lateral segments of the liver; and the nine Couinaud segments of the liver [8, 21–23]. The number of ROIs was a direct result of placing one ROI per anatomic designation set forth by the different ROI location approaches. Thus, the number of ROIs included two (one per left and

right liver lobe), four (one per anterior, posterior, medial, and lateral segment), and nine (one per Couinaud segment). The combinations of three different sizes, locations, and numbers of ROIs resulted in nine analysis paradigms (Fig. 1).

Three reviewers (an image analyst with 3 years of experience, a radiologist with 8 years of experience, and a radiologist with 5 years of experience) performed each of the nine analysis paradigms on both the PDFF and R2* maps of cohort A and then repeated every measurement to enable the assessment of intrareviewer agreement. Thus, each reviewer drew a total of 9540 independent ROIs for cohort A.

Three reviewers (an image analyst with 3 years of experience, a radiologist with 8 years of experience, and a medical student under supervision) performed each of the nine analysis paradigms on both the PDFF and R2* maps of cohort B and then repeated every measurement to enable the assessment of intrareviewer agreement. Thus, each reviewer drew a total of 6660 independent ROIs for cohort B.

All image analysis was performed using OsiriX Lite (version 7.0, OsiriX Foundation).

Assessment of Time Burden of Each Paradigm

Because the secondary purpose of this study was to establish guidelines for ROI-based liver PDFF and R2* measurements that provide a compromise between measurement performance and time burden for the image analyst, we recorded the time required for one reviewer to perform each paradigm. The period of time that was recorded included opening PDFF and R2* maps in OsiriX, drawing ROIs on PDFF and R2* maps, and exporting the ROI measurement data into Excel (version 2016, Microsoft).

Statistical Analysis

Each of the reviewers' measurements were processed as follows: for each ROI analysis paradigm and subject, we averaged all of the individual ROI measurements to obtain a single overall PDFF value and a single overall R2* value representing the entire liver as measured by the particular reviewer. These overall values were used to assess the inter- and intrareviewer agreement of all analysis paradigms. Bland-Altman analysis was used to assess inter- and intrareviewer agreement [32]. Because three reviewers performed the measurements, each paradigm underwent three Bland-Altman analyses.

Two-way ANOVA was performed to assess the significant effect of ROI characteristics (number and size) on the performance of the analysis paradigms (inter- and intrareviewer agreement and time burden). The significance threshold used was $p = 0.05$. All statistical analysis was performed using Excel and Matlab (version 9.0, MathWorks).

Results

Summary of Measured Proton Density Fat Fraction and R2*

The PDFF and R2* maps of all 53 subjects in cohort A were included in our analysis. PDFF measurements for this cohort had a mean ± SD of 5.9% ± 8.8% (range, −0.2% to 41.7%). R2* measurements had a mean of 32.3 s^{-1} ± 10.2 s^{-1} (range, 12.2–82.1 s^{-1}). In cohort A, 12

subjects (23%) had elevated liver PDFF (> 6%), and two (4%) had elevated liver R2* (> 60 s^{-1}) [17, 33]. No subjects in cohort A had both elevated PDFF and elevated R2*.

The PDFF and R2* maps of all 37 subjects in cohort B were included in our analysis. PDFF measurements had a mean of 5.0% ± 6.7% (range, −3.5% to 38.7%). R2* measurements for this cohort had a mean of 230.5 s^{-1} ± 155.5 s^{-1} (range, 16.8–704.6 s^{-1}). In cohort B, eight subjects (22%) had elevated liver PDFF (> 6%), and 34 (92%) had elevated liver R2* (> 60 s^{-1}) [17, 33]. Eight subjects in cohort B (22%) had both elevated PDFF and R2*.

Interreviewer Agreement

For cohort A, Bland-Altman analysis of interreviewer agreement showed that placing one largest-fit ROI in all Couinaud segments (paradigm 9) resulted in the narrowest limits of agreement (LOA) (± 2 SD) for liver PDFF measurements (Table 1). The three Bland-Altman analyses of interreviewer agreement for paradigm 9 had a mean difference ± LOA of −0.1% ± 0.8%, 0.2% ± 0.8%, and −0.1% ± 0.8%. However, paradigms 2–9 all had LOA less than ± 2%. Two-way ANOVA showed that both ROI number and ROI size had a significant effect on LOA width (ROI number, $p < 0.001$; ROI size, $p = 0.003$).

For liver R2* measurements in cohort A, Bland-Altman analysis of interreviewer agreement showed that paradigm 9 also resulted in the narrowest LOA (Table 1). The three Bland-Altman analyses of interreviewer agreement for paradigm 9 had a mean difference ± LOA of −0.7 ± 2.8 s^{-1}, 1.0 ± 3.3 s^{-1}, and −0.3 ± 3.0 s^{-1}. Also, only paradigm 9 resulted in all LOA being less than ± 3.4 s^{-1}. Two-way ANOVA showed that both ROI number and ROI size had a significant effect on LOA width (ROI number, $p = 0.02$; ROI size, $p < 0.001$).

For cohort B, Bland-Altman analysis of interreviewer agreement showed that paradigm 9 also resulted in the narrowest LOA for liver PDFF measurements (Table 2). The three Bland-Altman analyses of interreviewer agreement had a mean difference ± LOA of −0.2% ± 1.0%, 0.4% ± 2.2%, −0.3% ± 1.8%. Only paradigm 9 resulted in all LOA being less than ± 2.3%. Two-way ANOVA showed that both ROI number and ROI size had a significant effect on LOA width (ROI number, $p = 0.002$; ROI size, $p = 0.003$).

For liver R2* measurements in cohort B, Bland-Altman analysis of interreviewer agreement showed that paradigm 9 also resulted in the narrowest LOA (Table 2). The three Bland-Altman analyses of interreviewer agreement for paradigm 9 had a mean difference ± LOA of −2.2 s^{-1} ± 11.4 s^{-1}, 2.0 s^{-1} ± 21.6 s^{-1}, and 0.3 s^{-1} ± 2 0.6 s^{-1}. Only paradigm 9 resulted in all LOA being less than ± 22 s^{-1}. Two-way ANOVA showed that both ROI number and ROI size had a significant effect on LOA width (ROI number, $p = 0.009$; ROI size, $p = 0.003$).

Tables 1 and 2 summarize the Bland-Altman analysis of all paradigms for both subject cohorts. Our results show a general trend that the interreviewer agreement improves (i.e., LOA narrow) as the ROI number and ROI size increase (Figs. 2 and 3).

Intrareviewer Agreement

For cohort A, Bland-Altman analysis of intrareviewer agreement showed that placing one largest-fit ROI in all Couinaud segments (paradigm 9) resulted in the narrowest LOA for

liver PDFF measurements (Table 3). The three Bland-Altman analyses of intrareviewer agreement for paradigm 9 had a mean difference ± LOA of 0.0% ± 0.5%, 0.0% ± 0.6%, and 0.0% ± 0.5%. However, paradigms 2–9 had LOA less than ± 2.2%. Two-way ANOVA showed that both ROI number and ROI size had a significant influence on the width of the LOA (ROI number, $p < 0.001$; ROI size, $p = 0.003$).

For liver R2* measurements in cohort A, Bland-Altman analysis of intrareviewer agreement showed that paradigm 9 also resulted in the narrowest LOA (Table 3). The three Bland-Altman analyses of intrareviewer agreement for paradigm 9 had a mean difference ± LOA of 0.4 ± 1.9 s^{-1}, –0.6 ± 3.4 s^{-1}, and 0.0 ± 2.6 s^{-1}. Also, only paradigm 9 resulted in all LOA being less than ± 3.5 s^{-1}. Two-way ANOVA showed that both ROI number and ROI size had a significant effect on the LOA width (ROI number, $p = 0.01$; ROI size, $p = 0.001$).

For cohort B, Bland-Altman analysis of intrareviewer agreement showed that paradigm 9 resulted in the narrowest LOA for liver PDFF measurements (Table 4). The three Bland-Altman analyses of intrareviewer agreement for paradigm 9 had a mean difference ± LOA of −0.9% ± 0.8%, 0.2% ± 1.1%, and −0.2% ± 0.9%. Paradigm 9 resulted in all LOA being less than ± 1.2%. Two-way ANOVA showed that both ROI number and ROI size had a significant effect on the LOA width (ROI number, $p < 0.001$; ROI size, $p = 0.001$).

For liver R2* measurements in cohort B, Bland-Altman analysis of intrareviewer agreement showed that paradigm 9 also resulted in the narrowest LOA (Table 4). The three Bland-Altman analyses of intrareviewer agreement for paradigm 9 had a mean difference ± LOA of −1.0 s^{-1} ± 10.6 s^{-1}, 0.5 s^{-1} ± 13.2 s^{-1}, and 2.4 s^{-1} ± 20.0 s^{-1}. Two-way ANOVA showed that both ROI number and ROI size had a significant effect on the LOA width (ROI number, $p = 0.01$; ROI size, $p = 0.02$).

Tables 3 and 4 summarize the Bland-Altman analysis for all paradigms and subject cohorts. Our results show a general trend that the intrareviewer agreement improves (i.e., LOA narrow) as ROIs increase in number and size (Figs. 4 and 5).

Assessment of Time Burden for Each Paradigm

Time measurement data (Table S1, which can be viewed in the *AJR* electronic supplement to this article, available at www.ajronline.org) showed that acquiring liver PDFF and R2* measurements by placing one ROI in each of the left and right liver lobes (paradigms 1, 4, and 7) required, on average, less than 1 minute per subject (53.5 ± 6.5 s, 50.4 ± 3.2 s, and 58.6 ± 6.1 s, respectively). Paradigms that used four ROIs (paradigms 2, 5, and 7) required just over 1 minute per subject on average (81.2 ± 6.3 s, 68.5 ± 6.3 s, and 73.6 ± 7.6 s, respectively). Paradigms that used nine ROIs (paradigms 3, 6, and 9) required, on average, approximately 2 minutes per subject (113.2 ± 7.6 s, 126.3 ± 8.0 s, and 149.7 ± 8.6 s, respectively). Two-way ANOVA showed that both ROI number and ROI size had a significant effect on the time burden of each paradigm ($p < 0.001$ for both).

The effect of ROI size on time burden was moderate. Measurement time increased by 11.6% when ROI size increased from 1 cm^2 to 4 cm^2 when using nine ROIs (measurement time actually decreased, on average, by 10.7% when using two and four ROIs). Further,

measurement time increased, on average, by 14.1% when ROI size was increased from 4 cm^2 to the largest fit.

The number of ROIs had the largest impact on image analyst time burden. Analysis time increased by an average of 37.8% when ROI number increased from two to four and increased an additional 75.7% when ROI number increased from four to nine. According to our results, the time burden for the reviewer to perform liver PDFF and R2* measurements increases as both the number and size of ROIs increase.

Discussion

In this study we evaluated the reproducibility (interreviewer agreement) and repeatability (intrareviewer agreement) of different ROI size, location, and number for the analysis of quantitative CSE MRI–based liver PDFF and R2* measurements. Our results indicate that the reproducibility and repeatability of liver PDFF and R2* measurements improve as the liver sampling area increases through the use of ROIs that are large in both number and size. This trend was true for subjects with low and high liver iron content (cohorts A and B, respectively).

Our results show that placing one large ROI in each Couinaud segment was the best method (of those evaluated in this work) to achieve the most reproducible and repeatable liver PDFF and R2* measurements. This method includes either 4-cm^2 ROIs (paradigm 6) or largest-fit ROIs (paradigm 9). Our paradigm suggestion complements a recent study by Vu et al. [34], which reported that placing one large (~3 cm^2) ROI in each Couinaud segment results in the highest correlation, using whole-liver segmentation as the mean liver PDFF reference standard.

Our findings relate to other preliminary studies that have been conducted to evaluate the reproducibility of different ROI-based measurement approaches for liver PDFF and R2*. McCarville et al. [35] reported greater interreviewer agreement of R2* measurements when using whole-liver ROIs as opposed to small ROIs (at least 1 cm in diameter) [35]. However, this study did not investigate whether the location and number of ROIs affected reproducibility. Sofue et al. [36] also reported excellent interreviewer agreement of PDFF and R2* measurements when placing one ROI per anterior, posterior, and medial liver segment. However, this study did not compare other ROI location approaches nor did it investigate whether the size and number of ROIs affected reproducibility. Additionally, neither of these studies addressed intrareviewer agreement (repeatability).

Our study also sought to establish practical and standardized guidelines for ROI-based liver PDFF and R2* measurements that provide a compromise between the reproducibility and repeatability of measurements and the time burden for the image analyst. As expected, the time burden for the reviewer to acquire ROI-based PDFF and R2* measurements in the liver increases as the number and size of ROIs increases. Thus, despite the improved reproducibility and repeatability of paradigms 6 and 9, these paradigms are the most time consuming. Because ROI size has only a modest effect on time burden, a compromise between time burden and ROI analysis paradigm performance is most easily achieved by

reducing the number of ROIs. We suggest that placing one large ROI in the anterior, posterior, medial, and lateral segments of the liver, as indicated by paradigms 5 and 8, may be a reasonable compromise in some circumstances to balance reproducibility and repeatability of measurements with the time burden needed to perform the ROI analysis. Placing one largest-fit ROI in the left and right liver lobes (paradigm 7) may also be a reasonable compromise.

We believe that this work may provide researchers and clinicians with insight about the reproducibility, repeatability, and time burden of various ROI sampling methods. Depending on the specific application and requirements of a study, our results can be used to make decisions about the trade-offs between greater reproducibility and repeatability of liver PDFF and R2* measurements and the time burden for the image analyst. Further, we believe that this work may provide some practical and standardized approaches to MRI-based measurements of liver PDFF and R2*. The current literature contains considerable variability in the choice of ROI size, location, and number combinations. Such standardization is essential for multicenter research studies and clinical applications of these techniques, as well as for pooling results from different studies.

This study had several limitations. First, our analysis did not include whole-liver ROIs. Although whole-liver ROIs have shown good inter- and intrareviewer agreement for liver PDFF and good interreviewer agreement for liver R2*, manual segmentation of the whole liver is extraordinarily time consuming and not practical for most applications [35, 37]. Inter- and intrareviewer agreement improved in our study when the liver sampling area was increased, which indicates that whole-liver segmentation would further improve inter- and intrareviewer agreement of liver PDFF and R2*. Second, our study only focused on evaluating the reproducibility and repeatability of ROI sampling methods by assessing inter- and intrareviewer agreement, respectively; we did not evaluate the accuracy of the various proposed ROI sampling methods for the quantification of liver fat and iron content. Future work should focus on assessing the agreement between paradigms 5, 7, and 8 (which provide a good trade-off between repeatability and reproducibility and time burden) and accepted reference standards for liver PDFF and R2*. Third, just 22% (20/90) of the subjects evaluated in this study had elevated liver PDFF (> 6%). Thus, it is difficult to determine if our results would be consistent over a larger range of hepatic steatosis. Fourth, for paradigms that included multiple ROIs of different sizes (paradigms 7–9), we calculated the unweighted average of the measurements from the different ROIs. We note that weighted averaging based on the estimated variance of each ROI measurement might offer better performance. However, for simplicity, we did not use weighted averaging in this study. Finally, our study only included datasets acquired at 1.5 T; a similar evaluation at 3 T is needed. However, we expect that the reproducibility and repeatability of liver PDFF and R2* measurements obtained at 3 T will show qualitative trends similar to our results at 1.5 T.

In conclusion, this study evaluated the reproducibility and repeatability of different ROI size, location, and number combinations to analyze CSE MRI–based PDFF and R2* measurements in the liver. Inter- and intrareviewer agreement improved when the sampling area of the liver was increased by using multiple large ROIs to measure PDFF and R2*. Clinicians and researchers performing ROI-based measurements of liver PDFF and R2* can

improve the reproducibility and repeatability of PDFF and R2* measurements by sampling as much of the liver as possible with multiple large ROIs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by the Wisconsin Alumni Research Foundation Accelerator Program, National Institutes of Health grants (R01DK083380, R01DK088925, K24DK102595, R01DK100651, UL1TR00427, UL1RR025011), and research support to the University of Wisconsin from GE Healthcare and Bracco Diagnostics.

References

1. Reeder SB, Cruite I, Hamilton G, Sirlin CB. Quantitative assessment of liver fat with magnetic resonance imaging and spectroscopy. J Magn Reson Imaging. 2011; 34:729–749. [PubMed: 21928307]
2. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging. 2012; 36:1011–1014. [PubMed: 22777847]
3. Reeder SB. Emerging quantitative magnetic resonance imaging biomarkers of hepatic steatosis. Hepatology. 2013; 58:1877–1880. [PubMed: 23744793]
4. Rehm JL, Connor EL, Wolfgram PM, Eickhoff JC, Reeder SB, Allen DB. Predicting hepatic steatosis in a racially and ethnically diverse cohort of adolescent girls. J Pediatr. 2014; 165:319–325.e1. [PubMed: 24857521]
5. Yokoo T, Bydder M, Hamilton G, et al. Nonalcoholic fatty liver disease: diagnostic and fat-grading accuracy of low-flip-angle multiecho gradient-recalled-echo MR imaging at 1.5 T. Radiology. 2009; 251:67–76. [PubMed: 19221054]
6. Tang A, Tan J, Sun M, et al. Nonalcoholic fatty liver disease: MR imaging of liver proton density fat fraction to assess hepatic steatosis. Radiology. 2013; 267:422–431. [PubMed: 23382291]
7. Noureddin M, LAm J, Peterson MR, et al. Utility of magnetic resonance imaging versus histology for quantifying changes in liver fat in nonalcoholic fatty liver disease. Hepatology. 2013; 58:1930–1940. [PubMed: 23696515]
8. Henninger B, Kremser C, Rauch S, et al. Evaluation of liver fat in the presence of iron with MRI using T2* correction: a clinical approach. Eur Radiol. 2013; 23:1643–1649. [PubMed: 23334458]
9. Bellentani S, Scaglioni F, Marino M, Bedogni G. Epidemiology of non-alcoholic fatty liver disease. Dig Dis. 2010; 28:155–161. [PubMed: 20460905]
10. Kühn JP, Hernando D, Meffert PJ, et al. Protondensity fat fraction and simultaneous R2* estimation as an MR tool for assessment of osteoporosis. Eur Radiol. 2013; 23:3432–3439. [PubMed: 23812246]
11. Karçaaltincaba M, Idilman I, Celik A. Focal sparing of iron and fat in liver tissue in patients with hemosiderosis: diagnosis with combination of R2* relaxometry and proton density fat fraction calculation by MRI. Diagn Interv Radiol. 2011; 17:323–327. [PubMed: 21267943]
12. Hernando D, Cook RJ, Diamond C, Reeder SB. Magnetic susceptibility as a B0 field strength independent MRI biomarker of liver iron overload. Magn Reson Med. 2013; 70:648–656. [PubMed: 23801540]
13. Schwenzer NF, Machann J, Haap MM, et al. T2* relaxometry in the liver, pancreas, and spleen in a healthy cohort of one hundred twenty-nine subjects: correlation with age, gender, and serum ferritin. Invest Radiol. 2008; 43:854–860. [PubMed: 19002057]
14. St Pierre TG, Clark PR, Chua-anusom W, et al. Noninvasive measurement and imaging of liver iron concentrations using proton magnetic resonance. Blood. 2005; 105:855–861. [PubMed: 15256427]

15. Gandon Y, Olivié D, Guyader D, et al. Non-invasive assessment of hepatic iron stores by MRI. Lancet. 2004; 363:357–362. [PubMed: 15070565]
16. Gianesin B, Zefiro D, Musso M, et al. Measurement of liver iron overload: noninvasive calibration of MRI-R2* by magnetic iron detector susceptometer. Magn Reson Med. 2012; 67:1782–1786. [PubMed: 22135193]
17. 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. Blood. 2005; 106:1460–1465. [PubMed: 15860670]
18. Hankins JS, McCarville MB, Loeffler RB, et al. R2* magnetic resonance imaging of the liver in patients with iron overload. Blood. 2009; 113:4853–4855. [PubMed: 19264677]
19. Meisamy S, Hines CDG, Hamilton G, et al. Quantification of hepatic steatosis with T1-independent, T2*-corrected MR imaging with spectral modeling of fat: blinded comparison with MR spectroscopy. Radiology. 2011; 258:767–775. [PubMed: 21248233]
20. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB. Multiecho water-fat separation and simultaneous R2* estimation with multifrequency fat spectrum modeling. Magn Reson Med. 2008; 60:1122–1134. [PubMed: 18956464]
21. Motosugi U, Hernando D, Bannas P, et al. Quantification of liver fat with respiratory-gated quantitative chemical shift encoded MRI. J Magn Reson Imaging. 2015; 42:1241–1248. [PubMed: 25828696]
22. 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]
23. Bannas P, Kramer H, Hernando D, et al. Quantitative MR imaging of hepatic steatosis: validation in ex vivo human livers. Hepatology. 2015; 62:1444–1455. [PubMed: 26224591]
24. Hines CDG, Frydrychowicz A, Hamilton G, et al. T1 independent, T2* corrected chemical shift based fat–water separation with multi-peak fat spectral modeling is an accurate and precise measure of hepatic steatosis. J Magn Reson Imaging. 2011; 33:873–881. [PubMed: 21448952]
25. Achmad E, Yokoo T, Hamilton G, 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]
26. Hernando D, Wells SA, Vigen KK, Reeder SB. Effect of hepatocyte-specific gadolinium-based contrast agents on hepatic fat-fraction and R2*. Magn Reson Imaging. 2015; 33:43–50. [PubMed: 25305414]
27. Kramer H, Pickhardt PJ, Kliewer MA, et al. Accuracy of liver fat quantification with advanced CT, MRI, and ultrasound techniques: prospective comparison with MR spectroscopy. AJR. 2017; 208:92–100. [PubMed: 27726414]
28. Liu CY, McKenzie CA, Yu H, Brittain JH, Reeder SB. Fat quantification with IDEAL gradient echo imaging: correction of bias from T(1) and noise. Magn Reson Med. 2007; 58:354–364. [PubMed: 17654578]
29. Yu H, Shimakawa A, Hines CD, et al. Combination of complex-based and magnitude-based multiecho water-fat separation for accurate quantification of fat-fraction. Magn Reson Med. 2011; 66:199–206. [PubMed: 21695724]
30. Hernando D, Kellman P, Haldar JP, Liang ZP. Robust water/fat separation in the presence of large field inhomogeneities using a graph cut algorithm. Magn Reson Med. 2010; 63:79–90. [PubMed: 19859956]
31. Hernando D, Kramer JH, Reeder SB. Multipeak fat-corrected complex R2* relaxometry: theory, optimization, and clinical validation. Magn Reson Med. 2013; 70:1319–1331. [PubMed: 23359327]
32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1:307–310. [PubMed: 2868172]
33. Szczepaniak LS, Nurenberg P, Leonard D, et al. Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. Am J Physiol Endocrinol Metab. 2005; 288:E462–E468. [PubMed: 15339742]

34. Vu KN, Gilbert G, Chalut M, Chagnon M, Chartrand G, Tang A. MRI-determined liver proton density fat fraction, with MRS validation: comparison of regions of interest sampling methods in patients with type 2 diabetes. J Magn Reson Imaging. 2016; 43:1090–1099. [PubMed: 26536609]

35. McCarville MB, Hillenbrand CM, Loeffler RB, et al. Comparison of whole liver and small region-of-interest measurements of MRI liver R2* in children with iron overload. Pediatr Radiol. 2010; 40:1360–1367. [PubMed: 20333511]

36. Sofue K, Mileto A, Dale BM, Zhong X, Bashir MR. Interexamination repeatability and spatial heterogeneity of liver iron and fat quantification using MRI-based multistep adaptive fitting algorithm. J Magn Reson Imaging. 2015; 42:1281–1290. [PubMed: 25920074]

37. Tang A, Chen J, Le TA, et al. Cross-sectional and longitudinal evaluation of liver volume and total liver fat burden in adults with nonalcoholic steatohepatitis. Abdom Imaging. 2015; 40:26–37. [PubMed: 25015398]

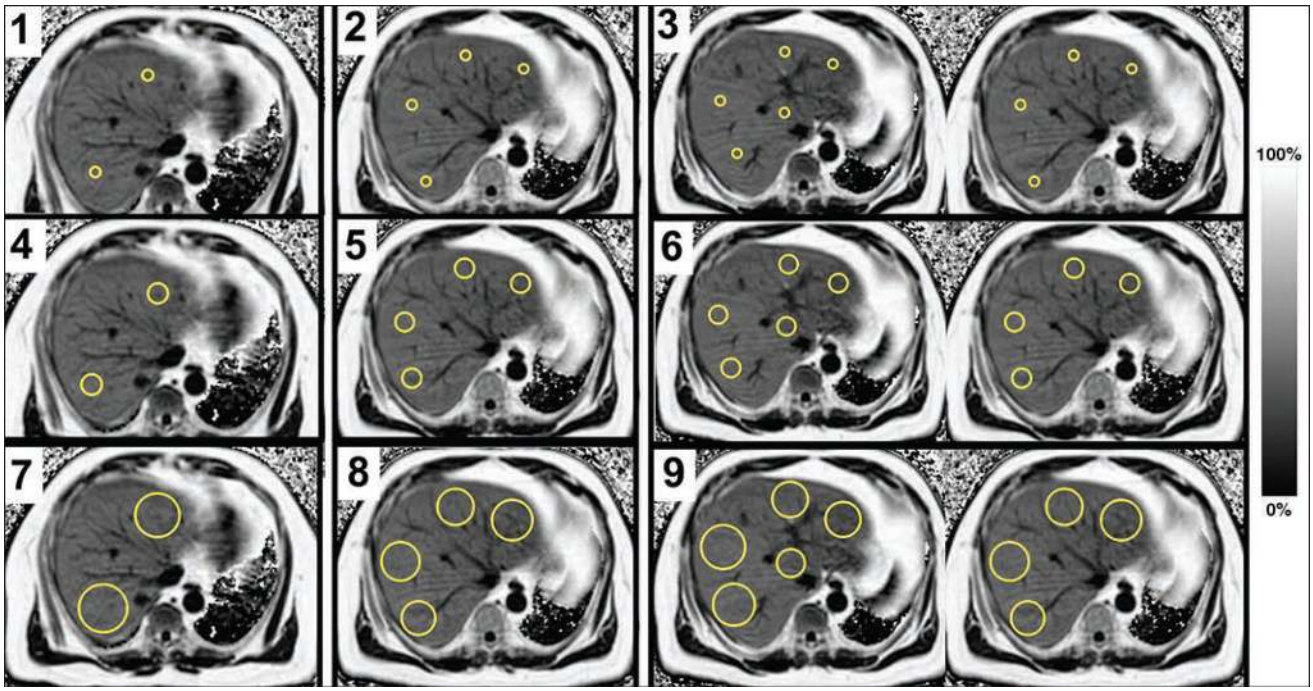

Fig. 1.

Visual representation of size, location, and number of ROIs for paradigms 1–9 on proton density fat fraction maps. ROIs for paradigms 1, 4, and 7 were placed in left and right liver lobes, and their sizes were 1 cm^2, 4 cm^2, and largest fit, respectively. Paradigms 2, 5, and 8 had ROIs in anterior, posterior, medial, and lateral liver segments, and their sizes were 1 cm^2, 4 cm^2, and largest fit, respectively. Paradigms 3, 6, and 9 had ROIs in nine Couinaud segments (at least two slices needed), and their sizes were 1 cm^2, 4 cm^2, and largest fit, respectively.

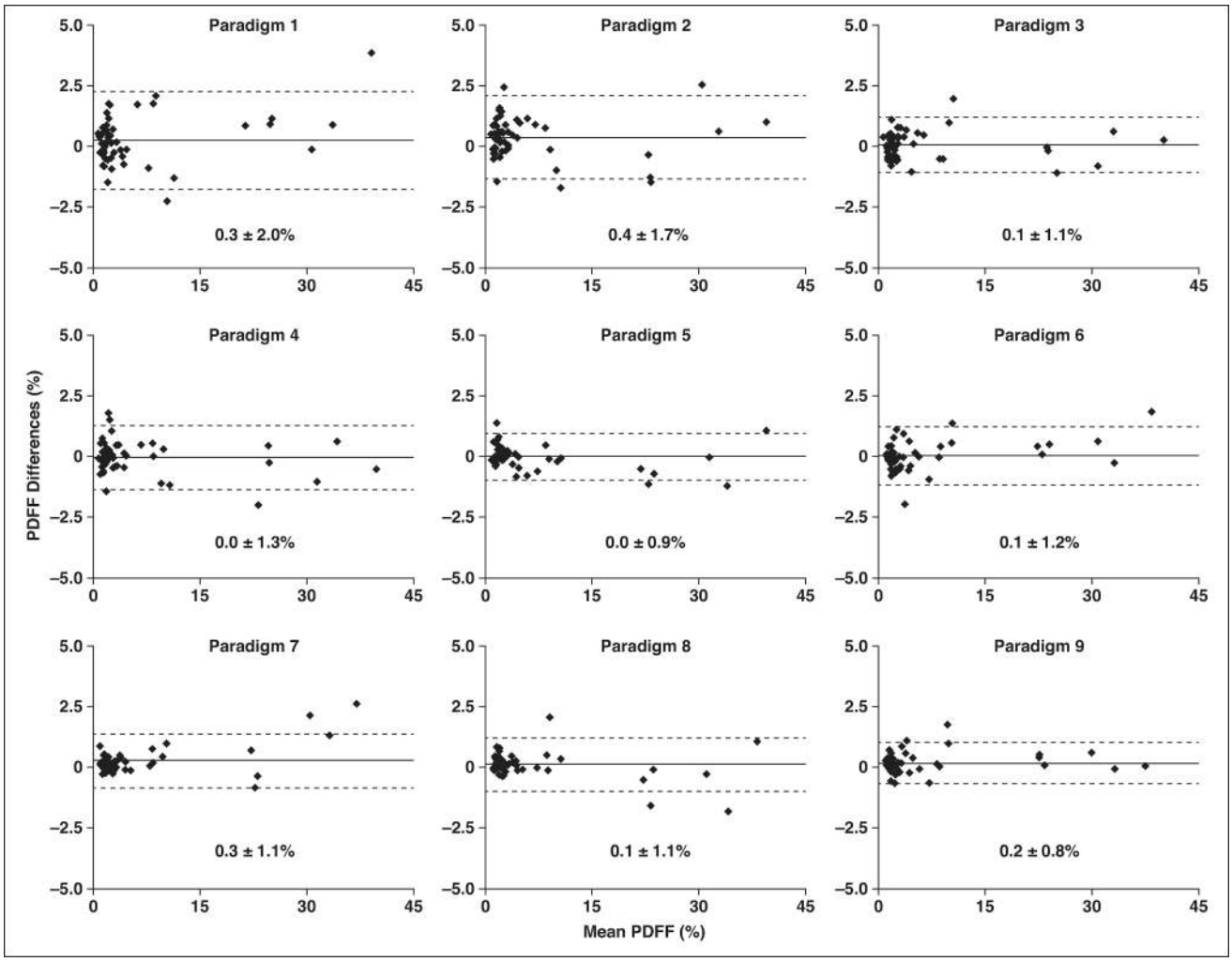

Fig. 2.

Bland-Altman plots show that interreviewer agreement for reviewers 2 and 3 increases (i.e., limits of agreement shown by dotted lines narrow) for proton density fat fraction (PDFF) in cohort A as ROI size increases (top to bottom) and ROI number increases (left to right). Solid lines indicate mean difference.

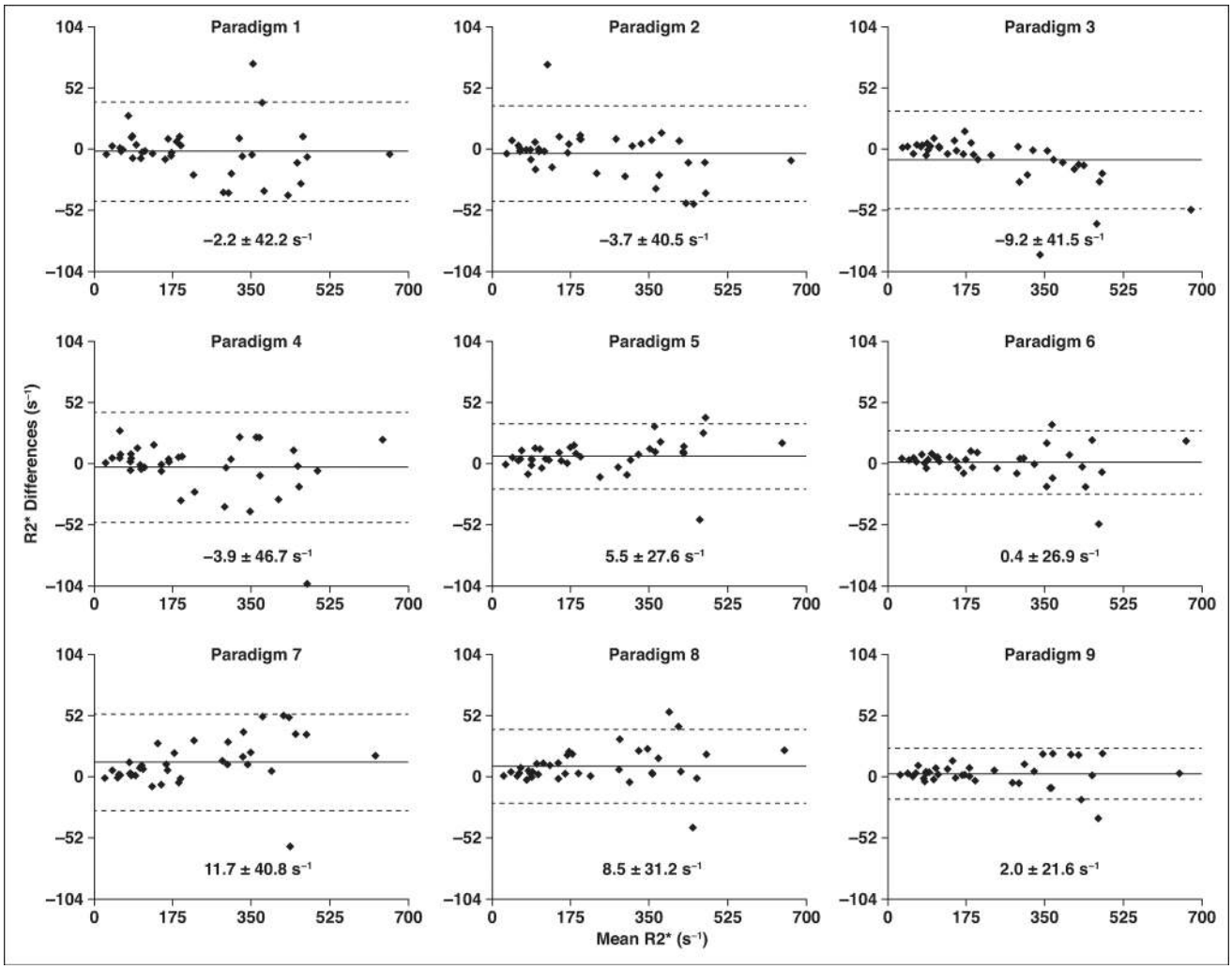

Fig. 3.

Bland-Altman plots show that interreviewer agreement for reviewers 2 and 3 increases (i.e., limits of agreement shown by dotted lines narrow) for R2* in cohort B as ROI size increases (top to bottom) and ROI number increases (left to right). Solid lines indicate mean difference.

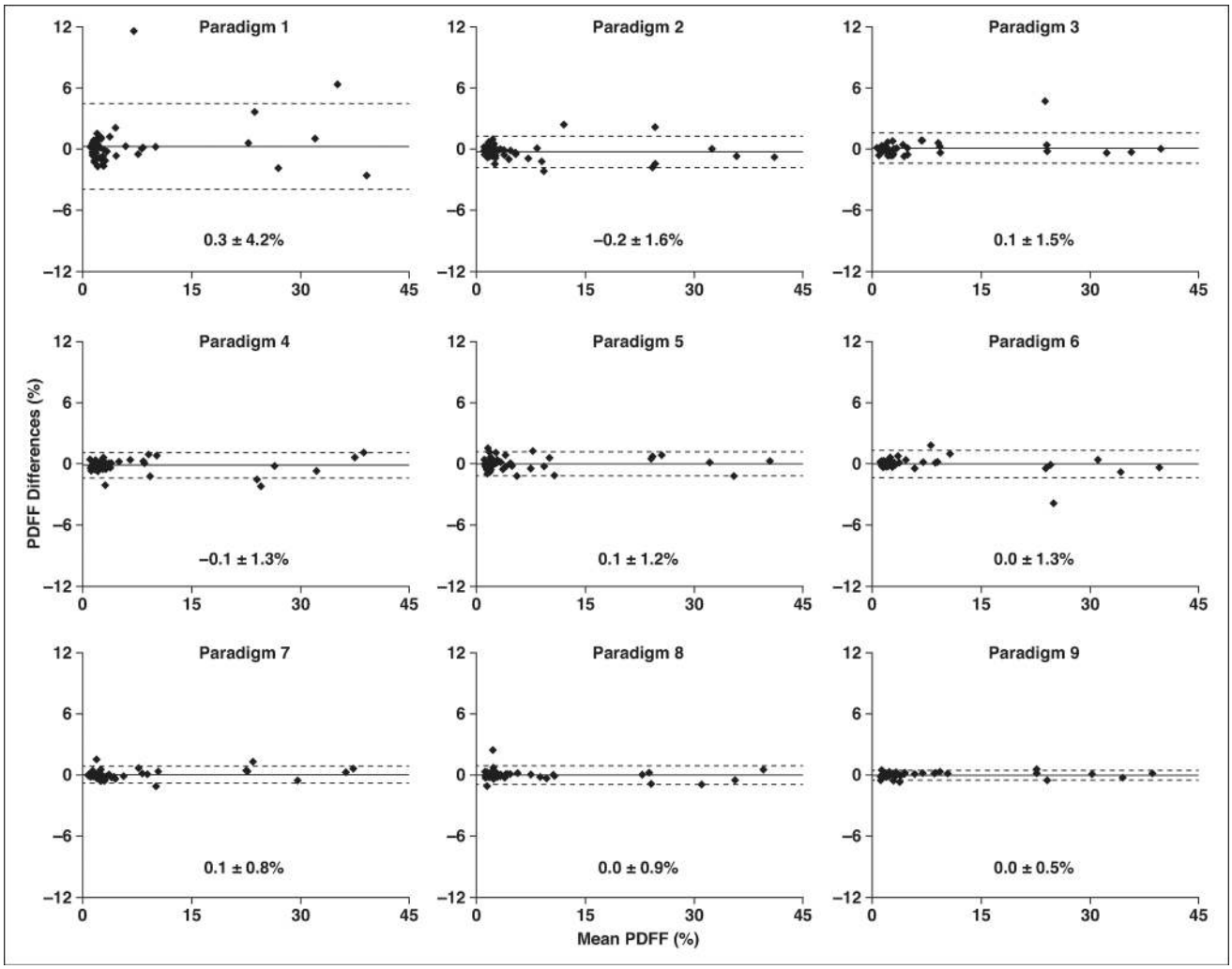

Fig. 4.

Bland-Altman plots show that intrareviewer agreement for reviewer 1 increases (i.e., limits of agreement shown by dotted lines narrow) for proton density fat fraction (PDFF) in cohort A as ROI size increases (top to bottom) and ROI number increases (left to right). Solid lines indicate mean difference.

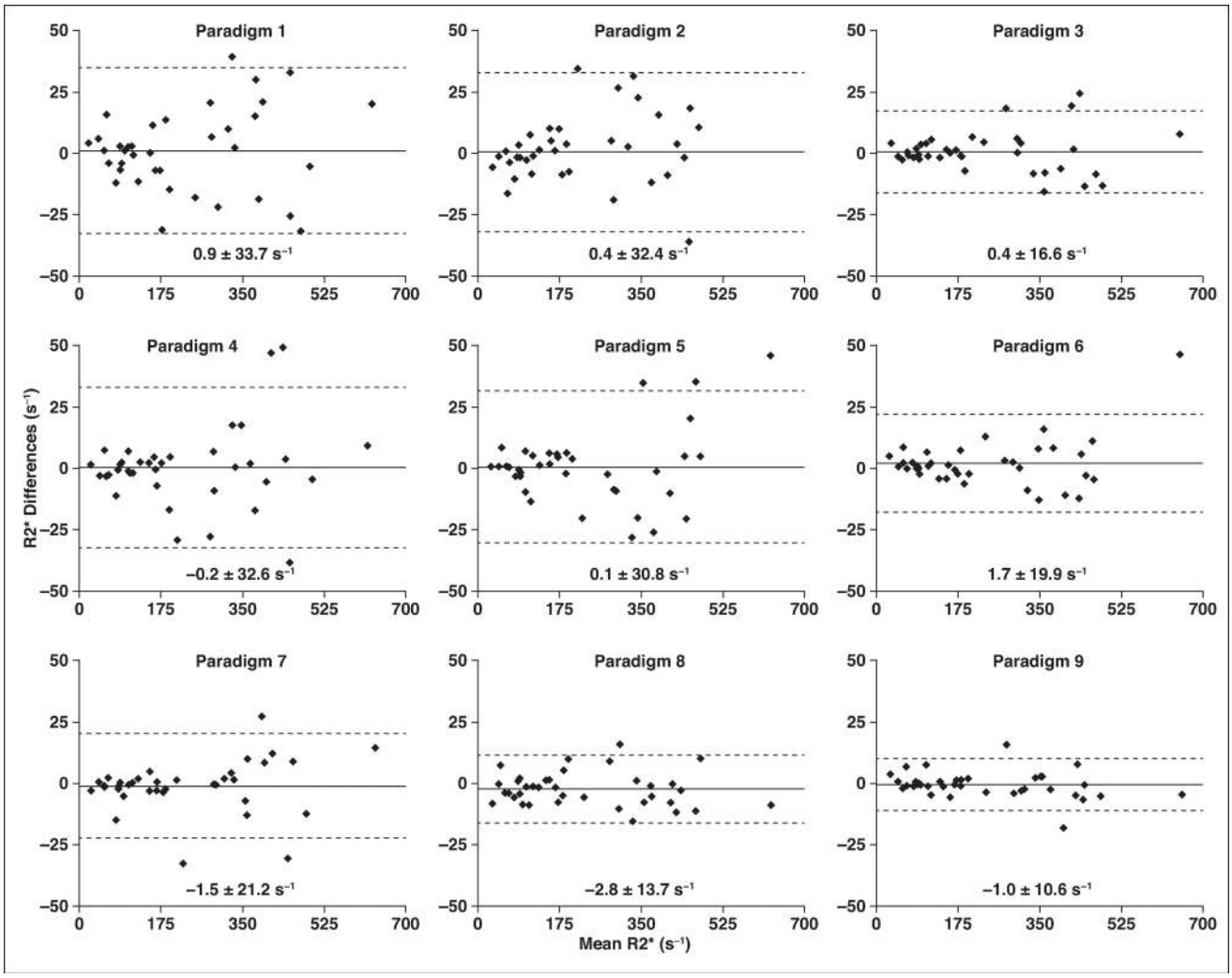

Fig. 5.

Bland-Altman plots show that intrareviewer agreement for reviewer 1 increases (i.e., limits of agreement shown by dotted lines narrow) for R2* in cohort B as ROI size increases (top to bottom) and ROI number increases (left to right). Solid lines indicate mean difference.

TABLE 1

Interreviewer Agreement of Proton Density Fat Fraction and R2* for Cohort A

Measurement, Paradigm	No. of ROIs	ROI Size	Reviewer 1 vs Reviewer 2	Reviewer 2 vs Reviewer 3	Reviewer 3 vs Reviewer 1	Mean LOA
Proton density fat fraction (%)						
1	2	1 cm^2	−0.2 ± 3.4	0.3 ± 2.0	−0.1 ± 7.9	± 4.4
2	4	1 cm^2	0.1 ± 1.8	0.4 ± 1.7	−0.5 ± 1.9	± 1.8
3	9	1 cm^2	−0.1 ± 1.5	0.1 ± 1.1	0.0 ± 1.5	± 1.4
4	2	4 cm^2	0.2 ± 1.5	0.0 ± 1.3	−0.2 ± 1.8	± 1.6
5	4	4 cm^2	0.0 ± 1.6	0.0 ± 0.9	0.0 ± 1.3	± 1.3
6	9	4 cm^2	0.0 ± 1.6	0.1 ± 1.2	0.0 ± 1.8	± 1.5
7	2	Largest fit	−0.1 ± 1.2	0.3 ± 1.1	−0.1 ± 1.3	± 1.2
8	4	Largest fit	0.0 ± 1.2	0.1 ± 1.1	−0.1 ± 1.0	± 1.1
9	9	Largest fit	−0.1 ± 0.8	0.2 ± 0.8	−0.1 ± 0.8	± 0.8
Effect on LOA width, *p*	0.0001	0.003				
R2* (s^{-1})						
1	2	1 cm^2	−0.3 ± 8.9	0.4 ± 7.5	−0.1 ± 9.3	± 8.5
2	4	1 cm^2	−2.5 ± 8.2	1.7 ± 9.7	0.8 ± 4.8	± 7.6
3	9	1 cm^2	−0.6 ± 5.7	−0.5 ± 5.0	1.1 ± 5.2	± 5.3
4	2	4 cm^2	0.9 ± 7.2	−1.0 ± 4.7	0.1 ± 7.4	± 6.5
5	4	4 cm^2	−1.7 ± 8.7	0.9 ± 8.5	0.8 ± 5.8	± 7.7
6	9	4 cm^2	−0.3 ± 4.6	0.0 ± 6.9	0.3 ± 4.0	± 5.2
7	2	Largest fit	−1.2 ± 7.3	1.0 ± 6.1	0.1 ± 5.4	± 6.3
8	4	Largest fit	−2.5 ± 7.2	1.9 ± 7.0	0.6 ± 5.3	± 6.5
9	9	Largest fit	−0.7 ± 2.8	1.0 ± 3.3	−0.3 ± 3.0	± 3.0
Effect on LOA width, *p*	0.02	0.0001				

Note—LOA = limits of agreement.

TABLE 2

Interreviewer Agreement of Proton Density Fat Fraction and R2* for Cohort B

Measurement, Paradigm	No. of ROIs	ROI Size	Reviewer 1 vs Reviewer 2	Reviewer 2 vs Reviewer 3	Reviewer 3 vs Reviewer 1	Mean LOA
Proton density fat fraction (%)						
1	2	1 cm^2	–0.4 ± 3.6	1.4 ± 5.1	–1.0 ± 4.5	± 4.4
2	4	1 cm^2	0.0 ± 2.2	0.4 ± 4.7	–0.4 ± 5.1	± 4.0
3	9	1 cm^2	0.5 ± 1.9	1.3 ± 4.2	–0.8 ± 3.6	± 3.2
4	2	4 cm^2	0.3 ± 3.8	0.9 ± 3.6	–0.6 ± 4.2	± 3.9
5	4	4 cm^2	0.2 ± 1.6	0.5 ± 2.4	–0.3 ± 2.0	± 2.0
6	9	4 cm^2	0.4 ± 1.5	0.5 ± 2.7	–0.2 ± 2.3	± 2.2
7	2	Largest fit	0.7 ± 3.2	1.1 ± 3.7	–0.4 ± 2.7	± 3.2
8	4	Largest fit	0.7 ± 1.6	0.8 ± 2.4	–0.1 ± 2.2	± 2.1
9	9	Largest fit	0.2 ± 1.0	0.4 ± 2.2	–0.3 ± 1.8	± 1.7
Effect on LOA width, *p*	0.002	0.003				
R2* (s^{-1})						
1	2	1 cm^2	1.8 ± 46.9	–2.2 ± 42.2	3.9 ± 46.7	± 45.3
2	4	1 cm^2	1.6 ± 28.7	–3.7 ± 40.5	5.3 ± 42.7	± 37.3
3	9	1 cm^2	–1.8 ± 18.3	–9.2 ± 41.5	7.3 ± 43.0	± 34.3
4	2	4 cm^2	–4.1 ± 27.6	–3.9 ± 46.7	8.1 ± 46.0	± 40.1
5	4	4 cm^2	4.2 ± 29.3	5.5 ± 27.6	–1.3 ± 31.9	± 29.6
6	9	4 cm^2	2.4 ± 19.0	0.4 ± 26.9	1.9 ± 25.6	± 23.8
7	2	Largest fit	4.7 ± 37.7	11.7 ± 40.8	–7.0 ± 24.0	± 34.2
8	4	Largest fit	2.9 ± 22.8	8.5 ± 31.2	–5.6 ± 27.9	± 27.3
9	9	Largest fit	2.2 ± 11.4	2.0 ± 21.6	0.3 ± 20.6	± 17.9
Effect on LOA width, *p*	0.009	0.003				

Note—LOA = limits of agreement.

AJR Am J Roentgenol. Author manuscript; available in PMC 2017 October 13.

TABLE 3

Intrareviewer Agreement of Proton Density Fat Fraction and R2* for Cohort A

Measurement, Paradigm	No. of ROIs	ROI Size	Reviewer 1	Reviewer 2	Reviewer 3	Mean LOA
Proton density fat fraction (%)						
1	2	1 cm^2	0.3 ± 4.2	−0.1 ± 2.0	0.0 ± 2.6	± 2.9
2	4	1 cm^2	−0.2 ± 1.6	−0.1 ± 1.5	0.1 ± 2.1	± 1.7
3	9	1 cm^2	0.1 ± 1.5	−0.0 ± 0.9	0.0 ± 1.6	± 1.3
4	2	4 cm^2	−0.1 ± 1.3	−0.1 ± 1.3	0.1 ± 2.0	± 1.5
5	4	4 cm^2	0.1 ± 1.2	0.0 ± 1.1	0.1 ± 1.1	± 1.1
6	9	4 cm^2	0.0 ± 1.3	−0.1 ± 1.0	0.0 ± 1.1	± 1.1
7	2	Largest fit	0.1 ± 0.8	−0.0 ± 1.2	−0.1 ± 0.9	± 1.0
8	4	Largest fit	0.0 ± 0.9	−0.0 ± 0.7	0.2 ± 1.1	± 0.9
9	9	Largest fit	0.0 ± 0.5	0.0 ± 0.6	0.0 ± 0.5	± 0.5
Effect on LOA width, *p*	0.0001	0.003				
R2* (s^{-1})						
1	2	1 cm^2	0.6 ± 6.1	1.1 ± 11.4	−0.3 ± 9.0	± 8.8
2	4	1 cm^2	−0.5 ± 4.5	−0.6 ± 7.2	0.4 ± 5.0	± 5.6
3	9	1 cm^2	−0.3 ± 4.5	−0.6 ± 5.2	0.5 ± 4.6	± 4.8
4	2	4 cm^2	−0.8 ± 5.2	0.6 ± 6.8	1.1 ± 6.7	± 6.2
5	4	4 cm^2	0.8 ± 4.6	−0.2 ± 6.9	1.0 ± 5.8	± 5.8
6	9	4 cm^2	−0.4 ± 2.7	−0.4 ± 4.8	1.2 ± 4.5	± 4.0
7	2	Largest fit	0.5 ± 4.3	−0.2 ± 6.3	0.8 ± 4.7	± 5.1
8	4	Largest fit	0.7 ± 3.3	−0.6 ± 6.7	0.3 ± 3.8	± 4.6
9	9	Largest fit	0.4 ± 1.9	−0.6 ± 3.4	0.0 ± 2.6	± 2.6
Effect on LOA width, *p*	0.01	0.001				

Note—LOA = limits of agreement.

TABLE 4

Intrareviewer Agreement of Proton Density Fat Fraction and R2* for Cohort B

Measurement, Paradigm	No. of ROIs	ROI Size	Reviewer 1	Reviewer 2	Reviewer 3	Mean LOA
Proton density fat fraction (%)						
1	2	1 cm^2	−0.2 ± 4.2	0.3 ± 3.2	0.1 ± 4.7	± 4.0
2	4	1 cm^2	−0.0 ± 2.1	0.4 ± 3.0	0.3 ± 5.0	± 3.4
3	9	1 cm^2	0.2 ± 1.9	−0.1 ± 1.7	−0.2 ± 3.1	± 2.2
4	2	4 cm^2	0.3 ± 2.3	0.3 ± 2.2	−0.3 ± 2.5	± 2.3
5	4	4 cm^2	0.0 ± 1.6	0.2 ± 1.6	0.2 ± 1.8	± 1.7
6	9	4 cm^2	0.1 ± 1.1	−0.1 ± 1.4	0.0 ± 0.9	± 1.1
7	2	Largest fit	0.1 ± 2.1	0.1 ± 1.8	−0.1 ± 1.7	± 1.9
8	4	Largest fit	0.2 ± 1.3	0.2 ± 1.6	0.0 ± 1.5	± 1.5
9	9	Largest fit	−0.9 ± 0.8	0.2 ± 1.1	−0.2 ± 0.9	± 0.9
Effect on LOA width, *p*	10^{-5}	0.001				
R2* (s^{-1})						
1	2	1 cm^2	0.9 ± 33.7	0.5 ± 31.0	−11.5 ± 52.6	± 39.1
2	4	1 cm^2	0.4 ± 32.4	3.9 ± 22.9	2.4 ± 51.4	± 35.6
3	9	1 cm^2	0.4 ± 16.6	0.2 ± 19.7	10.2 ± 38.8	± 25.0
4	2	4 cm^2	−0.2 ± 32.6	1.9 ± 27.5	1.5 ± 32.1	± 30.7
5	4	4 cm^2	0.1 ± 30.8	0.0 ± 16.8	1.2 ± 22.6	± 23.4
6	9	4 cm^2	1.7 ± 19.9	−0.9 ± 19.1	−1.2 ± 12.5	± 17.2
7	2	Largest fit	−1.5 ± 21.2	0.7 ± 20.6	−5.4 ± 31.9	± 24.6
8	4	Largest fit	−2.8 ± 13.7	2.0 ± 20.7	−7.6 ± 27.1	± 20.5
9	9	Largest fit	−1.0 ± 10.6	0.5 ± 13.2	2.4 ± 20.0	± 14.6
Effect on LOA width, *p*	0.01	0.02				

Note—LOA = limits of agreement.