Published in final edited form as:

Hepatology. 2015 November ; 62(5): 1444–1455. doi:10.1002/hep.28012.

Quantitative MR Imaging of Hepatic Steatosis: Validation in Ex Vivo Human Livers

Peter Bannas^{1,2,*}, Harald Kramer^{1,3,*}, Diego Hernando¹, Rashmi Agni⁴, Ashley M. Cunningham⁴, Rakesh Mandal⁴, Utaroh Motosugi¹, Samir D. Sharma¹, Alejandro Munoz del Rio¹, Luis Fernandez⁵, and Scott B. Reeder^{1,6,7,8,9}

Peter Bannas: PBannas@uwhealth.org; Harald Kramer: Harald.kramer@med.uni-muenchen.de; Diego Hernando: DHernando@wisc.edu; Rashmi Agni: RAgni@uwhealth.org; Ashley M. Cunningham: ACunningham@uwhealth.org; Rakesh Mandal: RMandal@uwhealth.org; Utaroh Motosugi: UMotosugi@uwhealth.org; Samir D. Sharma: SDsharma2@wisc.edu; Alejandro Munoz del Rio: AMunozDelRio@uwhealth.org; Luis Fernandez: Luisf@surgery.wisc.edu; Scott B. Reeder: SReeder@uwhealth.org

¹Department of Radiology, University of Wisconsin-Madison, Madison, WI

²Department of Radiology, University Hospital Hamburg-Eppendorf, Hamburg, Germany

³Department of Radiology, Ludwig-Maximilians-University Hospital, Munich, Germany

⁴Department of Pathology, University of Wisconsin-Madison, Madison, WI

⁵Department of Surgery, University of Wisconsin-Madison, Madison, WI

⁶Department of Medical Physics, University of Wisconsin-Madison, Madison, WI

⁷Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI

⁸Department of Medicine, University of Wisconsin-Madison, Madison, WI

⁹Department of Emergency Medicine, University of Wisconsin-Madison, Madison, WI

Abstract

Emerging magnetic resonance imaging (MRI) biomarkers of hepatic steatosis have demonstrated tremendous promise for accurate quantification of hepatic triglyceride concentration. These methods quantify the “proton density fat-fraction” (PDFF), which reflects the concentration of triglycerides in tissue. Previous *in vivo* studies have compared MRI-PDFF with histologic steatosis grading for assessment of hepatic steatosis. However, the correlation of MRI-PDFF with the underlying hepatic triglyceride content remained unknown. The aim of this *ex vivo* study was to validate the accuracy of MRI-PDFF as an imaging biomarker of hepatic steatosis. Using *ex vivo* human livers, we compared MRI-PDFF with magnetic resonance spectroscopy-PDFF (MRS-PDFF), biochemical triglyceride extraction and histology as three independent reference standards. A secondary aim was to compare the precision of MRI-PDFF relative to biopsy for the quantification of hepatic steatosis.

MRI-PDFF was prospectively performed at 1.5T in 13 explanted human livers. We performed co-localized paired evaluation of liver fat content in all nine Couinaud segments using single-voxel

Contact Information: Scott B. Reeder, M.D., Ph.D., Department of Radiology, University of Wisconsin-Madison, 600 Highland Avenue, Room E1/372, Madison, WI 53792–3252, Phone (608) 698-6121, sreeder@uwhealth.org.

*These authors contributed equally

MRS-PDFF (n=117), tissue wedges for biochemical triglyceride extraction (n=117), and five core biopsies performed in each segment for histologic grading (n=585). Accuracy of MRI-PDFF was assessed through linear regression with MRS-PDFF, triglyceride extraction and histology. Intra-observer agreement, inter-observer agreement and repeatability of MRI-PDFF and histologic grading were assessed through Bland-Altman analyses.

MRI-PDFF showed an excellent correlation with MRS-PDFF ($r=0.984$; CI: 0.978–0.989) and strong correlation with histology ($r=0.850$; CI: 0.791–0.894) and triglyceride extraction ($r=0.871$; CI: 0.818–0.909). Intra-observer agreement, inter-observer agreement and repeatability showed a significantly smaller variance for MRI-PDFF than for histologic steatosis grading (all $p<0.001$).

Conclusion—MRI-PDFF is an accurate, precise and reader-independent non-invasive imaging biomarker of liver triglyceride content, capable of steatosis quantification over the entire liver.

Keywords

Magnetic resonance imaging; Proton density fat fraction; Histologic fat quantification; Hepatic triglyceride analysis; Noninvasive biomarker of steatosis

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) affects an estimated 30% of the US population and is the most common cause of chronic liver disease (1–3). NAFLD is associated with the metabolic syndrome (4), higher rates of malignancy (5–7) and emerging evidence suggests it plays a causative role in type II diabetes (8). NAFLD encompasses a spectrum of liver diseases including isolated steatosis, non-alcoholic steatohepatitis (NASH), fibrosis and cirrhosis, which is a predisposing factor for hepatocellular carcinoma and liver failure (9–11). The hallmark of NAFLD is accumulation of triglycerides within hepatocytes, commonly referred to as hepatic steatosis (12).

Non-targeted liver biopsy with semi-quantitative histologic steatosis grading is the current reference standard for the diagnosis and grading of NAFLD (13, 14). Unfortunately, both the biopsy itself and the histologic grading of steatosis have several limitations. Liver biopsy is relatively expensive, and invasive, with risks that include pain, bleeding, and, very rarely, death (15). Perhaps more importantly, biopsy suffers from high sampling variability (16, 17) and semi-quantitative grading of hepatic steatosis is strongly observer-dependent (18, 19).

Advanced chemical shift-encoded water-fat separation MRI methods that estimate the proton density fat-fraction (PDFF) (20) are emerging as a promising non-invasive biomarker that can accurately map triglyceride concentration throughout the entire liver (21–26). MRI-PDFF has demonstrated equivalency with magnetic resonance spectroscopy–PDFF (MRS-PDFF) (21–24) and has demonstrated good correlation with histologic steatosis grading (27–29). Further, a recent longitudinal study has demonstrated that MRI-PDFF is more sensitive than histologic steatosis grading for the quantification of treatment-induced changes to hepatic steatosis (30).

However, in previous human *in vivo* studies the underlying hepatic triglyceride content remained unknown, and discrepancies between MRI-PDFF, MRS-PDFF and histologic

steatosis grading remained unresolved. Another shortcoming of previous comparative *in vivo* studies is that only a single non-targeted biopsy core from a random location throughout the entire liver was validated. Moreover, even within the validated samples, the true severity of steatosis remains unknown, since no direct comparison of MRI-PDFF with absolute triglyceride content has been performed in human livers.

Comprehensive and accurate assessment of triglyceride content within the liver requires multiple tissue samples from all liver segments with subsequent biochemical triglyceride extraction. The comparison of MRI-PDFF and triglyceride extraction has been performed up to now only in small animal models of steatosis (31, 32). Obviously, such an approach is not possible in living human subjects. However, such studies are needed to understand the relationship of PDFF derived from MRI and MRS relative to histologic steatosis grading and absolute triglyceride accumulation within the liver.

The purpose of this *ex vivo* study was to validate the technical accuracy of quantitative MRI-PDFF as a non-invasive imaging biomarker of hepatic steatosis in explanted human livers using MRS-PDFF, tissue triglyceride concentration and histology as reference standards. In addition, the variability of histologic steatosis grading was evaluated by assessment of intra- and interobserver agreement and by repeatability to determine the precision of biopsy, relative to quantitative MRI-PDFF.

EXPERIMENTAL PROCEDURES

Liver Specimen

Our Institutional Review Board provided a waiver for this prospective HIPAA compliant study. Human livers from cadaveric organ donors that were deemed unsuitable for transplantation and authorized for medical research were obtained either through our university hospital [blinded for review], the International Institute for the Advancement of Medicine, IIAM, (Edison, NJ), or the National Disease Research Interchange organization, NDRI (Philadelphia, PA). Per standard protocol, the identity of all donors was blinded to our institution.

Experimental Setup

Livers were preserved in University of Wisconsin (UW) solution by the providers and delivered cooled on ice as is standard practice for regular transplant organs. Each of the nine liver segments was labeled on the liver surface with an MR visible marker (vitamin D capsule) wrapped in surgical tape, using surgical suture (Fig. 1A). Photographs of entire livers with placed markers were taken for documentation. Labeling was performed to allow for precise co-localization of MRI regions of interest (ROIs), MRS voxels, core biopsies for histology and tissue wedges for triglyceride analyses. Livers remained in bags with phosphate buffered saline (PBS) and were placed into a plastic container. Containers with livers were then transferred to the MR scanner room where MRI-PDFF and MRS-PDFF was performed as described below. Thereafter, core biopsies and tissue wedges from each of the nine liver segments were obtained as described below.

MRI-PDFF Assessment

MR imaging was performed on a clinical 1.5 Tesla whole-body MR system (Signa HDxt, GE Healthcare, Waukesha, WI) using a single channel quadrature head coil. A coronal high-resolution T1-weighted 3D spoiled gradient echo (SGRE) sequence was acquired for planning of precise co-localization of MRS voxels with attached markers on the liver surface (Fig. 1B).

MRI-PDFF was performed using a 3D multi-echo chemical shift encoded SGRE sequence (22) with parameters equivalent to *in vivo* imaging, including 44×44cm² field of view, 256×160 matrix, 8mm slice thickness, 32 slices, 5° flip angle, ±125kHz receiver bandwidth, TR=13.6ms, and 6 echoes (TE_{init}=1.20ms, TE=1.98ms). Separated fat and water images were reconstructed using a graph-cut algorithm to avoid water-fat swapping (33), as well as spectral modeling of fat, and T2* correction (34, 35). Eddy current-related phase errors were addressed by using a mixed magnitude/complex fitting technique (36). Temperature-related resonance frequency shifts between water and fat were addressed by using a temperature-corrected fat-water reconstruction with offset 3.695 ppm between the water and the main methylene fat peak, as calibrated from the MRS spectra (37). MRI-PDFF imaging was repeated immediately following the first acquisition with a high-resolution acquisition to assess robustness and repeatability. Image parameters for the high resolution acquisition included: 32×24cm² field of view, 256×160 matrix, 3mm slice thickness, 52 slices, 5° flip angle, ±125kHz receiver bandwidth, TR=14.7ms, and 6 echoes (TE_{init}=1.06ms, TE=2.03ms).

MRI-PDFF maps were calculated from the separated water (W) and fat (F) images (PDFF=F/(W+F)) using a magnitude discrimination method (38), in order to remove the effects of B₁ coil-sensitivity (39). MRI-PDFF was measured from the calculated MRI-PDFF maps, with 2×2cm² regions of interest (ROI) placed in each of the nine liver Couinaud segments, placed ~1.5 cm beneath the liver surface using the attached markers for precise co-localization.

For assessment of intraobserver agreement, two ROI measurements were performed by the same radiologist ([blinded for review], 8 years of experience) with an interval of 4 weeks between the first and second measurement. For assessment of interobserver agreement, a second radiologist ([blinded for review], 15 years of experience) performed a third ROI measurement. For assessment of repeatability, the first radiologist ([blinded for review]) also performed ROI measurements on the second set of acquired MRI-PDFF maps. Radiologists were blinded to the results of MRS-PDFF, histology and triglyceride extraction.

MRS-PDFF Assessment

For MRS-PDFF, a single-voxel STEAM (Stimulated Echo Acquisition Mode) acquisition was performed without water suppression (40). A 2×2×2cm³ MRS voxel was placed in each of the nine liver segments beneath on the liver surface using the attached markers for exact co-localization while avoiding large vessels and bile ducts (Fig. 1B). Following a single pre-acquisition excitation, five spectra (TR=3500ms to avoid T₁-weighting) were acquired consecutively at progressively longer TEs of 10, 20, 30, 40 and 50ms, in order to enable

correction for T2 relaxation. A minimum mixing time (TM) of 5ms was chosen to minimize J-coupling effects (40). Frequency shifts between water and the main methylene fat peak were used to approximate the temperature of the liver specimen (37, 41). This was done only for livers (n=11) with an average of > 5% PDFF in order to obtain reliable estimates of the fat-water shift. A MR physicist ([blinded for review]) processed the multi-TE STEAM data automatically to provide an objective T2-corrected MRS-PDFF (42). The physicist was blinded to results of MRI-PDFF, histology, and triglyceride extraction.

Core Biopsies and Histological Assessment

Five core biopsies were taken with a 16-gauge biopsy gun (Max-Core, Bard Biopsy, Tempe, AZ) from each of the nine liver segments. The biopsies were taken in a fan-shaped distribution from the tissue located ~1.5 cm below the attached markers on the liver surface (Fig. 1C). This resulted in 45 cores from each liver. Cores were fixed in formalin and stained with hematoxylin and eosin (H&E).

Two liver pathologists at our institution assessed the degree of macrovesicular steatosis on digitized, randomized and blinded slides. The extent of macrosteatosis was evaluated by estimating the percentage of hepatocytes that contain intracellular macrovesicular fat droplets that displace the nucleus to the periphery of the cell (18, 19). In addition, the percentage of cells affected by fat vacuoles was separated by grade (14): grade 0 (<5%), grade 1 (5%–33%), grade 2 (>33%–66%), and grade 3 (>66%) for further statistical analyses.

For assessment of intraobserver agreement, one pathologist ([blinded for review], 3 years of experience) performed two blinded histologic steatosis ratings for the first of the five biopsies from each liver segment with an interval of 4 weeks between the first and second rating. For assessment of interobserver agreement, a second pathologist ([blinded for review], 3 years of experience) performed a third blinded histologic rating. For assessment of repeatability, the first pathologist ([blinded for review]) performed blinded histologic steatosis rating also on the second of the five biopsies from each liver segment. To demonstrate the variability of the results of bioptic sampling and histologic steatosis grading, the first pathologist ([blinded for review]) also rated the remaining blinded three cores from each liver segment. Readers were blinded to the experimental set up (i.e. readers were blinded to the fact that multiple biopsies were obtained from the same liver, and also that all biopsies were taken *ex vivo*) and to the results of MRI-PDFF, MRS-PDFF and triglyceride extraction.

Tissue Wedges and Triglyceride Assessment

One tissue wedge of ~1.5×1.5×1.0cm³ was excised ~1.5 cm beneath the markers on the liver surface from each segment using a scalpel, resulting in nine wedges from each liver (Fig. 1D). Tissue wedges were immediately placed on dry ice and shipped to AniLytics, Inc., Gaithersburg, MD for quantitative triglyceride extraction. Tissue wedges were weighed and diluted with 0.9% NaCl to achieve a concentration equivalent to 100 mg/ml. Wedges were sonicated to break down the tissue, well mixed and then filtered through a column filter (glass wool) to remove large particles. Samples were analyzed for triglyceride in the Roche

Hitachi 717 chemistry analyzer (Roche Diagnostics, Indianapolis, IN) using enzymatic triglyceride (GPO) reagent (Pointe Scientific Inc., Canton, MI). Results are expressed as a mass percent. The biochemists were blinded to all other results including MRI-PDFF, MRS-PDFF and histology.

Statistical Analysis

The accuracy of MRI-PDFF was assessed through linear regression with MRS-PDFF, triglyceride extraction and histological steatosis grading. Pearson's correlation coefficients (r) with 95% confidence intervals (CI) and significance levels (p-values) were computed to express the degree of linear association between measures. We defined correlation coefficients as "strong" if $r > 0.8$ and as "excellent" if $r > 0.9$.

Intra- and interobserver agreement as well as repeatability of MRI-PDFF and histologic grading were compared using Bland-Altman 95% limits of agreement analyses. The biases and variances derived from Bland-Altman analyses were compared with Wilcoxon signed rank test and with Fligner-Killen test, respectively.

Average histologic steatosis grade (Grade 0–3) was calculated for each liver from the grading of the 45 biopsies from each liver. To demonstrate the variability of the individual biopsies both within and between segments of the same liver, dot-plots were constructed for each liver and the number of over- or underestimated biopsies was calculated.

Statistical graphics and computations were performed by a statistician ([blinded for review]) in R (v3.0.2, R Foundation for Statistical Computing, Vienna, Austria). Differences were considered statistically significant if the 95% confidence intervals did not overlap or if p-values were less than 0.05.

RESULTS

Characteristics of Explanted Livers and Obtained Samples

Livers from thirteen consecutive donors (nine men, four women) were included. Mean donor age was 45.7 years (range, 23–69 years) and mean body mass index (BMI) was 40.2 kg/m² (range, 31–58 kg/m²). Two livers were healthy organs for which no suitable HLA-matched recipient could be identified; eleven livers were declined for transplantation due to the presence of hepatic steatosis as determined by biopsy (n=7), visual evaluation (n=2), ultrasound (n=1), or elevated aminotransferases (n=1). All organs reached our institution within 24 hours (mean, 12 hours; range, 5–24 hours) after portal vein and hepatic arterial cross clamping. MRI and MRS were successfully performed in all thirteen organs. MRS-derived temperature estimations indicated an average temperature of $11.1 \pm 5.0^{\circ}\text{C}$ (range 4.5°C – 19.1°C) over all livers. The variation within a given liver ranged from a minimum temperature range of 3°C (average $11.0 \pm 1.1^{\circ}\text{C}$) to a maximum temperature range of 10°C (average $9.7 \pm 4.7^{\circ}\text{C}$).

ROI-analyses of MRI-PDFF maps, MRS-PDFF analyses, triglyceride extraction and five core biopsies were successfully obtained for each of the nine liver segments of all thirteen

livers. This resulted in 117 data points for comparative assessment of MRI-PDFF with co-localized MRS-PDFF, histologic grading and triglyceride extraction.

Gross visual evaluation demonstrated the wide range of hepatic steatosis of the obtained explanted organs, ranging from healthy organs to organs with severe steatosis (Fig. 2A). These differences of hepatic steatosis were clearly observed on MRI-PDFF maps (Fig. 2B). Results of histologic grading, MRS-PDFF and triglyceride extraction were consistent with results of MRI-PDFF (Fig. 2C). Overall, the average MRI-PDFF, MRS-PDFF, histologic steatosis grade and triglyceride content were $11.3 \pm 8.0\%$ (range, 2–41%), 11.1 ± 7.7 (range, 2–39%), $23.5 \pm 19.3\%$ (range, 0–90%), and 15.2 ± 8.8 (range, 3–43%), respectively.

Further, the hepatic triglyceride content varied not only between livers, but also within livers. MRI-PDFF maps often depicted a heterogeneous distribution of hepatic steatosis (Fig. 3A). This heterogeneous triglyceride deposition was confirmed with the nine MRS spectra (Fig. 3B), nine biopsies (Fig. 3C) and nine tissue wedges for triglyceride analyses (not shown) in all Couinaud segments of the liver. In this example, volumetric MRI-PDFF and comprehensive sampling of the individual nine liver segments using either MRS-PDFF, histology or triglyceride extraction revealed that the steatosis degree in the right hepatic lobe is ~50% higher as compared to left hepatic lobe. For example, MRI-PDFF ranged from 20% in liver segment VII vs. 10% in liver segment IVa. These differences of liver segment VII vs. IVa were also observed using MRS (20 vs. 12%), histologic grading (20% vs. 10%) and confirmed by biochemical triglyceride extraction (17% vs. 7%).

Correlation of MRI-PDFF with MRS-PDFF, Histology and Triglycerides

Overall, MRI-PDFF showed an excellent correlation with MRS-PDFF ($r=0.984$; CI: 0.978–0.989, $p<0.001$) and a strong correlation with histology ($r=0.850$; CI: 0.791–0.894, $p<0.001$) and extracted triglycerides ($r=0.871$; CI: 0.818–0.909, $p<0.001$) (Fig. 4).

Interobserver and Intraobserver Agreement of MRI-PDFF and Histology

Bland-Altman analyses revealed a significantly smaller interobserver bias for repeated ROI measurements of MRI-PDFF maps compared to repeated histological steatosis grading (-0.7% vs. -15.8% , $p < 0.01$) (Fig. 5A). MRI-PDFF measurements also provided a significantly smaller variance as compared to histological steatosis grading as demonstrated by narrower 95% limits of agreement at Bland-Altman analyses (23.1% vs. 103.5%, $p < 0.001$). It should be noted that the bias and variance are expressed as percentages of the errors and not as absolute errors of the PDFF values and histologic steatosis grades, respectively. This representation has been chosen to facilitate direct comparison of the different metrics since the ranges are different for both metrics.

Similar results were found for intraobserver agreement (Fig. 5B). MRI-PDFF ROI measurements had a smaller bias compared to histologic steatosis grading (0.9% vs. 9.7%), however without being statistically significant ($p = 0.35$). Intraobserver MRI-PDFF measurements had a significantly smaller variance compared to histological steatosis grading as demonstrated by 95% limits of agreement at Bland-Altman analyses (19.1% vs. 99.9%, $p < 0.001$).

Repeatability of MRI-PDFF and Histology (Two Co-localized Biopsies)

Bland-Altman analyses revealed a smaller bias for repeated MRI-PDFF acquisition compared to repeated biopsy sampling (1.0% vs. -6.1%), however without being statistically significant ($p = 0.86$) (Fig. 6). Repeated MRI-PDFF acquisition provided a significantly smaller variance as compared to repeated biopsy sampling as demonstrated by narrower 95% limits of agreement seen with Bland-Altman analyses (10.9% vs. 109.9%, $p < 0.001$).

Variability of Histology

From each liver segment in all 13 livers ($n=117$) five biopsies were obtained, resulting in a total of 585 biopsies. From each of the 45 biopsies per liver the average percentage of cells affected by fat vacuoles and the steatosis grade (0–3) for each liver was calculated (Table 1). Based on the calculated average, two livers were grade 0 (average steatosis: 0.5%–1.7%), nine livers were grade 1 (average steatosis: 9.7% – 22.1%), two livers were grade 2 (average steatosis: 54.0% – 64.7%), and none of the livers were grade 3. High variability was observed between biopsies both within the same liver segment and between liver segments (Fig. 7). The variability was especially pronounced in the livers with high histological steatosis. Figure 7 illustrates not only the random sampling variability of biopsy, but also the natural variability of steatosis over the different liver segments. Thirty-seven of the 585 biopsies (6.3%) over- or underestimated the average degree of steatosis and consequently to a degree that resulted in a false categorization of the steatosis grade (Table 1).

DISCUSSION

In this study we have validated the technical performance of MRI-PDFF as a non-invasive imaging biomarker of hepatic steatosis in explanted human livers. Specifically, we have demonstrated excellent correlation with MRS-PDFF, tissue triglyceride concentration and histology that were used as three independent reference standards. Moreover, we demonstrated that MRI-PDFF had higher intra- and interobserver agreement and higher precision (repeatability) than histological steatosis grading. Co-localized sampling of multiple biopsies from the same liver segment also demonstrated high sampling variability. Overall, the results of this study demonstrate that MRI-PDFF is technically, an accurate and precise method to quantify fat in human liver tissue.

Previous *in vivo* studies have compared MRI-PDFF with MRS-PDFF or histological evaluation from a single biopsy. These studies showed excellent correlation of MRI with MRS (21–25, 28, 43) and good correlation with histology (27, 29, 44), albeit with varying results. This may be partly due to the fact that MRI ROI-measurements, MRS-voxels and locations of histological sampling were not exactly co-localized or time delays between MRI and biopsy may have introduced some degree of bias. More importantly, in all these studies the actual hepatic triglyceride content remained unknown and discrepancies between MRI-PDFF and histological results remained unresolved.

Our *ex vivo* study with MR-visible markers on the liver surface allowed for excellent co-localization of sampling sites for our comparison of MRI with MRS, histology and

triglyceride extraction. Our *ex vivo* results confirmed the excellent agreement of MRI-PDFF with MRS-PDFF and the good correlation with histology that has been observed *in vivo*. Importantly, we also found a good correlation of MRI-PDFF with extracted triglycerides. These results in *ex vivo* human livers confirm previous results from studies in rodent models of steatosis (31, 32), further validating the MRI-based methods for accurate quantification of hepatic steatosis.

When evaluating hepatic steatosis it is important to note that different biomarkers such as histological grading, PDFF and tissue triglyceride concentration are measuring fundamentally different properties of tissue (20), albeit all expressed as percentages. Specifically, histological assessment estimates the percentage of cells containing intracellular vacuoles of fat (14). Mass tissue triglyceride concentration measures the weight of triglycerides per gram of liver tissue, and is often expressed as a percentage. PDFF, however, is defined as the quotient of the number of protons from mobile triglycerides and the total number of protons from mobile water and mobile triglycerides (20). Using confounder-corrected MRS or chemical-shift encoded MRI methods, the PDFF can be measured accurately (21–24). Although PDFF is a fundamental measure of tissue triglyceride concentration and correlates closely with tissue triglyceride concentration, PDFF is not equivalent to the mass fat-fraction (45, 46). Liver contains NMR invisible material that is not measured using MR-based methods (47). This leads to a good correlation but fundamentally different measure of triglyceride concentration using PDFF or triglyceride extraction.

Our *ex vivo* findings suggest that discrepancies between MRI-PDFF and histology in previous *in vivo* studies may not be caused by variability of MRI-PDFF, but more likely from variability of histological steatosis grading. Histological steatosis grading has inherently high variability due to the natural heterogeneity of steatosis, and the challenge of evaluating an organ with a small sample of tissue. Variability of histological grading is also related to the subjective nature of semi-quantitative rating, and high inter-observer variability (18). Further, even if the grading is accurately performed, it is obtained from a single biopsy (16, 17), which may not be representative of the prevailing average steatosis degree in a given liver (29).

Our results confirmed the high inter- and intraobserver variability of reader-dependent histological steatosis grading and demonstrated a significantly lower variability for ROI-based measurements from PDFF maps. The direct comparison of the agreement of two repeated MRI-PDFF scans and of two co-localized biopsies demonstrated the influence of the sampling variability of biopsies on the final histological result. The effect of the sampling variability was further confirmed by comparison of five co-localized biopsies from each of the 117 assessed liver segments (total biopsies = 585). Our results showed that: a) repeated samples from one liver segment can vary dramatically, and b) individual liver segments might have much higher or lower steatosis degrees than the prevailing average steatosis degree.

Our findings also support previous *in vivo* studies demonstrating that MRI-PDFF is more accurate, more precise, less reader dependent and more sensitive than histological steatosis

grading to detect small changes of hepatic triglyceride content (30). Our study therefore further supports the notion that MRI-PDFF may not only be useful for diagnosis but also for long-term follow-up and monitoring of therapeutic interventions.

Despite the excellent technical performance of MRI-PDFF, we do not believe that MRI can or should replace biopsy for the diagnosis of diffuse liver disease. Rather, MRI-PDFF may complement histological assessment for monitoring of NAFLD. Biopsy is and should remain the clinical reference standard for the diagnosis, grading, and staging of diffuse liver disease. Although cost and the invasive nature of biopsy are relatively minor limitations, the most important limitation of biopsy is its known sampling variability (16). The high variability between repeated measurements limits the ability of biopsy to quantify longitudinal changes in features of diffuse liver disease during an intervention such as weight loss or drug therapy (30). Clearly, biopsy plays a central diagnostic role given its ability to evaluate other features of diffuse liver disease such as inflammation, fibrosis, ballooning degeneration, and degree of microsteatosis (19). This characterization is beyond the capability of current radiological methods such as MRI. However, after an initial biopsy that is accompanied by an MRI-PDFF scan for baseline diagnosis, MRI-PDFF could be used for further follow-up and monitoring of hepatic steatosis without the need for repeated biopsy.

There are some strengths of our study that are worth noting. One of the strengths of this study was the use of a three-way comparison of MRI-PDFF with MRS-PDFF, histological steatosis grading, and extracted triglycerides. Second, the use of MR-visible markers allowed accurate co-localization of all four techniques. Third, the radiologists (MRI-PDFF), physicists (MRS-PDFF), pathologists (histologic steatosis grading), and biochemists (triglyceride analyses) were blinded to the results of the other assessments. Fourth, the use of a state-of-the-art confounder corrected methods to quantify MRI-PDFF and MRS-PDFF.

However, our study also had several limitations. First, all livers were scanned at low temperatures and at different time points only *ex vivo* after explantation. It is not known how the transport, delay and temperature of the tissue will affect the measurement of MRI-PDFF and MRS-PDFF. To minimize the effect of potential bias of different temperatures we have used a temperature insensitive fitting method to construct MRI-PDFF maps (37). However, different times from organ harvest and differences in organ temperatures may have introduced bias into our results that may not have been present for *in vivo* experiments with perfused livers at body temperature. Unfortunately, a comparison of *ex vivo* MRI-PDFF with *in vivo* MRI-PDFF of the same subject was not feasible. Nonetheless, such a comparison would have been desirable to elucidate the good, but not perfect correlation of MRI-PDFF and results of triglyceride extraction.

The vast majority of clinical biopsy procedures are performed using percutaneous core biopsies. For patients undergoing intraoperative biopsy, however, wedge biopsy from the liver surface is sometimes preferred. It is unknown whether there are any significant differences in histological interpretations from core and wedge biopsies, which is a question this study could have addressed. Unfortunately, wedge biopsies were not performed as part

of this study and consideration should be made to compare wedge and core biopsies in future studies using *ex vivo* tissue.

A potential weakness of the statistical analyses is that we considered the nine replicates within a liver as independent. Since one would expect correlation to be present we are aware that estimations of precision might be overoptimistic but nevertheless allow for direct comparison of the methods in our study.

In conclusion, using *ex vivo* human livers, here we show that MRI-PDFF is an accurate imaging biomarker of liver triglyceride content that is more precise and less reader dependent than histological steatosis grading. MRI-PDFF allows non-invasive fat quantification of the entire liver and its utility should be further validated in clinical studies aimed at improved diagnosis and treatment monitoring.

Acknowledgments

Financial Support: The authors acknowledge the support of the NIH (R01DK083380, R01DK088925, R01DK100651, K24DK102595). We also wish to thank GE Healthcare and Bracco Diagnostics for their support.

List of abbreviations

MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
PDFF	Proton Density Fat-Fraction
NAFLD	Non-Alcoholic Fatty Liver Disease
NMR	Nuclear Magnetic Resonance
TR	Time of Repetition
TE	Time to Echo

References

1. Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, Grundy SM, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology*. 2004; 40:1387–1395. [PubMed: 15565570]
2. Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, Flegal KM. Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA*. 2006; 295:1549–1555. [PubMed: 16595758]
3. Schwimmer JB. Definitive diagnosis and assessment of risk for nonalcoholic fatty liver disease in children and adolescents. *Semin Liver Dis*. 2007; 27:312–318. [PubMed: 17682977]
4. Yki-Jarvinen H. Fat in the liver and insulin resistance. *Ann Med*. 2005; 37:347–356. [PubMed: 16179270]
5. Adams LA, Lymp JF, St Sauver J, Sanderson SO, Lindor KD, Feldstein A, Angulo P. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology*. 2005; 129:113–121. [PubMed: 16012941]
6. Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology*. 1999; 116:1413–1419. [PubMed: 10348825]

7. Sanyal AJ, Banas C, Sargeant C, Luketic VA, Sterling RK, Stravitz RT, Shiffman ML, et al. Similarities and differences in outcomes of cirrhosis due to nonalcoholic steatohepatitis and hepatitis C. *Hepatology*. 2006; 43:682–689. [PubMed: 16502396]
8. Adams LA, Waters OR, Knuiman MW, Elliott RR, Olynyk JK. NAFLD as a risk factor for the development of diabetes and the metabolic syndrome: an eleven-year follow-up study. *Am J Gastroenterol*. 2009; 104:861–867. [PubMed: 19293782]
9. McCullough AJ. The clinical features, diagnosis and natural history of nonalcoholic fatty liver disease. *Clin Liver Dis*. 2004; 8:521–533. viii. [PubMed: 15331061]
10. Brunt EM. Nonalcoholic steatohepatitis. *Semin Liver Dis*. 2004; 24:3–20. [PubMed: 15085483]
11. Smedile A, Bugianesi E. Steatosis and hepatocellular carcinoma risk. *Eur Rev Med Pharmacol Sci*. 2005; 9:291–293. [PubMed: 16231592]
12. Choi SS, Diehl AM. Hepatic triglyceride synthesis and nonalcoholic fatty liver disease. *Curr Opin Lipidol*. 2008; 19:295–300. [PubMed: 18460922]
13. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, Charlton M, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology*. 2012; 55:2005–2023. [PubMed: 22488764]
14. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol*. 1999; 94:2467–2474. [PubMed: 10484010]
15. Bravo AA, Sheth SG, Chopra S. Liver biopsy. *N Engl J Med*. 2001; 344:495–500. [PubMed: 11172192]
16. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, Grimaldi A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology*. 2005; 128:1898–1906. [PubMed: 15940625]
17. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology*. 2003; 38:1449–1457. [PubMed: 14647056]
18. El-Badry AM, Breitenstein S, Jochum W, Washington K, Paradis V, Rubbia-Brandt L, Puhan MA, et al. Assessment of hepatic steatosis by expert pathologists: the end of a gold standard. *Ann Surg*. 2009; 250:691–697. [PubMed: 19806055]
19. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, Ferrell LD, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*. 2005; 41:1313–1321. [PubMed: 15915461]
20. 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]
21. Yokoo T, Bydder M, Hamilton G, Middleton MS, Gamst AC, Wolfson T, Hassanein 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]
22. Meisamy S, Hines CD, Hamilton G, Sirlin CB, McKenzie CA, Yu H, Brittain JH, 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]
23. Hines CD, Frydrychowicz A, Hamilton G, Tudorascu DL, Vigen KK, Yu H, McKenzie CA, et al. T(1) independent, T(2) (*) corrected chemical shift based fat-water separation with multi-peak fat spectral modeling is an accurate and precise measure of hepatic steatosis. *J Magn Reson Imaging*. 2011; 33:873–881. [PubMed: 21448952]
24. 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 Med*. 2014; 72:1353–1365. [PubMed: 24323332]
25. Kuhn JP, Hernando D, Mensel B, Kruger PC, Ittermann T, Mayerle J, Hosten N, et al. Quantitative chemical shift-encoded MRI is an accurate method to quantify hepatic steatosis. *J Magn Reson Imaging*. 2013

26. Yokoo T, Shiehorteza M, Hamilton G, Wolfson T, Schroeder ME, Middleton MS, Bydder M, et al. Estimation of hepatic proton-density fat fraction by using MR imaging at 3.0 T. *Radiology*. 2011; 258:749–759. [PubMed: 21212366]
27. Kuhn JP, Hernando D, Munoz del Rio A, Evert M, Kannengiesser S, Volzke H, Mensel B, et al. Effect of multipeak spectral modeling of fat for liver iron and fat quantification: correlation of biopsy with MR imaging results. *Radiology*. 2012; 265:133–142. [PubMed: 22923718]
28. Hwang I, Lee JM, Lee KB, Yoon JH, Kiefer B, Han JK, Choi BI. Hepatic Steatosis in Living Liver Donor Candidates: Preoperative Assessment by Using Breath-hold Triple-Echo MR Imaging and H MR Spectroscopy. *Radiology*. 2014; 130863
29. Noworolski SM, Lam MM, Merriman RB, Ferrell L, Qayyum A. Liver steatosis: concordance of MR imaging and MR spectroscopic data with histologic grade. *Radiology*. 2012; 264:88–96. [PubMed: 22723561]
30. Nouredin M, Lam J, Peterson MR, Middleton M, Hamilton G, Le TA, Bettencourt R, 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]
31. Hines CD, Yu H, Shimakawa A, McKenzie CA, Warner TF, Brittain JH, Reeder SB. Quantification of hepatic steatosis with 3-T MR imaging: validation in ob/ob mice. *Radiology*. 2010; 254:119–128. [PubMed: 20032146]
32. Hines CD, Agni R, Roen C, Rowland I, Hernando D, Bultman E, Horng D, et al. Validation of MRI biomarkers of hepatic steatosis in the presence of iron overload in the ob/ob mouse. *J Magn Reson Imaging*. 2012; 35:844–851. [PubMed: 22127834]
33. 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]
34. 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]
35. Yu H, McKenzie CA, Shimakawa A, Vu AT, Brau AC, Beatty PJ, Pineda AR, et al. Multiecho reconstruction for simultaneous water-fat decomposition and T2* estimation. *J Magn Reson Imaging*. 2007; 26:1153–1161. [PubMed: 17896369]
36. Hernando D, Hines CD, Yu H, Reeder SB. Addressing phase errors in fat-water imaging using a mixed magnitude/complex fitting method. *Magn Reson Med*. 2012; 67:638–644. [PubMed: 21713978]
37. Hernando D, Sharma SD, Kramer H, Reeder SB. On the confounding effect of temperature on chemical shift-encoded fat quantification. *Magn Reson Med*. 2014; 72:464–470. [PubMed: 24123362]
38. 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]
39. Reeder SB, Wen Z, Yu H, Pineda AR, Gold GE, Markl M, Pelc NJ. Multicoil Dixon chemical species separation with an iterative least-squares estimation method. *Magn Reson Med*. 2004; 51:35–45. [PubMed: 14705043]
40. Hamilton G, Middleton MS, Bydder M, Yokoo T, Schwimmer JB, Kono Y, Patton HM, et al. Effect of PRESS and STEAM sequences on magnetic resonance spectroscopic liver fat quantification. *J Magn Reson Imaging*. 2009; 30:145–152. [PubMed: 19557733]
41. Soher BJ, Wyatt C, Reeder SB, MacFall JR. Noninvasive temperature mapping with MRI using chemical shift water-fat separation. *Magn Reson Med*. 2010; 63:1238–1246. [PubMed: 20432295]
42. Hernando, D.; Artz, N.; Hamilton, G.; Roldan, A.; Reeder, SB. Fully automated processing of multi-echo spectroscopy data for liver fat quantification. *ISMRM; Milan, Italy*. 2014. p. 2884
43. Reeder SB, Robson PM, Yu H, Shimakawa A, Hines CD, McKenzie CA, Brittain JH. Quantification of hepatic steatosis with MRI: the effects of accurate fat spectral modeling. *J Magn Reson Imaging*. 2009; 29:1332–1339. [PubMed: 19472390]
44. Tang A, Desai A, Hamilton G, Wolfson T, Gamst A, Lam J, Clark L, 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]

45. Reeder, SHC.; Yu, H.; McKenzie, C.; Brittain, J. Relationship between proton-density fat-fraction, true fat concentration for in vivo fat quantification with magnetic resonance imaging; Proceedings of the 19th Annual Meeting of ISMRM; Montreal, Canada. 2011. (abstract 805)
46. Reeder, SHC.; Yu, H.; McKenzie, C.; Brittain, J. On the definition of fat-fraction for in vivo fat quantification with magnetic resonance imaging; Proceedings of the 17th Annual Meeting of ISMRM; Honolulu, Hawaii. 2009. (abstract 207)
47. Hu, HLY.; Nagy, T.; Goran, M.; Nayak, K. Quantification of absolute fat mass: a validation study between chemical-shift MRI and chemical analysis; Proceedings of the 20th Annual MEeting of ISMRM; Melbourne, Australia. 2012. (abstract 1266)

Fig. 1. Experimental setup

(A) Photograph of an explanted liver with MR-visible markers labeling each liver segment for co-localization of MRI-PDFF measurements, MRS-PDFF voxels, core biopsies and tissue wedges. (B) High-resolution T1-weighted MR imaging allowed for positioning of MRS voxels ~1.5 cm below the markers as illustrated for liver segment II. (C) Five core biopsies (red lines) were obtained from each liver segment in a fan-shape sampling pattern in the same location using the surface markers for exact co-localization. (D) One co-localized tissue wedge (red-hatched wedge) was excised from each segment in the same

location. Numbers indicate liver segments. Note the MR-visible markers on the MR images next to the indication of liver segments.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 2. Examples of a healthy liver vs. livers with moderate and severe steatosis

(A) Photographs demonstrate increased size and yellow hue with increasing steatosis (left to right). (B) MRI-PDFF enabled volumetric quantification of steatosis over the entire liver with exemplary results of 4.3%, 13.6% and 33.2% PDF in segment VIII. (C) Histologic grading of co-localized core biopsies revealed that 8%, 30% and 75% of cells were affected by steatosis, respectively. Co-localized MRS-PDFF revealed fat-fractions of 4%, 14% and 33%, respectively. Triglyceride mass fractions from co-localized tissue wedges were 8%, 14% and 39%, respectively.

Fig. 3. Example of heterogeneous steatosis

(A) MRI-PDFF maps demonstrate highly heterogeneous steatosis, ranging from 10% in liver segment IVa to 20% in segment VII. Asterisks indicate MR visible markers. (B) Co-localized MRS-PDFF in all nine liver segments confirmed the inhomogeneous steatosis, ranging from 12% in segment IVa to 20% in segment VII. (C) Histologic grading from co-localized biopsies in each segment confirmed the inhomogeneous distribution of steatosis, ranging from 10% to 25% of cells affected by steatosis. Results of triglyceride analyses from co-localized tissue wedges for liver segments I to VIII were 8%, 10%, 13%, 7%, 7%, 11%, 11%, 17%, and 14%, respectively.

Fig. 4. Correlation of MRI-PDFF with MRS-PDFF, histological steatosis grade and extracted triglycerides

(A) MRI-PDFF showed an excellent correlation with MRS-PDFF ($r=0.98$, $p<0.001$). (B–C) MRI-PDFF showed a strong correlation with histological steatosis grading ($r=0.85$, $p<0.001$) and extracted triglycerides ($r=0.87$, $p<0.001$). Each of the 13 livers is indicated by nine symbols representing sampling from each of their 9 liver segments. Livers are sorted by increasing average steatosis grading.

Fig. 5. Inter- and intraobserver agreement of MRI-PDFF and histological steatosis grading
(A) Bland-Altman analyses of interobserver agreement of MRI-PDFF measurements and histological steatosis grading revealed a higher variance and larger bias of histological steatosis grading. **(B)** Results of intraobserver agreement also demonstrated a higher variance and larger bias of histological grading as compared to MRI-PDFF. Dotted lines indicate upper and lower 95% limit of agreement, solid line indicates bias. It is important to note that the bias and variance on the y-axis are expressed as percentages of the errors and not as absolute errors of the PDFF values and histologic steatosis grades, respectively. This representation has been chosen to allow of direct visual comparison of the variances and biases of the two techniques that have different metrics and different ranges. Note that due to this representation, the differences appear higher for smaller values of PDFF and histological grading than for larger values.

Fig. 6. Repeatability of MRI-PDFF and biopsy sampling

Bland-Altman analyses of (A) repeated MRI-PDFF measurements and (B) repeated histological steatosis grading revealed a higher interobserver variance and larger bias of histological steatosis grading. Dotted lines indicate upper and lower 95% limit of agreement, solid line indicates bias. It is important to note that the bias and variance on the y-axis are expressed as percentages of the errors and not as absolute errors of the PDFF values and histologic steatosis grades, respectively. This representation has been chosen to allow of direct visual comparison of the variances and biases of the two techniques that have different metrics and different ranges.

Fig. 7. Variability of histological fat assessment within segments and between segments

Five cores were obtained from each of the 117 liver segments in 13 livers, resulting in a total of 585 biopsies. Dot plots demonstrate the high variability of the 5 biopsies within individual liver segments (scatter in y-direction of a given liver segment). Moreover, dot plots indicate that actual differences in steatosis grade between individual liver segments exist (differences in y-direction between different segments). Dotted lines indicate the borders between steatosis grades: grade 0 (<5%), grade 1 (5%–33%), grade 2 (>33%–66%),

grade 3 (>66%). Solid line indicates averaged histopathologic steatosis degree in each liver. Red dots indicate biopsies that over- or underestimated the average

Table 1

Results of a total of 585 liver biopsies from thirteen livers (45 biopsies per liver)

Histological grading			no. and (%) of biopsies in each category						
Liver	Mean $\pm$ SD	range	Grade 0	<5%	Grade 1 33%	5- 66%	Grade 2 >33%- Grade 3	>66%	no. and (%) of false graded biopsies
13	0.5 $\pm$ 0.6	0-2	45	(100)	0	0	0	(0)	0
6	1.7 $\pm$ 1.5	1-5	37	(82.2)	8	(17.8)	0	(0)	8
5	9.7 $\pm$ 3.5	1-15	1	(2.2)	44	(97.8)	0	(0)	1
11	12.8 $\pm$ 4.5	5-20	0	(0)	45	(100)	0	(0)	0
2	13.0 $\pm$ 6.0	1-25	1	(2.2)	44	(97.8)	0	(0)	1
10	14.3 $\pm$ 5.9	5-40	0	(0)	44	(97.8)	1	(2.2)	1
8	14.9 $\pm$ 4.8	5-25	0	(0)	45	(100)	0	(0)	0
4	18.8 $\pm$ 6.0	5-30	0	(0)	45	(100)	0	(0)	0
9	21.6 $\pm$ 5.3	10-40	0	(0)	43	(95.6)	2	(4.4)	2
1	21.9 $\pm$ 3.9	5-25	0	(0)	45	(100)	0	(0)	0
12	22.1 $\pm$ 3.5	15-30	0	(0)	45	(100)	0	(0)	0
7	54.0 $\pm$ 10.9	20-70	0	(0)	2	(4.4)	39	(86.7)	6
3	64.7 $\pm$ 6.2	50-75	0	(0)	0	(0)	27	(60.0)	18
	20.8 $\pm$ 4.8	0-75	84	(14.4)	410	(70.1)	69	(11.8)	37

* Data are presented as means with standard deviations (SD) and ranges.

† Absolute numbers of biopsies, data in parentheses represent percentages. Livers are sorted by increasing average steatosis grading.