This copy is for personal use only. To order printed copies, contact *reprints@rsna.org*

Radiology

Quantifying Abdominal Adipose Tissue and Thigh Muscle Volume and Hepatic Proton Density Fat Fraction: Repeatability and Accuracy of an MR Imaging–based, Semiautomated Analysis Method[1]

Michael S. Middleton, MD, PhD
William Haufe, BSc
Jonathan Hooker, BSc
Magnus Borga, PhD
Olof Dahlqvist Leinhard, PhD
Thobias Romu, MSc
Patrik Tunón, MSc
Gavin Hamilton, PhD
Tanya Wolfson, MA
Anthony Gamst, PhD
Rohit Loomba, MD, MHSc
Claude B. Sirlin, MD

Purpose: To determine the repeatability and accuracy of a commercially available magnetic resonance (MR) imaging–based, semiautomated method to quantify abdominal adipose tissue and thigh muscle volume and hepatic proton density fat fraction (PDFF).

Materials and Methods: This prospective study was institutional review board—approved and HIPAA compliant. All subjects provided written informed consent. Inclusion criteria were age of 18 years or older and willingness to participate. The exclusion criterion was contraindication to MR imaging. Three-dimensional T1-weighted dual-echo body-coil images were acquired three times. Source images were reconstructed to generate water and calibrated fat images. Abdominal adipose tissue and thigh muscle were segmented, and their volumes were estimated by using a semiautomated method and, as a reference standard, a manual method. Hepatic PDFF was estimated by using a confounder-corrected chemical shift–encoded MR imaging method with hybrid complex-magnitude reconstruction and, as a reference standard, MR spectroscopy. Tissue volume and hepatic PDFF intra- and interexamination repeatability were assessed by using intraclass correlation and coefficient of variation analysis. Tissue volume and hepatic PDFF accuracy were assessed by means of linear regression with the respective reference standards.

Results: Adipose and thigh muscle tissue volumes of 20 subjects (18 women; age range, 25–76 years; body mass index range, 19.3–43.9 kg/m^2) were estimated by using the semiautomated method. Intra- and interexamination intraclass correlation coefficients were 0.996–0.998 and coefficients of variation were 1.5%–3.6%. For hepatic MR imaging PDFF, intra- and interexamination intraclass correlation coefficients were greater than or equal to 0.994 and coefficients of variation were less than or equal to 7.3%. In the regression analyses of manual versus semiautomated volume and spectroscopy versus MR imaging, PDFF slopes and intercepts were close to the identity line, and correlations of determination at multivariate analysis (R^2) ranged from 0.744 to 0.994.

Conclusion: This MR imaging–based, semiautomated method provides high repeatability and accuracy for estimating abdominal adipose tissue and thigh muscle volumes and hepatic PDFF.

©RSNA, 2017

[1] From the Liver Imaging Group, Department of Radiology (M.S.M., W.H., J.H., G.H., C.B.S.), Computational and Applied Statistics Laboratory, San Diego Supercomputing Center (T.W., A.G.), and Department of Medicine, Division of Gastroenterology and Hepatology (R.L.), University of California, San Diego, 9500 Gilman Dr, MC 0888, San Diego, CA 92093-0888; Advanced MR Analytics AB, Linköping, Sweden (M.B., O.D.L., T.R., P.T.); and Center for Medical Image Science and Visualization (M.B., O.D.L., T.R.), Department of Biomedical Engineering (M.B., T.R.), and Department of Medicine and Health (O.D.L.), Linköping University, Linköping, Sweden. Received March 12, 2016; revision requested May 6 and received November 27; accepted November 30; final version accepted December 20. **Address correspondence to** M.S.M. (e-mail: *msm@ucsd.edu*).

Study supported by Pfizer and by the National Institutes of Health (R01 DK088925).

©RSNA, 2017

Because of the increasing prevalence of obesity and its metabolic manifestations, there is a growing need to assess body composition, including estimation of adipose tissue and skeletal muscle volumes. Dual-energy x-ray absorptiometry (1) can be used for this but requires the use of modeling assumptions to differentiate visceral adipose tissue (VAT) from subcutaneous adipose tissue (SCAT), which introduces errors in the quantification of these compartments. Computed tomography allows accurate estimation of body compartment volumes but necessitates an undesirable radiation dose for whole-body assessment. Magnetic resonance (MR) imaging is also used to estimate SCAT (2,3), VAT (2,3), and muscle volumes (4). Advantages of MR imaging are that it permits reliable separation of SCAT from VAT, uses no ionizing radiation, and provides "one-stop" quantification of adipose tissue volumes and hepatic fat content.

Historically, MR imaging methods for adipose tissue volume quantification relied on manual segmentation, which is laborious, expensive, available only in specialized centers, and thus impractical for many research studies and not feasible for clinical care. Because of the limitations of manual segmentation, automated and semiautomated methods have been developed for adipose (3,5–15) and muscle (16–18) tissue segmentation. However, these methods are limited by binary tissue classification, two-dimensional segmentation, or both. Binary classification of each voxel as either adipose or nonadipose tissue leads to underestimation at tissue interfaces, with voxels containing subvoxel fat structures misclassified as nonadipose (14,15). Two-dimensional segmentation methods do not use information from neighboring sections, depending entirely on the information and image quality available in single sections. This leads to segmentation errors, especially in the presence of artifacts or morphologically complex fat structures such as adipose tissue enveloping visceral organs. When binary classification and two-dimensional segmentation are combined, these problems are magnified. In addition, none of these automated and semiautomated methods incorporate hepatic fat quantification and thus do not address this important manifestation of obesity.

To alleviate these problems, a commercial semiautomated image analysis method has been developed (AMRA Profiler; Advanced MR Analytics, Linköping, Sweden [19–23]), that uses quantitative fat and water separation to integrate the total fat signal intensity in a compartment of interest, thereby avoiding partial volume effects at interfaces and within voxels containing subvoxel fat. In addition, this method involves a three-dimensional approach to improve the robustness of the segmentation. Finally, this method incorporates estimation of hepatic proton density fat fraction (PDFF) by including a confounder-corrected, chemical shift–encoded sequence through the liver.

The primary aim of this study was to assess the repeatability and accuracy of this three-dimensional quantitative fat imaging–based method to estimate abdominal SCAT, VAT, total abdominal adipose tissue (TAT, the sum of SCAT and VAT) and thigh muscle volumes by using manual segmentation as a reference. A secondary aim of this study was to confirm the repeatability and, by using MR spectroscopy as the reference, accuracy of the included hepatic PDFF estimation sequence.

Advance in Knowledge

- Abdominal adipose and thigh muscle tissue volumes estimated by using this method showed high repeatability, with intra- and interexamination intraclass correlation coefficients ranging from 0.996 to 0.998, coefficients of variation ranging from 1.5% to 3.6%, and high accuracy compared with reference manual analysis, with regression slopes and intercepts not different from the identity line ($P > .05$).

Implication for Patient Care

- Current and future clinical and drug development studies may benefit from this semiautomated tissue volume and hepatic proton density fat fraction assessment MR imaging analysis method because there are many clinical settings in which monitoring these measures and changes in them may be desired.

Materials and Methods

Collaboration and Financial Support

Financial support for this study was provided by Pfizer Pharmaceuticals (New York, NY). None of the authors were employees of or consultants for Pfizer Pharmaceuticals. In addition, Advanced MR Analytics, or AMRA, provided in-kind support by performing some of the analyses for the study in a blinded fashion. Four coauthors of this article are employees of AMRA. Coauthors who are not employees of or consultants for AMRA had control of inclusion of all data and information submitted for publication.

Published online before print
10.1148/radiol.2017160606 **Content code:** GI

Radiology 2017; 283:438–449

Abbreviations:
ICC = intraclass correlation coefficient
PDFF = proton density fat fraction
SCAT = subcutaneous adipose tissue
SD = standard deviation
TAT = total adipose tissue
VAT = visceral adipose tissue

Author contributions:
Guarantors of integrity of entire study, M.S.M., C.B.S.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, M.S.M., M.B., O.D.L., R.L.; clinical studies, M.S.M., O.D.L., G.H., R.L., C.B.S.; experimental studies, J.H., T.R., P.T., G.H., C.B.S.; statistical analysis, M.B., T.W., A.G.; and manuscript editing, M.S.M., W.H., J.H., M.B., O.D.L., G.H., T.W., A.G., R.L., C.B.S.

Conflicts of interest are listed at the end of this article.

Figure 1

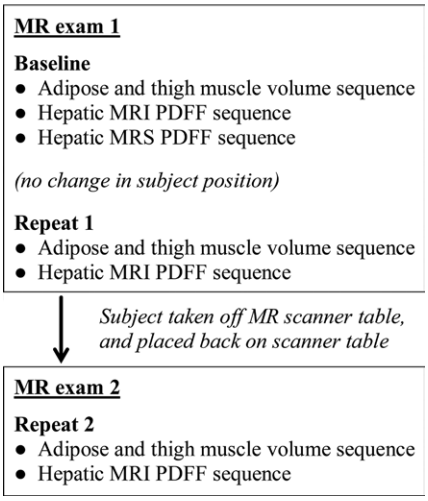

Figure 1: Schematic shows workflow for MR imaging examinations. MRS = MR spectroscopy.

Study Design

This was a single-site, prospective, cross-sectional, observational study approved by our institutional review board. Subjects were recruited at the University of California, San Diego, Nonalcoholic Fatty Liver Disease Translational Research Unit. Target enrollment was 20 subjects, based largely on feasibility. All study subjects were known or suspected to have nonalcoholic fatty liver disease. All subjects provided written informed consent and Health Insurance Portability and Accountability Act–compliant authorization for access to personal health information. Subject demographics were recorded, and body mass indexes were measured. Inclusion criteria were age of 18 years or older and willingness and ability to complete all research procedures. The exclusion criterion was contraindication to MR imaging.

MR Imaging Examination Workflow

A schematic of the imaging workflow is shown in Figure 1. Two unenhanced MR imaging examinations, each less than 20 minutes long, were performed per subject. To address the primary aim, in the first MR imaging examination the adipose tissue and thigh muscle volume sequence was performed twice (ie, baseline and repeat 1) in immediate succession without repositioning the subject on the table. Then, the subject was taken off the table, placed back on the table after a 10-minute interval, and the same sequence was performed again during this second MR imaging examination (ie, repeat 2). Hence, the adipose and thigh muscle volume sequence was performed a total of three times for each subject.

MR Image Acquisition

Subjects were imaged in the supine position with a 3-T MR imager (GE Signa EXCITE HDxt; GE Medical Systems, Milwaukee, Wis) by one of two MR imaging technologists (nonauthors, both with more than 10 years of experience). For any given subject, the MR imaging technologist was the same for both examinations. The body coil was used for all study sequences.

A commercially available water- and fat-separation sequence to estimate abdominal adipose tissue and thigh muscle volume was implemented as a dual-echo, T1-weighted, three-dimensional spoiled gradient-recalled-echo axial acquisition, with parameters listed in Table 1. Six overlapping image stacks were obtained in approximately 6 minutes from the base of the skull to the knees. Use of the body coil for all image stacks simplified image acquisition and ensured homogeneous signal intensity throughout acquired images. The top stack (base of skull to shoulders, 86 sections) and the bottom two stacks (thighs and knees, each 86 sections) were acquired during free-breathing sequences of approximately 29 seconds; the other three stacks (including the thorax, abdomen, and pelvis; each 50 sections) were acquired during separate breath-hold acquisitions of 18–27 seconds. To reduce total acquisition time, pre-imaging parameters were calibrated for the top stack and kept fixed for all subsequent stacks. In the rare instances when clipping (which manifests as uniformly high signal intensity in the background) was observed in a lower stack, the technologist repeated the stack acquisition with lower receiver gain.

To estimate hepatic MR imaging PDFF, axial images were acquired of the entire liver by using a multigradient-echo confounder-corrected chemical shift-encoded sequence with hybrid reconstruction known as iterative decomposition of water and fat with echo asymmetry and least-squares estimation, or IDEAL-IQ (24). Each acquisition produced water, fat, and PDFF map images. Parameters are listed in Table 1.

MR Spectroscopy Acquisition

Stimulated-echo acquisition mode, or STEAM, MR spectroscopic spectra of the liver were obtained to provide reference PDFFs for MR imaging PDFF estimates. Avoiding major vessels, bile ducts, and liver margins, a single $20 \times 20 \times 20$ mm^3 voxel was placed in the right hepatic lobe for the first MR spectroscopic acquisition of the first MR imaging examination. The MR spectroscopic voxel was shimmed automatically, and its location was overlaid on the corresponding axial localization images and saved for imaging-spectroscopy colocalization. Spectra were acquired by using parameters selected to minimize T1 weighting and to permit reliable T2 estimation while minimizing confounding effects of fat-peak J coupling (Table 1). No water, fat, or spatial saturation was applied. Signals recorded at eight array elements were combined by means of singular value decomposition and were saved for offline analysis (25).

Image Reconstruction and Composite Image Stack Creation

Acquired image stacks were reconstructed offline and converted to water- and fat-separated images by using a custom three-dimensional extension of the method described for phase-sensitive image reconstruction (26,27). The converted image stacks were automatically calibrated by using adipose tissue as internal signal intensity reference, where pure adipose tissue has unit value and absence of adipose tissue has zero value. This calibration allows for volumetric fat quantification while avoiding partial volume effects at interfaces by integrating the total fat signal intensity

Table 1

MR Imaging Examination Acquisition Parameters

Parameter	Sequence Type		
	MR Imaging Tissue Volume Sequence	MR Imaging PDFF	MR Spectroscopic PDFF
Sequence type	3D FGRE	2D FGRE	Proton STEAM
Coil type	Body	Torso phased array	Torso phased array
Field strength (T)	3	3	3
Repetition time (msec)	4.1–4.3	Approximately 7	3500
Echo time (msec)	Approximately 1.15, 2.3*	Approximately 1.0, 1.8, 2.6, 3.4, 4.2, 5.1†	10, 15, 20, 25, 30‡
Mixing time (msec)	NA	NA	5§
Flip angle (degrees)	10	3	90 (×3)
Section thickness (mm)	3.5	8	NA
Fractional echo sampling	0.8	0.8	NA
Receiver bandwidth (kHz)	± 125	± 125	5000
Base acquisition matrix	256 × 160	192 × 160	NA
Field of view (mm)	440–600‖	440–600‖	NA
Acceleration type	NA	2D ARC	NA
Acceleration factor	NA	2.2	NA
Duration (sec)	Approximately 18–27 for breath-hold acquisitions, approximately 29 otherwise	Approximately 21	Approximately 21
MR spectroscopy voxel size (mm^3)	NA	NA	20 × 20 × 20#
No. of presaturation pulses	default	default	one**

Note.—FGRE = fast gradient-recalled echo, 2D = two dimensional, 3D = three-dimensional, ARC = autocalibrating reconstruction for cartesian imaging, IDEAL-IQ = iterative decomposition of water and fat with echo asymmetry and least-squares estimation, NA = not applicable, STEAM = stimulated-echo acquisition mode

* For each acquisition, the two echoes were nominally out of phase and in phase.

† The exact repetition times and the exact echo times depend on the field of view, and hence, were not constant among all participants.

‡ This range of echo times permitted reliable T2 estimation while minimizing confounding effects of fat-peak J coupling.

§ To minimize both J coupling and T1 weighting.

‖ Adjusted to habitus and subject breath-hold capability.

Selected in right lobe of liver, away from vessels, bile ducts, liver edges, and any artifact.

** To balance T1 saturation on subsequent excitations.

within selected compartments (21,28). The calibrated image stacks were then merged into a single composite image stack by using the imager coordinates and a linear blending of the overlapping images to create a smooth, seamless transition between image stacks.

Semiautomated Image Segmentation, Image Selection, and Analysis

Each composite image stack (baseline and repeats 1 and 2) was segmented into its SCAT, VAT, TAT, and thigh muscle compartments, as described previously (19–23). This was done by using a three-dimensional, nonrigid, multiatlas segmentation method. Sixteen atlases, where the SCAT, VAT, and thigh muscle compartments had previously been labeled manually, were registered to the image volumes. The atlas library consisted of five women and 11 men with the following characteristics: mean age, 41.8 years (range, 28.0–54.9 years); mean weight, 88.3 kg (range, 45.0–134 kg); mean height, 176 cm (range, 151–192 cm); mean body mass index, 28.4 kg/m^2 (range, 19.7–46.4 kg/m^2). A voting scheme was then used to combine the 16 field labels into a three-dimensional segmentation of each compartment. SCAT was defined as adipose tissue outside the abdominal wall, while VAT was defined as all adipose tissue within the abdominal cavity; lipid-containing tissue within the abdominal wall was not included in either compartment. Thigh muscles were defined as gluteus, iliacus, adductor, hamstring, quadriceps femoris, and sartorius. Depending on the subject, a volume comprising 190–248 images was segmented for each repeated acquisition, giving a total of 570–744 segmented images per subject. After segmentation, an AMRA analyst (nonauthor, with 1.5 years of experience) reviewed the segmented volumes and manually corrected the segmentations where required on the basis of visual image assessment. To set the upper and lower bounds defining abdominal SCAT, the analyst identified the top of the T9 vertebral body and the top of the femoral head. Final three-dimensional segmentations of the abdominal SCAT, VAT, TAT, and thigh muscle compartments were generated, and the computed fat volumes within the bounded anatomic regions were recorded. Lean muscle volume was computed by subtracting the integrated fat volume under the muscle label mask from the volume of the mask (22,23). Total time for the

segmentation and analysis was less than 10 minutes per whole-body acquisition. Sample coronal, sagittal, and axial reconstructions of the segmented image stacks for each acquisition are shown in Figure 2.

A subset of 20 images obtained in the first acquisition (baseline) for each subject was selected for assessment of volume estimation accuracy. Of the 20 images, 15 were selected so that five covered the upper abdomen; five, the lower abdomen; and five, the pelvis; abdominal SCAT, VAT, and TAT volumes in these images were recorded. The remaining five images covered the thighs; thigh muscle volumes in these five images were recorded. The only consideration in selecting the 20 images were that images included the specified anatomic locations (ie, upper abdomen, lower abdomen, pelvis, and thigh muscles), and that major artifacts be avoided. These 20 sections also were segmented manually and analyzed as described in the sections that follow.

Manual Image Segmentation and Analysis

The 20 images were segmented manually by analysts from the University of California, San Diego (J.H. and W.H., each with more than 1 year of experience) by using a software package (Sliceomatic; Tomovision, Magog, Canada). VAT and SCAT compartments were segmented on the first 15 images, and the thigh muscles were segmented on the other five images (Fig 3). Volumes of these compartments were recorded and served as the reference standard for abdominal adipose tissue and thigh muscle volume accuracy analyses.

MR Imaging and MR Spectroscopic PDFF Analysis

Hepatic MR imaging PDFF was assessed for accuracy (baseline only) and intra- and interexamination repeatability (baseline and repeats 1 and 2) by placing a $20 \times 20 \times 20$ mm^3 region of interest on the water and fat MR images at a position corresponding to the (baseline) MR spectroscopic voxel location. The position of the region of interest was, if needed, slightly adjusted by the analyst to avoid regions of artifacts

Figure 2

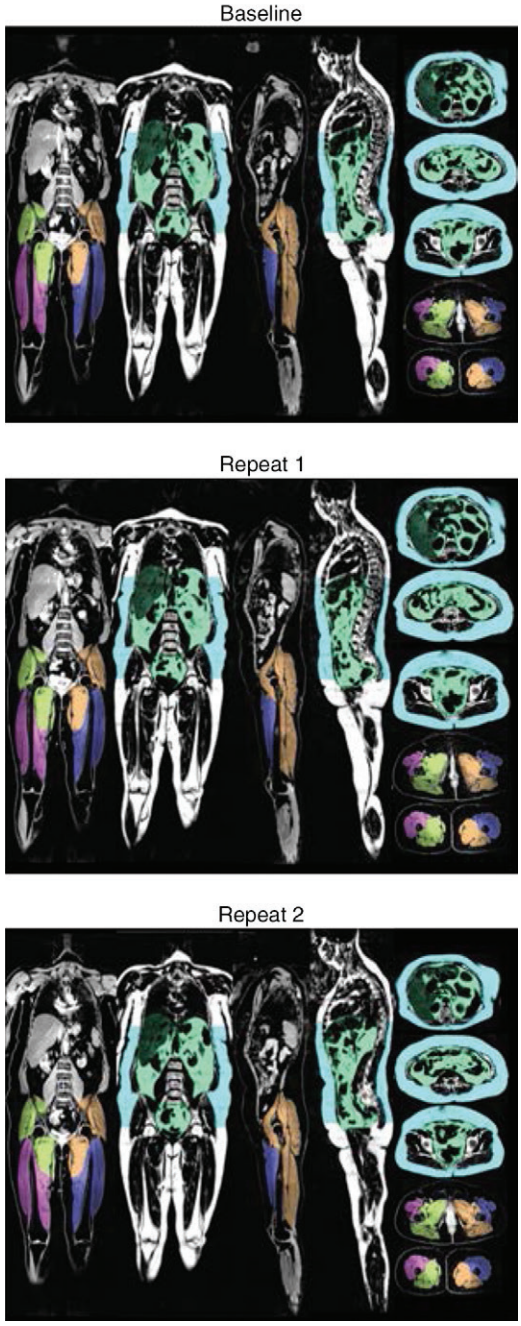

Figure 2: Coronal, sagittal, and axial reconstructions of segmented image stacks for each acquisition (baseline and repeats 1 and 2). SCAT (blue) volume was measured superiorly from T9 to top of the femoral head by using the semiautomated method. VAT (green) and thigh muscle volumes were measured in their entirety. Left anterior, left posterior, right anterior, and right posterior muscle groups are color coded for clarity.

Figure 3

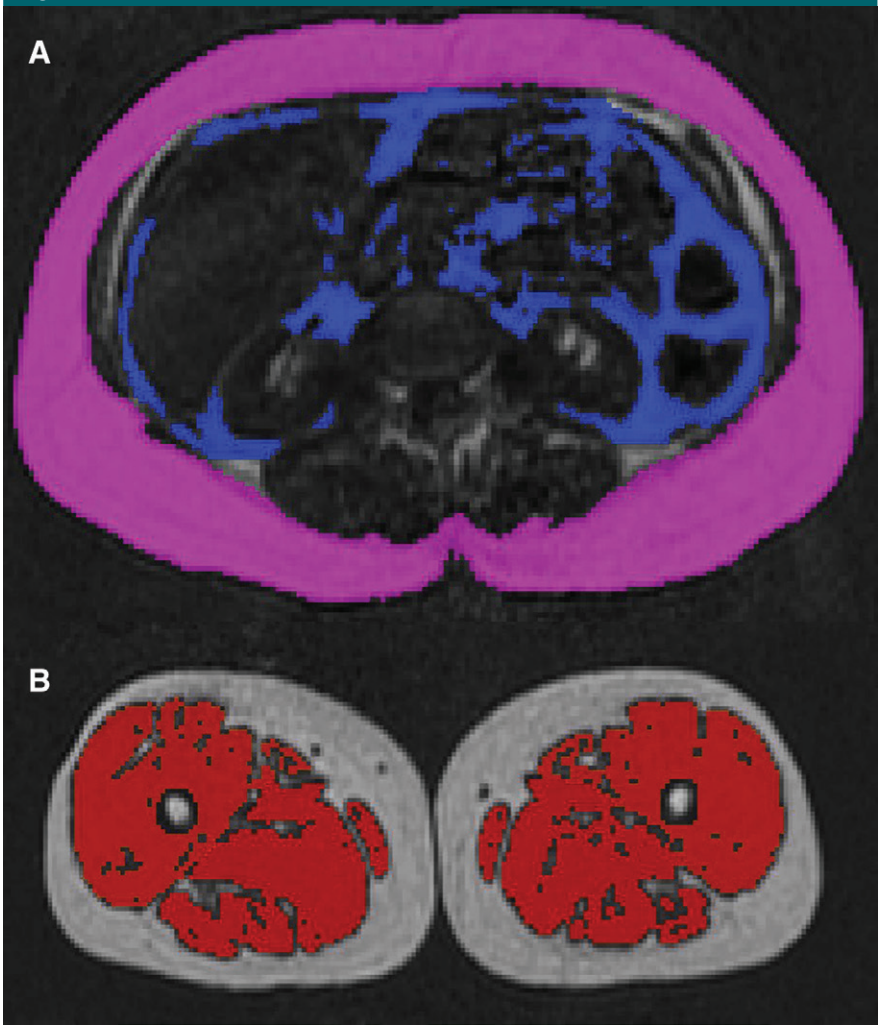

Figure 3: Axial MR images show manual segmentation, *A*, abdominal SCAT (pink) and VAT (purple); and, *B*, thigh muscle (red).

or vessels after inspection in a three-plane viewer. The mean value of pixels included in the region of interest was recorded as the estimated hepatic MR imaging PDFF value.

MR spectroscopic water (4.7 ppm) and fat (2.1, 1.3, 0.9 ppm) peaks were modeled as Gaussian resonances and measured. Non-linear least-square fitting was used to calculate spectral peak T2 values and T2-corrected peak areas. Fat peaks at 4.2 and 5.3 ppm, obscured by the water peak and not visible at in vivo field strengths in human liver were estimated from visible fat peaks on the basis of the known biochemical structure of human liver triglycerides (29). MR spectroscopic PDFF was calculated as the integrated sum of T2-corrected fat peaks divided by the sum of T2-corrected fat and water peaks.

Blinding

The AMRA analyst, who supervised the semiautomated whole-body segmentations, and the University of California, San Diego analysts, who performed the manual segmentations of the 20 selected sections, were blinded to each other's segmentations and results. Both sets of results were sent independently to a third author (M.S.M., with 17 years of experience), who then locked the study database before the statistical analyses were performed.

Statistical Analysis

Statistical analyses were performed by a staff statistician under the supervision of a faculty statistician (T.W. and A.G., respectively, both with more than 20 years of experience). Cohort population variables (age, sex, and body mass index) were summarized. For the 20 selected images for assessment of tissue volume accuracy, tissue volume estimates determined by using the semiautomated analysis method and reference tissue volume measurements determined by using the manual method were summarized.

Intra- (between baseline and repeat 1) and interexamination (between baseline and repeat 2) repeatability of abdominal SCAT, VAT, and TAT volume, and thigh muscle volume for the semiautomated analysis method were assessed by using the following metrics: the Bland-Altman bias (the mean of the differences between the two estimates), standard deviation (SD) of those differences, 95% limits of agreement (LOA = bias ± 1.96 SD, where LOA is limits of agreement), intraclass correlation coefficient (ICC) and its 95% confidence interval, and coefficient of variation for paired data (square root of the average of individual variances divided by the overall mean). Trends were assessed with correlation analysis and *P* values were computed.

Tissue volume accuracy for the semiautomated analysis method was assessed by using regression analysis. Each of the estimated tissue volume measures (abdominal SCAT, VAT, and TAT, and thigh muscle) obtained by means of the semiautomated analysis method was used as a single predictor in a univariate linear regression predicting its reference standard (corresponding manually determined volumes). From each regression, four metrics were obtained: the intercept of the regression line, the slope of the regression line, the average bias of the regression (defined as the square root of the average squared difference between

Table 2

Population and Cohort Characteristics

Variable	All Subjects	Women	Men
Sex*	20	18 (90)	2 (10)
Age (y)	50.1 ± 18.3 (25–76)	52.9 ± 17.1 (27–76)	25 ± 0 (25–25)
Body mass index (kg/m^2)	28.3 ± 7.2 (19.3–43.9)	27.4 ± 6.9 (19.3–43.9)	36.4 ± 4.2 (33.4–39.3)
Abdominal SCAT volume (T9 to femoral head) (cm^3)	9474 ± 5006 (3541–17 520)	8905 ± 4884 (3541–17 520)	14 600 ± 3 481 (12 400–17 060)
VAT volume (cm^3)	3190 ± 2130 (907–7183)	2952 ± 2106 (907–7183)	5329 ± 754 (4796–5863)
TAT volume (cm^3)	12 660 ± 6548 (4669–22 920)	11 860 ± 6322 (4669–22 870)	19 930 ± 4236 (16 930–22 920)
Thigh muscle volume (cm^3)	8562 ± 2067 (5725–12 110)	8377 ± 2097 (5725–12 110)	10 230 ± 517 (9860–10 590)
Hepatic PDFF (%)	7.1 ± 6.7 (1.8–27.9)	6.9 ± 7.0 (1.8–27.9)	8.9 ± 1.5 (7.8–10.0)

Note.—Unless otherwise indicated, data are mean ± SD, with the range in parentheses. Volume and PDFF measurements from the three acquisitions were averaged for each subject. PDFF was estimated by means of confounder-corrected, chemical shift-encoded MR imaging.

* Data are number of patients, with percentage in parentheses.

the regression line and the x = y identity line), and the R^2 of the regression model. Bootstrap-based confidence intervals were constructed around each of the accuracy metrics. Bootstrap-based P values for comparing the regression slope and intercept to the identity line slope and intercept were computed. Intra- and interexamination repeatability for hepatic MR imaging PDFF were assessed by using the same metrics as those described previously. Accuracy of the hepatic MR imaging PDFF relative to MR spectroscopic PDFF was assessed by using linear regression analysis as described previously.

Results

Twenty subjects (18 women and two men; mean age, 50.1 years [range, 25–76 years], mean body mass index, 28.3 kg/m^2 [range, 19.3–43.9 kg/m^2]) were recruited between July and November 2014. All recruited subjects met screening criteria and were enrolled in this study. Cohort characteristics are summarized in Table 2. Subjects ranged widely in anthropometric characteristics (body mass index, compartment volume) and hepatic PDFF.

Assessment of Volume Estimation Repeatability and Accuracy

Intra- and interexamination volume estimation repeatability based on all images in the selected composite image stacks are summarized in Table 3. Intra- and interexamination ICCs ranged from 0.996 to 0.998, and coefficients of variation ranged from 1.5% to 3.6%. Bland-Altman plots show that intra- and interexamination differences and 95% limits of agreement were

Table 3

Semiautomated Abdominal Adipose and Thigh Muscle Volume Intra- and Interexamination Repeatability

Measure	Intraexamination Values	Interexamination Values
Total abdominal SCAT volume		
Bias (cm^3)*	125 ± 295 (−453, 703)	74 ± 349 (−610, 758)
ICC	0.998 (0.994, 0.999)	0.998 (0.994, 0.999)
Coefficient of variation (%)	2.2	2.6
Mean (cm^3)	9984 ± 4947	9502 ± 5025
Total VAT volume		
Bias (cm^3)*	27 ± 158 (−283, 338)	13 ± 165 (−310, 337)
ICC	0.998 (0.993, 0.999)	0.997 (0.993, 0.999)
Coefficient of variation (%)	3.3	3.6
Mean (cm^3)	3320 ± 2204	3191 ± 2135
Total TAT volume		
Bias (cm^3)*	152 ± 370 (−573, 877)	88 ± 371 (−639, 815)
ICC	0.998 (0.994, 0.999)	0.998 (0.996, 0.999)
Coefficient of variation (%)	2.1	2.1
Mean (cm^3)	13 304 ± 6384	12 693 ± 6564
Total thigh muscle volume		
Bias (cm^3)*	63 ± 173 (−275, 401)	92 ± 155 (−212, 396)
ICC	0.997 (0.991, 0.999)	0.996 (0.988, 0.999)
Coefficient of variation (%)	1.5	1.5
Mean (cm^3)	8688 ± 2244	8568 ± 2087
Hepatic PDFF		
Bias (%)*	0.06 ± 0.65 (−1.23, 1.34)	−0.17 ± 0.74 (−1.62, 1.28)
ICC	0.995 (0.988, 0.998)	0.994 (0.986, 0.998)
Coefficient of variation (%)	6.5	7.3
Mean (%)	7.01 ± 6.67	7.20 ± 6.91

Note.—Unless otherwise indicated, data in parentheses are 95% confidence intervals. Means are the mean of the average of two intra- or interexamination measurement sets of values ± SD.

* Values for bias are ± SD, with 95% limits of agreement in parentheses.

Figure 4

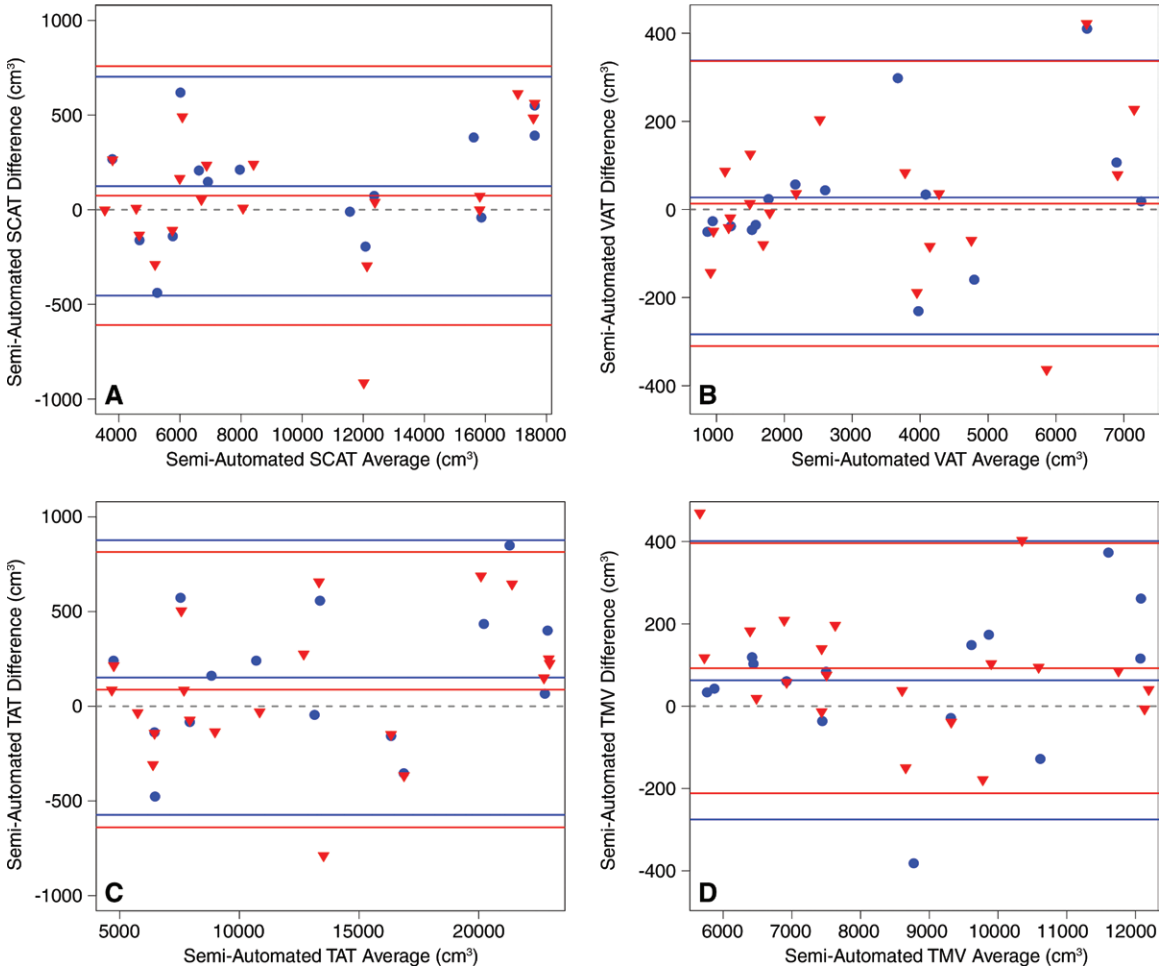

Figure 4: Bland-Altman plots for intraexamination (blue circle) and interexamination (red triangle) repeatability of semiautomated analysis method–determined volume estimates, respectively, for, *A*, abdominal SCAT; *B*, abdominal VAT; *C*, TAT; and, *D*, thigh muscle volume *(TMV)*. Intraexamination repeatability was evaluated by comparing baseline and repeat 1. Interexamination repeatability was evaluated by comparing baseline and repeat 3. Dashed gray line for each case represents zero bias. Central solid line is at bias level for each case, and upper and lower lines are 95% limits of agreement. Note that, for each compartment, intra- and interexamination results are similar. Bland-Altman metrics are summarized in Table 3.

small relative to measure volumes (Fig 4). There was no significant association between the difference and the average in any Bland-Altman intra- or interexamination plots (intraexamination, P = .233, .187, .250, and .306; interexamination, P = .304, .358, .217, and .210 for SCAT, VAT, TAT, and thigh muscle volumes, respectively).

Manually determined and semiautomated method–determined tissue volumes in the selected subsets of 20 images are summarized in Table 4. Regression scatterplots and all accuracy metrics are presented in Figure 5. Abdominal SCAT, VAT, TAT, and thigh muscle semiautomated estimated volumes for the 20 selected images for each subject were close to the corresponding manually determined reference measurements. Neither the slopes nor the intercepts of the regression lines were significantly different from those of the corresponding identity lines (slope, P = .33, .16, .29, and .27; intercept, P = .25, .62, .51, and .70; for SCAT, VAT, TAT, thigh muscle volumes, respectively). Regression bias was positive in all cases, ranging from 28.8 cm^3 to 100.5 cm^3, indicating slight underestimation according to results with the use of the semiautomated method compared with the manually determined reference measurements.

Assessment of PDFF Repeatability and Accuracy

Values for intra- and interexamination hepatic MR imaging PDFF estimation repeatability are summarized in Table 3 and presented in Figure 6, *A*. ICCs were 0.995 and 0.994, and coefficients

Table 4

Manually Determined and Semiautomated Method–determined Volume and PDFF Estimates (20 Selected Sections)

Parameter	Manually Determined Reference Data	Semiautomated Analysis Method–determined Data
Volume parameters		
Abdominal SCAT volume (cm^3)	1519.0 ± 696.5 (689.8–2,797.0)	1478.0 ± 689.3 (642.1–2756.0)
VAT volume (cm^3)	558.4 ± 354.5 (170.6–1398.0)	501.9 ± 321.8 (138.8–1164.0)
TAT volume (cm^3)	2078.0 ± 929.5 (921.7–3851.0)	1980.0 ± 899.1 (802.2–3411.0)
Thigh muscle volume (cm^3)	295.4 ± 86.2 (141.8–495.9)	321.4 ± 87.2 (224.0–568.4)
Hepatic PDFF (%)	7.04 ± 6.74 (1.57–28.12)	6.86 ± 6.49 (1.14–27.50)

Note.—Data shown are mean ± SD with ranges in parentheses. SCAT, VAT, and TAT volumes were calculated in 15 selected sections each, thigh muscle volumes were calculated in five selected sections.

Figure 5

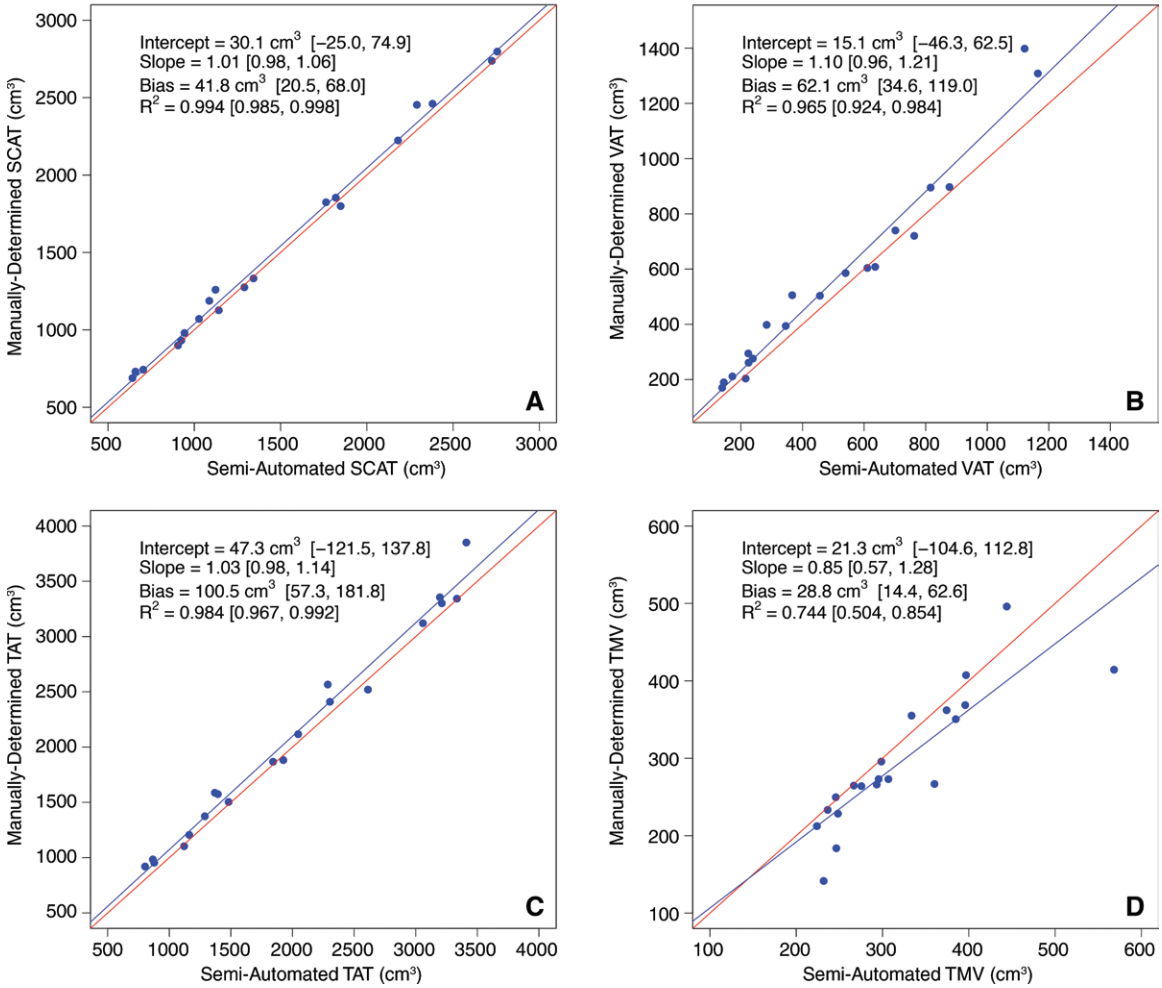

Figure 5: Regression plots of manually determined versus semiautomated tissue volumes for, *A*, SCAT; *B*, VAT; *C*, TAT; and *D*, thigh muscle volume *(TMV)*. Semiautomated estimates (predictors) are plotted on x-axes. Manually determined measurements (reference standards) are plotted on y-axes. Red line in each plot is identity line.

Figure 6

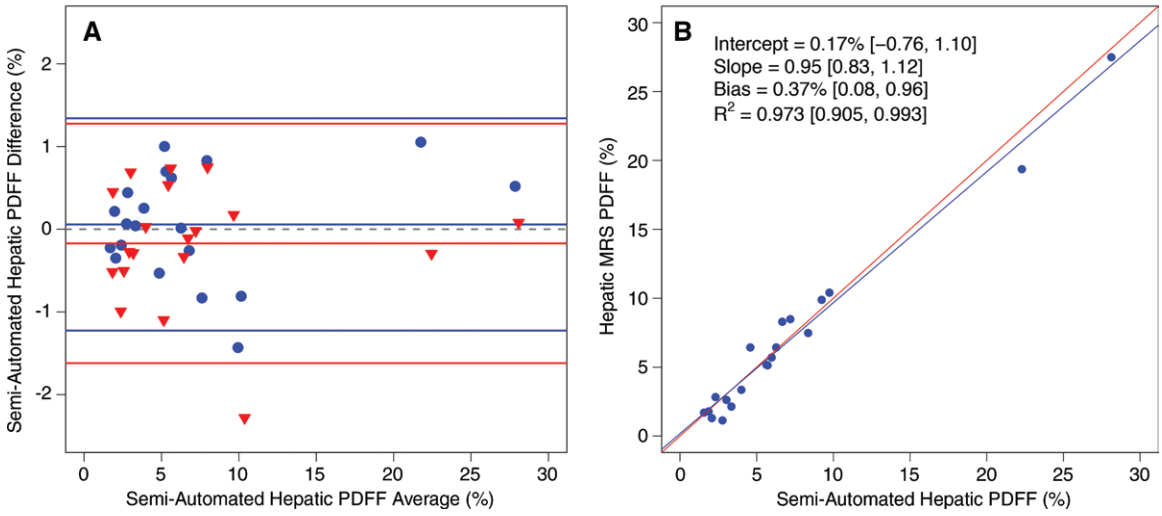

Figure 6: *A*, Bland-Altman plot for intraexamination (blue circles) and interexamination (red triangles) repeatability of semiautomated analysis method–estimated hepatic PDFF and, *B*, regression plot of MR spectroscopy versus semiautomated hepatic PDFF. Semiautomated PDFF (the predictor) is plotted on x-axis. MR spectroscopic PDFF (reference standard) is plotted on y-axis. Red line is identity line. Bland-Altman metrics are summarized in Table 3.

of variation were 6.5% and 7.3%, for intra- and interexamination repeatability, respectively. There was no significant association between the difference and the average in the Bland-Altman intraexamination (P = .388) or interexamination (P = .937) plot. Reference hepatic MR spectroscopic PDFF and semiautomated method–determined hepatic MR imaging PDFF values are summarized in Table 4. A regression scatterplot with all accuracy metrics is presented in Figure 6, *B*. Neither the slope nor the intercept of the regression line were significantly different from those of the identity line (slope, P = .27; intercept, P = .59).

Discussion

In this prospective, cross-sectional study of adults with variable anthropometric characteristics, a semiautomated MR imaging analysis method provides very good intra- and interexamination repeatability and accuracy of adipose and muscle tissue volumes and PDFF. Tissue volume intra- and interexamination ICCs and coefficients of variation ranged from 0.986 to 0.998 and 1.5%–3.6%, respectively. By using manual segmentation in a subset of images as the reference standard, neither the slope nor the intercept of the regression lines were significantly different from identity lines for any of the assessed measures. A consistent, small, and probably clinically nonmeaningful underestimation of tissue volumes with the semiautomated analysis method was noted relative to the manual method. The hepatic MR imaging PDFF measurements also were found to be repeatable and, with MR spectroscopic PDFF as the reference standard, accurate.

Authors of prior studies (2–18) have assessed abdominal SCAT, VAT, TAT, and thigh muscle volumes by using manual, semiautomated, and automated methods. However, most investigators who used prior methods have used two-dimensional segmentation, which can introduce quantification errors, and most have relied on signal intensity threshold–based binary classification, which can cause fat volume underestimation due to partial volume effects at tissue interfaces. Moreover, none have combined these tissue volumes and hepatic PDFF estimation in one package. In this study, volumes comprising up to 248 individual images were successfully segmented in less than 10 minutes, with included estimation of hepatic PDFF. The main implication of this study is that repeatability and accuracy of tissue volume estimation has now been assessed prospectively for a commercially available, semiautomated analysis method. Current and future clinical and drug development studies may benefit from this method, as may practitioners in clinical settings where monitoring change in these measures is desired. In addition, the high repeatability and accuracy of the hepatic MR imaging PDFF estimation sequence further supports ongoing efforts to qualify hepatic MR imaging PDFF as a biomarker of hepatic steatosis and helps to justify the incorporation of PDFF estimation in this semiautomated software package.

One limitation of this study was that only a single 3-T MR imager from a single manufacturer was used; thus, reproducibility in different types of imagers was not assessed in this study. However, Karlsson et al (20) previously reported good reproducibility between 1.5-T and 3-T imagers by using the proposed method for tissue compartment volume

measurements. In addition, Artz et al (30) reported good interfield-strength reproducibility for the incorporated PDFF estimation technique. The study sample consisted of 20 subjects, with the sample size determined according to feasibility rather than by means of an a priori power analysis. This was a limitation, because we potentially were underpowered to formally answer questions of accuracy. However, this sample size allowed an estimate of variability of the measures of interest, which in turn permitted preliminary assessment of performance. Nearly all subjects (18 of 20) were women. Recruitment was consecutive, and all potential subjects who were offered participation in this study were successfully enrolled; hence, there was no enrollment selection bias in the design, although there was an unintended majority of women included. The preferential inclusion of women in this study was unlikely to affect results, because subjects reflected a wide range of age, body mass index, and other anthropometric characteristics, and study aims were limited to repeatability and accuracy, and not to a clinical endpoint. An additional limitation was that all subjects were adults. This was unavoidable, because the recruitment was from an adult clinic specializing in treatment of fatty livers. Future research should include men, women, and children with geographic, racial, and ethnic and health style diversity to verify the robustness of this method.

Another limitation was that accuracy was examined in only a subset of nonrandomly selected images, but accuracy was not the major endpoint, and accuracy of the semiautomated analysis method has been assessed in prior studies (20,23). Although we did not address intra- and interoperator repeatability in our evaluation of the semiautomated analysis method, minimal operator supervision is required for its performance, and high intra- and interreader agreement have been reported. Results of a recent study (23) included intra- and interoperator coefficients of variation of 1.6% and 1.4% for VAT and 1.1% and 1.2% for abdominal SCAT, respectively.

In summary, this study showed that intra- and interexamination repeatability and accuracy with the use of a semiautomated analysis method and manual volume measurement as the reference standard, are very good. Future studies might address ways to further develop, qualify, and validate semiautomated tissue volumes as possible biomarkers of clinical endpoints. The results of this study also confirmed the expected high repeatability and accuracy of hepatic PDFF estimation with the use of a semiautomated analysis method.

Disclosures of Conflicts of Interest: M.S.M. Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: contracted work through university for Alexion, AstraZeneca, Bioclinica, Biomedical Systems, Bristol Myers Squibb, Galmed, GE, Genzyme, Gilead, Icon, Isis, Janssen, NuSirt, Pfizer, Profil, Sanofi, Shire, Siemens, Synageva, Takeda, and Viryoalscopics; grants from GE and Guerbet; patents pending, licensing, and royalties for University of California, San Diego. Other relationships: disclosed no relevant relationships. **W.H.** disclosed no relevant relationships. **J.H.** disclosed no relevant relationships. **M.B.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: stockholder of Advanced MR Analytics, patents issued. Other relationships: disclosed no relevant relationships. **O.D.L.** Activities related to the present article: employment with Advanced MR Analytics. Activities not related to the present article: stockholder for Advanced MR Analytics, patents issued. Other relationships: disclosed no relevant relationships. **T.R.** Activities related to the present article: work as project manager at Advanced MR Analytics. Activities not related to the present article: stockholder for Advanced MR Analytics. Other relationships: disclosed no relevant relationships. **P.T.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **G.H.** disclosed no relevant relationships. **T.W.** disclosed no relevant relationships. **A.G.** disclosed no relevant relationships. **R.L.** disclosed no relevant relationships. **C.B.S.** disclosed no relevant relationships.

References

1. Horan M, Gibney E, Molloy E, McAuliffe F. Methodologies to assess paediatric adiposity. Ir J Med Sci 2015;184(1):53–68.
2. Springer F, Ehehalt S, Sommer J, et al. Predicting volumes of metabolically important whole-body adipose tissue compartments in overweight and obese adolescents by different MRI approaches and anthropometry. Eur J Radiol 2012;81(7):1488–1494.
3. Hu HH, Chen J, Shen W. Segmentation and quantification of adipose tissue by magnetic resonance imaging. MAGMA 2016;29(2):259–276.
4. Valentin S, Yeates TD, Licka T, Elliott J. Inter-rater reliability of trunk muscle morphometric analysis. J Back Musculoskeletal Rehabil 2015;28(1):181–190.
5. Bonekamp S, Ghosh P, Crawford S, et al. Quantitative comparison and evaluation of software packages for assessment of abdominal adipose tissue distribution by magnetic resonance imaging. Int J Obes 2008;32(1):100–111.
6. Brennan DD, Whelan PF, Robinson K, et al. Rapid automated measurement of body fat distribution from whole-body MRI. AJR Am J Roentgenol 2005;185(2):418–423.
7. Kullberg J, Karlsson AK, Stokland E, Svensson PA, Dahlgren J. Adipose tissue distribution in children: automated quantification using water and fat MRI. J Magn Reson Imaging 2010;32(1):204–210.
8. Würslin C, Machann J, Rempp H, Claussen C, Yang B, Schick F. Topography mapping of whole body adipose tissue using a fully automated and standardized procedure. J Magn Reson Imaging 2010;31(2):430–439.
9. Müller HP, Raudies F, Unrath A, Neumann H, Ludolph AC, Kassubek J. Quantification of human body fat tissue percentage by MRI. NMR Biomed 2011;24(1):17–24.
10. Wald D, Teucher B, Dinkel J, et al. Automatic quantification of subcutaneous and visceral adipose tissue from whole-body magnetic resonance images suitable for large cohort studies. J Magn Reson Imaging 2012;36(6):1421–1434.
11. Poonawalla AH, Sjoberg BP, Rehm JL, et al. Adipose tissue MRI for quantitative measurement of central obesity. J Magn Reson Imaging 2013;37(3):707–716.
12. Thörmer G, Bertram HH, Garnov N, et al. Software for automated MRI-based quantification of abdominal fat and preliminary evaluation in morbidly obese patients. J Magn Reson Imaging 2013;37(5):1144–1150.
13. Addeman BT, Kutty S, Perkins TG, et al. Validation of volumetric and single-slice MRI adipose analysis using a novel fully automated segmentation method. J Magn Reson Imaging 2015;41(1):233–241.
14. Ludwig UA, Klausmann F, Baumann S, et al. Whole-body MRI-based fat quantifi-

cation: a comparison to air displacement plethysmography. J Magn Reson Imaging 2014;40(6):1437–1444.

15. Peng Q, McColl RW, Ding Y, Wang J, Chia JM, Weatherall PT. Automated method for accurate abdominal fat quantification on water-saturated magnetic resonance images. J Magn Reson Imaging 2007;26(3):738–746.
16. Broderick BJ, Dessus S, Grace PA, Olaighin G. Technique for the computation of lower leg muscle bulk from magnetic resonance images. Med Eng Phys 2010;32(8):926–933.
17. Brunner G, Nambi V, Yang E, et al. Automatic quantification of muscle volumes in magnetic resonance imaging scans of the lower extremities. Magn Reson Imaging 2011;29(8):1065–1075.
18. Baudin PY, Azzabou N, Carlier PG, Paragios N. Prior knowledge, random walks and human skeletal muscle segmentation. Med Image Comput Comput Assist Interv 2012;15(Pt 1):569–576.
19. Thomas MS, Newman D, Leinhard OD, et al. Test-retest reliability of automated whole body and compartmental muscle volume measurements on a wide bore 3T MR system. Eur Radiol 2014;24(9):2279–2291.
20. Karlsson A, Rosander J, Romu T, et al. Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. J Magn Reson Imaging 2015;41(6):1558–1569.
21. Andersson T, Romu T, Karlsson A, et al. Consistent intensity inhomogeneity correction in water-fat MRI. J Magn Reson Imaging 2015;42(2):468–476.
22. Dahlqvist Leinhard O, Johansson A, Rydell J, et al. Quantitative abdominal fat estimation using MRI. Presented at the 2008 19th International Conference on Pattern Recognition. New York, NY: IEEE, 2008; 1–4.
23. Borga M, Thomas EL, Romu T, et al. Validation of a fast method for quantification of intra-abdominal and subcutaneous adipose tissue for large-scale human studies. NMR Biomed 2015;28(12):1747–1753.
24. Meisamy S, Hines CD, Hamilton G, et al. Quantification of hepatic steatosis with T1-independent, T2-corrected MR imaging with spectral modeling of fat: blinded comparison with MR spectroscopy. Radiology 2011;258(3):767–775.
25. Bydder M, Hamilton G, Yokoo T, Sirlin CB. Optimal phased-array combination for spectroscopy. Magn Reson Imaging 2008;26(6):847–850.
26. Romu T, Dahlström N, Leinhard OD, Borga M. Robust water fat separated dual-echo MRI by phase-sensitive reconstruction. Magn Reson Med 2016 Oct 24. [Epub ahead of print]
27. Rydell J, Knutsson H, Pettersson J, et al. Phase sensitive reconstruction for water/fat separation in MR imaging using inverse gradient. Presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin, Germany: Springer, 2007; 210–218.
28. Peterson P, Romu T, Brorson H, Dahlqvist Leinhard O, Månsson S. Fat quantification in skeletal muscle using multigradient-echo imaging: Comparison of fat and water references. J Magn Reson Imaging 2016;43(1):203–212.
29. Hamilton G, Yokoo T, Bydder M, et al. In vivo characterization of the liver fat ^{1}H MR spectrum. NMR Biomed 2011;24(7):784–790.
30. Artz NS, Haufe WM, Hooker CA, et al. Reproducibility of MR-based liver fat quantification across field strength: Same-day comparison between 1.5T and 3T in obese subjects. J Magn Reson Imaging 2015;42(3):811–817.