Nonalcoholic Fatty Liver Disease: Diagnostic and Fat-Grading Accuracy of Low-Flip-Angle Multiecho Gradient-Recalled-Echo MR Imaging at 1.5 T[1]

Takeshi Yokoo, MD, PhD
Mark Bydder, PhD
Gavin Hamilton, PhD
Michael S. Middleton, MD, PhD
Anthony C. Gamst, PhD
Tanya Wolfson, MS
Tarek Hassanein, MD
Heather M. Patton, MD
Joel E. Lavine, MD, PhD
Jeffrey B. Schwimmer, MD
Claude B. Sirlin, MD

Purpose: To assess the accuracy of four fat quantification methods at low-flip-angle multiecho gradient-recalled-echo (GRE) magnetic resonance (MR) imaging in nonalcoholic fatty liver disease (NAFLD) by using MR spectroscopy as the reference standard.

Materials and Methods: In this institutional review board–approved, HIPAA-compliant prospective study, 110 subjects (29 with biopsy-confirmed NAFLD, 50 overweight and at risk for NAFLD, and 31 healthy volunteers) (mean age, 32.6 years ± 15.6 [standard deviation]; range, 8–66 years) gave informed consent and underwent MR spectroscopy and GRE MR imaging of the liver. Spectroscopy involved a long repetition time (to suppress T1 effects) and multiple echo times (to estimate T2 effects); the reference fat fraction (FF) was calculated from T2-corrected fat and water spectral peak areas. Imaging involved a low flip angle (to suppress T1 effects) and multiple echo times (to estimate T2* effects); imaging FF was calculated by using four analysis methods of progressive complexity: dual echo, triple echo, multiecho, and multiinterference. All methods except dual echo corrected for T2* effects. The multiinterference method corrected for multiple spectral interference effects of fat. For each method, the accuracy for diagnosis of fatty liver, as defined with a spectroscopic threshold, was assessed by estimating sensitivity and specificity; fat-grading accuracy was assessed by comparing imaging and spectroscopic FF values by using linear regression.

Results: Dual-echo, triple-echo, multiecho, and multiinterference methods had a sensitivity of 0.817, 0.967, 0.950, and 0.983 and a specificity of 1.000, 0.880, 1.000, and 0.880, respectively. On the basis of regression slope and intercept, the multiinterference (slope, 0.98; intercept, 0.91%) method had high fat-grading accuracy without statistically significant error ($P > .05$). Dual-echo (slope, 0.98; intercept, −2.90%), triple-echo (slope, 0.94; intercept, 1.42%), and multiecho (slope, 0.85; intercept, −0.15%) methods had statistically significant error ($P < .05$).

Conclusion: Relaxation- and interference-corrected fat quantification at low-flip-angle multiecho GRE MR imaging provides high diagnostic and fat-grading accuracy in NAFLD.

Supplemental material: *http://radiology.rsnajnls.org/cgi/content/full/2511080666/DC1*

[1] From the Departments of Radiology (T.Y., M.B., G.H., M.S.M., C.B.S.), Biostatistics (A.C.G., T.W.), Medicine (T.H., H.M.P.), and Pediatrics (J.E.L., J.B.S.), University of California, San Diego Medical Center, University of California at San Diego, MR3 Laboratory, 408 Dickinson St, San Diego, CA 92103-8226. Received April 12, 2008; revision requested June 16; revision received August 23; accepted September 26; final version accepted October 13. **Address correspondence to** C.B.S. (e-mail: *csirlin@ucsd.edu*).

Nonalcoholic fatty liver disease (NAFLD) is an emerging epidemic in Western countries and affects all ages and ethnicities (1,2). NAFLD can progress to nonalcoholic steatohepatitis, cirrhosis, and end-stage liver disease (3,4). Because early NAFLD may be reversible (5–10), screening and early intervention may be indicated.

Histologic visualization of hepatocellular fat vacuoles is the current standard for NAFLD assessment (3,11,12). However, biopsy is invasive and prone to sampling error (13–16). Furthermore, routine histologic examination is semiquantitative, observer dependent, and graded with broad severity brackets (3,11,12). A noninvasive, objective, and continuous-scale assessment may be preferable to biopsy in clinical practice and research.

Magnetic resonance (MR) spectroscopy directly enables measurement of fat and water proton signals and generally is considered the most accurate noninvasive technique for hepatic fat quantification (13,17,18). By using a long repetition time to minimize T1 effects and multiple echo times to correct for T2 effects, spectroscopy can help measure fat and water proton densities (PDs). As confirmed with biochemical assay of tissue specimens, fat fraction (FF) calculated from PDs determined at spectroscopy is equivalent to tissue triglyceride concentration (19,20). However, liver spectroscopy may be technically demanding and limited in spatial coverage and availability. By comparison, MR imaging is easier to perform, complete in liver coverage, and available at most centers. If comparable in accuracy, imaging would be an appealing practical alternative to spectroscopy.

A variety of imaging methods have been proposed for hepatic fat quantification. Previous human studies (21–26) have shown statistically significant correlations between imaging and spectroscopic FFs, but the accuracy of imaging relative to that of spectroscopy has been inconsistent. Imaging FF generally has been calculated from fat and water signal intensities (rather than PDs) and therefore has been affected by sequence type (23,27,28), imaging parameters (28–31), and tissue relaxation properties (32–34) due to confounding T1 and T2* (or T2) relaxation effects. The inconsistent accuracy of imaging compared with the accuracy of spectroscopy has limited the clinical utility of MR imaging as a primary diagnostic and monitoring tool (23).

To control the confounding relaxation effects, low-flip-angle multiecho gradient-recalled-echo (GRE) MR imaging has been proposed recently (33,35,36). A low flip angle suppresses T1 effects, and multiecho acquisition permits estimation and correction of T2* effects; imaging FF can then be calculated from fat and water PDs. Our purpose was to assess the accuracy of four fat quantification methods at low-flip-angle multiecho GRE MR imaging in NAFLD by using MR spectroscopy as the reference standard.

Advances in Knowledge

- Accurate hepatic fat quantification by using MR imaging requires suppression of T1 relaxation effects, correction of T2* relaxation effects, and correction of the interference effects between multiple spectral components.
- Correction of T2* effects improves diagnostic sensitivity for fatty liver disease, even in patients with no known hepatic iron-overload disorder.
- Correction of interference effects between multiple spectral components improves fat-grading accuracy over the full range of clinically relevant fat content.

Implications for Patient Care

- Low-flip-angle multiecho gradient-recalled-echo MR imaging with relaxation and interference correction is a rapid, safe, and highly accurate diagnostic and fat-grading modality for fatty liver disease.
- This study suggests that MR imaging can be used for initial diagnosis and long-term treatment of patients with fatty liver disease.

Materials and Methods

Study Design and Subjects

This prospective, cross-sectional, single-site clinical study was approved by an institutional review board and was compliant with the Health Insurance Portability and Accountability Act. Inclusion criteria were as follows: patients with biopsy-confirmed NAFLD, patients who were overweight and at risk for NAFLD (body mass index ≥ 25 kg/m^2), or healthy subjects of normal weight (body mass index < 25 kg/m^2). Subjects were recruited from the institutional hepatology and obesity clinics, as well as from the general public, and were enrolled consecutively. No formal

Published online before print
10.1148/radiol.2511080666

Radiology 2009; 251:67–76

Abbreviations:
FF = fat fraction
GRE = gradient-recalled echo
IP = in phase
NAFLD = nonalcoholic fatty liver disease
OP = opposed phase
PD = proton density
ROI = region of interest

Author contributions:
Guarantors of integrity of entire study, T.Y., 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; manuscript final version approval, all authors; literature research, T.Y., M.B., M.S.M., C.B.S.; clinical studies, T.Y., G.H., M.S.M., T.H., J.E.L., J.B.S., C.B.S.; statistical analysis, T.Y., M.S.M., A.C.G., T.W.; and manuscript editing, T.Y., M.B., G.H., M.S.M., A.C.G., T.W., T.H., H.M.P., J.E.L., C.B.S.

Funding:
This work was supported by a National Institute of Diabetes and Digestive and Kidney Diseases grant (no. R01 DK075128), a National Institute of Child Health and Human Development grant (no. U01DK061734), a San Diego EXPORT Center, National Center of Minority Health and Health Disparities grant (no. P60 MD00220), and a National Center for Research Resources for the General Clinical Research Center at the University of California, San Diego grant (no. M01 RR000827).

Authors stated no financial relationship to disclose.

power analysis was performed. Adult subjects gave informed consent. Pediatric subjects gave assent with parental informed consent. Subject age and sex were recorded. Other clinical data, including hepatic function panel results, were not recorded. One hundred fifteen subjects were enrolled between March and December 2007. Exclusion criteria were as follows: known history of other liver disease or contraindications to MR imaging (n = 0), inability to complete the research protocol (due to claustrophobia [n = 2] or due to imager malfunction [n = 2]), or spatial heterogeneity in fat distribution that made colocalization between imaging and single-voxel spectroscopy unreliable (n = 1). The remaining 110 subjects formed the study group. Twenty-nine subjects had biopsy-confirmed NAFLD, 50 were overweight (body mass index ≥ 25 kg/m^2), and 31 (healthy volunteers) were of normal weight. The demographic data are summarized in Table 1.

MR Imaging Examination

Subjects were examined in the supine position with a standard four-channel torso phased-array coil centered over the liver at 1.5 T (Symphony; Siemens Medical Systems, Erlangen, Germany).

Spectroscopy.—By using three-plane localizing images and avoiding vascular and biliary structures, a single 20 × 20 × 20-mm voxel was placed in a right hepatic segment, usually V or VIII. A single element of the array coil nearest to the voxel was selected. After shimming, stimulated-echo acquisition mode single-voxel proton spectroscopy was performed. To minimize T1 effects, repetition time was set at 3000 msec, and mixing time was set at 10 msec. To estimate T2, five single-average spectra were collected at echo times of 20, 30, 40, 50, and 60 msec in a single 15-second breath hold. Other parameters were 1000-Hz receiver bandwidth and 2048-point spectral resolution.

Imaging.—Multisection axial images were obtained by using a two-dimensional spoiled GRE sequence with all array coil elements. To minimize T1 effects (29,35,36), a low flip angle (10°) was used at a repetition time of 122 msec. To estimate fat-water signal interference and T2* effects (29,33,35), six echoes were obtained at serial opposed-phase (OP) and in-phase (IP) echo times (2.3, 4.6, 6.9, 9.2, 11.5, and 13.8 msec) during a single breath hold of 18 seconds or less. Other imaging parameters were 8-mm section thickness, 100% intersection gap, 500-Hz/pixel receiver bandwidth, one signal acquired, and rectangular field of view with a 256 × 160–256 matrix (adjusted to individual body habitus and breath-hold capacity).

To assess reproducibility, imaging and spectroscopy were performed again in 38 and 36 subjects, respectively, who agreed to a longer examination time.

Data Processing

A reference image was created by overlaying the spectroscopic voxel on the axial localizer image. Two trained technologists (both with 3 years of experience) reviewed MR images on a picture archiving and communication system workstation and, by using the saved overlay as a guide, manually placed a circular region of interest (ROI) of approximately 20 mm in diameter on one of the multiecho images at the spectroscopic voxel location. Picture archiving and communication system software automatically propagated the selected ROI to the rest of the multiecho images. The average ROI value at each echo time was recorded.

Data Analysis

Overview.—Fat and water PDs were estimated for spectroscopy and imaging (described below). FFs were calculated as the ratio of fat PD to total (fat and water) PD (37–39). T1 effects were assumed negligible with the above spectroscopic and imaging parameters; by using published fat and lean liver T1 values (40) in a computer simulation (36), the expected error upper limit due to residual T1 effects was 0.02% or less at spectroscopy and 1% or less at imaging.

Spectroscopy.—An MR physicist (G.H., 7 years of experience), who was blinded to imaging results, analyzed the spectra by using Advanced Method for Accurate, Robust, and Efficient Spectral fitting included in Java-based Magnetic Resonance User Interface (*http://sermn02.uab.es/mrui*) (41,42). At each echo time, the water (4.7 ppm) and fat (0.5–3 ppm) peak areas were calculated by using integration. The total fat peak area was defined as the sum of individual fat peaks. Water and total fat PD values were corrected for T2 by using a log-linear least-square fitting algorithm.

Imaging.—An MR imaging research fellow (T.Y., 2 years of experience), who was blinded to spectroscopic results, analyzed the imaging ROI values by using software (MATLAB; MathWorks, Natick, Mass). Four different image analysis methods (see Appendix E1, *http://radiology.rsnajnls.org/cgi/content/full/2511080666/DC1*) of increasing complexity were tested, which generated four FF estimates per subject from a single set of multiecho images (Table 2). The first

Table 1

Demographic Data of Subjects Undergoing MR Imaging of the Liver

Group	No. of Subjects	Age (y)*
Total	110	32.6 ± 15.6 (8–66)
Male	63	30.7 ± 16.6 (8–66)
Female	47	35.2 ± 14.0 (8–61)
Adult	80 (40 M, 40 F)	40.0 ± 11.3 (20–66)
Pediatric	30 (23 M, 7 F)	12.9 ± 3.3 (8–18)
Known NAFLD†	29 (21 M, 8 F)	20.6 ± 12.9 (8–54)
Risk of NAFLD‡	50 (27 M, 23 F)	38.6 ± 13.4 (9–66)
Healthy volunteer§	31 (15 M, 16 F)	34.5 ± 14.0 (8–61)

* Data are means ± standard deviations, with ranges in parentheses.

† Confirmed by biopsy.

‡ Body mass index 25 kg/m^2 or higher.

§ Body mass index less than 25 kg/m^2.

three methods were based on previously proposed strategies (33,35,36); the fourth method was new. The last three methods corrected for T2*; the first did not. For illustrative purposes, FF and, if appropriate, T2* maps were generated by applying each method pixel by pixel. The four methods used to analyze imaging ROI data were as follows:

1. Conventional dual-echo method (31,43): T2* signal decay was neglected, and only the phase interference between the water and dominant fat (methylene, 1.3 ppm) moieties was modeled. From the first pair of OP and IP (2.3 and 4.6 msec) echoes, FF was calculated (see Equation 12 in Appendix E1, *http://radiology.rsnajnls.org/cgi/content/full/2511080666/DC1*).

2. Triple-echo method (29): The water-methylene interference model described above was extended by incorporating tissue T2* relaxation effect as a monoexponential decay. T2* was estimated from the first pair of IP echoes (4.6 and 9.2 msec) (see Equation 9 in Appendix E1, *http://radiology.rsnajnls.org/cgi/content/full/2511080666/DC1*). From the first OP-IP pair (2.3 and 4.6 msec) and the estimated T2* value, FF then was calculated (see Equation 10 in Appendix E1, *http://radiology.rsnajnls.org/cgi/content/full/2511080666/DC1*).

3. Multiecho method (31,33): On the basis of the methylene-water interference model with monoexponential signal decay, the tissue T2* and FF were calculated simultaneously (see Equation 8 in Appendix E1, *http://radiology.rsnajnls.org/cgi/content/full/2511080666/DC1*) by using log-linear least-square fitting of all six echoes (2.3–13.8 msec).

4. Multiinterference method (new method): The signal model described above was further extended by constructing the fat spectrum as a weighted sum of the three main fat peaks centered at 2.1, 1.3, and 0.9 ppm (44,45)

Table 2

Summary of Features of Image Analysis Methods for PD Estimation

Feature	Dual Echo	Triple and Multiecho	Multiinterference
T1 suppression	Yes	Yes	Yes
T2* correction	No	Yes	Yes
Fat spectral modeling	No	No	Yes

Figure 1

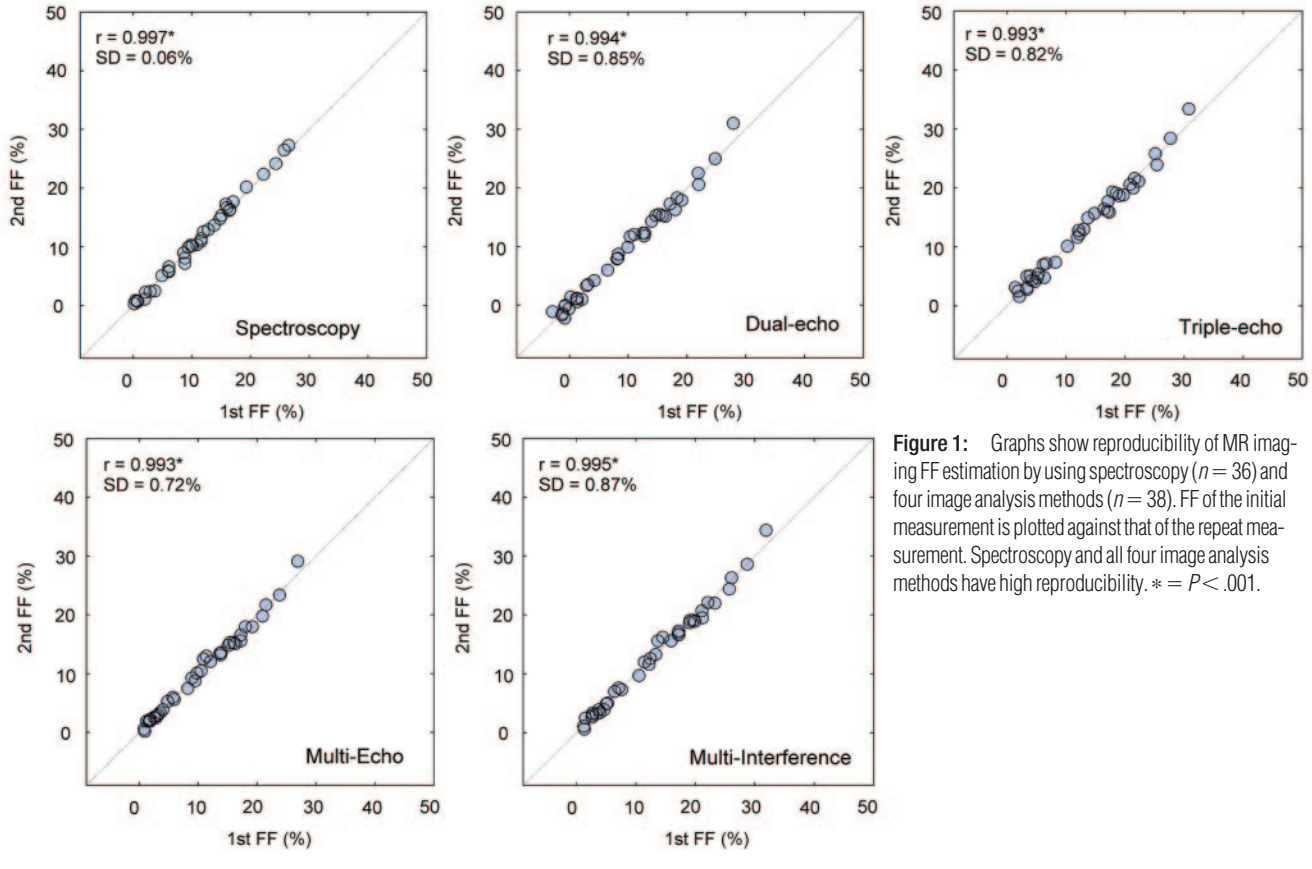

Figure 1: Graphs show reproducibility of MR imaging FF estimation by using spectroscopy ($n = 36$) and four image analysis methods ($n = 38$). FF of the initial measurement is plotted against that of the repeat measurement. Spectroscopy and all four image analysis methods have high reproducibility. $* = P < .001$.

Table 3

Diagnostic Sensitivity of Four Image Analysis Methods

Threshold	Dual Echo	Triple Echo	Multiecho	Multiinterference
5%	0.873 (55/63) [0.752, 0.926]	0.984 (62/63) [0.896, 0.986]	0.968 (61/63) [0.866, 0.985]	0.984 (62/63) [0.896, 0.986]
6%	0.820 (50/61) [0.694, 0.889]	0.984 (60/61) [0.879, 0.986]	0.934 (57/61) [0.832, 0.966]	0.984 (60/61) [0.879, 0.986]
6.25%	0.817 (49/60) [0.684, 0.888]	0.967 (58/60) [0.873, 0.984]	0.950 (57/60) [0.845, 0.971]	0.983 (59/60) [0.891, 0.985]
7%	0.759 (44/58) [0.599, 0.841]	1.000 (58/58) [0.977, 1.000]	0.914 (53/58) [0.810, 0.953]	1.000 (58/58) [0.977, 1.000]
8%	0.778 (42/54) [0.634, 0.869]	0.981 (53/54) [0.888, 0.983]	0.852 (46/54) [0.703, 0.909]	0.981 (53/54) [0.888, 0.983]

Note.—Numbers in parentheses were used to calculate sensitivities, and 95% confidence intervals are in brackets.

Table 4

Diagnostic Specificity of Four Image Analysis Methods

Threshold	Dual Echo	Triple Echo	Multiecho	Multiinterference
5%	1.000 (47/47) [0.973, 1.000]	0.872 (41/47) [0.728, 0.925]	0.936 (44/47) [0.808, 0.965]	0.851 (40/47) [0.704, 0.915]
6%	1.000 (49/49) [0.971, 1.000]	0.878 (43/49) [0.753, 0.935]	1.000 (49/49) [0.971, 1.000]	0.898 (44/49) [0.762, 0.946]
6.25%	1.000 (50/50) [0.976, 1.000]	0.880 (44/50) [0.753, 0.940]	1.000 (50/50) [0.976, 1.000]	0.880 (44/50) [0.752, 0.941]
7%	1.000 (52/52) [0.974, 1.000]	0.942 (49/52) [0.816, 0.966]	1.000 (52/52) [0.974, 1.000]	0.962 (50/52) [0.851, 0.982]
8%	1.000 (56/56) [0.979, 1.000]	0.964 (54/56) [0.854, 0.983]	1.000 (56/56) [0.979, 1.000]	0.929 (52/56) [0.815, 0.964]

Note.—Numbers in parentheses were used to calculate specificities, and 95% confidence intervals are in brackets.

(see Equation 7 in Appendix E1, *http://radiology.rsnajnls.org/cgi/content/full/2511080666/DC1*). The weighting was determined empirically from the spectroscopic data of subjects with an FF of more than 10%. (Subjects with sufficient fat content were selected to ensure reliable estimation of individual fat peak areas.) The normalized PDs of the respective fat peaks were as follows: 9.0% ± 2.6 (standard deviation), 77.9% ± 2.5, and 13.1% ± 2.7. By using all six echoes, the tissue T2* and FF were calculated by using a nonlinear least-square algorithm, lsqcurvefit (46).

Statistical Analysis

Reproducibility.—In subjects with repeat measurements, the two FF values from spectroscopy and from each image analysis method were plotted against each other. The standard deviation of the pairwise difference and correlation coefficient between the two measurements were calculated.

Diagnostic accuracy.—Performance statistics (sensitivity, specificity) were calculated for each image analysis method at a threshold value of 6.25% for spectroscopic FF. This value was equivalent (according to a conversion formula) to a 5.56% FF by wet weight, a previously suggested classification threshold for fatty liver (38,39). Two-sided 95% confidence intervals of the sensitivity and specificity estimates were calculated for each method, as well as the differences between each method pairs, as bias-corrected and accelerated intervals from the bootstrap distribution of 1000 resampled data sets (47). Because thresholds other than 6.25% also may be of interest, performance statistics were recalculated at threshold values of 5%, 6%, 7%, and 8%.

Grading accuracy.—Imaging and spectroscopic FF values were compared for each image analysis method by using linear regression. Statistically significant differences of the regression slopes and intercepts, with respect to null hypothesis of slope 1 and intercept 0, were determined by using two-tailed t tests at $\alpha = .05$. Because fat grading is clinically relevant only for diagnosed fatty liver, only subjects with a spectroscopic FF of more than 6.25% were included in this analysis.

Results

Subjects

The spectroscopic FF ranged from 0.2% to 34.5%, with a mean of 10.0% ± 9.1; 60 of 110 subjects had fatty liver according to the 6.25% spectroscopic criterion.

Reproducibility

High reproducibility of FF estimation was demonstrated in 36 and 38 subjects for spectroscopy and imaging, respectively (Fig 1). Correlation coefficients between measurements exceeded 0.99 for all methods ($P < .001$), and standard deviations of the differences between measurements were less than 1% in FF.

Diagnostic Accuracy

Tables 3 and 4 summarize the diagnostic performance of the four image analysis methods. At the diagnostic threshold of 6.25% for spectroscopic FF, the dual-echo method had the lowest sensitivity compared with all three T2*-corrected methods. The multiinterference spectral method had the highest sensitivity at 0.983, but differences in sensitivity between the three T2*-corrected methods were not significant (Table 5). The dual- and multiecho methods had a specificity of 1.000, which was significantly higher than that of the triple-echo and multiinterference methods (both

Table 5

Sensitivity and Specificity Comparison between Analysis Methods

Comparison	Sensitivity	Specificity
Dual echo versus triple echo	Less than*	More than*
Dual echo versus multiecho	Less than*	NS
Dual echo versus multiinterference	Less than*	More than*
Triple echo versus multiecho	NS	Less than*
Triple echo versus multiinterference	NS	NS
Multiecho versus multiinterference	NS	More than*

Note.—Comparisons were made at an FF diagnostic threshold of 6.25%. A significant difference was determined with two-sided 95% bootstrap confidence intervals of the sensitivity and specificity differences between two methods. NS = not significant.

* Indicates significant difference.

Figure 2

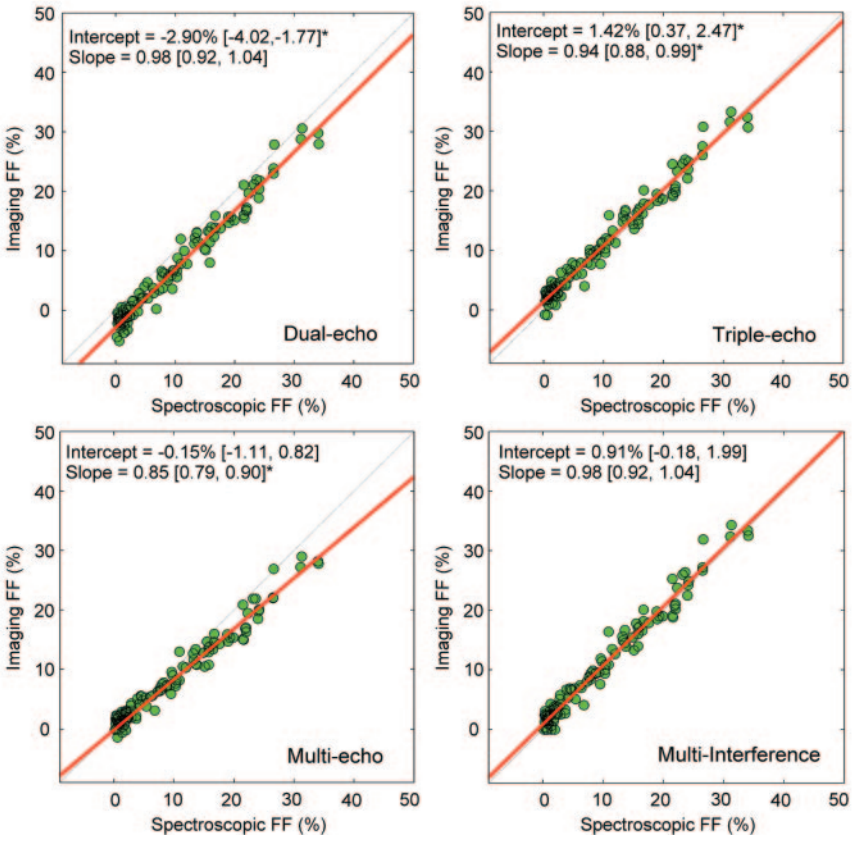

Figure 2: Scatterplots of MR spectroscopic versus imaging FFs calculated by using dual-echo, triple-echo, multiecho, and multiinterference methods. Red line represents the best fit through the data points whose spectroscopic FF is more than 6.25%, and gray line represents the null hypothesis (intercept = 0, slope = 1). The regression intercept and slope and their 95% confidence intervals are shown. * = Significant difference from intercept 0 or slope 1 according to two-tailed *t* test at $\alpha = .05$. All methods except for the multiinterference method have varying degrees of FF estimation error (ie, intercept significantly different from 0 and/or slope significantly different from 1.)

with specificity of 0.880). Similar trends in sensitivity and specificity were observed at 5%, 6%, 7%, and 8% threshold levels.

Grading Accuracy

As shown in the regression plots in Figure 2, the dual-echo method consistently resulted in underestimation of the spectroscopic FF by 2.90% over the entire FF range, but had no bias (systematic error) in slope. Estimated FF according to the triple-echo method had small, but statistically significant, bias in both slope and intercept. The multiecho method (without spectral modeling) was biased in slope only. The multiinterference method (with spectral modeling) had the highest grading accuracy, with no statistically significant bias. These observations are illustrated in Figure 3.

Discussion

This prospective clinical study compared hepatic fat quantification by using MR imaging and spectroscopy. Spectroscopic measurement of fat and water PDs was used to calculate the reference FF. Imaging data were acquired by using a low flip angle, at which T1 effects may be assumed negligible. Measurements were made at multiple echo times to permit T2* correction. Four image analysis methods were tested for diagnostic and grading accuracy.

The dual-echo method neglects T2* signal decay and assumes that the signal difference between IP and OP echoes is due to fat-water signal interference alone. With OP-then-IP sequential acquisition, the confounding T2* effect is known to cause FF underestimation (33,35,48). As expected, the dual-echo method systematically resulted in underestimation of spectroscopic FF over the full FF range, probably because of T2* relaxation effects intrinsic to liver tissue. As a result, this method had lower sensitivity than other T2*-corrected methods. Its high specificity is attributable to the systematic underestimation.

The triple-echo method resulted in improved diagnostic sensitivity (albeit at

Figure 3

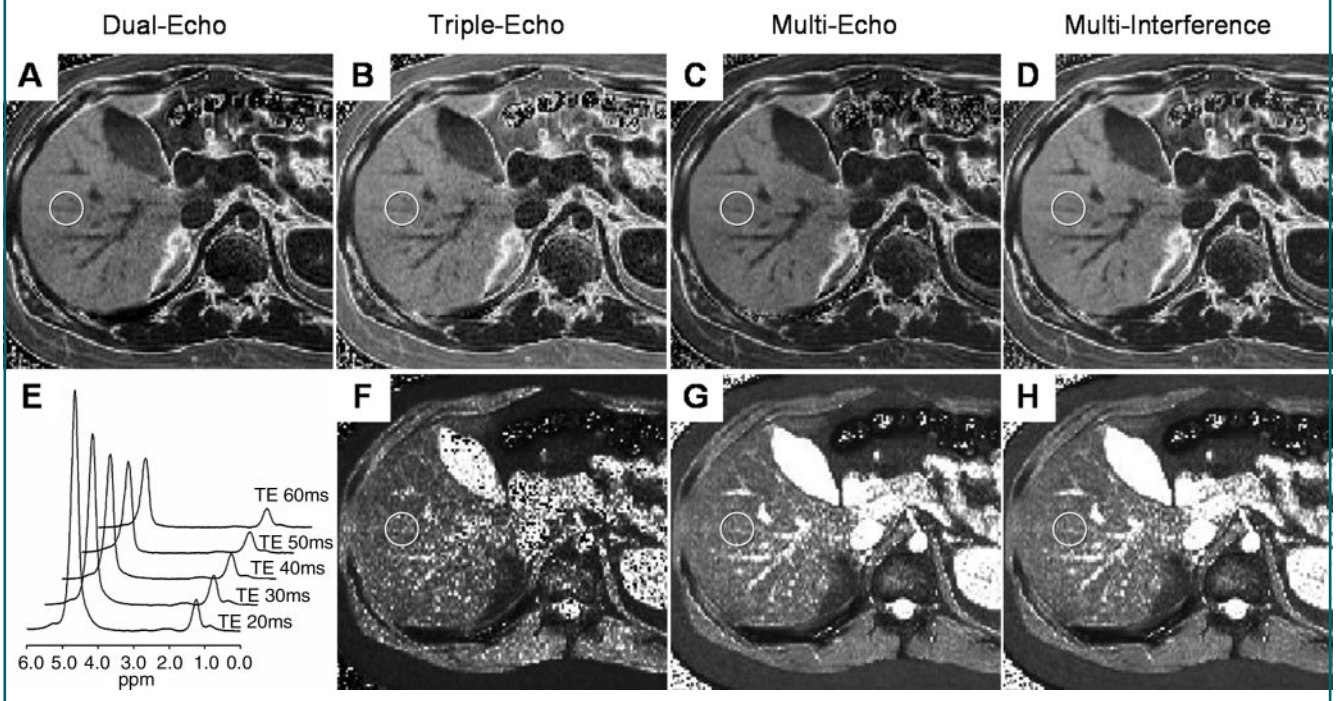

Figure 3: *A–D,* Estimated FF maps by using the four image analysis methods, *E,* multiecho MR spectra, and *F–H,* accompanying T2* maps in 18-year-old man with biopsy-confirmed NAFLD. The ROIs (circles) on imaging and spectroscopic voxel have been colocalized. The spectroscopic FF was 24.2% in *E.* Imaging FFs were 21.8% in *A,* 24.9% in *B,* 20.1% in *C,* and 24.9% in *D.* The estimated T2* values were 19.2 msec in *F,* 28.8 msec in *G,* and 24.7 msec in *H.* The triple-echo and multiinterference methods show higher quantification accuracy than the dual-echo and multiecho methods. *TE* = echo time.

lower specificity) and grading accuracy compared with the dual-echo method by T2* correction by using a single pair of IP echoes. While any pair of IP (or OP) echoes could be used for T2* correction, we favor the first IP pair because *(a)* earlier IP echoes tend to have superior signal-to-noise ratio and are less affected by fat-fat interference effects (described below) and *(b)* T2* estimation with OP echoes may be unreliable at high fat content because of near-complete signal cancellation.

The multiecho method estimates T2* from all six echoes. This method resulted in improved diagnostic specificity compared with that of the triple-echo method but caused systematic underestimation at high FFs. One likely explanation for the FF-dependent error is that the multiecho method assumes a simplified fat spectrum consisting of a single methylene peak at 1.3 ppm. The nominal OP and IP echo times are based on the phase-interference period between water and the methylene peak. However, the fat spectrum also contains nonmethylene peaks (eg, 2.1 and 0.9 ppm); the actual signal interference is therefore multicomponent and incompletely modeled by water-methylene interference alone. The resulting phase errors at nominal IP and OP echo times led to FF estimation error in proportion to the underlying fat content, as the signal contribution of the multiple fat peaks became a progressively larger proportion of the total (water and fat) signal.

Because the chemical shift between fat peaks is relatively small, fat-fat interference occurs at a slower rate than fat-water interference, and phase errors from fat-fat interference become more pronounced at later echo times. Thus, techniques that incorporate later echoes (eg, multiecho method) may be more susceptible to fat-fat interference effects than techniques that incorporate only early echoes (eg, dual- and triple-echo methods).

The multiinterference method corrects for the complex interference effect due to multiple fat peaks, in addition to correcting for T2* effects. This method had greater diagnostic sensitivity than the other methods and had higher fat grading accuracy over the full range of observed FF.

These results indicate that T2* correction improves the diagnostic sensitivity at a 6.25% spectroscopic threshold, and modeling the fat spectrum as a three-peak system improves quantification accuracy above this threshold. Thus, we recommend routine T2* correction for fatty liver screening and, if possible, spectral modeling for fatty liver grading. In our opinion, T2* correction is particularly important, as co-occurrence of hepatic steatosis and siderosis (iron deposition) is not uncommon (49,50). In hepatic siderosis, T2* may be so short (<10 msec) that the diagnostic sensitivity is profoundly compromised without T2* correction. With imagers without multiecho GRE capability, the triple-echo method may be a

reasonable practical alternative that can be implemented as two dual-echo sequences, with IP-then-OP and IP-then-IP acquisition, respectively.

One limitation of our study was that spectroscopy, rather than histologic examination, was used as the reference. While a spectroscopic reference may not reflect current clinical practice, histomorphologic assessment of intracellular fat vacuoles does not directly measure hepatic triglyceride concentration and may not be an appropriate reference standard for MR imaging. There also are challenges in imaging–histologic colocalization. These factors may account in part for the inconsistent accuracy of imaging relative to histology in previous reports (51–53). By comparison, spectroscopically determined PD FF is equivalent to tissue triglyceride concentration, and several clinical trials (6,54,55) on NAFLD have used spectroscopic FF as an outcome measure. Moreover, spectroscopy can be colocalized reliably with imaging. For these reasons, spectroscopy may be a more appropriate reference standard than histology for assessing accuracy of fat quantification.

The 6.25% FF diagnostic threshold was derived from T2-corrected point-resolved spatially localized spectroscopic data from the Dallas Heart Study (39). In our study, stimulated-echo acquisition mode spectroscopy was performed instead because of its relative insensitivity to J-coupling effects and more reliable T2 estimation for fat quantification (56). Although similar sequence parameters and spectroscopic analysis methods were used, it is conceivable that the diagnostic thresholds may differ slightly between stimulated-echo acquisition mode spectroscopy and point-resolved spatially localized spectroscopy. However, this potential difference probably did not affect our findings because the diagnostic performance was assessed at several threshold values (5%–8%) with qualitatively similar results.

One technical limitation of in vivo spectroscopy at clinical field strengths is the inability to resolve the three minor fat components (4.3, 5.2, and 5.3 ppm), which are obscured by the 4.7-ppm water peak. These fat peaks may contain up to 15% of the total fat content (38), leading to underestimation of triglyceride concentration. This is easily correctable with a conversion formula (38), as done by Szczepaniak et al (39). In our study, we did not apply the conversion formula because both MR spectroscopy and imaging are affected equally by this problem, and thus the relative accuracy between spectroscopy and imaging is unaltered.

A theoretical limitation of the image analysis methods presented here was the inability to accurately quantify fat content beyond 50% FF due to fat-water signal dominance ambiguity. In fat-dominant tissues (PD fat > water), additional techniques, such as phase unwrapping (57), fat suppression (28), or multiple flip angles (or multiple repetition times) (29), would be necessary. However, this theoretical limitation may not be an important clinical issue for hepatic fat quantification, as PD FF values exceeding 50% were not observed in any of the 2287 Dallas Heart Study subjects (39) or in any of our 110 subjects.

Random imaging noise has been suggested as a potential source of error in fat quantification imaging (25,36). When a complex image is rectified to a magnitude image, the complex zero-mean random noise is transformed into positive-mean random noise. This positive-mean noise is expected to introduce bias to fat and water PD estimates, and variance in the signals is expected to decrease reproducibility. While we did not observe gross evidence of these effects in our regression or reproducibility analyses, no formal assessments were performed. Future studies will be required to assess the effect of image noise in vivo.

Finally, we did not assess spatial heterogeneity of liver fat content, as spectroscopic and imaging FFs were compared within a single colocalized region. In principle, our MR imaging technique would enable assessment of the spatial heterogeneity, but such data were not analyzed.

In conclusion, this prospective clinical study showed close agreement between PD-based MR spectroscopy and imaging over the full range of clinically relevant fat content. Relaxation- and interference-corrected fat quantification at low-flip-angle multiecho GRE MR imaging provided high diagnostic and fat-grading accuracy in NAFLD.

Acknowledgments: We acknowledge contribution to this work by Richard Znaminowski (MR technologist) and Lillian Pacheco, Julie Collins, Alyssa Chavez, Iva Grabic, Joseph Aguilera, Zana Parman, Manuel Celedon, Lisa Clark, Lita Petchaporn, and Deanna Oliver (research coordinators).

References

1. Lavine JE, Schwimmer JB. Nonalcoholic fatty liver disease in the pediatric population. Clin Liver Dis 2004;8:549–558, viii–ix.
2. Farrell GC, Larter CZ. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. Hepatology 2006;43(2 suppl 1):S99–S112.
3. Sanyal AJ. AGA technical review on nonalcoholic fatty liver disease. Gastroenterology 2002;123:1705–1725.
4. McCullough AJ. Update on nonalcoholic fatty liver disease. J Clin Gastroenterol 2002;34:255–262.
5. Clark JM. Weight loss as a treatment for nonalcoholic fatty liver disease. J Clin Gastroenterol 2006;40(suppl 1):S39–S43.
6. Belfort R, Harrison SA, Brown K, et al. A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. N Engl J Med 2006;355:2297–2307.
7. Tilg H, Kaser A. Treatment strategies in nonalcoholic fatty liver disease. Nat Clin Pract Gastroenterol Hepatol 2005;2:148–155.
8. Promrat K, Lutchman G, Uwaifo GI, et al. A pilot study of pioglitazone treatment for nonalcoholic steatohepatitis. Hepatology 2004;39:188–196.
9. Harrison SA, Neuschwander-Tetri BA. Pharmacologic management of nonalcoholic fatty liver disease. Clin Liver Dis 2004;8:715–728, xii.
10. Angulo P. Treatment of nonalcoholic fatty liver disease. Ann Hepatol 2002;1:12–19.
11. 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.
12. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scor-

ing system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–1321.

13. Joy D, Thava VR, Scott BB. Diagnosis of fatty liver disease: is biopsy necessary? Eur J Gastroenterol Hepatol 2003;15:539–543.

14. Thampanitchawong P, Piratvisuth T. Liver biopsy: complications and risk factors. World J Gastroenterol 1999;5:301–304.

15. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898–1906.

16. Janiec DJ, Jacobson ER, Freeth A, Spaulding L, Blaszyk H. Histologic variation of grade and stage of non-alcoholic fatty liver disease in liver biopsies. Obes Surg 2005;15:497–501.

17. Charatcharoenwitthaya P, Lindor KD. Role of radiologic modalities in the management of non-alcoholic steatohepatitis. Clin Liver Dis 2007;11:37–54, viii.

18. Fischbach F, Bruhn H. Assessment of in vivo 1H magnetic resonance spectroscopy in the liver: a review. Liver Int 2008;28:297–307.

19. Szczepaniak LS, Babcock EE, Schick F, et al. Measurement of intracellular triglyceride stores by H spectroscopy: validation in vivo. Am J Physiol Endocrinol Metab 1999;276:E977–E989.

20. Thomsen C, Becker U, Winkler K, Christoffersen P, Jensen M, Henriksen O. Quantification of liver fat using magnetic resonance spectroscopy. Magn Reson Imaging 1994;12:487–495.

21. Brix G, Heiland S, Bellemann ME, Koch T, Lorenz WJ. MR imaging of fat-containing tissues: valuation of two quantitative imaging techniques in comparison with localized proton spectroscopy. Magn Reson Imaging 1993;11:977–991.

22. Irwan R, Edens MA, Sijens PE. Assessment of the variations in fat content in normal liver using a fast MR imaging method in comparison with results obtained by spectroscopic imaging. Eur Radiol 2008;18:806–813.

23. Kim H, Taksali SE, Dufour S, et al. Comparative MR study of hepatic fat quantification using single-voxel proton spectroscopy, two-point dixon and three-point IDEAL. Magn Reson Med 2008;59:521–527.

24. Kawamitsu H, Kaji Y, Ohara T, Sugimura K. Feasibility of quantitative intrahepatic lipid imaging applied to the magnetic resonance dual gradient echo sequence. Magn Reson Med Sci 2003;2:47–50.

25. Machann J, Thamer C, Schnoedt B, et al. Hepatic lipid accumulation in healthy subjects: a comparative study using spectral fat-selective MRI and volume-localized 1H-MR spectroscopy. Magn Reson Med 2006;55:913–917.

26. Kamman RL, Bakker CJ, van Dijk P, Stomp GP, Heiner AP, Berendsen HJ. Multi-exponential relaxation analysis with MR imaging and NMR spectroscopy using fat-water systems. Magn Reson Imaging 1987;5:381–392.

27. Bernard CP, Liney GP, Manton DJ, Turnbull LW, Langton CM. Comparison of fat quantification methods: a phantom study at 3.0T. J Magn Reson Imaging 2008;27:192–197.

28. Qayyum A, Goh JS, Kakar S, Yeh BM, Merriman RB, Coakley FV. Accuracy of liver fat quantification at MR imaging: comparison of out-of-phase gradient-echo and fat-saturated fast spin-echo techniques—initial experience. Radiology 2005;237:507–511.

29. Hussain HK, Chenevert TL, Londy FJ, et al. Hepatic fat fraction: MR imaging for quantitative measurement and display—early experience. Radiology 2005;237:1048–1055.

30. Bydder M, Middleton MS, Chavez AD, Sirlin CB. Effect of flip angle on fat quantification by Dixon techniques [abstr]. In: Proceedings of the Fourteenth Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, Calif: International Society for Magnetic Resonance in Medicine, 2006; 2300.

31. Fishbein MH, Gardner KG, Potter CJ, Schmalbrock P, Smith MA. Introduction of fast MR imaging in the assessment of hepatic steatosis. Magn Reson Imaging 1997;15:287–293.

32. Westphalen AC, Qayyum A, Yeh BM, et al. Liver fat: effect of hepatic iron deposition on evaluation with opposed-phase MR imaging. Radiology 2007;242:450–455.

33. Yu H, McKenzie CA, Shimakawa A, et al. Multiecho reconstruction for simultaneous water-fat decomposition and T2* estimation. J Magn Reson Imaging 2007;26:1153–1161.

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

35. Bydder M, Yokoo T, Hamilton G, et al. Relaxation effects in the quantification of fat using gradient echo imaging. Magn Reson Imaging 2008;26:347–359.

36. 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.

37. Rudin M, de Beer R. In-vivo magnetic resonance spectroscopy. Berlin, Germany: Springer-Verlag, 1992;

38. Longo R, Pollesello P, Ricci C, et al. Proton MR spectroscopy in quantitative in vivo determination of fat content in human liver steatosis. J Magn Reson Imaging 1995;5:281–285.

39. Szczepaniak LS, Nurenberg P, Leonard D, et al. Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. Am J Physiol Endocrinol Metab 2005;288:E462–E468.

40. de Bazelaire CM, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. Radiology 2004;230:652–659.

41. Naressi A, Couturier C, Castang I, de Beer R, Graveron-Demilly D. Java-based graphical user interface for MRUI, a software package for quantitation of in vivo/medical magnetic resonance spectroscopy signals. Comput Biol Med 2001;31:269–286.

42. Vanhamme L, van den Boogaart A, Van Huffel S. Improved method for accurate and efficient quantification of MRS data with use of prior knowledge. J Magn Reson 1997;129:35–43.

43. Dixon WT. Simple proton spectroscopic imaging. Radiology 1984;153:189–194.

44. Bydder M, Yokoo T, Hamilton G, Middleton MS, Chavez AD, Sirlin CB. Fat-fat interactions in Dixon-variant imaging [abstr]. In: Proceedings of the Fifteenth Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, Calif: International Society for Magnetic Resonance in Medicine, 2007; 1632.

45. Yokoo T, Barksdale J, Bydder M, et al. Effects of T1, T2, and spectral complexity on in- and out-of-phase imaging: a systematic approach by computer simulation [abstr]. In: Proceedings of the Fifteenth Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, Calif: International Society for Magnetic Resonance in Medicine, 2007; 1720.

46. Thomas FC, Yuying L. An interior trust region approach for nonlinear minimization subject to bounds. SIAM J Optim 1996;6:418–445.

47. Efron B. Better bootstrap confidence intervals. J Am Stat Assoc 1987;82:171–185.

48. Siegelman ES. MR imaging of diffuse liver disease: hepatic fat and iron. Magn Reson Imaging Clin N Am 1997;5:347–365.

49. George DK, Goldwurm S, MacDonald GA,

et al. Increased hepatic iron concentration in nonalcoholic steatohepatitis is associated with increased fibrosis. Gastroenterology 1998;114:311–318.

50. Bonkovsky HL, Jawaid Q, Tortorelli K, et al. Non-alcoholic steatohepatitis and iron: increased prevalence of mutations of the HFE gene in non-alcoholic steatohepatitis. J Hepatol 1999;31:421–429.

51. Kim SH, Lee JM, Han JK, et al. Hepatic macrosteatosis: predicting appropriateness of liver donation by using MR imaging—correlation with histopathologic findings. Radiology 2006;240:116–129.

52. Schuchmann S, Weigel C, Albrecht L, et al. Non-invasive quantification of hepatic fat fraction by fast 1.0, 1.5 and 3.0 T MR imaging. Eur J Radiol 2007;62:416–422.

53. Cho CS, Curran S, Schwartz LH, et al. Preoperative radiographic assessment of hepatic steatosis with histologic correlation. J Am Coll Surg 2008;206:480–488.

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

55. Petersen KF, Dufour S, Befroy D, Lehrke M, Hendler RE, Shulman GI. Reversal of nonalcoholic hepatic steatosis, hepatic insulin resistance, and hyperglycemia by moderate weight reduction in patients with type 2 diabetes. Diabetes 2005;54:603–608.

56. Hamilton G, Middleton MS, Bydder M, et al. Influence of MR spectroscopy sequence on the measurement of IHCL levels [abstr]. In: Radiological Society of North America Scientific Assembly and Annual Meeting Program. Oak Brook, Ill: Radiological Society of North America, 2007; 548.

57. Reeder SB, Pineda AR, Wen Z, et al. Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): application with fast spin-echo imaging. Magn Reson Med 2005;54:636–644.