Role of Advanced MRI Techniques in the Quantitative Assessment of Liver Fat—A Multimodality-Based Comparative Study of Diagnostic Performance in a Tertiary Care Institute

Santosh B. Patil1, Seetharaman Cannane2, Santhosh Poyyamoli1, Rinoy R. Anand1, Venkatesh Kasi3

1Department of Vascular and Interventional Radiology, Kovai Medical Center and Hospital, Coimbatore, India
2Department of Radiodiagnosis, Kovai Medical Center and Hospital, Coimbatore, India
3Department of Radiology, Kovai Medical Center and Hospital, Coimbatore, India

Address for correspondence Seetharaman Cannane, DMRD, DNB, Assistant Professor, Department of Radiodiagnosis, Kovai Medical Center and Hospital Institute of Medical Sciences and Research, 99, Avanashi Road, Coimbatore 641014, India (e-mail: drseetharaman@gmail.com).

Abstract

Objective To evaluate the accuracy of noninvasive imaging methods including gray-scale ultrasound, ultrasound shear wave elastography, unenhanced computed tomography (CT), and proton density fat fraction (PDFF) on magnetic resonance imaging (MRI) using three-dimensional (3D) multiecho multipoint chemical shift–encoded spoiled gradient echo (q-DIXON) sequence in the quantification of hepatic steatosis, with proton MR spectroscopy (H1-MRS) as the reference standard in Indian population.

Methods Our study included 100 consecutive adult patients referred to the department of radiology in our hospital for imaging of liver. Fat content of liver was recorded using MRI (H1-MRS and q-DIXON), unenhanced CT (average liver attenuation [ALI] and liver attenuation index [LAI]) and ultrasonography (USG) (gray-scale grading and shear wave elastography [SWE]). Data were analyzed by linear regression and Bland–Altman analysis for each technique compared with H1-MRS. The diagnostic performances of all the methods were compared using DeLong test, for detection of mild and moderate-to-severe hepatic steatosis, separately.

Results MRI q-DIXON PDFF showed excellent correlation (r = 0.917, r2 = 0.840) and strong agreement (1.48 ± 3.01) with H1-MRS-derived PDFF measurements. Unenhanced CT-based methods showed moderate correlation with modest agreement (r = −0.826, r2 = 0.681, −40.18 ± 16.05 for ALI and r = −0.858, r2 = 0.735, 13.4 ± 15.3 for LAI) whereas USG gray-scale assessment showed low correlation (weighted Kappa value 0.366) with H1-MRS PDFF. No correlation was found between USG-SWE results and PDFF measured with H1-MRS. Comparison of areas under curve (AUCs) using DeLong test revealed that MRI q-DIXON method performed the best for diagnosis of hepatic steatosis compared with rest. For moderate to severe steatosis, MRI q-DIXON and unenhanced CT-based methods had comparable diagnostic performance with AUCs not showing statistically significant differences.

Keywords ► H1-MR spectroscopy ► chemical shift MRI ► nonalcoholic fatty liver disease
Introduction

According to various population studies, at least 30% of the general adult population worldwide has fatty liver.\textsuperscript{1-3} A liver is considered fatty when the total fat content exceeds 5% of the wet weight of the liver.\textsuperscript{4,5} The most commonly recognized cause for fatty liver today is nonalcoholic fatty liver disease (NAFLD). Other conditions that are associated with hepatic steatosis include excessive alcohol consumption, chronic viral infection (hepatitis C), metabolic or storage disorders, and certain drugs/toxins.\textsuperscript{6}

Detection and quantification of hepatic steatosis is clinically important in several situations: In NAFLD, steatosis is recognized as the earliest biomarker and necessary feature for the development of nonalcoholic steato hepatitis (NASH). In hepatitis C, steatosis is associated with more severe fibrosis and rapid disease progression while an adequate response to antiviral treatment results in a decrease of steatosis.\textsuperscript{7} In addition, the regenerative capacity of liver in both donor and recipients in liver transplantation surgery is affected by hepatic steatosis which can result in primary nonfunction of the liver graft. The maximum amount of fatty infiltration for liver grafts accepted by most transplantation centers varies between 10 and 30%.\textsuperscript{8}

A quantitative biomarker, if developed, can be used for diagnosis of disease and clinical decision making by applying cutoff values for binary classification of disease presence or absence or disease severity. It is necessary that a biomarker should be accurate and precise.

Liver biopsy has historically been considered the reference standard for assessment of hepatic steatosis. However, biopsy is an invasive procedure. The histopathology report is subject to interobserver variability and also prone to sampling errors.\textsuperscript{9}

Generic serum markers of liver disease (e.g., aminotransferases) are relatively insensitive and nonspecific for the detection of hepatic steatosis.\textsuperscript{10}

The imaging modalities which can detect and quantify the hepatic fat content noninvasively include ultrasound, unenhanced computed tomography (CT)- and magnetic resonance imaging (MRI)-based methods (chemical shift MRI and proton MR-spectroscopy \([H^1\text{-MRS}]\)). \([H^1\text{-MRS}]\) is presently considered the most accurate technique among all.\textsuperscript{11}

To date, only a few studies have evaluated all three imaging modalities for quantification of hepatic fat simultaneously in a single study design.\textsuperscript{12,13} Moreover, there is no published evidence in ethnic Indian population. Our study aims at the validation of noninvasive imaging methods including gray-scale ultrasound, ultrasound shear wave elastography (SWE), unenhanced CT, and proton density fat fraction (PDFF) using multiecho three-dimensional (3D) Dixon sequence in the quantification of hepatic steatosis, with \([H^1\text{-MRS}]\) as the reference standard in Indian population.

Materials and Methods

This was a cross-sectional study conducted in a tertiary care hospital in Southern India. Hospital ethical and scientific committee approvals were obtained prior to the conduct of the study. The patients were volunteers from a regular screening program for liver diseases in our institute, which included healthy people and those with previously undetected liver disease. Prior informed consents were obtained from all the patients except for standard exclusion criteria for CT and MRI examinations such as pregnancy, claustrophobia, and presence of cardiac pacemakers or other implanted electronic devices; no specific exclusion criteria were applied. A pilot study was done initially where we collected data of 20 patients from May 1 to June 30, 2016. The sample size of the principal study was then estimated based on the findings of pilot study using the following formula:

\[
n = 4Z_\alpha^2 \left( \frac{S}{W} \right)^2 \tag{1}
\]

Minimum required sample size of the study was estimated to be 73, based on the confidence level \((\alpha)\) of 95%, desired width of confidence interval \((W)\) of 5, and standard deviation \((S)\) of 10.9. A total of 100 successive adult patients were included in this study conducted from May 1, 2016, to May 31, 2017. None of the patients were excluded from the study based on the exclusion criteria.

Age, sex, height, and weight of each patient were recorded before doing imaging. Body mass index was calculated by using the formula weight in kilograms/height in meters\textsuperscript{2}. MRI, unenhanced CT, and USG examinations were performed...
in succession for all the patients. All the examinations of a particular patient were performed within a period of 2 hours, usually between 8.00 a.m. and 12.00 p.m. to reduce diurnal variation.

MRA was performed on a MAGNETOM Skyra 3.0T MRI system (Siemens Healthcare, Erlangen, Germany) using a specialized application package called LIVER LABORATORY. The package included T1 VIBE-based dual-echo and multiecho chemical shift (Dixon) sequences as well as proton MRS.

T1 VIBE e-DIXON (dual-echo spoiled gradient echo Dixon) images were routinely acquired in 18 seconds (single breath-hold). The imaging parameters included TR 7.17 milliseconds, TE1 2.4 milliseconds, TE2 4.8 milliseconds; 288 × 162 matrix; 3 mm thickness; 80 slices; and 5° flip angle (to reduce T1-related bias). This sequence screened the liver for fat deposition, hence called "first look" Dixon.

T1 VIBE q-DIXON (multiecho multipoint spoiled gradient–echo Dixon) images were routinely acquired in 18 seconds (single breath-hold). The imaging parameters were TR 9.7 milliseconds; multiple (6) TE = (1.2, 2.4, 4.8, 5.9, 7.1, 8.3) milliseconds (to correct T2* bias); 256 × 154 matrix; 3.5 mm thickness; 64 slices; and 5° flip angle (to reduce T1-related bias). Multifat peak model was used for robust lipid estimation. The final report sheet was created by automated processing and proton density fat fraction was expressed as text and color bar.

The HISTO (proton MRS) was a 15-second single breath-hold, stimulated echo acquisition mode (STEAM) spectroscopy sequence. It used a single voxel spectroscopy with typical voxel size being 3 × 3 × 3 cm³. A series of five sequences were concatenated with a fixed TR of 3,000 milliseconds (to reduce T1 bias), resulting in acquired TEs of 12, 24, 36, 48, and 72 milliseconds (multiple TEs were used to correct for T2 relaxation effects). Reconstruction and postprocessing of HISTO-MRS was inline and a report sheet was created. The PDFF was expressed as text and color bar.

During each of these sequences, the system automatically suggested a position of region of interest (ROI); however, it could also be placed manually. Major vessels/ducts and gall bladder were avoided during positioning of the ROIs. Care was taken to position the ROIs in similar regions of liver tissue for all the sequences. The ROI in MRS sequence can only be placed during preprocessing, whereas the ROI in q-DIXON can be placed in postprocessing also. In cases where automatic liver segmentation in q-DIXON sequence was not accurate, manual segmentation was done by drawing ROIs for each slice using free hand ROI tool and then the values were averaged together. MRI images and results were interpreted by a radiologist with 12 years of experience and he was blinded to the results of other modalities.

CT examinations were performed on a 64 slice MDCT–GE Lightspeed VCT (GE Healthcare, Milwaukee, Wisconsin, United States). Unenhanced CT scan was performed with the following parameters—120 kV, 200 mA with dose modulation, matrix 512 × 512, section thickness of 5 mm, and an increment of 4 mm. CT images were analyzed on AW server (Advanced Workstation 2.0, GE Healthcare, Milwaukee, Wisconsin, United States). Average liver attenuation (Lavg) was calculated by manually placing circular ROIs of at least 1 cm² area at multiple places in liver, covering all the hepatic segments and avoiding major vessels, ducts, and fissures. Average splenic attenuation was measured by placing ROIs at its upper, mid, and lower poles. The liver attenuation index (LAI) was calculated as difference between average hepatic and average splenic attenuations. CT images were interpreted by a radiologist of 10 years of experience and he was blinded to the results of other modalities.

Ultrasound was performed using Logiq E9 (GE healthcare, Milwaukee, Wisconsin, United States) ultrasonography (USG) scanner with dedicated SWE software by a radiologist with 3 years of experience who was blinded to the results of other modalities. All the patients were examined in the supine position in fasting state. A 3.6 Hz convex transducer was used. On gray-scale ultrasound, a 4-point scale was used for classifying hepatic steatosis: 0, no evidence of steatosis; 1, mild steatosis with changes in echogenicity; 2, moderate steatosis with obscuration of portal venous walls; and 3, severe steatosis with marked attenuation and no visualization of the diaphragm or any vessels. SWE scores in kilopascals were obtained by placing ROIs of at least 1 cm² area in each of the eight segments of the liver and avoiding major vessels and ducts. A report sheet was generated finally with a mean elastography score in kilopascals.

Correlation and simple linear regression analysis was performed to correlate and calibrate the hepatic fat content obtained from all modalities against H1-MRS PDFF. Further, Bland–Altman analysis was performed to find out the agreement between hepatic fat content obtained by each of the modality and H1-MRS PDFF. Because gray-scale USG results were qualitative, weighted kappa analysis was done to find its correlation with H1-MRS PDFF. Receiver operating curve (ROC) analysis was performed for each of the imaging method for diagnosing all grades of hepatic steatosis (H1–MRS PDFF ≥ 5.6%) as well as for moderate to severe steatosis (H1–MRS PDFF ≥ 10%), separately.11,15–17 Corresponding areas under curve (AUC) and optimum cutoff values for diagnosis were obtained for all the imaging methods. The AUCs of different methods were compared using DeLong test.18 p-Values less than 0.05 were considered as statistically significant.

Results

The study population comprised of 88 men and 12 women. The mean age of the study population was 47 years (range 21–77 years). The body mass index values for the study population ranged from a minimum of 19.47 to a maximum of 33.15 (mean of 26.75).

The PDFF values for the study population obtained by H1-MRS ranged from a minimum of 1.10% to a maximum of 25.50% (mean 9.90). Out of the total 100 patients in the study population, 36% had normal liver with PDFF values on H1-MRS being <5.56%. Twenty-three percent patients had grade I (mild) steatosis (H1-MRS PDFF ≥ 5.56% and <10%) whereas 41% had grade II–III (moderate to severe) steatosis (H1-MRS PDFF ≥ 10%).
Three cases showed unusually high PDFF values on MRI q-DIXON sequence due to fat–water swap phenomenon. Hence, these three values were not considered for statistical analysis. The R2* values in all the cases were on a lower side (less than 88 second^-1), suggesting no iron overload.

The PDFF values for the rest of 97 patients in the study population ranged from a minimum of 1.70% to a maximum of 22.90% (mean 8.65%).

Results of correlation and linear regression analysis showed excellent correlation between PDFF values obtained by H1-MRS and MR q-DIXON methods (r = 0.917, r² = 0.840, p = 0.00) (►Fig. 1A). Bland–Altman analysis revealed strong agreement between the PDFF values by two methods (mean of limits of agreement 1.48 ± SD 3.01) (►Fig. 2A).

Based on ROC analysis, we derived an optimal cutoff value of 5.3% for diagnosis of hepatic steatosis by MR q-DIXON method in our study population (AUC 0.923, sensitivity 89%, specificity 70%, and accuracy 83%) (►Fig. 3A). Similarly, we derived an optimal cutoff of 8.5% for diagnosis of moderate to severe hepatic steatosis by MR q-Dixon method (AUC 0.981, sensitivity 97%, specificity 89%, and accuracy 92%) (►Fig. 4A).
Unenhanced CT
The average liver attenuation values for the study population obtained by unenhanced CT ranged from a minimum of 25.90 HU to a maximum of 64.60 HU (mean 50.08 HU).

The average attenuation of liver on unenhanced CT scan showed moderate negative correlation (r = −0.826, p = 0.00) with H¹-MRS PDFF. Linear regression showed good correlation between the two methods (r² = 0.681, p = 0.00) (►Fig. 1B). The limits of Bland-Altman analysis were modest, with mean for the limits of agreement being −40.18 (± SD 16.05) (►Fig. 2B). We further derived an optimal cutoff of 47 HU for the diagnosis of moderate to severe steatosis (AUC 0.956, sensitivity 78%, specificity 95%, and accuracy 88%) (►Fig. 4B).

Unenhanced CT
The average liver attenuation values for the study population obtained by unenhanced CT ranged from a minimum of 25.90 HU to a maximum of 64.60 HU (mean 50.08 HU).

The average attenuation of liver on unenhanced CT scan showed moderate negative correlation (r = −0.826, p = 0.00) with H¹-MRS PDFF. Linear regression showed good correlation between the two methods (r² = 0.681, p = 0.00) (►Fig. 1B). The limits of Bland-Altman analysis were modest, with mean for the limits of agreement being −40.18 (± SD 16.05) (►Fig. 2B). We further derived an optimal cutoff of 47 HU for the diagnosis of moderate to severe steatosis (AUC 0.956, sensitivity 78%, specificity 95%, and accuracy 88%) (►Fig. 4B).

The CT LAI values for the study population obtained by unenhanced CT ranged from a minimum of −30.00 HU to a maximum of 12.80 HU (mean −3.58 HU).

The LAI on unenhanced CT scan showed strong negative correlation (r = −0.858, p = 0.00) with H¹-MRS PDFF. Linear regression showed good correlation between the two methods (r² = 0.735, p = 0.00) (►Fig. 1C). Bland-Altman analysis in our study revealed modest agreement between H¹ MRS PDFF and CT LAI values (mean for limits of agreement 13.4± SD 15.3) (►Fig. 2C). We further derived an optimal cutoff of −2 HU for the diagnosis of moderate to severe hepatic steatosis (AUC of 0.853 was noted for diagnosing moderate to severe hepatic steatosis).
and a cutoff value of −7 HU for the diagnosis of moderate to severe steatosis (AUC 0.985, sensitivity 95%, specificity 96%, and accuracy 96%) (►Fig. 4C).

Ultrasound
Estimation of hepatic steatosis by gray-scale ultrasound showed low correlation with $^{1}{H}$-MRS PDFF results with weighted kappa value of 0.366 ($p = 0.00$).

The sensitivity, specificity, and accuracy for diagnosing hepatic steatosis on gray-scale USG were 70.3%, 52.8%, and 62%, respectively. Similarly the sensitivity, specificity, and accuracy for diagnosing moderate to severe steatosis were found to be 48.8%, 98.3%, and 76%, respectively. ROC analysis showed an AUC of 0.693 for diagnosing all grades of hepatic steatosis (►Fig. 3D) and 0.853 for diagnosing moderate to severe hepatic steatosis (►Fig. 4D).

The ultrasound elastography scores for the study population obtained by SWE ranged from a minimum of 2.96 to a maximum of 13.89 (mean 4.79). The SWE scores showed poor correlation with the PDFF values obtained by $^{1}{H}$-MRS ($r = -0.0310, r^2 = 0.001, p = 0.763$) (►Fig. 1D).

Comparison of MRI, Unenhanced CT, and USG Methods
The comparison of AUCs by ROC analysis revealed that MRI q-DIXON method performed the best for diagnosis of all grades of hepatic steatosis compared with rest. The difference in the AUC of MRI q-DIXON with other methods was statistically significant ($p < 0.05$ for all comparisons by DeLong test).

Unenhanced CT–based methods had a better AUC compared with USG gray scale with $p < 0.05$ (statistically significant) (►Table 1A).

On the other hand, for moderate to severe steatosis, MRI q-DIXON and unenhanced CT–based methods had comparable diagnostic performance with AUC not showing statistically significant differences ($p > 0.05$ for all comparisons). Comparison of CT-based methods showed that CT LAI performed better as compared with CT $L_{avg}$ method with $p = 0.04$ for comparison of AUCs. Gray-scale USG performed inferiorly as compared with MRI- and CT-based methods ($p < 0.05$ for all comparisons) (►Table 1B).

Discussion
After the findings of Dallas heart study (a multiethnic, probability-based population sample) were published in 2005, $^{1}{H}$-MRS is increasingly being used as reference standard instead of liver biopsy in clinical trials, diagnostic studies, and observational studies. In this study, the upper limit of normal (95th percentile) for hepatic triglyceride content in 345 healthy subjects without identifiable risks for hepatic steatosis was found to be 5.56%.

Chemical shift MRI utilizes the difference in resonance frequency of water and lipid to differentiate tissues containing only water from those containing both water and lipid. The most advanced chemical shift technique is single breath-hold T1-independent, T2-corrected, spectrally corrected complex–based technique.

Our study showed that 3D multiecho multipoint chemical shift–encoded spoiled gradient echo sequence (MRI q-DIXON) had excellent correlation with $^{1}{H}$-MRS in hepatic fat quantification. This method also had good diagnostic performance in detecting all grades of hepatic steatosis. Similar results were reported in various other studies conducted in adults as well as in children. Adding this sequence to the routine standard-of-care liver or abdominal MRI examination will minimally lengthen the examination duration and provide the benefit of quantifying liver fat content over the entire organ which can help in early diagnosis of fatty liver. Further this sequence can also be repeated several times to assess treatment response without any fear of radiation exposure. Several studies have shown that distribution of fat in liver can be inhomogeneous. The image dataset obtained by single breath-hold chemical shift imaging sequence can be used to calculate fat content from different segments of liver by drawing ROIs manually during postprocessing.

In our study, unenhanced CT–based methods (average attenuation of liver and liver attenuation index) showed good correlation with $^{1}{H}$-MRS. CT was accurate for the diagnosis of grade 2 to 3 steatosis but was not as accurate for detecting grade 1 steatosis. CT quantification of fat seems logical in prospective liver donors who undergo preoperative CT evaluation for assessment of liver volume and vasculature. Also, unenhanced CT is performed routinely as a standard of care for many indications (like urolithiasis, CT colonography), so liver fat measurements can be easily obtained in these subjects from the already existing datasets.

Ultrasound is often used as a first-line noninvasive method for evaluating fatty liver disease. The traditional four-grade visual assessment on gray scale ultrasound showed low correlation with $^{1}{H}$-MRS in our study. This was also reflected in the wide range of sensitivity and specificity values reported for diagnosis of steatosis using USG in various studies. Nevertheless, considering the ease of the procedure, better accessibility, and relatively low cost of USG, USG still has merits for a screening tool for hepatic steatosis.

Though USG elastography has been found useful for assessing liver stiffness associated with fibrotic or cirrhotic changes, only little is known about the application and accuracy of this technique for evaluating liver fat content. Our study tried to evaluate the correlation between $^{1}{H}$-MRS PDFF scores and two-dimensional (2D) SWE scores in its present form. USG elastography in its present form was found to have no role in quantification of hepatic fat content. However further research should focus on development and clinical validation of elastography-based investigational techniques (such as calculation of controlled attenuation parameter) in the detection of hepatic steatosis. Such a technique can help in simultaneous assessment of hepatic steatosis and fibrosis, thus allowing early diagnosis of and screening for NASH and NAFLD-related cirrhosis.

We acknowledge a few important limitations of our study. $^{1}{H}$-MRS was used as the reference standard without histopathological correlation. Similar to tissue biopsy, the spatially heterogeneous distribution of hepatic fat is also a
confounding factor which limits the use of H¹-MRS as a reference standard. However, this factor is not believed to have much influence on our study, because none of our patients had geographic pattern of fat distribution on other fat imaging methods. Also, the results of H¹-MRS correlated well with other imaging methods in almost all cases. Our study did not evaluate whether the presence of cirrhosis or fibrosis influences the accuracy of the MRI, CT, or USG results.

Conclusion
q-DIXON shows strongest correlation with MRS and should be preferred for estimation of hepatic fat, especially when MRS is not available. Although unenhanced CT shows limited diagnostic performance in detecting mild steatosis, it certainly has a role in diagnosing moderate-to-severe hepatic steatosis, such as evaluating donor candidates for living donor liver transplantation. USG, using both the traditional four-grade visual assessment and elastography in the present form, appears to have limited role in liver fat quantification.

Conflict of Interest
None declared.

References