Assessment of Accuracy, Repeatability, Reproducibility and Robustness of Fat Quantification in a Water-Fat Phantom

H. Yu[1], C. D. Hines[2], A. Shimakawa[1], C. A. McKenzie[3], S. B. Reeder[4], and J. H. Brittain[5]

[1]Global Applied Science Laboratory, GE Healthcare, Menlo Park, CA, United States, [2]Departments of Radiology, Biomedical Engineering, University of Wisconsin, Madison, Wisconsin, United States, [3]Department of Medical Biophysics, University of Western Ontario, London, Ontario, Canada, [4]Departments of Radiology, Medical Physics, Biomedical Engineering, University of Wisconsin, Madison, WI, United States, [5]Global Applied Science Laboratory, GE Healthcare, Madison, Wisconsin, United States

Introduction Fat quantification using MRI has demonstrated clinical value in a variety of applications including fatty liver disease. Typically multi-echo chemical shift based water-fat separation methods are used. Corrections for various confounding factors such as spectral complexity of fat [1, 2], T2* relaxation [1, 3], and T1 bias [4] have been considered. Full validation of quantitative techniques requires evaluation of accuracy, precision (repeatability and reproducibility) and robustness. *In vivo* studies with comparison to reference standards such as biopsy require substantial effort, time and cost.

Water-fat phantoms are an alternative to provide preliminary evaluation of fat quantification techniques. In this work, we study the performance of a chemical shift based, T2* corrected and T1 independent fat quantification technique, quantitative IDEAL [2-5], using a water-fat phantom with fat-fractions ranging from 0 to 100%. In quantitative IDEAL, six echoes are collected from a 3D-SPGR acquisition, reconstructed to produce water, fat images and fat-fraction maps.

Methods An emulsified water-fat phantom was built using the procedure described by Hines et al [6], including 8 bottles with fat-fractions of 0%, 5%, 10%, 20%, 30%, 40%, 50%, 100%. The following scans were performed.

- Accuracy: the phantom was scanned with a "standard" abdomen protocol in the axial plane on a 3T scanner (HDx, GE Healthcare, Waukesha, WI): six echoes, 224x160x24, 36x29cm FOV, 10mm slice thickness, 8 channel cardiac coil, parallel imaging (R=1.78). Fat-fraction was measured by placing an ROI in each bottle in the fat-fraction images. The mean value in the ROI was recorded for all 8 bottles.
- Linearity: linear regression of the fat-fractions measured in all 8 bottles was performed, resulting in slope, intercept and r^2.
- Robustness: Robustness describes the invariance of the measured fat-fractions with respect to the imaging parameters. Based on the "standard" protocol, a number of scans were performed with a change of one imaging parameter at a time. Altered parameters included resolution, parallel imaging acceleration factor, number of echoes, scan orientation and receiver coils (Table 1).
- Repeatability: a scan with the "standard" protocol was repeated twice with the phantom repositioned and re-landmarked between the scans. The difference in fat-fraction measurements between the repeated scans and the "accuracy" scan was obtained to study the repeatability of the technique.
- Reproducibility: the fat-fractions were measured in a second 3T scanner (MR750, GE Healthcare, Waukesha, WI) as well as a 1.5T scanner (HDx, GE Healthcare). The fat-fraction values measured on all 3 scanners were compared to study the reproducibility.

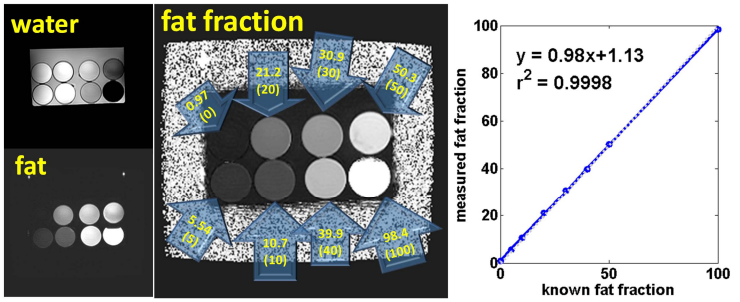

Figure 1: Quantitative IDEAL separated images of a 3T scan using the "standard protocol". The measured and known (in parenthesis) fat-fraction (%) are labeled. Correlation between measured and known fat-fraction is plotted.

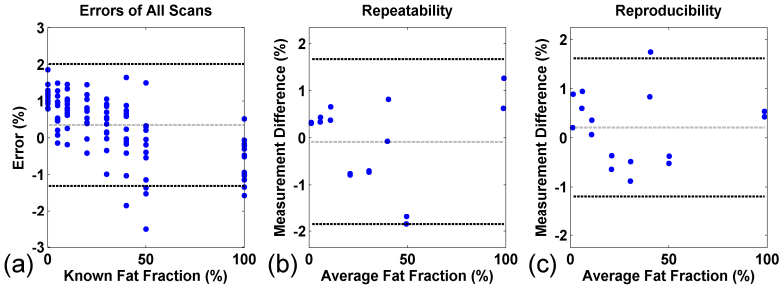

Figure 2: (a) the distribution of errors from all scans; (b) Bland-Altman plot for the repeatability study (the difference of repeated scans and the "accuracy" scan); (c) Bland-Altman plot for the reproducibility study (the difference in measurements between different scanners). In all plots, dotted line represents the mean of the distribution (0.35, -0.09, 0.21 respectively) and the dashed line denotes Limit of Agreement, which are [-1.31, 2.01], [-1.85, 1.68] and [-1.19, 1.62] respectively.

Results Figure 1 shows water, fat and fat-fraction images obtained from the "accuracy" scan with the "standard" protocol. The measured and known (in parentheses) fat-fraction values in all 8 bottles demonstrated excellent agreement. Table 1 lists the fat-fractions measured in all scans in the format of mean absolute errors among the 8 bottles, max absolute errors and the correlation parameters. The distribution of errors (difference of measured and known fat-fraction) from all bottles in all scans is plotted in Figure 2(a). To study the precision of the measurement, the difference of fat-fractions between the "standard" scan and the repeated scans is calculated to evaluate repeatability, while the difference in fat-fraction measured on two different scanners (3T scanner 1 vs 3T scanner 2, 3T scanner 1 vs. 1.5T scanner) is calculated for evaluating reproducibility. The Bland-Altman plots of repeatability and reproducibility demonstrated Limit of Agreement for precision well within 2%.

Scan	Mean Error (%)	Max Error (%)	slope	Intercept (%)	r^2
Accuracy: standard	0.78	1.57	0.98	1.13	0.9998
Repeatability: standard 1	0.88	1.52	0.98	1.05	0.9996
Repeatability: standard 2	0.74	1.36	0.98	0.91	0.9996
Robustness: 2D PI accelerated	0.69	1.45	0.98	1.04	0.9998
Robustness: 8 echoes	0.60	1.84	1.00	-0.24	0.9995
Robustness: 192x192 matrix	0.59	1.34	0.98	0.75	0.9998
Robustness: slice thickness = 8mm	0.66	1.00	0.98	0.96	0.9999
Robustness: coronal	0.96	1.46	0.98	1.26	0.9996
Robustness: sagittal	0.78	1.11	0.98	1.08	0.9999
Robustness: oblique	0.90	2.51	0.99	0.33	0.9987
Robustness: 8 Channel CTL coil	0.78	1.50	0.99	0.94	0.9999
Reproducibility: 3T scanner 2	0.92	1.86	0.98	1.42	0.9996
Reproducibility: 1.5T	0.83	1.18	0.98	1.26	1.0000

Table 1: Measured fat-fractions in the format of mean absolute errors in 8 bottles, maximum absolute errors and correlation parameters for all scans.

Discussion and Conclusion Accuracy, precision (including repeatability and reproducibility) and robustness are important measures of a quantitative imaging method. In this work, we used phantom studies to evaluate the performance of a quantitative IDEAL technique. While such studies do not eliminate the need for in vivo validation, well designed phantom experiments allow easy assessment of the "ground truth", manageable scan repeats and flexible scan configurations such as use of different coils and scanners, thus they provide valuable initial validation. We have demonstrated that quantitative IDEAL is highly accurate, repeatable, reproducible and robust in water-fat phantoms. It has potential to offer an accurate and precise MR imaging method to measure fat-fraction in a 0 ~ 100% range, independent of imaging parameters.

References

[1]. Bydder et al, MRM 2008;26(3):347-359.
[2]. Yu et al, MRM 2008: 60(5):1122-1134
[3]. Yu et al, JMRI 2007;26(4):1153-1161.
[4]. Liu et al, MRM 2007:58(2):354-64.
[5]. Hines et al, ISMRM 2010:263
[6]. Hines et al, JMRI 2009, 30:1215–1222