Multi-site, multi-vendor validation of accuracy, robustness and reproducibility of fat quantification on an oil-water phantom at 1.5T and 3T

Diego Hernando[1], Mustafa R. Bashir[2], Gavin Hamilton[3], Jean M. Shaffer[2], Samir D. Sharma[1], Claude B. Sirlin[3], Keitaro Sofue[2,4], Nikolaus M. Szeverenyi[3], Takeshi Yokoo[5,6], Qing Yuan[5], and Scott B. Reeder[1,7]

[1]Radiology, University of Wisconsin-Madison, Madison, WI, United States, [2]Radiology, Duke University, Durham, NC, United States, [3]Radiology, University of California, San Diego, San Diego, CA, United States, [4]Radiology, Kobe University Graduate School of Medicine, Kobe, Hyogo, Japan, [5]Radiology, University of Texas Southwestern Medical Center, Dallas, TX, United States, [6]Advanced Imaging Research Center, University of Texas Southwestern Medical Center, Dallas, TX, United States, [7]Medical Physics, University of Wisconsin-Madison, Madison, WI, United States

Target Audience: Scientists and clinicians interested in fat quantification and quantitative imaging biomarkers.

Purpose: Chemical shift-encoded techniques for MRI-based quantification of triglyceride concentration have shown great promise for diagnosis, quantitative staging and treatment monitoring of nonalcoholic fatty liver disease (NAFLD). Although multiple clinical validation studies have been conducted in recent years, demonstrating accuracy[1-5], precision[4,6] and robustness[7] of these techniques on different platforms, direct validation of reproducibility in phantoms across multiple sites, vendors, platforms and field strengths has yet to be performed. ***The purpose of this work*** was to test the accuracy, robustness and reproducibility of proton-density fat-fraction (PDFF) measurements acquired on the same phantom at four sites and using three MRI vendors, both at 1.5T and 3T.

Methods: *Phantom and imaging sites:* An agar-based oil-water phantom consisting of 11 cylindrical vials (diameter=25 mm, height=90 mm) with multiple peanut oil concentrations (PDFF = 0%, 2.6%, 5.3%, 7.9%, 10.5%, 15.7%, 20.9%, 31.2%, 41.3%, 51.4%, 100%, adjusting for water volume loss) was constructed as described in previous work[8]. The same phantom was sequentially shipped to and scanned at four different imaging sites. Scanner vendors included GE, Philips, and Siemens, each with 1.5T and 3T platforms (Table 1).

Site (vendor)	1.5T scanner	3T scanner
Site 1 (GE)	HDxt	MR750
Site 2 (Siemens)	Aera	Tim Trio
Site 3 (Philips)	Achieva	Ingenia
Site 4 (GE)	HDxt	HDxt

Table 1: *Scanners used in this study*

Imaging protocol: Vials were placed contiguously on the scanner table, parallel to the main magnetic field. Fat quantification acquisitions were performed at both 1.5T and 3T using each site's version of a multi-echo 3D spoiled gradient echo (SGRE) sequence, including two different protocols (to test robustness to varying acquisition parameters), as described in Table 2.

	1.5T Protocol 1	1.5T Protocol 2	3T Protocol 1	3T Protocol 2
Number of echoes	6	6	6	6
TE_{init}	2.30ms	1.20ms	1.15ms	1.20ms
ΔTE	2.30ms	2.00	1.15ms	1.00ms
Slice thickness	4mm	4mm	4mm	4mm
Flip angle	3°-5°	3°-5°	2°-3°	2°-3°

Table 2: *Protocols (3D multi-echo SGRE) used in this study*

PDFF reconstruction and analysis: Complex images were sent to a central site (Site 1) for reconstruction of PDFF maps, including correction for multi-peak fat (liver fat spectrum derived by Hamilton et al)[9,10], T2* relaxation[6], eddy currents[11,12], and temperature-related frequency shifts[8]. The same offline reconstruction algorithm was applied to the data acquired on each scanner at each of the four sites. For each PDFF map, measurements were performed by placing an ROI (~3cm^2) on each of the vials, averaging over the three central slices within the phantom. PDFF measurements from each vial were compared to the true PDFF using linear regression analysis. Further, PDFF measurements were compared across sites, vendors, field strengths and protocols (to test reproducibility and robustness) using the two-way random, single-measure intra-class correlation coefficient (ICC).

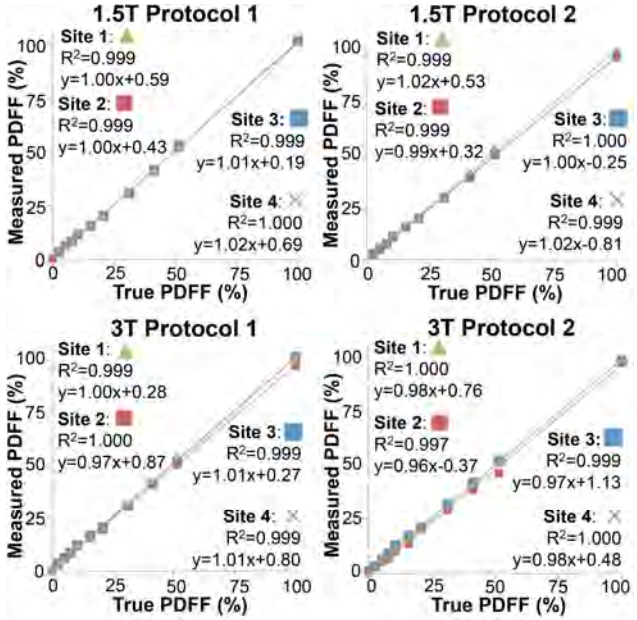

Figure 1: *Phantom PDFF measurement results, showing accurate, robust and reproducible fat quantification.*

Results: PDFF maps from all sites were reconstructed successfully. Correlation results are shown in Figure 1. PDFF measurements agreed closely with the true PDFF, demonstrating accuracy at all sites. The measured ICC over all sites, vendors, field strengths and protocols was ICC=0.9987, with 95% CI=0.9972-0.9996. Close agreement in PDFF was observed across protocols (demonstrating robustness to changes in acquisition parameters) as well as across sites, platforms and field strengths (demonstrating reproducibility), over the entire range of PDFF (0-100%).

Discussion & Conclusion: Development of quantitative imaging biomarkers requires validation across different vendors, sites and platforms. This work demonstrates excellent accuracy, robustness and reproducibility of confounder-corrected fat quantification techniques in an oil-water phantom across three vendors, four sites, two field strengths, and eight magnets. In vivo multi-center studies are needed to extend these results in patients.

References: [1]Yokoo et al, Radiology 251:67-76, 2009. [2]Yokoo et al, Radiology 258:749-759, 2011. [3]Meisamy et al, Radiology 258:767-775, 2011. [4]Hines et al, JMRI 33: 873–881. 2011. [5]Zhong et al, MRM 72: 1353–1365, 2014. [6]Negrete et al, JMRI 39:1265-1271, 2014. [7]Hansen et al, JMRI 30:151-157, 2012. [8]Hernando et al, MRM 72:464-470, 2013. [9]Yu et al, MRM 60:1122-1134, 2008. [10]Hamilton et al, NMR in Biomed 24:784-790, 2011. [11]Bydder, MRM 26:347-59, 2008. [12]Yu et al, MRM, 66:199-206, 2011.

Acknowledgements: We acknowledge the following NIH grants: R01DK083380, R01DK088925, R01 DK100651, K24DK102595. We also thank GE Healthcare, Siemens, and Philips for their support.