Reproducibility and diagnostic accuracy of in vivo proton magnetic resonance spectroscopy in detection of hepatic steatosis

J. Qi[1], M. Gönen[2], J. Shia[3], L. Schwartz[4], N. Kemeny[5], M. D'Angelica[6], Y. Fong[6], J. Koutcher[1], and K. Zakian[1]

[1]Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY, United States, [2]Epidemiology-Biostatistics, Memorial Sloan-Kettering Cancer Center, [3]Pathology, Memorial Sloan-Kettering Cancer Center, [4]Radiology, Columbia University Medical School, [5]Medicine, Memorial Sloan-Kettering Cancer Center, [6]Surgery, Memorial Sloan-Kettering Cancer Center

OBJECTIVES: Accumulation of fat in liver is associated with pathologies including obesity, diabetes, metabolic syndrome, and liver toxicity (1). Non-invasive assessment of hepatocellular triglyceride content (HTGC) is essential for serial analysis of steatosis associated with pathology. ^{1}H MRS is an accurate *in vivo* technique for measuring HTGC (2). The aim of this study was to assess the reproducibility and accuracy of HTGC measurements on a 1.5T scanner using non-invasive, single-voxel ^{1}H MRS during a single breath-hold.

METHODS: Experiments were performed on a 1.5 Tesla GE Signa scanner using a 4-channel torso phased array coil (G.E., Milwaukee, WI). A non-water-suppressed ^{1}H MRS was acquired from a $20 \times 20 \times 20$ mm^3 voxel in the right liver using PRESS with TR=5s, TE=40ms, spectral width (SW) = 2000 Hz, points = 512, and 4 acquisitions acquired during a single breath-hold. The voxel size was chosen so that major vessels and ducts could be easily avoided. Automated shimming on the voxel was also performed during a breath-hold. Water and lipid peak areas were measured using jMRUI (3). The fat peak area (PA) was measured after water was filtered using singular-value decomposition. Measured peak areas were corrected for water and methylene lipid T_2 values calculated from 10 subjects. HTGC was reported as: FFW = $PA_{fat}/(PA_{water} + PA_{fat})$. **Reproducibility:** To investigate the reproducibility of FFW measurements, 16 heathy adults (11 men and 5 women) were studied twice with a 2 week interval. Median subject age was 35 years (range 22 to 60 years), and median subject BMI was 22.2 kg/m^2 (range 19.7 to 34.7 kg/m^2). Within a single exam, two spectra were acquired within one minute of each other to provide reproducibility data under optimum conditions. Intra- and inter- exam coefficients of variance (CV) were calculated. **Accuracy:** To evaluate the accuracy of ^{1}H MRS in quantitation of HTGC, FFW from 16 colorectal cancer patients were correlated with liver histology assessments. Tissue specimens were obtained either from wedge or segmental resection (n=12) or from liver biopsy (n=4). Technical limitations prevented spatial matching of the liver tissue sampled to the MRS voxel location. Median patient age was 51 years (range 34 to 65 years), and median patient BMI was 26.5 kg/m^2 (range 20.5 to 35.3 kg/m^2). Histologic steatosis was graded on a 0-5 scale: 0) no fatty cells, 1) fatty cells in the centrilobular region constituting <25% of the tissue, 2) fatty liver cells involve 25-50% of the tissue, 3) fatty cells in 50-75% of the tissue, and 4) fatty cells in >75% of the tissue.

RESULTS: The mean FFW in healthy volunteers was 0.013 with a range from undetectable lipid to 0.060. All but 1 subject had FFW less than 0.02. As expected, the intra-exam reproducibility was very high (Fig.1a) with a coefficient of variance (CV) of 6%. The 2-week inter-examination data also showed a high degree of correlation (R = 0.963) (Fig.1b). The inter-exam coefficient of variation measured from 16 subjects was 18% which was substantially higher than the intra-exam CV. This suggests that moving the subject and repositioning the voxel as well as inherent heterogeneity in liver lipid can influence serial FFW measurements. We cannot rule out possible limitations caused by low lipid SNR. The subject with the largest percentage variation at 2 weeks had FFW1 = 0.014 and FFW2 = 0.006. The latter value only slightly exceeded the threshold of measurability due to SNR considerations. Shimming may have played a part here since water FWHM1 was 0.18 ppm while FWHM2 was 0.24 ppm. With this outlier removed, the inter-exam CV was 14.9% and R = 0.976. Fig. 2 relates histologic steatosis grade to the corresponding ^{1}H MRS-measured FFW. A positive correlation was observed between FFW and steatosis grade (r = 0.78, *P* < 0.001). ^{1}H MRS could not differentiate grades 0 and 1. Overlap in grades 2 and 3 could be partially due to heterogeneity of steatosis and tissue sampling limitations. If clinically significant steatosis is defined as grade 2 or above, a FFW cutoff value of 0.025 results in 86% sensitivity and 89% specificity for detection of clinically significant steatosis.

DISCUSSION and CONCLUSIONS: The inter-exam CV of FFW measured in an 8 cm^3 voxel during a breath-hold on a 1.5T magnet is approximately 15% in healthy subjects with low HTGC. Previous studies reported inter-examination CV values in the 7-11% range and intra-examination CV values similar to ours (4,5). It's possible that non-breath-held techniques used in those studies may "average in" some lipid spatial heterogeneity, enhancing reproducibility. ^{1}H MRS-measured FFW correlates with histologic steatosis grade and is an accurate indicator of clinically relevant steatosis in colorectal cancer patients.

1 Marchesini G, et al. Am J Med 1999; 107: 450-5.   2 Van Werven JR, et al. Radiology 2010; 245: 159-68.
3 Vanhamme L, et al. J Magn Reson 1997; 129:35-43.   4 Thomas EL, et al. Gut 2005; 54: 122-7.
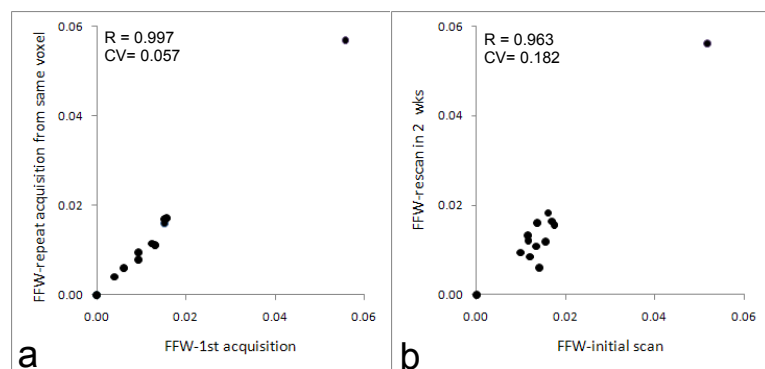5 Szczepaniak LS, et al. Am J Physiol Endocrinol Metab 2005; 288: E462-8

Fig. 1. FFW Intra- examination (1a) and inter-examination (1b) reproducibility in 16 healthy volunteers. Subjects with lipid undetectable above the noise threshold were assigned FFW=0.
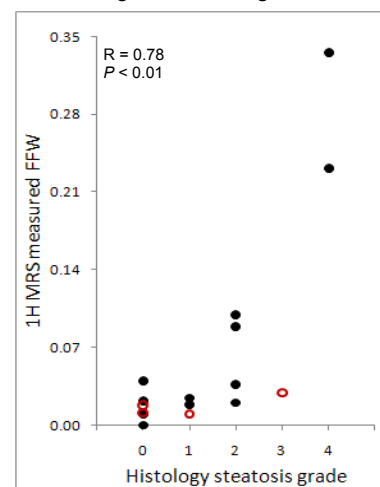

Fig. 2. FFW vs. liver steatosis grade in 16 patients. Open red circles indicate tissue from needle biopsy.