Reliability Analysis of Liver Apparent Diffusion Coefficient Measurement: Importance of ROI Size and Image Threshold

M. Eatesam[1], M. Nystrom[1], S. M. Noworolski[1,2], J. L. Dodge[3], R. B. Merriman[4], and A. Qayyum[1]
[1]Radiology and Biomedical Imaging, UCSF, San Francisco, CA, United States, [2]Graduate Group in Bioengineering, UC San Francisco and Berkeley, San Francisco and Berkeley, CA, United States, [3]Department of Internal Medicine, UCSF, Fresno, CA, United States, [4]Department of Medicine, California Pacific Medical Center, San Francisco, CA, United States

INTRODUCTION: Liver biopsy is used to detect and stage liver fibrosis, but it is invasive and limited by sampling error due to the heterogeneous nature of fibrosis [1,2]. Because of such limitations, alternative, non-invasive techniques are under investigation. The feasibility of using diffusion-weighted imaging (DWI) has been of recent interest but in order to perform a quantitative biological study, it is necessary to assess the reliability of the technique [3]. To such an end studies addressing repeatability and reproducibility are performed. Repeatability reflects "intra-observer agreement" and is a prerequisite to reproducibility. Liver ADC maps suffer from bright and dark artifacts due to blood vessels which may degrade measurements. The purpose of this study was to compare liver ADC measurement repeatability with multiple image analysis techniques designed to reduce the impact of the blood vessels in the liver ADC maps while studying patients with biopsy-proven nonalcoholic fatty liver disease (NAFLD) and healthy volunteers.

METHODS: Two successive breath-hold DWI scans were performed at 1.5T in 14 subjects: 6 patients with biopsy proven NAFLD and 8 healthy volunteers, (9 women, 5 men; mean age 41 years; range, 23-60). An echo-planar DWI sequence was used: TR/TE, 2000/min; matrix size, 128x128; FOV, 40-48 cm; slice/gap 10/2mm; number of signals averaged, 3; 9 slices; ASSET acceleration, 2; b-values, 0, 600. Subjects were removed from the scanner for 45 minutes between scans. ADC was derived from 6 small and 3 large regions of interest (ROIs) and with 4 different image threshold applications. Average right liver lobe ADC was obtained at 3 levels: superior, inferior and at the level of the portal vein using small 1.5cm diameter ROIs placed in the anterior and posterior segments avoiding major vessels and also using a single large ROI occupying most of the right lobe including vessels [Fig 1]. For large ROIs, 4 thresholding techniques were applied to ADC maps: 1) exclusion of ADC values below the 10th percentile or above the 90th percentile for each subject; 2) exclusion of ADC values falling beyond 1 standard deviation (SD) for each subject; 3) exclusion of ADC values falling beyond 2 SD for each subject; 4) exclusion of ADC values below the average dark (low signal intensity, SI) vessel ADC and above the average bright (high SI) vessel ADC [Fig 1 & 2]. Intraclass correlation coefficients (ICC) were determined with Shrout-Fleiss reliability test. Technique bias was assessed using Bland-Altman plots and ordinary least-squares regression [4].

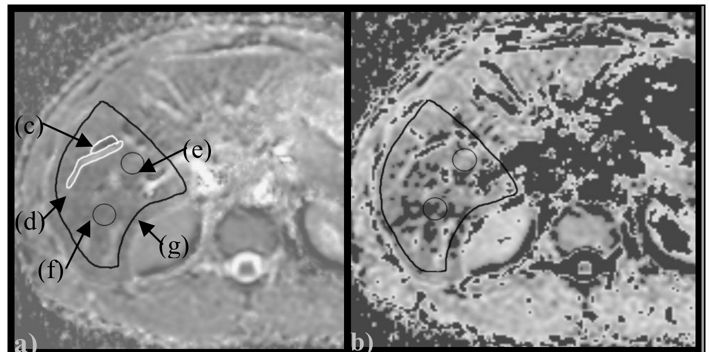

Figure 1. ADC maps a) before and b) after thresholding with ROIS for the dark (c) and bright (d) vessels, and the anterior (e), posterior (f), and large right lobe (g) liver ROIs

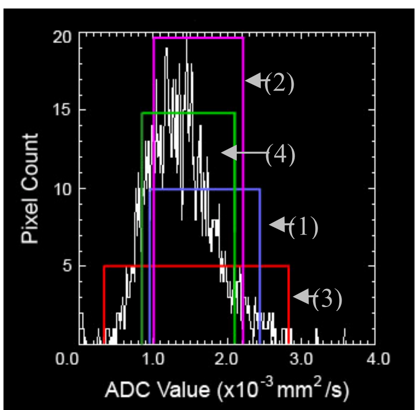

Figure 2: Histogram of a large ROI displaying the ADC ranges for the different thresholding techniques

Table 1. ICC, fixed bias (FB) per Bland-Altman plots, and proportional bias (PB) for threshold applications.

Analysis	ICC	FB	95% CI	PB	p-value
Small ROI	0.79	No	-0.23 - 0.25	No	0.40
Large ROI	0.88	No	-0.22 - 0.20	No	0.10
10th-90th percentile	0.89	No	-0.21 - 0.18	No	0.09
1 SD	0.89	No	-0.21 - 0.18	No	0.09
2 SD	0.90	No	-0.20 - 0.18	No	0.07
Vessel correction	0.90	No	-0.13 - 0.11	Yes	0.01

RESULTS: ADC repeatability between successive scans was lower with small ROIs (ICC 0.79) compared to large ROIs (ICC 0.88). Repeatability of large ROI ADC measurements was good with and without threshold applications (ICC 0.89 to 0.90); Bland-Altman plots showed no significant disagreement between successive scans for all threshold applications, although vessel correction demonstrated the narrowest range of differences between measurements, implying greatest repeatability [Table 1]. A proportional bias was detected with vessel threshold correction.

DISCUSSION: Our results showed greater repeatability with large rather than small ROI ADC measurements. Vessel correction was associated with the narrowest range of ADC difference between scans, suggestive of highest repeatability. However a proportional bias was observed which should be explored further. Additionally, assessment of the relative repeatability of low and high b value ADC is required due to the different relative role of perfusion in these measures and the potentially differing concomitant impact of a thresholding technique.

REFERENCES: 1- Friedman, SL. J Hepatol, 2003. 38(S1): S38-53. 2- Bedossa P, Dargere D, Paradis V. Hepatology 2003;38(6):1449-1457. 3- P.F. Watson, A. Petrie. Theriogenology 2010; 73:1167–1179. 4- Bland, Altman. Intl J Nurs Stud. 2010.