Reproducibility, intra- and inter-observer variability of ADC measurement by volumetric segmentation of bone marrow in whole body diffusion-weighted imaging (WB-DWI)

Sharon L Giles[1], Catherine J Simpkin[1], David J Collins[1,2], Christina Messiou[1], and Nandita M deSouza[1,2]

[1]*MRI Department, Royal Marsden Hospital, Sutton, Surrey, United Kingdom,* [2]*Clinical Magnetic Resonance, Institute of Cancer Research, Sutton, Surrey, United Kingdom*

Target Audience: Radiologists, radiographers, physicists and clinicians with interest in Whole Body Diffusion Weighted MRI (WB-DWI).

Purpose: Current response criteria in myeloma do not recommend a role for imaging, although it is acknowledged that further functional techniques should be evaluated (1). Whole-body diffusion weighted imaging (WB-DWI) is emerging as a quantitative tool to assess response to treatment in myeloma (2) but analysis methods, including observer dependent segmentation of the skeleton, remains without consensus. In patient studies, changes in ADC indicative of treatment response may be relatively small and therefore the degree of variability in the measurement method is crucial. The purpose of this study was to evaluate the reproducibility, intra- and interobserver variability of the analysis using a segmentation technique in a cohort of healthy volunteers.

Methods: 8 healthy volunteers were scanned at 2 time-points within one week (TP1 and TP2). Volunteers were selected on the basis of having visible marrow on DWI. Scans were undertaken with a Siemens Magnetom Avanto 1.5T system with Total Imaging Matrix to allow multiple surface coil elements to be used in conjunction. The acquisition covered the abdomen and pelvis in 2 stations of the same single shot double spin echo echo-planar DW sequence using STIR fat suppression in free-breathing in blocks of 50 slices (slice thickness 5mm, no gap, FOV 430mm, phase direction AP, GRAPPA factor 2, TR 14800 ms, TE 66 ms, TI 180 ms, voxel size 2.9x2.9x5 mm, 4 NSA, matrix 150 x 150, bandwidth 1960 Hz per pixel). Diffusion gradients with b values of 50 and 900 s/mm^2 were applied in 3 orthogonal directions and averaged to provide isotropic trace images.).In patient studies we have previously used this same sequence over 5 stations to provide whole skeleton coverage. Siemens OncoTreat software was used to generate Apparent Diffusion Coefficient (ADC) histograms for TP1 and TP2 from co-registered volumetric segmentations of visible marrow within the thoraco-lumbar spine and pelvis on the b900 s/mm^2 images for each volunteer. Any change in shape or position of histograms was summarised by calculating the percentage mean ADC change and the t statistic (difference in means ÷ standard error of the means) for each volunteer. These segmentations were first undertaken by Observer 1 (1a), then again on a separate occasion by the same observer (1b) to evaluate intra-observer variability. Inter-observer variability was assessed by a third segmentation of the data by a second observer (2a). Reproducibility of the ADC measurements was assessed by Bland Altman analysis, whilst intra- and inter-observer variability were assessed with intra-class correlation coefficients.

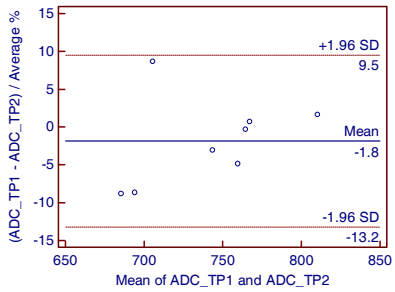
Results: Data presented in ***Table 1*** show that there was some variation in the volume of visible narrow that could be segmented within the spine and pelvis between volunteers. The ADC measurement was, in some volunteers, found to be highly reproducible, with a coefficient of repeatability of 81.7% for the whole cohort, illustrated in ***Figure 1a***. The measurements were also highly consistent, with intraclass correlation coefficients (ICC) of 0.981 (95% confidence interval (CI): 0.907-0.996) and 0.991 (95% CI: 0.956-0.998) for intra- and inter-observer comparisons respectively. Graphical illustration of the inter-observer reproducibility is given in the Bland Altman plot at ***Figure 1b***. The t statistic was also calculated consistently, with an ICC of 0.995 (95% CI: 0.976-0.999).

Table 1: ADC measurements

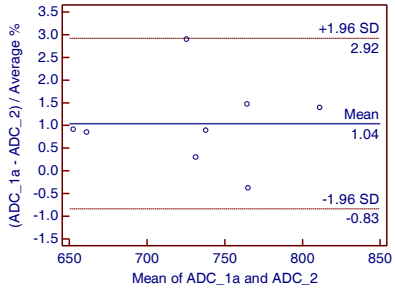
Volunteer	Observer	No. pixels	Mean ADC TP1 (x10^{-6}mm^2/s)	Mean ADC TP2 (x10^{-6}mm^2/s)	% Change Mean ADC	t
1	1a	51943	655.5	715.6	9.2	-48.8
	1b	50099	658.9	716.6	8.8	-43.4
	2	48557	649.5	726.1	11.8	-56.8
2	1a	29424	816.7	803.7	-1.6	4.9
	1b	36516	806.5	793.2	-1.6	5.6
	2	39315	805.4	793.7	-1.5	5.2
3	1a	44622	741.3	777.9	4.9	-20.9
	1b	42181	731.9	767.3	4.8	-21.5
	2	43752	734.7	770.2	4.8	-21.8
4	1a	39292	664.0	724.2	9.1	-33.0
	1b	41586	660.9	722.9	9.4	-36.3
	2	37809	658.4	715.5	8.7	-32.8
5	1a	46233	763.5	765.8	0.3	-1.1
	1b	49178	785.6	787.2	0.2	-0.8
	2	48597	766.4	770.1	0.5	-1.9
6	1a	63393	769.9	764.5	-0.7	3.6
	1b	59739	762.6	757.2	-0.7	3.6
	2	61964	758.6	754.0	-0.6	3.5
7	1a	44947	732.3	754.7	3.1	-14.6
	1b	44474	728.8	752.0	3.2	-15.5
	2	43465	730.1	757.0	3.7	-17.8
8	1a	55416	736.1	675.1	-8.3	37.8
	1b	51906	731.5	671.3	-8.2	35.8
	2	46222	715.1	657.6	-8.0	35.7

Figure 1: Bland Altman Plots

(a) ADC Reproducibility



(b) Inter-observer Reproducibility



Discussion and Conclusions: The reproducibility of the ADC measurement was mainly extremely good; in 3 volunteers results were in line with phantom measurements made at this institution over the same time period (mean coefficient of variation 1.2%), in 2 there was <5% variation in the measurement and a <10% variation in the remaining 3. It is unknown whether these larger variations were due to inconsistencies in data acquisition or some biological factor in the volunteers. The segmentation process used did not introduce additional variation across or between observers and provided a consistent and reliable method to generate ADC histograms. This technique is therefore potentially of value in clinical assessment of treatment response in myeloma patients.

References: (1) Rajkumar et al, 2011, *Blood* ,117(18), 4691–4695 ; (2) Horger et al, 2011, *AJR*, 196, 790-795

Acknowledgments: CRUK and EPSRC Cancer Imaging Centre in association with the MRC and Department of Health (England) grant C1060/A10334, also NHS funding to the NIHR Biomedical Research Centre, Dr. P. Gall & Dr. T Feiweier at Siemens Healthcare for provision of the OncoTreat software and WIP DWI sequence & Dr. M. Blackledge at Institute of Cancer Research for optimisation of shimming techniques used in the acquisition.