MULTI-SITE LIVER TUMOUR ADC REPRODUCIBILITY AT 1.5 T

Ryan Pathak[1], Hossein Ragheb[2], Neil A Thacker[2], David Morris[2], and Alan Jackson[1]

[1]*The Wolfson Molecular Imaging Centre, University of Manchester, Manchester, United Kingdom,* [2]*Centre for Imaging Sciences, University of Manchester, Manchester, United Kingdom*

Target audience – This work is relevant to: a) investigators who wish to use Apparent Diffusion Coefficient (ADC) measurements as a quantitative biomarker for tumour response following drug therapy; and b) clinical oncologists who need to monitor disease progression in an individual patient.

Purpose – The QuIC-ConCePT project is a multi-site European study to develop clinical imaging biomarkers. ADC, calculated from Diffusion-Weighted MRI (DWI), measures the water mobility within tissue, an indirect quantification of cellular density, and degree of liquefaction. A tumour undergoes cell death and necrosis following successful treatment. Significant ADC change may indicate a shift in cell density towards necrosis, and therefore be a biomarker candidate for early treatment response. For use in the clinical setting of personalised medicine, there must be universal confidence in the accuracy of ADC and its interpretation. This requires development of standardised acquisition methods and an understanding of ADC measurement reproducibility, in the setting of multi-site trials, so that significant change in ADC is measured with confidence. Variables with influence on measurement accuracy must be identified and possibly reduced. This will reduce the cost of future drug trials and improve decision making for patients. Colorectal liver metastasis is an ideal candidate to examine ADC, as within the setting of a single-site study with treatment monitoring, reproducibility has been poor[1]. Within the liver, a proportion of this can be attributed to respiratory motion, and is yet to be quantified.

Methods – Following ethical approval and with informed consent, 4 sites acquired patient data using equivalent scanning protocols on 1.5T scanners, Seimens, GE and Philips[2]. ADC was calculated using mono-exponential fits from 3 b-value DWI images (b-100, 500, 900 s/mm^2). 19 patients (4 from site A, 5 from B, C and E) with untreated colorectal liver metastasis were scanned twice within 14 days. Parametric maps of voxel-by-voxel ADC measurements (with Rician noise correction) were generated using the project software[3]. A trainee radiologist manually drew and matched regions of interest (ROIs) over metastatic tumour on b-100 images (Osirix platform), which were then stacked to generate a 3D volume. Mean ADC values were calculated from ROI ADC histograms. For comparison, 2D ROIs were also identified from the single slice that best represented tumour (2D-T), for the largest single slice region through a tumour (2D-L), and for a fixed-size ROI within normal liver parenchyma (2D-N). Previous evaluation studies have used coefficient of variation (CoV) and Bland-Altman plots for absolute values[4]. However, the interpretation of absolute mean ADC, in terms of degree of tissue cell density may be different from one data set to another. This needs to be taken into account if we do not wish to interpret instabilities in measurement as significant biological change. An error model was developed to predict the reproducibility of percentage change in ADC (ΔADC%) in individual tumours. Reproducibility may be affected by different sources of error such as signal to noise, imaging artifact, motion, and fixed errors from gradient non-linearity. ROI selection and ROI size can also affect reproducibility. In our model, the measurement error of ΔADC% is estimated from error propagation of the expected errors in each original mean ADC measurement. We used likelihood optimisation to fit the parameters of the error model. By factoring out differences in ADC corresponding to measurement errors, we can get a better understanding of reproducibility, and explore the effect of other parameters, e.g. motion, ROI size and vendor. After scaling ΔADC% by its estimated error, data points are directly comparable and we expect them to follow a t-distribution.

Results – 20% of data sets were visibly affected by motion artifact. These were identified as outliers when modeling the error on ΔADC%. In Fig. 1, CoV was lowest for 3D ROIs with the exception of site A (3.3%-17%). Considering all 3D data together, CoV was 7.5%. 2D ROIs with subjective slice selection to represent solid tumour exhibit the lowest performance. In Fig. 2, ΔADC% corresponding to 4 ROI methods and all patients are plotted against the corresponding ROI size (Log of number of voxels). The scatter of ΔADC% is as much as 20% depending on the ROI method. For all methods, as ROI size increases, ΔADC% improves. Fig. 3 shows our model of uncertainty in ΔADC% for all data sets. Larger ROIs, even with visible motion artifact, are up to 4 times more reproducible than some smaller ROIs. The number of voxels at which the improvement in reproducibility flattens out is about 2000 voxels, where errors smaller than 4% are achievable. In Fig. 4, ΔADC% has been scaled using the error model of Fig. 3. This improves our ability to quantify significant changes in ΔADC%, showing that the reproducibility measurements are quantitatively well predicted when samples with gross motion are excluded. It becomes clear that motion artifact is the major cause of poor reproducibility once other factors of measurement error are taken into account. The same way, we also found that there was little difference between vendors in terms of level of accordance with the average error model (unlike Fig. 1 which might suggest significant performance differences). Our findings for 2D-N ROIs are consistent with previous results[3].

Discussion – Tumours exhibit heterogeneity. To increase confidence in representative ADC metrics, factors affecting measurement of their reproducibility must be identified and quantified. For an individual, to detect significant change in ADC (e.g. 10% shift in the mean ADC within a ROI towards necrosis) the target reproducibility needs to be less than 4%. Previous single-site studies suggest measurement variability, at best of 10%, which is much larger than any early therapeutic effect[3]. Basic analysis of our multi-site, multi-vendor study suggests average CoV of 7.5% is achievable. We have shown however, that incorporation of a statistical model of expected errors is a sophisticated method to study reproducibility and understanding factors involved. ROI size is more important than the method used (3D, 2D-T or 2D-L). Volumes of around 2000 voxels exhibit improvement in reproducibility. By factoring out other sources of measurement error, we quantified reproducibility more explicitly, and so, improved interpretation of the results. Specifically, squares in Fig. 4 suggest that respiratory motion is the major factor causing poor reproducibility, especially at smaller volumes. Excluding outliers, 95% of data points fall within t=2, consistent with a t-distribution with 15 degrees of freedom.

Conclusion – Using mean ADC in the liver tumour, reproducibility of less than 4% is achievable but currently only for larger ROIs. Therefore, use of ROIs with a minimum number of voxels should be considered. Use of motion correction methods should make measurement of smaller ROIs (and smaller tumours) more viable.

Acknowledgements – Thanks to our collaborators who collected the data: the Institute of Cancer Research (London), Radboud University of Nijmegan, VU University Medical Centre of Amsterdam and the Chrisite Hospital Manchester. With thanks to David Collins, Naomi Douglas, Arend Heerschap and Joost PA Kuijer. Our study received support from the Innovative Medicines Initiative Joint Undertaking (www.imi.europa.eu) under grant agreement number 115151, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

Fig 1. CoV reproducibility for 4 ROI methods.

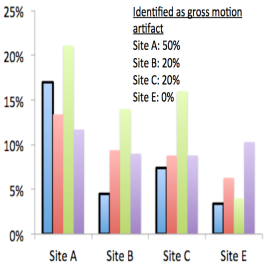


Fig 2. ΔADC% vs Log ROI size.

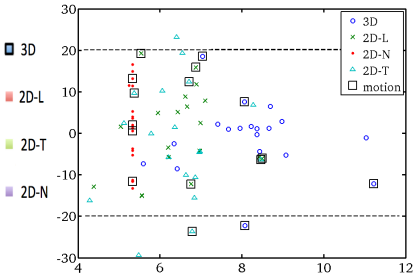


Fig 3. Estimated error on ΔADC% vs Log ROI size.

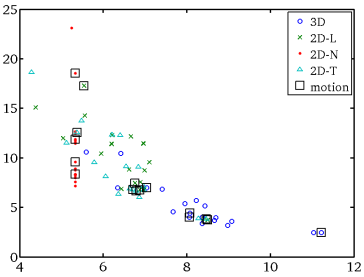
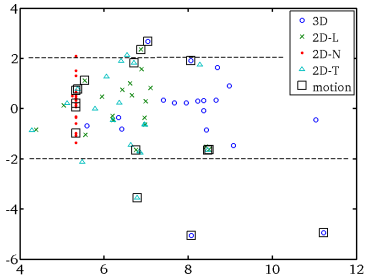


Fig 4. ΔADC% scaled by its error vs Log ROI size.



Reference: (1). Padhani AR et al. Neoplasia. 2009, 11(2):102-25. **(2).** Collins D. IMI WP3 Task 2.3 (Internal), 11/05/2012. **(3).** Ragheb H et al. ISMRM 2014, 22: 2650. **(4).** Deckers F et al. J Magn Reson Imaging. 2014, 40(2):448-56.