Published in final edited form as:

J Magn Reson Imaging. 2015 December ; 42(6): 1759–1764. doi:10.1002/jmri.24955.

Apparent Diffusion Coefficient Is Highly Reproducible on Preclinical Imaging Systems: Evidence From a Seven-Center Multivendor Study

Sabrina Doblas, PhD^{1,*}, Gilberto S. Almeida, PhD², François-Xavier Blé, PhD³, Philippe Garteiser, PhD¹, Benjamin A. Hoff, PhD⁴, Dominick J.O. McIntyre, PhD⁵, Lydia Wachsmuth, PhD⁶, Thomas L. Chenevert, PhD⁴, Cornelius Faber, PhD⁶, John R. Griffiths, FRCP, D Phil⁵, Andreas H. Jacobs, MD⁷, David M. Morris, PhD⁸, James P.B. O'Connor, FRCR, PhD⁹, Simon P. Robinson, PhD², Bernard E. Van Beers, MD, PhD¹, and John C. Waterton, PhD⁸

¹Laboratory of imaging biomarkers, UMR 1149 – CRI, Inserm, Paris Diderot University, Paris, France

²Division of Radiotherapy and Imaging, Institute of Cancer Research, London, UK

³Personalised Healthcare & Biomarkers, AstraZeneca, Macclesfield, UK

⁴Department of Radiology, University of Michigan, Ann Arbor, Michigan, USA

⁵CRUK Cambridge Institute, University of Cambridge, Cambridge, UK

⁶Department of Clinical Radiology, Münster University Hospital, Westfälische Wilhelms-Universität, Münster, Germany

⁷European Institute for Molecular Imaging, Westfälische Wilhelms-Universität, Münster, Germany

⁸Centre for Imaging Sciences, University of Manchester, Manchester, UK

⁹Institute of Cancer Sciences, University of Manchester, Manchester, UK

Abstract

Purpose—To evaluate between-site agreement of apparent diffusion coefficient (ADC) measurements in preclinical magnetic resonance imaging (MRI) systems.

Materials and Methods—A miniaturized thermally stable ice-water phantom was devised. ADC (mean and interquartile range) was measured over several days, on 4.7T, 7T, and 9.4T Bruker, Agilent, and Magnex small-animal MRI systems using a common protocol across seven sites. Day-to-day repeatability was expressed as percent variation of mean ADC between acquisitions. Cross-site reproducibility was expressed as $1.96 \times$ standard deviation of percent deviation of ADC values.

Results—ADC measurements were equivalent across all seven sites with a cross-site ADC reproducibility of 6.3%. Mean day-to-day repeatability of ADC measurements was 2.3%, and no site was identified as presenting different measurements than others (analysis of variance

* Address reprint requests to: S.D., CRI – UMR 1149 Inserm – Paris Diderot University, Faculté de médecine X. Bichat, 16 rue Henri Huchard, 75018 Paris, France. sabrina.doblas@inserm.fr.

[ANOVA] $P = 0.02$, post-hoc test n.s.). Between-slice ADC variability was negligible and similar between sites ($P = 0.15$). Mean within-region-of-interest ADC variability was 5.5%, with one site presenting a significantly greater variation than the others ($P = 0.0013$).

Conclusion—Absolute ADC values in preclinical studies are comparable between sites and equipment, provided standardized protocols are employed.

Biomarkers are objective measurements that inform about patient benefit, patient harm, or lack of benefit or harm.¹ Critical to their successful implementation in healthcare is the ability to validate and qualify the biomarker to a level where it can be used robustly to alter patient management.² That is the case for imaging biomarkers, whose value relies on their noninvasiveness and the extensive range of anatomical and functional information they provide. Some magnetic resonance imaging (MRI) techniques—such as diffusion-weighted MRI—are used routinely in healthcare as qualitative tests, but the same data can be used to derive quantitative parameters.³ Translation of these putative biomarkers has generally been hindered by lack of technical or biological validation and many have failed to alter clinical care.⁴ This technical validation largely relies on the assessment of parameter measurement repeatability (ie, agreement between measurements separated by a short period of time, with all experimental conditions kept the same) and reproducibility (ie, agreement between measurements with varying conditions: different observers, different machine vendors, different field strengths).⁵

A large number of preclinical and clinical studies are performed every year that require measurement of both pretreatment parameter values and a change in values from baseline. Changes in parameter values must be ascribed to either true physiopathological changes or the inherent biomarker variability. However, there has been little attempt to investigate reproducibility of MRI biomarkers, especially in a multisite multivendor context.⁴ Measurement consistency is extremely important for successful biomarker translation. Otherwise, the use of these imaging biomarkers will be restricted (at best) to single-center studies of early biomarker discovery or evaluation of phase I therapies at a single site.

Apparent diffusion coefficient (ADC) obtained from diffusion-weighted imaging (DWI) has been proposed as an imaging biomarker in several disease areas including cerebrovascular ischemia and oncology⁶ because of its potential added value for tumor characterization and therapy follow-up.³ For example, ADC has been shown to increase significantly in combretastatin-A4-phosphate-treated solid tumors,⁷ and DWI data can provide metrics for the assessment of early therapy response in brain tumors.⁸ For biomarker validation on clinical systems, standard phantoms such as ice-water flood phantoms have been devised, and consistent ADC measurements across sites have been reported.⁹ Similar validation is crucial on small-bore preclinical MRI systems, since clinical studies often require preclinical validation on animal models, usually done on these types of MRI scanners. A preclinical setting is optimal for refining the imaging timing and dose of the therapeutic agent, for the identification of confounding factors, and for understanding the relationship between the imaging biomarker and the physiopathological changes occurring in the target tissue. Besides, an imaging biomarker validated in a preclinical context (ie, with sufficient repeatability and reproducibility) will be more easily translated to clinical studies and

clinical practice, and this preclinical validation would ensure that absolute values can be employed and compared between sites rather than relative parameter changes. However, a full evaluation of preclinical cross-site agreement has never been performed. Hence, the goal of this study was to evaluate the test–retest repeatability and cross-site reproducibility of ADC measurements on several preclinical MRI scanners.

Materials and Methods

Phantom Setup

The phantom implemented was adapted from a previously validated ice-water clinical design.¹⁰ It consists of two concentric tubes: a 7-ml polypropylene tube with a plug-type push-in cap glued to the cap of a 50-ml polypropylene tube, and centrally located within it (Fig. 1). The phantom was prepared as follows: the outer tube was filled with distilled water and frozen overnight, with the inner tube positioned but not sealed. On the day of the MRI experiment, the inner tube was filled with distilled water kept at 4°C and the phantom sealed and put back in the refrigerator to stabilize for 20–30 minutes. The presence of ice in the outer tube was verified both before and after scanning. To assess the stability of the inner tube water temperature, sites 1 and 4 took temperature measurements using a probe (Small Animal Instruments, Stony Brook, NY) placed at the center of the inner tube, away from the wall. The temporal stability of ADC was also checked by site 1 over six imaging sessions with DWI acquisitions running over 30 minutes.

Sites Instrumentation

Seven sites participated in the study, as collaborators in the Innovative Medicines Initiative Consortium “QuIC-ConCePT”. Sites were equipped with preclinical MRI systems from three vendors, Bruker (Bruker BioSpin, Ettlingen, Germany), Agilent (Agilent Technologies, Santa Clara, CA), and Magnex (Magnex Scientific, Oxford, UK), and three field strengths were represented: 4.7T, 7T, and 9.4T (details in Table 1).

DWI Acquisition

After initial localization scans, DWI acquisition was conducted as follows. Sites were instructed to use the same coil configuration as they usually do for small-animal DWI studies of tumor models, and to choose their acquisition parameters to satisfy the following criteria: use of a spin-echo sequence preceded by shimming on the volume of interest, one slice centered on the magnet isocenter, volume of acquisition voxels of at most 0.2 mm³, echo time (TE) between 20 and 40 msec, repetition time (TR) greater than 1000 msec, and between two and six b values with at least two b values above 100 s/mm², for a scan duration of around 8 minutes (depending on the chosen sequence parameters). Acquisition parameters for each site are shown in Table 1. Each site performed between three and five imaging sessions over several days.

ADC Calculation

A circular region of interest (ROI) was drawn in the inner tube with an average margin of 6 pixels away from the edges of the tube, for a mean ROI size of 150 mm² across the sites. As ice can form in some cases in the inner tube, the minimum-b image was also checked to

ensure that the ROI did not include any ice, which would create susceptibility variations and anomalous ADC measurements. Each center was permitted to use its preferred postprocessing software and algorithm: a pixel-by-pixel monoexponential fitting method using a least square approach with ROOT (ROOT, CERN, Geneva, Switzerland; sites 1 and 2), R (R, Vienna, Austria; site 4), or MatLab (MathWorks, Natick, MA, USA; site 6), a log linearization approach using the Agilent console standard ADC fitting program (site 3), or MatLab (site 7) and a pixel-by-pixel Bayesian maximum a posteriori approach using ImageView (ImageView, IDL, ITT Visual Information Systems, Boulder, CO, USA; site 5). However, it was mandatory to calculate ADC values using only the b values above 100 s/mm^2 following recommendations from Padhani et al.³

Data Reporting and Statistics

For each imaging session an ADC averaged over the ROI was obtained, as well as the within-ROI interquartile range (IQR, expressed as percent variation of ADC). The signal-to-noise ratio (SNR) was also measured for each b image as the mean signal intensity within the phantom ROI divided by the standard deviation of signal intensity in the background in an artifact-free region. Within-site day-to-day repeatability was calculated as the percent variation of ADC between imaging sessions, and cross-site reproducibility was calculated as $1.96 \times$ standard deviation (SD) percent deviation of ADC between sites (Gaussian distribution of all ADC checked with a d'Agostino-Pearson omnibus normality test).⁹ To evaluate between-slice ADC variability, five sites (detailed in Table 2) acquired multislice DWI scans and, for each acquired scan, the coefficient of variation (CoV) of ADC between slices was computed (expressed as percent mean ADC). Analysis of variance (ANOVA) with Bonferroni post-hoc tests were conducted to test if sites differed by their ADC measurements, their %ADC variation within ROI, or their CoV of ADC between slices. Relationships between field strength and three variability parameters (day-to-day repeatability, within-ROI %ADC variation, and between-slice CoV of ADC), as well as between maximum SNR and these same three variability parameters were investigated using Spearman's correlations. Statistical significance was considered for $P < 0.05$. Statistical tests were performed using GraphPad Prism 5 (GraphPad Software, La Jolla, CA).

Results

Stability of the Phantom

At the beginning of the DWI acquisition, the average temperature measured in the inner tube water was $0.3 \pm 0.25^\circ\text{C}$ (measurements from sites 1 and 4). The monitoring of temperature for the whole duration of DWI acquisition (5 min 36 sec for site 1, 13 min 20 sec for site 4) showed good stability of the inner tube water temperature, with a mean increase of $0.1 \pm 0.1^\circ\text{C}$. A mean ADC variation of $0.8 \pm 0.5\%$ was observed after 20 minutes, and of $1.3 \pm 1\%$ after 30 minutes.

Overall ADC Reliability

The overall cross-site reproducibility of ADC was 6.3%, with ADC ranging from $1.15 \times 10^{-3} \text{ mm}^2/\text{s}$ to $1.33 \times 10^{-3} \text{ mm}^2/\text{s}$ (Table 2, Fig. 2). Mean day-to-day repeatability of ADC measurements was 2.3%, ranging from 0.6% to 5.9%. Overall, variances of ADC

measurements were different between sites but no site measurements were identified as significantly different from the others (ANOVA $P=0.02$, no significance at post-hoc Bonferroni test). Day-to-day repeatability was not affected by field strength ($P=0.84$, Spearman's correlation) or SNR ($P=0.5$).

Between-Slice and Within-ROI ADC Variability

Across all sites the IQR ranged from $0.02 \times 10^{-3} \text{ mm}^2/\text{s}$ to $0.19 \times 10^{-3} \text{ mm}^2/\text{s}$, corresponding to within-ROI ADC variations ranging from 1.5% to 14.7% (Table 2). ADC variations within the ROI were minimal (mean overall ADC variation of 5.5%), while being more pronounced at some sites (maximum mean within-ROI ADC variation of 10.1% observed at site 5, which exhibited significantly higher within-ROI ADC variation than sites 1, 2, 3, and 7 ($P=0.0013$)). ADC variability between slices was negligible, ranging from 0.1% to 2.1%, and was similar between sites ($P=0.15$). Field strength had no detectable effect on within-ROI ADC variation ($P=0.16$) or between-slice ADC variation ($P=0.83$). Likewise, SNR had no effect on between-slice ADC variation ($P=0.75$). However, within-ROI ADC variation was negatively correlated with SNR ($P<0.001$, $r=-0.68$).

Discussion

Investigators commonly measure relative change in ADC in response to therapy, to take into account systematic differences induced by hardware and software.³ However, multisite studies cannot be conducted using relative values, and absolute ADC values could be used for treatment stratification.³ The use of absolute ADC values requires this imaging biomarker to present good repeatability and reproducibility. DWI repeatability studies have been conducted on 1.5T and 3T clinical systems using phantoms,^{9,11} animal models of disease,^{12,13} and in both healthy volunteers¹⁴ and patients.^{15,16} However, such repeatability studies in preclinical systems remain scarce.^{17,18} Besides, there is little evidence that absolute ADC values, as opposed to relative changes, are comparable between sites, acquisition protocols, processing protocols, and MRI systems (ie, reproducibility studies).⁴ Only a few multisite DWI reproducibility studies have been conducted on clinical systems using phantoms and healthy volunteers.^{9,11,14}

Large, multicompartiment phantoms¹¹ are unsuitable for preclinical systems, and phantoms using viable isolated tissue or cells,¹⁹ while relevant for single-site studies, are not cost-effective and technically too complex to be shared between sites. Here we chose to build a smaller version of the iced water phantom used by Malyarenko et al⁹ for our multisite reproducibility study. We observed that this simple ice-water phantom provided stable ADC measurements, with very little interslice variation and an acceptable day-to-day repeatability, well within the range of what is reported in the few existing single-site reliability studies on preclinical systems.^{17,18} The magnitude of ADC changes usually reported in preclinical studies is on the order of 15% to 50%, when ADC is calculated over high b values only.^{20,21} This is above the highest test-retest reliability value observed within our study, which shows that absolute ADC values can be used in preclinical studies within a single site.

Seven sites conducted ADC measurements of this ice-water phantom following the same protocol for preparing the phantom and using similar DWI acquisition parameters on diverse

MRI systems. Cross-site reproducibility of ADC measurements was 6.3%, which is slightly higher than the values reported by Malyarenko et al⁹ (within 3% excluding a clear outlier, while none of the seven sites in this study were outliers), but remains above the usual range (15%–50%) of ADC changes reported in preclinical evaluations.^{20,21}

This level of variation, both intra- and intersite, would include factors such as variations in operator performance, hardware/software characteristics, and environment. We were less stringent on the choice of the coil, DWI acquisition parameters, and of the postprocessing method by the sites than Malyarenko et al were. The sites were to use acquisition parameters in an acceptable range, and were otherwise advised to do as they routinely do for preclinical DWI evaluations of a tumor model, to mimic the everyday situation. Acquisition protocols were overall quite similar. Sites exhibited a clear difference in SNR, but this did not correlate with day-to-day repeatability, and the only effect of SNR was on within-ROI ADC variation.

One limitation of our study would be that ice-water phantoms are more prone to thermal instability, which impacts ADC variability. While preliminary tests and temperature measurements showed that a relative stability could be obtained, this factor may have played a role in test–retest variability seen in some sites and in the overall cross-site reproducibility. Besides, we have not investigated other factors that may contribute to ADC variability, such as homogeneity and stability over time of the MRI magnetic field,²² and the spatial linearity of gradients.⁹ The latter would have implied doing off-center measurements, which is not practical due to the small diameter size of most preclinical MR scanners and coils. Finally, ADC variability measured using this type of phantom study has to be corrected for in vivo preclinical evaluations to take into account variability factors specific to the in vivo setting, such as the physiological variation between animals and the effect of respiratory and cardiac motion.

From the reproducibility measured here, we conclude that absolute ADC measurements in preclinical studies are comparable with a high degree of precision between sites and equipment, provided that standardized protocols and appropriate phantoms are employed. These results suggest that absolute ADC values (in addition to relative ADC changes) can be employed in animal models as a translational imaging biomarker in stratification and response assessment.

Acknowledgments

On behalf of the QuIC-ConCePT Consortium whose participants include: AstraZeneca, European Organisation for Research and Treatment of Cancer (EORTC), University of Cambridge, University of Manchester, Westfälische Wilhelms-Universität Münster, Radboud University Nijmegen Medical Center, Institut National de la Santé et de la Recherche Médicale, Stichting Maastricht Radiation Oncology "Maastricht Clinic," VUmc Amsterdam, King's College London, Universitair Ziekenhuis Antwerpen, Institute of Cancer Research – Royal Cancer Hospital, Erasmus Universitair Medisch Centrum Rotterdam, Imperial College of Science Technology and Medicine, Keosys S.A.S., Eidgenössische Technische Hochschule Zürich, Amgen NV, Eli Lilly and Company Ltd, GlaxoSmithKline Research & Development Limited, Merck KGa, Pfizer Limited, F. Hoffmann – La Roche Ltd, Sanofi-Aventis Research and Development. The research leading to these results has 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. This presentation reflects the authors' view and neither IMI JU nor the European Union and EFPIA are liable for any use that may be made of the information contained herein. This work was in part supported by the Münster interdisciplinary centre for clinical research (core

unit PIX), the QIN grant U01CA166104, the Manchester Cancer Research Centre, the Cancer Research UK Clinician Scientist grant R115246, the University of Cambridge, CRUK (grant C14303/A17197) and Hutchison Whampoa Limited. The Cancer Imaging Centre at ICR, and RMH was supported by Cancer Research UK, in association with MRC & Dept of Health, C1060/A16464, and NHS funding to the NIHR Biomedicine Research Centre and the Clinical Research Facility in Imaging.

References

1. Biomarkers-Definitions-Working-Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001; 69:89–95. [PubMed: 11240971]
2. Dorfman GS, Sullivan DC, Schnall MD, Matrisian LM. The Translational Research Working Group developmental pathway for image-based assessment modalities. *Clin Cancer Res.* 2008; 14:5678–5684. [PubMed: 18794075]
3. Padhani AR, Liu G, Koh DM, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia.* 2009; 11:102–125. [PubMed: 19186405]
4. Waterton JC, Pytkkanen L. Qualification of imaging biomarkers for oncology drug development. *Eur J Cancer.* 2012; 48:409–415. [PubMed: 22226478]
5. Kessler LG, Barnhart HX, Buckler AJ, et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Stat Methods Med Res.* 2015; 24:9–26. [PubMed: 24919826]
6. Sinkus R, Van Beers BE, Vilgrain V, Desouza N, Waterton JC. Apparent diffusion coefficient from magnetic resonance imaging as a biomarker in oncology drug development. *Eur J Cancer.* 2012; 48:425–431. [PubMed: 22226479]
7. Koh DM, Blackledge M, Collins DJ, et al. Reproducibility and changes in the apparent diffusion coefficients of solid tumours treated with combretastatin A4 phosphate and bevacizumab in a two-centre phase I clinical trial. *Eur Radiol.* 2009; 19:2728–2738. [PubMed: 19547986]
8. Moffat BA, Chenevert TL, Lawrence TS, et al. Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. *Proc Natl Acad Sci U S A.* 2005; 102:5524–5529. [PubMed: 15805192]
9. Malyarenko D, Galban CJ, Londy FJ, et al. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. *J Magn Reson Imaging.* 2013; 37:1238–1246. [PubMed: 23023785]
10. Chenevert TL, Galban CJ, Ivancevic MK, et al. Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies. *J Magn Reson Imaging.* 2011; 34:983–987. [PubMed: 21928310]
11. Lavdas I, Miquel ME, McRobbie DW, Aboagye EO. Comparison between diffusion-weighted MRI (DW-MRI) at 1.5 and 3 Tesla: a phantom study. *J Magn Reson Imaging.* 2014; 40:682–690. [PubMed: 24925470]
12. Song YS, Park CM, Lee SM, et al. Reproducibility of histogram and texture parameters derived from intravoxel incoherent motion diffusion-weighted MRI of FN13762 rat breast carcinomas. *Anticancer Res.* 2014; 34:2135–2144. [PubMed: 24778015]
13. Joo I, Lee JM, Yoon JH, Jang JJ, Han JK, Choi BI. Nonalcoholic fatty liver disease: intravoxel incoherent motion diffusion-weighted MR imaging-an experimental study in a rabbit model. *Radiology.* 2014; 270:131–140. [PubMed: 24091358]
14. Ye XH, Gao JY, Yang ZH, Liu Y. Apparent diffusion coefficient reproducibility of the pancreas measured at different MR scanners using diffusion-weighted imaging. *J Magn Reson Imaging.* 2014; 40:1375–1381. [PubMed: 24222019]
15. Corona-Villalobos CP, Pan L, Halappa VG, et al. Agreement and reproducibility of apparent diffusion coefficient measurements of dual-b-value and multi-b-value diffusion-weighted magnetic resonance imaging at 1.5 Tesla in phantom and in soft tissues of the abdomen. *J Comput Assist Tomogr.* 2013; 37:46–51. [PubMed: 23321832]
16. Heijmen L, Ter Voert EE, Nagtegaal ID, et al. Diffusion-weighted MR imaging in liver metastases of colorectal cancer: reproducibility and biological validation. *Eur Radiol.* 2013; 23:748–756. [PubMed: 23001604]

17. Whisenant JG, Ayers GD, Loveless ME, Barnes SL, Colvin DC, Yankeelov TE. Assessing reproducibility of diffusion-weighted magnetic resonance imaging studies in a murine model of HER2+ breast cancer. *Magn Reson Imaging*. 2014; 32:245–249. [PubMed: 24433723]
18. Suh JY, Cho G, Song Y, et al. Is apparent diffusion coefficient reliable and accurate for monitoring effects of antiangiogenic treatment in a longitudinal study? *J Magn Reson Imaging*. 2012; 35:1430–1436. [PubMed: 22314928]
19. Richardson S, Siow B, Batchelor AM, Lythgoe MF, Alexander DC. A viable isolated tissue system: a tool for detailed MR measurements and controlled perturbation in physiologically stable tissue. *Magn Reson Med*. 2013; 69:1603–1610. [PubMed: 22821404]
20. Joo I, Lee JM, Han JK, Choi BI. Intravoxel incoherent motion diffusion-weighted MR imaging for monitoring the therapeutic efficacy of the vascular disrupting agent CKD-516 in rabbit VX2 liver tumors. *Radiology*. 2014; 272:417–426. [PubMed: 24697148]
21. Schelhaas S, Wachsmuth L, Viel T, et al. Variability of proliferation and diffusion in different lung cancer models as measured by 3'-deoxy-3'-18F-fluorothymidine PET and diffusion-weighted MR imaging. *J Nucl Med*. 2014; 55:983–988. [PubMed: 24777288]
22. El-Sharkawy AM, Schar M, Bottomley PA, Atalar E. Monitoring and correcting spatio-temporal variations of the MR scanner's static magnetic field. *MAGMA*. 2006; 19:223–236. [PubMed: 17043837]

FIGURE 1. ADC map ($\times 10^{-3} \text{ mm}^2/\text{s}$) and design of the phantom (A: 7-ml distilled water tube, B: 50-ml iced outer tube). A representative ROI is shown in red.

FIGURE 2.

Scatterplot of ADC reported by each site. All sites are in good overall agreement, but within-site day-to-day repeatability varies from site to site.

TABLE 1

Instrumentation and DWI Acquisition Parameters for Each Site

Site	MR make and model	Field strength (T)	Type of coil	TE (msec)	TR (msec)	b values (s/mm ²)	FOV (cm)	Acquisition matrix	Acquisition time
1	Bruker PharmaScan	7	Volume	20	2000	150, 300, 500, 700	4	128 × 128	5 min 36 s
2	Bruker BioSpec	9.4	Volume	19	1000	100, 200, 300, 400, 600, 700	3.5	128 × 128	9 min 36 s
3	Agilent	9.4	Volume	21	2000	150, 300, 500, 700	4	128 × 128	6 min 20 s
4	Bruker BioSpec	4.7	Surface	20	2000	150, 300, 500, 700	3.2	80 × 80	13 min 20 s
5	Bruker AV300 Microimaging	7	Volume	32	1500	200, 300, 500, 750, 1000	3	128 × 128	4 min
6	Magnex	7	Surface	20	2250	150, 500, 1000	2.56	128 × 64	8 min 10 s
7	Agilent	7	Volume	37	4000	120, 1200	3	128 × 64	8 min 31 s

Sites: 1: Inserm, 2: WWU Münster, 3: University of Cambridge, 4: AstraZeneca, 5: Institute of Cancer Research, 6: University of Manchester, 7: University of Michigan.

TABLE 2

ADC Calculation and Repeatability

Site	Mean SNR on b_{low} image	Mean SNR on b_{high} image	ADC($\times 10^{-3}$ mm ² /s) ^a	Mean within-ROI ADC variation (%) ^a	Between-slice ADC CoV (%) ^a	Day-to-day repeatability (%)
1	179	105	1.22 ± 0.01	4.7 ± 1.4	1.0 ± 0.6	0.8
2	105	78	1.26 ± 0.01	5.1 ± 1.4	0.5 ± 0.1	0.6
3	292	171	1.22 ± 0.07	3.9 ± 0.6	1.3 ± 1.1	5.9
4	31	16	1.30 ± 0.02	6.8 ± 0.9	1.2 ± 0.9	1.8
5	62	29	1.28 ± 0.02	10.1 ± 4.2 ^b	n/a	1.4
6	123	45	1.30 ± 0.05	6.2 ± 2.0	n/a	3.8
7	264	86	1.24 ± 0.02	1.7 ± 0.2	0.3 ± 0.1	1.5
ANOVA			$P = 0.02$; post-hoc test not significant	$P = 0.0013$	$P = 0.15$	
Overall mean			1.26	5.5	0.9	2.3

b_{low} : the lowest b value used for ADC calculation; b_{high} : the highest b value used for ADC calculation. n/a: sites having acquired monoslice DWI data.

^aData are expressed as mean ± standard deviation.

^bMean within-ROI ADC variation is significantly different between site 5 and sites 1, 2, 3, and 7.