



University
of Glasgow



CALIBRATION OF ADC VALUES FOR QUANTITATIVE MRI

Bethany Aylward

2173297A

MSc Medical Physics

Date of submission: 08-Aug-2022

First supervisor: John McLean

Second supervisor: Stacey McGowan Holloway

*This research project is submitted in partial fulfilment of the requirements for the degree of
MSc in Medical Physics*

CONTENTS

ABSTRACT	III
ACKNOWLEDGEMENTS	IV
ACRONYMS AND ABBREVIATIONS.....	V
1 INTRODUCTION AND BACKGROUND.....	1
1.1 DIFFUSION-WEIGHTED MRI (DWI).....	1
1.2 APPARENT DIFFUSION COEFFICIENT (ADC)	4
1.3 QUANTITATIVE IMAGING AND BIOMARKERS	6
1.4 QUALITY ASSURANCE (QA).....	7
1.5 QUANTITATIVE IMAGING BIOMARKERS ALLIANCE (QIBA).....	8
1.6 PROJECT AIMS AND REPORT OVERVIEW.....	9
2 LITERATURE REVIEW	10
2.1 CLINICAL APPLICATIONS OF ADC VALUES	10
2.2 QUALITY ASSURANCE OF ADC VALUES.....	11
3 METHODS.....	18
3.1 CALIBERMRI DIFFUSION PHANTOM.....	18
3.2 SET-UP.....	21
3.3 MRI HARDWARE.....	22
3.4 SCANNING PROTOCOL	22
3.5 QCAL SOFTWARE	23
3.6 BIOMARKER PERFORMANCE ASSESSMENT.....	23
4 SEQUENCE OPTIMISATION FOR THE LC THERMOMETER	27
4.1 RESULTS	27
4.2 DISCUSSION	30
5 ADC ACCURACY AND PRECISION	32
5.1 TEMPERATURE AND QUALITATIVE RESULTS.....	32
5.2 CONFORMANCE TO THE QIBA DWI PROFILE.....	33
5.3 BIAS AND LINEARITY OF ADC MEASUREMENTS.....	35
5.4 PRECISION OF ADC MEASUREMENTS.....	40
6 THE EFFECT OF HEAD COIL CHOICE ON ADC VALUES	45
6.1 RESULTS	45
6.2 DISCUSSION	50
7 ASSESSMENT OF CLINICAL RESEARCH PROTOCOLS.....	51
7.1 RESULTS	51
7.2 DISCUSSION	53
8 OVERALL DISCUSSION	54
8.1 EVALUATION OF CALIBERMRI DIFFUSION PHANTOM AND QCAL SOFTWARE.....	54
8.2 ADC CALIBRATION AT ROOM TEMPERATURE.....	54
8.3 DISCUSSION OF THE QIBA CONFORMANCE PROCESS.....	56
9 CONCLUSIONS	57
9.1 SUGGESTIONS FOR FUTURE WORK.....	57
9.2 ADDENDUM	58

REFERENCES	59
APPENDICES	65
APPENDIX A: INSTRUCTIONS FOR PERFORMING AN ICE BATH WITH THE CALIBERMRI DIFFUSION PHANTOM ...	65
APPENDIX B: RECOMMENDED SEQUENCE PARAMETERS PROVIDED BY CALIBERMRI	68
APPENDIX C: FULL SEQUENCE PARAMETERS	69
APPENDIX D: SEQUENCE REQUIREMENTS FOR ANALYSIS IN QCAL-MR	79
APPENDIX E: ROOM TEMPERATURE AND ICE BATH QA RESULTS FOR SIEMENS PRISMA 3T	80
APPENDIX F: ROOM TEMPERATURE AND ICE BATH QA RESULTS FOR SIEMENS SOLA 1.5T	81
APPENDIX G: SNR VARIATION WITH PHANTOM VIAL FOR ICE BATH SCANS	82
APPENDIX H: LONG-TERM REPEATABILITY OF ROOM TEMPERATURE ADC MEASUREMENTS.....	83
APPENDIX I: REPRODUCIBILITY OF ADC MEASUREMENTS BETWEEN THE TWO SCANNERS AT 0°C	84
APPENDIX J: QA RESULTS FOR ALL VIALS USING THREE HEAD COILS.....	85

ABSTRACT

Background: The apparent diffusion coefficient (ADC) is a measurable quantity from diffusion-weighted MRI (DWI). Changes in ADC have been shown to enable diagnosis, monitor disease, and predict treatment outcomes. However, poor standardisation limits clinical use. The accuracy and precision of ADC measurements vary between sites and scanners, decreasing confidence in results. The Quantitative Imaging Biomarkers Alliance (QIBA) aim to promote standardisation of biomarkers, such as ADC, through use of phantoms and standardised protocols. QIBA specify performance requirements within profiles, and conformance to a profile gives confidence in clinical measurements. ADC values are temperature dependent, therefore, the QIBA DWI profile requires phantom measurements to be made at 0°C. However, due to this being logistically challenging, quality assurance of ADC values at room temperature is appealing.

Project Aims:

- Evaluate a new quantitative DWI phantom and its associated analysis software
- Explore ADC calibration at room temperature
- Assess the technical performance of two MRI scanners for conformance to the QIBA DWI Profile, to confirm that high-quality ADC measurements can be delivered by NHS GGC.
- Investigate the variability of ADC values depending on the choice of head coil and DWI parameters

Methods: The CaliberMRI diffusion phantom contains 0%, 10%, 20%, 30%, 40%, and 50% PVP solutions. An increased %PVP corresponds to a decreased ADC. The phantom allows integrated room temperature measurements with a novel MR-readable thermometer. Traceable reference ADC values are provided for each solution at 0°C and 16-24°C, so either the thermometer or an ice bath can be used. Technical performance of two scanners was assessed following the QIBA DWI profile. Bias and linearity were assessed by comparison of measured ADC values to reference values. Short-term repeatability was assessed by four immediate repeat measurements, and long-term repeatability was assessed by two room temperature sessions 72 hours apart, either side of an ice bath scan. Effect of head coil choice on ADC measurements was investigated by performing room temperature QA on each coil.

Results: Both scanners were able to achieve conformance to the QIBA profile using an ice bath scan. Room temperature measurements showed good accuracy and precision for the central water vial, within the QIBA tolerances defined for 0°C, supporting that room temperature ADC calibration is practicable. Across ice bath and room temperature scans, bias of the central water vial ranged from -1.6 to 1.2% on the 3T Siemens Prisma and -3.1 to 1.3% on the 1.5T Siemens Sola. Vials with lowest ADC had inferior performance metrics, associated with an increase in artefacts with increasing %PVP. Variations between head coils were small (0-30% PVP measured within +/-2.5% bias for all coils), suggesting results obtained from each coil are comparable. Two clinical research protocols were successfully run on the phantom and demonstrated good performance at biologically relevant ADC values.

Conclusions: Characterisation of scanner performance is essential to promote the use of ADC measurements. The CaliberMRI diffusion phantom offers a convenient method of standardising ADC measurements due to the MR-readable thermometer that allows non-invasive room temperature measurements. Technical performance of both scanners was conformant to the QIBA DWI Profile, and QIBA performance tolerances were achievable for water vials in both scanners, at room temperature and 0°C. This work demonstrates that room temperature validation of ADC measurement is possible and covers a wide physiological range that matches the range of ADC values found in the human body. Through assessment of variability with choice of head coil and diffusion parameters, ADC has been shown to be a robust biomarker that can be reliably measured, and protocols have been established to be able to confirm conformance of other coils and scanners.

ACKNOWLEDGEMENTS

I would like to give thanks to John McLean for supervising this project, providing guidance and support throughout, and to Stacey McGowan Holloway for providing valuable radiotherapy insight.

Thanks also goes out to the CaliberMRI team for ongoing discussions and continued interest in this project. An additional thank you to the developers of qCal, for quick and comprehensive responses to queries.

Finally, thank you to Julia for continuous love and support throughout this MSc.

ACRONYMS AND ABBREVIATIONS

ADC	Apparent Diffusion Coefficient
DWI	Diffusion-weighted (magnetic resonance) Imaging
EIBALL	European Imaging Biomarker Alliance
EPI	Echo Planar Imaging
LC	Liquid crystal
LT	Long term
MRI	Magnetic Resonance Imaging
NIST	National Institute of Standards and Technology
QA	Quality assurance
QIB	Quantitative Imaging Biomarker
QIBA	Quantitative Imaging Biomarker Alliance
qMRI	Quantitative magnetic resonance imaging
RC	Repeatability coefficient
ROI	Region of Interest
SNR	Signal-to-noise ratio
ST	Short-term
TE	Echo time
TR	Repetition time
TWIST	Time-resolved angiography with Interleaved Stochastic Trajectories (MR sequence)
VIBE	Volumetric Interpolated Breath-hold Examination (MR sequence)
VOI	Volume of Interest
wCV	Within-subject coefficient of variation

1 INTRODUCTION AND BACKGROUND

This first chapter gives an overview of diffusion-weighted MRI and the concept of imaging biomarkers, specifically the apparent diffusion coefficient (ADC). MRI quality assurance in this context is also discussed before the aims of this project are outlined.

1.1 DIFFUSION-WEIGHTED MRI (DWI)

MRI can be used in the diagnosis and monitoring of a wide range of conditions and pathologies. Several different contrast mechanisms are available, providing both anatomical and functional information. In pathology, sub-cellular changes often occur before gross anatomical changes are seen and, therefore, being able to detect these changes can be beneficial in several clinical applications. One example of these techniques is diffusion-weighted MRI (DWI).

DWI uses molecular diffusion as a contrast mechanism in MRI to identify areas where diffusion is more or less restricted than in normal tissue. In DWI, diffusion refers to the probabilistic, random process by which water molecules move gradually over time [1]. This random motion, also known as Brownian motion, is due to thermal energy of the molecules [2]. In isotropic free diffusion, the movement of water molecules is completely unrestricted, and the molecules can move equally in all directions. Increased diffusion results in a greater spread of molecules, however, their average location does not change [3]. For these freely diffusing molecules, the probability of a molecule's location over a given amount of time (t) follows a Gaussian distribution.

The Gaussian function of displacement (x) has mean $\langle x \rangle = 0$ and variance $2Dt$ [4] where D is the self-diffusion coefficient and t is the diffusion time. Therefore, the root-mean-squared displacement is equal to the standard deviation:

$$\sqrt{\langle x^2 \rangle} = \sqrt{2Dt}$$

Equation 1

This is known as the Einstein relationship and shows that the root-mean-squared displacement of a molecule at a given time, t , is dependent upon its diffusion coefficient, D [4].

The Stokes-Einstein expression for D [4] reveals the dependence of diffusion on the temperature, viscosity, and particle size:

$$D = \frac{k_B T}{6\eta R}$$

Equation 2

Where T is the absolute temperature, k_B is Boltzmann's constant, η is the viscosity of the solution, and R is the hydrodynamic radius of the molecule.

Biological tissues can be considered to consist of compartments of aqueous solution, as they are made up of both membrane-bound cells and organelles and interstitial space between these [4]. Physical barriers, such as cell membranes, macromolecules, and other tissue microstructures, prevent or limit motion, causing diffusion to be restricted [3]. Restricted diffusion can be either isotropic or anisotropic. In the former, diffusion is equally restricted in all directions, whilst in the latter, diffusion is preferentially limited along one direction. Different tissues restrict water to different degrees, allowing the diffusivity of water to act as a contrast mechanism [1].

In biological tissue, the rate of diffusion of water molecules is characterised using the apparent diffusion coefficient (ADC) rather than the true diffusion coefficient. The use of the ADC reflects the uncertainties in the measurement method due to the indirect measurement of diffusion, which is an average diffusivity over all tissue microenvironments within the voxel and contains contributions from other sources of tissue motion [4].

1.1.1 Measuring diffusion using MRI

Stejskal and Tanner [5] developed a method of measuring diffusion using MRI that still underlies the basic method of DWI employed today. The sequence is commonly referred to as Pulsed Gradient Spin Echo (PGSE) and is based on a standard spin-echo sequence with the addition of a symmetrical pair of diffusion sensitising gradients on either side of the 180° refocusing pulse. These gradients make the MR signal sensitive to molecular motion.

If no diffusion occurs, there is no net movement of spins, meaning the spins are re-phased by the second gradient, which cancels out the effects of the first, and consequently there is little change to the net magnetisation vector due to the diffusion gradients and no loss of signal. Conversely, if diffusion of the spins occurs between the application of the two diffusion gradients, there will be dephasing of the spins as they do not experience equal dephasing and rephasing gradients. This decreases the net magnetisation vector in the transverse plane and consequently results in a loss of signal from the tissue [4]. The greater the diffusivity of the tissue (higher ADC), the greater the reduction in signal. This process is shown in **Figure 1**.

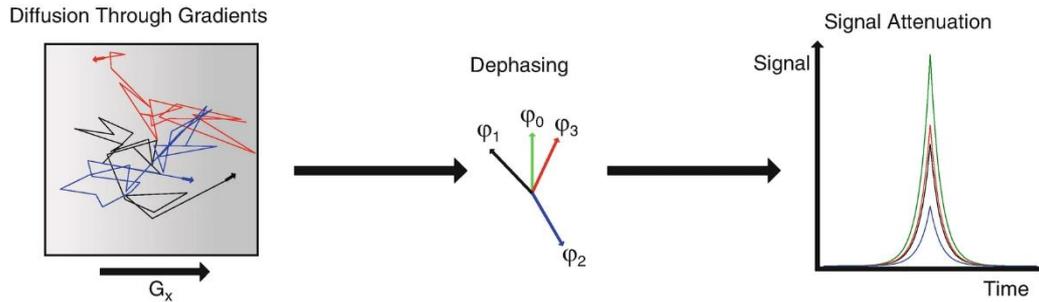


Figure 1. Left: three spins diffusing randomly in a one-dimensional magnetic gradient, G_x . Centre: Depending on the motion of the spins, they will have dephased by different amounts, ϕ_0 represents the phase of a spin with no diffusion. Right: The greater the dephasing of the spins, the greater the signal attenuation. Taken from Venkataraman & Zhong, 2021 [1].

The signal intensity is described by the following equation:

$$S = S_0 e^{-(\gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3})) \cdot ADC}$$

Equation 3

Where G , δ , and Δ are properties of the diffusion gradients, S_0 is the signal intensity with no diffusion weighting applied, and γ is the gyromagnetic ratio. For high ADC, the signal intensity on a DWI image is low. For tissues with a high ADC value, indicating more restricted diffusion, the signal intensity on a DWI image is much higher [3].

1.1.2 DWI Pulse sequences

Based on a standard spin-echo pulse sequence, the PGSE sequence consists of a pair of 90° and 180° RF pulses with additional diffusion-weighting pulses on either side of the 180° pulse (**Figure 2**). These gradients control the sensitivity to diffusion [3]. The most common readout method for DWI is a single-shot echo-planar imaging (SS-EPI) acquisition. In an EPI sequence, k-space is filled in a zig-zag pattern

by blipping the phase, encoding gradient to fill all of k-space within a single application of the pulse sequence. This sequence is chosen due to its speed, which freezes bulk motion that would otherwise obscure the diffusion contrast. However, due to the increased speed, EPI sequences have lower resolution and are more susceptible to artefacts [3]. The main artefacts seen with SS-EPI acquisitions are Eddy current artefacts, susceptibility artefacts, ghosting artefacts, chemical shift artefacts, and motion artefacts [6]. Alternative readout mechanisms have been investigated [7, 8]. For example, fast spin echo methods have been shown to improve image quality and reduce motion artefacts, however, acquisition times are longer, and SAR is higher [7].

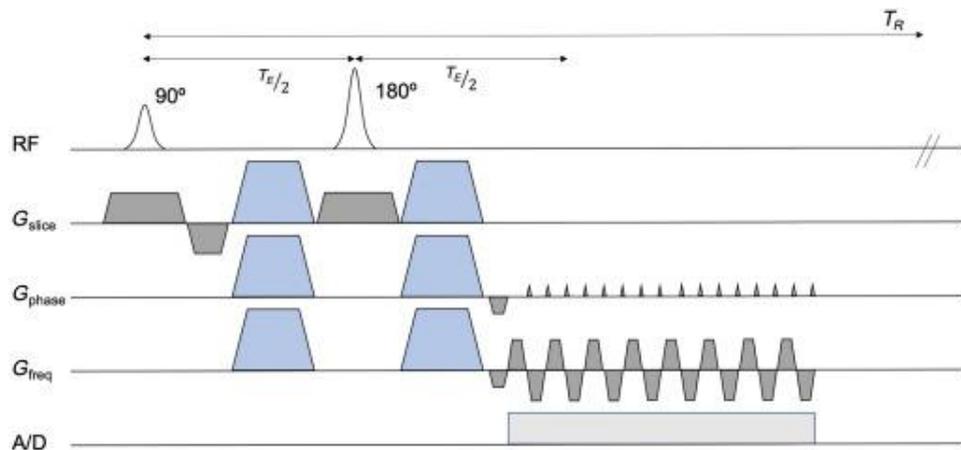


Figure 2. SS-EPI DWI pulse sequence. The blue gradients are the diffusion-weighting pulses. A 2D EPI readout sequence is shown. Taken from Yang & McNab, 2020 [9].

1.1.3 The b-value

The sensitivity to diffusion is controlled by the b-value (s/mm^2), which was defined by Stejskal and Tanner [5] as:

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right)$$

Equation 4

Where γ is the gyromagnetic ratio of a proton (42.6 MHz/T), G is the gradient amplitude (mT/m), δ is the duration of each gradient (s) and Δ is the centre-to-centre separation of the two gradients (s) (**Figure 3**). Any of the latter three parameters can be varied to change the strength of the diffusion weighting. In practice, the b-value is selected by the operator and one of the factors is automatically changed by the scanner software. Most commonly, it is the gradient amplitude that will vary [4], as this does not change the minimum TE [3]. It can be noted that $\left(\Delta - \frac{\delta}{3} \right)$ is known as the diffusion time, which appears in Einstein's relationship for the root-mean-squared displacement (Equation 1).

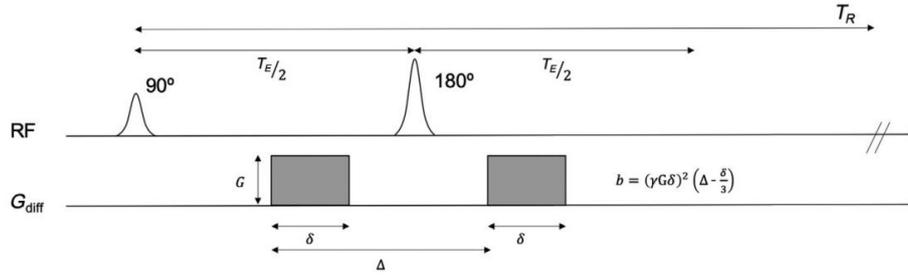


Figure 3. Diffusion gradients labelled with the parameters in Equation 4 that control the sensitivity to diffusion. Taken from Yan & McNab, 2020 [9].

A high b-value increases sensitivity to diffusion, while a low b-value decreases it. Substituting the expression for the b-value, the diffusion weighting equation simplifies to:

$$S(b) = S_0 e^{-b \cdot ADC}$$

Equation 5

S_0 is the signal intensity when no diffusion gradients are applied (i.e., $b = 0$). This is equivalent to a T2-weighted image obtained from a spin-echo sequence. $S(b)$ is the signal intensity for a particular b-value.

1.1.4 Choice of gradients

As MRI systems have three orthogonal magnetic field gradient coils, gradients can be produced along the x, y, or z-axes. Alternatively, by combining coils, a magnetic field gradient can be produced along any arbitrary direction [4].

The composite of three orthogonal DWIs is known as a trace DWI, obtained by taking the geometric average of three orthogonal direction DWI obtained at the same b-value [10]:

$$SI = \sqrt[3]{SI_x SI_y SI_z}$$

Equation 6

SI is the signal intensity of the isotropic trace DWI and SI_i is the signal intensity of the diffusion-weighted image with the gradients applied in the i^{th} direction. Trace DWI eliminates anisotropic diffusion and creates an image that is independent of patient orientation by averaging the diffusion measurements from the three orthogonal directions [1].

1.2 APPARENT DIFFUSION COEFFICIENT (ADC)

Due to the long TE required, DWI is inherently T2-weighted, which can modify the appearance of the diffusion-weighted image [3, 6]. The most widely known effect is ‘T2 shine-through’, in which a region of elevated T2 can be misinterpreted as an area of restricted diffusion [3]. To overcome some of the uncertainties associated with diffusion-weighted images, ADC maps are created. By obtaining diffusion-weighted images at two or more b-values, calculating the ADC value for each voxel, and displaying these as a map of greyscale values corresponding to the strength of isotropic diffusion, the diffusion can be separated from the relaxation effects such as T2 shine-through [1].

With two b-values, the ADC value can be calculated by the following equation:

$$ADC = \frac{1}{b_2 - b_1} \ln(S_{b_1}/S_{b_2})$$

Equation 7

Alternatively, if DWI is obtained over a range of b-values, a least-squares fit can be performed on a plot of the natural logarithm of the relative signal intensity, $\ln(S_b)$, against the b-value [3] (**Figure 4**). The ADC value is obtained from the slope of the fitted exponential curve. More b-values provide a better fit and a more accurate ADC value [3].

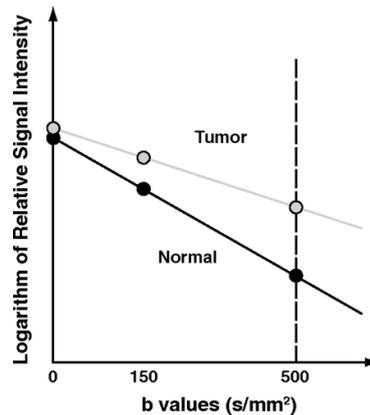


Figure 4. To determine the ADC value from multiple b-values, the logarithm of relative signal intensity is plotted against the b-value. The slope of the line is the ADC value, which is different in different tissues. Taken from Koh & Collins, 2007 [11].

A bright region on an ADC map corresponds to a high ADC value, which indicates free or less restricted diffusion that would appear dark on DWI. A dark region on an ADC map corresponds to a low ADC value, indicating more restricted diffusion that would appear bright on DWI (**Figure 5**). DWI and ADC maps are generally interpreted qualitatively alongside one another, allowing T2 effects and true diffusion changes to be distinguished.

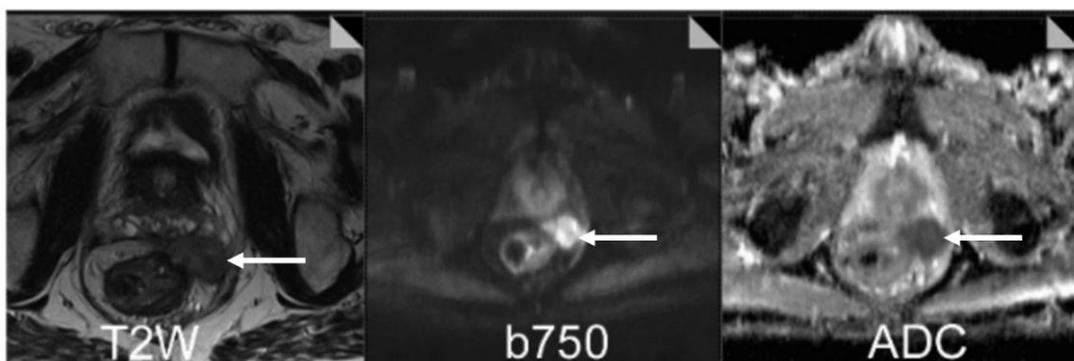


Figure 5. T2-weighted image (left), DWI (centre), and ADC map (right) showing a lesion with restricted diffusion, appearing bright on the DWI image and dark on the ADC map. Taken from Padhani et al., 2009 [12].

1.2.1 Mono-exponential or bi-exponential?

From multiple different b-value DW images, it is found that the signal decay is bi-exponential in most organs with more rapid signal attenuation at very low b-values (**Figure 6**), rather than the mono-exponential assumption commonly used [3, 11]. The signal from multiple b-values can be fitted to a two-component model known as the IVIM (intravoxel incoherent motion) model [13]. In this model, the shorter diffusion coefficient (known as the ‘pseudo-diffusion coefficient’) is due to capillary perfusion, while the longer diffusion coefficient is the true diffusion coefficient [3]. However, it has proved difficult to achieve good accuracy and reproducibility with bi-exponential fitting and, therefore, the mono-exponential model remains common. Occasionally, low b-values ($b = 50-100$) are used instead of $b=0$ to remove the perfusion fraction [3].

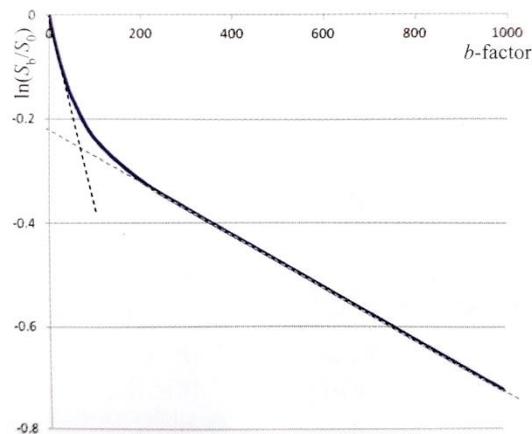


Figure 6. Bi-exponential signal decay shows more rapid signal attenuation at low b-values due to sensitivity to capillary perfusion. Image from McRobbie et al., 2017 [3].

1.2.2 Limitations of ADC maps

Qualitative assessment can have limitations. For example, a tumour surrounded by oedema will appear dark on an ADC map even in the absence of more restricted diffusion and, therefore, quantitative ADC comparison to normal brain is recommended [14]. Through direct voxel values or, more commonly, region of interest (ROI) analysis, the ADC values of a specific tissue or region can be determined. This has led to the proposition of the use of ADC as a biomarker within quantitative imaging and has driven the development of quantitative MRI.

1.3 QUANTITATIVE IMAGING AND BIOMARKERS

Quantitative imaging has been defined as “the extraction of quantifiable features from medical images for the assessment of normal or the severity, degree of change, or status of a disease, injury, or chronic condition relative to normal” [15]. Successful quantitative imaging requires standardised methods, qualified imaging devices, validated algorithms, as well as standardised analysis and reporting [16]. Quantitative imaging biomarkers (QIBs) are a measurable quantity obtained from quantitative imaging, defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [17]. Biomarkers need to be reliable, reproducible, and sensitive to the pathophysiological conditions being imaged [16].

There have been several biomarkers proposed for MRI. These include relaxivities (T1, T2, T2*), fat fraction, iron content, and diffusivity. In a 2021 European survey of radiologists, it was found that the most commonly applied qMRI sequence was DWI with 82% of responding radiologists currently using the technique clinically [18]. In the results of another survey, this time focusing specifically on glioma imaging throughout Europe, it was found that although DWI is almost always performed, in the majority of cases (78.2%), ADC is assessed by visual comparison to a normal brain rather than assessed quantitatively [14]. Currently, variability of ADC values is holding it back from more widespread clinical use and, in particular, reproducibility is a major challenge [19, 20].

1.3.1 Common quantitative imaging performance metrics

For any quantitative value, the uncertainty must be known. This is characterised by the bias, linearity, and precision of the measurement. *Bias* is an estimate of systematic measurement error, calculated as the difference between the measured value and the true value or a reference value with suitably small uncertainty [15]. *Linearity*, in metrology, is the ability of a measurement method to provide measured values that are directly proportional to the quantity being measured [15].

Precision is the closeness of agreement between measured values obtained by replicate measurements under specified conditions. Precision includes repeatability and reproducibility. *Repeatability* is the measurement precision when the measurement conditions are unchanged between replicate measurements, meaning the same measurement procedure, operators, measuring system, and operating conditions are used [15]. In contrast, *reproducibility* is the measurement precision when the measurement conditions vary between replicate measurements, for example, different measuring systems or operators [15]. Each of these concepts underpins confidence in ADC measurements and will be focused on throughout this work.

1.3.2 ADC Calibration

While qMRI requires standardisation of all aspects of the process, from acquisition to analysis and reporting, there is a large focus on the MRI scanners themselves, and reproducibility of results across scanners. Tofts [21] summarised that “A perfect quantitative MRI machine is one that, in making a measurement, contributes no significant extra variation to that which already exists from biological variation”. NHS GGC has a wide range of MRI systems of varying age, vendor, model, and field strength. It is important to be able to identify which systems are suitable for qMRI applications. One step towards this is implementation of quality assurance procedures using traceable phantoms and standardised sequences.

1.4 QUALITY ASSURANCE (QA)

Biomarker QA provides confidence that an observed change is a physiological change rather than due to differences in technical performance between systems, or drift in a single scanner over time. QA should be practical, scalable, and affordable. Furthermore, it should not be too time or resource consuming. QA on a single scanner allows assessment of intra-scanner repeatability, system stability, and comparisons before and after maintenance or upgrades. When the same QA is performed on multiple scanners, inter-scanner reproducibility can be assessed and hardware and software variations between scanners identified. Harmonisation between scanners through calibration of their technical performance is essential for comparison of quantitative measurements.

QA allows traceability and standardisation for biomarkers through standardised acquisitions and analysis of a traceable phantom. Traceability can be defined as a “chain of comparisons which directly relates any given measurement to the primary standard determination of that unit” [22]. Traceability requires involvement of national measurement centres, such as the National Physics Laboratory (NPL) and the National Institute of Science and Technology (NIST) to provide or validate reference materials

and primary standards. Currently, there are groups around the world trying to improve the standardisation of biomarkers, such as the European Imaging Biomarkers Alliance (EIBALL), iMet-MRI, and the Quantitative Imaging Biomarkers Alliance (QIBA). QIBA focus on technical validation of biomarkers, considering their bias, repeatability, reproducibility, and sources of variability.

1.5 QUANTITATIVE IMAGING BIOMARKERS ALLIANCE (QIBA)

QIBA was launched by the Radiological Society of North America (RSNA) in 2007 with the mission to “improve the value and practicality of QIBs by reducing variability across devices, patients, and time.” [23]. Currently, QIBA is comprised of 22 biomarker committees [23]. Each biomarker committee aims to deliver a profile, specific to that biomarker [24]. Profiles are technical performance documents that aim to standardise the full imaging chain related to the biomarker and minimise underlying sources of uncertainty [24].

A profile acts as an implementation guide and includes a checklist specifying what each person and device (including both hardware and software) involved in the imaging traceability chain must be capable of achieving. All profiles follow a standard structure, and the requirements are focused on achieving biomarker values with minimal systematic bias and measurement variability [10]. Conformance to a profile means all people and devices conform to all the specifications assigned to them in the profile [10]. If the system operates within the conformance limits stated in the profile, then contribution of technical errors to the confidence intervals measured for tissue will be negligible [10].

Claim statements are organ-specific and indicate the reproducibility of the quantitative measurement assuming adequate technical performance requirements are met. They are derived from scientific literature and updated as more data becomes available [10, 25]. If conformance to the profile is achieved, the biomarker performance described in the profile claim should be expected [10]. Claims are provided for brain, liver, prostate, and breast in the DWI Profile because these organ systems have high clinical utilisation of ADC and sufficient statistical evidence within peer-reviewed literature to support the profile claims [10]. The four claims for DWI are provided below:

“Claim 1a: A measured change in the ADC of a brain lesion of 11% or larger indicates that a true change has occurred with 95% confidence.

Claim 2a: A measured change in the ADC of a liver lesion of 26% or larger indicates that a true change has occurred with 95% confidence.

Claim 3a: A measured change in the ADC of a prostate lesion of 47% or larger indicates that a true change has occurred with 95% confidence.

Claim 4a: A measured change in the ADC of a breast lesion of 13% or larger indicates that a true change has occurred with 95% confidence.” [10]

1.5.1 QIBA DWI Profile Requirements

This research project will focus on the QIBA DWI profile’s requirements for the acquisition device, assessed using a quantitative DWI phantom at the magnet isocentre. The QIBA requirements for the acquisition device are summarised in **Table 1**. Methods for determining each performance metric can be found in Section 3.6.

Table 1. QIBA DWI Profile requirements for the acquisition device. Using an ice-water phantom or other quantitative DWI phantoms at/near the isocentre of the scanner. RC = Repeatability coefficient, wCV = within-subject coefficient of variation, ST = short-term, LT = long-term, SNR = Signal to Noise Ratio.

PERFORMANCE METRIC	QIBA DWI PROFILE TOLERANCE
%BIAS	$\leq 3.60\%$ or $\leq 40 \mu\text{m}^2/\text{s}$
RC _{ST}	$\leq 15 \mu\text{m}^2/\text{s}$
WCV _{ST}	$\leq 0.5\%$
RC _{LT}	$\leq 65 \mu\text{m}^2/\text{s}$
WCV _{LT}	$\leq 2.2\%$
95% CONFIDENCE INTERVAL OF THE SLOPE (β_1)	$0.95 \leq 95\% \text{ CI for } \beta_1 \leq 1.05$
R ² OF THE LINEAR MODEL FIT	> 0.9
MAX B-VALUE DEPENDENCE	$< 2\%$
RANDOM ERROR	$< 2\%$
SNR	$\geq 50 \pm 5$
%BIAS WITH OFFSET FROM ISOCENTER: WITHIN 4 CM IN ANY DIRECTION ¹	$< 4\%$
R/L OFFSET < 10 CM (WITH A/P AND S/I < 4 CM) ¹	$< 10\%$
A/P OFFSET < 10 CM (WITH R/L AND S/I < 4 CM) ¹	$< 10\%$
S/I OFFSET < 5 CM (WITH R/L AND A/P < 4 CM) ¹	$< 10\%$

¹ requires a uniform DWI phantom

Conformance to the QIBA DWI profile is desirable because it confirms a clinical site is delivering high-quality imaging measurements and allows performance issues to be quantitatively identified, optimised, and monitored [23]. Additionally, it promotes standardisation of biomarker measurements and improves future quantitative imaging guidelines.

1.6 PROJECT AIMS AND REPORT OVERVIEW

The aims of this project are as follows:

- Evaluate a new quantitative DWI phantom and its associated analysis software
- Explore ADC calibration at room temperature
- Assess the technical performance of two MRI scanners for conformance to the QIBA DWI profile, to confirm that high-quality ADC measurements can be delivered by NHS GGC.
- Investigate the variability of ADC values depending on the choice of head coil and DWI parameters

In the next chapter, a review of the literature on both the clinical applications of ADC as a biomarker and methods to calibrate ADC values is performed. The CaliberMRI diffusion phantom is introduced, and more details are provided in chapter 3, which outlines the methods used within the experimental work. The novel aspect of the phantom, the LC thermometer which allows ADC calibration at room temperature, requires a T1-weighted image to determine the phantom temperature. Optimisation of this sequence is discussed in chapter 4. With an appropriate QA protocol established, the accuracy and precision of ADC measurements on two scanners is presented in chapter 5 and within this chapter ice bath and room temperature ADC calibration are compared. The effect of head coil choice on ADC measurement is investigated in chapter 6. With ADC QA using a standardised protocol established, in chapter 7 the phantom is used to evaluate two clinical research protocols. Finally, in chapter 8, the aims are reviewed in an overall discussion of the work presented.

2 LITERATURE REVIEW

This chapter presents a broad literature review, split into two sections. The first focuses on the clinical applications of ADC values, highlighting current advantages and limitations of ADC measurements. The second section presents methods of calibrating ADC values, looking at the development of a suitable phantom.

2.1 CLINICAL APPLICATIONS OF ADC VALUES

Use of ADC values as a quantitative biomarker has been increasingly investigated over the last two decades. The QIBA profile provides claims for brain, breast, liver, and prostate because there is sufficient test-retest literature available for these areas. However, there is growing literature on other body sites, including head and neck [26-28], pancreas [29, 30], and other abdominal organs [31].

ADC has been investigated as a biomarker for diagnosis and classification, for example, ADC values have been shown to be able to differentiate between subtypes of breast cancer [32] and classify liver lesions [19]. Lesion characterisation has also been possible in the brain [33] and head and neck cancers [27]. ADC values have also been proposed as a biomarker for predicting response to treatment [34], including to chemotherapy [35] and during radiotherapy [36], as well as predicting clinical outcomes [37].

ADC values measured in human tissue cover a wide range of values. In head and neck lesions alone, Wang et al. [28] found malignant lymphomas had a mean ADC of $0.66 \pm 0.17 \times 10^{-3} \text{ mm}^2/\text{s}$, carcinomas $1.13 \pm 0.43 \times 10^{-3} \text{ mm}^2/\text{s}$, solid tumours $1.56 \pm 0.51 \times 10^{-3} \text{ mm}^2/\text{s}$, and benign cystic lesions $2.05 \pm 0.62 \times 10^{-3} \text{ mm}^2/\text{s}$, with a significant difference between each tumour type ($p < 0.05$).

Across the literature, several key advantages of quantitative DWI have been agreed upon. MRI offers superior soft-tissue contrast compared to CT, allowing better definition of tumours and organs at risk for treatment planning [26]. Quantitative ADC values are objective and therefore beneficial for influencing clinical decisions and evaluating response to treatment [32], compared to qualitative visual assessment. Additionally, the non-ionising nature of MRI means a tumour can be monitored more frequently during treatment as there is no additional radiation dose to the patient. This allows for adaptive treatment based upon changes seen.

In breast cancer treatment, Galban et al. [38] found that a large change in ADC is possible with only a small change in tumour volume. In one patient, mean ADC increased by 21% ($1.1 \times 10^{-3} \text{ mm}^2/\text{s}$ to $1.4 \times 10^{-3} \text{ mm}^2/\text{s}$) between pre- and mid-treatment while the tumour volume only decreased by 2% [38]. ADC values, therefore, allow therapeutic effects to be detected earlier in some diseases. It has also been demonstrated that it is possible to use quantitative DWI to differentiate tumour recurrence and post-radiotherapy changes [39]. By identifying post-radiotherapy changes, more conservative measures can be taken compared to if tumour recurrence was suspected, which may preserve more organ function [39].

However, there remains limitations to ADC measurements clinically. Although ADC values have been shown to be clinically useful in a wide range of applications, they can suffer from poor repeatability and reproducibility. Several different factors have been found to influence ADC values. Biological factors such as patient demographics and tumour heterogeneity can influence the reported ADC value [29, 38], as well as factors such as patient set-up and implementation of breath-hold techniques [40, 41]. However, there is considerable variation in ADC values due to technical factors. Technical variations between scanners and DWI sequence parameters result in poor harmonisation of ADC values and substantial variation between scanners has been seen [20, 41].

Although several authors have found no significant difference in ADC values obtained at different field strengths [11, 29, 42], differences were found by Dale et al. [43] when measuring ADC values in the liver. Also looking at abdominal organs, Rosenkrantz et al. [31] found both 1.5T and 3T scanners provided similar ADC values, however, image quality was worse at 3T. Galban et al. [38] highlighted that use of different coils may introduce additional variability as noise levels in ADC measurements can vary with coil design, due to differences in signal sensitivity and homogeneity between coils. Furthermore, gradient non-linearity increases variability of ADC values if the lesion location relative to iso-centre varies between scans and/or patients [38].

The ideal number, range, and position of b-values used depends on the organ. Merisaari et al. [44] tested a range of b-value for prostate tumour classification. They found the choice of b-values affected the repeatability of the ADC values, however, did not impact the characterisation performance. The choice of readout sequence can influence quantitative DWI. Although EPI sequences are most common, they are associated with several artefacts [19, 32]. Artefacts can prevent evaluation of ADC values. In a study by Bickel et al. [32], 7.6% of patients had to be excluded due to imaging artefacts. Alternative readout sequences can reduce geometric distortion, however, Schouten et al. [45] reported this may result in different ADC values and poorer response prediction.

To reduce the limitations of using ADC values as a quantitative biomarker, quality assurance and guidance by organisations such as QIBA are essential. Phantoms can be used to calibrate ADC values and investigate the different factors that influence them.

2.2 QUALITY ASSURANCE OF ADC VALUES

ADC values can be utilised as a biomarker through standardised, regular, and robust quality assurance procedures using a phantom with known, stable ADC values. The ideal properties for a diffusion phantom include accurate and traceable components that are chemically stable, viscous, cost-effective, and non-toxic, with a range of ADC values that cover the range seen in biological tissues as well as T1 and T2 values similar to tissue [46-48]. The phantom should be transportable, easy to set up, and allow reproducible positioning. While the human body stays at a stable temperature, for calibration of a temperature-dependent biomarker such as ADC, phantoms require external temperature control or additional temperature calibration [49]. Phantoms allow estimation of bias with reference to known values, however, repeatability in phantoms is generally better than repeatability in human subjects due to the inherent variability within subjects (due to variation in tissue structure, geometry, and patient motion). Imaging healthy volunteers, therefore, improves characterisation of the scanner repeatability and is often performed alongside phantom studies [50-55].

The QIBA DWI profile requires ADC measurement bias and precision to be assessed near isocentre using a quantitative DWI phantom with known diffusion properties [10]. Various phantoms have been developed and reported in the literature. An overview of some of these phantoms is presented and their advantages and disadvantages are discussed.

2.2.1 Agar and sucrose DWI phantom

In 1998, Laubach [56] developed a diffusion phantom using agar gels and sucrose solutions. The phantom (**Figure 7**) consisted of two compartments and the difference in ADC values between the two compartments mimicked the difference in ADC values between normal grey matter and acute stroke in humans. Agar and sucrose were both identified as inexpensive, widely accessible, and easy to prepare. Both solutions had similar T2, so appear similar on the b0 image, however, the sucrose compartment appears brighter on the b1000 image. The difference between ADC values was imitated due to the high concentration of sucrose required to achieve an accurate absolute ADC value. The phantom was not

sealed, and the relaxation properties of the gel decreased beyond two weeks. Due to the instability, and the lack of temperature control, the phantom is not a suitable design for reproducible ADC QA.

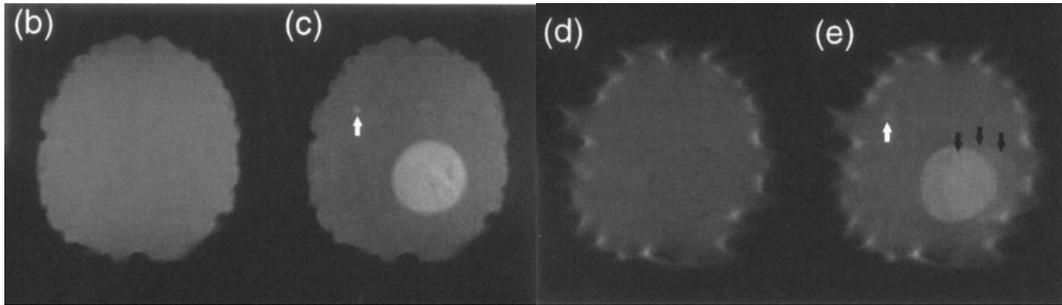


Figure 7. Agar-filled brain mould with a sucrose compartment, created by Laubach et al. [56] **b)** T2-weighted spin echo, **c)** spin echo image with $b = 1000 \text{ s/mm}^2$, **d)** DWI EPI, $b = 0 \text{ s/mm}^2$, **e)** DWI EPI, $b = 1000 \text{ s/mm}^2$. Black arrows show chemical shift artefact of the sucrose solution. Taken from Laubach et al. (1998) [56].

2.2.2 Alkane-based DWI phantom

Tofts et al. [48] investigated the suitability of a range of organic liquids, including cyclic alkanes, n-alkanes, and alcohols. The ADC values of each liquid were characterised as a function of temperature from 15°C to 30°C in 5°C, to find liquids that covered a range of $0.3 - 2.1 \cdot 10^{-3} \text{ m}^2/\text{s}$, corresponding to ADC values seen in white matter, ischemic brain, and multiple sclerosis. Using different liquids, no weighing, mixing, or volume measurements were required to generate a range of ADC values. Tofts et al. compared fitting methods to model the change in diffusion coefficient with temperature for 15°C to 30°C and chose a quadratic model to interpolate the data. They found the liquids had temperature coefficients ranging from 1.7 to 3.2 %/°C. All liquids are well-defined, stable, and readily available; however, handling requires protective equipment and a well-ventilated room or fume cupboard. Pierpaoli et al. [47] noted acceptance of alkanes in a clinical setting is limited due to concerns regarding their flammability and potential toxicity.

2.2.3 Water as a DWI phantom

Water is a popular phantom material because it is safe, cheap, and readily available. However, low viscosity increases the likelihood of ADC values being affected by bulk movements due to vibration [57]. Additionally, a T2 value higher than values in the human body can impact precision of ADC measurements [58]. To reduce the T2 value of water, several authors have proposed doping the water [55, 58-60].

Delakis et al. [58] proposed the use of two aqueous test solutions, CuSO_4 and sucrose, to develop a quality control protocol for ADC measurements. Doping water with a small concentration of CuSO_4 can reduce T2 without changing the ADC value [58]. Delakis et al. measured the temperature of the scanner bore during scanning and corrected the ADC values to a standard temperature of 21°C. Both solutions displayed an approximately linear relationship between ADC and temperature. Comparing to the work of Laubach et al. [56] and Tofts et al. [48], Delakis et al. concluded their solutions are cheaper and safer to handle than alkanes and, unlike agar gels, do not require special storage conditions for long-term stability.

2.2.4 Ice-water DWI phantoms

In 2011, Chenevert et al. [61] proposed an ice-water phantom for assessment of ADC measurements. The phantom consisted of a 50 mL conical tube filled with distilled water within a 1000 mL jar filled with an ice-water bath for temperature control (**Figure 8a, b**). The ADC value of water at 0°C is well

characterised in the literature as $1.1 \times 10^{-3} \text{ mm}^2/\text{s}$. Ice ensures the phantom remains at a stable temperature, and is cheap, safe, and widely available. Thermal equilibrium is achieved within 30 minutes and the ice bath can provide temperature control for several hours (**Figure 8c**). The phantom was used to investigate variability between 20 MRI scanners at 6 institutions. Protocols were not standardised, only b-values specified, and hence TR and TE had high variation between sites. All measurements were within 10% of the literature value, with 86% of measurements within 5% (**Figure 8d**). Differences in ADC values between manufacturers were observed. Pooled over all systems, there was no impact of field strength on ADC value. Repeatability of a single scanner over 25 days was found to be within $\pm 5\%$. Chenevert et al. demonstrated that an ice-water phantom is suitable for calibration of ADC values in a multi-centre trial.

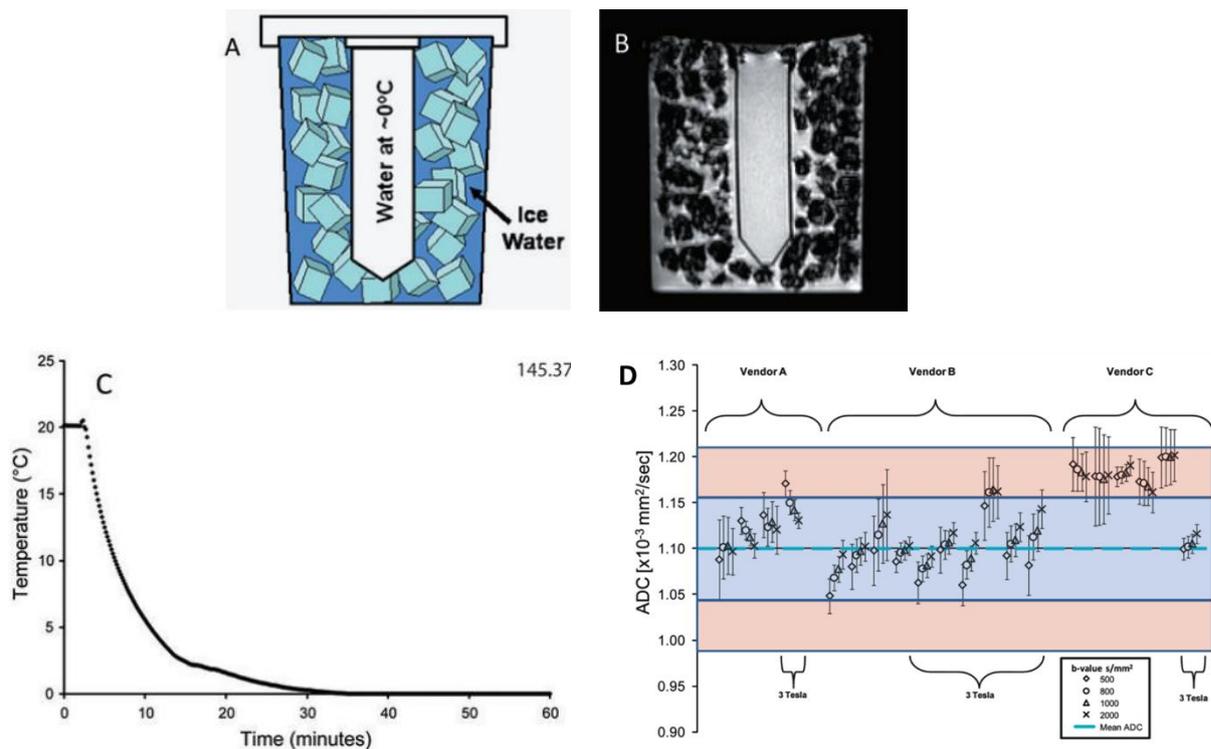


Figure 8. a) schematic and b) axial MR image of the ice-water phantom. c) time taken for the water vial to reach thermal equilibrium when placed in the ice bath. d) ADC measurements on each scanner, grouped by vendor, magnetic field strength, and b-value. The blue shaded region indicates $\pm 5\%$ bias, and the orange region indicates $\pm 10\%$ bias. Images and graphs taken directly from Chenevert et al. (2011) [61].

Following this work, Malyarenko et al. [62] used an ice bath phantom consisting of five vials of water (one central and four outer) to evaluate agreement of ADC values from 35 scanners at 18 sites. In contrast to Chenevert et al. [61], Malyarenko et al. implemented a common acquisition protocol across sites to reduce variability. Each site completed four immediate repeats using a head coil. ADC values of the central water vial at isocentre were within 3% of the literature value for 95% of the systems, average intra-exam repeatability was within 1%, and day-to-day repeatability within 4.5%. DWI SNR was lower at 1.5T than 3T but was adequate to not introduce additional bias on all scanners. These results agree with or are better than the results by Chenevert et al. [61], suggesting the standardised protocol across site improved harmonisation. Outer vials were not used for reproducibility statistics due to variations in positioning between sites. Using a torso coil, off-isocentre measurements were performed to assess spatial variability and demonstrated lower reproducibility across scanners. ADC

error increased approximately quadratically with distance from the isocentre and higher field strength corresponded to higher off-centre variability.

Presently, ice-water phantoms are commonly used in calibration of quantitative DWI [52-54, 59, 63] and can be used to demonstrate conformance with the QIBA DWI profile [10]. However, significant preparation is required before every session and only a single ADC value is available.

2.2.5 Polyvinylpyrrolidone (PVP) DWI phantoms

In 2009, Pierpaoli et al. [47] proposed polyvinylpyrrolidone (PVP) as a suitable material for a diffusion phantom. PVP is non-toxic with good chemical stability and used in many medical applications. By varying the concentration of PVP in water, a range of ADC values can be produced. Solutions of 15% to 65% w/v were measured at 22°C on a 3T Scanner. Pierpaoli et al. determined ADC decreases linearly with increased %PVP up to 50% (**Figure 9a**). Above 50%, the linear relationship does not hold.

Boss et al. [64] presented a PVP-based diffusion phantom developed by the National Cancer Institute, RSNA QIBA, and NIST. Following the work by Chenevert et al. [61], the phantom is designed to be filled with ice-water, allowing temperature-controlled measurements. Two rings of PVP solutions from 0 to 50% (**Figure 10**) allow assessment of ADC measurement over a range of ADC values. An additional central water vial allows comparison to ice-water phantoms. Boss et al. [65] found over the range of ADC values (0.12 to 1.12×10^{-3} mm²/s), reproducibility across scanners and imaging sites was good with coefficient of variation (CV) between 1.1 to 2.2% for 0 to 40% PVP, however, the 50% PVP vial showed a much higher variation (11.3%) (**Figure 9b**). Average CV between inner and outer vials of the same concentration was less than 5% for all vials. This phantom is recommended by the QIBA DWI profile and will be referred to as the ‘QIBA DWI phantom’ in this text.

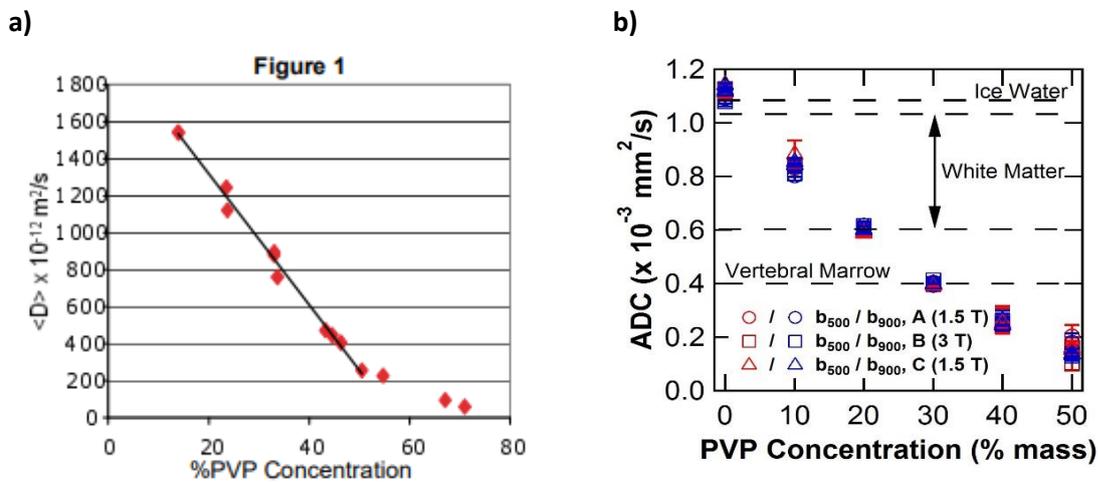


Figure 9. a) Relationship between %PVP and the measured diffusion coefficient at room temperature, taken directly from Pierpaoli et al. (2009) [47]. b) Measured ADC value for 0 to 50% PVP solutions at 0°C, with comparison to physiological ADC values, taken directly from Boss et al. (2014) [64].

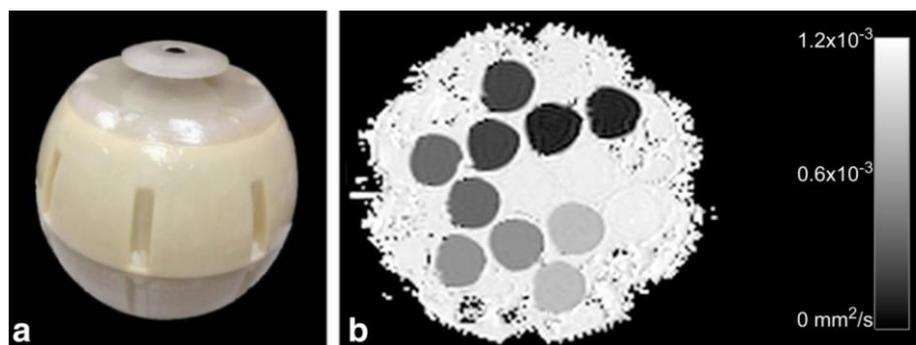


Figure 10. PVP-based diffusion phantom developed by the National Cancer Institute, RSNA QIBA, and NIST, presented by Boss et al. [64]. **a)** photo of the phantom. **b)** ADC map showing the array of solutions with different diffusivities. Images taken from Keenan et al. (2019) [66].

2.2.6 QIBA DWI Phantom

From 2019, the most widely used phantom for calibration of DWI in literature is the QIBA DWI phantom [7, 8, 57, 67-72]. The phantom is commercially available with NIST traceable solutions [73]. The range of PVP solutions allows linearity to be estimated over a physiologically relevant ADC range, however, at 0°C, this does not cover the full physiological range. Additionally, significant preparation is required to scan the phantom at 0°C. In a 2018 review, Keenan et al. [25] note that users find temperature control of the phantom difficult and measuring the phantom temperature would be preferable.

Lewis et al. [70] used the QIBA DWI phantom to evaluate DWI on a 0.35 T MR-LINAC, a diagnostic 3T scanner, and a 1.5T radiotherapy MR-Simulator. At 1.5T, bias ranged from 0.1% for water to -7.8% for the 50% PVP solution and -0.08% to -4.62% at 3T. Linearity and other performance metrics were not assessed. A significant difference was found between ADC values at 0.35T compared to 1.5T and 3T. Consequently, through use of the QIBA DWI phantom, they identified that further characterisation of ADC bias is required before the MR-LINAC can be used for qMRI. However, by focusing purely on bias, other equally important performance metrics for quantitative DWI are overlooked. Lewis et al. found the requirement for an ice bath led to significant susceptibility-related heterogeneities which produced large artefacts at both 1.5T and 3T. They highlighted it is advantageous to scan the phantom in coronal orientation to prevent air bubbles within the imaging plane; however, for patient scans, axial orientation is more typical.

Comparing institutional DWI protocols to a standardised benchmark protocol across 15 sites and both 1.5T and 3T scanners, Van Houdt et al. [69], found %bias of the central water vial of 0 to 3% for the benchmark protocol and 1 to 4% for the institutional protocols. They concluded that measurement of consistent ADC values is feasible despite protocol differences (**Figure 11**), including differences in the b-values used. Using the benchmark protocol, random error was higher than the QIBA requirements in 12 sites and RC was higher in 7. With institutional protocols, both the random error and RC were higher at only 3 institutes.

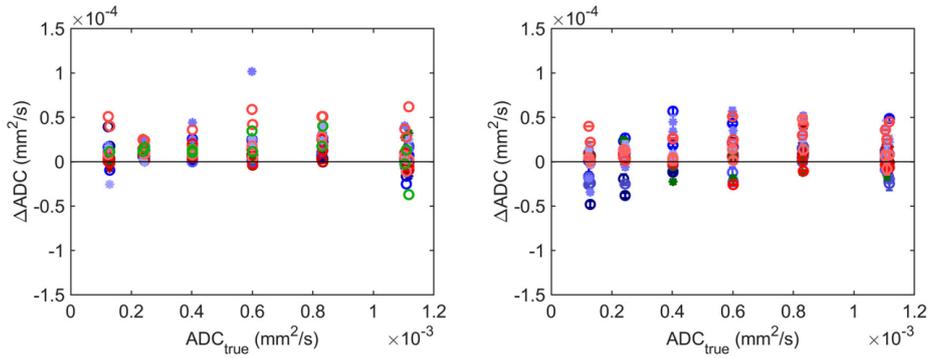


Figure 11. Absolute bias of ADC measurements against true ADC values for the standardised (left) and institutional (right) protocols, colours represent different vendors and field strengths. Taken from van Houdt et al. (2020) [69].

Wang et al. [57] assessed the bias, repeatability, and reproducibility of a clinical prostate protocol. Only 4 out of the 13 vials were used (central water vial, and inner ring 10-30% PVP vials), justified by the chosen vials covering a clinically relevant range for the prostate. For these vials, bias was between -8.0% and 2.7% with repeatability CV <2.40% and reproducibility CV <3.98%. McDonald et al. [8] calculated the bias for all vials; however, other performance metrics were only calculated for the central water vial. Results on an MR-Sim showed ADC bias within $0.1 \times 10^{-3} \text{ mm}^2/\text{s}$ for all vials using different DWI sequences. The central water vial precision results met the QIBA DWI profile tolerances when using an EPI sequence, however, not when using other DWI sequences.

To fully utilise the phantom, Carr et al. [72] performed repeat measurements of the QIBA DWI phantom over 12 months on a 3T scanner using the QIBA recommended QA protocol. All performance requirements in the QIBA profile were assessed in three orthogonal imaging directions (axial, coronal, and sagittal). This provided a long-term system stability assessment and allowed the recommended frequency of future QA to be determined. There was minimal imaging direction dependence on ADC performance. No significant difference was found between the inner and outer ring vials. For all orientations, the central water vial results were within the QIBA specification. Vials with lower diffusivities (40% and 50% PVP) had inferior performance metrics (**Figure 12**). Carr et al. reported considerable time required to prepare the ice bath and an associated increase in susceptibility-induced distortions, agreeing that the future direction is towards ADC performance assessment at room temperature with room temperature reference values.

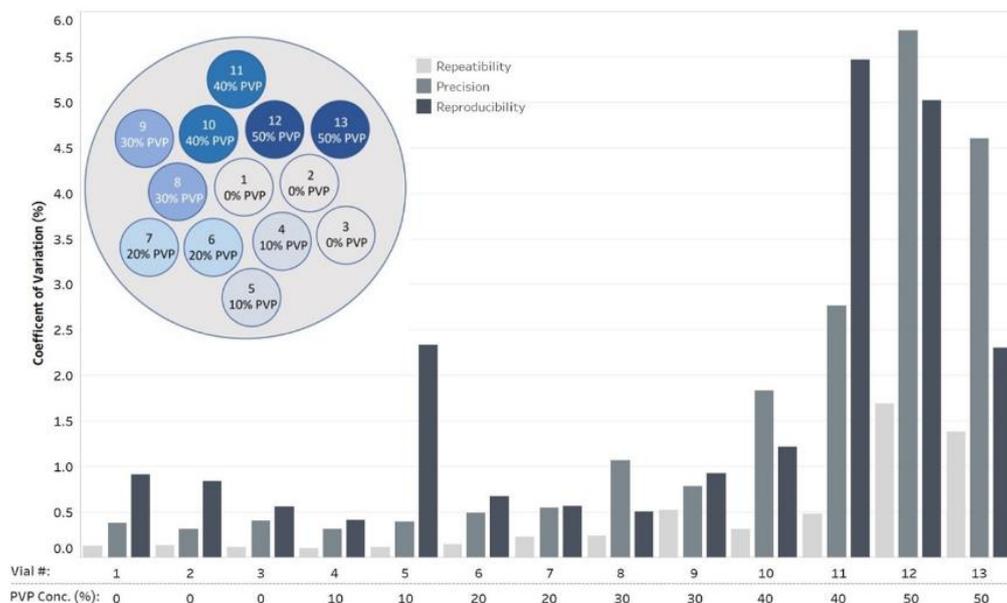


Figure 12. Average axial orientation repeatability coefficient of variation (CV_{ST}), precision (CV_P), and reproducibility (CV_{LT}) for each vial. Included in the top left is a diagram indicating the vial arrangement and PVP concentrations within the phantom. Taken from Carr et al. (2022) [72].

2.2.7 CaliberMRI Diffusion Phantom

The QIBA DWI phantom is functional for standardising DWI QA, allowing characterisation of the uncertainties associated with ADC measurements. However, requirements for an ice bath make the measurement process logistically challenging for routine QA. Additionally, although the 0-50% PVP covers a wide range of ADC values, it does not cover the full physiological range expected in the body at 0°C [25]. The CaliberMRI diffusion phantom has identical configuration to the QIBA DWI phantom described above, however, it also contains a liquid crystal (LC) MR-readable thermometer designed to be used at room temperature with NIST-traceable room temperature reference ADC values. This phantom is presented in more detail in chapter 3, in which the phantom configuration, preparation, and set-up are discussed. The CaliberMRI diffusion phantom can also be filled with ice-water and used identically to the QIBA DWI phantom to assess conformance to the QIBA DWI profile. At the time of writing, there is no literature published on room temperature ADC QA using the CaliberMRI DWI phantom with the LC thermometer. Therefore, sequence optimisation for imaging the LC thermometer is explored in chapter 4, and a comparison of ice bath and room temperature ADC calibration is explored in chapter 5.

3 METHODS

The CaliberMRI diffusion phantom allows calibration of ADC values at room temperature and assessment of scanner conformance to the QIBA DWI profile using an ice bath. Within this chapter, the phantom is presented, and the preparation required for both room temperature and ice bath scans are outlined. The DWI QA protocol, following QIBA requirements, is defined, and finally, the performance metrics used to assess conformance to the QIBA profile are discussed.

3.1 CALIBERMRI DIFFUSION PHANTOM

The CaliberMRI diffusion phantom (Serial number 128-0163, CaliberMRI, Boulder, CO, USA) contains a range of polyvinylpyrrolidone (PVP) solutions within a water-filled spherical plastic shell with a 194 mm outer diameter (**Figure 13**). A central vial plate separates the top and bottom hemispheres.

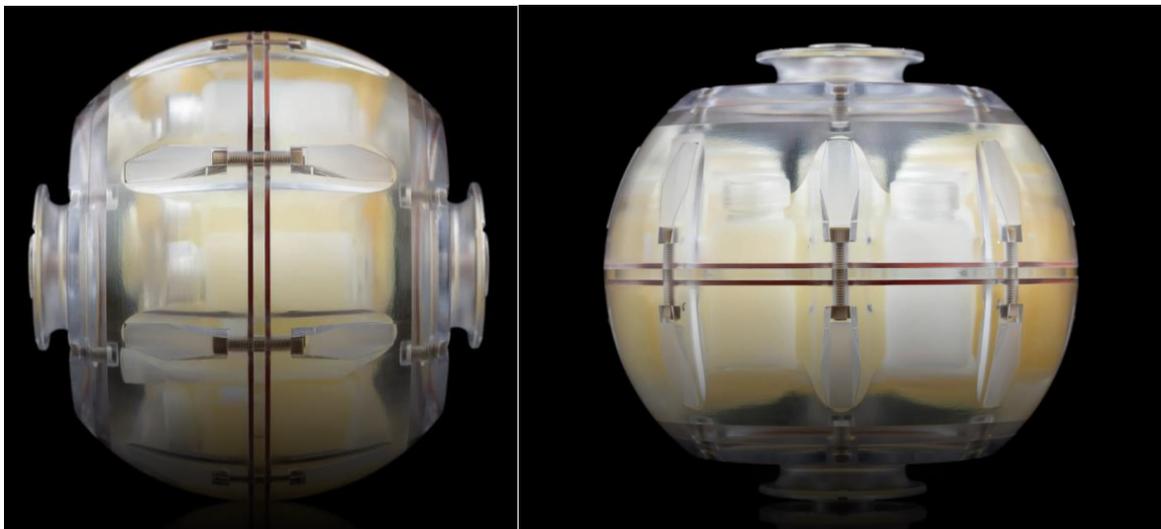


Figure 13. CaliberMRI diffusion phantom. Images taken directly from qmri.com/diffusion [26-May-2022].

The central vial plate houses thirteen 30 mL vials containing 0, 10, 20, 30, 40, and 50% w/w PVP and 10 liquid crystal (LC) thermometer vials for measuring temperatures from 15°C to 24°C. The PVP vials form two concentric rings containing one of each solution, organised so the PVP concentration is decreasing counterclockwise around each ring. A central 0% vial (distilled water) is designed to sit at the isocentre of the scanner when the phantom is positioned as intended. Three 5 mL vials of water extend above and below the 30 mL vials to act as fiducial markers when reviewing the MR images. **Figure 14** shows a labelled cross-sectional T1-weighted view of the vials.

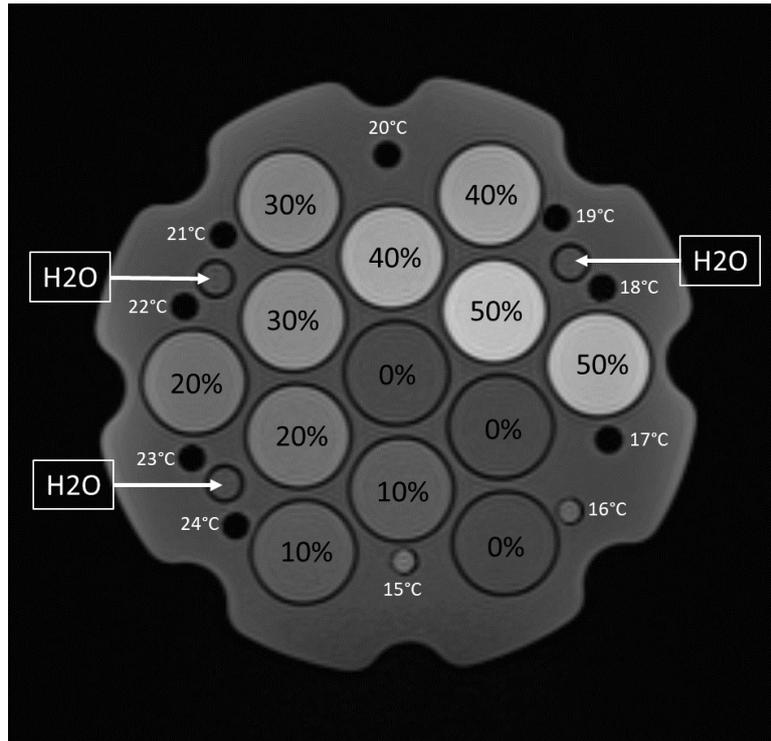


Figure 14. T1-weighted image of the CaliberMRI diffusion phantom. The 30 mL solution vials are labelled with the %PVP w/w, and the thermometer vials are labelled with their approximate transition temperature. The three 5 mL water vials are also labelled.

3.1.1.1 PVP Solutions

PVP is a water-soluble polymer that has been widely adopted as an ideal material for quantitative DWI phantoms (Section 2.2). CaliberMRI provide NIST traceable ADC values for the 0-50% PVP vials over a range of temperatures at 3T, shown in **Table 2**. Tofts et al. [48] stated that an appropriate range of ADC values for a test object at room temperature would be $0.3 - 2.1 \times 10^{-9} \text{ m}^2/\text{s}^{-1}$. At 22°C, the CaliberMRI phantom covers this full range.

For room temperature measurements, reference ADC values are obtained by linear interpolation between the NIST ADC values for the closest temperatures. For example, with a thermometer reading of $18.6 \pm 0.8^\circ\text{C}$, reference ADC values for 18.6°C are found by linear interpolation between the NIST data for 18°C and the NIST data for 20°C .

Table 2. ADC measurements ($\times 10^{-3} \text{ mm}^2/\text{s}$) of 0-50% PVP solutions at different temperatures at 3T by NIST Boulder. The ADC uncertainty ($\times 10^{-3} \text{ mm}^2/\text{s}$) is provided in brackets. Values provided by CaliberMRI [74].

PVP CONCENTRATION (%)	0 °C	16 °C	18 °C	20 °C	22 °C	24 °C	26 °C
0	1.109 (0.025)	1.799 (0.039)	1.898 (0.041)	2.000 (0.043)	2.106 (0.045)	2.232 (0.048)	2.335 (0.050)
10	0.817 (0.019)	1.376 (0.031)	1.462 (0.032)	1.551 (0.034)	1.640 (0.036)	1.742 (0.038)	1.825 (0.039)
20	0.579 (0.015)	1.047 (0.024)	1.112 (0.025)	1.183 (0.027)	1.258 (0.028)	1.322 (0.029)	1.390 (0.031)
30	0.380 (0.011)	0.722 (0.017)	0.760 (0.018)	0.832 (0.020)	0.886 (0.021)	0.929 (0.022)	0.989 (0.023)
40	0.220 (0.007)	0.440 (0.012)	0.471 (0.012)	0.507 (0.013)	0.545 (0.014)	0.584 (0.015)	0.625 (0.016)
50	0.110 (0.005)	0.231 (0.008)	0.246 (0.008)	0.267 (0.008)	0.293 (0.009)	0.323 (0.009)	0.352 (0.010)

3.1.1.2 LC MR-Readable thermometer

The MR-visible thermometer, developed by NIST and CaliberMRI, is comprised of the 10 LC vials located around the inside edge of the phantom (**Figure 14, Figure 15**). The thermometer vials are filled with a cholesteric liquid crystal that has temperature-dependent isotropic liquid and crystalline phases [75]. The temperature at which the LC changes phase is known as the transition temperature, and each LC is formulated so the transition occurs at a specified temperature [76]. The transition temperatures of the vials in the phantom range from 15°C to 24°C.

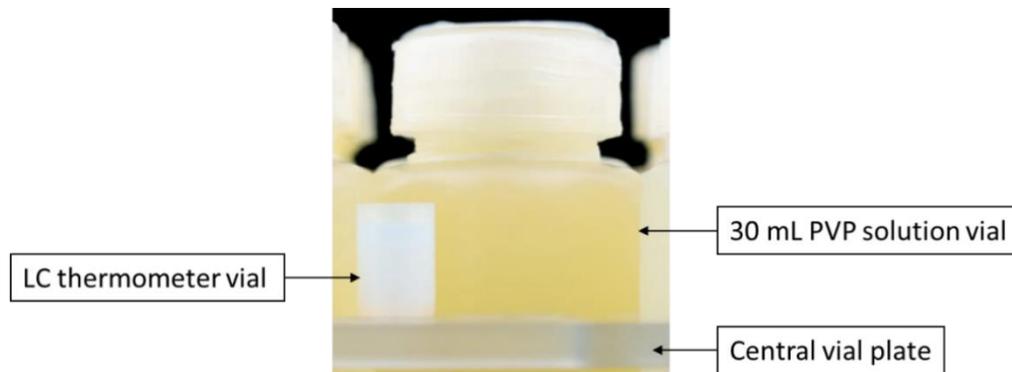


Figure 15. The thermometer vials are located on top of the central vial plate. When the LC is in a crystalline state, it appears a pearlescent white as shown in the image. Image taken from [73].

To determine the temperature of the phantom, the thermometer vials are imaged using a T1-weighted sequence. If the LC is in an ordered crystalline phase, that is, if the temperature is below its transition temperature, then the MR signal decays rapidly. Consequently, no signal is detected, and the vial appears dark. In contrast, in the isotropic liquid phase, there is signal from the vial, and it appears bright relative to vials in the crystalline phase, indicating the phantom temperature is above its transition temperature. Temperature data is embedded in the images during the measurement protocol, rather than being recorded separately, giving an integrated temperature measurement that can be automatically detected by analysis software.

3.2 SET-UP

3.2.1 Phantom preparation: room temperature

For room temperature measurements, CaliberMRI recommends that the phantom sits in the scanner room for a minimum of six hours to allow for temperature stabilisation. For all measurements at room temperature, the phantom was placed in the scanner room a minimum of one day before scanning to allow the temperature to stabilise overnight.

3.2.2 Phantom preparation: Ice-water temperature control

At this current time, ADC calibration must be performed at 0°C for QIBA conformance. The CaliberMRI diffusion phantom can be filled with a crushed ice and water bath to allow measurements at 0°C instead of using the LC thermometer. However, it should be noted that due to the LC thermometer CaliberMRI highly discourages users from filling this model with ice. After discussion with the CaliberMRI team, it was found that there is no risk of exposing the thermometer to 0°C, however, there is a small risk of damage to the LC vials if the ice shards are sharp and loaded into the phantom with force. Full instructions for ice bath measurements are provided in the QIBA DWI Supplement 1 [77] and summarised in Appendix A. The temperature of the phantom was measured before and after scanning using a digital thermometer (DTM3000-spezial, LKM Electronic GmbH, resolution: 0.01°C, accuracy: +/- 0.04°C).

3.2.3 Phantom positioning

The phantom can be positioned in coronal, axial, or sagittal orientation. To characterise all three imaging planes, the phantom must be physically rotated to each orientation. In all orientations, the centre of the central vial should be at the scanner isocentre.

CaliberMRI provides guidance on the positioning of the phantom in each orientation, using the serial number sticker as a marker and describing the rotations required. To improve set-up efficiency, coloured stickers were placed onto the phantom, using a different colour for each orientation. For each orientation, the triangle created from three stickers of the corresponding colour should be anteriorly, with the top of the triangle pointing into the scanner bore, shown in **Figure 16**. The phantom was stabilised using cushioning and the positioning lasers were aligned with external features on the phantom to allow a repeatable set-up for each scan. The same cushioning was used on each scanner throughout the experimental work.

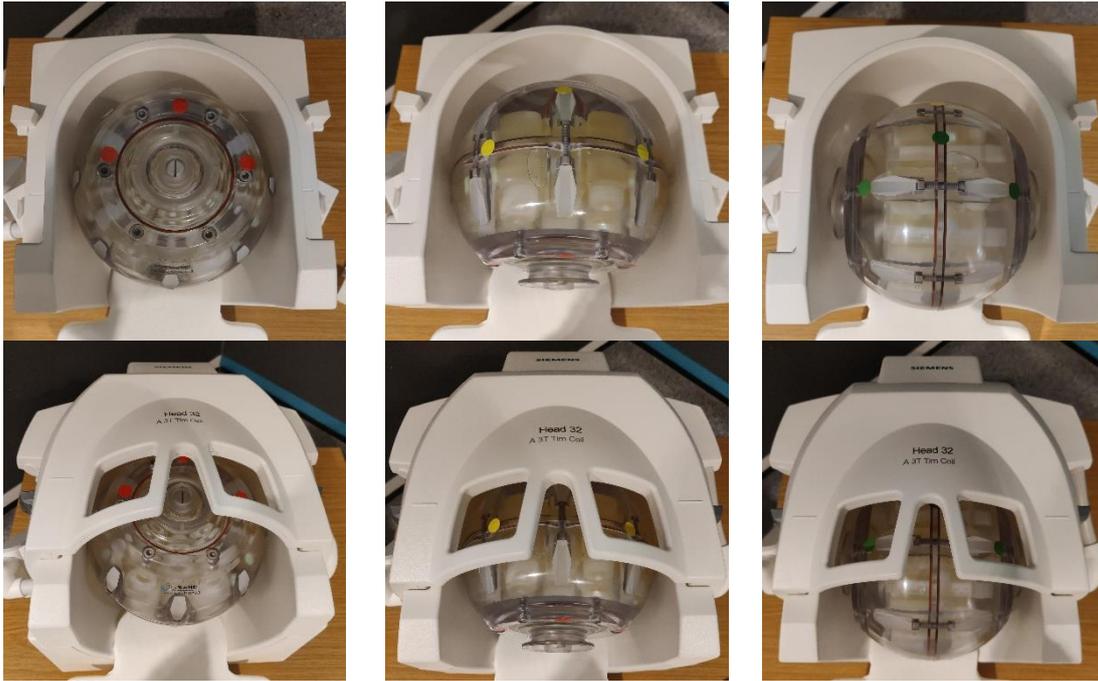


Figure 16. Coronal (left), axial (centre), and sagittal (right) orientations of the phantom to obtain a cross-section through the vials in the plane of interest, with (top) and without (bottom) the upper half of the 32-channel head coil attached.

3.3 MRI HARDWARE

Two MR scanners are focused on in this work, as these scanners were identified as ones on which quantitative DWI was likely to be used. The first is a Siemens Prisma, 3T, with a 60 cm bore. This is a research scanner located at the Clinical Research Facility (CRF), Queen Elizabeth University Hospital (Glasgow, UK). Its imaging gradients have a peak gradient strength of 80 mT/m per axis and slew rate of 200 T/m/s per axis. The second is a dedicated radiotherapy MR-simulator, located at the Beatson West of Scotland Cancer Centre (BWOSCC) (Glasgow, UK). The scanner is a Siemens MAGNETOM Sola, 1.5T, with a 70 cm bore. Its imaging gradients have a peak gradient strength of 45 mT/m per axis and slew rate 200 T/m/s per axis.

Three different head coils are used in this project, a 20-channel head and neck coil, a 32-channel head coil and a 64-channel head coil. All three coils were available on the 3T Siemens Prisma but only the 20-channel coil was available on the 1.5T Siemens Sola. Unless stated otherwise, the 20-channel coils were used on both scanners.

3.4 SCANNING PROTOCOL

Within the QIBA DWI profile, there is a generic phantom QA protocol as well as some scanner model-specific QA protocols and clinical protocols for brain, liver, prostate, and breast. Many parameters have acceptable, target, and ideal values. Acceptable requirements must be met to conform to the profile. Target and ideal requirements can improve results, however, ideal requirements may require more effort and non-standard hardware and software to achieve [10].

Initially, a protocol was created on each scanner following the recommended protocol in the CaliberMRI diffusion phantom manual (Rev G, August 2021) [74]. This protocol met the acceptable criteria of the QIBA QA protocol. However, for compatibility with their analysis software, CaliberMRI

later provided an updated protocol (Appendix B). Therefore, parameters were updated to reduce acquisition conformance warnings within the software.

For full assessment of one imaging plane, four DWI repeats in immediate succession with no changes between scans were performed. The DWI sequence was a single-shot EPI sequence with a 3-scan trace and five b-values: 0, 500, 1000, 1500, 2000 s/mm². The repetition time (TR) for both field strengths was 8000 ms and the echo time (TE) was the minimum achievable: 81 ms and 96 ms for 3T and 1.5T, respectively. Both the number of averages and parallel imaging factor were 2 and scan time was 234 s. Slices were perpendicular to the long-axis of the vials. Specific scan protocols used are specified within each results section. The final diffusion phantom QA scan protocols and parameters used are included in Appendix C, and available to download on the department website [78].

To use the LC thermometer, a T1-weighted scan is required. Optimisation of this T1-weighted sequence is discussed in chapter 4. A T1-weighted image was performed at the beginning and end of each session to assess any temperature change during scanning.

3.5 QCAL SOFTWARE

qCal-MR[®] (CaliberMRI, Boulder, CO, USA). is an automated quantitative MRI QC software developed by CaliberMRI, designed for analysis of their phantoms. The analysis has been reviewed and approved by the QIBA DWI Biomarker Committee, meaning it can be used to assess profile conformance. DICOM files are uploaded, and the phantom temperature can be manually entered, detected from a T1-weighted image, or ‘Ice water’ can be selected if using an ice bath. qCal-MR took less than 5 minutes to analyse a dataset. Conformance warnings occur if the protocol does not match the recommended parameters (Appendix D), and calculation warnings are given if any potential problems with calculations were identified. Both graphical and table outputs are created.

For series to be recognised as repeats and therefore analysed together, the ‘Protocol name’ DICOM header must be the same, as well as the slice position. Regions of interest (ROIs) are automatically found, and performance metrics are calculated (Section 3.6). ROI placement cannot be manually edited; however, volumes of interest (VOIs) can be modified to exclude specified slices, allowing results to be recalculated without aberrant data. If an ice bath scan is performed, results are checked against the QIBA DWI profile tolerances and CaliberMRI offers QIBA conformance certification for the DWI profile.

If the LC thermometer option is chosen and a set of T1 images with the required parameters are uploaded to the software, the temperature is determined from the number of bright LC vials compared to the number that are dark. If more than one temperature scan is uploaded, qCal determines the temperature at the time of each DWI scan by linear interpolation between the temperatures based on the time the images were acquired. Therefore, if the temperature increases by a measurable amount during the scans, earlier scans use ADC reference values at a lower temperature than later scans.

3.6 BIOMARKER PERFORMANCE ASSESSMENT

3.6.1 Accuracy

Per the QIBA DWI profile, ADC measurement accuracy is assessed via the bias and linearity of the measurements. b-value dependence and SNR can also impact ADC measurement bias and therefore also have tolerances specified within the profile (**Table 1**).

3.6.1.1 ADC Value

qCal uses the Trace-DWI DICOM images to calculate the ADC maps using a mono-exponential signal decay model, introduced in Section 1.1:

$$S_{bx} = S_{b0} \cdot e^{(-bx \cdot ADC)}$$

Equation 8

S_{bx} is the value of the pixel in the DWI image with b-value $x \text{ s/mm}^2$ and S_{b0} is the value of the pixel in the DWI image with b-value 0 s/mm^2 . With more than two b-values, the ADC value is determined using a least squares fit. [79]

Mean ADC value per VOI is reported for each series, as well as the average ADC value per vial for the four repeats. *ADC bias* is defined as the difference between the measured ADC value and the NIST reference ADC value. The *%bias* is the ADC bias expressed as a percentage of the NIST reference value.

To assess linearity, the 95% confidence interval (CI) of the slope of a plot of measured ADC value against the NIST reference is obtained. qCal produces this plot and states the slope, slope error, intercept, and R^2 value.

To determine if statistically significant differences exist between ADC values measured under different conditions, and to compare to values in the literature, IBM SPSS Statistics (Version 28.0.0.0, IBM Corp., New York, USA) was used. To assess if data followed a normal distribution, a Shapiro-Wilks test for normality was performed. If the normality assumption was met, a one-sample t-test was used to determine if a statistically significant difference existed between the measured mean ADC values and NIST ADC values. Independent samples t-test was performed to determine if a statistically significant difference exists between different sessions. Statistical significance is assessed at the 5% level such that a p-value <0.05 is a statistically significant result. Comparing results between three head coils, ANOVA was performed, with a Tukey post hoc test used to determine between which pairs a significant difference existed. SPSS was also used to obtain 95% CIs for mean ADC values and linearity.

3.6.1.2 B-value dependence

Ideally, there should be minimal change in measured ADC value with choice of b-value. From ADC maps created between the available pairs of b-values, b-value dependence is calculated using the following formula:

$$ADC \text{ b value Dependence} = 100\% \cdot \left| \frac{ADC_{b_{min}b_j} - ADC_{b_{min}b_i}}{ADC_{b_{min}b_i}} \right|$$

Equation 9

Where b_1 and b_2 are not equal, $(b_2 - b_{min})$ and $(b_1 - b_{min})$ should be $\geq 400 \text{ s/mm}^2$ for adequate diffusion contrast, and b_{min} is the lowest b-value [10]. The maximum b-value dependence for each VOI is reported in qCal.

3.6.1.3 DWI SNR

DWI SNR is evaluated to confirm that ADC bias could be measured without incremental bias due to low SNR [10]. To calculate DWI SNR, qCal follows the guidelines in the QIBA DWI profile. A study should contain four series of diffusion-weighted images from immediate repeats. For each series, the DWI image with b-value 0 is selected. From the b_0 image for all repetitions, a signal image is defined as the mean value of each pixel across the b_0 scan and a noise image is defined as the standard deviation of each pixel across the b_0 scan. The b_0 DWI SNR is the ratio of the ROI mean value from the signal

image to the ROI mean value from the noise image [79]. The same procedure is repeated for each b-value.

3.6.2 Precision

For conformance to the QIBA DWI profile, ADC measurement precision of immediate repeats (short-term repeatability) is assessed via random error, RC_{ST} , and wCV_{ST} . Precision between session on the same scanner (long-term repeatability) is assessed using the RC_{LT} and wCV_{LT} . QIBA tolerances for these metrics are stated in **Table 1**.

3.6.2.1 Random measurement error

qCal calculates random error by defining a signal image and a noise image as the mean value and standard deviation of each pixel, respectively, across the ADC maps from each series. The random error is the ratio of the ROI mean value from the noise image to the ROI mean value from the signal image [10, 79]:

$$\%CV = 100\% \cdot \frac{\sigma}{\mu}$$

Equation 10

3.6.2.2 Repeatability coefficient (RC)

The RC represents the measurement precision when conditions of the measurement procedure are unchanged between measurements [10, 80]. It defines the least significant difference between two repeated measurements taken under identical conditions at a confidence level of 95% [81]. If the measured difference is less than the RC, it may be due to measurement imprecision rather than a true change [81]. The short-term (intra-session) RC (RC_{ST}) is calculated per vial in qCal from the standard deviation of the mean ADC values for each VOI across the different series (σ_W):

$$RC = 1.96 \cdot \sqrt{2 \cdot \sigma_W^2} = 2.77 \cdot \sigma_W$$

Equation 11

3.6.2.3 Within-subject Coefficient of Variation (wCV)

wCV is commonly used to assess repeatability in test-retest studies [10]. A small wCV indicates high precision, whereas a large wCV indicates low precision and a large change in ADC would be required to be confident a real change has occurred [81]. qCal uses the mean (μ) and standard deviation (σ_W) of the ADC values calculated for each VOI across the different series are used to calculate the wCV:

$$wCV = 100\% \cdot \frac{\sigma_W}{\mu}$$

Equation 12

3.6.2.4 Long-term Repeatability

qCal is not currently able to analyse data from different studies and therefore cannot be used to calculate reproducibility or long-term repeatability metrics. Long-term system repeatability is characterised by RC_{LT} and wCV_{LT} , which are equivalent to RC_{ST} and wCV_{ST} , however, use multiple longitudinal phantom scans rather than immediate repeats. Following the method by Wang et al. [82], mean ADC values of repeated measurements within a session are used for calculation of between-session repeatability, and wCV_{LT} and RC_{LT} are calculated per the QIBA DWI profile guidance for test-retest studies [10].

3.6.2.5 Reproducibility

An additional assessment of precision, reproducibility between measurement methods, can be characterised using the reproducibility coefficient (RDC). RDC, similar to the RC, is the 95% precision

limit [83], giving the minimum detectable difference when different imaging methods are used, and is calculated as

$$RDC = 2.77 \cdot \sigma_D$$

Equation 13

Where σ_D is the within-subject standard deviation under reproducibility conditions [83].

4 SEQUENCE OPTIMISATION FOR THE LC THERMOMETER

The LC thermometer is a novel method of allowing integrated temperature measurement of an MR phantom. A T1-weighted scan is required to view LC vials and determine the temperature of the phantom. Within this chapter, sequence optimisation is investigated, focusing on the contrast between the vials and the background and the scan time to maximise the efficiency of the QA image acquisition process.

4.1 RESULTS

4.1.1 Initial findings

The T1-weighted scan for temperature determination was mainly investigated on the 3T scanner. However, initial scans were also performed on the 1.5T scanner. A T1-weighted scan with TR 550 ms and TE 7.4 ms was used, with a slice thickness of 5 mm. Using this sequence, it was possible to view the LC thermometer (**Figure 17**).

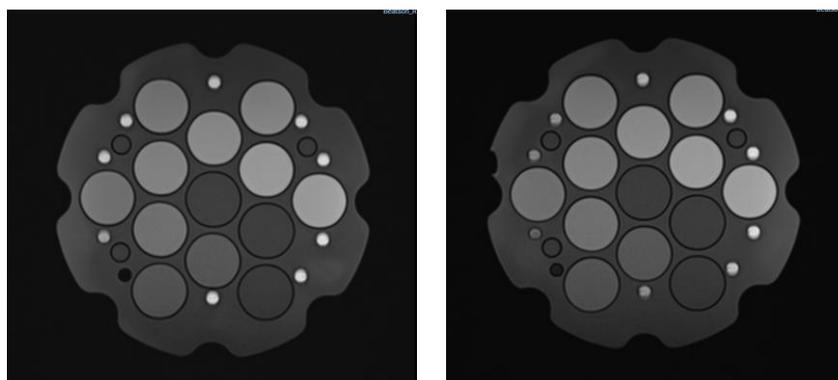


Figure 17. T1-weighted images of the LC thermometer using a long TR (550 ms) and large slice width (5mm). 9 vials appear bright. Start of session (left) and end of session (right).

For the image to be accepted and analysed by qCal, TR must be less than 25 ms (full requirements in Appendix D). Consequently, the software was not able to determine the temperature from the above images. The software requirements are in opposition with the example scans specified in the patent for the LC thermometer, which suggests TR of 200 ms or 60.48 ms when using a gradient echo or spoiled gradient echo, respectively [76].

In updated documentation provided by CaliberMRI (Appendix B), a 3D Spoiler Gradient Echo (TR 6 ms and TE minimum) was recommended. This is equivalent to a Siemens FLASH (Fast Low Angle Shot) sequence. FLASH is a spoiled gradient echo sequence that uses a low flip angle and a short TR, however, the 3D FLASH sequence available on the scanner had limited TR (21 ms) and TE (5.38 ms), which were not short enough to reach the recommended parameters. VIBE (Volumetric Interpolated Breath-hold Examination) is a modified form of FLASH, that can give a higher SNR [84]. Zero-filling interpolation of k-space allows acquisition time to be decreased whilst maintaining image quality. The VIBE sequence allowed all parameters to be set within the required ranges. A TWIST (Time-resolved angiography With Interleaved Stochastic Trajectories) sequence, which allowed a shorter scan time, was also tested. TWIST is a 3D magnetic resonance angiography (MRA) technique. TWIST achieves

fast scan times by sampling the centre of k-space, which provides information on contrast, more frequently than the periphery, which contributes to high resolution [85].

4.1.2 VOI Placement

qCal automatically locates the LC vial VOIs and the placement on each slice can be seen within the software. It was found that occasionally, VOI placement would be incorrect. An example is shown in **Figure 18**. If the VOI included the surrounding water rather than the thermometer vial, the software would conclude that the ‘vial’ is bright as the VOI signal is higher than a dark, crystalline-phase LC vial. This leads to an incorrect temperature and consequently incorrect ADC reference values.

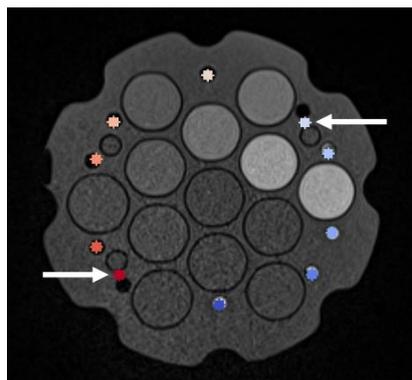


Figure 18. Incorrect ROI placement for the LC Vials can lead to errors in the temperature measurement. The fifth vial is ‘dark’; however, the VOI is misplaced and therefore ‘bright’, increasing the reported temperature.

4.1.3 Sequences

An example slice obtained from each sequence is shown in **Figure 19**. Using FLASH, no vials appear bright, despite two being in the liquid state during scanning and qCal was unable to identify the thermometer in the scan. The TWIST sequence gives poorer image quality, however, the contrast of the bright LC vials to the surrounding water in the phantom is very good. TWIST has the shortest scan time of the three (39 s) which is advantageous to reduce the time taken for QA. Although VIBE was developed as a breath hold sequence, when modifying parameters in accordance with the requirements provided by CaliberMRI, scan time was 205 s. Using the VIBE sequence, a shift of the signal from the LC vials in the liquid state is seen in the frequency encoding direction. This could result in an incorrect temperature reading and consequently incorrect reference ADC values.

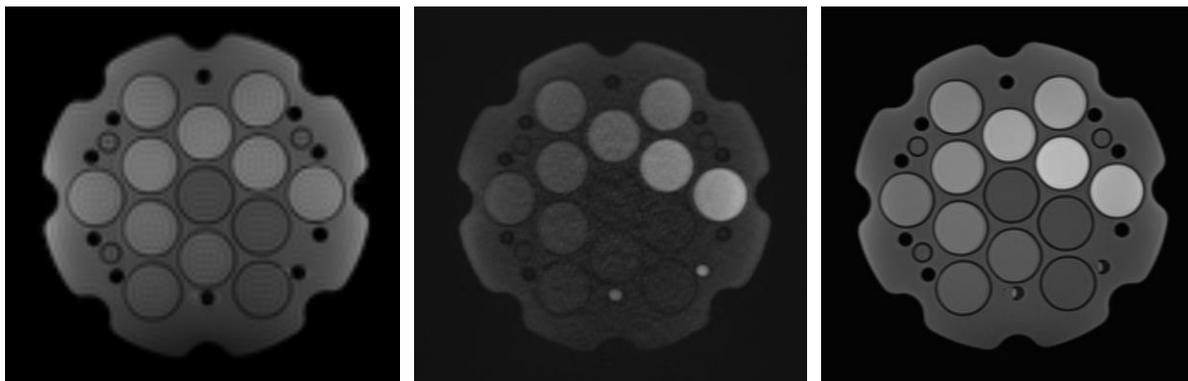


Figure 19. *T1-weighted images of the LC thermometer using three different sequences: FLASH (left), TWIST (centre), and VIBE (right).*

Chemical shift artefacts occur due to different molecular environments of nuclei causing local variations in magnetic field strength, which result in small changes in the resonant frequencies of molecules. Most commonly, they are associated with interfaces between tissues with significantly different fat and water content, although they can occur at the interface of any two substances with different chemical shifts [86]. Depending on the number of pixels shifted, the artefact can appear as light and dark bands on opposite sides of a structure, or as a ghost image [3].

The presence of the artefact suggests the LC resonates at a different frequency to the water. Even though the LC and its surroundings are excited with the same frequency, the frequency response of the LC is different, causing misregistration of the signals from the liquid-state vials so their location in the image is spatially mismatched. A lower bandwidth or higher magnetic field strength causes a bigger shift and therefore a more significant artefact. The artefact can be reduced by increasing bandwidth, at the cost of reduced SNR [86].

The VIBE sequence was repeated with three different bandwidth values. It can be seen in **Figure 20** that the shift is reduced as the bandwidth is increased. All three VIBE scans were uploaded to qCal. When the vibe series with a bandwidth of 180 Hz/Px was analysed, qCal incorrectly concluded that there were 0 bright vials and therefore that the temperature was outside the thermometer range. This results in incorrect NIST ADC reference values, and the bias and linearity of the scanner cannot be assessed if no other temperature information is available. There is no noticeable change in image quality across the three bandwidths, and therefore, a bandwidth of 810 Hz/Px is recommended.

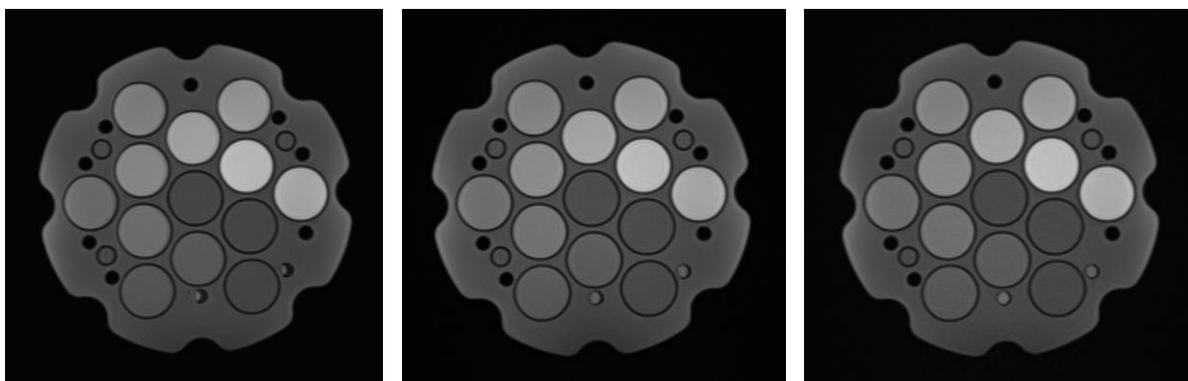


Figure 20. *Increasing bandwidth of the VIBE sequence: 180 Hz/Px (left), 455 Hz/Px (centre), and 810 Hz/Px (right). All images show the same slice and the phantom not moved between scans.*

4.1.4 Field Strength

VIBE and TWIST sequences were implemented on the 1.5T scanner, replicating parameters when possible. Images are shown in **Figure 21**. Both images appear noisier at 1.5T compared to 3T and the contrast between the vials and the background for the TWIST sequence is poorer at 1.5T. However, contrast of the vials on the VIBE sequence is excellent and a shorter scan time for the VIBE sequence (163 s) was achieved by reducing TR to 6 ms. qCal was unable to detect the thermometer in the TWIST images, and therefore, the VIBE sequence is recommended for the 1.5T scanner.

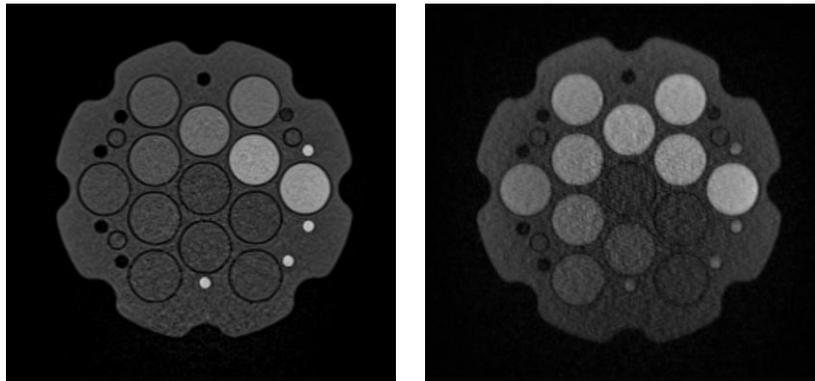


Figure 21. T1-weighted images of the phantom on the 1.5T scanner using a VIBE sequence (left) and TWIST sequence (right).

4.2 DISCUSSION

While the TWIST sequence provided a short scan time on the 3T system, at 1.5T image quality was inadequate for qCal to detect the LC vials. VIBE produced images suitable for qCal on both scanners and since scan time was less than the time of one DWI scan, it does not add a significant time to the full QA protocol. If qCal did not require the full phantom to be imaged in order to identify the LC vials then the slab thickness could be significantly reduced, reducing the scan time.

The main advantage of the LC thermometer is it allows non-invasive temperature measurement. Temperature readings can be performed immediately before and after DWI scans without moving or disturbing the phantom, reducing variability. Furthermore, temperature information is embedded in the image rather than stored separately, allowing all information to be recovered from the images.

For temperatures between those that NIST reference values are provided for, linear interpolation is used. Linear relationships have been demonstrated for water [58]; however, quadratic and exponential models have been proposed for other solutions [48, 87]. To check the suitability of the linear model, the data in **Table 2** was plotted and a linear model fits the data for each %PVP well over a temperature range from 16°C to 24°C ($R^2 > 0.98$).

A limitation of the thermometer is its accuracy, with a temperature uncertainty of $\pm 0.7^\circ\text{C}$ to $\pm 1.2^\circ\text{C}$ depending on the LC vial. For more accurate temperature measurements, more vials would be required [75]. Alternatively, if qCal could distinguish between partially transitioned and fully transitioned vials then the intermediate state of the LCs could be utilised as the average signal over the VOI of a transitioning vial will be between that of a fully liquid or fully crystalline vial (**Figure 22**).

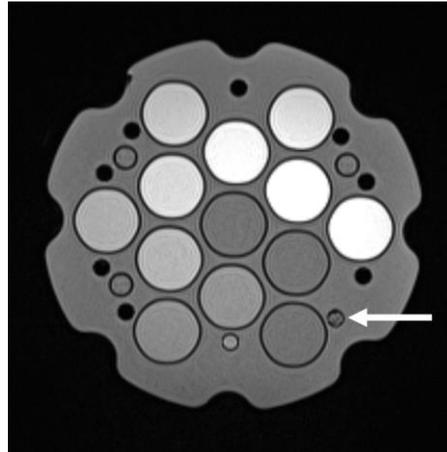


Figure 22. The second LC vial (indicated by the white arrow) has a signal level between that of the LC vials on either side, suggesting it is in an intermediate, transitioning state.

In summary, the LC thermometer offers a promising method for non-invasive phantom temperature measurement. However, both artefacts and poor ROI placement can cause incorrect temperature measurement, leading to incorrect reference ADC values. Bandwidth was found to be an important sequence parameter, to reduce a shift in signal from the liquid-phase LC vials. Two sequences are proposed for a 3T scanner, and a VIBE sequence with bandwidth 810 Hz/Px was transferable to 1.5T. This sequence was subsequently used as part of the QA protocol.

5 ADC ACCURACY AND PRECISION

As discussed in chapter 2, previous literature mentions the benefits of room temperature measurements over use of an ice bath [25]. However, for QIBA DWI profile conformance certification, a temperature-controlled ice-water phantom is required to compare the measured ADC value to the well-supported ADC value of water at 0°C, $1.1 \times 10^{-3} \text{ mm}^2/\text{s}$. Due to the lack of literature using the QIBA DWI phantom or CaliberMRI diffusion phantom at room temperature, it was of interest to not only perform a 0°C scan to assess QIBA conformance and compare scanner performance to the literature, but also to compare results obtained at 0°C and those obtained at room temperature using the LC thermometer to establish if room temperature ADC QA is feasible.

The phantom was used to characterise the bias and repeatability of ADC measurements made on the 3T Siemens Prisma and 1.5T Siemens Sola. To compare results at room temperature to results at 0°C, an ice bath scan was performed between two room temperature scans. At each scanner, the phantom was left in the scanner room overnight to reach a stable room temperature and scanned the next morning. Following scanning, an ice bath was prepared (Appendix A) and the phantom was left to equilibrate to 0°C for approximately 2 hours, ready to be scanned in the afternoon. After scanning the chilled phantom, it was left in the scanner room to return to room temperature. Due to scheduling, the phantom was left over the weekend, and consequently, the room temperature scan was repeated approximately 72 hours after the first scan.

The full QIBA DWI profile scan procedure requires image acquisition in coronal, axial, and sagittal planes [10] and therefore the phantom was scanned with 4 immediate repeats in each orientation. A T1-weighted scan was performed at the start and end of all sessions, so full data acquisition was the same for both 0°C and room temperature scans.

Within this chapter, first, visual assessment of each session is performed. The results for the central water vial are discussed in Section 5.2 because it is these results at 0°C that are required for QIBA conformance certification. Following this, the accuracy (Section 5.3) and precision (Section 5.4) of all vials are assessed using the performance metrics described in Section 3.6, to characterise ADC measurements over a wide, clinically relevant range of ADC values. Tabulated results from the four repeats in coronal orientation for each session are shown in Appendix E and F.

5.1 TEMPERATURE AND QUALITATIVE RESULTS

On the 3T scanner, immediately before scanning, ice bath phantom temperature was $-0.02 \pm 0.04^\circ\text{C}$ in the top half and $0.00 \pm 0.04^\circ\text{C}$ in the bottom half. After scanning, in a region where no ice remained, the temperature had increased to $0.10 \pm 0.04^\circ\text{C}$, however, in the majority of the phantom, temperature was $0.02 \pm 0.04^\circ\text{C}$. Both room temperature scans were at $15.6 \pm 0.9^\circ\text{C}$. The corresponding NIST reference value of water is $1782 \pm 40.4 \mu\text{m}^2/\text{s}$.

On the 1.5T scanner, initial ice bath phantom temperature was $0.04 \pm 0.04^\circ\text{C}$ in the top half and $0.01 \pm 0.04^\circ\text{C}$ in the bottom half. After scanning, phantom temperature was $0.01 \pm 0.04^\circ\text{C}$ in both halves. For both room temperature scans, phantom temperature was $20.7 \pm 0.9^\circ\text{C}$, corresponding to a NIST reference value for water of $2037 \pm 47.5 \mu\text{m}^2/\text{s}$.

Figure 23 and **Figure 24** show the same slice across all b-values from one repeat of the room temperature and ice bath measurements at 3T and 1.5T, respectively. Within literature, one suggested advantage of room temperature scans is a reduction in the occurrence of large artefacts due to significant susceptibility-related heterogeneities created by the ice bath [70, 72]. Although artefacts were seen

across all scans, the presence of ice increased the susceptibility artefacts seen and showed increased distortion of the vials, supporting this suggestion.

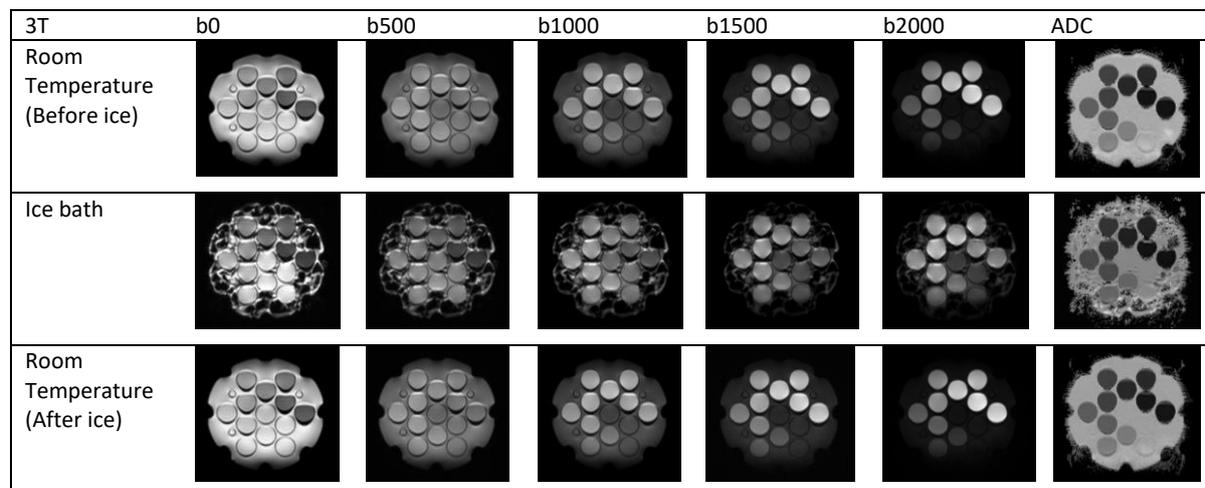


Figure 23. Trace DWI images at each b-value for room temperature and ice bath sessions on the 3T Siemens Prisma. All images show the same slice (two slices below the central vial plate) from the second coronal orientation repeat.

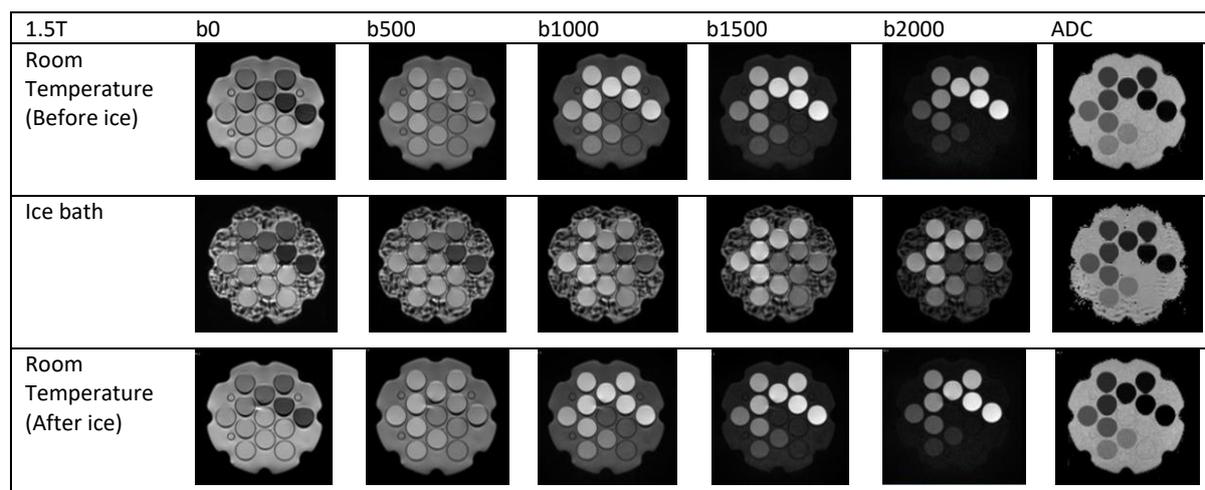


Figure 24. Trace DWI images at each b-value for room temperature and ice bath sessions on the 1.5T Siemens Sola. All images show the same slice (two slices below the central vial plate) from the second coronal orientation repeat.

5.2 CONFORMANCE TO THE QIBA DWI PROFILE

5.2.1 Results

Results for the central water vial for each session and orientation are provided in **Table 3** and **Table 4** for the 3T and 1.5T scanner, respectively.

On the 3T scanner, at 0°C, only one value for the centre water vial did not meet the QIBA requirement: the sagittal orientation 95% CI for the slope (0.937, 0.998) is not within the tolerance (**Table 3**). Very high bias (average bias 71.3%) for the outer 40% PVP vial was identified as the cause. Examining the images, artefacts from the 40% vial and adjacent 50% vial result in a non-homogeneous ADC map,

including a band of high intensity in the ROI on some slices. Excluding this vial, the 95% CI for slope is 0.987, 1.009 ($R^2=0.999$) which meets the QIBA requirements and is consistent with the other orientations. This highlights how the presence of artefacts within an ROI can introduce large biases to the measurement of ADC values, especially for vials with low diffusivity.

On the 1.5T scanner (**Table 4**), at 0°C, all coronal and axial results for the central water vial are within the QIBA DWI Profile tolerances. In sagittal orientation, the maximum b-value dependence (2.52% for b-value pair 500, 1500) is greater than the QIBA limit of 2%. A change in ADC value dependent upon the choice of b-values may suggest output gradient amplitude is not linear with input demand, background gradients are contributing to the actual b-value, there is spurious signal in b0, or SNR is inadequate at high b-values [10].

At 0°C, the central water vial measurement on both scanners has excellent agreement with the literature value of $1.1 \times 10^{-3} \text{ mm}^2/\text{s}$. However, a one-sample t-test comparing the mean ADC value using results from all orientations to the NIST value for the vial contents ($1109 \mu\text{m}^2/\text{s}$) shows a statistically significant difference ($p < 0.001$) for both scanners. At 3T, the mean bias from the NIST value is $2.63 \mu\text{m}^2/\text{s}$ (95% CI: $1.98 \mu\text{m}^2/\text{s}$, $3.28 \mu\text{m}^2/\text{s}$). At 1.5T, the mean bias from the NIST value is $11.11 \mu\text{m}^2/\text{s}$ (95% CI: $9.17 \mu\text{m}^2/\text{s}$, $13.05 \mu\text{m}^2/\text{s}$). The 95% CIs for both scanners are within the QIBA tolerance of $\pm 40 \mu\text{m}^2/\text{s}$. Therefore, the biases can be considered very good and unlikely to cause a clinically significant contribution to a measured change in ADC.

Both scanners are also able to meet all the profile requirements at room temperature in at least one orientation.

Table 3. Results for the central water vial on the 3T scanner. An ice bath ($-0.02 - 0.10 \pm 0.04^\circ\text{C}$) was performed between room temperature sessions ($15.6 \pm 0.9^\circ\text{C}$). Each session assessed 3 orientations, each with 4 immediate repeats. Red values indicate values are outside of QIBA tolerances.

SIEMENS PRISMA 3T	ROOM TEMPERATURE (BEFORE ICE)			ICE BATH			ROOM TEMPERATURE (AFTER ICE)		
	Coronal	Axial	Sagittal	Coronal	Axial	Sagittal	Coronal	Axial	Sagittal
NIST ADC ($\mu\text{m}^2/\text{s}$)	1782	1782	1782	1109	1109	1109	1782	1782	1782
MEAN ADC ($\mu\text{m}^2/\text{s}$)	1752.6	1759.0	1781.5	1110.4	1112.4	1111.9	1783.1	1790.8	1802.8
BIAS ($\mu\text{m}^2/\text{s}$)	-29.1	-22.7	-0.232	1.44	3.50	2.96	1.43	9.08	21.0
BIAS (%)	-1.64	-1.27	-0.0130	0.130	0.315	0.267	0.0803	0.510	1.18
MAX B-VALUE DEPENDENCE (%)	0.098	0.084	2.83	0.196	0.438	1.75	0.277	0.315	1.89
MAX B-VALUE PAIR	(1000, 2000)	(1000, 2000)	(500, 2000)	(1000, 1500)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)
SNR B0	157.5	181.5	264.3	259.9	325.6	185.3	150.2	143.7	239.4
RANDOM ERROR (%)	1.28	1.32	1.44	0.592	0.469	0.713	1.41	1.13	1.57
RC ($\mu\text{m}^2/\text{s}$)	4.46	8.09	10.9	1.07	1.50	1.76	0.254	1.96	3.60
WCV (%)	0.092	0.17	0.22	0.035	0.049	0.057	0.0052	0.040	0.072
LINEARITY: SLOPE	0.992	0.994	0.980	0.990	1.002	0.968^a	1.001	1.001	0.999
LINEARITY: SLOPE ERROR	0.0133	0.0080	0.0098	0.0096	0.0085	0.0304	0.0164	0.0077	0.0081
LINEARITY: R²	0.998	0.999	0.999	0.999	0.999	0.988	0.997	0.999	0.999

^aIf the outer PVP40 vial is excluded from the Sagittal orientation Ice bath linearity calculation, the slope is 0.998 with error 0.005, intercept 12.681 and R^2 0.999.

Table 4. Results for the central water vial on the 1.5T scanner. An ice bath (0.01 - 0.04 +/- 0.04°C) was performed between room temperature sessions (20.7 +/- 0.9°C). Each session assessed 3 orientations, each with 4 immediate repeats.

SIEMENS SOLA 1.5T	ROOM TEMPERATURE (BEFORE ICE)			ICE BATH			ROOM TEMPERATURE (AFTER ICE)		
	<i>Coronal</i>	<i>Axial</i>	<i>Sagittal</i>	<i>Coronal</i>	<i>Axial</i>	<i>Sagittal</i>	<i>Coronal</i>	<i>Axial</i>	<i>Sagittal</i>
ORIENTATION	<i>Coronal</i>	<i>Axial</i>	<i>Sagittal</i>	<i>Coronal</i>	<i>Axial</i>	<i>Sagittal</i>	<i>Coronal</i>	<i>Axial</i>	<i>Sagittal</i>
NIST ADC ($\mu\text{m}^2/\text{s}$)	2037	2037	2037	1109	1109	1109	2037	2037	2037
MEAN ADC ($\mu\text{m}^2/\text{s}$)	2001.6	1991.2	1974.4	1123.0	1116.2	1121.2	2005.9	2005.7	1980.2
BIAS ($\mu\text{m}^2/\text{s}$)	-35.5	-45.9	-62.7	14.0	7.15	12.2	-31.2	-31.4	-56.9
BIAS (%)	-1.74	-2.25	-3.08	1.26	0.645	1.10	-1.53	-1.54	-2.79
MAX B-VALUE DEPENDENCE (%)	2.14	1.65	3.05	1.50	1.17	2.52	1.69	1.11	3.22
MAX B-VALUE PAIR	(500, 1500)	(1000, 2000)	(1500, 2000)	(500, 2000)	(500, 2000)	(500, 1500)	(1500, 2000)	(1500, 2000)	(1500, 2000)
SNR B0 RANDOM ERROR (%)	133.4	144.8	105.5	144.6	164.2	133.1	125.3	147.5	116.8
RC ($\mu\text{m}^2/\text{s}$)	3.68	3.41	4.02	1.25	1.04	1.32	3.67	3.49	3.96
WCV (%)	5.75	9.58	11.29	1.21	1.25	1.75	11.02	7.11	9.44
LINEARITY: SLOPE	0.10	0.17	0.21	0.039	0.040	0.056	0.20	0.13	0.17
LINEARITY: SLOPE ERROR	0.953	0.986	0.944	0.985	1.010	1.001	0.952	0.986	0.944
LINEARITY: R²	0.00773	0.0109	0.00839	0.0115	0.00886	0.00710	0.00762	0.00909	0.00785
	0.999	0.998	0.999	0.998	0.999	0.999	0.999	0.999	0.999

5.2.2 Discussion

The QA protocol implemented on both scanners had full acquisition conformance, indicating it met the QIBA protocol requirements. Using this protocol, both scanners were able to achieve technical conformance to the QIBA DWI Profile. Conformance to the QIBA profile indicates the scanner has the technical performance required to deliver the claims defined in the profile and can provide high-quality measurement of ADC values that allow ADC to be used as a biomarker. Further, these results showed that the QIBA tolerances can be met at room-temperature using the LC thermometer and NIST-traceable reference values provided by the phantom manufacturer.

5.3 BIAS AND LINEARITY OF ADC MEASUREMENTS

Although the QIBA profile focuses on a single water vial, it is advantageous to be able to characterise bias over a range of physiologically relevant ADC values. The CaliberMRI diffusion phantom allows the magnitude of the bias to be determined.

5.3.1 Results

Combining all results within a session for each vial, the %bias on the 3T scanner is shown in **Figure 25**, and **Figure 26** for the 1.5T scanner. The 40% and 50% PVP vials have higher %bias for both room temperature and ice bath scans. At 3T, three outliers are identified. Upon inspection in qCal, these were due to poor ROI placement (**Figure 27**).

QA for Assessing Quantitative Biomarkers in Patients using MRI

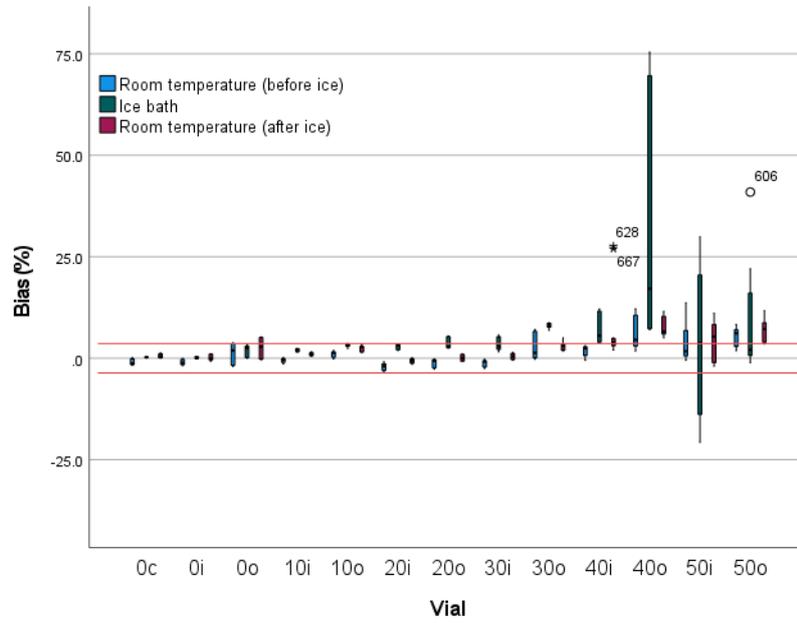


Figure 25. %Bias per vial on the 3T scanner with results from all orientations for the same vial pooled. c = central, i = inner, o = outer.

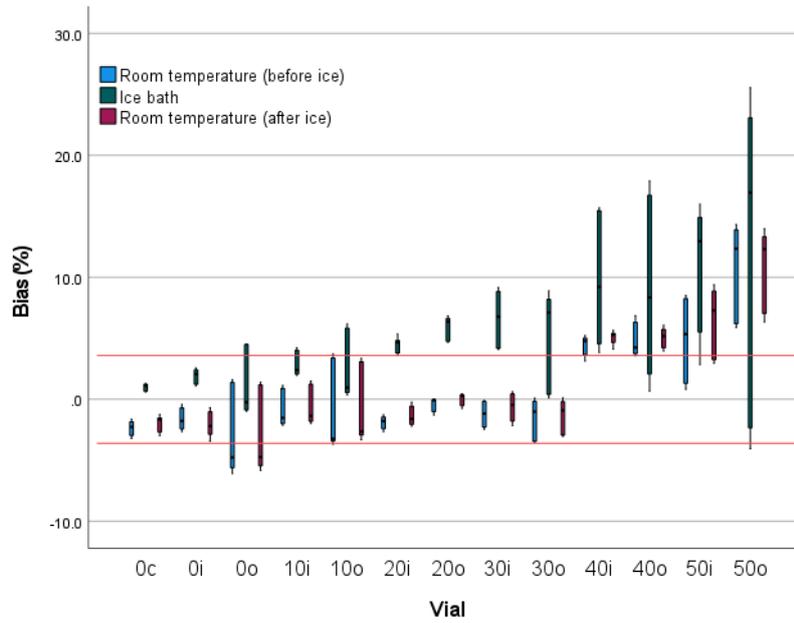


Figure 26. %Bias per vial at 1.5T with results from all orientations for the same vial pooled. c = central, i = inner, o = outer.

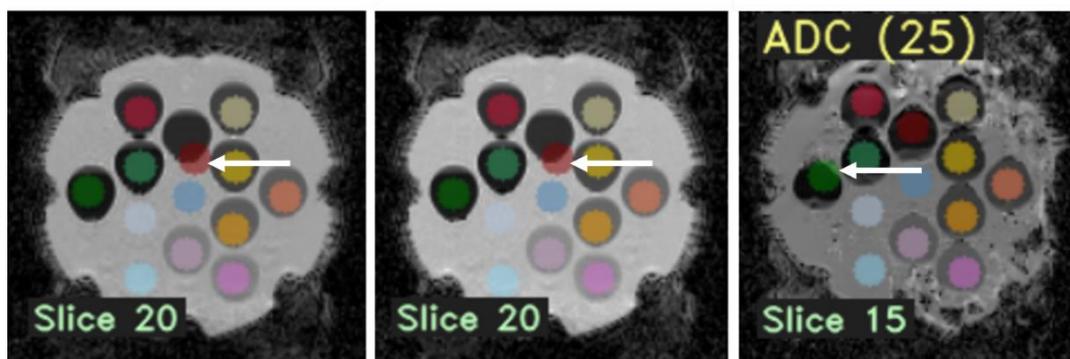


Figure 27. Poor ROI placement leads to high bias and standard deviation. Examples show poor placement for the inner 40% PVP vial at room temperature (left and centre) and the outer 50% PVP vial at 0°C (right).

Excluding the 50% PVP vial at 0°C (ADC 110 $\mu\text{m}^2/\text{s}$), all b_0 SNR is >50 . At b_{2000} , all SNR is >5 for all vials. This means SNR on both scanners is adequate to measure ADC using b -values up to b_{2000} without introducing additional bias due to low SNR (defined as <5 at high b -values in the QIBA DWI profile [10]). DWI SNR is higher at 3T than 1.5T across all vials and b -values (Appendix G)

Combining the ice and room temperature results at 3T (**Figure 28**), linearity over a range of 110 $\mu\text{m}^2/\text{s}$ to 1782 $\mu\text{m}^2/\text{s}$ is excellent (95% CI for the slope: (0.987, 0.995), $R^2 = 0.998$). Due to the higher room temperature at 1.5T, a total range of 110 $\mu\text{m}^2/\text{s}$ to 2037 $\mu\text{m}^2/\text{s}$ was achieved (**Figure 29**). The 95% CI for the slope of mean measured ADC against NIST ADC is 0.959 to 0.967 with $R^2 = 0.998$, which is also within the QIBA DWI Profile requirement (**Table 1**).

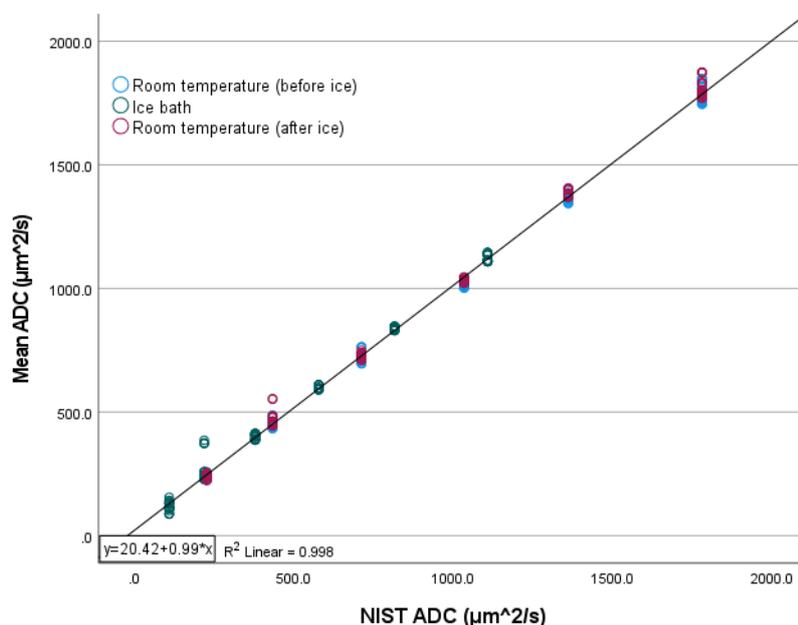


Figure 28. Linearity assessment on the 3T scanner. Mean ADC of the VOI is plotted against the NIST reference ADC value for the vial. Each data point is the mean ADC for a vial from an individual DWI scan. Each session has 4 repeats in each orientation.

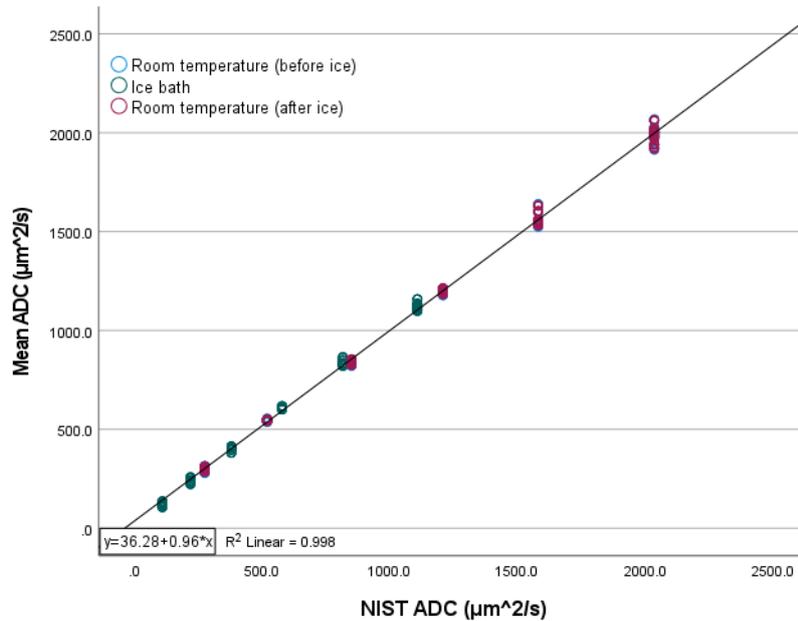


Figure 29. Linearity assessment on the 1.5T scanner. Mean ADC of the VOI is plotted against the NIST reference ADC value for the vial. Each data point is the mean ADC for a vial from an individual DWI scan. Each session has 4 repeats in each orientation.

5.3.2 Discussion

There was generally a statistically significant difference between the measured ADC value and the NIST ADC value, suggesting that the technical performance of the scanner introduces systematic bias into the ADC measurement, however, the bias is small. At 0°C, absolute bias ranges from -10.9 to +50.3 $\mu\text{m}^2/\text{s}$ on the 1.5T scanner, and from -22.7 to +45.0 $\mu\text{m}^2/\text{s}$ (excluding the outer PVP40 vial in sagittal orientation) on the 3T scanner. This is within the bias of $0.1 \times 10^{-3} \text{ mm}^2/\text{s}$ found for all vials of the QIBA DWI Phantom (0°C) by McDonald et al. [8]. Collecting results at both 0°C and room temperature allowed assessment of ADC values over a range nearly double that of performing the ice bath scan alone. This covers the range of ADC values seen in the human body well, increasing confidence in clinical ADC measurements.

For all sessions, % bias was greatest and outside of the QIBA tolerance for the lowest diffusivities (high PVP concentrations), and these vials were found to be the most affected by artefacts. Artefacts within ROIs and poor ROI placement were both found to contribute to high bias in ADC measurements. Air bubbles can become trapped between vials, creating susceptibility artefacts which may alter ADC values if the artefact occurs within a vial (**Figure 30**). It is advisable to check both the phantom prior to scanning and the initial T1-weighted image for bubbles. If any bubbles are present, the phantom should be topped up through its fine fill port using the pipette provided.

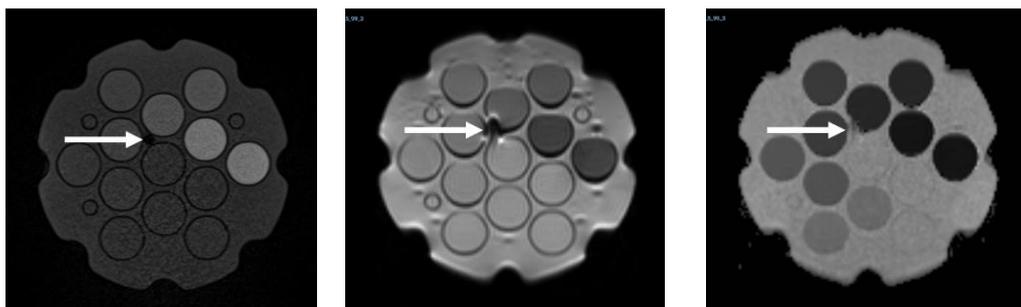


Figure 30. Example artefact created by an air bubble between vials in coronal orientation. T1 (left), DWI b_0 TRACE (centre) and ADC map (right).

The PVP vials themselves experience artefacts. As well as truncation artefacts caused by the vial walls, an enhanced signal is seen on one side of the vial and a signal void on the other. The artefact increases with increasing PVP concentration, contributing to the higher bias of the 40 and 50% PVP solutions, and is worse in coronal and sagittal orientation (**Figure 31**). This apparent shift of the signal from the PVP vials is only seen on the DWI sequence and occurs in the phase encoding (PE) direction. This is consistent with the EPI sequence used [9] as the continuous sampling in the PE direction to fill k-space within a single echo results in a reduced bandwidth in the PE direction, leading to increased susceptibility effects and geometric distortion [88].

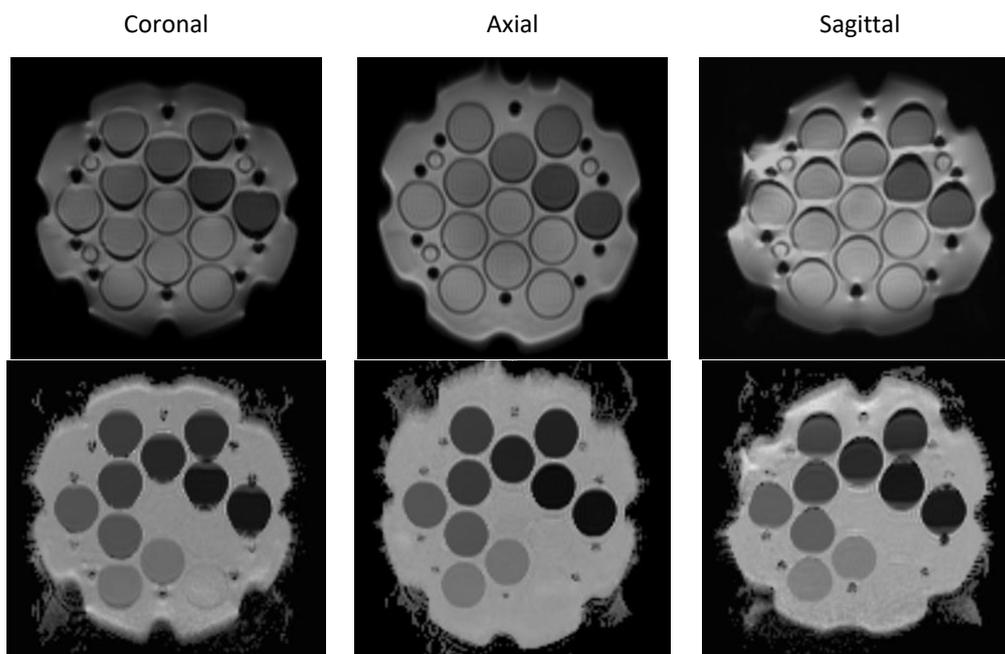


Figure 31. Truncation artefacts are seen within all vials and susceptibility-related artefacts increase with increasing PVP concentration, causing a signal void on one side of each vial, and a signal pile up on the other. Seen most significantly in sagittal and coronal orientations.

At 3T, artefacts resulted in saturation within some images in areas where susceptibility effects caused signal enhancement, both at the edge of PVP vials and around the LC vials (**Figure 32**). Saturation occurred mostly on the b_0 images. A few b_{500} slices were also affected, however, these slices were outside of the VOIs. Saturation was also reported by Carr et al.[72], who found saturation affected the SNR and b-value dependence, with minimal effects on bias, agreeing with the results in **Figure 32** that

show the saturation is unlikely to occur within the centre of the vials. However, if the VOI does contain an area of saturation, the measured ADC values can be significantly increased. No saturation was detected at 1.5T.

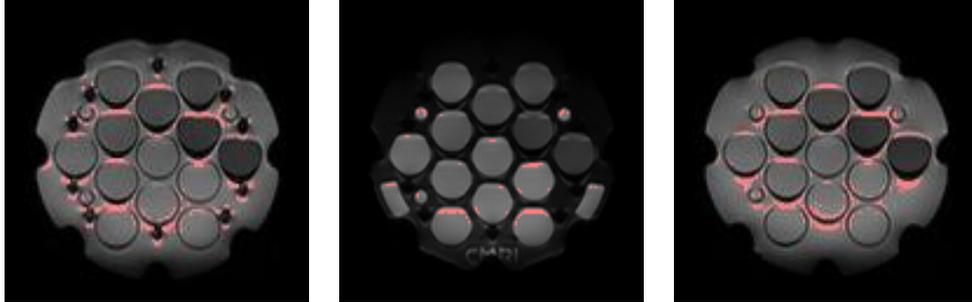


Figure 32. Examples of saturation occurring (coronal orientation). Saturated pixels are coloured in red. Credit to David Van Wie (Boulder Labs, CO, USA) for highlighting the saturated pixels.

A small variation in mean ADC was found between measurements in different orientations. However, this is expected as the phantom is manually repositioned between measurements, causing a small change in position of the centre of the phantom from isocentre with each orientation, introducing additional variation in the ADC measurement. There may also be a change in temperature seen during scanning, which would change the ADC values of the vials. Repeat scans with the phantom removed and then replaced in the same orientation, would allow clearer identification of any changes due to the choice of imaging plane. Averaging results over several imaging sessions, Carr et al. [72] found minimal imaging directional dependence on ADC performance.

One major limitation for room temperature measurements at 1.5T is a calculation warning given by qCal that “No calibration data for phantom type 128 with serial number 163 at field 1.5T. Using calibrations for 3T.”, meaning that the NIST reference values were obtained at 3T and have not been validated for 1.5T. Although it is widely agreed in the literature that ADC values are independent of field strength between 1.5T and 3T [61], a verified reference value is essential for reliable assessment of bias and therefore calibration data for 1.5T would be strongly desirable.

5.4 PRECISION OF ADC MEASUREMENTS

Accuracy alone is not enough to characterise biomarker measurement. Precision of the measurement must also be considered.

5.4.1 Results

For each scanner, per the QIBA profile, three types of precision are considered: random measurement error, short-term repeatability, and long-term repeatability. Precision results are discussed for the coronal orientation, as this is the orientation specified by CaliberMRI in their QA protocol. Results are presented for all vials in **Figure 33** and **Figure 34** for the 3T and 1.5T scanners, respectively. For both field strengths, there is good consistency between the random error at room temperature on different days whereas RC_{ST} and wCV_{ST} show more variability between the two room temperature sessions.

Excluding 50% PVP vials at 0°C, all 3T measurements over all vials are within the QIBA tolerance for random error. For the 1.5T scanner, random error at 0°C shows little variation from 0 to 30% PVP vials and is within the QIBA tolerance, however, it increases for the 40% vials and is significantly higher for the 50% vials. At room temperature, random error of water vials is much greater compared to at 0°C.

Across both scanners, all RC_{ST} values are within the QIBA tolerance for all vials and temperatures and are generally less than $5 \mu\text{m}^2/\text{s}$. wCV_{ST} of 0-30% PVP vials are also within the QIBA tolerance on both scanners. At 3T, vials with high wCV_{ST} are those identified as having poor ROI placement (**Figure 27**), demonstrating the importance of ROI placement. For most vials, the RC_{ST} and wCV_{ST} at room temperature are greater than at 0°C .

From **Figure 25** and **Figure 26** in Section 5.3, generally, there appears to be very good agreement between the two room temperature sessions, especially at 1.5T. For measurements on the 1.5T scanner, long-term repeatability values for all vials are within the QIBA tolerance with wCV_{LT} values from 0.02 to 1.19% and RC_{LT} 0.9 to $13.2 \mu\text{m}^2/\text{s}$. At 3T, excluding the inner PVP40 vial, both RC_{LT} and wCV_{LT} (ranges 3.3 to $59.9 \mu\text{m}^2/\text{s}$ and 0.47% to 1.39%, respectively) are within the QIBA tolerance over the full ADC. Variations in long-term repeatability across all vials are included in Appendix H.

QA for Assessing Quantitative Biomarkers in Patients using MRI

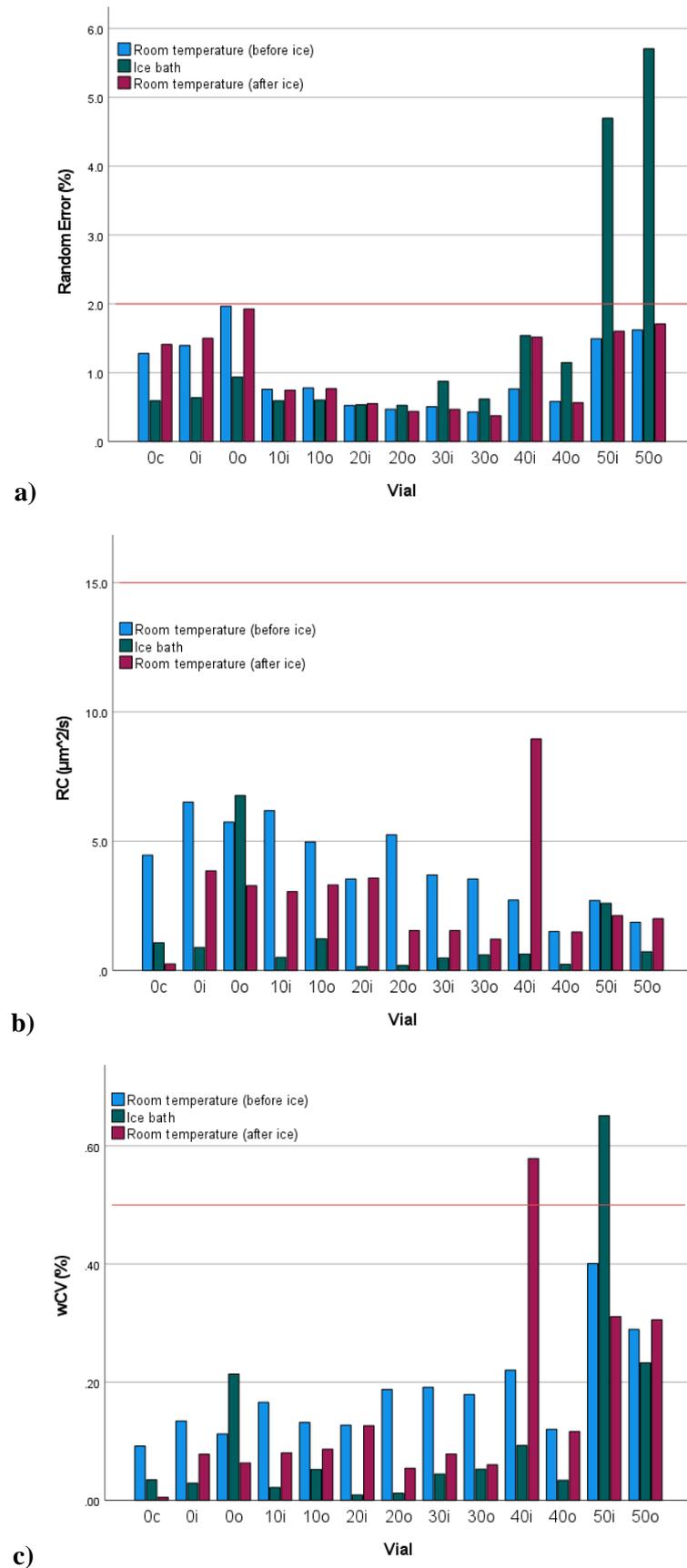


Figure 33. Precision of the 3T scanner, comparing room temperature and ice bath measurements with the phantom in coronal orientation. a) Random error b) RC_{ST} and c) wCV_{ST} . On all, the red line indicates QIBA tolerance, number indicates PVP percentage, c = centre, i = inner, o = outer.

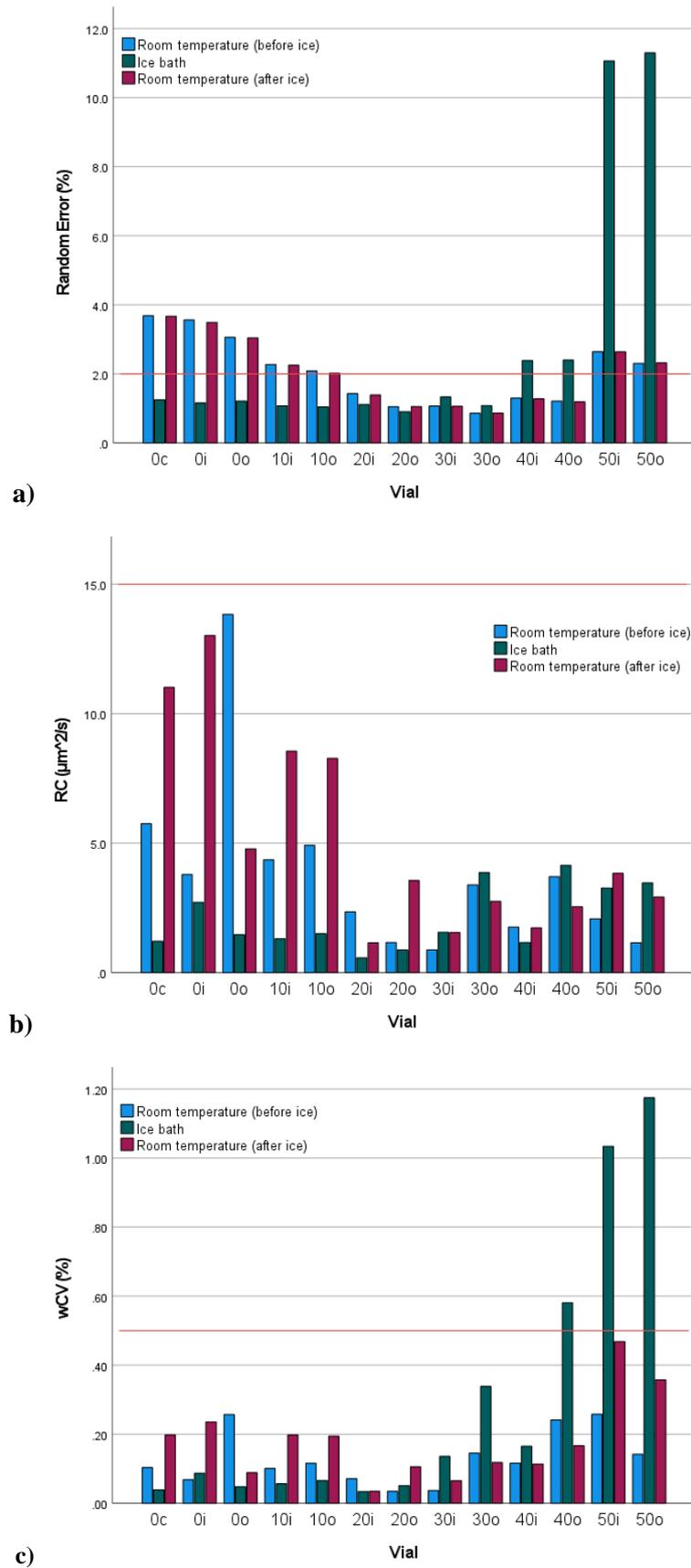


Figure 34. Precision of the 1.5T scanner, comparing room temperature and ice bath measurements with the phantom in coronal orientation. a) Random error b) RC_{ST} and c) wCV_{ST} . On all, the red line indicates QIBA tolerance, number indicates PVP percentage, c = centre, i = inner, o = outer.

5.4.2 Discussion

Repeatability within QIBA tolerances confirms variability due to technical performance of the scanner is small enough to allow a true biological change to be identified. Often, only the central water vial is assessed [10], assuming that precision is constant across all measured ADC values. These results find that assumption reasonable for high diffusivities, however, precision was worse for the 40 and 50% vials, although this is likely to be related to the PVP solutions due to the artefacts observed. Carr et al. [72] reported results across all vials at 0°C and found very similar results (**Figure 12**, Section 2.2.6).

Comparing ice bath and room temperature results, random error was noticeably greater for water at room temperature. This increase may be related to the lower viscosity of water with increased temperature. The ideal DWI phantom material has high viscosity to reduce the impact of bulk motion on ADC measurement [58]. Decreased viscosity increases sensitivity to vibration and convective motion, impacting precision. Additionally, temperature gradients are more likely to exist and not detected at room temperature, increasing convective motion.

For long-term repeatability metrics, ideally the phantom should be at the same temperature for all sessions. Ice water is therefore ideal as the temperature is well controlled, however, multiple ice-bath scans are logistically challenging. With a stable scanner room environment, phantom temperature should not fluctuate significantly between room temperature measurements, as seen here, however, the higher temperature measurement uncertainty may decrease repeatability between measurements. Despite this, results for both scanners at room temperature are consistent with those found in literature for 0°C, of $wCV_{LT} < 2.40\%$ for water to PVP30 vials [57, 72].

It is not always possible for the same patient to be scanned on the same scanner, and therefore reproducibility between different scanners can be important to characterise. A standardised QA protocol allows harmonisation between different scanners. Reproducibility between the two scanners was good. At 0°C, RDC between the two scanners across all vials in coronal orientation at 0°C is 0.47 to 89 $\mu\text{m}^2/\text{s}$ (Appendix I). It is expected that if two scanners conform to the requirements of the QIBA profile, reproducibility between the two scanners should be good, and this is supported by these values.

Within this chapter, it has been shown both scanners are able to conform to the technical performance required within the QIBA DWI profile. Further investigation into accuracy and precision using all vials has shown poorer quality measurement of vials with low ADC values, likely caused by increased artefacts on high PVP concentration vials. The results provide confidence in quantitative ADC measurements obtained from both scanners and demonstrate that evaluation of scanner performance is possible at room temperature. Following these findings, variability of ADC measurements using different receive coils is investigated by applying the QA protocol to three head coils at room temperature (Chapter 6).

6 THE EFFECT OF HEAD COIL CHOICE ON ADC VALUES

One hardware factor that can easily be varied on an MR scanner is the choice of receive coil. Ideally, the ADC value obtained should be independent of the choice of coil used. Not all sites will have the same coils available for all scanners or a single site may have multiple variations of coils available and, therefore, effect of coil choice should be characterised. There has been some comparison in the literature, for example, between a head coil and a torso coil [62], however, clinically these two coils would not be used to measure the same organs and, therefore, ADC values from these coils would not be directly compared. In healthy volunteers, intra-scanner variability was estimated to be up to 8% depending on the coil system used [41]. Coil design has been suggested to affect noise in ADC measurements [38].

The three receive head coils available within the scanner room for the 3T Siemens Prisma were used to scan the phantom at room temperature following the CaliberMRI QA protocol [74], with four immediate repeats in coronal orientation, one scan in axial and sagittal orientation, and a T1 scan before and after. Results from all vials were compared to tolerances stated in the QIBA DWI Profile (**Table 1**). Scans were performed within the same week to minimise variability.

6.1 RESULTS

6.1.1 Temperature and Qualitative results

Phantom temperature was $19.6 \pm 0.8^\circ\text{C}$ for all sessions, however, for the 32 channel coil repeats, the slice position changed within the immediate repeats. While the ADC bias could still be determined from this data, SNR, random error, RC_{ST} , and wCV_{ST} require unchanged slice position. Therefore, another 32-channel coil acquisition at $18.6 \pm 0.8^\circ\text{C}$ within the same week is used for comparison of these parameters.

Qualitatively, results look consistent across head coils. Artefacts seen are those mentioned in chapter 5 and are similar for each head coil across all series. Signal saturation occurred only in sagittal orientation with the 64-channel head coil.

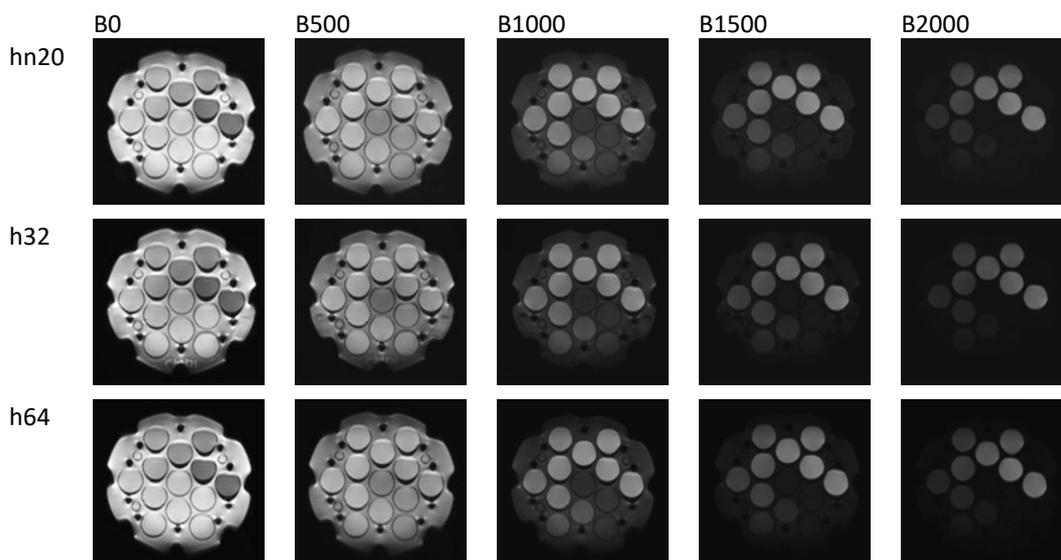


Figure 35. Trace DWI images at each b-value for the three head coils. All images show the same slice, one slice above the central vial plate, from the second coronal orientation repeat.

6.1.2 Accuracy

The QIBA technical performance metrics are presented for the central water vial for each head coil in **Table 5**. Applying the QIBA DWI tolerances, all values except random error for the 64-channel coil are within Profile tolerances. Results for all vials are included in Appendix J.

Table 5. Results for central water vial with each head coil, including linearity results derived from all 13 vials. Red value indicates it is outside of the QIBA tolerance.

PERFORMANCE METRIC	HEAD/NECK 20 (T = 19.6 +/- 0.8°C)	HEAD 32 (T = 18.6 +/- 0.8°C)	HEAD 64 (T = 19.6 +/- 0.8°C)
NIST ADC [$\mu\text{m}^2/\text{s}$]	1980	1929	1980
MEAN ADC [$\mu\text{m}^2/\text{s}$]	1969.0	1928.6 ^a	1952.0
BIAS [$\mu\text{m}^2/\text{s}$]	-10.60	-0.021	-27.59
BIAS [%]	-0.536	-0.00109 ^a	-1.39
MAX B-VALUE	0.466	0.352	0.495
DEPENDENCE [%] (B-VALUE PAIR)	(500, 2000)	(500,2000)	(500, 2000)
DWI SNR B0	167	143	174
DWI SNR B2000	10.9	13.4	9.19
RANDOM ERROR [%]	1.94	1.59	2.31
RC _{ST} [$\mu\text{m}^2/\text{s}$]	2.64	2.47	3.93
wCV _{ST} [%]	0.049	0.046	0.073
SLOPE (B ₁)	0.981	0.992	0.981
(SLOPE ERROR)	(0.008)	(0.008)	(0.011)
R ²	0.999	0.999	0.999

^aAt 19.6 +/- 0.8°C, the mean ADC value for the 32-channel coil was 1950.9, corresponding to -1.448% bias.

There is a statistically significant difference between mean ADC values of the central water vials at the same temperature from each head coil ($p < 0.01$). Tukey's method indicates a significant difference between the 20 and 64 channel and the 20 and 32 channel coils ($p < 0.01$) but not between the 32 and 64 channel coils ($p = 0.882$). However, since all measurements are within the temperature-related uncertainty ($1980 \pm 41.2 \mu\text{m}^2/\text{s}$), difference in the mean ADC could be due to variation in temperature alone. The 20-channel coil data was collected immediately following the 32-channel coil data, so a small temperature increase during scanning would be expected, however, the increase may be too small to be detected by the LC thermometer. The results support this as an increase in mean ADC value of $18 \mu\text{m}^2/\text{s}$ is seen between the two scans.

Looking at all vials, all 0 to 30% PVP vials were within +/- 3.6% bias across all head coils (**Figure 36**). All head coils demonstrated an increase in %bias outside of the QIBA tolerance for 40 and 50% PVP vials. Absolute bias across all VOIs of all vials ranged from -41.8 to $68.3 \mu\text{m}^2/\text{s}$, which remains in agreement with the +/- $100 \mu\text{m}^2/\text{s}$ observed in literature [8]. The choice of head coil does not affect the linearity of the ADC measurement. The slope and error are consistent between coils and with the results in chapter 5 that only used the 20-channel coil.

There is minimal variability in SNR from different head coils (**Figure 37**). b0 SNR was over 100 for all vials using all head coils, suggesting all head coils can provide adequate SNR to measure the ADC values without introducing additional bias.

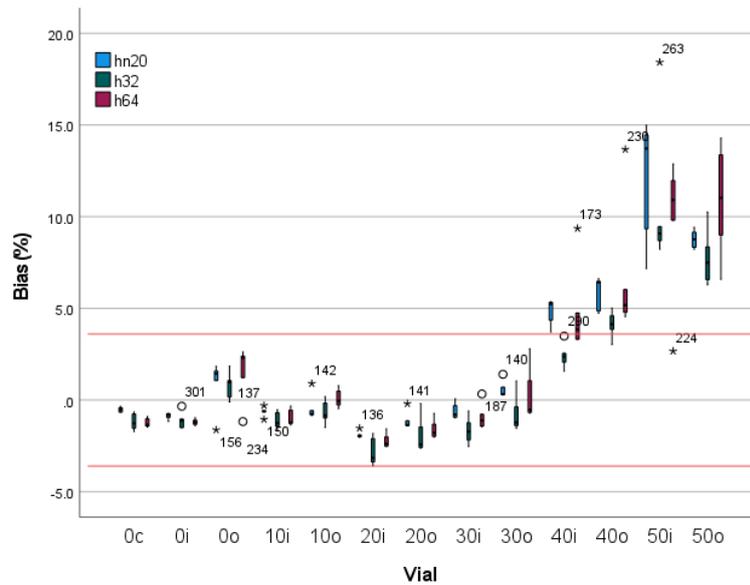


Figure 36. %Bias per vial for each head coil. Phantom temperature 19.6 +/- 0.8°C. The red line indicates QIBA tolerance, number indicates %PVP, c = centre, i = inner, o = outer.

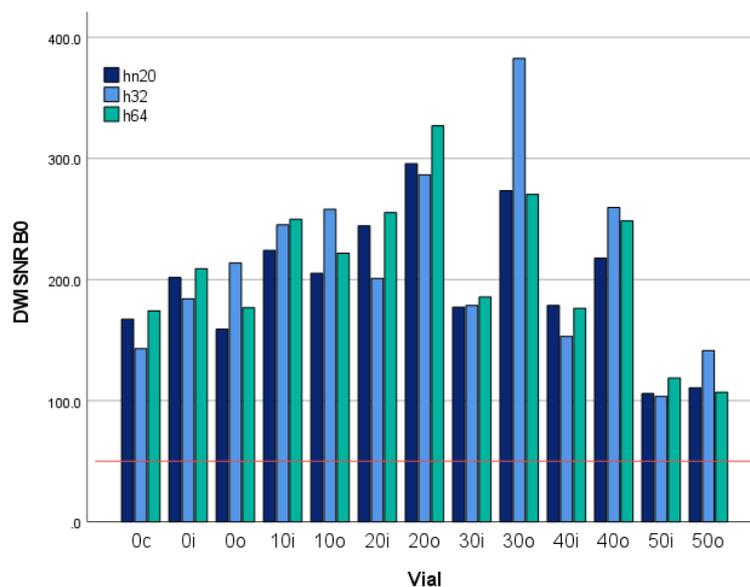


Figure 37. DWI b0 SNR per vial for each head coil. The red line indicates QIBA tolerance, number indicates %PVP, c = centre, i = inner, o = outer.

6.1.3 Precision

Random error is lowest using the 32-channel coil for all vials except the inner PVP50 vial, however, the majority of measurements from all coils remain within the QIBA tolerance (Figure 38a). For the 64-channel coil, all water vials have random error greater than 2%, consistent with findings in Section 5.4 that random error is higher for room temperature measurements of water than those at 0°C.

RC_{ST} is within QIBA tolerance for all vials except outer the water vial using the 20-channel coil (Figure 38b). Looking at Figure 38c, all wCV_{ST} values are within the QIBA tolerance except the outer PVP50

vial using the 64-channel coil. Generally, short-term repeatability is worst using the 64-channel vial, suggesting greater intra-session standard deviation.

Comparing mean ADC obtained at the same temperature, RDC between the three coils is within $40 \mu\text{m}^2/\text{s}$ for all vials (**Table 6**). Values are greater than the short-term repeatability for each head coil due to the repositioning of the phantom in each coil. Although the QIBA DWI profile does not include tolerances for reproducibility, the values for all vials are within the QIBA tolerance for test-retest RC_{LT} performed under the same conditions, suggesting that changing the head coil used between measurements would not violate the profile claims.

Table 6. RDC for all vials in coronal orientation and temperature $19.6 \pm 0.8^\circ\text{C}$.

VIAL	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
RDC ($\mu\text{m}^2/\text{s}$)	28.1	13.9	39.4	18.0	18.1	22.1	18.4	14.8	18.7	26.2	17.4	11.1	15.3

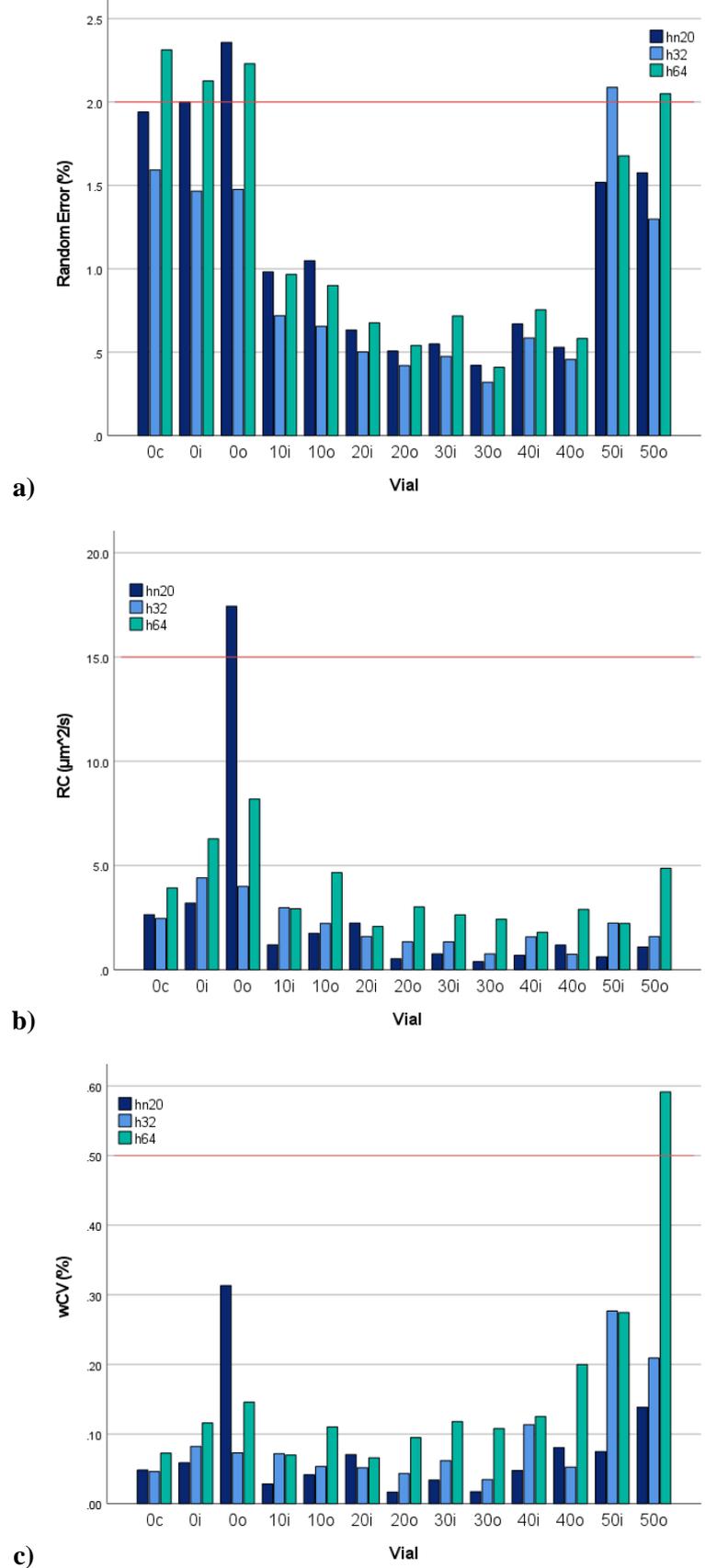


Figure 38. Precision results across all vials, comparing the three head coils. a) random error, b) RC_{ST} , c) wCV_{ST} . On all, the red line indicates QIBA tolerance, number indicates PVP percentage, c = centre, i = inner, o = outer.

6.2 DISCUSSION

Although specified for a central water vial at 0°C, the room temperature results from all vials were compared to the QIBA DWI tolerances as a starting baseline for accuracy and precision. For the central water vial, all results except random error were within the profile's specifications. It was found in Section 5.4 that random error is increased for water at room temperature compared to 0°C. There was a statistically significant difference between the mean ADC value measured at the same temperature, however, all results were within the temperature uncertainty of the reference ADC value and, therefore, can be considered consistent. There was no noticeable difference in image quality or artefacts between coils. Despite there being some trend in repeatability based on the coil used, the repeatability is excellent across all coils and the variation between coils is unlikely to impact clinical results.

These results showed less variation than the 8% intra-scanner variation dependent on the coil system stated in literature [41], however, those measurements were in healthy volunteers and included coils with and without sensitivity correction, introducing more variability than is seen with a phantom and between the three coils used here.

Clinically, these results indicate that results from different head coils can be compared without introducing significant variability to the measured ADC values. Accuracy is consistent across head coils and repeatability is very good for all coils, suggesting the three coils can be used interchangeably. It is important to perform ADC QA on each coil that will be used for quantitative DWI, to be able to assess the suitability of the coil and reliability of the measurements. By allowing room temperature ADC calibration, the CaliberMRI DWI phantom provides a quick process to establish a baseline performance for each coil.

7 ASSESSMENT OF CLINICAL RESEARCH PROTOCOLS

It is known that ADC measurements have a dependence on the DWI parameters used [89] and should be tailored to the organ being imaged. After gaining confidence ADC calibration under idealised conditions in the previous chapters, validation of clinical research protocols using the phantom was investigated. This allowed the robustness of ADC values to changes in sequence parameters, most notably changes to the b-values, to be assessed. This generates confidence in ADC values used in clinical trials and research where changes in ADC are used as an endpoint for monitoring, diagnosis, or prediction.

A T1-weighted scan was performed before and after 4 repeats of the research DWI sequence with the phantom in the appropriate orientation for the sequence used. If the DWI sequence met the minimal requirements of qCal (specified in Appendix D), the data can be analysed within qCal, however, warnings for acquisition conformance are given. Two different clinical research protocols that met the minimum qCal requirements were investigated:

- The BRITER protocol is a DWI sequence for the brain, performed on the 3T scanner. The BRITER DWI sequence is a 3-scan trace with 3 b-values: 0, 500, and 1000 s/mm², with 2 averages for the lowest two b-values and 3 averages for b1000. TR and TE are 6100 ms and 69 ms, respectively. An axial orientation is used, and scan time is 124 s.
- The AIRPANC protocol has been developed for an upcoming study on the 1.5T scanner to measure ADC values in the pancreas. The 3-scan trace sequence uses 4 b-values: 0, 150, 400, and 800 s/mm² with 1, 2, 3, and 3 averages respectively. TR is 4000 ms and TE 49 ms and scan time is 122 s.

7.1 RESULTS

7.1.1 BRITER Protocol

At an average phantom temperature of 16.15°C, absolute bias ranges from -7.22 to 34 μm²/s across all vials, covering an ADC range of 232 to 1807 μm²/s. %bias is less than 3.5% over an ADC range 725 – 1807 μm²/s. At lower ADC values, %bias increases to up to 15% for 50% PVP; however, linearity remains excellent (slope 0.991, R² = 0.9995). b0 SNR >143 and b1000 SNR >97 for all vials is high enough to measure ADC values without introducing additional bias. Excluding 50% PVP, all precision metrics are within the QIBA tolerance, with random error <1%, RC_{ST} <4.4 μm²/s and wCV_{ST} <0.33%, suggesting very good repeatability for immediate repeats of the sequence. Performance metrics for all vials are presented in **Table 7**.

Comparing the protocol to the QIBA recommended protocol for brain, this sequence meets at least the ‘acceptable’ criteria for each parameter specified. Therefore, as both the protocol and the technical performance of the scanner conform with the profile, the claim that “A measured change in the ADC of a brain lesion of 11% or larger indicates that a true change has occurred with 95% confidence” can be applied.

Table 7. Performance metrics for all vials of the CaliberMRI diffusion phantom imaged using the BRITER clinical research protocol on the 3T scanner (average temperature 16.15°C). Red values are outside of the QIBA tolerance.

VIAL	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value ($\mu\text{m}^2/\text{s}$)	1807	1807	1807	1383	1383	1052	1052	725	725	442	442	232	232
Mean ADC ($\mu\text{m}^2/\text{s}$)	1815.9	1837.2	1821.3	1400.0	1395.7	1044.8	1056.8	742.7	749.8	475.0	471.3	265.5	259.8
Bias ($\mu\text{m}^2/\text{s}$)	9.30	30.61	14.70	17.30	13.05	-7.22	4.82	17.96	25.01	32.60	28.94	33.40	27.68
Bias (%)	0.518	1.70	0.817	1.26	0.947	-0.683	0.461	2.48	3.46	7.38	6.55	14.40	11.93
Max b-value dependence	0.245 (500, 1000)	0.173 (500, 1000)	0.289 (500, 1000)	0.004 (500, 1000)	0.226 (500, 1000)	0.049 (500, 1000)	0.961 (500, 1000)	0.086 (500, 1000)	1.24 (500, 1000)	1.64 (500, 1000)	0.718 (500, 1000)	7.85 (500, 1000)	0.779 (500, 1000)
SNR b0	116.9	157.6	313.6	159.2	360.9	152.1	288.0	151.5	259.3	149.6	241.3	143.7	214.1
SNR b500	112.3	179.8	256.1	169.2	336.6	155.8	294.8	171.6	295.5	177.9	293.8	161.0	327.0
SNR b1000	97.4	131.0	151.0	169.7	232.1	152.3	232.0	169.5	300.6	198.7	315.5	181.9	368.4
Random error (%)	0.454	0.451	0.381	0.412	0.342	0.466	0.451	0.606	0.545	0.925	0.918	1.91	1.84
RC_{ST} ($\mu\text{m}^2/\text{s}$)	3.22	9.69	1.83	5.67	1.90	1.75	1.56	2.50	4.00	4.38	4.25	1.89	4.32
wCV_{ST} (%)	0.0641	0.190	0.0362	0.146	0.0492	0.0604	0.0533	0.121	0.193	0.333	0.326	0.257	0.600

7.1.2 AIRPANC Protocol

Results for all vials are presented in **Table 8**. Over an ADC range of 276 to 2037 $\mu\text{m}^2/\text{s}$ (phantom temperature 20.7 +/- 0.9°C), absolute bias ranged from -30.6 to 69.1 $\mu\text{m}^2/\text{s}$. The %bias of the 0 to 30% PVP vials is within 3.4% and therefore within the QIBA tolerance. Consistent with the QA protocol results, 40% and 50% PVP vials have higher bias, up to 17.3%. b0 SNR >135 and b800 SNR >82 at b800 indicates adequate SNR. Excluding the 50% PVP vials and the RC_{ST} of the outer water vial, random error <1.5%, RC_{ST} <14 $\mu\text{m}^2/\text{s}$ and wCV_{ST} <0.48% are all within the QIBA tolerances for precision. In agreement with the QA protocol results, linearity is excellent (slope 1.000, R² = 0.998). Max b-value dependence is >2% for all vials, however, use of b150 violates the QIBA requirement for b-value dependence (Section 3.6), which states $b_1 - b_{\min} > 400 \text{ s/mm}^2$ for adequate diffusion contrast. qCal takes $b_{\min}=b_0$ giving a difference of only 150 s/mm^2 . However, low b-values are commonly used in the abdomen to remove perfusion effects or apply a bi-exponential ADC model[3], which is not possible within qCal.

In a review of quantitative DWI pancreas studies, from 18 studies with 792 patients, Barral et al. [29] found mean ADC of the whole pancreas to be $1.611 \times 10^{-3} \text{ mm}^2/\text{s}$. At 20.7°C the vial closest to this is the 10% PVP vial. The inner and outer PVP10 vials have bias 0.93% and 1.77% respectively, indicating that the sequence can measure ADC values at a clinically relevant ADC value with minimal error. Differences between the phantom and human anatomy are important to note. The pancreas is located deep within the abdomen and therefore will be further from coil elements than the phantom vials, decreasing SNR. Scans of healthy volunteers alongside phantom studies are important to verify performance.

One limitation of these results is the use of the head coil, which is not the coil that would be used clinically to image the pancreas. However, it was chosen here for consistency with other results in this report, to isolate variability in ADC measurement due to the sequence only, rather than introducing additional sources of variation.

Table 8. Performance metrics for all vials of the CaliberMRI diffusion phantom imaged using the AIRPANC clinical research protocol on the 1.5T scanner (temperature 20.7 +/- 0.9°C). Red values are outside of the QIBA tolerance.

VIAL	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value ($\mu\text{m}^2/\text{s}$)	2037	2037	2037	1582	1582	1209	1209	851	851	520	520	276	276
Mean ADC ($\mu\text{m}^2/\text{s}$)	2039.8	2064.7	2106.2	1596.8	1610.1	1178.6	1192.8	832.1	840.8	546.1	542.4	323.8	318.6
Bias ($\mu\text{m}^2/\text{s}$)	2.70	27.6	69.1	14.7	28.0	-30.6	-16.4	-18.8	-10.1	25.8	22.1	47.7	42.5
Bias (%)	0.132	1.35	3.39	0.929	1.77	-2.53	-1.36	-2.21	-1.18	4.95	4.26	17.3	15.4
Max b-value dependence	5.14 (150, 400)	5.07 (150, 800)	4.60 (150, 800)	9.71 (150, 800)	9.16 (150, 800)	11.0 (150, 800)	14.4 (150, 800)	9.19 (150, 800)	10.7 (150, 800)	4.58 (150, 400)	7.18 (150, 400)	9.82 (150, 800)	9.39 (150, 800)
SNR b0	135.4	145.1	163.9	162.9	177.5	162.2	187.7	152.6	157.5	143.9	149.4	114.9	138.0
SNR b150	238.7	223.0	209.8	276.4	324.1	313.6	337.7	295.1	321.9	288.6	303.7	231.8	222.5
SNR B400	178.0	165.1	143.8	222.2	252.0	274.0	327.0	300.0	303.3	292.4	288.1	248.5	216.9
SNR b800	83.8	82.7	87.3	129.3	147.2	184.5	206.9	228.5	231.3	256.3	243.0	232.5	212.9
Random error (%)	0.824	0.811	0.710	0.725	0.632	0.743	0.643	0.946	0.929	1.45	1.44	2.97	2.73
RC_{ST} ($\mu\text{m}^2/\text{s}$)	9.71	13.78	17.91	8.02	4.14	2.96	5.11	1.10	5.15	4.47	7.18	5.94	8.99
wCV_{ST} (%)	0.172	0.241	0.307	0.181	0.0928	0.0907	0.155	0.0476	0.221	0.295	0.478	0.663	1.03

7.2 DISCUSSION

Consistent with the QA protocol results, both studies had poorest repeatability of the lowest diffusivity vials. This, however, may indicate the choice of b-values are not high enough to characterise the ADC values of these vials [8]. ADC values of the 40 and 50% PVP vials are below those commonly seen clinically in both brain and pancreas. At clinically relevant ADC values, both sequences demonstrate good bias and repeatability, in accordance with tolerances in the QIBA DWI profile.

Not all protocols can be evaluated on the phantom without modification. The FOV must cover all vials and ideally cover the full phantom. The AIRPANC protocol did not consist of enough slices to cover the full phantom, however, by ensuring the central slice was aligned with the central vials plate, qCal was able to place VOIs within each vial.

Room temperature measurements allow a larger range of ADC values to be measured. This means protocols can be tested on the phantom over the range of values expected clinically to check the choice of b-values, SNR, and bias is appropriate, giving confidence in the ADC values obtained. For example, for the pancreas, if an ice bath scan was performed, the highest ADC value would be lower than the mean ADC expected in the pancreas. In comparison, at room temperature, this value was achieved with 10% PVP. These results demonstrate that the CaliberMRI diffusion phantom can be used to evaluate the accuracy and precision of clinical research protocols at physiologically relevant ADC values.

8 OVERALL DISCUSSION

Within this chapter, discussion focuses on the aims presented in Section 1.6.

8.1 EVALUATION OF CALIBERMRI DIFFUSION PHANTOM AND QCAL SOFTWARE

The first aim of this project was to evaluate a new quantitative DWI phantom and its associated analysis software. The CaliberMRI diffusion phantom offers a robust and safe phantom for measurement of a physiologically relevant range of ADC values. The phantom was a suitable size for all three head coils available on the 3T scanner and both 0°C and room temperature ADC calibration was possible.

The phantom contains 10 LC thermometer vials to measure temperatures over approximately 1°C intervals. However, the uncertainty in the temperature measurement is relatively large, up to +/- 1.2°C at 21.8°C. Generally, the uncertainty in the reference ADC values due to the uncertainty in the temperature was greater than the variation between measurements. If applying the QIBA tolerance of +/- 3.6% bias to all vials, at low ADC values, the temperature uncertainty becomes close to the tolerance level.

Despite identification of room temperature measurements as a promising option for ADC calibration [25, 70, 72], the only ADC values for all the PVP solutions above 0°C are those provided in the phantom manual. This means there is an absence of verification, and performance metrics must be compared to those obtained at 0°C rather than a direct comparison. This has additional consequences at 1.5T, as CaliberMRI has not confirmed traceability to the reference values at 1.5T. Current values are measured at 3T and, therefore, the software gives a calculation warning for 1.5T.

qCal-MR enables fast and convenient analysis of the CaliberMRI DWI phantom, largely a result of the automatic ROI placement within the PVP and LC thermometer vials. The cloud platform means the data can be accessed anywhere and easily shared between users. Automatic ROI placement is not always optimal, occasionally including artefacts and surrounding water, increasing the reported ADC value. However, large VOIs using several slices act to minimise the effects of heterogeneity in the vials. A current disadvantage of qCal is the inability to perform calculations across two different studies for easy long-term repeatability and reproducibility assessment.

Currently, the software requires slices to cover the entire phantom to correctly place VOIs. Per feedback to the CaliberMRI team, they confirmed future intentions to use the central vial plate to determine vial locations, which would enable the number of slices to be reduced, improving scan time. Additionally, the phantom must be in correct orientation or vials are mislabelled as the incorrect PVP concentration, giving incorrect bias results. Placing the positioning stickers onto the phantom (as shown in Section 3.2) helped to avoid this occurring.

8.2 ADC CALIBRATION AT ROOM TEMPERATURE

To explore the second aim of this project, ADC calibration at room temperature, room temperature scans were performed either side of an ice bath scan. An ice bath allows accurate and repeatable temperature control and minimises temperature variations during scanning. However, it requires significant preparation and decreasing the temperature decreases the range of ADC values available. In contrast, room temperature measurements are more convenient, requiring minimal preparation, and allow bias, linearity, and repeatability to be assessed over a much larger range of ADC values. However, uncertainty in the temperature measurement is greater and temperature variations during scanning are more likely.

Room-temperature scans demonstrated good accuracy and precision in all imaging directions with central water vial bias $<1.64\%$, and random error $<1.57\%$, $wCV_{ST} <0.22\%$, $wCV_{LT} <1.26\%$ for the 3T Siemens Prisma and $<3.08\%$, $<4.02\%$, $<0.21\%$ and $<0.51\%$, respectively, for the 1.5T Siemens Sola. These values are within the QIBA tolerances defined for a water vial at 0°C , supporting that room temperature ADC calibration is feasible.

Worst performance was seen for vials with lowest diffusivity, in agreement with results in the literature [65, 72]. Analysis of the images shows higher concentration PVP vials experience greater artefacts, including distortion and saturation, which can have a significant impact on ADC bias and repeatability. Signal saturation on the 3T scanner should be investigated further, to see if the QA protocol can be optimised to reduce its occurrence.

In this report, the QIBA tolerances for a 0°C water vial at the scanner isocentre were used for room temperature measurements and applied to all vials in the phantom, and sequences other than the QA protocol. It has been found that although not directly applicable, they are generally achievable and therefore serve as appropriate starting points for evaluation of quantitative DWI on a scanner. Baseline values per scanner and actionable limits can be developed from these values through further repeat measurements to establish expected longer-term variability.

Proving ADC calibration at room temperature is possible, sources of variability in ADC measurements were investigated. Using the established QA protocol, impact of head coil choice on ADC measurement performance was assessed. Accuracy and precision results were comparable across coils and the RDC for the three coils is $<40\ \mu\text{m}^2/\text{s}$ for all vials, suggesting choice of head coil on the 3T scanner does not introduce large additional variability to ADC measurements. Establishing that high-quality ADC measurements were possible using the QA protocol, two clinical research protocols were used to obtain ADC measurements from the phantom. Results demonstrated that ADC measurements on both scanners are robust to variations in sequence parameters. Both sequences demonstrate good bias and repeatability, giving confidence in the ADC values obtained from them.

There are many sources of variation not considered within this report; only the technical performance requirements for the acquisition device were considered, however, the QIBA DWI profile states requirements for other factors such as image distribution and patient-specific factors (positioning, ROI determination, spatial registration) [10] that were not evaluated here. Only two scanners were assessed, and these were both from the same vendor. Furthermore, only head coils were assessed.

Use of a phantom provides repeatable conditions in a controlled environment. Therefore, repeatability performance results found using a phantom will be better than the repeatability within a human. This is considered within QIBA claim statements, however, characterisation of repeatability in healthy volunteers would provide further confidence in the full imaging chain.

There are some limitations within the data. Measurements to compare the head coils could not all be performed on the same day due to limited time on the scanner, meaning environment, software, or hardware changes between measurements could introduce additional sources of variation. Similarly, it is not possible to perform ice bath and room temperature measurements immediately following each other. However, a room temperature scan was performed both before and after the ice bath to be able to detect any significant changes during this time, and none were observed.

Although performing only one ice bath per scanner is ideal from a logistical perspective, no test-retest or longer-term system stability assessment is performed that is directly comparable to published results using an ice bath. Multiple test-retest repeats would be preferable to be able to confidently establish a baseline for the scanner performance.

8.3 DISCUSSION OF THE QIBA CONFORMANCE PROCESS

The project aim ‘to assess the technical performance of two MRI scanners for conformance to the QIBA DWI Profile to confirm that high-quality ADC measurements can be delivered by NHS GGC’, was met by performing an ice bath scan on two scanners. The measurements performed on both scanners showed excellent conformance with the QIBA DWI profile and are comparable to results in the literature. Ultimately, conformance to the QIBA profile confirms that the scanners have the technical ability to reliably measure ADC. This promotes the use of ADC as a biomarker in studies and trials, consequently developing research and promoting precision medicine.

The publishing of the QIBA DWI profile has significantly standardised how ADC bias, repeatability, and reproducibility have been reported in the literature. However, some discrepancies still exist between the choice of reporting the mean or the median ADC of an ROI, and in the definitions of repeatability and reproducibility. One major challenge in harmonisation of biomarkers is making groups such as QIBA more widely known. In a 2021 survey, only 1.5% and 11.0% of responding radiologists were aware of QIBA and EIBALL, respectively [18].

Because QIBA profiles rely on available literature, limitations in the QIBA profile are generally due to limitations in the literature currently available for quantitative DWI. For example, the QIBA DWI profile only covers four organ systems and does not consider ultra-high (>3T) or low (<1.5T) field strengths due to insufficient literature. The performance tolerances are all specified for water at 0°C so only consider a single ADC value, and room temperature measurements cannot be used for conformance certification. Even if two scanners conform to the QIBA DWI profile, the claims only hold when the subject is scanned on the same scanner [10]. Although the profile requires the ADC calibration at 0°C, it is hoped that QIBA may update the profile for room temperature calibration once more data and confidence in the traceability of room temperature reference values are available.

The QIBA DWI Profile does not state a frequency at which periodic QA should be performed due to insufficient data, only specifying that QA procedures should follow any hardware or software upgrades. In a prostate reproducibility trial described by Boss et al. [90], a repeat phantom scan was required for any new scanner introduced to the study, after any major changes to scanner hardware or software, or if it had been one year since initial scanner qualification to the study. Carr et al. [72] also recommended annual ADC QA outside of QA before and after commencing trials and surrounding major scanner upgrades. In contrast, Paterson et al. [36] specified their phantom will be scanned monthly to verify accuracy of ADC measurement.

Although CaliberMRI provides QIBA conformance certification using their phantom and qCal software, it is still in the initial stages of being set up and as such required an appreciable degree of iteration of experimental work and discussion with CaliberMRI and the qCal software team. At the time of writing, we have received verbal confirmation that both scanners have achieved conformance certification, however, the process of formally administering this status is still in development. To reduce the overhead in getting started, more detailed protocols in scanner vendor format would be ideal, particularly for the T1-weighted scan for the thermometer. Consequently, the protocols established here have been placed on the departmental website [78] with the intention that others may benefit from this work.

9 CONCLUSIONS

Calibration of ADC values underpins confidence in research and clinical trials that use quantitative DWI, a technique that has promising applications in precision medicine and adaptive treatment. The technical performance of the scanner, characterised by bias, linearity, and precision, determines the reliability of the ADC measurements and their use in diagnosis, monitoring disease and predicting outcomes.

Two scanners within NHSGGC on which quantitative DWI was considered likely were assessed for conformance to the QIBA DWI Profile and both scanners were able to achieve conformance. This indicates that the variation in ADC measurements caused by the technical performance of the scanner can be considered negligible compared to observed biological variation.

Room temperature calibration using a traceable phantom and standardised QA protocol is possible, with results comparable to those obtained at 0°C while reducing phantom preparations and allowing assessment of scanner performance at a wider range of ADC values that better cover the range seen physiologically. The LC MC-visible thermometer offers a novel method of non-invasive temperature measurement for MR phantoms. Optimised protocols for viewing the thermometer and obtaining QA DWI scans that meet the requirements of the qCal analysis software have been published on the department website.

At both room temperature and 0°C, vials with the highest PVP concentration, corresponding to the lowest diffusivity, had inferior performance metrics and a higher occurrence of artefacts. However, performance remained consistent with findings in literature. Variations between three different head coils on the same scanner were found to be small, indicating all are suitable for quantitative DWI. Furthermore, the CaliberMRI diffusion phantom can be used to evaluate clinical research DWI sequences for ADC bias, linearity, SNR, and repeatability to assess the suitability of the sequence for quantitative DWI.

9.1 SUGGESTIONS FOR FUTURE WORK

Following this work, ongoing QA would provide information on longer-term repeatability to identify any scanner drift and assess the impact of upgrades, increasing confidence in ADC measurements. Additionally, healthy volunteer scans should be performed prior to any trials to fully characterise the repeatability of in vivo measurements of the anatomy of interest.

qCal uses the trace DWI images and calculates the ADC values internally for standardisation of the ADC calculation independent of the scanner model and its software. ADC maps can also be automatically generated by the scanner, and these are the ADC values that would be used clinically, however, they are not verified within the QA protocol outlined in this work. As a part of future QA, ADC values derived from the ADC map created by each scanner should be compared to the ADC values obtained using qCal and any poor agreement should be investigated.

ADC calibration was investigated at 7T, however, significant artefacts occurred and development of a 7T QA protocol would be required. An additional barrier to 7T calibration is the phantom size. The only coil the phantom fit in had limited field of view that did not cover the full phantom in all orientations. Similar to 1.5T, NIST ADC values are not traceable to 7T. However, it would be of interest to develop a 7T ADC QA protocol to be able to assess the possibility of quantitative MRI at 7T. This may include investigation of other readout sequences, as artefacts associated with EPI sequences are enhanced with increasing magnetic field strength.

9.2 ADDENDUM

On 04-Aug-2022, after finalisation of data within this report, we were notified by CaliberMRI of an error within qCal. Temperatures for the LC thermometer were incorrect by up to 0.3°C. This affects the room temperature NIST reference values, and, therefore, the bias and linearity reported. B-value dependence, SNR, and all precision metrics are unaffected.

Key temperatures used in Chapters 5 – 7 and the corresponding corrected temperatures are stated in **Table 9**, along with their associate NIST reference value for water and 50% PVP. All corrected temperatures are within the previous temperature uncertainty for each LC vial, and therefore, this update only has a small impact on the results. However, as bias and linearity are key performance metrics, this change is important to acknowledge.

Bias measured at 0°C is unaffected as the LC thermometer is not used, and, therefore, no doubt is cast on the performance of either scanner. Additionally, all head coil bias data used for analysis is unaffected as the temperature for these sessions (19.6 +/- 0.8°C) is unchanged.

Table 9. Temperatures used within Chapters 5-7 and corresponding corrected temperatures, provided by CaliberMRI.

Number of bright vials	Temperature used in this report (°C)	NIST reference value for water ($\mu\text{m}^2/\text{s}$)	NIST reference value for 50% PVP ($\mu\text{m}^2/\text{s}$)	Corrected temperature (°C)	NIST reference value for water ($\mu\text{m}^2/\text{s}$)	NIST reference value for 50% PVP ($\mu\text{m}^2/\text{s}$)
1	15.6 +/- 0.9	1782 +/- 40.4	228 +/- 6.79	15.3 +/- 0.8	1769 +/- 34.8	226 +/- 6.05
5	19.6 +/- 0.8	1980 +/- 41.2	263 +/- 8.90	19.6 +/- 0.8	1980 +/- 41.2	263 +/- 8.90
6	20.7 +/- 0.9	2037 +/- 47.5	276 +/- 11.4	20.5 +/- 1.1	2026 +/- 57.7	274 +/- 13.6

REFERENCES

- [1] A. Venkataraman and J. Zhong, "Basics of Diffusion Measurements by MRI," in *Diffusion-Weighted MR Imaging of the Brain, Head and Neck, and Spine*, T. Moritani and A. A. Capizzano Eds., 3 ed.: Springer International Publishing, 2021, ch. 1, pp. 3-10.
- [2] L. R. Arlinghaus and T. E. Yankeelov, "Diffusion-Weighted MRI," in *Quantitative MRI in Cancer*, T. E. Yankeelov, D. R. Pickens, and R. R. Price Eds. Boca Raton, FL: CRC Press, 2012, ch. 7, pp. 81-97.
- [3] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince, *MRI from Picture to Proton*, 3 ed. Cambridge: Cambridge University Press, 2017.
- [4] C. D. Kroenke, "Physical and Physiological Principles of Diffusion," vol. 1, N. Seiberlich *et al.* Eds.: *Advances in Magnetic Resonance Technology and Applications*, 2020, ch. 18, pp. 457-476.
- [5] E. O. Stejskal and J. E. Tanner, "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient," *The Journal of Chemical Physics*, vol. 42, no. 1, pp. 288-292, 1965, doi: 10.1063/1.1695690.
- [6] T. Moritani, A. Hiwatashi, and J. Zhong, "Pitfalls and Artifacts of Diffusion-Weighted Imaging," in *Diffusion-Weighted MR Imaging of the Brain, Head and Neck, and Spine*, T. Moritani and A. A. Capizzano Eds., 3 ed.: Springer International Publishing, 2021, ch. 3, pp. 29-41.
- [7] N. Tyagi, M. Cloutier, K. Zakian, J. O. Deasy, M. Hunt, and A. Rimner, "Diffusion-weighted MRI of the lung at 3T evaluated using echo-planar-based and single-shot turbo spin-echo-based acquisition techniques for radiotherapy applications," *Journal of Applied Clinical Medical Physics*, vol. 20, no. 1, pp. 284-292, 2019, doi: 10.1002/acm2.12493.
- [8] B. A. McDonald *et al.*, "In Vivo and Phantom Repeatability of Diffusion-Weighted MRI Sequences on 1.5T MRI-Linear Accelerator (MR-Linac) and MR Simulator Devices for Head and Neck Cancers: Results from a Prospective R-IDEAL Stage 2a Evaluation of Tumor and Normal Tissue Apparent Diffusion Coefficients as Quantitative Imaging Biomarkers," *medRxiv*, Preprint p. 2022.05.28.22275724, 2022, doi: 10.1101/2022.05.28.22275724.
- [9] G. Yang and J. A. McNab, "Acquisition of Diffusion MRI Data," in *Advances in Magnetic Resonance Technology and Applications*, vol. 1, N. Seiberlich *et al.* Eds., 1 ed.: Academic Press, 2020, ch. 19, pp. 477-507.
- [10] Quantitative Imaging Biomarkers Alliance, "QIBA Profile: Diffusion-Weighted Magnetic Resonance Imaging (DWI)," 2019. Accessed: July 28, 2022. [Online]. Available: https://qibawiki.rsna.org/images/6/63/QIBA_DWIPProfile_Consensus_Dec2019_Final.pdf
- [11] D.M. Koh and D. J. Collins, "Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology," *AJR Am J Roentgenol*, vol. 188, no. 6, pp. 1622-1635, 2007, doi: 10.2214/AJR.06.1403.
- [12] Padhani, "Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations," *Neoplasia*, vol. 11, no. 2, pp. 102-125, Feb 2009, doi: 10.1593/neo.81328.
- [13] Bihan, "MR Imaging of Intravoxel Incoherent Motions: Application to Diffusion and Perfusion in Neurological Disorders " 1986.
- [14] S. C. Thust *et al.*, "Glioma imaging in Europe: A survey of 220 centres and recommendations for best clinical practice," *European Radiology*, vol. 28, no. 8, pp. 3306-3317, 2018-08-01 2018, doi: 10.1007/s00330-018-5314-5.
- [15] D. C. Sullivan *et al.*, "Metrology Standards for Quantitative Imaging Biomarkers," *Radiology*, vol. 277, no. 3, pp. 813-825, 2015, doi: 10.1148/radiol.2015142202.
- [16] R. J. Nordstrom and J. F. Eary, "Introduction and Background," in *Quantitative Imaging in Medicine: Background and Basics*, R. J. Nordstrom Ed., 1 ed. Melville, NY: AIP Publishing, 2021, ch. 1, pp. 1-8.

- [17] Biomarkers and D. W. Group, "Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework," *Clinical Pharmacology & Therapeutics*, vol. 69, no. 3, pp. 89-95, 2001, doi: 10.1067/mcp.2001.113989.
- [18] E. Manfrini *et al.*, "From research to clinical practice: a European neuroradiological survey on quantitative advanced MRI implementation," *European Radiology*, vol. 31, no. 8, pp. 6334-6341, 2021, doi: 10.1007/s00330-020-07582-2.
- [19] A. Shenoy-Bhangle, V. Baliyan, H. Kordbacheh, A. R. Guimaraes, and A. Kambadakone, "Diffusion weighted magnetic resonance imaging of liver: Principles, clinical applications and recent updates," *World Journal of Hepatology*, vol. 9, no. 26, p. 1081, 2017, doi: 10.4254/wjh.v9.i26.1081.
- [20] J. Virostko *et al.*, "Dynamic contrast-enhanced magnetic resonance imaging and diffusion-weighted magnetic resonance imaging for predicting the response of locally advanced breast cancer to neoadjuvant therapy: a meta-analysis," *Journal of Medical Imaging*, vol. 5, no. 1, p. 011011, 2017, doi: 10.1117/1.jmi.5.1.011011.
- [21] P. S. Tofts, "Concepts: Measurement in MRI," in *Quantitative MRI of the Brain: Principles of Physical Measurement*, M. Cercignani, N. G. Dowell, and P. S. Tofts Eds., 2 ed. Boca Raton, FL: CRC Press, 2018, ch. 1.
- [22] M. Cashmore, A. McCann, S. Wastling, C. McGrath, J. Thornton, and M. Hall, "Clinical quantitative MRI and the need for metrology," *British Journal of Radiology*, vol. 94, no. 1120, p. 20201215, 2021.
- [23] Radiological Society of North America. "Quantitative Imaging Biomarkers Alliance." Radiological Society of North America. <https://www.rsna.org/research/quantitative-imaging-biomarkers-alliance> (accessed 28 July, 2022).
- [24] A. R. Guimaraes, "Quantitative Imaging Biomarker Alliance (QIBA): Protocols and Profiles," in *Quantitative Imaging in Medicine: Background and Basics* R. Nordstrom Ed., 1 ed. Melville, NY: AIP Publishing, 2021, ch. 3, pp. 3-1 - 3-21.
- [25] K. E. Keenan *et al.*, "Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom," *Magnetic Resonance in Medicine*, vol. 79, no. 1, pp. 48-61, 2018, doi: 10.1002/mrm.26982.
- [26] K. H. Wong, R. Panek, S. A. Bhide, C. M. Nutting, K. J. Harrington, and K. L. Newbold, "The emerging potential of magnetic resonance imaging in personalizing radiotherapy for head and neck cancer: an oncologist's perspective," *British Journal of Radiology*, vol. 90, no. 1071, 2017, doi: 10.1259/bjr.20160768.
- [27] I. Serifoglu, I. I. Oz, M. Damar, O. Tokgoz, O. Yazgan, and Z. Erdem, "Diffusion-weighted imaging in the head and neck region: usefulness of apparent diffusion coefficient values for characterization of lesions," *Diagnostic and Interventional Radiology*, vol. 21, no. 3, pp. 208-214, 2015, doi: 10.5152/dir.2014.14279.
- [28] J. Wang *et al.*, "Head and Neck Lesions: Characterization with Diffusion-weighted Echo-planar MR Imaging," *Radiology*, vol. 220, no. 3, pp. 621-630, 2001.
- [29] M. Barral *et al.*, "Diffusion-weighted MR Imaging of the Pancreas: Current Status and Recommendations," *Radiology*, vol. 274, no. 1, pp. 45-63, 2015.
- [30] N. C. Balci, W. H. Perman, S. Saglam, F. Akisik, R. Fattahi, and M. Bilgin, "Diffusion-Weighted Magnetic Resonance Imaging of the Pancreas," *Top Magn Reson Imaging*, vol. 20, no. 1, pp. 43-47, 2009, doi: 10.1097/RMR.0b013e3181b48667.
- [31] A. B. Rosenkrantz, M. Oei, J. S. Babb, B. E. Niver, and B. Taouli, "Diffusion-Weighted Imaging of the Abdomen at 3.0 Tesla: Image Quality and Apparent Diffusion Coefficient Reproducibility Compared With 1.5 Tesla," *Journal of Magnetic Resonance Imaging*, vol. 33, no. 1, pp. 128-135, 2011, doi: 10.1002/jmri.22395.
- [32] H. Bickel *et al.*, "Quantitative Apparent Diffusion Coefficient as a Noninvasive Imaging Biomarker for the Differentiation of Invasive Breast Cancer and Ductal Carcinoma In Situ," *Investigative Radiology*, vol. 50, no. 2, pp. 95-100, 2015.
- [33] E. Aliotta *et al.*, "Automated apparent diffusion coefficient analysis for genotype prediction in lower grade glioma: association with the T2-FLAIR mismatch sign," *Journal of Neuro-Oncology*, vol. 149, no. 2, pp. 325-335, 2020, doi: 10.1007/s11060-020-03611-8.

- [34] X.-R. Li *et al.*, "DW-MRI ADC values can predict treatment response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy," *Medical Oncology*, vol. 29, no. 2, pp. 425-431, 2012, doi: 10.1007/s12032-011-9842-y.
- [35] S. C. Partridge *et al.*, "Diffusion-weighted MRI Findings Predict Pathologic Response in Neoadjuvant Treatment of Breast Cancer: The ACRIN 6698 Multicenter Trial," *Radiology*, vol. 289, no. 3, pp. 618-627, 2018, doi: 10.1148/radiol.2018180273.
- [36] C. Paterson *et al.*, "Study of diffusion weighted MRI as a predictive biomarker of response during radiotherapy for high and intermediate risk squamous cell cancer of the oropharynx: The MeRIInO study," *Clinical and Translational Radiation Oncology*, vol. 2, pp. 13-18, 2017, doi: 10.1016/j.ctro.2016.12.003.
- [37] T. L. Chenevert *et al.*, "Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors," *Journal of the National Cancer Institute*, vol. 92, no. 24, pp. 2029-2036, 2000, doi: 10.1093/jnci/92.24.2029.
- [38] C. J. Galbán *et al.*, "Multi-Site Clinical Evaluation of DW-MRI as a Treatment Response Metric for Breast Cancer Patients Undergoing Neoadjuvant Chemotherapy," *PLOS ONE*, vol. 10, no. 3, p. e0122151, 2015, doi: 10.1371/journal.pone.0122151.
- [39] V. Vandecaveye *et al.*, "Evaluation of the larynx for tumour recurrence by diffusion-weighted MRI after radiotherapy: initial experience in four cases," *The British Journal of Radiology*, vol. 79, no. 944, pp. 681-687, 2006.
- [40] T. C. Kwee, T. Takahara, D. M. Koh, R. A. J. Nivelstein, and P. R. Luijten, "Comparison and Reproducibility of ADC Measurements in Breathhold, Respiratory Triggered, and Free-Breathing Diffusion-Weighted MR Imaging of the Liver," *Journal of Magnetic Resonance Imaging*, vol. 28, no. 5, pp. 1141-1148, 2008, doi: 10.1002/jmri.21569.
- [41] M. Sasaki *et al.*, "Variability in Absolute Apparent Diffusion Coefficient Values across Different Platforms May Be Substantial: A Multivendor, Multi-institutional Comparison Study," *Radiology*, vol. 249, no. 2, pp. 624-630, 2008.
- [42] D. C. Newitt *et al.*, "Multisite concordance of apparent diffusion coefficient measurements across the NCI Quantitative Imaging Network," *Journal of Medical Imaging*, vol. 5, no. 1, p. 011003, 2018, doi: 10.1117/1.JMI.5.1.011003.
- [43] B. M. Dale, A. C. Braithwaite, D. T. Boll, and E. M. Merkle, "Field Strength and Diffusion Encoding Technique Affect the Apparent Diffusion Coefficient Measurements in Diffusion-Weighted Imaging of the Abdomen," *Investigative Radiology*, vol. 45, no. 2, pp. 104-108, 2010.
- [44] H. Merisaari *et al.*, "Diffusion-weighted imaging of prostate cancer: effect of b-value distribution on repeatability and cancer characterization," *Magnetic Resonance Imaging*, vol. 33, no. 10, pp. 1212-1218, 2015.
- [45] C. S. Schouten *et al.*, "Diffusion-weighted EPI- and HASTE-MRI and 18F-FDG-PET-CT early during chemoradiotherapy in advanced head and neck cancer," *Quant Imaging Med Surg*, vol. 4, no. 4, pp. 239-250, 2014, doi: 10.3978/j.issn.2223-4292.2014.07.15.
- [46] N. P. Jerome *et al.*, "Development of a temperature-controlled phantom for magnetic resonance quality assurance of diffusion, dynamic, and relaxometry measurements," *Medical Physics*, vol. 43, no. 6, pp. 2998-3007, 2016, doi: 10.1118/1.4948997.
- [47] C. Pierpaoli, J. Sarlls, U. Nevo, P. J. Basser, and F. Horkay, "Polyvinylpyrrolidone (PVP) water solutions as isotropic phantoms for diffusion MRI studies," *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 17, p. 1414, 2009.
- [48] P. S. Tofts *et al.*, "Test liquids for quantitative MRI measurements of self-diffusion coefficient in vivo," *Magnetic Resonance in Medicine*, vol. 43, no. 3, pp. 368-374, 2000.
- [49] P. E. Kinahan, T. L. Chenevert, and D. Malyarenko, "Standards, Phantoms, and Site Qualification," in *Quantitative Imaging in Medicine: Background and Basics*, R. J. Nordstrom Ed. Melville, NY: AIP Publishing, 2021, pp. 2-1- 2-26.
- [50] N. M. deSouza *et al.*, "Implementing diffusion-weighted MRI for body imaging in prospective multicentre trials: current considerations and future perspectives," *European Radiology*, vol. 28, no. 3, pp. 1118-1131, 2018, doi: 10.1007/s00330-017-4972-z.

- [51] E. M. Palacios *et al.*, "Toward Precision and Reproducibility of Diffusion Tensor Imaging: A Multicenter Diffusion Phantom and Traveling Volunteer Study," *American Journal of Neuroradiology*, vol. 38, no. 3, pp. 537-545, 2017, doi: 10.3174/ajnr.a5025.
- [52] B. Moreau, A. Iannessi, C. Hoog, and H. Beaumont, "How reliable are ADC measurements? A phantom and clinical study of cervical lymph nodes," *European Radiology*, vol. 28, no. 8, pp. 3362-3371, 2018, doi: 10.1007/s00330-017-5265-2.
- [53] E. Giannotti, S. Waugh, L. Priba, Z. Davis, E. Crowe, and S. Vinnicombe, "Assessment and quantification of sources of variability in breast apparent diffusion coefficient (ADC) measurements at diffusion weighted imaging," *European Journal of Radiology*, vol. 84, no. 9, pp. 1729-1736, 2015, doi: 10.1016/j.ejrad.2015.05.032.
- [54] M. Grech-Sollars *et al.*, "Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain," *NMR in Biomedicine*, vol. 28, no. 4, pp. 468-485, 2015, doi: 10.1002/nbm.3269.
- [55] M. E. Miquel, A. D. Scott, N. D. MacDougall, R. Boubertakh, N. Bharwani, and A. G. Rockall, "In vitro and in vivo repeatability of abdominal diffusion-weighted MRI," *The British Journal of Radiology*, vol. 85, pp. 1507-1512, 2012, doi: 10.1259/bjr/32269440.
- [56] H. J. Laubach *et al.*, "A phantom for diffusion-weighted imaging of acute stroke," *Journal of Magnetic Resonance Imaging*, vol. 8, no. 6, pp. 1349-1354, 1998, doi: 10.1002/jmri.1880080627.
- [57] Y. Wang *et al.*, "Quantitative MRI: Defining repeatability, reproducibility and accuracy for prostate cancer imaging biomarker development," *Magnetic Resonance Imaging*, vol. 77, pp. 169-179, 2021, doi: 10.1016/j.mri.2020.12.018.
- [58] I. Delakis, E. M. Moore, M. O. Leach, and J. P. De Wilde, "Developing a quality control protocol for diffusion imaging on a clinical MRI system," *Physics in Medicine and Biology*, vol. 49, pp. 1409-1422, 2004, doi: 10.1088/0031-9155/49/8/003.
- [59] T. L. Chenevert *et al.*, "Errors in Quantitative Image Analysis due to Platform-Dependent Image Scaling," *Translational Oncology*, vol. 7, no. 1, pp. 65-71, 2014, doi: 10.1593/tlo.13811.
- [60] G. Belli *et al.*, "Quality assurance multicenter comparison of different MR scanners for quantitative diffusion-weighted imaging," *Journal of Magnetic Resonance Imaging*, vol. 43, no. 1, pp. 213-219, 2015, doi: 10.1002/jmri.24956.
- [61] T. L. Chenevert *et al.*, "Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies," *Journal of Magnetic Resonance Imaging*, vol. 34, no. 4, pp. 983-987, 2011, doi: 10.1002/jmri.22363.
- [62] D. Malyarenko *et al.*, "Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom," *Journal of Magnetic Resonance Imaging*, vol. 37, no. 5, pp. 1238-1246, 2013, doi: 10.1002/jmri.23825.
- [63] D. I. Malyarenko *et al.*, "Demonstration of nonlinearity bias in the measurement of the apparent diffusion coefficient in multicenter trials," *Magnetic Resonance in Medicine*, vol. 75, no. 3, pp. 1312-1323, 2016, doi: 10.1002/mrm.25754.
- [64] M. A. Boss *et al.*, "Temperature-controlled Isotropic Diffusion Phantom with Wide Range of Apparent Diffusion Coefficients for Multicenter Assessment of Scanner Repeatability and Reproducibility," *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 22, p. 4505, 2014.
- [65] M. Boss *et al.*, "TU-C-12A-08: Thermally-Stabilized Isotropic Diffusion Phantom for Multisite Assessment of Apparent Diffusion Coefficient Reproducibility," *Medical Physics*, vol. 41, no. 6, pp. 464-464, 2014, doi: 10.1118/1.4889298.
- [66] K. E. Keenan *et al.*, "Recommendations towards standards for quantitative MRI (qMRI) and outstanding needs," *Journal of Magnetic Resonance Imaging*, vol. 49, no. 7, pp. e26-e39, 2019, doi: 10.1002/jmri.26598.
- [67] R. Paudyal *et al.*, "Repeatability of Quantitative Diffusion-Weighted Imaging Metrics in Phantoms, Head-and-Neck and Thyroid Cancers: Preliminary Findings," *Tomography*, vol. 5, no. 1, pp. 15-25, 2019, doi: 10.18383/j.tom.2018.00044.
- [68] E. S. Kooreman *et al.*, "Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator," *Radiotherapy and Oncology*, vol. 133, pp. 156-162, 2019, doi: 10.1016/j.radonc.2019.01.011.

- [69] P. J. van Houdt *et al.*, "Phantom-based quality assurance for multicenter quantitative MRI in locally advanced cervical cancer," *Radiotherapy and Oncology*, vol. 153, pp. 114-121, 2020, doi: 10.1016/j.radonc.2020.09.013.
- [70] B. Lewis *et al.*, "Evaluation of diffusion-weighted MRI and geometric distortion on a 0.35T MR-LINAC at multiple gantry angles," *Journal of Applied Clinical Medical Physics*, vol. 22, no. 2, pp. 118-125, 2021, doi: 10.1002/acm2.13135.
- [71] P. J. van Houdt *et al.*, "Integration of quantitative imaging biomarkers in clinical trials for MR-guided radiotherapy: Conceptual guidance for multicentre studies from the MR-Linac Consortium Imaging Biomarker Working Group," *European Journal of Cancer*, vol. 153, pp. 64-71, 2021, doi: 10.1016/j.ejca.2021.04.041.
- [72] M. E. Carr *et al.*, "Conformance of a 3T radiotherapy MRI scanner to the QIBA Diffusion Profile," *Medical Physics*, vol. 49, no. 7, pp. 4508-4517, 2022, doi: 10.1002/mp.15645.
- [73] CaliberMRI. "DWI Standardisation Realized through Facilitated Diffusion." <https://qmri.com/diffusion-phantom-containing/> (accessed 05 Aug, 2022).
- [74] CaliberMRI, "Diffusion phantom for ADC qMRI Standardisation (Rev G)," 2021.
- [75] K. E. Keenan, K. F. Stupic, S. E. Russek, and E. Mirowski, "MRI-visible liquid crystal thermometer," *Magnetic Resonance in Medicine*, vol. 84, no. 3, pp. 1552-1563, 2020-09-01 2020, doi: 10.1002/mrm.28224.
- [76] E. Mirowski, M. Snow, and K. Keenan, "Liquid crystal thermometer for MRI," US Patent Appl. 15/912,420, 2020. [Online]. Available: <https://patents.justia.com/patent/10809331>
- [77] Quantitative Imaging Biomarkers Alliance, "QIBA profile conformance testing DWI MR Supplement 1," 2021.
- [78] NHS GGC MRI Physics. "Diffusion MRI, ADC values for QIBA using Caliber MRI diffusion phantom." <https://www.mriphysics.scot.nhs.uk/diffusion-mri-adc-values-for-qiba-using-caliber-mri-diffusion-phantom/> (accessed 07 Aug, 2022).
- [79] CaliberMRI. "qCal-MR Protocols." https://qcalsoftware.com/protocol_definitions (accessed 28 July, 2022).
- [80] D. L. Raunig *et al.*, "Quantitative imaging biomarkers: A review of statistical methods for technical performance assessment," *Statistical Methods in Medical Research*, vol. 24, no. 1, pp. 27-67, 2015, doi: 10.1177/0962280214537344.
- [81] A. Shukla-Dave *et al.*, "Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials," *Journal of Magnetic Resonance Imaging*, vol. 49, no. 7, pp. e101-e121, 2019, doi: 10.1002/jmri.26518.
- [82] X. Wang, S. B. Reeder, and D. Hernando, "An acetone-based phantom for quantitative diffusion MRI," *Journal of Magnetic Resonance Imaging*, vol. 46, no. 6, pp. 1683-1692, 2017, doi: 10.1002/jmri.25727.
- [83] L. G. Kessler *et al.*, "The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions," *Statistical Methods in Medical Research*, vol. 24, no. 1, pp. 9-26, 2015, doi: 10.1177/0962280214537333.
- [84] N. M. Rofsky *et al.*, "Abdominal MR Imaging with a Volumetric Interpolated Breath-hold Examination," *Radiology*, vol. 212, no. 3, pp. 876-884, 1999.
- [85] R. P. Lim *et al.*, "3D Time-Resolved MR Angiography (MRA) of the Carotid Arteries with Time-Resolved Imaging with Stochastic Trajectories: Comparison with 3D Contrast-Enhanced Bolus-Chase MRA and 3D Time-Of-Flight MRA," *American Journal of Neuroradiology*, vol. 29, no. 10, pp. 1847-1854, 2008, doi: 10.3174/ajnr.a1252.
- [86] V. M. Runge and J. T. Heverhagen, "Chemical Shift: Sampling Bandwidth," in *The Physics of Clinical MR Taught Through Images*: Springer International Publishing, 2022, pp. 246-247.
- [87] P. Pullens, P. Bladt, J. Sijbers, A. I. R. Maas, and P. M. Parizel, "Technical Note: A safe, cheap, and easy-to-use isotropic diffusion MRI phantom for clinical and multicenter studies," *Medical Physics*, vol. 44, no. 3, pp. 1063-1070, 2017, doi: 10.1002/mp.12101.
- [88] D. Le Bihan, C. Poupon, A. Amadon, and F. Lethimonnier, "Artifacts and pitfalls in diffusion MRI," *Journal of Magnetic Resonance Imaging*, vol. 24, no. 3, pp. 478-488, 2006, doi: 10.1002/jmri.20683.

- [89] S. C. A. Steens *et al.*, "Reproducibility of brain ADC histograms," *European Radiology*, vol. 14, no. 3, pp. 425-430, 2004, doi: 10.1007/s00330-003-2121-3.
- [90] M. A. Boss *et al.*, "Repeatability and Reproducibility Assessment of the Apparent Diffusion Coefficient in the Prostate: A Trial of the ECOG-ACRIN Research Group (ACRIN6701)," *Journal of Magnetic Resonance Imaging*, 2022, doi: 10.1002/jmri.28093.

APPENDICES

APPENDIX A: INSTRUCTIONS FOR PERFORMING AN ICE BATH WITH THE CALIBERMRI DIFFUSION PHANTOM

Full instructions are provided in the QIBA DWI Supplementary 1 [77]. A summary of the method followed in this report is included below.

Required materials:

- Bucket, large enough to completely submerge the phantom in
- Jug for pouring chilled water into the phantom
- Ice, a mixture of both ice cubes and crushed ice is ideal
- Chilled water
- Phantom tools: torque wrench with a 5 mm hex driver bit & straight screwdriver bit (provided with the phantom)
- Pipette (provided with the phantom)
- Digital thermometer

Optional materials:

- Insulated bag and/or blanket aid in temperature stabilisation
- Bowl to hold the phantom in the fridge
- Clear plastic bag to place the phantom in while scanning to prevent condensation on the scanner
- Rolling pin to crush ice cubes into smaller pieces

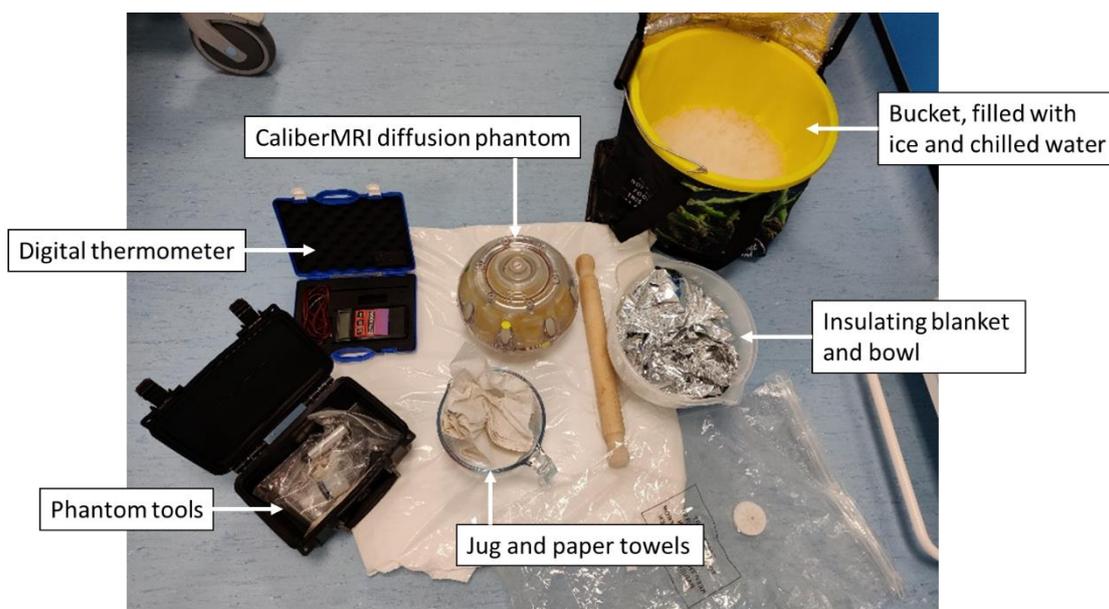


Figure A1. Required materials for performing an ice bath scan with the CaliberMRI phantom.

Procedure:

- Part 1: Pre-chill
 - Empty the phantom by removing both end covers of the phantom using the provided torque wrench.
 - Pre-chill the phantom by submerging the open phantom in a bucket containing a crushed ice and chilled water mix.
 - Leave the phantom submerged for approximately 15 minutes.



Figure A2. The phantom is emptied by removing the end covers (left) and then submerged in a bucket filled with ice and chilled water for 15 minutes (right).

- Part 2: Filling with ice-water
 - Remove the phantom from the bucket, emptying any water inside.
 - Fill one half of the phantom with ice and replace the cover on that half of the phantom.
 - Turn the phantom over and fill the other half of the phantom with ice. The phantom should be filled with as much ice as possible.
 - Pour ice-chilled water from the bucket into the phantom, filling any gaps in the ice, then replace the end cover.
 - With both ends replaced, top up the phantom through the fine fill port to remove any air bubbles.

- Part 3: Temperature stabilisation
 - Place the phantom in the fridge for approximately 2 hours. If available, wrap the phantom in an insulating blanket.
 - Before scanning, check for air bubbles and top up the phantom with ice-chilled water through the fine fill port if required.
 - Measure the temperature of the phantom through the fine fill port on both sides to confirm 0°C has been reached.



Figure A3. Left: The phantom is left in the fridge for approximately 2 hours, wrapped in an insulating blanket. Right: Phantom temperature is measured through the fine fill port.

APPENDIX B: RECOMMENDED SEQUENCE PARAMETERS PROVIDED BY CALIBERMRI

Table B1. Updated sequence parameter tables provided by CaliberMRI for the T1 thermometer scan and the DWI scan.

	Isotropic Volume	DWI
Sequence	3D/SPGR	DWI EPI, Single Spin-Echo or Mono-Polar
Scan Options	GE: EDR Philips: 3D FFE; Fast = none	
Scan Plane	Coronal	Coronal
Feature / AP Offset from Center (mm LPS)	Full Volume Scan LC vials / -18 to -5	Full Volume Scan Diffusion mimics / -35 to 40
B Values (s/mm ²)	-	0, 500, 1000, 1500, 2000
TR (ms)	< 10 (~6 preferred)	8000
TE (ms)	Minimum (< 4)	Minimum (< 130)
T1 Values (ms)	-	-
Flip Angle (deg)	10	90
Fat Suppression	False	False
ETL / Turbo Factor	-	-
Number of Averages	1	2
Parallel Imaging Factor	-	2
FOV (FE, mm)	250	220
FOV (PE, mm)	250	220
Pixel Size (mm x mm)	0.98 x 0.98	<= 1.72 x 1.72
Section Thickness / Gap (mm)	1 / 0	4 / 1
Matrix (FE)	256	>= 128
Matrix (PE)	256	>= 128
Matrix (SE) / # of Slices	192	25
Pixel Bandwidth (Hz)	> 400	> 1000
PE Direction	S/I	S/I
Notes		Perform 4 passes. Full prep for pass 1; auto prep or manual prep with no adjustments for pass 2, 3, and 4.
Approx. Acq. Time per Setting (min)		4
# Settings	1	4
Total Time for this Series (min)		16

APPENDIX C: FULL SEQUENCE PARAMETERS

.pdf and .exar files are available to download at: <https://www.mriphysics.scot.nhs.uk/diffusion-mri-adc-values-for-qiba-using-caliber-mri-diffusion-phantom/> [78]

C.1 DWI Sequence, coronal orientation, 3T Siemens Prisma

SIEMENS MAGNETOM Prisma

\\Physics\John Mc\Diffusion_QA\Diffusion_qa_qiba_v3\ep2d_diff_cor	
TA: 3:54 PM: REF Voxel size: 1.7×1.7×4.0 mmPAT: 2 Rel. SNR: 1.00 : epse	
Properties	
Prio recon	Off
Load images to viewer	On
Inline movie	Off
Auto store images	On
Load images to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Auto close inline display	Off
Start measurement without further preparation	Off
Wait for user to start	Off
Start measurements	Single measurement
Routine	
Slice group	1
Slices	35
Dist. factor	25 %
Position	L0.0 P0.0 H2.4 mm
Orientation	C > T-0.3
Phase enc. dir.	F >> H
AutoAlign	---
Phase oversampling	0 %
FoV read	220 mm
FoV phase	100.0 %
Slice thickness	4.0 mm
TR	8000 ms
TE	81.0 ms
Averages	2
Concatenations	1
Filter	Raw filter, Prescan Normalize
Coil elements	HEA;HEP
Contrast - Common	
TR	8000 ms
TE	81.0 ms
MTC	Off
Magn. preparation	None
Fat suppr.	None
Contrast - Dynamic	
Averages	2
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	1
Delay in TR	0 ms
Resolution - Common	
FoV read	220 mm
FoV phase	100.0 %
Slice thickness	4.0 mm
Base resolution	130
Phase resolution	100 %
Phase partial Fourier	6/8
Interpolation	Off
Resolution - iPAT	
Accel. mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	40
Resolution - iPAT	
Reference scan mode	EPI/separate
Resolution - Filter Image	
Distortion Corr.	Off
Prescan Normalize	On
Dynamic Field Corr.	Off
Resolution - Filter Rawdata	
Raw filter	On
Elliptical filter	Off
Geometry - Common	
Slice group	1
Slices	35
Dist. factor	25 %
Position	L0.0 P0.0 H2.4 mm
Orientation	C > T-0.3
Phase enc. dir.	F >> H
FoV read	220 mm
FoV phase	100.0 %
Slice thickness	4.0 mm
TR	8000 ms
Multi-slice mode	Interleaved
Series	Interleaved
Concatenations	1
Geometry - AutoAlign	
Slice group	1
Position	L0.0 P0.0 H2.4 mm
Orientation	C > T-0.3
Phase enc. dir.	F >> H
AutoAlign	---
Initial Position	L0.0 P0.0 H2.4
L	0.0 mm
P	0.0 mm
H	2.4 mm
Initial Rotation	90.00 deg
Initial Orientation	C > T
C > T	-0.3
> S	0.0
Geometry - Saturation	
Fat suppr.	None
Special sat.	None
Geometry - Navigator	
System - Miscellaneous	
Positioning mode	REF
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Adaptive Combine
Matrix Optimization	Off
AutoAlign	---
Coil Select Mode	On - AutoCoilSelect

QA for Assessing Quantitative Biomarkers in Patients using MRI

SIEMENS MAGNETOM Prisma

System - Adjustments

B0 Shim mode	Standard
B1 Shim mode	TrueForm
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Dominant Fat	Off
Assume Silicone	Off
Adjustment Tolerance	Auto

System - Adjust Volume

! Position	Isocenter
! Orientation	Coronal
! Rotation	0.00 deg
! R >> L	216 mm
! F >> H	216 mm
! A >> P	35 mm
Reset	Off

System - pTx Volumes

B1 Shim mode	TrueForm
Excitation	Standard

System - Tx/Rx

Frequency 1H	123.205026 MHz
Correction factor	1
Gain	High
Img. Scale Cor.	1.000
Reset	Off
? Ref. amplitude 1H	0.000 V

Physio - Signal1

1st Signal/Mode	None
TR	8000 ms
Concatenations	1

Physio - PACE

Resp. control	Off
Concatenations	1

Diff - Neuro

Diffusion mode	3-Scan Trace
Diff. directions	3
Diffusion Scheme	Bipolar
Diff. weightings	5
b-value 1	0 s/mm ²
b-value 2	500 s/mm ²
b-value 3	1000 s/mm ²
b-value 4	1500 s/mm ²
b-value 5	2000 s/mm ²
b-value 1	2
b-value 2	2
b-value 3	2
b-value 4	2
b-value 5	2
Diff. weighted images	Off
Trace weighted images	On
ADC maps	On
FA maps	Off
Mosaic	Off
Tensor	Off
Noise level	40

Diff - Body

Diffusion mode	3-Scan Trace
----------------	--------------

Diff - Body

Diff. directions	3
Diffusion Scheme	Bipolar
Diff. weightings	5
b-value 1	0 s/mm ²
b-value 2	500 s/mm ²
b-value 3	1000 s/mm ²
b-value 4	1500 s/mm ²
b-value 5	2000 s/mm ²
b-value 1	2
b-value 2	2
b-value 3	2
b-value 4	2
b-value 5	2
Diff. weighted images	Off
Trace weighted images	On
ADC maps	On
Exponential ADC Maps	Off
FA maps	Off
Invert Gray Scale	Off
Calculated Image	Off
b-Value >=	0 s/mm ²
Noise level	40

Diff - Composing

Distortion Corr.	Off
------------------	-----

Sequence - Part 1

Introduction	On
Optimization	None
Multi-slice mode	Interleaved
Free echo spacing	Off
Echo spacing	0.93 ms
Bandwidth	1202 Hz/Px

Sequence - Part 2

EPI factor	130
RF pulse type	Normal
Gradient mode	Performance*
Excitation	Standard

Sequence - pTX Pulses

C.2 DWI Sequence, coronal orientation, 1.5T Siemens Sola

SIEMENS MAGNETOM 1.5T XQ Numaris/X VA31A-012Y

\ZZ_DWI_QIBA\DWI_QIBA\QA_MSC\QA_200622\QIBA_diff_qa_cor	
TA: 3:54 min Coil Selection: Auto Voxel Size: 1.7x1.7x4.0 mm ³ Acc:: 2 Rel. SNR: 1.00	
Properties	
Start measurement without further preparation	Off
Wait for User to Start	Off
Start measurements	Single Measurement
Prio Recon	Off
Auto Open Inline Display	Off
Auto Close Inline Display	Off
Load Images to MR View&GO	On
Auto Store Images	On
Load Images to Stamp Segments	Off
Load Images to Graphic Segments	Off
Graphic segment	Default
Inline Movie	Off
Routine	
Slice Group	1
Slices	35
Distance Factor	25 %
Position	L0.0 P18.6 H0.6 mm
Orientation	C > T-0.7 > S-0.6
Phase Encoding Dir.	F >> H
Phase Oversampling	0 %
FoV Read	220 mm
FoV Phase	100.0 %
Slice Thickness	4.0 mm
TR	8000.0 ms
TE	96.00 ms
Concatenations	1
AutoAlign	---
Contrast - Common	
TR	8000.0 ms
TE	96.00 ms
MTC	Off
Magn. Preparation	None
Fat-Water Contrast	Standard
Reconstruction	Magnitude
Contrast - Dynamic	
Dynamic Mode	Standard
Delay in TR	0.00 ms
Resolution - Common	
FoV Read	220 mm
FoV Phase	100.0 %
Slice Thickness	4.0 mm
Base Resolution	130
Phase Resolution	100 %
Interpolation	Off
Resolution - Acceleration	
Acceleration mode	GRAPPA
Reference Scans	EPI/Separate
Acceleration Factor PE	2
Reference Lines PE	40
Phase Partial Fourier	6/8
Resolution - Filter	
Raw Filter	On
Elliptical Filter	Off
Resolution - Filter	
Distortion Correction	2D
Normalize	Prescan
Noise Masking	Off
Geometry - Common	
Slice Group	1
Slices	35
Distance Factor	25 %
Position	L0.0 P18.6 H0.6 mm
Orientation	C > T-0.7 > S-0.6
Phase Encoding Dir.	F >> H
Phase Oversampling	0 %
FoV Read	220 mm
FoV Phase	100.0 %
Slice Thickness	4.0 mm
TR	8000.0 ms
Multi-Slice Mode	Interleaved
Series	Interleaved
Concatenations	1
Geometry - AutoAlign	
Slice Group	1
Position	L0.0 P18.6 H0.6 mm
Orientation	C > T-0.7 > S-0.6
Phase Encoding Dir.	F >> H
AutoAlign	---
Initial Position	L0.0 P18.6 H0.6
L	0.0 mm
P	18.6 mm
H	0.6 mm
Initial Orientation	C > T
C > T	-0.70
> S	-0.60
Initial Rotation	90.00 deg
Geometry - Navigator	
Geometry - Saturation	
Special Saturation	None
Geometry - Tim Planning Suite	
Set-n-Go Protocol	Off
Table Position	0 mm
Table Position	H
Inline Composing	Off
System - Miscellaneous	
Coil Selection	Auto Coil Select
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combination	Adaptive Combine
Matrix Optimization	Off
Coil Focus	Flat
System - Adjustments	
Adjustment Strategy	Standard
B0 Shim	Standard
CoilShim	Off

QA for Assessing Quantitative Biomarkers in Patients using MRI

SIEMENS MAGNETOM 1.5T XQ Numaris/X VA31A-012Y

System - Adjustments

Adjustment Tolerance	Auto
Adjust with Body Coil	Off
Confirm Frequency	Never
Assume Silicone	Off

System - Adjust Volume

Position	L0.0 P18.6 H0.6 mm
Orientation	C > T-0.7 > S-0.6
Rotation	90.00 deg
F >> H	220 mm
R >> L	220 mm
A >> P	174 mm
Reset	Off

System - Tx/Rx

Frequency 1H	63.680140 MHz
? Ref. Amplitude 1H	0.000 V
Reset	Off
Image Scaling	1.000

Physio - Signal

1st Signal/Mode	None
TR	8000.0 ms
Concatenations	1

Physio - PACE

Resp. Control	Off
Concatenations	1

Diff

Diffusion Mode	3-Scan Trace
Diff. Directions	3
Diffusion Scheme	Bipolar
Diff. Weightings	5
b-value 1	0 s/mm ²
b-value 2	500 s/mm ²
b-value 3	1000 s/mm ²
b-value 4	1500 s/mm ²
b-value 5	2000 s/mm ²
Averages 1	2
Averages 2	2
Averages 3	2
Averages 4	2
Averages 5	2
Dynamic Field Correction	Off
Invert Gray Scale	Off
Diff. Weighted Images	Off
Trace Weighted Images	On
Tensor	Off
FA Maps	Off
ADC Maps	On
Exponential ADC Maps	Off
b-value >=	0 s/mm ²
ADC Noise Threshold	40
Noise Masking	Off
Calculated Image	Off

Sequence - Part 1

Sequence Name	epse
Excitation	Standard
RF Pulse Type	Normal
Gradient Mode	Fast
Bandwidth	1132 Hz/Px
Echo Spacing	0.97 ms

Sequence - Part 1

Free Echo Spacing	Off
Optimization	None
EPI Factor	130

Sequence - Part 2

Introduction	On
Phase Correction	Internal

C.3 T1-weighted VIBE Sequence for viewing the LC thermometer, 3T Siemens Prisma

SIEMENS MAGNETOM Prisma

\\Physics\John Mc\Diffusion_QA\Diffusion_qa_qiba_v3t1_vibe_fs_cor_p2_at_start TA: 3:25 PM: REF Voxel size: 1.0x1.0x1.0 mmPAT: 2 Rel. SNR: 1.00 : fl

Properties

Prio recon	Off
Load images to viewer	On
Inline movie	Off
Auto store images	On
Load images to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Auto close inline display	Off
Start measurement without further preparation	Off
Wait for user to start	Off
Start measurements	Single measurement

Routine

Slab group	1
Slabs	1
Dist. factor	20 %
Position	R0.6 A1.1 H3.6 mm
Orientation	C > T-0.4
Phase enc. dir.	F >> H
AutoAlign	Head > Orbits
Phase oversampling	0 %
Slice oversampling	0.0 %
Slices per slab	192
FoV read	256 mm
FoV phase	100.0 %
Slice thickness	1.0 mm
TR	7.58 ms
TE	2.88 ms
Averages	1
Concatenations	1
Filter	Prescan Normalize, Elliptical filter
Coil elements	HEA;HEP

Contrast - Common

TR	7.58 ms
TE	2.88 ms
Flip angle	10.0 deg
Fat suppr.	None
Water suppr.	None
Dixon	Off
Dixon evaluation	Off

Contrast - Dynamic

Averages	1
Averaging mode	Long term
Reconstruction	Magnitude
Measurements	1
Multiple series	Each measurement

Resolution - Common

FoV read	256 mm
FoV phase	100.0 %
Slice thickness	1.0 mm
Base resolution	256
Phase resolution	100 %
Slice resolution	100 %
Phase partial Fourier	Off
Slice partial Fourier	Off

Resolution - Common

Trajectory	Cartesian
View sharing	Off
Interpolation	Off

Resolution - iPAT

PAT mode	GRAPPA
Accel. factor PE	2
Ref. lines PE	24
Accel. factor 3D	1
Reference scan mode	Integrated

Resolution - Filter Image

Image Filter	Off
Distortion Corr.	Off
Prescan Normalize	On
Unfiltered images	Off
Normalize	Off
B1 filter	Off

Resolution - Filter Rawdata

Raw filter	Off
Elliptical filter	On
POCS	Off

Geometry - Common

Slab group	1
Slabs	1
Dist. factor	20 %
Position	R0.6 A1.1 H3.6 mm
Orientation	C > T-0.4
Phase enc. dir.	F >> H
Slice oversampling	0.0 %
Slices per slab	192
FoV read	256 mm
FoV phase	100.0 %
Slice thickness	1.0 mm
TR	7.58 ms
Multi-slice mode	Sequential
Series	Ascending
Concatenations	1

Geometry - AutoAlign

Slab group	1
Position	R0.6 A1.1 H3.6 mm
Orientation	C > T-0.4
Phase enc. dir.	F >> H
AutoAlign	Head > Orbits
Initial Position	R0.6 A1.1 H3.6
R	0.6 mm
A	1.1 mm
H	3.6 mm
Initial Rotation	88.58 deg
Initial Orientation	C > T
C > T	-0.4
> S	0.0

Geometry - Saturation

Fat suppr.	None
Water suppr.	None
Dixon	Off

QA for Assessing Quantitative Biomarkers in Patients using MRI

SIEMENS MAGNETOM Prisma

Geometry - Saturation

Dixon evaluation	Off
Special sat.	None

System - Miscellaneous

Positioning mode	REF
Table position	H
Table position	0 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Adaptive Combine
Save uncombined	Off
Matrix Optimization	Off
AutoAlign	Head > Orbits
Coil Select Mode	Off - AutoCoilSelect

System - Adjustments

B0 Shim mode	Standard
B1 Shim mode	TrueForm
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Dominant Fat	Off
Assume Silicone	Off
Adjustment Tolerance	Auto

System - Adjust Volume

Position	R0.6 A1.1 H3.6 mm
Orientation	C > T-0.4
Rotation	88.58 deg
F >> H	256 mm
R >> L	256 mm
A >> P	192 mm
Reset	Off

System - pTx Volumes

B1 Shim mode	TrueForm
Excitation	Slab-sel.

System - Tx/Rx

Frequency 1H	123.205026 MHz
Correction factor	1
Gain	Low
Img. Scale Cor.	1.000
Reset	Off
? Ref. amplitude 1H	0.000 V

Physio - PACE

Resp. control	Off
Concatenations	1

Inline - Common

View sharing	Off
Flip angle	10.0 deg
Measurements	1
Burn time-to-center	Off
Temporal interpolation	1
3D centric reordering	Off
Time to center	103.9 s

Inline - Inline

Subtract	Off
Measurements	1

Inline - Inline

StdDev	Off
Liver registration	Off
Save original images	On

Inline - MIP

MIP-Sag	Off
MIP-Cor	Off
MIP-Tra	Off
MIP-Time	Off
Save original images	On

Inline - Soft Tissue

Wash - In	Off
Wash - Out	Off
TTP	Off
PEI	Off
MIP - time	Off
Measurements	1

Inline - Composing

Distortion Corr.	Off
------------------	-----

Sequence - Part 1

Introduction	On
Dimension	3D
Elliptical scanning	Off
Asymmetric echo	Weak
Contrasts	1
Optimization	None
Multi-slice mode	Sequential
Bandwidth	810 Hz/Px

Sequence - Part 2

RF pulse type	Normal
Gradient mode	Normal
Excitation	Slab-sel.
RF spoiling	On
Incr. Gradient spoiling	On

Sequence - Assistant

Mode	Off
Allowed delay	30 s

C.4 T1-weighted TWIST Sequence for viewing the LC thermometer, 3T Siemens Prisma

SIEMENS MAGNETOM Prisma

\\Physics\John Mc\Diffusion_QA\Diffusion_qa_qiba_v3_twist\twist_cor_p4_iso_at_start	
TA: 0:39 PM: ISO Voxel size: 1.0x1.0x1.0 mmPAT: 4 Rel. SNR: 1.00 : fl	
Properties	
Prio recon	Off
Load images to viewer	On
Inline movie	Off
Auto store images	On
Load images to stamp segments	Off
Load images to graphic segments	Off
Auto open inline display	Off
Auto close inline display	Off
Start measurement without further preparation	Off
Wait for user to start	On
Start measurements	Single measurement
Routine	
Slab group	1
Slabs	1
Position	R1.9 A2.3 H3.9 mm
Orientation	Coronal
Phase enc. dir.	R >> L
AutoAlign	---
Phase oversampling	0 %
Slice oversampling	16.7 %
Slices per slab	192
FoV read	250 mm
FoV phase	100.0 %
Slice thickness	1.00 mm
TR	3.01 ms
TE	1.15 ms
Filter	Distortion Corr.(2D), Prescan Normalize, Elliptical filter
Coil elements	HEA;HEP
Resolution - iPAT	
Accel. factor PE	4
Ref. lines PE	24
Accel. factor 3D	1
Reference scan mode	Integrated
Resolution - Filter Image	
Image Filter	Off
Distortion Corr.	On
Mode	2D
Unfiltered images	Off
Prescan Normalize	On
Unfiltered images	Off
Normalize	Off
B1 filter	Off
Resolution - Filter Rawdata	
Raw filter	Off
Elliptical filter	On
Geometry - Common	
Slab group	1
Slabs	1
Position	R1.9 A2.3 H3.9 mm
Orientation	Coronal
Phase enc. dir.	R >> L
Slice oversampling	16.7 %
Slices per slab	192
FoV read	250 mm
FoV phase	100.0 %
Slice thickness	1.00 mm
TR	3.01 ms
Geometry - AutoAlign	
Slab group	1
Position	R1.9 A2.3 H3.9 mm
Orientation	Coronal
Phase enc. dir.	R >> L
AutoAlign	---
Initial Position	R1.9 A2.3 H3.9
R	1.9 mm
A	2.3 mm
H	3.9 mm
Initial Rotation	0.00 deg
Initial Orientation	Coronal
Geometry - Saturation	
Fat suppr.	None
Special sat.	None
System - Miscellaneous	
Positioning mode	ISO
Table position	H
Table position	4 mm
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combine Mode	Sum of Squares
Save uncombined	Off
Matrix Optimization	Off
Contrast - Common	
TR	3.01 ms
TE	1.15 ms
Flip angle	10 deg
Fat suppr.	None
Contrast - Dynamic	
Reconstruction	Magnitude
Measurements	1
Multiple series	Each measurement
Resolution - Common	
FoV read	250 mm
FoV phase	100.0 %
Slice thickness	1.00 mm
Base resolution	256
Phase resolution	100 %
Slice resolution	100 %
Phase partial Fourier	6/8
Slice partial Fourier	6/8
View sharing	TWIST
Central region A	15 %
Sampling density B	20 %
Interpolation	Off
Resolution - iPAT	
PAT mode	GRAPPA

QA for Assessing Quantitative Biomarkers in Patients using MRI

SIEMENS MAGNETOM Prisma

System - Miscellaneous

AutoAlign	---
Coil Select Mode	On - AutoCoilSelect

System - Adjustments

B0 Shim mode	Tune up
B1 Shim mode	TrueForm
Adjust with body coil	Off
Confirm freq. adjustment	Off
Assume Dominant Fat	Off
Assume Silicone	Off
Adjustment Tolerance	Maximum

System - Adjust Volume

Position	Isocenter
Orientation	Transversal
Rotation	0.00 deg
A >> P	263 mm
R >> L	350 mm
F >> H	350 mm
Reset	Off

System - pTx Volumes

B1 Shim mode	TrueForm
Excitation	Slab-sel.

System - Tx/Rx

Frequency 1H	123.204786 MHz
Correction factor	1
Gain	Low
Img. Scale Cor.	1.000
Reset	Off
? Ref. amplitude 1H	0.000 V

Angio - Common

View sharing	TWIST
Central region A	15 %
Sampling density B	20 %
Dynamic recon mode	Symmetric Share
Flip angle	10 deg
Temporal resolution	12.07 s
Measurements	1
Burn time-to-center	Off
Temporal interpolation	1
Time to center	3.8 s

Angio - Inline

Subtract	Off
Measurements	1
StdDev	Off
Save original images	On

Angio - MIP

MIP-Sag	Off
MIP-Cor	Off
MIP-Tra	Off
MIP-Time	Off
Save original images	On

Angio - Composing

Distortion Corr.	On
Mode	2D
Unfiltered images	Off

Sequence - Part 1

Dimension	3D
Elliptical scanning	Off
Asymmetric echo	Strong
Optimization	Min. TE TR
Bandwidth	810 Hz/Px

Sequence - Part 2

RF pulse type	Normal
Gradient mode	Fast
Excitation	Slab-sel.
RF spoiling	On
Phase Enc. Rewinder	On

Sequence - Assistant

Mode	Min flip angle
Min flip angle	18 deg
Allowed delay	0 s

C.5 T1-weighted VIBE Sequence for viewing the LC thermometer, 1.5T Siemens Sola

SIEMENS MAGNETOM 1.5T XQ Numaris/X VA31A-012Y

\ZZ_DWI_QIBA\DWI_QIBA\QA_MSC\QA_200622\QIBA_qa_vibe_cor_iso	
TA: 2:43 min Coil Selection: Auto Voxel Size: 1.0x1.0x1.0 mm ³ Acc:: 2 Rel. SNR: 1.00	
Properties	
Start measurement without further preparation	On
Wait for User to Start	Off
Start measurements	Single Measurement
Prio Recon	Off
Auto Open Inline Display	Off
Auto Close Inline Display	Off
Load Images to MR View&GO	On
Auto Store Images	On
Load Images to Stamp Segments	On
Load Images to Graphic Segments	Off
Graphic segment	Default
Inline Movie	Off
Routine	
Slab Group	1
Slabs	1
Distance Factor	20 %
Position	R1.4 P18.6 H0.7 mm
Orientation	C > T-1.4
Phase Encoding Dir.	F >> H
Slices per Slab	192
Phase Oversampling	0 %
Slice Oversampling	0.0 %
FoV Read	250 mm
FoV Phase	100.0 %
Slice Thickness	1.0 mm
TR	6.0 ms
TE	2.38 ms
Averages	1
Concatenations	1
AutoAlign	Head > Orbits
Contrast - Common	
TR	6.0 ms
TE	2.38 ms
Flip Angle	10 deg
Fat-Water Contrast	Standard
Contrasts	1
Reconstruction	Magnitude
Contrast - Dynamic	
Dynamic Mode	Standard
Temporal Interpolation	1
Measurements	1
Multiple Series	Off
3D Reordering	Standard
Time to Center	82.6 s
Burn Time to Center	Off
Resolution - Common	
FoV Read	250 mm
FoV Phase	100.0 %
Slice Thickness	1.0 mm
Base Resolution	256
Phase Resolution	100 %
Slice Resolution	100 %
Trajectory	Cartesian
Interpolation	Off
Resolution - Acceleration	
Acceleration mode	GRAPPA
Total Factor	2
Reference Scans	Integrated
Acceleration Factor PE	2
Reference Lines PE	24
Acceleration Factor 3D	1
Phase Partial Fourier	Off
Slice Partial Fourier	Off
Asymmetric Echo	Weak
Elliptical Scanning	Off
Resolution - Filter	
Raw Filter	Off
Elliptical Filter	On
POCS	Off
Distortion Correction	2D
Normalize	Prescan
Noise Masking	Off
Image Filter	Off
Geometry - Common	
Slab Group	1
Slabs	1
Distance Factor	20 %
Position	R1.4 P18.6 H0.7 mm
Orientation	C > T-1.4
Phase Encoding Dir.	F >> H
Slices per Slab	192
Phase Oversampling	0 %
Slice Oversampling	0.0 %
FoV Read	250 mm
FoV Phase	100.0 %
Slice Thickness	1.0 mm
TR	6.0 ms
Series	Ascending
Concatenations	1
Geometry - AutoAlign	
Slab Group	1
Position	R1.4 P18.6 H0.7 mm
Orientation	C > T-1.4
Phase Encoding Dir.	F >> H
AutoAlign	Head > Orbits
Initial Position	R1.4 P18.6 H0.7
R	1.4 mm
P	18.6 mm
H	0.7 mm
Initial Orientation	C > T
C > T	-1.40
> S	0.00
Initial Rotation	90.00 deg
Geometry - Saturation	
Special Saturation	None
Geometry - Tim Planning Suite	
Set-n-Go Protocol	Off
Table Position	0 mm
Table Position	H
Inline Composing	Off

QA for Assessing Quantitative Biomarkers in Patients using MRI

SIEMENS MAGNETOM 1.5T XQ Numaris/X VA31A-012Y

System - Miscellaneous

Coil Selection	Auto Coil Select
MSMA	S - C - T
Sagittal	R >> L
Coronal	A >> P
Transversal	F >> H
Coil Combination	Adaptive Combine
Matrix Optimization	Off
Coil Focus	Flat

System - Adjustments

Adjustment Strategy	Standard
B0 Shim	Standard
CoilShim	Off
Adjustment Tolerance	Auto
Adjust with Body Coil	Off
Confirm Frequency	Never
Assume Silicone	Off

System - Adjust Volume

Position	R1.4 P18.6 H0.7 mm
Orientation	C > T-1.4
Rotation	90.00 deg
F >> H	250 mm
R >> L	250 mm
A >> P	192 mm
Reset	Off

System - Tx/Rx

Frequency 1H	63.680140 MHz
? Ref. Amplitude 1H	0.000 V
Reset	Off
Image Scaling	1.000

Physio - PACE

Resp. Control	Off
Concatenations	1

Inline - Liver

Liver Registration	Off
Save Original Images	On

Inline - Subtraction

Subtract	Off
Measurements	1
StdDev	Off
Save Original Images	On

Inline - Cardiac

Save Original Images	On
Contrasts	1
TE	2.38 ms
TR	6.0 ms

Inline - MIP

MIP Sag	Off
MIP Cor	Off
MIP Tra	Off
MIP Time	Off
Radial MIP	Off
Save Original Images	On
MPR Sag	Off
MPR Cor	Off
MPR Tra	Off

Inline - Soft Tissue

Wash-in	Off
Wash-out	Off
TTP	Off
PEI	Off
MIP Time	Off
Measurements	1

Inline - Composing

Inline Composing	Off
------------------	-----

Sequence - Part 1

Sequence Name	fl
Dimension	3D
Excitation	Slab-sel.
RF Pulse Type	Normal
Gradient Mode	Normal
Bandwidth	810 Hz/Px
Asymmetric Echo	Weak
Optimization	None

Sequence - Part 2

Introduction	On
RF Spoiling	On
Incr. Gradient Spoiling	On
Breast Application	Off

Sequence - Assistant

SAR Assistant	Off
Optimization	None

APPENDIX D: SEQUENCE REQUIREMENTS FOR ANALYSIS IN QCAL-MR*Table D1. Required (essential) and expected (preferable) parameters for the T1-weighted images. From [79].*

Requirements
<p>Across all series, there shall be a single value for field of view mm. Across all series, there shall be a single value for pixel bandwidth. Across all series, there shall be a single value for repetition time. Across all series, there shall be a single value for b value. Across all series, there shall be a single value for echo time. Across all series, there shall be a single value for inversion time. Fat suppression shall be False. Repetition time shall be at most 25. Slice thickness shall be at most 5. Across all series, there shall be a single value for flip angle.</p>
Expected
<p>Magnetic field strength should be any of (1.5, 3). B value should be 0. Flip angle should be between 10 and 20. Repetition time should be between 4 and 10. Slice thickness should be at most 1. Echo time should be at most 4. Acquisition matrix freq dim should be at least 256. Acquisition matrix phase dim should be at least 256. Pixel bandwidth should be at least 400.</p>

Table D2. Required (essential) and expected (preferable) parameters for the DWI images. From [79]

Requirements
<p>Across all series, there shall be a single value for field of view mm. Across all series, there shall be a single value for pixel bandwidth. Across all series, there shall be a single value for repetition time. Across all series, there shall be a single value for flip angle. Across all series, there shall be a single value for echo time. Across all series, there shall be a single value for fat suppression. Repetition time shall be at least 2000. Philips series reconstruction number shall be 1. Each value for b value must occur the same number of times. Across all series, there shall be at least 2 distinct values for b value. Number of frames (series * temporal positions) shall be at least 2.</p>
Expected
<p>Magnetic field strength should be any of (1.5, 3). Flip angle should be 90. Fat suppression should be False. B value should be 0, 500, 1000, 1500, 2000. Number of frames (series * temporal positions) should be at least 4. DWI Images should be trace DWI images or be a complete set of directional DWI images that can be used to produce an anisotropic image. All series should be in immediate succession. Repetition time should be between 7000 and 10000. Echo time should be between 50 and 150. Number of averages should be 2. Pixel bandwidth should be between 1000 and 2500. Acquisition matrix freq dim should be between 128 and 160. Acquisition matrix phase dim should be between 128 and 160. Slice thickness should be between 3.8 and 4.2. Spacing between slices mm should be between 4 and 6. In plane phase encoding direction should be any of ('COL', 'COLUMN'). Parallel imaging factor should be between 1.9 and 2.1.</p>

APPENDIX E: ROOM TEMPERATURE AND ICE BATH QA RESULTS FOR SIEMENS PRISMA 3T

Table E1. 3T Siemens Prisma room temperature (before ice) results from four immediate DWI repeats in coronal orientation. Phantom temperature 15.6 +/- 0.9 °C. Green values are within QIBA tolerances.

Vial	Oc	Oi	Oo	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	1782	1782	1782	1362	1362	1035	1035	713	713	435	435	228	228
Mean ADC (µm²/s)	1752.6	1752.0	1847.7	1346.6	1363.0	1003.3	1009.7	697.1	713.2	446.6	453.8	247.8	233.7
Bias (µm²/s)	-29.15	-29.73	65.99	-15.39	0.97	-31.98	-25.56	-16.39	-0.28	12.11	19.26	19.87	5.70
Bias (%)	-1.64	-1.67	3.70	-1.13	0.0716	-3.09	-2.47	-2.30	-0.0398	2.79	4.43	8.71	2.50
Max b-value dependence	0.0983	0.170	7.74	0.303	2.42	0.656	0.213	2.21	0.985	3.95	0.949	3.10	3.46
Max b-value pair	(1000, 2000)	(1000, 1500)	(500, 2000)	(1000, 2000)	(500, 2000)	(500, 1500)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 1000)	(1000, 2000)	(500, 2000)
SNR B0	157.5	209.5	235.6	250.7	296.7	224.7	321.9	178.6	318.7	160.2	250.8	127.0	150.1
SNR B2000	18.4	16.9	12.0	41.5	39.8	84.1	95.8	137.1	154.5	149.9	193.5	155.1	167.0
RC _{ST} (µm²/s)	4.46	6.51	5.74	6.18	4.97	3.54	5.25	3.70	3.54	2.73	1.51	2.71	1.87
Random error (%)	1.28	1.40	1.97	0.760	0.780	0.526	0.470	0.507	0.429	0.765	0.584	1.49	1.62
wCV _{ST} (%)	0.092	0.13	0.11	0.17	0.13	0.13	0.19	0.19	0.18	0.22	0.12	0.40	0.29

Table E2. 3T Siemens Prisma ice bath results from four immediate DWI repeats in coronal orientation. Green values are within QIBA tolerances.

Vial	Oc	Oi	Oo	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	1109	1109	1109	817	817	579	579	380	380	220	220	110	110
Mean ADC (µm²/s)	1110.4	1109.0	1144.1	830.7	839.7	590.9	593.9	391.1	413.0	246.1	257.6	137.0	110.1
Bias (µm²/s)	1.44	0.01	35.07	13.68	22.67	11.89	14.90	11.09	32.96	26.06	37.63	26.98	0.09
Bias (%)	0.130	0.000856	3.16	1.67	2.77	2.05	2.57	2.92	8.67	11.8	17.1	24.5	0.0844
Max b-value dependence	0.196	0.367	5.15	0.740	2.89	0.175	0.312	4.24	1.50	8.48	7.05	15.6	9.98
Max b-value pair	(1000, 1500)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(1500, 2000)	(500, 1500)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)
SNR B0	259.9	277.5	194.6	225.8	183.0	212.3	204.0	166.8	205.7	148.7	169.4	74.5	75.6
SNR B2000	64.4	59.3	42.1	90.7	90.0	155.7	177.0	189.5	210.4	190.2	179.6	130.2	123.2
RC _{ST} (µm²/s)	1.07	0.891	6.77	0.502	1.22	0.152	0.198	0.482	0.606	0.634	0.243	2.60	0.727
Random error (%)	0.592	0.637	0.936	0.593	0.603	0.534	0.526	0.875	0.619	1.54	1.15	4.70	5.71
wCV _{ST} (%)	0.035	0.029	0.21	0.022	0.052	0.0093	0.012	0.044	0.053	0.093	0.034	0.65	0.23

Table E3. 3T Siemens Prisma room temperature (after ice) results from four immediate DWI repeats in coronal orientation. Phantom temperature 15.6 +/- 0.9 °C. Green values are within QIBA tolerances.

Vial	Oc	Oi	Oo	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	1782	1782	1782	1362	1362	1035	1035	713	713	435	435	228	228
Mean ADC (µm²/s)	1783.2	1780.9	1874.2	1370.2	1382.6	1021.7	1027.6	710.8	727.1	504.4	462.8	249.5	236.7
Bias (µm²/s)	1.43	-0.89	92.43	8.16	20.54	-13.58	-7.74	-2.62	13.61	69.86	28.26	21.53	8.69
Bias (%)	0.0803	-0.0498	5.19	0.599	1.51	-1.31	-0.748	-0.367	1.91	16.1	6.50	9.44	3.81
Max b-value dependence	0.277	0.204	7.40	0.226	1.97	0.457	0.616	2.98	1.10	3.02	1.07	2.87	1.17
Max b-value pair	(500, 2000)	(1000, 1500)	(500, 2000)	(1000, 2000)	(1000, 2000)	(500, 1500)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 1000)	(500, 1000)	(500, 2000)
SNR B0	150.2	174.7	155.2	220.1	216.9	174.4	291.4	218.3	351.8	123.8	256.6	124.5	126.8
SNR B2000	16.4	15.5	12.2	41.4	39.7	76.6	95.2	136.7	167.2	127.9	194.2	161.2	146.3
RC _{ST} (µm²/s)	0.254	3.85	3.28	3.05	3.31	3.57	1.55	1.54	1.21	8.95	1.49	2.13	2.01
Random error (%)	1.41	1.50	1.93	0.747	0.768	0.552	0.438	0.468	0.376	1.52	0.565	1.60	1.71
wCV _{ST} (%)	0.0052	0.078	0.063	0.080	0.086	0.13	0.054	0.078	0.060	0.58	0.12	0.31	0.31

APPENDIX F: ROOM TEMPERATURE AND ICE BATH QA RESULTS FOR SIEMENS SOLA 1.5T

Table F1. 1.5T Siemens Sola room temperature (before ice) results from four immediate DWI repeats in coronal orientation. Phantom temperature 20.7 +/- 0.9 °C. Green values are within QIBA tolerances.

Vial	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	2037	2037	2037	1582	1582	1209	1209	851	851	520	520	276	276
Mean ADC (µm²/s)	2001.6	2000.9	1941.2	1558.2	1530.9	1188.0	1208.9	849.7	842.3	546.7	554.2	291.0	292.9
Bias (µm²/s)	-35.48	-36.17	-95.88	-23.95	-51.27	-21.28	-0.37	-1.21	-8.57	26.44	33.92	14.94	16.78
Bias (%)	-1.74	-1.78	-4.71	-1.51	-3.24	-1.76	-0.0306	-0.142	-1.01	5.08	6.52	5.41	6.08
Max b-value dependence	2.14	2.41	2.98	1.21	0.739	1.07	1.40	1.62	0.915	2.63	5.24	110	22.7
Max b-value pair	(500, 1500)	(500, 1500)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 1000)	(500, 2000)	(500, 2000)	(500, 2000)
SNR B0	133.4	127.7	110.5	119.6	85.3	114.4	133.8	122.6	75.4	96.2	70.7	73.0	77.4
SNR B2000	5.7	5.9	7.0	11.7	12.9	25.5	33.0	48.8	54.7	67.9	61.5	82.3	92.3
RC _{ST} (µm²/s)	5.75	3.79	13.8	4.36	4.92	2.35	1.16	0.874	3.39	1.76	3.71	2.08	1.15
Random error (%)	3.68	3.56	3.06	2.27	2.09	1.43	1.05	1.07	0.860	1.30	1.21	2.64	2.30
wCV _{ST} (%)	0.10	0.068	0.26	0.10	0.12	0.071	0.035	0.037	0.15	0.12	0.24	0.26	0.14

Table F2. 1.5T Siemens Sola ice bath results from four immediate DWI repeats in coronal orientation. Green values are within QIBA tolerances.

Vial	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	1109	1109	1109	817	817	579	579	380	380	220	220	110	110
Mean ADC (µm²/s)	1123.0	1122.0	1098.9	833.6	821.0	606.6	616.0	414.0	412.2	254.2	257.9	114.9	106.5
Bias (µm²/s)	13.96	13.01	-10.12	16.56	3.97	27.63	37.03	34.05	32.25	34.22	37.87	4.92	-3.48
Bias (%)	1.26	1.17	-0.913	2.03	0.486	4.77	6.40	8.96	8.49	15.6	17.2	4.47	-3.16
Max b-value dependence	1.50	2.31	0.080	1.16	1.10	4.33	2.49	5.21	1.73	5.39	5.23	205	38.0
Max b-value pair	(500, 2000)	(500, 1500)	(1000, 1500)	(500, 2000)	(500, 1500)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 1000)	(500, 1000)	(500, 2000)	(500, 2000)	(500, 2000)
SNR B0	144.6	125.2	120.9	120.0	106.8	111.9	130.7	111.6	62.1	80.4	57.8	39.5	40.7
SNR B2000	30.3	32.2	31.5	49.2	52.7	73.8	88.4	96.4	71.4	87.2	68.9	59.7	68.2
RC _{ST} (µm²/s)	1.21	2.71	1.46	1.31	1.50	0.574	0.876	1.56	3.87	1.16	4.14	3.27	3.47
Random error (%)	1.25	1.16	1.21	1.07	1.04	1.11	0.907	1.33	1.08	2.39	2.40	11.1	11.3
wCV _{ST} (%)	0.039	0.087	0.048	0.057	0.066	0.034	0.051	0.14	0.34	0.17	0.58	1.0	1.2

Table F3. 1.5T Siemens Sola room temperature (after ice) results from four immediate DWI repeats in coronal orientation. Phantom temperature 20.7 +/- 0.9 °C. Green values are within QIBA tolerances.

Vial	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	2037	2037	2037	1582	1582	1209	1209	851	851	520	520	276	276
Mean ADC (µm²/s)	2005.9	1994.2	1940.8	1559.7	1533.7	1189.8	1212.7	855.1	843.1	548.6	550.9	296.0	294.8
Bias (µm²/s)	-31.21	-42.89	-96.32	-22.50	-48.40	-19.40	3.47	4.19	-7.85	28.26	30.56	19.89	18.70
Bias (%)	-1.53	-2.11	-4.73	-1.42	-3.06	-1.60	0.287	0.492	-0.922	5.43	5.87	7.20	6.77
Max b-value dependence	1.69	2.48	3.03	0.957	0.773	1.99	1.58	2.75	0.835	1.48	5.93	70.1	21.8
Max b-value pair	(1500, 2000)	(1500, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 1000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 1000)	(500, 2000)	(500, 1500)	(500, 2000)
SNR B0	125.3	121.6	82.9	117.4	81.2	116.1	122.4	108.0	55.6	78.3	54.2	71.4	78.7
SNR B2000	5.8	6.1	7.0	11.8	13.6	25.5	32.3	47.2	45.9	62.7	51.1	77.1	88.4
RC _{ST} (µm²/s)	11.0	13.0	4.78	8.55	8.27	1.15	3.56	1.55	2.75	1.73	2.55	3.84	2.92
Random error (%)	3.67	3.49	3.04	2.25	2.02	1.39	1.05	1.06	0.866	1.28	1.19	2.64	2.32
wCV _{ST} (%)	0.20	0.24	0.089	0.20	0.19	0.035	0.11	0.065	0.12	0.11	0.17	0.47	0.36

APPENDIX G: SNR VARIATION WITH PHANTOM VIAL FOR ICE BATH SCANS

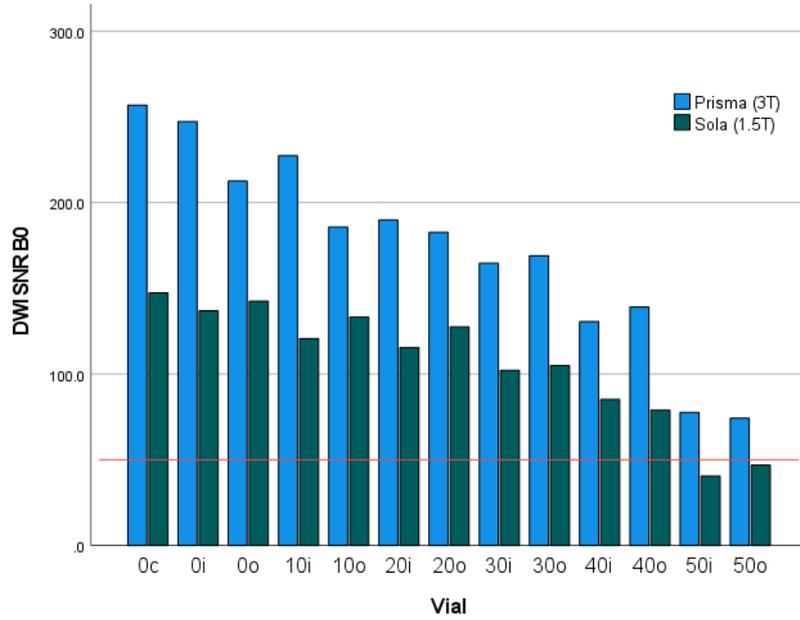


Figure G1. Average SNR of all orientations at b_0 . Both scans are at 0°C . The red line indicates QIBA tolerance for b_0 SNR.

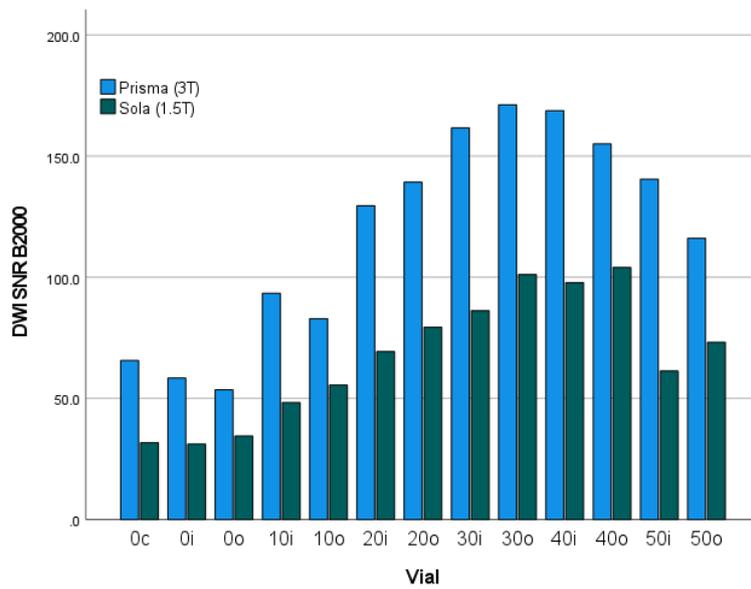


Figure G2. Average SNR of all orientations at b_{2000} . Both scans are at 0°C .

APPENDIX H: LONG-TERM REPEATABILITY OF ROOM TEMPERATURE ADC MEASUREMENTS

Table H1. Long-term repeatability metrics for the 3T scanner. Phantom in coronal orientation. Phantom temperature 15.6 +/- 0.9 C.

Vial	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
%wCV_{LT}	1.22	1.15	1.00	1.23	1.01	1.28	1.24	1.38	1.36	8.59	1.39	0.47	0.90
RC_{LT}	59.9	56.5	51.8	46.1	38.3	36.0	34.9	27.0	27.2	113.1	17.6	3.3	5.8

Table H2. Long-term repeatability metrics for the 1.5T scanner. Phantom in coronal orientation. Phantom temperature 15.6 +/- 0.9 C.

Vial	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
%wCV_{LT}	0.15	0.24	0.02	0.07	0.13	0.11	0.22	0.45	0.06	0.24	0.43	1.19	0.46
RC_{LT}	8.4	13.2	0.9	2.8	5.6	3.7	7.5	10.6	1.4	3.6	6.6	9.7	3.8

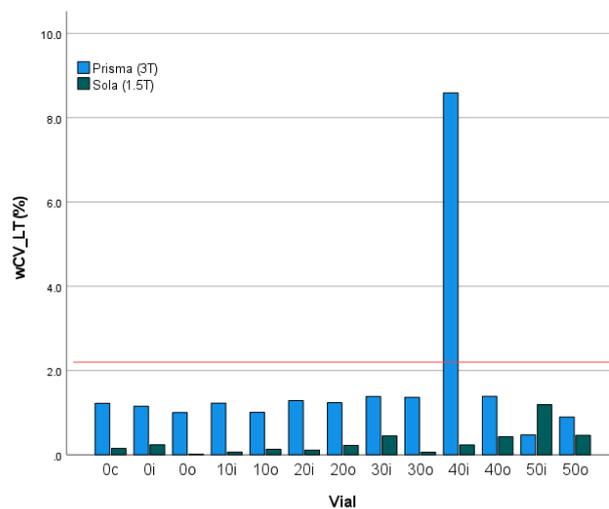
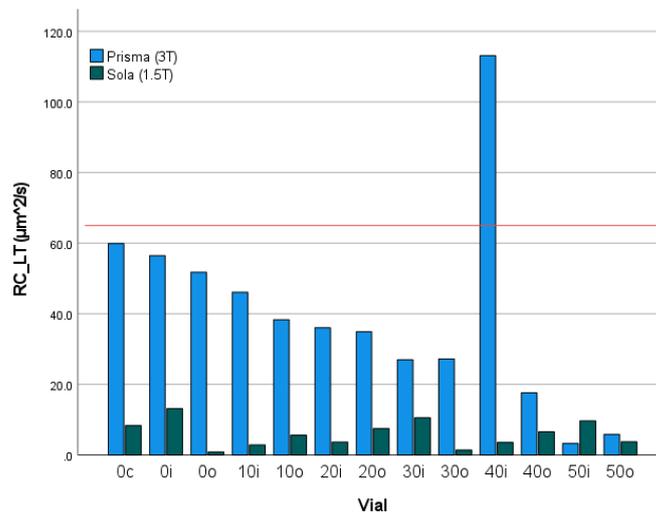


Figure H1. RC_{LT} (top) and wCV_{LT} (bottom) in coronal orientation between the two room temperature scans on each scanner. Red lines indicate QIBA tolerances.

APPENDIX I: REPRODUCIBILITY OF ADC MEASUREMENTS BETWEEN THE TWO SCANNERS AT 0°C

Table II. RDC of ADC measurements at 0°C between the two scanners. Phantom in coronal orientation.

VIAL	0c	0i	0o	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
RDC ($\mu\text{m}^2/\text{s}$)	24.5	25.5	88.5	5.66	36.6	30.8	43.4	45.0	1.40	16.0	0.472	43.2	7.00

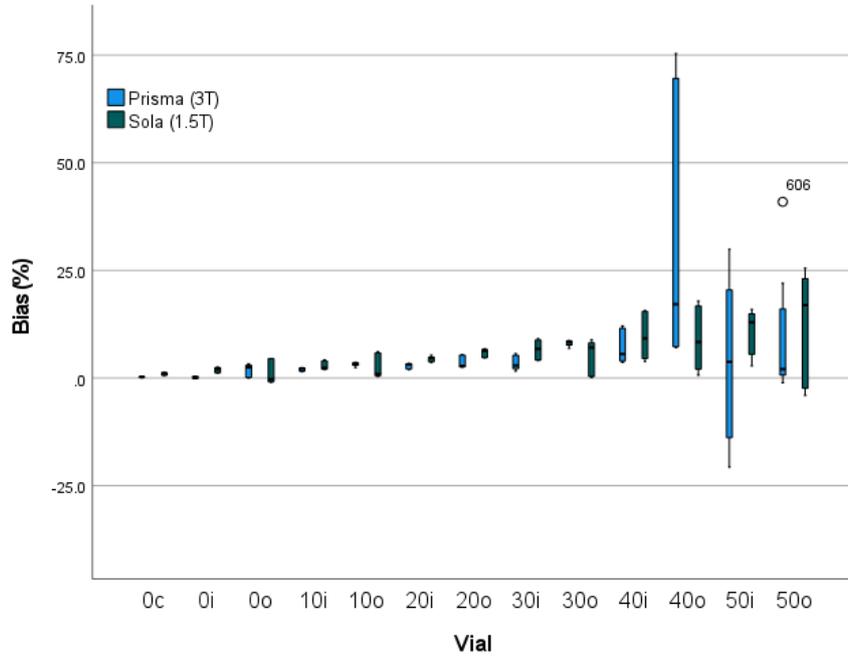


Figure II. % ADC bias of the Siemens Prisma (3T) and Siemens Sola (1.5T) measured at 0°C.

APPENDIX J: QA RESULTS FOR ALL VIALS USING THREE HEAD COILS

Table J1. 20-channel head and neck coil results from four immediate DWI repeats in coronal orientation (19.6 +/- 0.8°C). Green values are within QIBA tolerances.

Vial	Oc	Oi	Oo	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	1980	1980	1980	1533	1533	1169	1169	818	818	500	500	263	263
Mean ADC (µm²/s)	1969.0	1964.2	2009.4	1523.8	1521.2	1145.8	1152.7	810.5	820.1	526.3	532.2	300.2	285.2
Bias (µm²/s)	-10.60	-15.38	29.80	-9.40	-12.03	-22.99	-16.11	-7.11	2.45	26.45	32.41	37.37	22.40
Bias (%)	-0.536	-0.777	1.51	-0.613	-0.785	-1.97	-1.38	-0.870	0.300	5.29	6.48	14.2	8.52
Max b-value dependence	0.466	0.541	3.53	0.524	2.06	0.618	0.253	0.426	1.54	3.11	0.625	5.04	3.21
Max b-value pair	(500, 2000)	(500, 2000)	(500, 2000)	(1500, 2000)	(500, 2000)	(500, 1500)	(500, 2000)	(500, 1000)	(500, 2000)	(500, 2000)	(1000, 2000)	(500, 2000)	(500, 2000)
SNR B0	167.3	201.8	159.0	224.1	205.2	244.4	295.8	177.2	273.4	178.6	217.7	105.9	110.7
SNR B2000	10.9	10.5	8.9	27.8	26.0	59.8	73.6	96.5	131.1	139.5	174.4	120.6	129.2
RC _{ST} (µm²/s)	2.64	3.20	17.4	1.20	1.75	2.24	0.533	0.764	0.392	0.696	1.19	0.622	1.09
Random error (%)	1.94	2.00	2.36	0.982	1.05	0.634	0.507	0.550	0.422	0.670	0.530	1.52	1.58
wCV _{ST} (%)	0.049	0.059	0.31	0.028	0.042	0.071	0.017	0.034	0.017	0.048	0.081	0.075	0.14

Table J2. 32-channel head coil results from four immediate DWI repeats in coronal orientation (18.6 +/- 0.8°C). Green values are within QIBA tolerances.

Vial	Oc	Oi	Oo	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	1929	1929	1929	1489	1489	1133	1133	782	782	482	482	252	252
Mean ADC (µm²/s)	1928.6	1932.8	1977.0	1496.2	1503.6	1116.2	1121.2	786.7	795.7	502.7	511.0	295.9	276.7
Bias (µm²/s)	-0.02	4.20	48.38	7.52	14.93	-17.10	-12.15	5.05	14.15	20.94	29.24	43.63	24.37
Bias (%)	-0.00109	0.218	2.51	0.505	1.00	-1.51	-1.07	0.646	1.81	4.35	6.07	17.3	9.66
Max b-value dependence	0.352	0.598	3.37	0.388	1.38	0.524	0.574	0.738	0.542	0.971	0.393	5.73	4.65
Max b-value pair	(500, 2000)	(500, 1000)	(500, 2000)	(1000, 2000)	(500, 2000)	(500, 2000)	(500, 2000)	(500, 1000)	(500, 2000)	(500, 2000)	(500, 1000)	(500, 2000)	(500, 2000)
SNR B0	143.0	184.1	213.7	245.2	258.1	201.0	286.5	178.7	382.6	153.1	259.5	103.6	141.4
SNR B2000	13.4	14.5	14.3	38.7	41.9	75.7	92.5	114.0	182.6	156.8	217.7	129.8	178.5
RC _{ST} (µm²/s)	2.47	4.41	4.00	2.98	2.22	1.60	1.35	1.34	0.763	1.58	0.745	2.24	1.60
Random error (%)	1.59	1.47	1.48	0.718	0.655	0.502	0.420	0.475	0.320	0.586	0.457	2.09	1.30
wCV _{ST} (%)	0.046	0.082	0.073	0.072	0.053	0.052	0.043	0.062	0.035	0.11	0.053	0.28	0.21

Table J3. 64-channel head coil results from four immediate DWI repeats in coronal orientation (19.6 +/- 0.8°C). Green values are within QIBA tolerances.

Vial	Oc	Oi	Oo	10i	10o	20i	20o	30i	30o	40i	40o	50i	50o
NIST Value (µm²/s)	1980	1980	1980	1533	1533	1169	1169	818	818	500	500	263	263
Mean ADC (µm²/s)	1952.0	1955.2	2027.3	1513.7	1530.1	1140.1	1146.6	809.9	812.6	527.0	525.8	293.5	292.6
Bias (µm²/s)	-27.59	-24.38	47.72	-19.46	-3.12	-28.68	-22.16	-7.71	-4.97	27.25	25.98	30.67	29.80
Bias (%)	-1.39	-1.23	2.41	-1.27	-0.203	-2.45	-1.90	-0.943	-0.608	5.45	5.20	11.7	11.3
Max b-value dependence	0.495	0.681	5.40	0.300	1.92	0.203	0.399	0.492	0.370	3.71	0.524	5.58	11.1
Max b-value pair	(500, 2000)	(500, 2000)	(500, 2000)	(1000, 2000)	(1000, 2000)	(500, 1000)	(500, 1500)	(1000, 2000)	(500, 1000)	(500, 2000)	(1000, 2000)	(500, 2000)	(500, 2000)
SNR B0	174.2	209.0	176.8	249.8	221.8	255.3	327.0	185.6	270.4	176.3	248.4	118.8	107.0
SNR B2000	9.2	9.9	9.4	28.7	30.1	56.1	68.6	79.6	123.7	128.5	153.5	127.3	119.7
RC _{ST} (µm²/s)	3.93	6.28	8.19	2.93	4.66	2.08	3.02	2.64	2.43	1.80	2.90	2.22	4.87
Random error (%)	2.31	2.13	2.23	0.967	0.900	0.676	0.540	0.717	0.409	0.754	0.582	1.68	2.05
wCV _{ST} (%)	0.073	0.12	0.15	0.070	0.11	0.066	0.095	0.12	0.11	0.13	0.20	0.27	0.59