

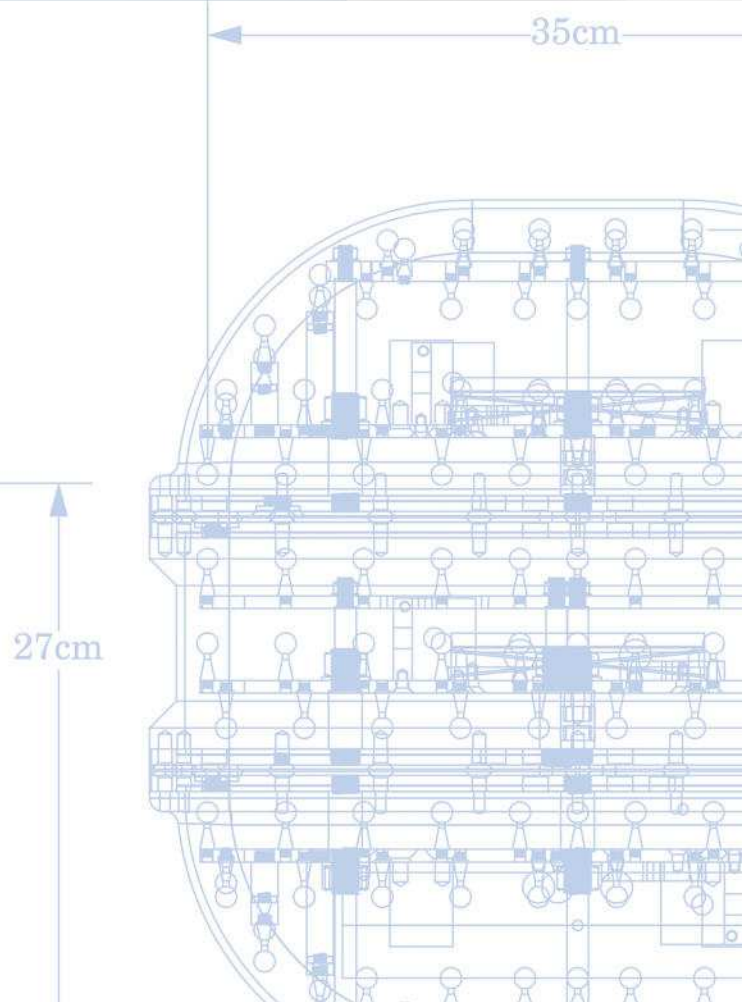
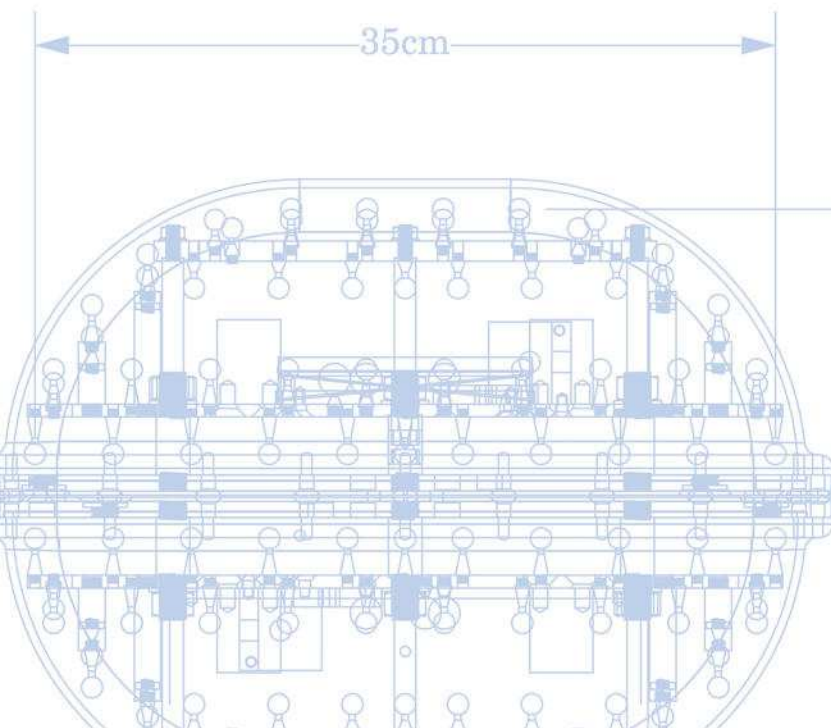
Magphan® RT

MR Imaging QA Designed for Radiotherapy

An integrated phantom and analysis system with a modular, easy-to-handle design

The growing use of MRI for radiation therapy promises many benefits for treatment accuracy and, ultimately, patient outcomes. However, along with these benefits come challenges in accurately characterizing distortion and other critical imaging parameters.

The Phantom Laboratory proudly introduces the Magphan RT system that combines a unique, easy-to-handle phantom design with intuitive and accurate analysis from Image Owl for a tool that will become an indispensable part of your QA suite.



A framework for optimization of diffusion-weighted MRI protocols for large field-of-view abdominal-pelvic imaging in multicenter studies

Jessica M. Winfield^{a)} and David J. Collins

MRI Unit, Royal Marsden Hospital, Downs Road, Sutton, Surrey SM2 5PT, United Kingdom and Cancer Research UK Cancer Imaging Centre, Division of Radiotherapy and Imaging, The Institute of Cancer Research, 123 Old Brompton Road, London SW7 3RP, United Kingdom

Andrew N. Priest

Department of Radiology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, United Kingdom

Rebecca A. Quest and Alan Glover

Imperial College Healthcare NHS Trust, Imaging Department, Hammersmith Hospital, Du Cane Road, London W12 0HS, United Kingdom

Sally Hunter

Department of Radiology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, United Kingdom

Veronica A. Morgan

MRI Unit, Royal Marsden Hospital, Downs Road, Sutton, Surrey SM2 5PT, United Kingdom and Cancer Research UK Cancer Imaging Centre, Division of Radiotherapy and Imaging, The Institute of Cancer Research, 123 Old Brompton Road, London SW7 3RP, United Kingdom

Susan Freeman

Department of Radiology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, United Kingdom

Andrea Rockall

Imperial College Healthcare NHS Trust, Imaging Department, Hammersmith Hospital, Du Cane Road, London W12 0HS, United Kingdom

Nandita M. deSouza

MRI Unit, Royal Marsden Hospital, Downs Road, Sutton, Surrey SM2 5PT, United Kingdom and Cancer Research UK Cancer Imaging Centre, Division of Radiotherapy and Imaging, The Institute of Cancer Research, 123 Old Brompton Road, London SW7 3RP, United Kingdom

(Received 3 August 2015; revised 22 October 2015; accepted for publication 16 November 2015; published 21 December 2015)

Purpose: To develop methods for optimization of diffusion-weighted MRI (DW-MRI) in the abdomen and pelvis on 1.5 T MR scanners from three manufacturers and assess repeatability of apparent diffusion coefficient (ADC) estimates in a temperature-controlled phantom and abdominal and pelvic organs in healthy volunteers.

Methods: Geometric distortion, ghosting, fat suppression, and repeatability and homogeneity of ADC estimates were assessed using phantoms and volunteers. Healthy volunteers (ten per scanner) were each scanned twice on the same scanner. One volunteer traveled to all three institutions in order to provide images for qualitative comparison. The common volunteer was excluded from quantitative analysis of the data from scanners 2 and 3 in order to ensure statistical independence, giving $n = 10$ on scanner 1 and $n = 9$ on scanners 2 and 3 for quantitative analysis. Repeatability and interscanner variation of ADC estimates in kidneys, liver, spleen, and uterus were assessed using within-patient coefficient of variation (wCV) and Kruskal–Wallis tests, respectively.

Results: The coefficient of variation of ADC estimates in the temperature-controlled phantom was 1%–4% for all scanners. Images of healthy volunteers from all scanners showed homogeneous fat suppression and no marked ghosting or geometric distortion. The wCV of ADC estimates was 2%–4% for kidneys, 3%–7% for liver, 6%–9% for spleen, and 7%–10% for uterus. ADC estimates in kidneys, spleen, and uterus showed no significant difference between scanners but a significant difference was observed in liver ($p < 0.05$).

Conclusions: DW-MRI protocols can be optimized using simple phantom measurements to produce good quality images in the abdomen and pelvis at 1.5 T with repeatable quantitative measurements in a multicenter study. © 2016 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [<http://dx.doi.org/10.1118/1.4937789>]

Key words: magnetic resonance imaging, diffusion, multicenter studies, optimization, phantoms

1. INTRODUCTION

Diffusion-weighted magnetic resonance imaging (DW-MRI) is well established as a qualitative and quantitative imaging technique in oncologic applications in the body. The high contrast between solid tumors and normal tissues, arising from the restricted diffusion of water molecules in many solid tumors compared with many normal tissues, aids detection of disease and monitoring of response to treatment. Moreover, the most common quantitative parameter, the apparent diffusion coefficient (ADC), has been widely investigated to monitor treatment response, for example, response to chemotherapy in hepatic metastases,^{1,2} breast cancer,^{3,4} advanced ovarian cancer,^{5,6} and nonsmall-cell lung cancer,⁷ and to chemoradiation therapy in cervical cancer⁸ and squamous cell carcinoma of the head and neck.^{9,10} Pretreatment estimates of ADC have also been shown to be predictive of response to chemotherapy in pancreatic cancer¹¹ and to chemoradiation therapy in nonsmall-cell lung cancer.¹²

The ADC is estimated by fitting a monoexponential function to the signal (S) measured using DW-MRI at two or more diffusion weightings (b -values), as described by

$$S(b) = S_0 \exp(-b \cdot \text{ADC}). \quad (1)$$

In the estimation of ADC, the variation in acquisition parameters across scanners from different manufacturers and between models and software versions from the same manufacturer potentially affects quantitation; the effects of some of these variations have been investigated in healthy volunteers. ADC estimates have been shown to depend on the choice of b -values in kidneys¹³ and in liver, spleen, and pancreas.¹⁴ The latter study also showed no significant difference in ADC estimates between orthogonal diffusion encoding and three-scan trace weighted encoding.¹⁴

Consensus recommendations for DW-MRI as a cancer biomarker recommend that protocols should be optimized to maximize signal-to-noise ratio (SNR), minimize artifacts from ghosting and distortion, optimize fat suppression, ensure ADC values can be measured accurately and reproducibly, and, ideally, use parameters which can be replicated on other platforms;¹⁵ this is essential in order to achieve comparable quantitative data. In the first instance, protocol development using specialized phantoms to interrogate the effects of distortion and fat suppression circumvents the ethical and timing constraints of volunteer studies and may be particularly valuable in multicenter projects.

The aims of this study were, therefore, to develop methods for optimization of DW-MRI protocols on multiple platforms using specialist phantoms [polydimethylsiloxane (PDMS), corn oil, and temperature-controlled sucrose] and volunteers, implement the protocols on 1.5 T scanners from different manufacturers, and assess repeatability of the measurements in phantoms and normal tissues in healthy volunteers. A formal comparison between scanners was not a focus of this study.

2. METHODS

We developed the following phantoms and methods for assessment of geometric distortion and ghosting, fat suppression, and repeatability of ADC estimates.

2.A. Geometric distortion and ghosting

A large cylindrical phantom filled with PDMS was used to assess geometric distortion and ghosting.¹⁶ PDMS was used as it has a very low diffusion coefficient at room temperature and therefore allows comparison of geometric properties of images at different diffusion weightings without loss of signal due to diffusion. Subtraction images were calculated by subtracting the $b = 0 \text{ s mm}^{-2}$ image from the corresponding heavily diffusion-weighted image ($b = 900\text{--}1000 \text{ s mm}^{-2}$) in order to assess geometrical differences between the images, assuming no signal decay with diffusion-weighting in the PDMS phantom (MATLAB 2014a, MathWorks, Inc., Natick, MA). The impact of the choice of diffusion gradient scheme, parallel imaging, and receiver bandwidth on the degree of geometric distortion was assessed. A maximum b -value of 1000 s mm^{-2} was used in assessments of the diffusion gradient scheme and parallel imaging in order to clearly demonstrate eddy current effects. Assessment of the effects of bandwidth on geometric distortion and ghosting was carried out using a maximum b -value of 900 s mm^{-2} in order to reflect the final protocol as the optimal bandwidth may be protocol-dependent. A semiquantitative distortion index (DI) was calculated by taking the ratio of the mean absolute pixel value in the subtraction images to the mean pixel value in the background (BG) noise, estimated using a region of interest (ROI) drawn in the corresponding $b = 900 \text{ s mm}^{-2}$ image in a region away from the phantom and ghosts, as described in the previous studies.¹⁷ The strength of the ghosting was quantified using a ghost-to-signal ratio, estimated from the mean pixel value in a ROI in the ghost region (G) to the mean pixel value in a ROI in the phantom (S) in the $b = 0 \text{ s mm}^{-2}$ images.

2.B. Fat suppression

Fat suppression was assessed using images of healthy volunteers and using a corn oil phantom described previously.¹⁸ The phantom was placed on the couch at 45° to the z -axis in order to present an elliptical cross section, which is of similar shape and size to axial abdominal images obtained in human subjects. Images were acquired axially and reformatted sagittally in order to assess fat suppression along the length of the imaging volume.

2.C. Homogeneity of ADC estimates

Large cylindrical phantoms filled with water (manufacturers' water bottle phantoms) were used to assess variation in ADC estimates across a large field-of-view (FOV), assuming the isocenter of the magnet provides the "true" ADC estimate. Images were acquired axially. Homogeneity of ADC estimates in the z -direction was assessed using ROIs drawn in the center

of the image on all slices. Homogeneity of ADC estimates in the right–left and anterior–posterior directions was assessed using line profiles drawn through the center of the central slice. ADC estimates were calculated for each pixel using a least-squares fit to all b -values (trust-region-reflective algorithm, fit to monoexponential curve, MATLAB 2014a, MathWorks, Inc., Natick, MA).

2.D. Repeatability of ADC estimates in an ice-water phantom

An ice-water phantom (Fig. 1) was used to compare ADC estimates between scanners and to monitor stability of the ADC estimates from each scanner over the period of the volunteer study. The design of the phantom was based on the previous studies using ice-water phantoms in DW-MRI (Ref. 19) and has been described in detail elsewhere.²⁰ The phantom consisted of a Perspex cylinder containing five sucrose solutions in polypropylene tubes (0%–20% sucrose, Sigma-Aldrich) surrounded by a mixture of ice and water. The

phantom was allowed to stabilize for 45 min to ensure that samples were at 0 °C before scanning. The protocols used for scanning the ice-water phantom were as used for the healthy volunteers (Table 1) using a smaller FOV (320 mm read FOV) with correspondingly smaller pixels (2.5 × 2.5 mm acquired pixel size) and 5 mm slice thickness. No fat suppression was applied in the images of the ice-water phantom, unless a water-selective excitation was inherent to the diffusion-weighted echo-planar imaging (EPI) sequence. ADC estimates were calculated using in-house software (Levenberg–Marquardt algorithm, least squares fit to monoexponential curve using all b -values, ADEPT, Institute of Cancer Research, London, UK). Phantoms were scanned on multiple occasions during the course of the volunteer studies (scanner 1: $n = 14$, over a period of 22 months; scanner 2: $n = 6$, over 17 months; scanner 3: $n = 8$, over 20 months). After inspection for systematic variation in ADC estimates over time, coefficient of variation (CV = standard deviation/mean) were used to assess repeatability of ADC estimates for each tube on each scanner.

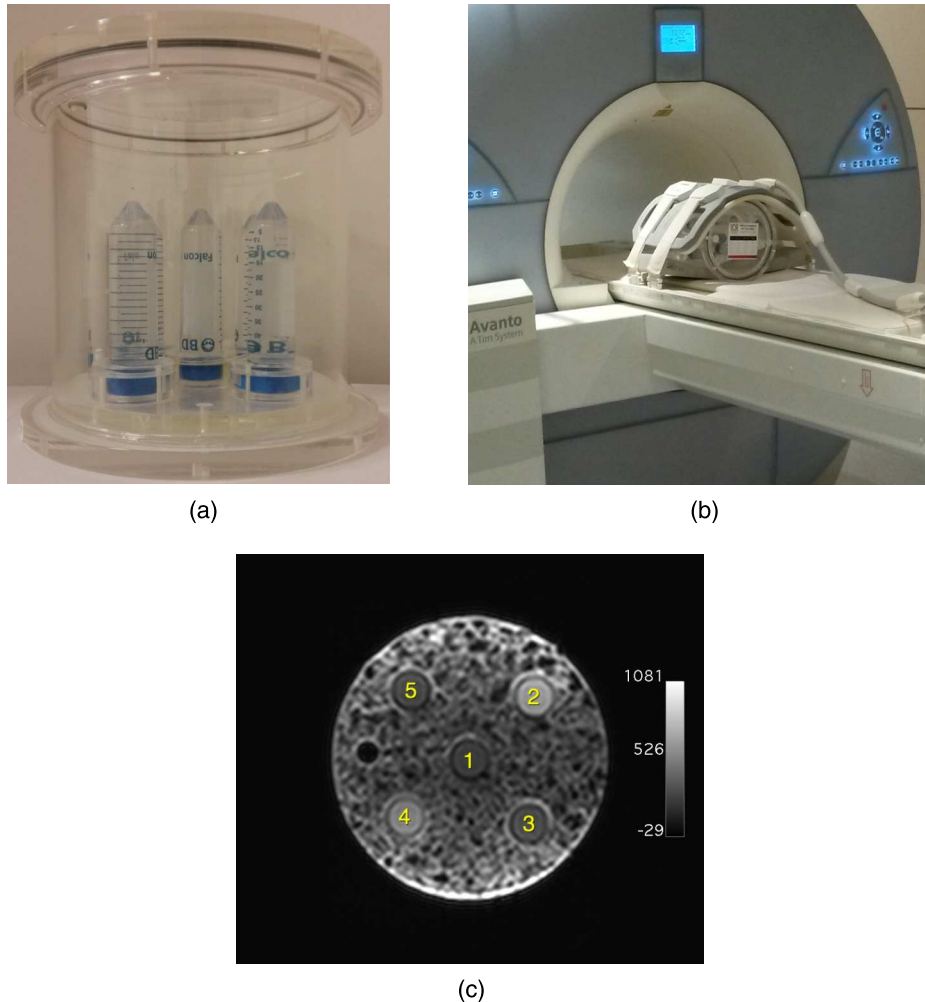


FIG. 1. (a) Ice-water phantom containing five tubes of sucrose solutions. The phantom was filled with a mixture of crushed ice and water and allowed to cool before use. (b) Ice-water phantom positioned on couch using body matrix coil (scanner 1, Avanto, Siemens). (c) Axial image of ice-water phantom showing locations of five tubes of sucrose solutions surrounded by a mixture of ice and water. A Perspex marker rod is also visible on the left-hand side of the figure. Scanner 1; ss-EPI; axial slices; 5 mm slice thickness; FOV 320 × 280 mm; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 76 ms; TR = 8000 ms; parallel imaging reduction factor 2; receive bandwidth 1776 Hz/pixel; bipolar gradients; trace image; $b = 100 \text{ s mm}^{-2}$; 1 NSA.

TABLE I. DW-MRI protocols used to scan healthy volunteers on three scanners. Parameters in bold were standardized across the three scanners.

	Scanner 1	Scanner 2	Scanner 3
Manufacturer	Siemens AG, Healthcare Sector, Erlangen, Germany	GE Healthcare, Waukesha, WI, USA	Philips Healthcare, Best, The Netherlands
Model	MAGNETOM Avanto	Discovery MR450	Intera
Field strength (T)	1.5	1.5	1.5
Maximum gradient amplitude (mT m ⁻¹)	45	50	33
Maximum slew rate (mT m ⁻¹ s ⁻¹)	200	200	160
Receive coil (s)	2×anterior body matrix and posterior spine matrix	12 channel body array	XL torso coil
Patient position	Feet first supine	Feet first supine	Head first supine
Sequence	Single-shot EPI	Single-shot EPI	Single-shot EPI
Slice orientation	Axial	Axial	Axial
FOV (read) (mm)	380	380	380
FOV (phase) (mm)	332	334	333
Slice thickness (mm)	6	6	6
Slices per station	26	26	26
Number of stations	3	3	3
PE direction	AP	AP	AP
Acquired matrix (read)	128	128	128
Reconstructed matrix (read)	256	256	256
Acquired pixel size (mm × mm)	3.0 × 3.0	3.0 × 3.0	3.0 × 3.0
Reconstructed pixel size (mm × mm)	1.5 × 1.5	1.5 × 1.5	1.5 × 1.5
Echo time (TE) (ms)	75	81	96
Δ (ms) ^a	25.9 (effective Δ)	Gradient pulse widths 3.7 and 15.4; ramp time 0.5	47.3
δ (ms) ^a	17.7 (effective δ)		18.3
Repetition time (TR) (ms)	8000	8000	8000
Receive bandwidth ^b	1776 Hz/pixel	Receiver bandwidth ±125 kHz (pixel bandwidth 1953 Hz/pixel)	Water-fat shift 11.1 pixels (bandwidth 19.6 Hz; bandwidth in EPI frequency direction 1583 Hz/pixel)
Number of signal averages (NSA)	4	4	4
Parallel imaging ^c	GRAPPA, reduction factor 2; 36 ACS lines	ASSET, reduction factor 2	SENSE, reduction factor 2
Partial Fourier	No	Yes (16 overscans)	No
Fat suppression	SPAIR	Water-selective excitation	SPAIR, delay time 110 ms
Diffusion gradient scheme ^d	Bipolar	DSE	Monopolar
Number of diffusion directions	3	3	3
Diffusion encoding scheme ^e	3-scan trace	ALL	Gradient overplus
Diffusion-weighted images	Trace	Trace	Trace
Acquired b-values (s mm⁻²)	0, 100, 500, 900	0, 100, 500, 900	0, 100, 500, 900
Breathing instructions	Free-breathing	Free-breathing	Free-breathing
Navigator/gating	None	None	None
Acquisition time per station	5 min 44 s	5 min 28 s	5 min 28 s

^aDiffusion gradient separation (Δ) and duration (δ), or related parameters, are quoted using the terminology used by the three manufacturers.

^bReceive bandwidth is quoted using the terminology and units used by the three manufacturers.

^cGeneralized autocalibrating partially parallel acquisition (GRAPPA); autocalibration signal (ACS); sensitivity encoding (SENSE); array spatial sensitivity encoding technique (ASSET).

^dBipolar and double spin-echo (DSE) are implementations of the twice-refocused spin-echo encoding scheme.

^eThree-scan trace and gradient overplus use three mutually orthogonal diffusion gradient directions, which are not aligned with the cardinal directions of the scanner; ALL uses three gradient directions aligned with axes of magnet. In all cases, images from three diffusion gradient directions were combined to produce trace images.

Phantoms were scanned at three institutions using scanners from three manufacturers (a different manufacturer at each institution, Table I). The PDMS phantom and corn oil phantom were transported to each institution and the same phantom was therefore imaged on each scanner. For the assessment of homogeneity of ADC estimates, the manufacturers' water

bottle phantoms were used; hence, a different phantom was used at each institution. For the assessment of repeatability of ADC estimates, a set of ice-water phantoms were manufactured under identical conditions and tested at the lead site before being issued to each site for repeated measurements.

2.E. Healthy volunteers

Healthy volunteers were scanned twice on the same scanner, with Institutional Review Board Approval and with their written informed consent, on one of the three 1.5 T MR scanners at different institutions using the protocols described in Table I ($n = 10$ on each scanner). DW images were acquired using three contiguous stations covering the abdomen and pelvis. The couch was moved between stations in order to position the center of each station at the magnet isocenter. The bandwidths, parallel imaging, fat suppression methods, diffusion gradient schemes, and number of slices used in the protocols were determined from the results of the phantom experiments. One volunteer traveled to all three institutions in order to provide images for qualitative comparison. As logistical constraints prevented all ten volunteers from traveling to all three institutions, we therefore recruited the remaining nine volunteers as a separate cohort at each site with identical inclusion and exclusion criteria at each institution. The common volunteer was excluded from quantitative analysis of the data from scanners 2 and 3 in order to ensure statistical independence, giving $n = 10$ on scanner 1 and $n = 9$ on scanners 2 and 3 for quantitative analysis. As this study formed part of the work-up for a multicenter trial in ovarian cancer, all volunteers were female (scanner 1: median age 30 yr, range 26–50 yr; scanner 2: median age 35 yr, range 23–58 yr, for the nine volunteers included in quantitative analysis; scanner 3: median age 31 yr, range 26–50 yr, for the nine volunteers included in quantitative analysis). The median interval between scans was 5 days (range 1–8 days). Volunteers were fasted for 4 h before each scan. Premenopausal volunteers were scanned between day 6 and 14 of a menstrual cycle. All DW-MRI scans were carried out in free-breathing as respiratory triggering or navigators were not available on all platforms.

All analyses were carried out at the lead site by a single observer. ROIs were drawn by region growing (kidneys and spleen) or freehand (liver) or by placement of small circular ROIs (uterus) on computed DW images using in-house software.²¹ Computed DW-MRI is a postprocessing step that enables the user to generate DW images at arbitrary b -values. The user may, therefore, generate a DW image at a b -value that was not acquired in the imaging sequence in order to provide improved contrast-to-noise for visualization

of lesions or tissues.²¹ Diffusion weightings for computed DW images were chosen for each organ to give optimal contrast between the organ and surrounding tissues ($b = 500 \text{ s mm}^{-2}$ for kidneys; $b = 800 \text{ s mm}^{-2}$ for liver; $b = 1000 \text{ s mm}^{-2}$ for spleen; $b = 1000 \text{ s mm}^{-2}$ for uterus). For each organ, ROIs were drawn on three contiguous slices and were combined to give a volume of interest (VOI) that contained a large number of pixels. ROIs for the kidneys were drawn around the whole area of the left and right kidneys at the level of the renal hilum (median 4014 pixels in VOI, range 3029–5505). ROIs in the liver were drawn around the whole area of the right lobe of the liver, avoiding motion artifacts near the dome of the liver (median 8958 pixels, range 5804–13 740). ROIs for the spleen were drawn on the slices where the largest area of spleen was visible, avoiding partial volume effects with the kidneys (median 3635 pixels, range 1508–7669). ROIs in the uterus were drawn by placing three circles (32 pixels/circle) in the myometrium on each slice (288 pixels in VOI). Figure 2 shows examples of ROIs in kidneys, liver, spleen, and uterus in images of one volunteer on scanner 1. ADC estimates for each pixel in the ROIs were calculated using in-house software (Levenberg–Marquardt algorithm, least squares fit to monoexponential curve, ADEPT, Institute of Cancer Research, London, UK) using b -values 100, 500, and 900 s mm^{-2} . ($b = 0 \text{ s mm}^{-2}$ was excluded from the analysis in order to reduce the influence of perfusion.^{22–24})

In order to reduce the sensitivity to outlier values, the median ADC estimate from all pixels in the VOI was used for analysis of repeatability. The repeatability of ADC estimates for each organ and each scanner was assessed using the method of Bland and Altman.²⁵ Bland–Altman plots of untransformed data showed a relationship between the differences in the repeated measurements and their means that was improved by using the natural logarithm of the data.²⁵ The within-patient coefficient of variation (wCV) of the log-transformed data was used to describe the repeatability of the ADC estimates ($\text{wCV} = \sqrt{[\exp(\Sigma d^2/2N) - 1]}$, where Σd^2 is the sum of squared differences between pairs of measurements and N is the number of volunteers on each scanner).²⁶ The difference between ADC estimates from three scanners was investigated using a Kruskal–Wallis test for each organ, taking the mean of two visits from each volunteer (MATLAB 2014a, Mathworks, Inc., Natick, MA).

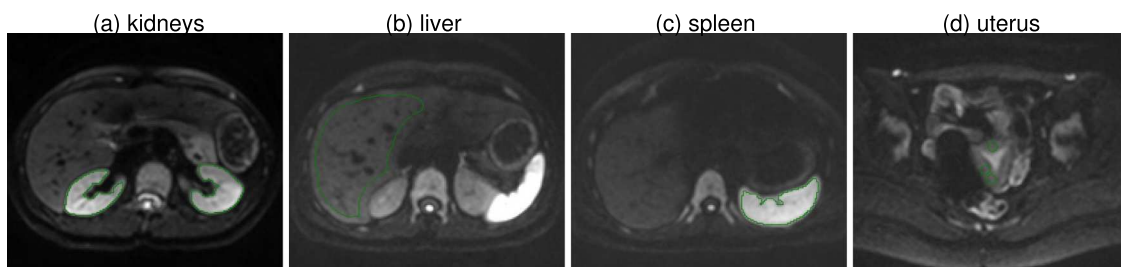


FIG. 2. Images of one volunteer on scanner 1 showing delineations of ROIs in (a) kidneys, shown on computed b -value 500 s mm^{-2} , (b) liver, shown on computed b -value 800 s mm^{-2} , (c) spleen, shown on computed b -value 1000 s mm^{-2} , (d) uterus, shown on computed b -value 1000 s mm^{-2} .

3. RESULTS

Phantoms were imaged on all three scanners. Assessment of distortion, ghosting, fat suppression, and homogeneity is described using images from one selected scanner in each case to illustrate each result.

3.A. Distortion

Distortion, assessed using the PDMS phantom, is greater in the $b = 1000 \text{ s mm}^{-2}$ images than in the $b = 0 \text{ s mm}^{-2}$ images, particularly when acquired using the monopolar sequence compared with the DSE sequence (Fig. 3). This is likely due to eddy current effects. Regions where distortion has occurred between the $b = 0$ and $b = 1000 \text{ s mm}^{-2}$ images appear white or black in the subtraction images. Distortion is more marked in the phase-encode (anterior–posterior) direction, indicating that eddy current effects are more marked in this direction. A parallel imaging artifact may also be seen. The

signal-to-background ratio, estimated from the mean signal in the phantom (S) and BG regions in the $b = 0 \text{ s mm}^{-2}$ images, was higher in the monopolar sequence compared with the DSE sequence (730 and 650, respectively).

When investigating the effects of parallel imaging, there was less distortion evident in the subtraction images using a reduced acquisition (parallel imaging reduction factor $R = 2$) than in images acquired without the use of parallel imaging (Fig. 4). Again, distortion was greater in the anterior–posterior direction, indicating that eddy current effects are more marked in this direction. A parallel imaging artifact may also be seen in the images with reduced acquisition.

In investigating the effect of bandwidth, images acquired using monopolar diffusion gradients, without the use of parallel imaging, were used to clearly illustrate the distortion, although our results show that use of bipolar diffusion gradients and parallel imaging is advantageous in reduction of geometric distortion (Figs. 3 and 4). Distortion visualized at the anterior and posterior edges of the phantom was

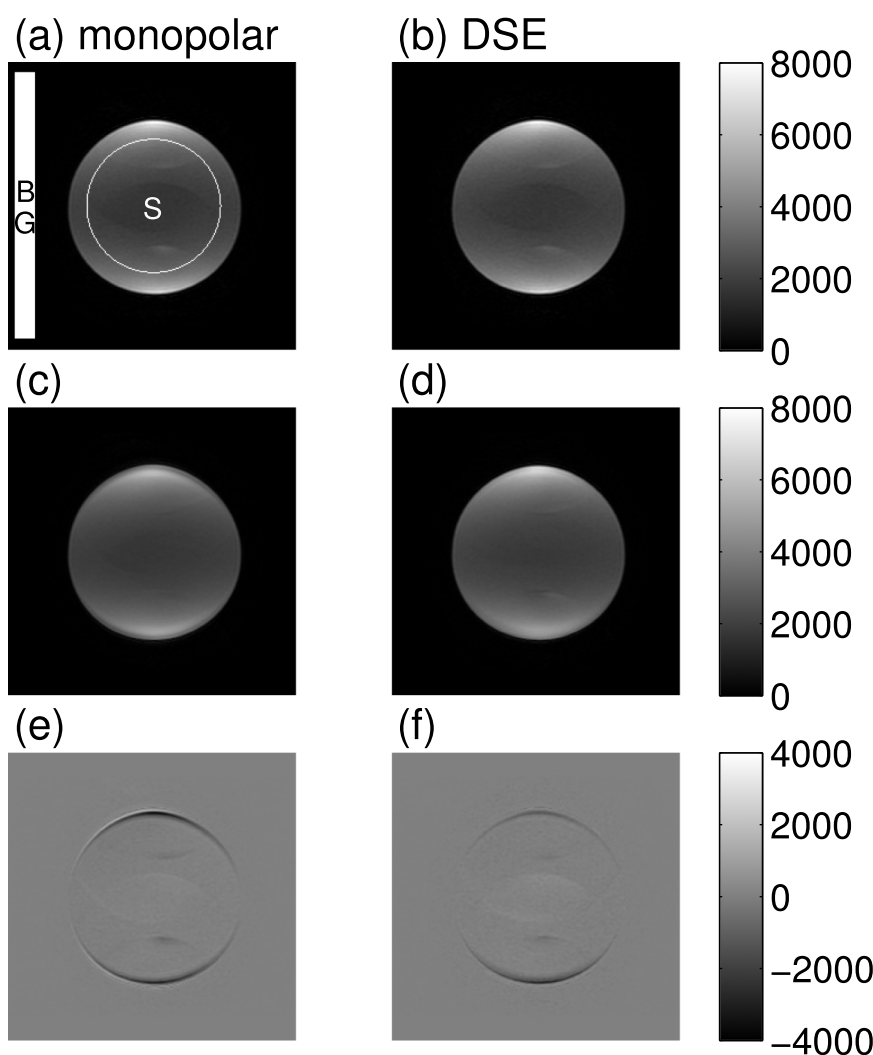


FIG. 3. Axial images of a PDMS phantom: (a) and (b) $b = 0 \text{ s mm}^{-2}$, (c) and (d) $b = 1000 \text{ s mm}^{-2}$, and (e) and (f) subtraction images. Images [(a) and (c)] acquired using a monopolar sequence and [(b) and (d)] acquired using DSE. All other measurement parameters were consistent between the two acquisitions. Scanner 2; ss-EPI; axial slices; 5 mm slice thickness; FOV $320 \times 280 \text{ mm}$; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 84 ms; TR = 8000 ms; parallel imaging reduction factor 2; receiver bandwidth 111 kHz; trace images; 1 NSA. Subtraction images show the $b = 0 \text{ s mm}^{-2}$ image subtracted from the $b = 1000 \text{ s mm}^{-2}$ image. The ROIs used to estimate BG noise and signal (S) in the $b = 0 \text{ s mm}^{-2}$ images are shown in (a).

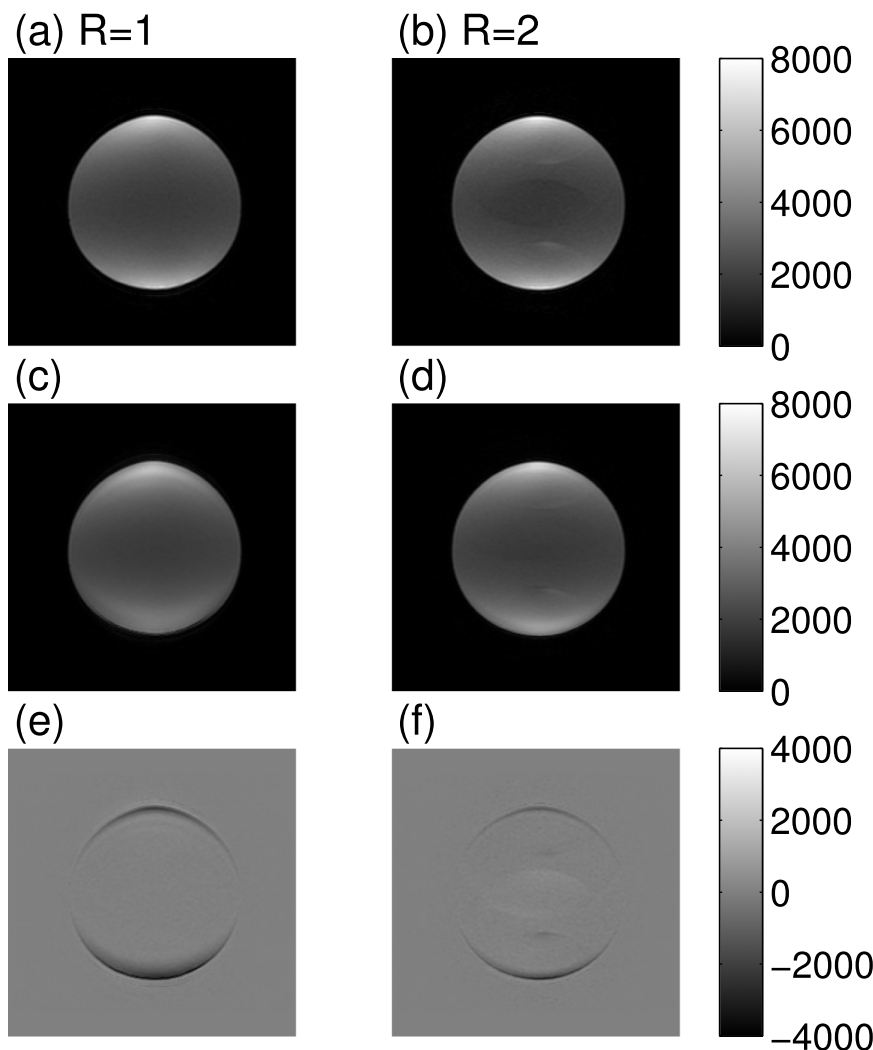


FIG. 4. Axial images of a PDMS phantom: (a) and (b) $b = 0 \text{ s mm}^{-2}$, (c) and (d) $b = 1000 \text{ s mm}^{-2}$, and (e) and (f) subtraction images showing distortion as black or white regions. Images [(a) and (c)] acquired without parallel imaging and [(b) and (d)] using parallel imaging reduction factor $R = 2$. Images (a) and (c) were acquired by setting the reduction factor to 1 in order to use the same reconstruction method as the reduced acquisition. All other measurement parameters were consistent between the two acquisitions. Scanner 2; ss-EPI; axial slices; 5 mm slice thickness; FOV $320 \times 280 \text{ mm}$; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 84 ms; TR = 8000 ms; receiver bandwidth 111 kHz; double spin-echo; trace images; 1 NSA. Subtraction images show the $b = 0 \text{ s mm}^{-2}$ image subtracted from the $b = 1000 \text{ s mm}^{-2}$ image.

more marked at lower bandwidths (Fig. 5). Subtraction images confirmed that distortion was the greatest in the anterior–posterior direction, indicating more marked eddy current effects in this direction. Higher receive bandwidths (lower interecho spacing) resulted in less signal in the subtraction images, indicating a reduction in geometric distortion.

3.B. Ghosting

Visual inspection shows that the ghosts are stronger at higher bandwidths (Fig. 6). The intensity of the ghost as a percentage of the signal (G/S) was (a) 1.0%, (b) 1.4%, (c) 1.7%, (d) 2.4%, (e) 2.3%, (f) 3.7%, and (g) 8.2%; these results are also shown graphically in Fig. 7.

Figure 7 shows DI and ghost-to-signal ratio plotted against receiver bandwidth, measured from the images shown in Figs. 5 and 6. There is a reduction in DI and increase in

ghost-to-signal ratio with increasing bandwidth, as shown qualitatively in Figs. 5 and 6. The small increase in DI at 3005 Hz/pixel may be attributable to the large increase in noise and ghosting at the highest bandwidth. Bandwidths in the range 1500–2000 Hz/pixel represent an appropriate trade-off between geometric distortion and ghosting in this case.

3.C. Fat suppression

Good fat suppression along the full length of the imaging volume (15.6 cm in z -direction) was achieved using SPAIR, illustrated in a phantom and in a healthy volunteer [Figs. 8(a) and 8(b)]. On some scanners, further optimization is required to select the SPAIR delay time, which may be protocol dependent. Using SPAIR at a range of delay times indicated an optimal SPAIR delay of 110 ms for this protocol on scanner 3 [Figs. 8(c)–8(e)]. The SPAIR delay was a user-defined parameter only on scanner 3.

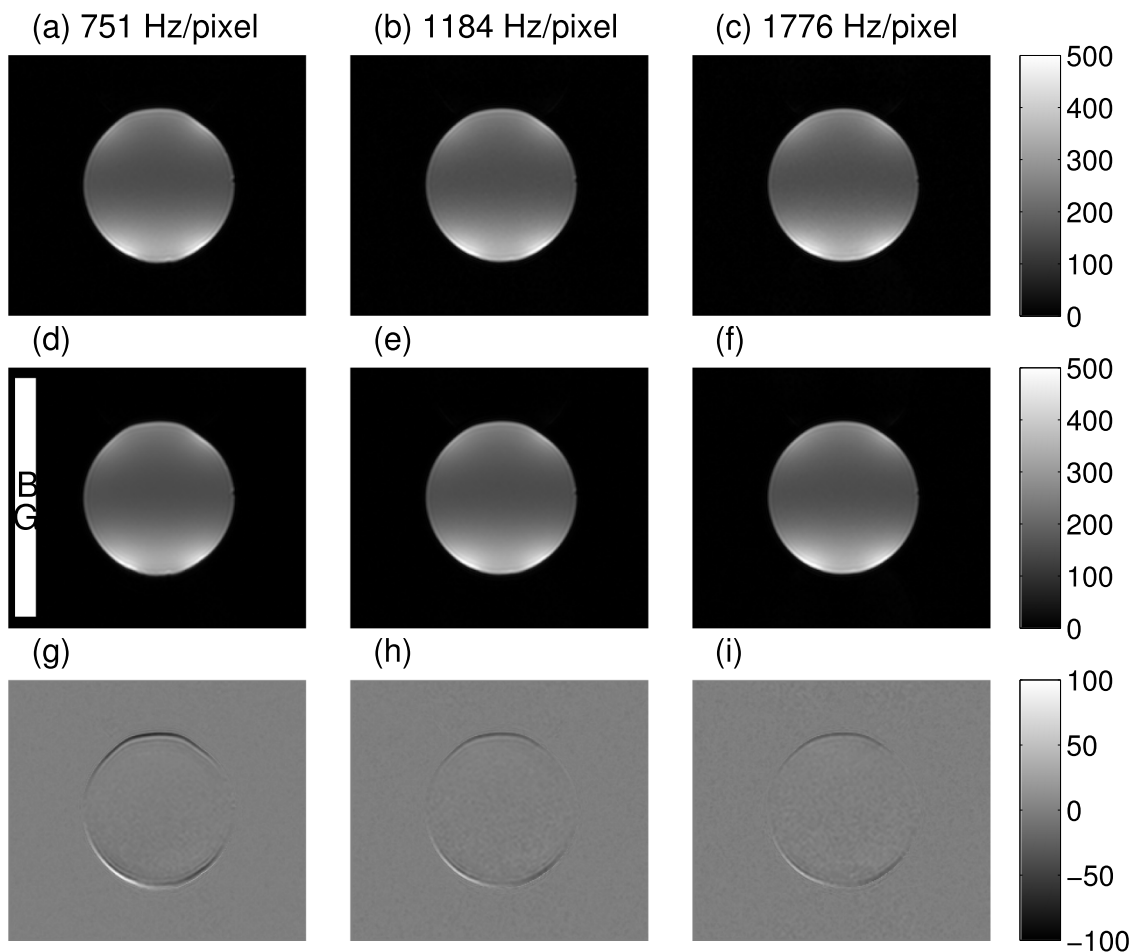


FIG. 5. Axial images of a PDMS phantom: (a)–(c) $b = 0 \text{ s mm}^{-2}$, (d)–(f) $b = 900 \text{ s mm}^{-2}$, and (g)–(i) subtraction images. Images were acquired at different receive bandwidths [(a), (d), and (g)] 751 Hz/pixel, [(b), (e), and (h)] 1184 Hz/pixel, and [(c), (f), and (i)] 1776 Hz/pixel. All other measurement parameters were consistent between the acquisitions. Scanner 1; ss-EPI; axial slices; 6 mm slice thickness; FOV $380 \times 332 \text{ mm}$; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 178 ms; TR = 8000 ms; no parallel imaging; monopolar diffusion gradients; trace images; 1 NSA. Subtraction images show the $b = 0 \text{ s mm}^{-2}$ image subtracted from the $b = 900 \text{ s mm}^{-2}$ image. The ROI used to estimate BG noise in each $b = 900 \text{ s mm}^{-2}$ image is shown in (d).

3.D. Homogeneity of ADC estimates

ADC estimates from square ROIs (81 pixels) drawn at the center of each slice of a stack of axial slices through a water bottle phantom showed no noticeable deviation in median ADC estimates between the central slice and end slices (<5%) in a sequence with 26 slices [Fig. 9(a)]. The effects of increasing the number of slices are shown in Fig. 9(b), which demonstrates a reduction in ADC estimates (~10%) in the end slices compared with the central slice. Except for edge-effects at the water-plastic-air interfaces, ADC estimates from line profiles (width 9 pixels) drawn through the center of the central slice showed no noticeable deviation in median ADC estimates across the FOV in the right–left and anterior–posterior directions (<5%) for either sequence [Figs. 9(c)–9(f)].

3.E. Repeatability of ADC estimates in temperature-controlled phantom

The repeatability of the ADC estimates from five samples in an ice-water phantom was good, with CVs below 4% for all tubes on all three scanners (Table II). There was

good agreement between ADC estimates from three scanners, with all scanners showing <5% deviation from the mean ADC estimate across all scanners for all tubes. None of the scanners showed systematic variations over the course of the study.

3.F. Diffusion-weighted images of healthy volunteers

Figure 10 shows axial images ($b = 100 \text{ s mm}^{-2}$ and $b = 900 \text{ s mm}^{-2}$) and ADC maps from the same healthy volunteer from three scanners. Images from all three scanners showed good fat suppression, no marked ghosting artifacts, no gross geometric distortion, and no marked parallel imaging artifacts, although the SNR was noticeably lower in the images from scanner 3.

3.G. Repeatability of ADC estimates in healthy volunteers

One volunteer from scanner 2 was excluded from analysis of the uterus due to susceptibility artifacts arising from bowel gas but was included in analysis of other organs. Bland–Altman plots for ADCs measured in kidneys, liver,

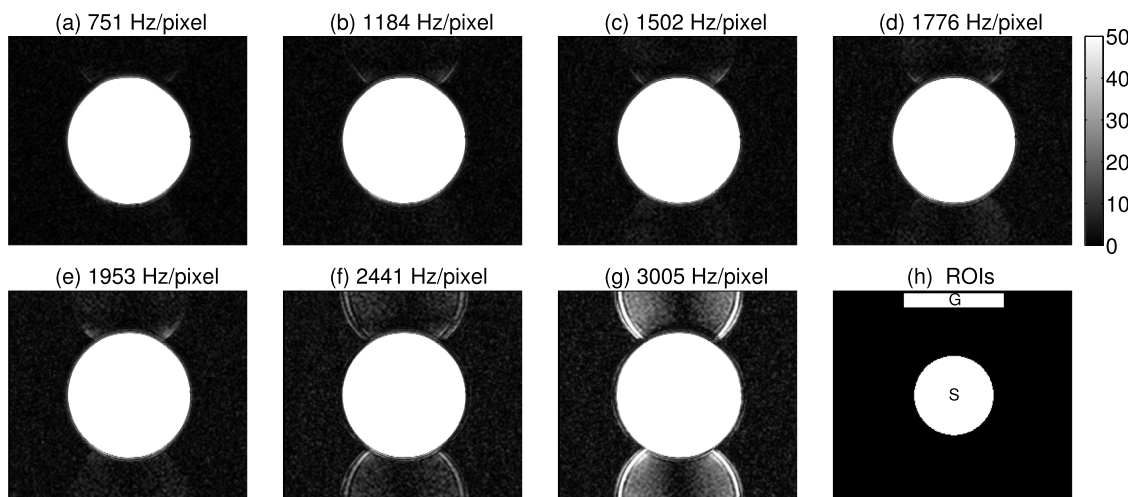


FIG. 6. Axial images of a PDMS phantom acquired using receive bandwidths (a) 751 Hz/pixel, (b) 1184 Hz/pixel, (c) 1502 Hz/pixel, (d) 1776 Hz/pixel, (e) 1953 Hz/pixel, (f) 2441 Hz/pixel, and (g) 3005 Hz/pixel. All other measurement parameters were consistent between the acquisitions. Scanner 1; ss-EPI; axial slices; 6 mm slice thickness; FOV 380×332 mm; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 178 ms; TR = 8000 ms; no parallel imaging; monopolar diffusion gradients; trace images; $b = 0 \text{ s mm}^{-2}$; 1 NSA. Images have been windowed to enhance the appearance of the ghosts. All images are shown at the same window width and level. ROIs used to estimate intensities in ghosts (G) and signal (S) are shown in (h).

spleen, and uterus are shown in Fig. 11. The mean difference between pairs of ADC estimates was not significantly different from zero for any of the organs on any of the scanners (single-sample t -test on differences in natural logarithm of ADC, using a p -value < 0.05 to indicate significance). The repeatability of ADC estimates was good, with wCV 2%–4% for kidneys, 3%–7% for liver, 6%–9% for spleen, and 7%–10% for uterus (Table III). Kruskal–Wallis tests showed no significant difference in ADC estimates in kidneys, spleen, or uterus between three scanners. There was, however, a significant difference between ADC estimates in the liver from the three scanners.

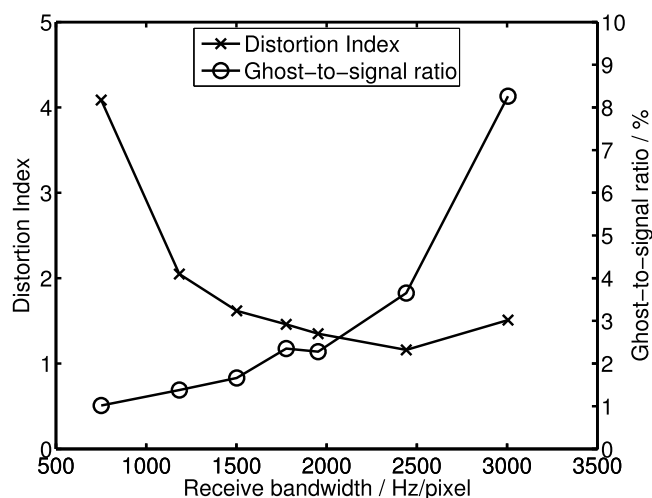


FIG. 7. Distortion index and ghost-to-signal ratio measured from axial images of a PDMS phantom acquired using receive bandwidths from 751 to 3005 Hz/pixel. All other measurement parameters were consistent between the acquisitions. Scanner 1; ss-EPI; axial slices; 6 mm slice thickness; FOV 380×332 mm; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 178 ms; TR = 8000 ms; no parallel imaging; monopolar diffusion gradients; trace images; maximum b -value = 900 s mm^{-2} ; 1 NSA. The ROI used to estimate background is shown in Fig. 5(d). ROIs used to estimate intensities in ghosts (G) and signal (S) are shown in Fig. 6(h).

4. DISCUSSION

These data demonstrate that measurements using simple phantoms may be used to optimize DW-MRI protocols for abdominal/pelvic imaging using a large FOV on scanners from three manufacturers and that these optimized protocols result in ADC values with low wCV across multiple sites and vendors. The high quality images obtained from such a quality-assured process mean that it is possible to implement a DW-MRI protocol for abdominal and pelvic imaging with a good degree of standardization across scanners from three manufacturers and obtain quantitative data in multicenter trials that can be pooled for analysis.

In development of abdominal/pelvic imaging, it is crucial to use large phantoms in order to assess behavior across the FOV that will be employed in the patient population. Small phantoms may fail to reveal artifacts that are apparent at a larger FOV, for example, as a result of poorer shimming across a large FOV. The phantoms that we developed in this study allow qualitative and quantitative assessment of geometric distortion, ghosting, and fat suppression, as well as repeatability and homogeneity of ADC estimates across a large FOV. In many cases, the images can be assessed on the scanner console, which facilitates rapid protocol development, which is particularly valuable in multicenter studies with limited time available in busy clinical departments. A large PDMS phantom, which has previously been used to assess geometric distortion in single-center studies,¹⁶ allows assessment of the effects of bandwidth, parallel imaging, and choice of diffusion gradient scheme in one imaging session using simple subtraction images. Assessment of ghosting, which has been shown to be scanner dependent in another study using a water bottle phantom,²⁷ can be combined with the same set of measurements.

Optimization of DW-MRI protocols requires trade-offs to be made between many imaging properties, for example, between geometric distortion and ghosting. B_0 distortion can

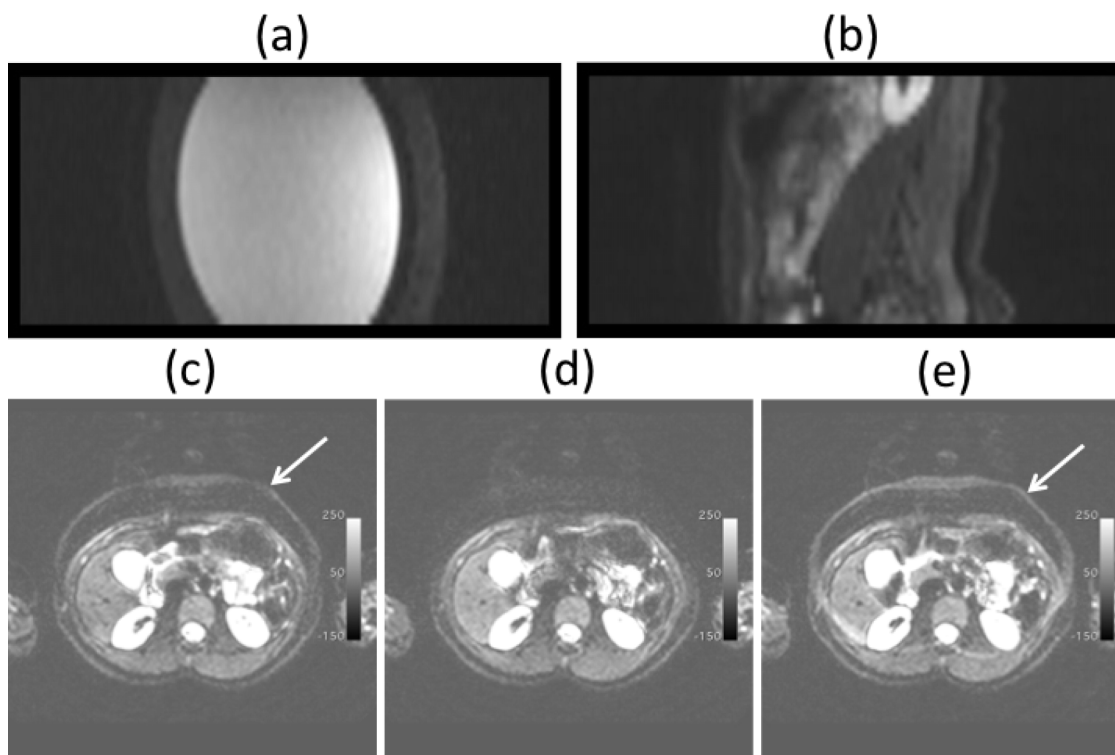


FIG. 8. [(a) and (b)] Sagittally reformatted images of (a) phantom and (b) abdominal image of healthy volunteer. Scanner 1; ss-EPI; axial slices; 6 mm slice thickness; FOV 380×332 mm; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 75 ms; TR = 8000 ms; parallel imaging reduction factor 2; receive bandwidth 1776 Hz/pixel; bipolar gradients; trace images; $b = 100 \text{ s mm}^{-2}$; 1 NSA for phantom; 4 NSA for volunteer; SPAIR fat suppression. [(c)–(e)] Optimization of SPAIR delay time (c) 90 ms, (d) 110 ms, and (e) 130 ms. Unsuppressed fat (arrow) is minimized at delay time of 110 ms. Scanner 3; ss-EPI; axial slices; 6 mm slice thickness; FOV 380×333 mm; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 96 ms; TR = 8000 ms; parallel imaging reduction factor 2; water-fat shift 11.1 pixels; monopolar gradients; trace images; $b = 100 \text{ s mm}^{-2}$; 1 NSA; SPAIR fat suppression.

be reduced by increasing the read bandwidth but this is accompanied by an increase in ghosting at higher bandwidths. Quantitative estimates of distortion and ghosting in images of the PDMS phantom allowed us to find an optimal range of bandwidths where both distortion and ghosting were low. In our experience, the preferred read bandwidth varies between scanners and may vary between protocols on the same scanner. Bandwidths in the range ~ 1500 – 2000 Hz/pixel gave acceptable results in the scanners assessed as part of this study. It is also important to note, however, that ghosts may be difficult to assess using protocols that apply SENSE or ASSET due to suppression of signals in background regions.

Development of protocols for multicenter studies necessitates additional trade-offs between optimization and standardization of protocols on different platforms. Some parameters could not be standardized, notably diffusion gradient scheme, diffusion gradient strengths and timings, and methods of fat suppression and parallel imaging. The twice-refocused spin-echo gradient scheme was used on scanners 1 and 2 as assessment of geometric distortion using a PDMS phantom showed that distortion was reduced by employing twice-refocused spin-echo (called “bipolar” and “DSE” on scanners 1 and 2, respectively) diffusion gradient schemes, compared with monopolar gradients, due to reduction in eddy current effects. Twice-refocused spin-echo schemes are, however, not available on all platforms, for example,

scanner 3 where a monopolar scheme was used. It would have been possible to standardize the protocols further by using a monopolar scheme on all scanners but this would result in impaired image quality compared with bipolar or DSE schemes. It is, however, important to note that the minimum TE may be higher using twice-refocused gradients, possibly leading to loss of SNR. Twice-refocused schemes may also result in lower SNR than monopolar gradients even at the same TE, which may be due to imperfect refocusing. Geometric distortion was also reduced by employing parallel imaging, which reduces the length of the echo train, and has the additional advantage of reducing the minimum available TE, thus improving SNR. Parallel imaging was employed in this study, despite the lack of standardization between platforms. It would be possible to standardize the acquisition further by not using parallel imaging but this would result in more distortion, due to longer echo trains, and lower SNR, due to longer TE.

TE was reduced on scanners 1 and 3 by employing diffusion encoding schemes (three-scan trace or gradient overplus) that use three orthogonal diffusion gradients that are not aligned with the cardinal directions of the scanner. These schemes allow reduction in TE compared with three gradient directions aligned with axes of magnet (called orthogonal or ALL on some platforms) due to reduction in ramp times. It would have been possible to standardize the protocols further by using orthogonal gradients or single gradient directions on

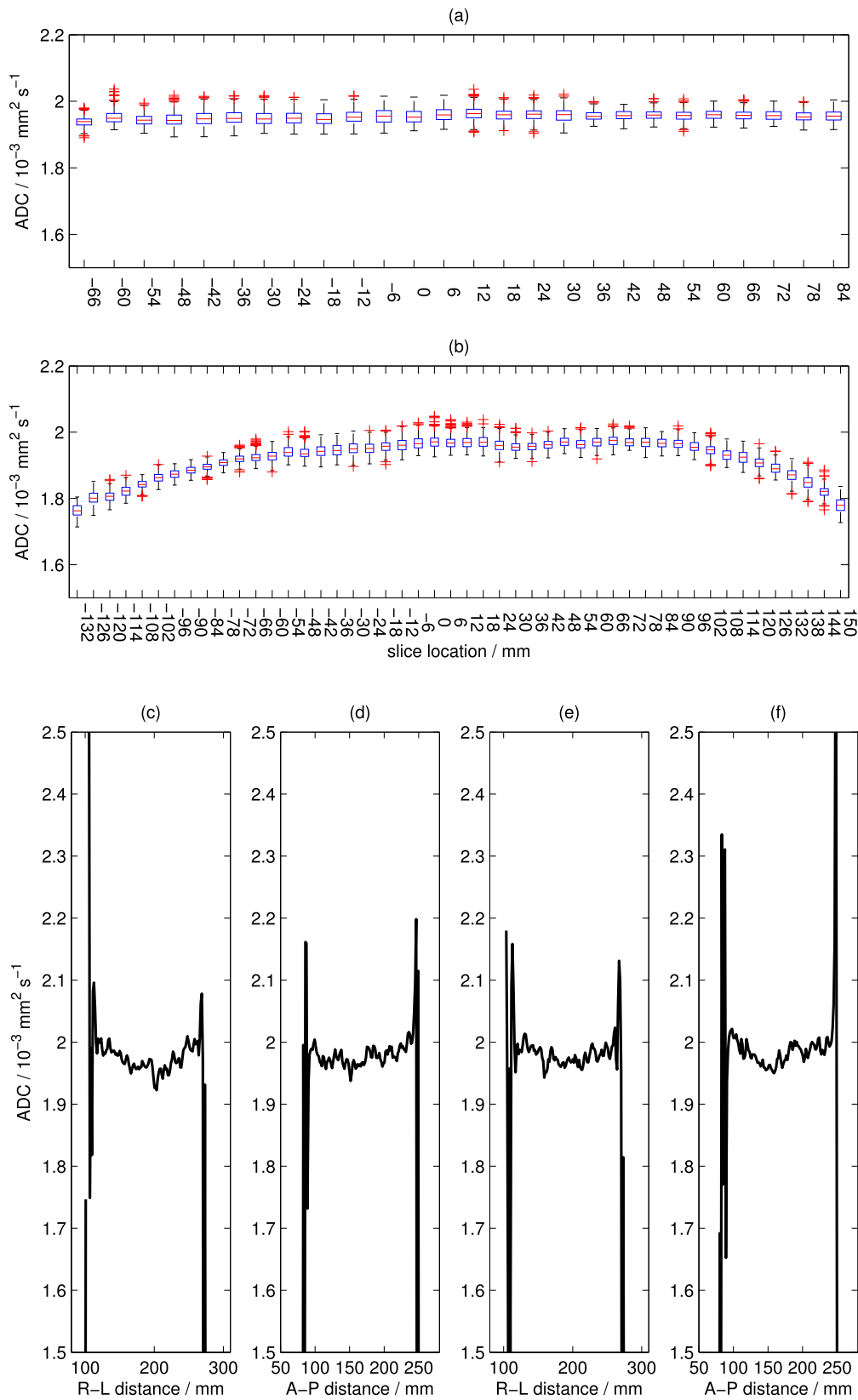


FIG. 9. ADC estimates from square ROIs (81 pixels) drawn at the center of each slice of a stack of axial slices through a water bottle phantom and ADC estimates from line profiles drawn through the center of the central slice. (a), (c), and (d) 26 slices, (b), (e), and (f) 48 slices. Line profiles drawn in (c) and (e) right–left and (d) and (f) anterior–posterior directions. All other measurement parameters were consistent between the acquisitions. Scanner 1; ss-EPI; axial slices; 6 mm slice thickness; FOV 380×332 mm; acquired matrix 128; reconstructed matrix 256; PE direction = AP; TE = 75 ms; TR = 8000 ms; parallel imaging reduction factor 2; receive bandwidth 1776 Hz/pixel; bipolar diffusion gradients; trace images; $b = 0, 100, 500, 900 \text{ s mm}^{-2}$; 1 NSA. Boxplots show ADC estimates from all pixels in the ROI for each slice. Horizontal lines (red) represent the median ADC, boxes (blue) represent the interquartile range, whiskers represent the range, excluding outliers, and crosses (red) represent values judged to be outliers. Line profiles represent median ADCs at each point in the line (width 9 pixels).

TABLE II. Mean, standard deviation, and coefficient of variation of ADC estimates from repeated measurements of five tubes of sucrose solutions in an ice-water phantom on three scanners. The mean ADC estimates from three scanners were calculated by taking the mean of the three mean ADCs for each tube.

Sample	Sucrose concentration/% w/v	Scanner	Mean (\pm SD) of ADC estimates/ 10^{-3} mm ² s ⁻¹	Coefficient of variation (CV)/%	Mean ADC estimate from three scanners/ 10^{-3} mm ² s ⁻¹
1	0	1	1.14 (\pm 0.02)	2.1	1.14
		2	1.14 (\pm 0.03)	2.6	
		3	1.13 (\pm 0.01)	0.7	
2	0	1	1.15 (\pm 0.02)	1.6	1.13
		2	1.12 (\pm 0.01)	0.8	
		3	1.11 (\pm 0.01)	0.3	
3	10	1	0.96 (\pm 0.02)	1.7	0.94
		2	0.94 (\pm 0.02)	1.9	
		3	0.92 (\pm 0.02)	2.3	
4	10	1	0.96 (\pm 0.02)	2.5	0.94
		2	0.93 (\pm 0.01)	0.7	
		3	0.93 (\pm 0.02)	2.6	
5	20	1	0.77 (\pm 0.01)	1.1	0.76
		2	0.78 (\pm 0.03)	3.5	
		3	0.72 (\pm 0.01)	0.9	

all scanners but this would have resulted in longer TE, and hence lower SNR, compared with three-scan trace or gradient overplus. Despite the use of gradient overplus, the shortest TE available for this protocol on scanner 3 was still longer than on scanners 1 and 2 (96 ms on scanner 3 compared with 75 and 81 ms on scanners 1 and 2, respectively). The longer TE on scanner 3 is a possible explanation for the lower SNR on this scanner.

The timings of the diffusion gradients (Δ and δ) vary between different implementations of diffusion gradient schemes and were not standardized between scanners in this study. Furthermore, the definition of b -values in terms of Δ and δ is different for double spin-echo and monopolar gradient schemes.²⁸ The values of Δ and δ may not be readily available on all platforms and the effects of varying Δ and δ on the ADC estimates have not been fully explored. In this study, the large differences in Δ and δ between scanners do not appear to have manifested themselves in the ADC results as only small differences were observed between ADC estimates from the three scanners.

Previous studies have described the development of a large phantom containing corn oil, which was used in this study to assess fat suppression over a large FOV.¹⁸ Corn oil has similar spectral properties and T_1 relaxation times to subcutaneous abdominal fat observed *in vivo* and the large size of the phantom allows assessment of fat suppression at the edges of the FOV, for example, where B_0 homogeneity may be poorer. Fat suppression requires extensive optimization on some scanners. The preferred method (or methods, since combinations of methods are possible on some platforms) may depend on the scanner and application. It is not possible to standardize the method of fat suppression across scanners from different manufacturers although spectral methods may be employed on all scanners. Spectral methods were chosen

over inversion recovery (IR) in the protocols described here as IR reduces the overall signal and introduces T_1 weighting.

Ice-water phantoms have previously been used to compare ADC estimates between scanners.^{19,29} Phantoms containing sucrose solutions have also been used to assess long-term repeatability of ADC estimates.³⁰ Ice-water provides a simple and inexpensive method to control temperature. Sucrose restricts diffusion of water molecules and can be used to reduce the ADC of water to values comparable to ADCs observed *in vivo*. Our ADC estimates from samples 1 and 2 (0% sucrose) were in good agreement ADC estimates in distilled water at $\sim 0^\circ\text{C}$ reported in other studies.¹⁹ Inclusion of samples containing 10%–20% sucrose allowed us to compare ADC estimates between the three scanners at the range of ADCs observed in tumors and normal tissues, confirming that there was good agreement between the three scanners across the relevant range of ADCs. The use of a large FOV imposes an additional requirement for good homogeneity of ADC estimates across the FOV in addition to accurate ADC estimates near the magnet isocenter. Homogeneity of ADC estimates is particularly relevant in patients with multiple lesions spread across a large volume where some lesions may be far from the magnet isocenter. By using an ice-water phantom to obtain absolute ADC estimates near the magnet isocenter, in combination with relative measurements across the FOV using a water bottle, we were able to assess the performance of our protocols for quantitative DW-MRI across large imaging volumes. Assessments of homogeneity of ADC estimates using a large water bottle were used to determine the maximum feasible extent of the FOV in the z -direction. Deviation in ADC estimates at the ends of a large FOV may lead to errors in measurements of tumors at the ends of the FOV in the case of very large tumors or multiple

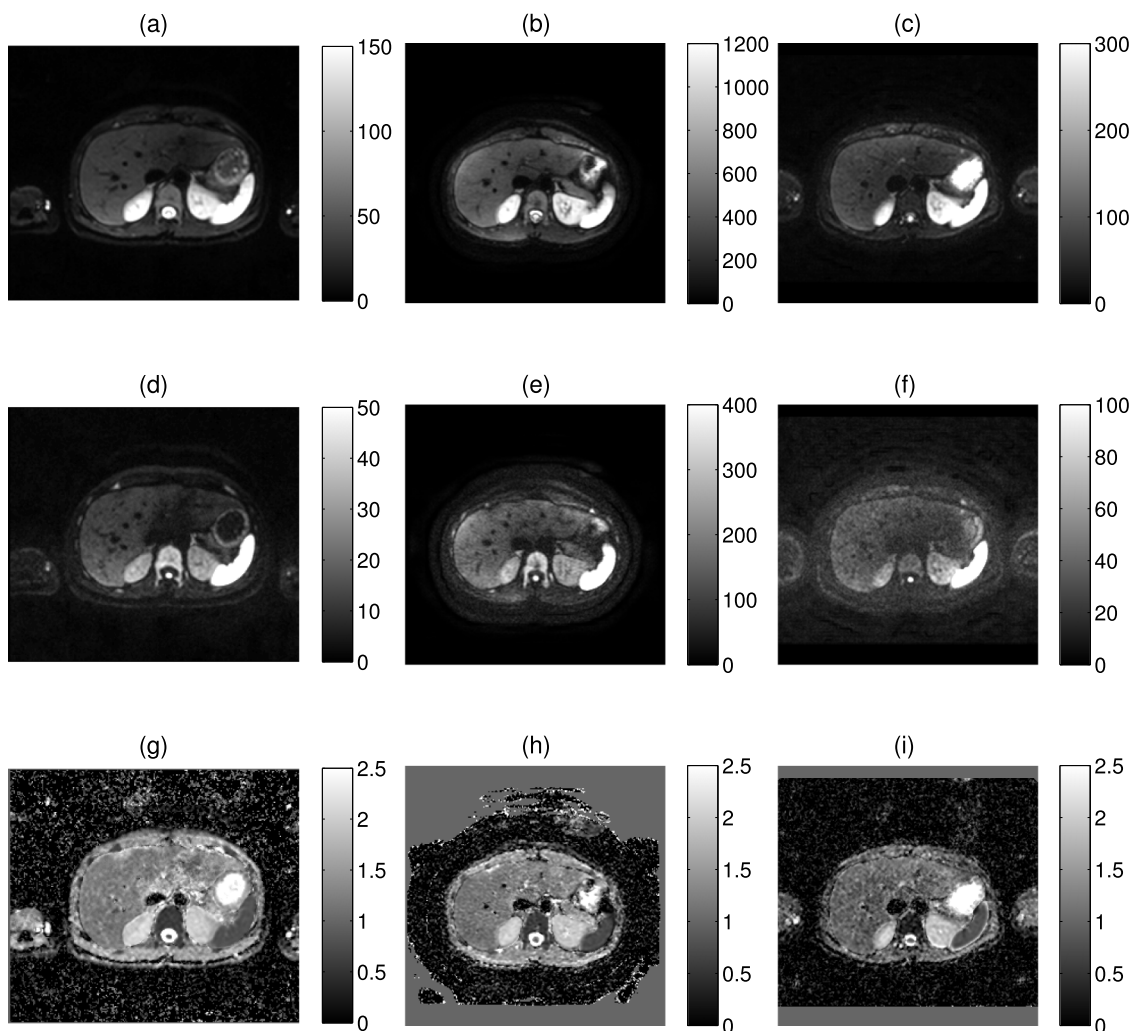


FIG. 10. Axial images of the same healthy volunteer from three scanners [(a)–(c)] $b = 100 \text{ s mm}^{-2}$; [(d)–(f)] $b = 900 \text{ s mm}^{-2}$; [(g)–(i)] ADC maps. Protocols described in Table I (a), (d), and (g) scanner 1; (b), (e), and (h) scanner 2; (c), (f), and (i) scanner 3.

lesions. The results described in this study indicate that the ADC may be underestimated by 10% at the ends of a large FOV ($>200 \text{ mm}$) in the z -direction. Inhomogeneity of ADC estimates across a large FOV may be due to nonlinearity of the diffusion-weighting gradients.³¹ In this study, 26 slices of 6 mm slice thickness were used, providing a 156 mm FOV in the z -direction, as this was achievable on all scanners in the study. The good homogeneity of ADC estimates in the right–left and anterior–posterior directions confirms the feasibility using a large in-plane; this test can also be used to detect artifacts, for example, due to inappropriate image filters or susceptibility artifacts.

The good repeatability of ADC estimates in the ice-water phantom and in healthy volunteers was in agreement with the previous studies,^{5,19,32,33} indicating that the protocols described here are robust and suitable for quantitative studies. There were no marked differences in repeatability of ADC estimates between the three scanners, which is an important requirement for comparison of ADC changes in multicenter studies, for example, in assessment of response to treatment.

A limitation of this study was the use of different volunteers on each scanner as logistical constraints prevented

ten volunteers from traveling to all three institutions. It was, therefore, only possible to compare ADC estimates from the cohorts, rather than comparing individual measurements. A further limitation is that only 1.5 T scanners were included in this study. However, the phantoms described can also be used at 3 T where similar optimization is required although the challenges presented are greater.¹⁸

A previous multicenter study, which looked at ADC estimates in abdominal organs in DW-MRI data from healthy volunteers imaged on scanners from three manufacturers at 1.5 and 3 T, showed good agreement between ADC estimates in gall bladder, pancreas, spleen, and kidneys from different scanners at 1.5 T but poor agreement in pancreas and kidneys at 3 T and poor agreement in liver at both field strengths.³⁴ The DW-MRI protocols used the same sequence (single-shot EPI), FOV, matrix, number of averages, slice thickness, and slice gap in all scanners but the TE, TR, b -values, and methods of parallel imaging were not standardized (although a standard subset of b -values were used for analysis). Another multicenter study, which looked at gray matter and white matter in healthy volunteers, found significant variation in ADC estimates between scanners from

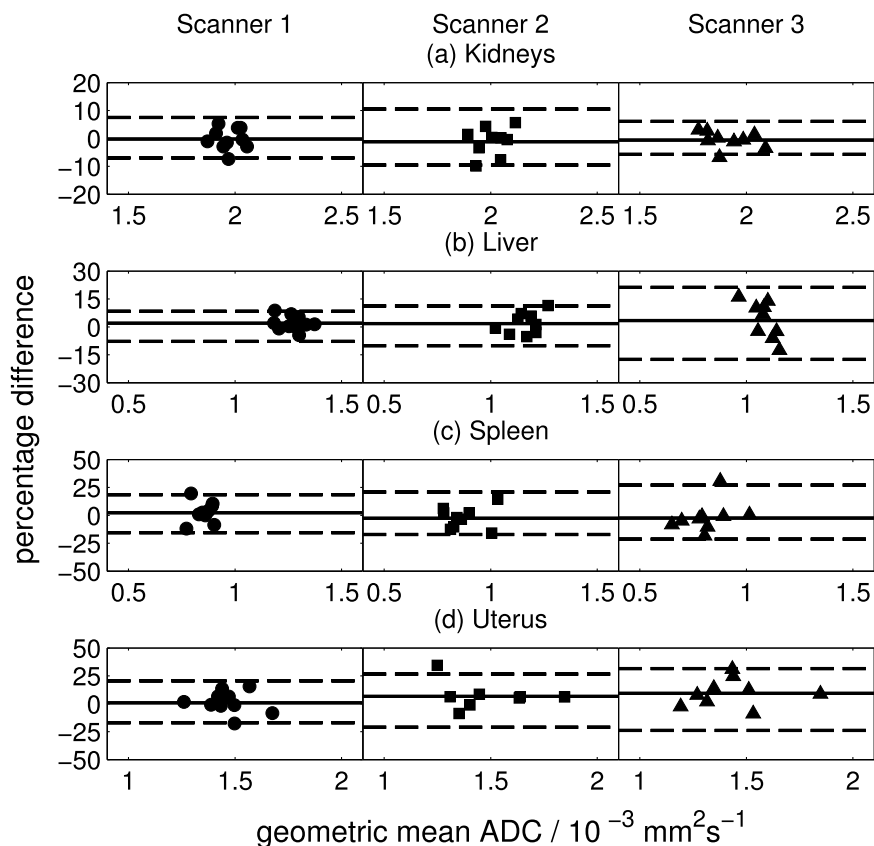


FIG. 11. Bland–Altman plots for median ADC estimates in (a) kidneys, (b) liver, (c) spleen, and (d) uterus of ten healthy volunteers on three scanners using protocols described in Table I. Key to symbols: circles scanner 1; squares scanner 2; triangles scanner 3.

the same manufacturer and between scanners from different manufacturers.³⁵ The FOV, matrix, number of averages, slice thickness, slice gap, number of motion probing gradients, and b -values were kept constant between scanners but again TE, TR, and methods of parallel imaging were not standardized.

In both of these studies, as in our protocols, timing parameters and methods of parallel imaging differed between scanners, demonstrating the difficulty of producing fully standardized DW-MRI protocols across multiple platforms.

In our study, the good agreement of ADC estimates in the ice-water phantom between the three scanners and the

TABLE III. Coefficients of variation, median, and range of ADC estimates from abdominal and pelvic organs measured on three scanners. p -values show results of Kruskal–Wallis test for each organ.

Organ	Scanner	Within-patient coefficient of variation (wCV)/%	Median (range) of ADC estimates/ $10^{-3} \text{ mm}^2 \text{ s}^{-1}$	p -value (Kruskal–Wallis)
Kidneys	1	2.6	1.97 (1.87–2.06)	0.1
	2	3.6	2.01 (1.89–2.12)	
	3	2.1	1.88 (1.78–2.09)	
Liver	1	2.9	1.27 (1.18–1.37)	2×10^{-4a}
	2	3.9	1.17 (1.02–1.27)	
	3	6.9	1.08 (0.97–1.16)	
Spleen	1	6.1	0.86 (0.77–0.90)	0.4
	2	6.8	0.84 (0.78–1.04)	
	3	8.7	0.81 (0.65–1.01)	
Uterus	1	6.8	1.46 (1.26–1.68)	0.8
	2	8.5	1.43 (1.26–1.85)	
	3	9.9	1.45 (1.19–1.85)	

^a $p < 0.05$.

absence of any significant difference in ADC estimates in the kidneys, spleen, and uterus in healthy volunteers show that it is possible to obtain similar ADC estimates on different scanners despite the inherent differences in the acquisitions. The significant difference in ADC estimates in liver between the three scanners in our study corroborates the results from an earlier study, which also demonstrated poor agreement in ADC estimates in the liver between 1.5 T scanners from different manufacturers.³⁴ These results suggest that multicenter studies using ADC estimates in normal liver may require tighter protocol matching, including TE and other timing parameters, and higher SNR. The lack of agreement in ADC estimates in the liver may be related to the lower signal in the liver compared with other organs or phantom measurements, owing to the relatively short T_2 of the liver, compared with other organs measured in this study ($T_2 \sim 46$ ms at 1.5 T, compared with 85–87 ms for kidney, 79 ms for spleen, and 117 ms for myometrium),³⁶ leading to greater sensitivity to the noise characteristics of the three scanners. It is important to note, however, that tumors generally have high SNR on diffusion-weighted images and would therefore be less sensitive to noise. Tumors may, therefore, be expected to show good agreement in ADC estimates between scanners, as observed in kidneys, spleen, and uterus. Furthermore, the change in ADC between pre- and post-treatment measurements carried out on the same scanner may be acceptable in multicenter studies if absolute values cannot be compared. This remains to be verified in patient studies.

5. CONCLUSIONS

A range of simple phantoms can be used to optimize protocols for DW-MRI of the abdomen and pelvis on 1.5 T scanners from three manufacturers. The optimized protocols give good quality images across a large FOV in healthy volunteers with excellent repeatability of ADC estimates in abdominal and pelvic organs (wCV = 2%–10% for median ADCs). The methods described here serve as a framework for setting up multicenter DW-MRI studies.

ACKNOWLEDGMENTS

The authors thank the generous volunteers for their time and patience. The authors are grateful to Kathy May and Tiffany Rushen for their input in trial setup and maintenance. The authors acknowledge funding from Cancer Research UK BIDD Grant No. C1353/A12762 and Cancer Research UK and Engineering and Physics Sciences Research Council Cancer Imaging Centre in association with Medical Research Council and Department of Health No. C1060/A10334 and National Health Service funding to the Clinical Research Facility in Imaging and the National Institute for Health Research Biomedical Research Centres at both Royal Marsden Hospital/Institute of Cancer Research and Cambridge. The authors acknowledge the support of the National Institute for Health Research, through the Cancer

Research Network. The authors are also grateful for financial support from Addenbrooke's Charitable Trust.

^{a)} Author to whom correspondence should be addressed. Electronic mail: jessica.winfield@icr.ac.uk

¹D.-M. Koh, E. Scurr, D. Collins, B. Kanber, A. Norman, M. O. Leach, and J. E. Husband, "Predicting response of colorectal hepatic metastasis: Value of pretreatment apparent diffusion coefficients," *AJR, Am. J. Roentgenol.* **188**, 1001–1008 (2007).

²Y. Cui, X.-P. Zhang, Y.-S. Sun, L. Tang, and L. Shen, "Apparent diffusion coefficient: Potential imaging biomarker for prediction and early detection of response to chemotherapy in hepatic metastases," *Radiology* **248**, 894–900 (2008).

³L. R. Jensen, B. Garzon, M. G. Heldahl, T. F. Bathen, S. Lundgren, and I. S. Gribbestad, "Diffusion-weighted and dynamic contrast-enhanced MRI in evaluation of early treatment effects during neoadjuvant chemotherapy in breast cancer patients," *J. Magn. Reson. Imaging* **34**, 1099–1109 (2011).

⁴S. H. Park, W. K. Moon, N. Cho, I. C. Song, J. M. Chang, I.-A. Park, W. Han, and D.-Y. Noh, "Diffusion-weighted MR imaging: Pretreatment prediction of response to neoadjuvant chemotherapy in patients with breast cancer," *Radiology* **257**, 56–63 (2010).

⁵S. Kyriazi, D. J. Collins, C. Messiou, K. Pennert, R. L. Davidson, S. L. Giles, S. B. Kaye, and N. M. deSouza, "Metastatic ovarian and primary peritoneal cancer: Assessing chemotherapy response with diffusion-weighted MR imaging—Value of histogram analysis of apparent diffusion coefficients," *Radiology* **261**, 182–192 (2011).

⁶E. Sala, M. Y. Kataoka, A. N. Priest, A. B. Gill, M. A. McLean, I. Joubert, M. J. Graves, R. A. F. Crawford, M. Jimenez-Linan, H. M. Earl, C. Hodgkin, J. R. Griffiths, D. J. Lomas, and J. D. Brenton, "Advanced ovarian cancer: Multiparametric MR imaging demonstrates response- and metastasis-specific effects," *Radiology* **263**, 149–159 (2012).

⁷H. Yabuuchi, M. Hatakenaka, K. Takayama, Y. Matsuo, S. Sunami, T. Kamitani, M. Jinnouchi, S. Sakai, Y. Nakanishi, and H. Honda, "Non-small cell lung cancer: Detection of early response to chemotherapy by using contrast-enhanced dynamic and diffusion-weighted MR imaging," *Radiology* **261**, 598–604 (2011).

⁸V. N. Harry, S. I. Semple, F. J. Gilbert, and D. E. Parkin, "Diffusion-weighted magnetic resonance imaging in the early detection of response to chemoradiation in cervical cancer," *Gynecol. Oncol.* **111**, 213–220 (2008).

⁹A. D. King, F. K. F. Mo, K.-H. Yu, D. K. W. Yeung, H. Zhou, K. S. Bhatia, G. M. K. Tse, A. C. Vlantis, J. K. T. Wong, and A. T. Ahuja, "Squamous cell carcinoma of the head and neck: Diffusion-weighted MR imaging for prediction and monitoring of treatment response," *Eur. Radiol.* **20**, 2213–2220 (2010).

¹⁰V. Vandecavaye, P. Dirix, F. De Keyser, K. Op de Beeck, V. V. Poorten, I. Roebben, S. Nuyts, and R. Hermans, "Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma," *Eur. Radiol.* **20**, 1703–1714 (2010).

¹¹T. Niwa, M. Ueno, S. Ohkawa, T. Yoshida, T. Douichi, K. Ito, and T. Inoue, "Advanced pancreatic cancer: The use of apparent diffusion coefficient to predict response to chemotherapy," *Br. J. Radiol.* **82**, 28–34 (2009).

¹²Y. Ohno, H. Koyama, T. Yoshikawa, K. Matsumoto, N. Aoyama, Y. Onishi, and K. Sugimura, "Diffusion-weighted MRI versus ¹⁸F-FDG PET/CT: Performance as predictors of tumor treatment response and patient survival in patients with non-small cell lung cancer receiving chemoradiotherapy," *AJR, Am. J. Roentgenol.* **198**, 75–82 (2012).

¹³J. L. Zhang, E. E. Sigmund, H. Chandarana, H. Rusinek, Q. Chen, P.-H. Vivier, B. Taouli, and V. S. Lee, "Variability of renal apparent diffusion coefficients: Limitations of the monoexponential model for diffusion quantification," *Radiology* **254**, 783–792 (2010).

¹⁴B. 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," *Invest. Radiol.* **45**, 104–108 (2010).

¹⁵A. R. Padhani, G. Liu, D.-M. Koh, T. L. Chenevert, H. C. Thoeny, T. Takahara, A. Dzik-Jurasz, B. D. Ross, M. Van Cauteren, D. Collins, D. A. Hammoud, G. J. S. Rustin, B. Taouli, and P. L. Choyke, "Diffusion-weighted magnetic resonance imaging as a cancer biomarker: Consensus and recommendations," *Neoplasia* **11**, 102–125 (2009).

¹⁶D. J. Collins and M. Blackledge, "Techniques and optimization," in *Diffusion-Weighted MR Imaging*, 1st ed., edited by D.-M. Koh and H. C. Thoeny (Springer, Berlin, 2010), pp. 20–25.

- ¹⁷S. Kyriazi, M. Blackledge, D. J. Collins, and N. M. deSouza, "Optimising diffusion-weighted imaging in the abdomen and pelvis: Comparison of image quality between monopolar and bipolar single-shot spin-echo echo-planar sequences," *Eur. Radiol.* **20**, 2422–2431 (2010).
- ¹⁸J. M. Winfield, N. H. M. Douglas, N. M. deSouza, and D. J. Collins, "Phantom for assessment of fat suppression in large field-of-view diffusion-weighted magnetic resonance imaging," *Phys. Med. Biol.* **59**, 2235–2248 (2014).
- ¹⁹D. Malyarenko, C. J. Galban, F. J. Londy, C. R. Meyer, T. D. Johnson, A. Rehemtulla, B. D. Ross, and T. L. Chenevert, "Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom," *J. Magn. Reson. Imaging* **37**, 1238–1246 (2013).
- ²⁰N. H. M. Douglas, J. M. Winfield, N. M. deSouza, D. J. Collins, and M. R. Orton, "Development of a phantom for quality assurance in multicentre clinical trials with diffusion-weighted MRI," *Proceedings of the International Society of Magnetic Resonance in Medicine, 2013*, presentation number 3114.
- ²¹M. D. Blackledge, M. O. Leach, D. J. Collins, and D.-M. Koh, "Computed diffusion-weighted MR imaging may improve tumor detection," *Radiology* **261**, 573–581 (2011).
- ²²D. Le Bihan, R. Turner, C. T. W. Moonen, and J. Pekar, "Imaging of diffusion and microcirculation with gradient sensitization: Design, strategy, and significance," *J. Magn. Reson. Imaging* **1**, 7–28 (1991).
- ²³D.-M. Koh, D. J. Collins, and M. R. Orton, "Intravoxel incoherent motion in body diffusion-weighted MRI: Reality and challenges," *AJR, Am. J. Roentgenol.* **196**, 1351–1361 (2011).
- ²⁴K. G. Hollingsworth and D. J. Lomas, "Influence of perfusion on hepatic MR diffusion measurement," *NMR Biomed.* **19**, 231–235 (2006).
- ²⁵M. J. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet* **327**, 307–310 (1986).
- ²⁶X. He and S. O. Oyadiji, "Application of coefficient of variation in reliability-based mechanical design and manufacture," *J. Mater. Process. Technol.* **119**, 374–378 (2001).
- ²⁷M. Giannelli, S. Diciotti, C. Tessa, and M. Mascalchi, "Characterization of Nyquist ghost in EPI-fMRI acquisition sequences implemented on two clinical 1.5 T MR scanner systems: Effect of readout bandwidth and echo spacing," *J. Appl. Clin. Med. Phys.* **11**, 170–180 (2010).
- ²⁸M. Koch and D. G. Norris, "An assessment of eddy current sensitivity and correction in single-shot diffusion-weighted imaging," *Phys. Med. Biol.* **45**, 3821–3832 (2000).
- ²⁹T. L. Chenevert, C. J. Galban, M. K. Ivancevic, S. E. Rohrer, F. J. Londy, T. C. Kwee, C. R. Meyer, T. D. Johnson, A. Rehemtulla, and B. D. Ross, "Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies," *J. Magn. Reson. Imaging* **34**, 983–987 (2011).
- ³⁰I. 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," *Phys. Med. Biol.* **49**, 1409–1422 (2004).
- ³¹D. I. Malyarenko, D. Newitt, L. J. Wilmes, A. Tudorica, K. G. Helmer, L. R. Arlinghaus, M. A. Jacobs, G. Jajamovich, B. Taouli, T. E. Yankeelov, W. Huang, and T. L. Chenevert, "Demonstration of nonlinearity bias in the measurement of the apparent diffusion coefficient in multicenter trials," *Magn. Reson. Med.* (2015). [Epub ahead of print.]
- ³²D.-M. Koh, M. Blackledge, D. J. Collins, A. R. Padhani, T. Wallace, B. Wilton, N. J. Taylor, J. J. Stirling, R. Sinha, P. Walicke, M. O. Leach, I. Judson, and P. Nathan, "Reproducibility and changes in the apparent diffusion coefficient of solid tumours treated with combretastatin A4 phosphate and bevacizumab in a two-centre phase I clinical trial," *Eur. Radiol.* **19**, 2728–2738 (2009).
- ³³A. Andreou, D.-M. Koh, D. J. Collins, M. Blackledge, T. Wallace, M. O. Leach, and M. R. Orton, "Measurement reproducibility of perfusion fraction and pseudodiffusion coefficient derived by intravoxel incoherent motion diffusion-weighted MR imaging in normal liver and metastases," *Eur. Radiol.* **23**, 428–434 (2013).
- ³⁴O. F. Donati, D. Chong, D. Nanz, A. Boss, J. M. Froehlich, E. Andres, B. Seifert, and H. C. Thoeny, "Diffusion-weighted MR imaging of upper abdominal organs: Field strength and intervendor variability of apparent diffusion coefficients," *Radiology* **270**, 454–463 (2014).
- ³⁵M. Sasaki, K. Yamada, Y. Watanabe, M. Matsui, M. Ida, S. Fujiwara, and E. Shibata, "Variability in absolute apparent diffusion coefficient values across different platforms may be substantial: A multivendor, multi-institutional comparison study," *Radiology* **249**, 624–630 (2008).
- ³⁶C. M. J. de Bazelaire, G. D. Duhamel, N. M. Rofsky, and D. C. Alsop, "MR imaging relaxation times of abdominal and pelvic tissues measured *in vivo* at 3.0 T: Preliminary results," *Radiology* **230**, 652–659 (2004).