10.1148/radiol.2020200221

Appendix E1

Additional Methods Details

Table E1 includes the acquisition details, including scan order and acquisition times.

Data acquired with the axially reformatted simultaneous multislice imaging (AR-SMS) protocol were reconstructed offline in MATLAB 2017b (Mathworks, Natick, MA) using an in-house pipeline. Across five random cases, on a single server (quad AMD Opteron 6140, 256 GB RAM), the average total reconstruction time, excluding data transfer, was 6 hours 36 minutes per case of 28 acquisitions each; approximate times for each step are listed. The reconstruction was developed for high image quality and was not optimized for efficient reconstruction. Future work would be necessary to speed up the reconstruction for clinical use.

First, each acquisition of undersampled data and reference lines was independently corrected for Nyquist ghosts with a slice-and coil-specific, first-order correction estimated by ghost/object minimization (14) (40 minutes). Single band and fully sampled reference data were used to calculate weights for phase encoding unaliasing using GRAPPA (1) and SMS separation using Slice-GRAPPA (11) (10 minutes). SMS unaliasing was performed on each acquisition (35 minutes), followed by a second iteration of slice-specific Nyquist ghost correction on each unaliased slice using ghost/object minimization (63 minutes). GRAPPA weights were applied for phase encoding unaliasing (39 minutes). The partial Fourier edge was filtered with a $\sin^2$ window (7 seconds). Coil combination was performed using a root sum-of-square (16 minutes). The magnitude images were then corrected for geometric distortion using topup (21) based on fully sampled reference data with reversed phase encoding (3 hours 8 minutes). Averaging was performed to combine equal b-values, and then the ADC map was calculated using a log-linear fit with masking of negative values (0.5 seconds). Finally, the separate axial b-value images and ADC maps were written in Digital Imaging and Communications in Medicine (DICOM) format for viewing.

Subanalysis Comparing Standard-A and Standard-B

The standard SE-EPI protocol used was determined by the clinical indication; I-SPY 2 patients received Standard-A, while clinical, nontrial patients received the shorter Standard-B. To have a larger number of lesions, the main analysis treated both variations of standard SE-EPI (A and B) as a single group, expecting negligible differences between the scans. To confirm this choice, a linear mixed model was retrospectively performed in a subanalysis to fit the image quality ratings after separating Standard-A and Standard-B. The quality of Standard-A and Standard-B were not statistically different, modeled as 2.11 and 1.92, respectively ($P > .354$). Moreover, Table E2 shows that both Standard-A and Standard-B were rated significantly lower than RS-EPI and AR-SMS. This confirms that the previous interpretation of method comparison is consistent for both Standard-A and Standard-B.

Phantom ADC Quantification

The breast phantom (CaliberMRI) contains twelve compartments of varying polyvinylpyrrolidone (0%–40%) and water to imitate in vivo tissue with various ADC values. The manufacturer reported nominal b-values for these compartments between 533 and 1456 mm^2/s measured at 18.2°C ± 0.4°C. The breast phantom was scanned in the 16-channel breast coil (ie, outside of isocenter) using each method, repeated four times in one scan session. The experiment was not temperature controlled and did not include gradient nonlinearity correction. The temperature of the scan room was approximately 70°F (21.1°C). The nominal true ADC values reported by the manufacturer were adjusted by 2.4%/°C following (2). Standard and RS-EPI ADC maps were generated online such that details on fitting, filtering, masking etc are unknown to the authors. AR-SMS was reconstructed by an inhouse pipeline, including a log-linear fit to produce ADC maps. Regions of interest were manually drawn (initials) on Digital Imaging and Communications in Medicine (DICOM) images, in a center slice of each compartment to measure the mean ADC.

The mean ADC values from each compartment, averaged across four measurements is plotted in Figure E1 against the adjusted nominal ADC, where the dotted line represents concordance. Despite using varying methods, b-values, averages, and resolutions, all three methods measure similar ADC values, which correspond closely to the nominal ADC values estimated at 70°F (dotted line). Standard and RS-EPI measure ADC values that are slightly higher than those measured by AR-SMS.

This ADC comparison is limited by the lack of temperature control, potential differences in ADC fitting algorithms, and uncorrected gradient nonlinearities, especially outside of iso-center. However, the results suggest that both advanced DWI methods measure reasonable ADC values without substantial bias. Future work is needed to define ADC thresholds and explore the diagnostic value of AR-SMS.

Lesion Size Comparison

In the reader study, radiologists were asked to measure the longest diameter (LD) of the lesion on a CE-subtraction image, followed by $b = 800\ s/mm^2$ images from each DWI method in random order. The radiologist was free to pick a representative slice of their choosing or decline to answer based on their perceived inability to measure. The root-mean-square error (RMSE) was measured between LDs measured on CE-MRI and $b = 800\ s/mm^2$ images from each method. Lesions for which no measurement was recorded were excluded.

The RMSE for each DWI method is listed in Figure E2. The lesion LD measured on AR-SMS was closest to that measured on the CE-subtraction image, followed by RS-EPI. The lesion was considered “unmeasureable” on 5 reads for Standard SE-EPI and 2 for each RS-EPI and AR-SMS.

The comparison of lesion length measurements between CE-MRI and high b-value images gives us some valuable insight as to how fully DWI captures the story of a lesion with respect to the gold standard. This comparison may be especially important in the context of contrast-free screening. As anatomic images are acquired with high spatial resolution, one would expect that the longest diameter measurements would increase in agreement as DWI resolution improves, which is reflected in these results. However, this analysis is inherently limited as the contrast of CE-subtraction images is not always consistent with that of $b = 800\ s/mm^2$ images.

Furthermore, the slice positions did not match exactly across DWI methods as the radiologists were each free to choose a representative slice. Thus, the measurements may not reflect the same region of the lesion.

References

1. Griswold MA, Jakob PM, Heidemann RM, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med 2002;47(6):1202–1210.

2. Le Bihan D. Diffusion and perfusion magnetic resonance imaging: applications to functional MRI. New York, NY: Raven, 1995.

Table E1. Acquisition details

Clinical Protocol		I-SPY 2 Protocol	
Localizer	0:17	Localizer	0:17
T_1-weighted 3D GRE (without fat suppression)	1:24	T_1-weighted 3D GRE (without fat suppression)	1:24
Interactive shimming	~2:00	Interactive shimming	~2:00
T_1-weighted 3D GRE (with fat suppression)	0:32	T_1-weighted 3D GRE (with fat suppression)	0:40
DWI (Standard-B)	3:46	DWI (Standard-A)	4:48
T_1-weighted CE (1 precontrast, 3 postcontrast)	7:52	T_1-weighted CE (1 precontrast, 6 postcontrast)	10:44
T_2-weighted TIRM	4:59	T_2-weighted TIRM	4:59
Spoiled 3D GRE (VIBE)	2:45	Spoiled 3D GRE (VIBE)	2:45
Axially reformatted SMS (scan 1)	4:54*	Axially reformatted SMS (scan 1)	4:54*
Axially reformatted SMS (scan 2)	1:13*	Axially reformatted SMS (scan 2)	1:13*
RS-EPI	4:58	RS-EPI	4:58
TOTAL	**34:40**	**TOTAL**	**38:52**

* AR-SMS was acquired with two additional reference scans because the optimal reconstruction was not known a priori. However, the data used for image reconstruction (including SMS, GRAPPA, and topup reference scans) were acquired in 4:52.

All clinical scans were completed before the advanced DWI, which were performed for research only. Times are listed in min:sec. DWI = diffusion-weighted imaging, EPI = echo-planar imaging, RS = readout-segmented, SMS = simultaneous multislice imaging, GRAPPA = generalized autocalibrating partially parallel acquisitions, CE = contrast enhanced.

Table E2. Subanalysis separating Standard-A and Standard-B. Data were separated into two categories based on what standard SE-EPI protocol was used (Standard-A or Standard-B)

Model	Method Comparison	Effect (95% CI)	*P* value
Model 1–including Standard-A n = 22 lesions	RS-EPI vs Standard-A	0.56 (0.32, 0.80)	<0.001*
	AR-SMS vs Standard-A	1.33 (1.09, 1.58)	<0.001*
	AR-SMS vs RS-EPI	0.77 (0.53, 1.02)	<0.001*
Model 2–including Standard-B n = 8 lesions	RS-EPI vs Standard-B	0.58 (0.18, 0.98)	0.017
	AR-SMS vs Standard-B	1.25 (0.85, 1.65)	<0.001*
	AR-SMS vs RS-EPI	0.67 (0.27, 1.07)	0.005*

* statistically significant

The overall image quality was fit to a linear mixed model, including effects from reader and participant, for each group. Results confirm that the comparison of image quality is consistent across both Standard-A and Standard-B. Asterix indicate statistical significance after Tukey adjustment. RS = readout segmented, EPI = echo-planar imaging, AR = axially reformatted, SMS = simultaneous multislice imaging, ADC = apparent diffusion coefficient, DWI = diffusion-weighted imaging, SE = spin-echo.