Received: 4 January 2019 | Revised: 14 May 2019 | Accepted: 27 May 2019

DOI: 10.1002/hbm.24691

RESEARCH ARTICLE

WILEY

Towards an optimised processing pipeline for diffusion magnetic resonance imaging data: Effects of artefact corrections on diffusion metrics and their age associations in UK Biobank

Ivan I. Maximov[1,2] | Dag Alnæs[2] | Lars T. Westlye[1,2]

[1]Department of Psychology, University of Oslo, Oslo, Norway

[2]Department of Mental Health and Addiction, Norwegian Centre for Mental Disorders Research spiepr132 (NORMENT), Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway

Correspondence
Ivan I. Maximov, Department of Psychology, University of Oslo, Oslo, Norway.
Email: ivan.maximov@psykologi.uio.no

Funding information
Research Concil of Norway, Grant/Award Number: 249795; South-Eastern Norway Regional Health Authority, Grant/Award Numbers: 2014097, 2015073, 2016083

Abstract

Increasing interest in the structural and functional organisation of the human brain encourages the acquisition of big data sets comprising multiple neuroimaging modalities, often accompanied by additional information obtained from health records, cognitive tests, biomarkers and genotypes. Diffusion weighted magnetic resonance imaging data enables a range of promising imaging phenotypes probing structural connections as well as macroanatomical and microstructural properties of the brain. The reliability and biological sensitivity and specificity of diffusion data depend on processing pipeline. A state-of-the-art framework for data processing facilitates cross-study harmonisation and reduces pipeline-related variability. Using diffusion magnetic resonance imaging (MRI) data from 218 individuals in the UK Biobank, we evaluate the effects of different processing steps that have been suggested to reduce imaging artefacts and improve reliability of diffusion metrics. In lack of a ground truth, we compared diffusion metric sensitivity to age between pipelines. By comparing distributions and age sensitivity of the resulting diffusion metrics based on different approaches (diffusion tensor imaging, diffusion kurtosis imaging and white matter tract integrity), we evaluate a general pipeline comprising seven postprocessing blocks: noise correction; Gibbs ringing correction; evaluation of field distortions; susceptibility, eddy-current and motion-induced distortion corrections; bias field correction; spatial smoothing and final diffusion metric estimations. Based on this evaluation, we suggest an optimised processing pipeline for diffusion weighted MRI data.

KEYWORDS
diffusion pipeline, diffusion weighted imaging, UK biobank data

1 | INTRODUCTION

Increasing interest in the role of individual differences in human brain architecture in health and disease has stimulated the neuroscience community to initiate a number of large brain data projects. Due to the attractive combination of increasing availability, low costs, its noninvasive nature and high sensitivity magnetic resonance imaging (MRI) including T_1/T_2-weighted images, functional MRI and diffusion weighted imaging

has become the preferred and standard brain imaging modality in these large efforts, including the UK Biobank (UKB; Miller et al., 2016).

Diffusion MRI is based on the effect of the Brownian motion of water molecules in biological tissue (Basser, Mattiello, & Le Bihan, 1994) and allows one to probe and visualise brain organisation at the micrometre scale (Johansen-Berg & Behrens, 2014). Recent advances in theoretical and experimental diffusion MRI approaches (Novikov, Kiselev, & Jespersen, 2018) have offered various diffusion models and sequences allowing for a detailed description of the signal decay due to water diffusion. Advanced diffusion measurements are technically challenging and optimal data quality places high demands on practical implementation and protocol, including hardware gradient system and coil. Due to limited time and technical constraints, researchers designing imaging studies face various trade-offs, influencing, for example, signal-to-noise ratio (SNR) and options related to the specific pulse sequences such as mono or bipolar diffusion encoding gradients.

Beyond MRI sequence and acquisition parameters, various sources of distortions influence the resulting diffusion metrics, and different approaches for quality control (QC) and corrections have been suggested (Alfaro-Almagro et al., 2018; Esteban et al., 2017; Farzinfar et al., 2013; Hasan, 2007; Oguz et al., 2014). Ideally, the QC methods should reliable to identify and correct typical artefacts originating from subject head motion, discarded volumes and low SNR, which may be particularly present at high diffusion weightings, also known as *b*-values. Despite recent major developments and improvements (Alfaro-Almagro et al., 2018; Cui, Zhong, Xu, He, & Gong, 2013; Miller et al., 2016; Roalf et al., 2016), automated procedures for QC and artefact reduction largely represent unresolved challenges in the imaging community.

Various postprocessing steps have been suggested to correct specific sources of noise and distortions, including thermal noise evaluation (Veraart, Novikov, et al., 2016; Veraart, Fieremans, & Novikov, 2016), Gibbs ringing correction (Kellner, Dhital, Kiselev, & Reisert, 2016; Veraart, Fieremans, Jelescu, Knoll, & Novikov, 2016), susceptibility distortion correction (Andersson & Sotiropoulos, 2016), motion correction (Andersson, Graham, Zsoldos, & Sotiropoulos, 2016; Andersson & Sotiropoulos, 2016), correction of physiological noise and outliers (Maximov et al., 2015; Maximov, Grinberg, & Shah, 2011; Sairanen, Leemans, & Tax, 2018; Walker et al., 2011) and eddy current induced geometrical distortions (Taylor et al., 2016). However, although the application of even part of the postprocessing steps such as noise correction has been demonstrated to improve sensitivity and provide additional information about absolute diffusion metrics (Kochunov et al., 2018), systematic evaluations of the effects of the different steps on the diffusion metrics are scarce.

Several minimal postprocessing pipelines have been recommended to prepare structural, functional and diffusion MRI data (Alfaro-Almagro et al., 2018; Cui et al., 2013; Glasser et al., 2013; Sotiropoulos et al., 2013). For example, the UKB diffusion pipeline first employs fieldmap generation using the anterior–posterior (AP) and posterior–anterior (PA) images of original diffusion data. The selection of AP-PA images is done by an estimation of correlations across AP-PA pairs to find the most accurate reference. Thus, the UKB pipeline include only one diffusion-specific step based on *eddy* (Andersson & Sotiropoulos, 2016; Andersson et al., 2016, 2017), correcting the eddy currents and head motion, susceptibility artefacts and identification and replacement of outlier slices. Providing a comprehensive approach for artefact correction of diffusion data, a recent publication introduced the Diffusion parameter ESTimation with Gibbs and NoisE Removal (DESIGNER) pipeline, which allows one to identify and minimise thermal noise, Gibbs ringing artefacts, Rician noise bias, eddy-current and B_0-induced spatial distortions and motion-related artefacts, and which was demonstrated to improve accuracy of common diffusion MRI metrics (Ades-Aron et al., 2018). Improvements in accuracy and precision of the derived diffusion metrics were evaluated using numerical and diffusion phantoms and a small number of in vivo brain imaging datasets. However, how different pipeline steps influence the associations between diffusion metrics and phenotypes with relevance for studies of individual differences, such as, for instance, age-related differences, is still open. In lack of a ground truth reference sensitivity to age is a relevant criterion for pipeline comparisons, partly due to frequently observed associations between data quality characteristics (e.g., due to subject motion) and age.

With the aim to identify the most efficient and adequate pipeline steps and to assess their influence on diffusion data analysis, we tested the effects of various postprocessing steps on different diffusion scalar metrics, based on diffusion tensor imaging (DTI; Basser et al., 1994), diffusion kurtosis imaging (DKI; Jensen, Helpern, Ramani, Lu, & Kaczynski, 2005) and white matter tract integrity (WMTI; Fieremans, Jensen, & Helpern, 2011) using UKB diffusion MRI data. We assessed the direct influence of pipeline on conventional QC metrics by comparing estimated temporal signal-to-noise ratio (tSNR; Roalf et al., 2016) of the diffusion-weighted volumes between pipelines for each of the two

TABLE 1 Demographic data of the used UK Biobank sample

Subgroups (years)	Number of subjects	Age (mean/std) years	Sex (F/M)
"40"	12	40.40/0.08	6/6
"42"	13	42.08/0.29	6/7
"44"	16	43.92/0.28	8/8
"46"	14	46.01/0.31	7/7
"48"	13	48.00/0.32	7/6
"50"	15	50.06/0.29	8/7
"52"	15	52.07/0.27	7/8
"54"	11	54.11/0.35	7/4
"56"	15	55.97/0.26	7/8
"58"	13	57.98/0.29	6/7
"60"	15	59.99/0.27	8/7
"62"	14	61.95/0.26	7/7
"64"	12	64.11/0.25	7/5
"66"	14	66.11/0.25	8/6
"68"	14	68.08/0.27	7/7
"70"	12	69.81/0.18	6/6
Total	218	54.95/9.09	112/106

TABLE 2 Summary of used and alternative algorithms with links to the numerical implementations

Pipeline step	Algorithm/utility	Link to the software	Comments
Noise correction	*MP-PCA*	https://github.com/NYU-DiffusionMRI/mppca_denoise www.mrtrix.org	Implementations are accessible as original code (Veraart et al.) as well as a part of mrtrix3
Gibbs ringing correction	*Local sub-voxel shift*	https://bitbucket.org/reisert/unring www.mrtrix.org	Implementations are accessible as original code (Kellner et al.) as well as a part of mrtrix3
EPI-geometrical distortions	*TOPUP-EDDY* (FSL)	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki	Modifications of B_0 and eddy-current/motion corrections can be found from other authors
	TORTOISE	https://science.nichd.nih.gov/confluence/display/nihpd/DIFF_PREP+Main	
Bias field corrections	*N4BiasFieldCorrection* (ANTs)	http://stnava.github.io/ANTs/	Any methods can be used here by calling an original utility or over matrix wrapper.
	FAST (FSL)	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki	
Spatial smoothing	*fslmaths* (FSL) *convert3d* (itksnap)	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki http://www.itksnap.org/pmwiki/pmwiki.php?n=Downloads.C3D	There are many different utilities allowing one to do a spatial smoothing
Diffusion metric estimations	*IRLLS* (DKI)	https://github.com/NYU-DiffusionMRI/Diffusion-Kurtosis-Imaging	One can find another code for estimation of the diffusion metrics
	DKE	https://www.nitrc.org/projects/dke/	

Abbreviations: DKI, diffusion kurtosis imaging; MP-PCA, principle component analysis of Marchenko–Pastur.

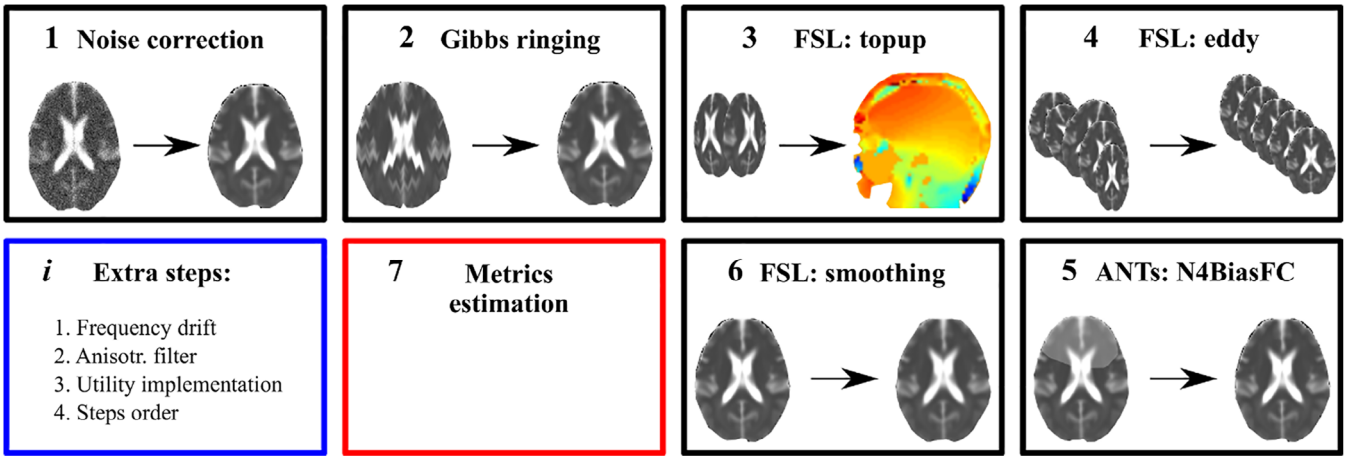

FIGURE 1 Schematic representation of a general pipeline. Numbers in the upper left corner correspond to the step order. The Step 7 is an estimation of final diffusion metrics depending on used diffusion model. The Step *i* presents a possible variability in the pipeline but omitted in the present work, for example, a frequency drift correction, application of different spatial filters (isotropic vs. anisotropic), difference in algorithmic utility implementations (ANTs vs. FSL), permutations in the step orders (Steps 5 vs. 4) [Color figure can be viewed at wileyonlinelibrary.com]

shells (b = 1,000, 2,000 s/mm^2). To assess which degree pipeline influences across-subject analysis and corresponding interpretations, we compared estimated age-curves (Grinberg et al., 2017; Tamnes, Roalf, Goddings, & Lebel, 2017; Westlye et al., 2010; Westlye, Reinvang, Rootwelt, & Espeseth, 2012) of the diffusion metrics between pipelines using voxel-wise analysis based on tract-based spatial statistics (Smith et al., 2006, 2007) and multiple linear regression analysis on diffusion metrics averaged across the TBSS skeleton.

2 | METHODS AND MATERIALS

2.1 | Subjects and data

Table 1 summarises the demographics of the 218 UKB participants included in the present work. We computed diffusion scalar metrics using two different pipelines including various intermediate steps (see Figure 1). An accurate overview of the UKB data acquisition, protocol parameters and image validation can be found elsewhere (Alfaro-Almagro et al., 2018;

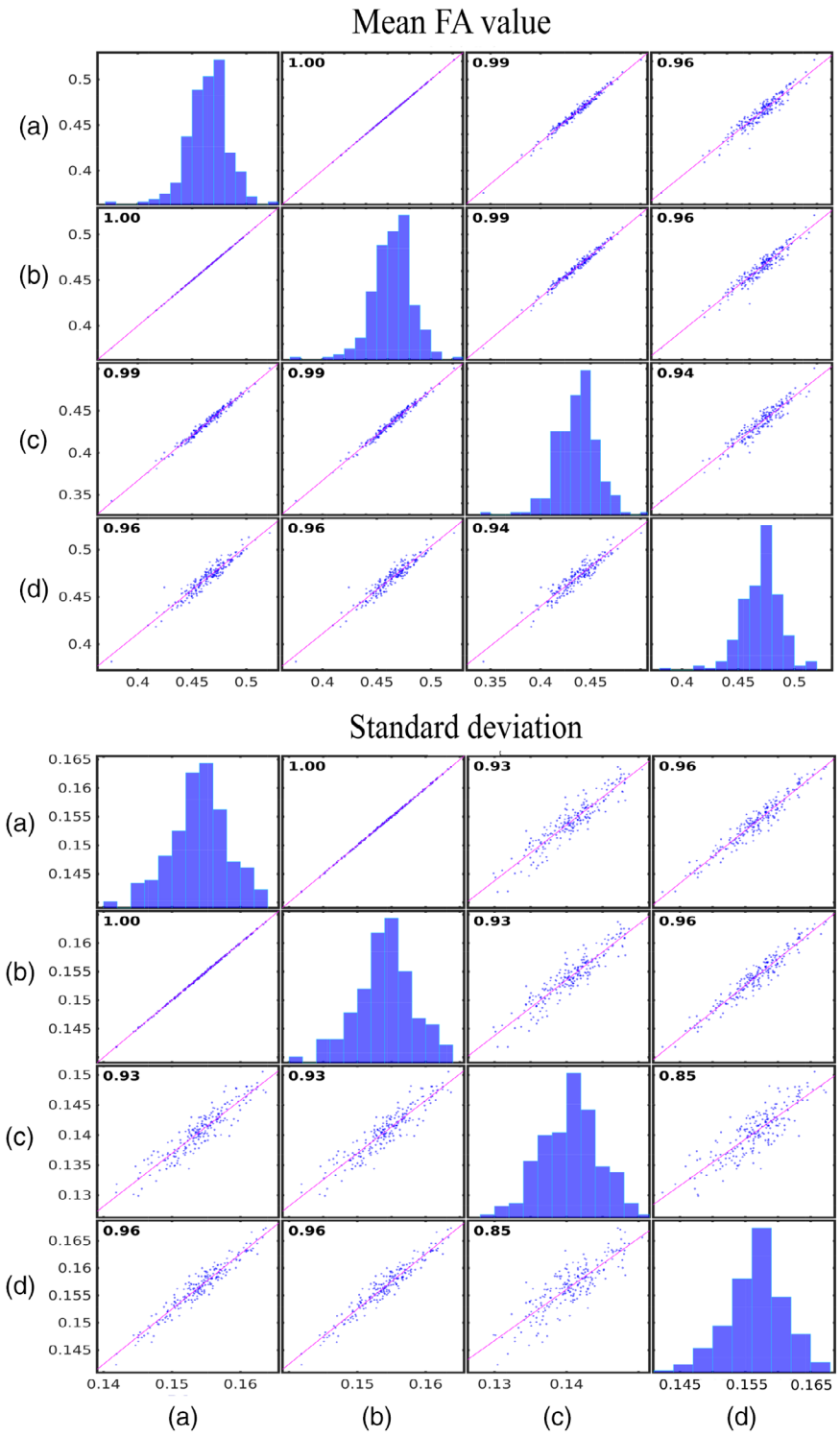

FIGURE 2 Correlation plots for FA based on DKI fitting obtained for four different datasets (see Figure 1) (a) up to Step 5; (b) up to Step 4; (c) up to Step 7; (d) original UK Biobank pipeline. Diffusion metrics were averaged over estimated subject skeletons in the case of each pipeline in accordance with the TBSS preparation pipeline. DKI, diffusion kurtosis imaging; FA, fractional anisotropy [Color figure can be viewed at wileyonlinelibrary.com]

Miller et al., 2016). In brief, the diffusion sequence was a conventional Stejskal–Tanner monopolar spin-echo echo-planar imaging (Stejskal & Tanner, 1965) with multiband factor 3, *b*-values were 1,000 and 2,000 s/mm^2 and 50 noncoplanar diffusion directions per each *b*-shell. The spatial resolution was isotropic 2 mm^3, and 5 AP versus 3 PA images with $b = 0$ s/mm^2 were acquired. All subjects were scanned at a single 3T Siemens Skyra (VD13A SP4) with a standard Siemens 32-channel head coil, in Cheadle Manchester. The original UKB postprocessing pipeline is described in details online (http://biobank.ctsu.ox.ac.uk/crystal/docs/brain_mri.pdf) and includes susceptibility, eddy-current and head motion corrections accompanied with slice outlier detection and replacement, all performed with *topup* and *eddy*.

Figure 1 gives an overview of the current pipeline. We divided the postprocessing flow into seven general blocks. Additional block *i* (marked by blue frame in Figure 1) consists of extra steps allowing one to substitute or extend used steps or algorithms. An advantage of

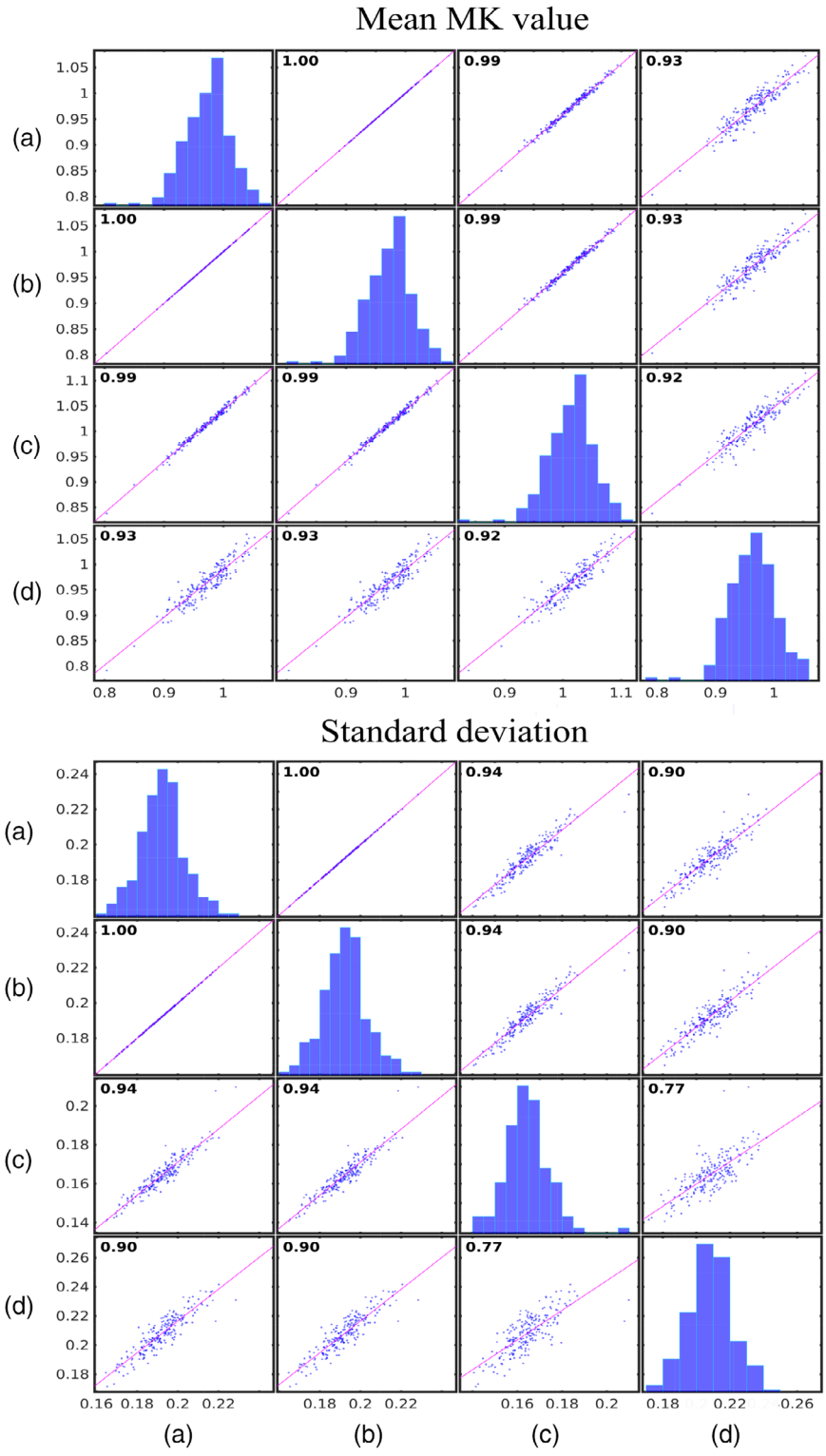

FIGURE 3 Correlation plots for MK based on DKI fitting obtained for four different data sets: (a) up to Step 5; (b) up to Step 4; (c) up to Step 7; (d) original UK Biobank pipeline. Diffusion metrics were averaged over estimated subject skeletons in the case of each pipeline in accordance with the TBSS preparation pipeline. DKI, diffusion kurtosis imaging; MK, mean kurtosis [Color figure can be viewed at wileyonlinelibrary.com]

the proposed pipeline is freely accessible code for all processing steps. Table 2 summarises possible alternatives and links to the software implementations for each pipeline step. Whereas the current assessment of the pipeline was based on UK Biobank only, we assume that the recommendations generalise to other diffusion MRI data sets with conventional acquisition parameters. Below we briefly describe each step in the suggested order.

2.2 | Noise correction

The noise in diffusion data is spatially dependent in the case of multichannel receive coils (Aja-Fernandez, Vegas-Sanchez-Ferrero, & Tristan-Vega, 2014; Andre et al., 2014; Maximov, Farrher, Grinberg, & Shah, 2012). Principle component analysis of Marchenko–Pastur (MP-PCA) noise-only distribution provides an accurate and fast method of noise

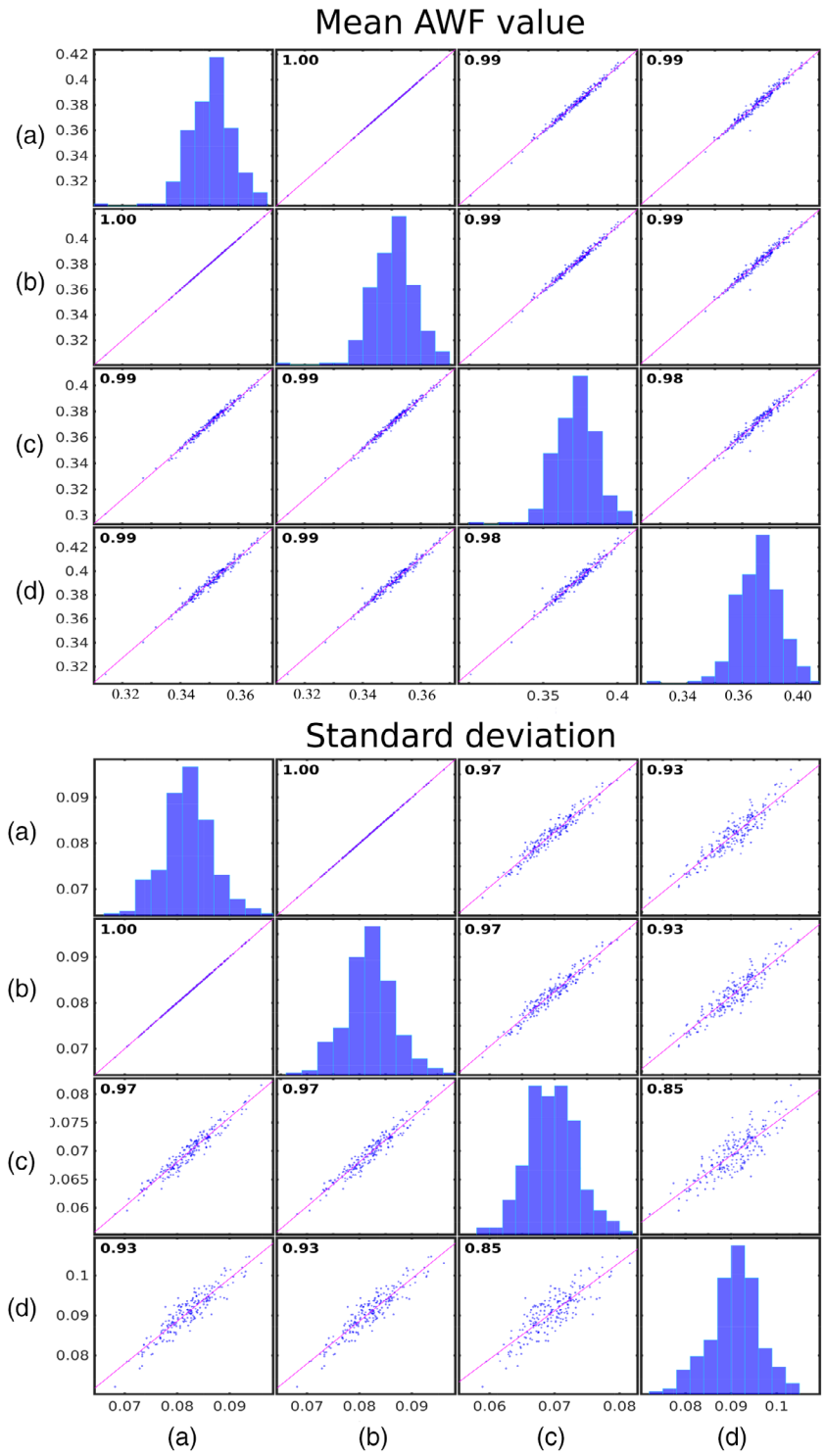

FIGURE 4 Correlation plots for AWF based on WMTI fitting obtained for four different data sets: (a) up to Step 5; (b) up to Step 4; (c) up to Step 7; (d) original UKB pipeline. Diffusion metrics were averaged over estimated subject skeletons in the case of each pipeline in accordance with the TBSS preparation pipeline. AWF, axonal water fraction; WMTI, white matter tract integrity [Color figure can be viewed at wileyonlinelibrary.com]

evaluation and reduction (Veraart, Fieremans, & Novikov, 2016; Veraart, Novikov, et al., 2016). The thermal noise correction using the MP-PCA method should be the first step in data analysis due to an assumption about uncorrelated noise both spatially and across the diffusion space. In the present work, we used the original Veraart's MATLAB code (The Mathworks, Natick, MA): https://github.com/NYU-DiffusionMRI/mppca_denoise. Note that the noise correction methods are regularly improved and in future application the current implementation may be substituted by a more efficient approach (see e.g., Aja-Fernandez et al., 2014; Manjon et al., 2015). The noise in MR images can be described by a Rician distribution. To avoid a possible bias affecting the diffusion data, we use MP-PCA estimated standard deviation at each voxel and an analytical approach developed by Koay and Basser (2006) for evaluation of the true signal (Ades-Aron et al., 2018).

2.3 | Gibbs ringing correction

Various artefacts appearing in the raw data due to table vibration (Gallichan et al., 2010), radio-frequency-based distortions, incorrect

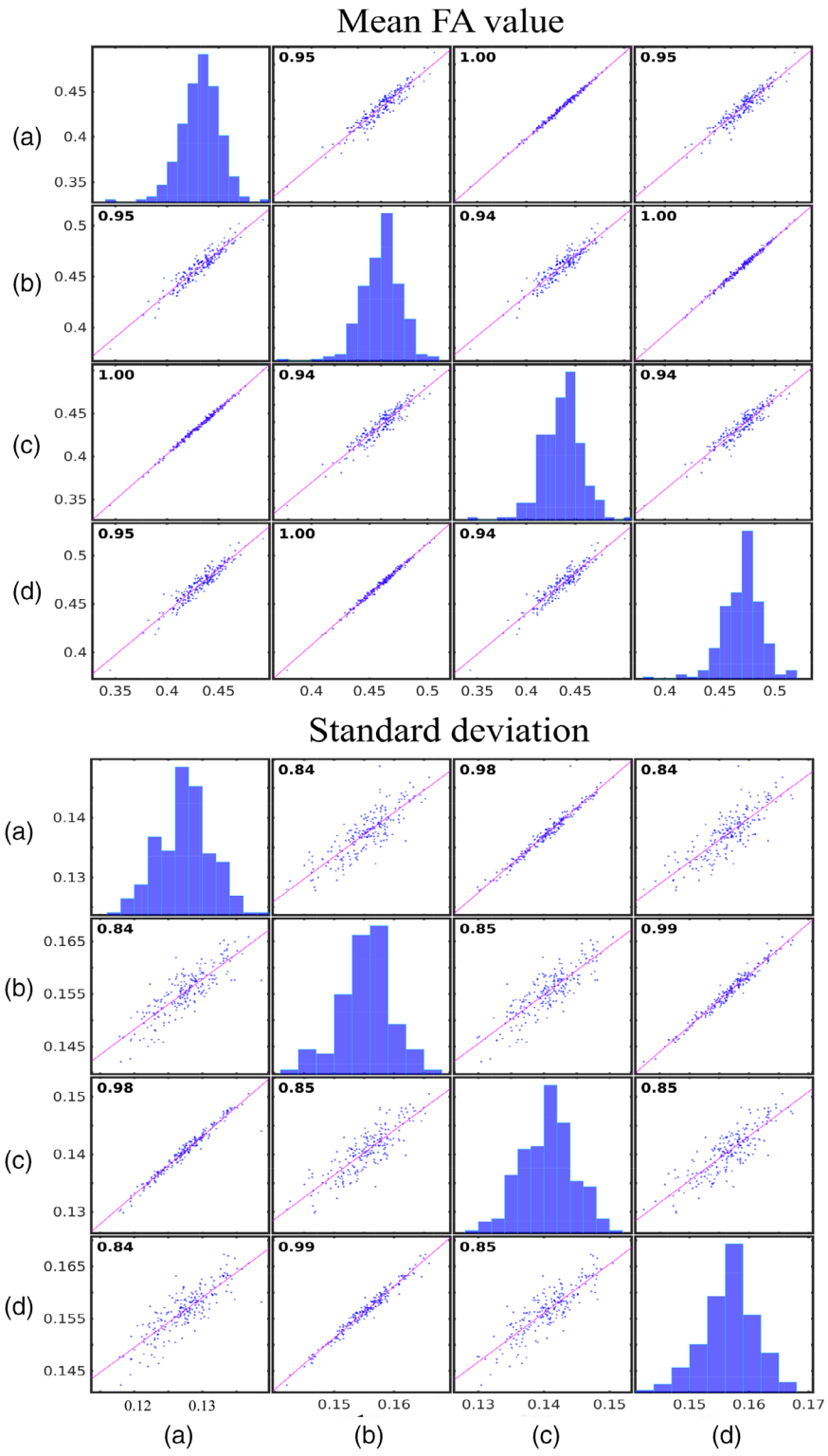

FIGURE 5 Correlation plots for FA based on conventional DTI fitting (1 *b*-shell) obtained for four different data sets: (a) up to Step 5; (b) up to Step 4; (c) up to Step 7; (d) original UKB pipeline. Diffusion metrics are averaged over estimated subject skeletons in the case of each pipeline in accordance with TBSS preparation pipeline. FA, fractional anisotropy; DTI, diffusion tensor imaging; UKB, UK Biobank [Color figure can be viewed at wileyonlinelibrary.com]

magnetic field gradient calibration (McRobbie, Moore, Graves, & Prince, 2006) can significantly degrade the diffusion data. One of the most frequent artefacts is known as the Gibbs ringing artefact. This appears due to a *k*-space truncation along finite image sampling and can be suppressed by post hoc methods (Kellner et al., 2016; Perrone et al., 2015; Veraart, Fieremans, Jelescu, et al., 2016). Here, we used the approach developed by Kellner et al. (2016) and the original MATLAB code: https://bitbucket.org/reisert/unring.

2.4 | EPI distortions

Diffusion data acquisition is based on echo-planar imaging, which is susceptible to multiple distortions originating from magnetic field inhomogeneity. A few approaches have been developed to correct field inhomogeneities: a simple and robust method based on field mapping; a method based on evaluation of point spread function; and reversed gradient approach (Wu et al., 2008). *FSL* (Smith et al., 2004) offers an excellent

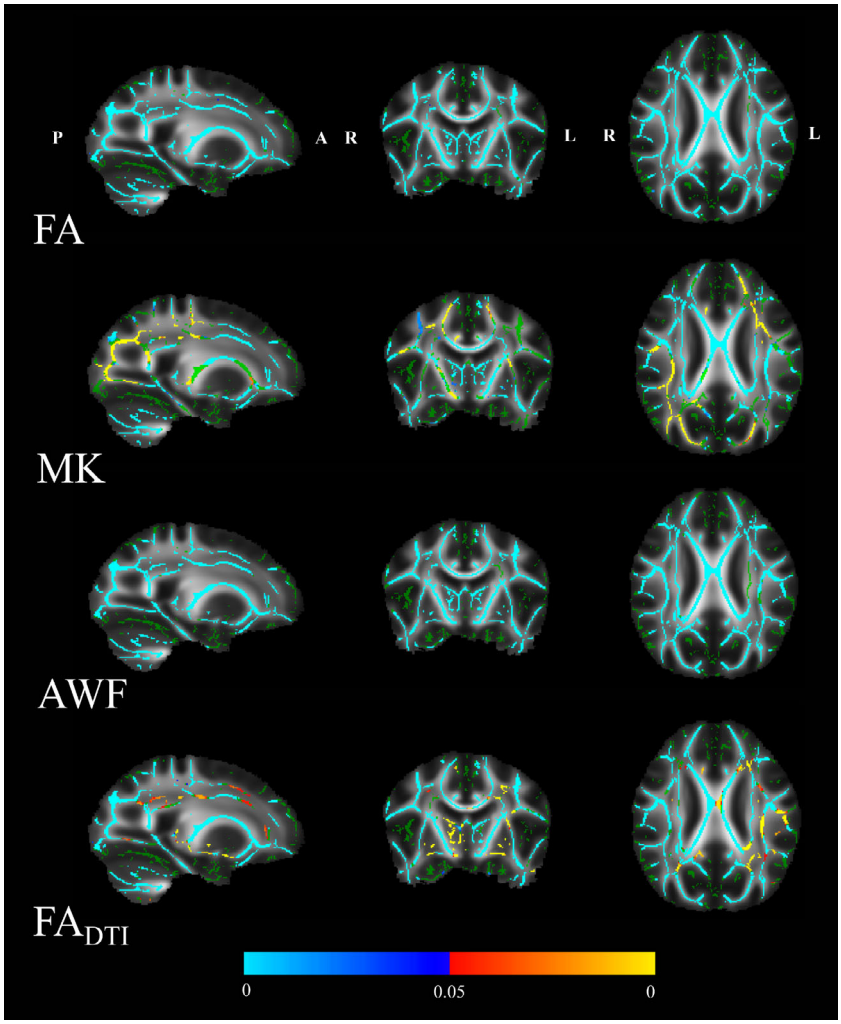

FIGURE 6 Results of TBSS analysis between the pipelines: Original UKB and the proposed here (S7). TBSS analysis for diffusion metrics based on FA, MK and AWF (2 *b*-shells); TBSS analysis for diffusion metric based on FA_{DTI} fitting (1 *b*-shell). All images are in standard MNI space and correspond to the coordinates: $x = 26$; $y = -8$; $z = 24$. The red-yellow colour means that metrics from S7 pipeline are significantly higher than from UKB ($p < .05$); the blue-light-blue colour means an opposite situation. The presented TBSS results are TFCE corrected. ANTs, Advanced Normalization Tools; AWF, axonal water fraction; FA, fractional anisotropy; FSL, FMRIB Software Library; MK, mean kurtosis; MNI, Montreal Neurological Institute; TBSS, tract-based spatial statistics; UKB, UK Biobank [Color figure can be viewed at wileyonlinelibrary.com]

utility for the EPI geometric distortion correction (*topup*, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/topup; Andersson, Skare, & Ashburner, 2003). *Topup* requires data with opposite phase-encoding directions for the nondiffusion weighted images, for example, anterior–posterior and posterior–anterior pair or left–right and right–left pair.

2.5 | Motion, eddy current and susceptibility distortion correction

Topup and *eddy* works together for correcting distortions appeared due to eddy currents, head motion and susceptibility originated artefacts. The GPU accelerated version of *eddy* (*eddy_cuda*) allows one to significantly speed up the computations as well as providing additional options such as in slice alignments, improved outlier detection and multiband dataset estimations (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy/UsersGuide; Andersson et al., 2016, 2017; Andersson & Sotiropoulos, 2016).

2.6 | Field nonuniformity

MR images possess a low frequency intensity shift appearing as intensity inhomogeneity over the image. Several studies have evaluated its influence on the intrasubject and intersubject reproducibility of T_1-weighted structural MRI data (Banerjee & Maji, 2015; Ganzetti, Wenderoth, & Mantini, 2016), but less has been published regarding effects of nonuniformity correction on diffusion data. To avoid bias based on the field nonuniformity, we applied a bias field correction for $b = 0$ s/mm^2 image. Then, the estimated field map was applied to all diffusion images to decrease the field inhomogeneity. We used the *N4BiasFieldCorrection* utility from ANTs (Tustison et al., 2010). The order of the bias field correction step is discussed below.

2.7 | Spatial smoothing

After all the above mentioned steps, the diffusion data, in theory, are ready for the estimation of diffusion scalar metrics. To increase SNR, which may be particularly beneficial for the numerical stability of advanced diffusion models (Maximov, Tonoyan, & Pronin, 2017; Maximov & Vellmer, 2019; Vellmer, Tonoyan, Suter, Pronin, & Maximov, 2018), we applied spatial smoothing of the raw diffusion data. For simplicity, we used isotropic smoothing with a Gaussian kernel 1 mm^3 implemented in the FSL function *fslmaths*.

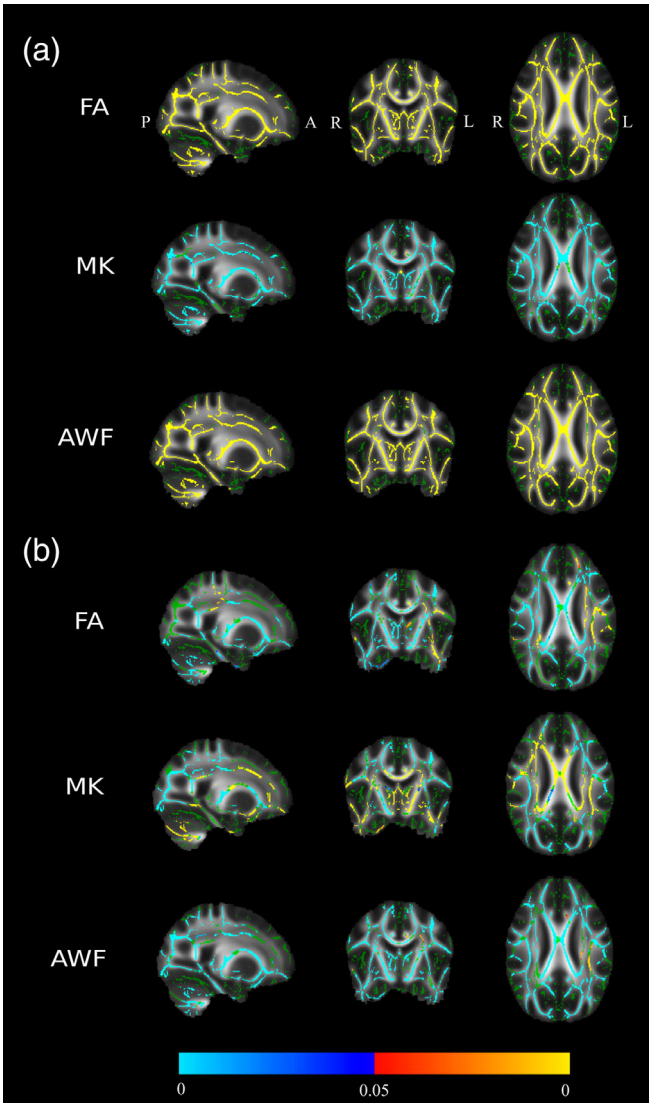

FIGURE 7 Results of TBSS analysis between data sets produced by (a) S5 and S7; (b) S5 and original UKB pipeline. All images are in standard MNI space and correspond to the coordinates: $x = 26$; $y = -8$; $z = 24$. The red-yellow colour means that the first pipeline is significantly higher than the second one ($p < .05$); the blue-light-blue colour means an opposite situation. The presented TBSS results are TFCE corrected. AWF, axonal water fraction; FA, fractional anisotropy; MK, mean kurtosis; TFCE, threshold-free cluster enhancement; UKB, UK Biobank [Color figure can be viewed at wileyonlinelibrary.com]

2.8 | Metric estimation

UKB diffusion data acquisition was done using a multishell protocol with b = 1,000 and 2,000 s/mm^2 in addition to b = 0 s/mm^2. We based our evaluation on various diffusion metrics derived using three different approaches: Conventional DTI (Basser et al., 1994), namely, fractional anisotropy (FA), mean, axial and radial diffusivity (MD, AD and RD, respectively); DKI (Jensen et al., 2005) with FA, MD, AD, RD, mean, axial and radial kurtosis (MK, AK, RK, respectively); and WMTI (Fieremans et al., 2011) metrics with axonal water fraction (AWF), extra-axonal axial and radial diffusivities (AE and RE) and tortuosity (Tort). These metrics are based on a cumulant expansion of the diffusion propagation function, that is, strictly speaking, they do not represent a comprehensive diffusion biophysical model (Novikov et al., 2018). Nevertheless, these maps are very popular and easy to obtain in clinical studies. For DKI, we used an approach proposed by Veraart, Sijbers, Sunaert, Leemans, and Jeurissen (2013) and the original MATLAB code (https://github.com/NYU-DiffusionMRI/Diffusion-Kurtosis-Imaging). The DTI metrics were estimated using *DTIFIT* in FSL, by means of a linear weighted least squares option in command line for the shell b = 1,000 s/mm^2. We assume that the original UKB DTI metrics were estimated with the same option, although it was not mentioned in the description (Miller et al., 2016).

2.9 | Additional options

Some of the steps can be substituted by other approaches or implementations. For example, nonuniformity field corrections used in functional MRI and brain tissue segmentation may increase the accuracy of the motion correction (Ganzetti et al., 2016). The applied isotropic spatial filtering even with a quite small Gaussian kernel introduces blurring of tissue borders and increase partial voluming. A classical anisotropic diffusion filter based on the Perona–Malik algorithm (Perona & Malik, 1990) may provide an alternative with less blurring (Van Hecke et al., 2010; Vellmer et al., 2018). Therefore, we suggest to carefully consider the influence of different degradation factors on the diffusion image quality and to choose a reliable and robust tool for the correction step tailored to the study (see, e.g., considerations related to neonatal neuroimaging: Bastiani et al., 2019).

2.10 | Temporal SNR

To quantify the effects of the pipelines using a conventional data quality metric, we estimated tSNR (Roalf et al., 2016) for each pipeline, which allows one to present a single metric characterising the whole brain diffusion weighted data set and to perform comparative estimations of data quality (Tønnesen et al., 2018). The average temporal SNR is defined as $\text{tSNR} = \text{mean}_{\text{volume}}(\text{mean}_{\text{voxel}}/\text{std}_{\text{voxel}})$, along diffusion dimension. For each shell (b = 1,000 and b = 2,000 s/mm^2), we calculated temporal SNR after steps S4, S5, S7 of the developed pipeline and the UKB pipeline (see below).

2.11 | Statistical analysis

To compare diffusion metrics obtained using different pipelines, we used TBSS (Smith et al., 2006). Initially, all volumes were aligned to the FMRI58_FA template, supplied by FSL, using nonlinear transformation realised by FNIRT (Andersson, Jenkinson, & Smith, 2007). Next, a mean FA image of all subjects was obtained and thinned to create mean FA skeleton. Afterward, the maximal FA values for each subject were projected onto the skeleton to minimise confounding effects due to partial voluming and any residual alignment problems.

We performed voxel-wise comparisons between diffusion metrics obtained from the different pipelines using general linear models (GLMs). For simplicity, we used individual level difference maps (UKB—S7) when comparing pipelines. For each pipeline, we also

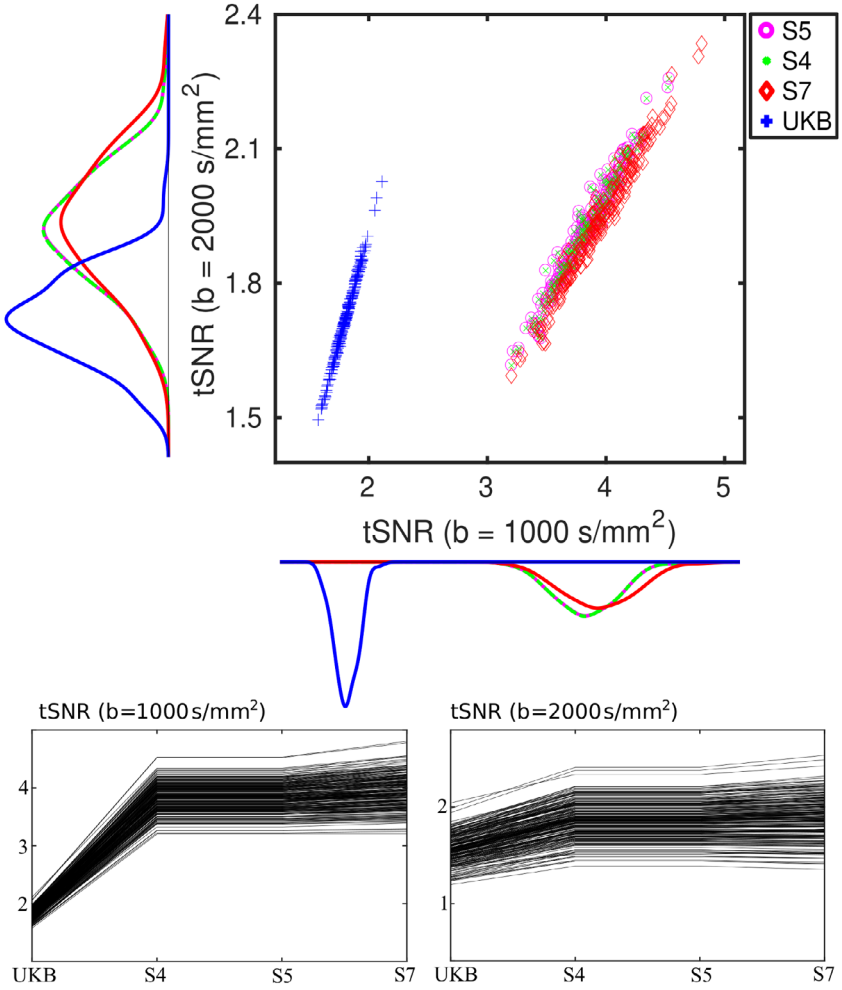

FIGURE 8 Scatter plots of different postprocessing steps for evaluated tSNR values using *b*-values equal to 1,000 (*x* axis) and 2,000 (*y* axis) s/mm^2. tSNR, temporal signal-to-noise ratio; UKB, UK Biobank [Color figure can be viewed at wileyonlinelibrary.com]

tested for associations with age using GLMs, including sex as covariate. For all contrasts, statistical analysis was performed using permutation-based inference implemented in *randomise* (Winkler, Ridgway, Webster, Smith, & Nichols, 2014) with 5,000 permutations. Threshold-free cluster enhancement (TFCE, Smith & Nichols, 2009) was used. Statistical p value maps were thresholded at $p < .05$ corrected for multiple comparisons across space.

In addition to voxel-wise statistics, we used diffusion metrics averaged over the skeletons for estimating age differences using linear models, as well as for visualisation of age curves and differences between pipelines using the *corrplot* function in MATLAB. Linear regressions were performed using two models: Model 1: $y = b_0 + b_1{*}Age + b_2{*}Sex$ and Model 2: $y = b_0 + b_1{*}Age + b_2{*}Age^2 + b_3{*}Sex$, where *Age* is age in years and *Sex* corresponds to male or female. These two models allowed us to test for linear and quadratic associations between diffusion metrics and age for each pipeline. Regression parameters were estimated using the MATLAB function *fitlm* with a robust estimator based on the Welsch function (Holland & Welsch, 1997) to decrease the influence of possible outliers. Comparison of the variance between groups was performed using repeated measures ANOVA test as implemented in the MATLAB function *ranova* and standard deviations estimated over the skeleton for each subject. Comparison of regression parameters for age (slopes) from the linear models for each pipeline was done using the R package *cocor* (Diedenhofen & Musch, 2015).

3 | RESULTS

3.1 | Voxel-wise comparisons of diffusion metrics between pipelines

Figure 2 shows the scatter plots for FA obtained from the DKI fitting and S4, S4, S7 and UKB pipelines. Top correlation plot corresponds to mean skeleton FA, the bottom plot represents the individual standard deviations across the skeleton. Briefly, the results revealed high correlations of the diffusion metrics between all pipelines, however, ANOVA test revealed significantly ($p < 10^{-30}$) lower variance in S7 pipeline (histogram peak values and mean/std are at MD—0.18 (0.16/0.03); AD—0.29 (0.30/0.03); RD—0.15 (0.17/0.03), see Figure S1) compared to the original UKB pipeline (histogram peak values and mean/std are at MD—0.21 (0.24/0.05); AD—0.35 (0.35/0.03); RD—0.23 (0.24/0.04), see Figure S1). Figure 3 shows the correlation plots for MK from the DKI model. The metrics were highly correlated between pipelines. The variances were also lower (ANOVA test, $p < 10^{-30}$) in the S7 pipeline compared to UKB (see Figure S2). Figure 4 shows the correlation plots

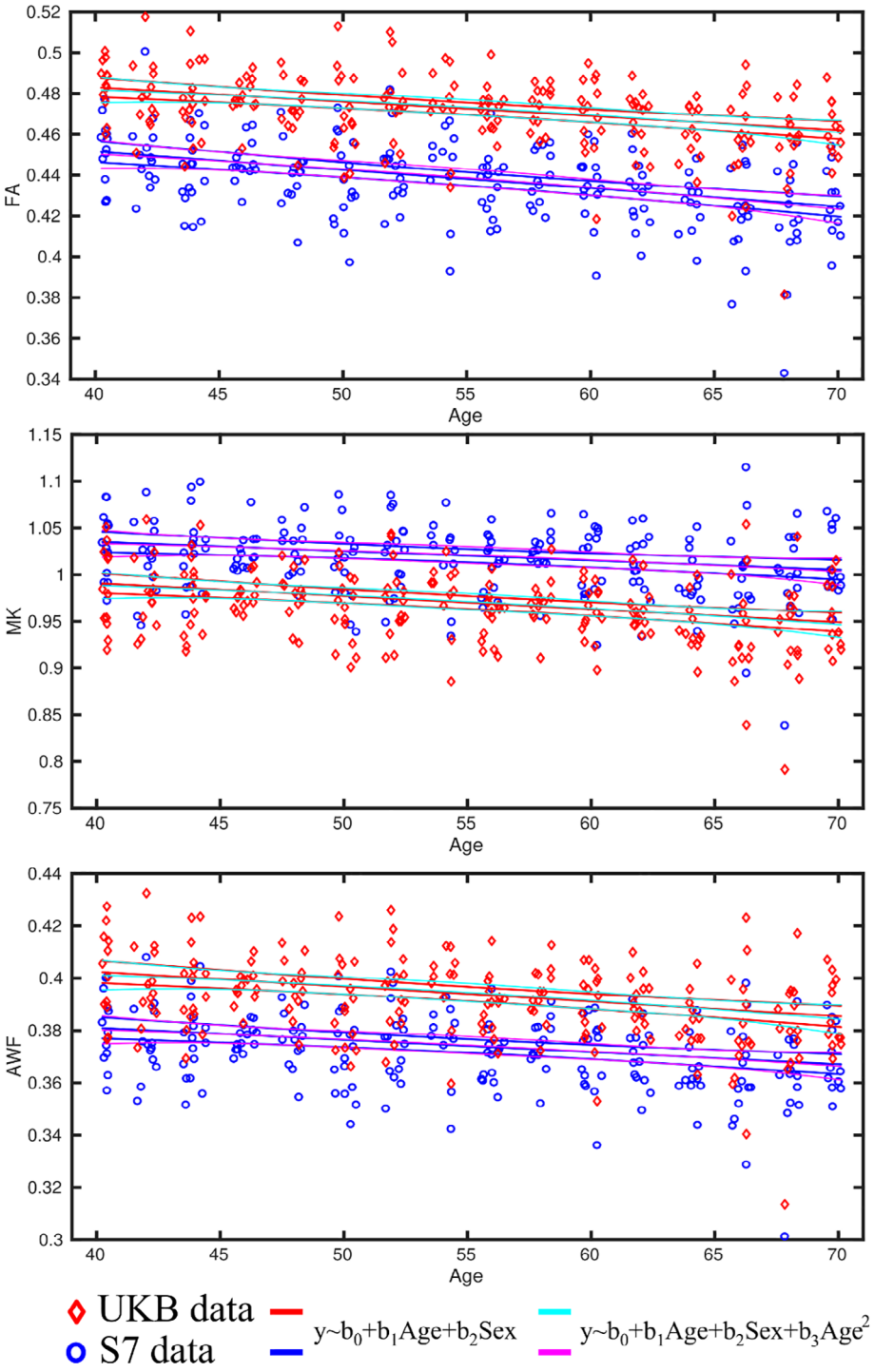

FIGURE 9 Linear age correlations of diffusion metrics obtained from two pipelines (S7 and original UKB). Regressions were performed for two linear models with included Sex as a covariant. Regression lines are plotted together with a 95% interval of confidence. AWF, axonal water fraction; FA, fractional anisotropy; MK, mean kurtosis; UKB, UK Biobank [Color figure can be viewed at wileyonlinelibrary.com]

of AWF from the WMTI model. Since the estimation of WMTI metrics were based on the DKI values, the WMTI diffusion metrics exhibited quite high correlations for all pipelines similar to Figure 3. The variance of S7 pipeline was lower (ANOVA test, $p < 10^{-30}$) compared to all other pipelines (histogram peak values and mean/std are at AWF—0.07 (0.07/0.004); AE—0.40 (0.41/0.02); RE—0.18 (0.19/0.03); Tort—0.65 (0.66/0.07); see Figure S3). Figure 5 shows the scatter plots for FA based on single-shell (b = 1,000 s/mm^2) DTI, suggesting similar relationships between the pipelines as for FA based on the DKI models, with lower variance in the S7 pipeline compared to the original UKB pipeline. Similar results using other conventional DTI metrics (MD, AD and RD) can be seen in Figure S4.

Figure 6 shows the results from the voxel-wise comparison between the original UKB pipeline and S7. Both DKI/WMTI and DTI revealed significant differences ($p < .05$, corrected using permutation testing and TFCE) between pipelines, where the S7 pipeline metrics revealed both higher and lower values compared to UKB pipeline, in particular, see MK and conventional FA. Figure 7 shows the results of the analysis based on S5 versus UKB and S5 versus S7. Briefly, the results revealed significant difference ($p < .05$, corrected using

TABLE 3 Estimated regression intercepts/slopes/root mean square error (RMSE), and R^2 for two linear models (Model 1: $y = b_0 + b_1{*}Age + b_2{*}Sex$; Model 2: $y = b_0 + b_1{*}Age + b_2{*}Age^2 + b_3{*}Sex$) in age-curves in Figure 8

Pipeline/model		FA	MD	AD	RD
	Model 1				
	Slope (b_1)	$-0.895{\cdot}10^{-3}$	$1.525{\cdot}10^{-3}$	$1.055{\cdot}10^{-3}$	$1.774{\cdot}10^{-3}$
	Intercept (b_0)	0.487	0.799	1.284	0.555
S7	RMSE	0.017	0.027	0.028	0.030
	R^2	0.175	0.214	0.142	0.227
	Model 2				
	Slope (b_0)	$-0.905{\cdot}10^{-3}$	$1.561{\cdot}10^{-3}$	$1.075{\cdot}10^{-3}$	$1.811{\cdot}10^{-3}$
	Intercept (b_1)	0.488	0.797	1.283	0.553
	RMSE	0.017	0.027	0.028	0.030
	R^2	0.173	0.216	0.142	0.229
	Model 1				
	Slope (b_1)	$-0.696{\cdot}10^{-3}$	$1.256{\cdot}10^{-3}$	$0.667{\cdot}10^{-3}$	$1.503{\cdot}10^{-3}$
	Intercept (b_0)	0.511	0.818	1.351	0.552
	RMSE	0.015	0.024	0.024	0.027
UKB	R^2	0.140	0.188	0.087	0.209
	Model 2				
	Slope (b_1)	$-0.701{\cdot}10^{-3}$	$1.274{\cdot}10^{-3}$	$0.688{\cdot}10^{-3}$	$1.518{\cdot}10^{-3}$
	Intercept (b_0)	0.511	0.817	1.350	0.551
	RMSE	0.015	0.025	0.025	0.027
	R^2	0.136	0.186	0.087	0.206
		MK	AK	RK	
	Model 1				
	Slope (b_1)	$-0.995{\cdot}10^{-3}$	$-0.586{\cdot}10^{-3}$	$-1.978{\cdot}10^{-3}$	
	Intercept (b_0)	1.075	0.810	1.493	
	RMSE	0.037	0.019	0.068	
S7	R^2	0.06	0.106	0.063	
	Model 2				
	Slope (b_1)	$-1.000{\cdot}10^{-3}$	$-0.591{\cdot}10^{-3}$	$-1.989{\cdot}10^{-3}$	
	Intercept (b_0)	1.075	0.811	1.493	
	RMSE	0.037	0.019	0.068	
	R^2	0.056	0.101	0.059	
	Model 1				
	Slope (b_1)	$-1.395{\cdot}10^{-3}$	$-0.663{\cdot}10^{-3}$	$-2.600{\cdot}10^{-3}$	
	Intercept (b_0)	1.047	0.774	1.487	
	RMSE	0.036	0.018	0.068	
UKB	R^2	0.115	0.126	0.106	
	Model 2				
	Slope (b_1)	$-1.405{\cdot}10^{-3}$	$-0.678{\cdot}10^{-3}$	$-2.606{\cdot}10^{-3}$	
	Intercept (b_0)	1.047	0.777	1.487	
	RMSE	0.036	0.018	0.069	
	R^2	0.111	0.128	0.101	
		AWF	AE	RE	
	Model 1				
	Slope (b_1)	$-0.464{\cdot}10^{-3}$	$0.616{\cdot}10^{-3}$	$2.158{\cdot}10^{-3}$	

(Continues)

TABLE 3 (Continued)

Pipeline/model		FA	MD	AD	RD
	Intercept (b_0)	0.400	1.820	0.852	
	RMSE	0.013	0.036	0.035	
S7	R^2	0.091	0.048	0.239	
	Model 2				
	Slope (b_1)	$-0.467 \cdot 10^{-3}$	$0.625 \cdot 10^{-3}$	$2.190 \cdot 10^{-3}$	
	Intercept (b_0)	0.400	1.819	0.851	
	RMSE	0.014	0.036	0.035	
	R^2	0.087	0.045	0.239	
	Model 1				
	Slope (b_1)	$-0.566 \cdot 10^{-3}$	$0.110 \cdot 10^{-3}$	$1.435 \cdot 10^{-3}$	
	Intercept (b_0)	0.425	1.937	0.863	
	RMSE	0.014	0.034	0.028	
UKB	R^2	0.118	0.013	0.176	
	Model 2				
	Slope (b_1)	$-0.569 \cdot 10^{-3}$	$0.110 \cdot 10^{-3}$	$1.456 \cdot 10^{-3}$	
	Intercept (b_0)	0.425	1.937	0.862	
	RMSE	0.014	0.034	0.029	
	R^2	0.114	0.011	0.175	

Abbreviations: AD, axial diffusivity; FA, fractional anisotropy; MD, mean diffusivity; RD, radial diffusivity; UKB, UK Biobank.

permutation tests and TFCE) between S5 versus S7, and S5 versus UKB. Interestingly, that analysis of S5 versus UKB reproduces similar patterns as for S7 versus UKB (see Figure 6).

3.2 | Temporal signal-to-noise ratio

Figure 8 shows the tSNR distributions for each pipeline by scatter plots for b = 1,000 and 2,000 s/mm^2. For b = 1,000 s/mm^2 mean estimated tSNR (std) = 3.83 (0.25), 3.83 (0.25), 3.94 (0.29) and 1.81 (0.10) for S4, S5, S7 and UKB, respectively. For b = 2,000 s/mm^2 mean estimated tSNR (std) = 1.92 (0.12), 1.92 (0.12), 1.93 (0.14) and 1.73 (0.10). These results indicate 2.2 times (for b = 1,000 s/mm^2) and 1.1 time (for b = 2,000 s/mm^2) higher tSNR in the S7 pipeline compared to the UKB pipeline.

3.3 | Age-related differences across pipelines

Figure 9 shows the estimated linear fits with age for the various diffusion metrics and Table 3 shows the summary stats from the regression models, including the intercepts and slopes. Cocor revealed no significant differences in the estimated slopes between pipelines (in all cases $z < 0.005$, $p > .99$). The DTI metrics exhibited expected age-related differences with lower FA and higher MD, AD and RD with higher age. MK, AK and RK showed age-related reductions, that is, reduced non-Gaussianity of the water diffusion with increased age. Metrics based on WMTI (AWF, AE and RE) demonstrated reduction of the AWF and extension of the extra-axonal water diffusivity with increased age for both regression models. Note, that in the regression Model 2, the associations between diffusion metrics and age^2 and sex were not significant ($p > .05$) and the total explained variance in Model 2 was very similar to Model 1. Estimated linear fits with age for MD, AD, RD, AK, RK, AE and RE metrics are presented in Figure S8.

Figure 10 shows the results from the voxel-wise GLM testing for linear age associations with the different diffusion metrics across the skeleton, and a scatter plot of the uncorrected voxel-wise t-stats from each of the two pipelines. The spatial correlation between the uncorrected t-stats maps were 0.84, 0.79, 0.84 for FA, MK and AWF, respectively (scatter plots are presented in Figure 9c). Table 4 summarises the number of voxels identified in each pipeline. Whilst the two pipelines revealed highly overlapping results, a few regions showed significant age differences in one of the two pipelines only. For example, for MK and AWF S7 identified significant ($p < .05$, corrected) correlations in splenium and genu, and UKB revealed significant age-associations for MK in occipital region of the skeleton.

4 | DISCUSSION

A growing interest in utilising advanced diffusion MRI to study the human brain motivated us to test the effects of various data processing pipelines on different diffusion metrics. Differences in data post-processing steps such as the current S4, S5 and S7 are likely to influence reliability and subsequent interpretation of results. Thus, a harmonised diffusion pipeline may prove valuable for increasing sensitivity, reliability

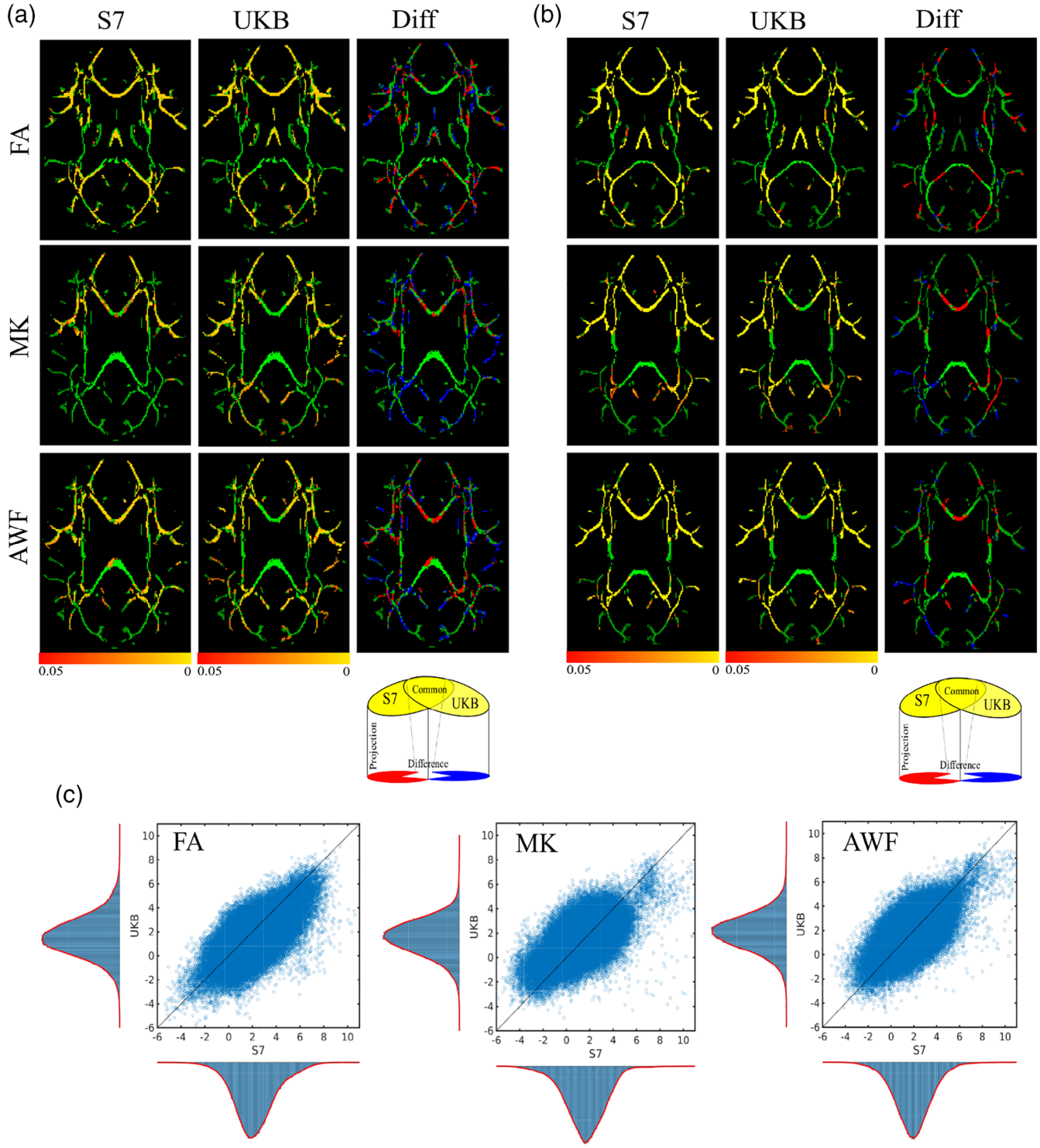

FIGURE 10 Results of general linear model (GLM) tests of diffusion metric versus age across the skeleton. "Diff" columns visualise the spatial difference between the GLM results: The red colour marked the regions with significant difference ($p < .05$) detected in S7 pipeline but not in UKB. Blue voxels showed significant age correlation in UKB pipeline but not in S7. The mean skeleton is visualised by the green colour. The presented TBSS results are TFCE corrected. (a) Voxel-wise analysis was performed using individual skeletons derived for each pipeline separately; (b) voxel-wise analysis was performed using common skeleton derived for the merged datasets; (c) scatter plots of *t*-stats for FA, MK and AWF derived using the merged data sets. AWF, axonal water fraction; FA, fractional anisotropy; MK, mean kurtosis; TFCE, threshold-free cluster enhancement; UKB, UK Biobank [Color figure can be viewed at wileyonlinelibrary.com]

and generalisability across studies. We suggest a general framework with the following postprocessing steps: (a) noise correction, (b) Gibbs-ringing correction, (c) field mapping, (d) susceptibility, eddy current and head motion distortion corrections, (e) B_1 field correction, (f) spatial smoothing and (g) final metrics estimation. Our comparison between three post-processing steps in the current diffusion pipeline demonstrated that the

TABLE 4 The number of voxels (N_{vox}) depending on the pipeline with significant age-correlation specific for the given pipeline versus a total number of voxels with significant age correlation ($p < .05$, corrected)

Pipeline specific/total N_{vox}	Individual skeletons: skeleton size (N_{S7}/N_{UKB}) 117,623/132,025			Common skeleton: skeleton size N_{vox} 120,958		
	FA	MK	AWF	FA	MK	AWF
S7	35,151/70,335	10,816/26,896	21,452/52,974	29,130/69,152	12,691/47,408	14,752/62,434
UKB	23,744/58,928	43,297/59,377	38,180/69,702	5,995/45,911	20,508/55,225	15,130/62,812

Note: Voxel estimations were performed for analyses based on individual skeletons derived by each pipeline and for the skeleton derived from the merged data sets. See also Figure 9 for explanation.
Abbreviations: AWF, axonal water fraction; FA, fractional anisotropy; MK, mean kurtosis; UKB, UK Biobank.

general pipeline suggested here yield a substantially higher tSNR compared to the original UKB pipeline, and also influence the estimated age curves with potentially relevant implications.

Overall, the diffusion metrics derived after the different steps in the current pipeline demonstrated high correlations and similar distributions. In some cases, S7 resulted in lower variance of the diffusion metrics than others, for example, for WMTI. Although we interpret the reduced variance to indicate higher precision in the current context, it should be emphasised that lower variance does not necessarily indicate higher accuracy of the diffusion metric estimation. Nevertheless, S7 exhibited quite high correlations with S5 for the conventional DTI metrics. This strong correlation suggests a small effect of the Gaussian smoothing on the metric estimations to additional data interpolations introduced by the Gibbs ringing (S2) and eddy corrections (S4) and can be interpreted as a metric shift. Results from the UKB pipeline showed relatively high correlations with S4 and S5 for the DTI metrics, and slightly lower for DKI and WMTI. For DTI the UKB pipeline showed stronger correlations with S4 than to S5. The correlations between the S7 and UKB pipelines were lower than for other pipeline pairs. Overall, the results support that all steps of the proposed S7 pipeline might lead to relevant improvements in the estimations of diffusion metrics.

TBSS revealed a significant difference between the S7 and UKB pipelines for all diffusion metrics. Interestingly, the differences between pipelines did not reflect simple global shifts of the diffusion parameters across the skeleton, but rather spatially variable differences across several metrics, including DTI, DKI and WMTI. The observed differences between S4 and S5 pipelines suggest significant effects of bias field corrections across a large part of the skeleton. The comparison between S5 and UKB pipelines revealed similar results as those seen when comparing the S7 and UKB pipelines (Figures 6 and 7b), suggesting that spatial smoothing in the S7 pipeline yields a reasonable improvement which did not remove important information from the data set. The regionally specific increases or decreases in diffusion metrics for different pipelines might partly be explained by an effect of the noise correction step on physiological noise around the large arteries or strong susceptibility-induced artefacts close to air cavities in the brain. Such artefacts might introduce spatially variable distortions, which could lead to spurious findings. This could explain the previously demonstrated higher sensitivity to group differences after noise correction (Kochunov et al., 2018). Moreover, our comparison between pipelines demonstrated that noise and Gibbs ringing corrections (corresponds to S4) influenced tSNR both in the case of conventional (b = 1,000 s/mm^2) and at higher diffusion weightings (b = 2,000 s/mm^2; see Figure 8). Interestingly, the bias field correction step (S5) did not change tSNR compared to S4. Similar, the spatial smoothing step (S7) did not introduce a strong shift in tSNR compared to S4 and S5, however, has been shown to influence further diffusion metric estimations by reducing the number of "bad" voxels (Veraart et al., 2013).

To assess possible practical consequences of the different correction steps, we compared the estimated age slopes in DKI and WMTI metrics between pipelines. Age-related differences are abundant in the relevant age span (Grinberg et al., 2017; Westlye et al., 2010). All included metrics showed an effect of pipeline and, although similar signs, some variability in estimated age-slopes between the S7 and UKB pipelines. The voxel-wise comparisons revealed a higher number of voxels showing significant age associations in the S7 compared to the UKB pipeline for DTI and WMTI metrics. On the contrary, the UKB pipeline demonstrated a higher number of significant voxels for DKI metrics. Although subtle, pipeline related global and spatially varying differences in diffusion metrics will have consequences for subsequent analyses, for example, for machine-learning-based age prediction or diagnostic classification or prediction of clinical traits (Alnaes et al., 2018; Doan et al., 2017; Kuhn et al., 2018; Richard et al., 2018).

In conclusion, our analysis of UKB data demonstrated that temporal SNR and estimated diffusion metrics are sensitive to processing pipeline and might benefit from the proposed sequential advanced post-processing steps. Although applied in UKB, the current pipeline offers an example of a general approach for harmonisation of postprocessing steps across diffusion MRI studies. Whereas artefact correction may be particular important when applying complex diffusion models to multishell diffusion data (Galdi et al., 2019) such as a difference between conventional and kurtosis-derived diffusion metrics or novel experimental setups (Vellmer et al., 2017). The current pipeline can be adapted to other diffusion scheme such as conventional single-shell diffusion acquisitions and isotropic diffusion weighting (Maximov & Vellmer, 2019; Vellmer et al., 2017).

ACKNOWLEDGMENTS

This work was funded by the Research Council of Norway (249795) and the South-Eastern Norway Regional Health Authority (2014097, 2015073, 2016083).

CONFLICT OF INTEREST

Authors declared no conflict of interest.

ORCID

Ivan I. Maximov https://orcid.org/0000-0001-6319-6774
Lars T. Westlye https://orcid.org/0000-0001-8644-956X

REFERENCES

Ades-Aron, B., Veraart, J., Kochunov, P., McGuire, S., Sherman, P., Kellner, E., … Fieremans, E. (2018). Evaluation of the accuracy and precision of the diffusion parameter EStImation with Gibbs and NoisE removal pipeline. *NeuroImage*, *183*, 532–543.

Aja-Fernandez, S., Vegas-Sanchez-Ferrero, G., & Tristan-Vega, A. (2014). Noise estimation in parallel MRI: GRAPPA and SENSE. *Magnetic Resonance Imaging*, *32*, 281–290.

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., … Smith, S. M. (2018). Image processing and quality control for the first 10000 brain imaging datasets from UK Biobank. *NeuroImage*, *166*, 400–424.

Alnaes, D., Kaufmann, T., Doan, N. T., Cordova-Palomera, A., Wang, Y., Bettella, F., … Westlye, L. T. (2018). Association of heritable cognitive ability and psychopathology with white matter properties in children and adolescents. *JAMA Psychiatry*, *75*, 287–295.

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., Filippini, N., & Bastiani, M. (2017). Towards a comprehensive framework for movement and distortion correction of diffusion MR imaging: Within volume movement. *NeuroImage*, *152*, 450–466.

Andersson, J. L. R., Graham, M. S., Zsoldos, E., & Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR imaging. *NeuroImage*, *141*, 556–572.

Andersson, J. L. R., Jenkinson, M., & Smith, S. (2007). *Non-linear optimisation*. In: FMRIB Technical Report TR07JA1. FMRIB Analysis Group, Oxford University. Retrieved from www.fmrib.ox.ac.uk/analysis/techrep

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, *20*, 870–888.

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effect and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078.

Andre, E. D., Grinberg, F., Farrher, E., Maximov, I. I., Shah, N. J., Meyer, C., … Balteau, E. (2014). Influence of noise correction on intra- and inter-subject variability of quantitative metrics in diffusion kurtosis imaging. *PLoS One*, *9*, e94531.

Banerjee, A., & Maji, P. (2015). Rough sets and stomped normal distribution for simultaneous segmentation and bias field correction in brain MR images. *IEEE Image Processing*, *24*, 5764–5776.

Basser, P. J., Mattiello, J., & Le Bihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, *66*, 259–267.

Bastiani, M., Andersson, J. L. R., Cordero-Grande, L., Murgasova, M., Hutter, J., Price, A. N., … Sotiropoulos, S. N. (2019). Automated processing pipeline for neonatal diffusion MRI in the developing Human Connectome Project. *NeuroImage*, *185*, 750–763.

Cui, Z., Zhong, S., Xu, P., He, Y., & Gong, G. (2013). PANDA: A pipeline toolbox for analyzing brain diffusion images. *Frontiers in Human Neuroscience*, *7*, 42.

Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS One*, *10*, e0121945.

Doan, N. T., Persson, A. E. K., Alnaes, D., Kaufmann, T., Rokicki, J., Cordova-Palomera, A., … Westlye, L. T. (2017). Dissociable diffusion MRI patterns of white matter microstructure and connectivity in Alzheimer's disease spectrum. *Scientific Reports*, *7*, 45131.

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLoS One*, *12*, e0184661.

Farzinfar, M., Oguz, I., Smith, R. G., Verde, A. R., Dietrich, C., Gupta, A., … Styner, M. A. (2013). Diffusion imaging quality control via entropy of principal direction distribution. *NeuroImage*, *82*, 1–12.

Fieremans, E., Jensen, J. H., & Helpern, J. A. (2011). White matter characterization with diffusion kurtosis imaging. *NeuroImage*, *58*, 177–188.

Galdi, P., Blesa, M., Stoye, D. Q., Sullivan, G., Lamb, G. J., Quigley, A. J., … Boardman, J. P. (2019). Neonatal morphometric similarity mapping for predicting brain age and characterizing neuroanatomic variation associated with preterm birth. *BioRxiv*. https://doi.org/10.1101/569319

Gallichan, D., Scholz, J., Bartsch, A., Behrens, T. E., Robson, M. W., & Miller, K. L. (2010). Addressing a systematic artifact in diffusion-weighted MRI. *Human Brain Mapping*, *31*, 193–202.

Ganzetti, M., Wenderoth, N., & Mantini, D. (2016). Quantitative evaluation of intensity inhomogeneity correction methods for structural MR brain images. *Neuroinformatics*, *14*, 5–21.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., … for the WU-Minn HCP Consortium. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, *80*, 105–124.

Grinberg, F., Maximov, I. I., Farrher, E., Neuner, I., Amort, L., Thoennessen, H., … Shah, N. J. (2017). Diffusion kurtosis metrics as biomarkers of microstructural development: A comparative study of a group of children and a group of adults. *NeuroImage*, *144*(Pt. A), 12–22.

Hasan, K. M. (2007). A framework for quality control and parameter optimization in diffusion tensor imaging: Theoretical analysis and validation. *Magnetic Resonance Imaging*, *25*, 1196–1202.

Holland, P. W., & Welsch, R. E. (1997). Robust regression using iteratively reweighted least-squares. *Communications in Statistics: Theory and Methods*, *A6*, 813–827.

Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., & Kaczynski, K. (2005). Diffusion kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, *53*, 1432–1440.

Johansen-Berg, H., & Behrens, T. E. J. (2014). *Diffusion MRI: From quantitative measurement to in-vivo neuroanatomy*. Amsterdam, The Netherlands: Elsevier Academic Press.

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, *76*, 1574–1581.

Koay, C. G., & Basser, P. J. (2006). Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. *Journal of Magnetic Resonance*, *179*, 317–322.

Kochunov, P., Dickie, E. W., Viviano, J. D., Turner, J., Kingley, P. B., Jahanshad, N., … Voineskas, V. N. (2018). Integration of routine QA data into meta-analysis may improve quality and sensitivity of multisite diffusion tensor imaging studies. *Human Brain Mapping*, *39*, 1015–1023.

Kuhn, T., Kaufmann, T., Doan, N. T., Westlye, L. T., Jones, J., Nunez, R. A., … Thames, A. D. (2018). An augmented aging process in brain white matter in HIV. *Human Brain Mapping*, *39*, 2532–2540.

Manjón, J. V., Coupé, P., & Buades, A. (2015). MRI noise estimation and denoising using non-local PCA. *Medical Image Analysis*, *22*(1), 35–47. https://doi.org/10.1016/j.media.2015.01.004

Maximov, I. I., Farrher, E., Grinberg, F., & Shah, N. J. (2012). Spatially variable Rician noise in magnetic resonance imaging. *Medical Image Analysis*, *16*, 536–548.

Maximov, I. I., Grinberg, F., & Shah, N. J. (2011). Robust tensor estimation in diffusion tensor imaging. *Journal of Magnetic Resonance*, *213*, 136–144.

Maximov, I. I., Thoennessen, H., Konrad, K., Amort, L., Neuner, I., & Shah, N. J. (2015). Statistical instability of TBSS analysis based on DTI fitting algorithm. *Journal of Neuroimaging*, *25*, 883–891.

Maximov, I. I., Tonoyan, A. S., & Pronin, I. N. (2017). Differentiation of glioma malignancy grade using diffusion MRI. *Physica Medica*, *40*, 24–32.

Maximov, I. I., & Vellmer, S. (2019). Isotropically weighted intravoxel incoherent motion brain imaging at 7T. *Magnetic Resonance Imaging*, *57*, 124–132.

McRobbie, D. W., Moore, E., Graves, M. J., & Prince, M. R. (2006). *MRI from picture to proton*. Cambridge, UK: Cambridge University Press.

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., … Smith, S. M. (2016). Multimodal population brain imaging in the UKbiobank prospective epidemiological study. *Nature Neuroscience*, *19*, 1523–1536.

Novikov, D. S., Kiselev, V. G., & Jespersen, S. N. (2018). On modeling. *Magnetic Resonance in Medicine*, *79*, 3172–3193.

Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., … Styner, M. (2014). DTIPrep: Quality control of diffusion-weighted images. *Frontiers in Neuroinformatics*, *8*, 4.

Perona, P., & Malik, J. (1990). *Scale-space and edge detection using anisotropic diffusion*. (Vol. *12*, Issue 7). IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/34.56205

Perrone, D., Aelterman, J., Puzurica, A., Jeurissen, B., Philips, W., & Leemans, A. (2015). The effect of Gibbs ringing artifacts on measures derived from diffusion MRI. *NeuroImage*, *120*, 441–455.

Richard, G., Kolskaar, K., Sanders, A. M., Kaufmann, T., Petersen, A., Doan, N. T., … Westlye, L. T. (2018). Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. *PeerJ*, *6*, 5908.

Roalf, D. R., Quarmley, M., Elliott, M. A., Satterthwaite, T. D., Vandekar, S. N., Ruparel, K., … Gur, R. E. (2016). The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. *NeuroImage*, *125*, 319–903.

Sairanen, V., Leemans, A., & Tax, C. M. W. (2018). Fast and accurate slisewise outlier detection (SOLID) with informed model estimation for diffusion MRI data. *NeuroImage*, *181*, 331–346.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., … Behrens, T. E. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*, 1487–1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, 208–219.

Smith, S. M., Johansen-Berg, H., Jenkinson, M., Rueckert, D., Nichols, T. E., Miller, K. L., … Behrens, T. E. (2007). Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. *Nature Protocols*, *2*, 499–503.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*, 83–98.

Sotiropoulos, S. N., Jbabdi, S., Xu, J., Andersson, J. L., Moeller, S., Auerbach, E. J., … for the WU-Minn HCP Consortium. (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage*, *80*, 125–143.

Stejskal, E. O., & Tanner, J. E. (1965). Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, *42*, 288.

Tamnes, C. K., Roalf, D. R., Goddings, A. L., & Lebel, C. (2017). Diffusion MRI of white matter microstructure development in childhood and adolescence: Methods, challenges and progress. *Developmental Cognitive Neuroscience*, *33*, 161–175. https://doi.org/10.1016/j.dcn.2017.12.002

Taylor, P. A., Alhamud, A., van der Kouwe, A., Saleh, M. G., Laughton, B., & Meintjes, E. (2016). Assessing the performance of different DTI motion correction strategies in the presence of EPI distortion correction. *Human Brain Mapping*, *37*, 4405–4424.

Tønnesen, S., Kaufmann, T., Doan, N. T., Alnæs, D., Córdova-Palomera, A., Meer, D. V., … Westlye, L. T. (2018). White matter aberrations and age-related trajectories in patients with schizophrenia and bipolar disorder revealed by diffusion tensor imaging. *Scientific Reports*, *8*, 14129.

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*, 1310–1320.

Van Hecke, W., Leemans, A., De Backer, S., Jeurissen, B., Parizel, P. M., & Sijbers, J. (2010). Comparing isotropic and anisotropic smoothing for voxel-based DTI analysis: A simulation study. *Human Brain Mapping*, *31*, 98–114.

Vellmer, S., Stirnberg, R., Edelhoff, D., Suter, D., Stoecker, T., & Maximov, I. I. (2017). Comparative analysis of isotropic diffusion weighted imaging sequences. *Journal of Magnetic Resonance*, *275*, 137–147.

Vellmer, S., Tonoyan, A. S., Suter, D., Pronin, I. N., & Maximov, I. I. (2018). Validation of DWI pre-processing procedures for reliable differentiation between human brain gliomas. *Zeitschrift für Medizinische Physik*, *28*, 14–24.

Veraart, J., Fieremans, E., Jelescu, I. O., Knoll, F., & Novikov, D. S. (2016). Gibbs ringing in diffusion MRI. *Magnetic Resonance in Medicine*, *76*, 301–314.

Veraart, J., Fieremans, E., & Novikov, D. S. (2016). Diffusion MRI noise mapping using random matrix theory. *Magnetic Resonance in Medicine*, *76*, 1582–1593.

Veraart, J., Novikov, D. S., Christiaens, D., Ades-Aron, B., Sijbers, J., & Fieremans, E. (2016). Denoising of diffusion MRI using random matrix theory. *NeuroImage*, *142*, 394–406.

Veraart, J., Sijbers, J., Sunaert, S., Leemans, A., & Jeurissen, B. (2013). Weighted linear least squares estimation of diffusion MRI parameters: Strengths, limitations, and pitfalls. *NeuroImage*, *81*, 335–346.

Walker, L., Chang, L. C., Koay, C. G., Sharma, N., Cohen, L., Verma, R., & Pierpaoli, C. (2011). Effect of physiological noise in population analysis of diffusion tensor MRI data. *NeuroImage*, *54*, 1168–1177.

Westlye, L. T., Reinvang, I., Rootwelt, H., & Espeseth, T. (2012). Effects of APOE on brain white matter microstructure in healthy adults. *Neurology*, *79*, 1961–1969.

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., … Fjell, A. M. (2010). Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. *Cerebral Cortex*, *20*, 2055–2068.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*, 297–381.

Wu, M., Chang, L. C., Walker, L., Lemaitre, H., Barnett, A. S., Marenco, S., & Pierpaoli, C. (2008). Comparison of EPI distortion correction methods in diffusion tensor MRI using a novel framework. *Medical Image Computing and Computer Assisted Intervention*, *11*, 321–329.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Maximov II, Alnæs D, Westlye LT. Towards an optimised processing pipeline for diffusion magnetic resonance imaging data: Effects of artefact corrections on diffusion metrics and their age associations in UK Biobank. *Hum Brain Mapp*. 2019;40:4146–4162. https://doi.org/10.1002/hbm.24691