Sensitivity of Diffusion MRI to White Matter Pathology: Influence of Diffusion Protocol, Magnetic Field Strength, and Processing Pipeline in Systemic Lupus Erythematosus

Evgenios N. Kornaropoulos 1,2*, Stefan Winzeck 2,3, Theodor Rumetshofer 1, Anna Wikstrom 1, Linda Knutsson 4,5,6, Marta M. Correia 7, Pia C. Sundgren 1,8,9† and Markus Nilsson 1†

1 Clinical Sciences, Diagnostic Radiology, Lund University, Lund, Sweden, 2 Division of Anaesthesia, University of Cambridge, Cambridge, United Kingdom, 3 BioMedIA Group, Department of Computing, Imperial College London, London, United Kingdom, 4 Department of Medical Radiation Physics, Lund University, Lund, Sweden, 5 Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, School of Medicine, Baltimore, MD, United States, 6 F.M. Kirby Research Center, Kennedy Krieger Institute, Baltimore, MD, United States, 7 MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom, 8 Lund University BioImaging Center, Lund University, Lund, Sweden, 9 Department of Medical Imaging and Physiology, Skåne University Hospital, Lund, Sweden

There are many ways to acquire and process diffusion MRI (dMRI) data for group studies, but it is unknown which maximizes the sensitivity to white matter (WM) pathology. Inspired by this question, we analyzed data acquired for diffusion tensor imaging (DTI) and diffusion kurtosis imaging (DKI) at 3T (3T-DTI and 3T-DKI) and DTI at 7T in patients with systemic lupus erythematosus (SLE) and healthy controls (HC). Parameter estimates in 72 WM tracts were obtained using TractSeg. The impact on the sensitivity to WM pathology was evaluated for the diffusion protocol, the magnetic field strength, and the processing pipeline. Sensitivity was quantified in terms of Cohen’s $d$ for group comparison. Results showed that the choice of diffusion protocol had the largest impact on the effect size. The effect size in fractional anisotropy (FA) across all WM tracts was 0.26 higher when derived by DTI than by DKI and 0.20 higher in 3T compared with 7T. The difference due to the diffusion protocol was larger than the difference due to magnetic field strength for the majority of diffusion parameters. In contrast, the difference between including or excluding different processing steps was near negligible, except for the correction of distortions from eddy currents and motion which had a clearly positive impact. For example, effect sizes increased on average by 0.07 by including motion and eddy correction for FA derived from 3T-DTI. Effect sizes were slightly reduced by the incorporation of denoising and Gibbs-ringing removal (on average by 0.011 and 0.005, respectively). Smoothing prior to diffusion model fitting generally reduced effect sizes.
In summary, 3T-DTI in combination with eddy current and motion correction yielded the highest sensitivity to WM pathology in patients with SLE. However, our results also indicated that the 3T-DKI and 7T-DTI protocols used here may be adjusted to increase effect sizes.

Keywords: diffusion MRI, DTI, DKI, ROI-based analysis, ultra-high magnetic field strength (7T), diffusion processing, white matter fiber-tracts, effect sizes

1. INTRODUCTION

Diffusion MRI (dMRI) can be used to characterize the microstructure of white matter (WM) fiber tracts by parameters obtained with for example diffusion tensor imaging (DTI). Examples of such parameters include the mean, axial, and radial diffusivity (MD, AD, and RD, respectively) and the fractional anisotropy (FA). Changes in these parameters have been detected in numerous conditions, including aging (1), traumatic brain injury (TBI) (2, 3), schizophrenia (4, 5), Parkinson’s disease (6, 7), multiple sclerosis (MS) (8), and systemic lupus erythematosus (SLE) (9–12), (13–15). Diffusion kurtosis imaging (DKI) is an extension to DTI that provides information complementary to DTI (16–18), but requires a more comprehensive acquisition protocol and, thus, longer scan times. Whether to accept the longer scan times of a DKI protocol or to opt for a shorter DTI protocol is just one of the many questions scientists face when designing a dMRI protocol. Other questions may be what magnetic field strength to use, as it can also influence the outcome of a study (19). In addition, there are many questions concerning the choice of the processing pipeline, which can also impact the sensitivity of dMRI to pathology (20). An evaluation of all these aspects would enable a more informed choice of methods. Here, we evaluated three aspects: the diffusion protocol (DTI vs. DKI), the magnetic field strength (3T vs. 7T), and the processing pipeline (seven different options). The evaluation was based on a groupwise comparison of dMRI data from patients with SLE. This is a disease with a broad variety of symptoms of both a neurologic and psychiatric nature (21). Previous studies have reported reduced FA in the corpus callosum and a wide range of association fibers (14, 22, 23). In this study, our goal was to analyze the degree to which the more resource-intensive approaches such as DKI, 7T, or computationally expensive processing bring benefits in terms of increased sensitivity to WM pathology in patients with SLE, and analyze the degree to which the results align with other studies on different dMRI protocols and processing pipelines.

Concerning the diffusion protocol, the main difference between DTI and DKI is that the latter allows for the estimation of the mean, axial, and radial kurtosis (MK, AK, and RK, respectively) in addition to the parameters obtained with DTI (MD, AD, RD, and FA) (24). The kurtosis parameters characterize the diffusional heterogeneity that might be present in tissues consisting of compartments with different diffusivities (18, 25, 26). This benefit comes at a cost: DKI needs a multi-shell acquisition protocol with at least two non-zero and different $b$-values, in contrast to DTI where a single-shell acquisition is sufficient (27). Moreover, DKI requires the acquisition of images with higher $b$-values (in the range of 2,000–2,500 $s/mm^2$). This, in turn, necessitates diffusion encoding with longer gradient pulses, and therefore, DKI is performed with longer echo times than DTI, which reduces the baseline signal-to-noise-ratio (SNR). This is often compensated for by reducing the spatial resolution in DKI compared with DTI. DKI and DTI have been applied together before (28–38). Generally, these studies point to a rise in FA, MK, AK, and RK and a decrease in MD, AD, and RD in the early development of cerebral WM (29, 30). Then a reverse process takes place later either due to aging (28) or due to neurodegenerative disease in conditions such as schizophrenia (32), MS (31), Alzheimer’s disease (33, 34, 36, 37), and Parkinson’s disease (38, 39). However, it is not clear whether DTI or DKI is most sensitive to WM pathology (40). For example, in Alzheimer’s disease, MD and MK seem to be most sensitive, but some studies highlight the former (e.g., (41)) and others the latter (e.g., (42)). Overall, MD is often reported to have high sensitivity to neurodegeneration, followed by MK and to a lesser degree FA, RD, and RK (40).

How the magnetic field strength influences the sensitivity of dMRI to pathology has been less extensively investigated than the effect of the diffusion protocol, possibly because ultra-high-field (UHF) imaging (e.g., 7T) has only recently become relatively widely available for clinical research (43–45). For similar image resolution, 7T-DTI offers increased contrast-to-noise-ratio and SNR compared to 3T-DTI (46–48). However, 7T-DTI suffers from increased spatial heterogeneity in brain regions such as the temporal lobes (49, 50). A recent investigation on the impact of the magnetic field strength in a small population of seven MS patients and six healthy controls (HC) showed that both 3T and 7T are viable options for imaging WM tissue change in MS (31).

Apart from the diffusion protocol and the magnetic field strength, image processing can also affect sensitivity to pathology (20, 51–53). Optimizing the processing pipeline has the potential to increase the sensitivity to pathology (54–57). For example, age-related WM changes seem to be best revealed when a combination of all the state-of-the-art processing steps are applied (20).

In this study, we investigated the dependence of the sensitivity to WM pathology in patients with SLE on the diffusion protocol (DTI or DKI), the magnetic field strength (3T or 7T), and the inclusion of various processing steps (denoising, Gibbs-ringing removal, eddy-current and motion correction, and smoothing, in different combinations). Data was acquired with three protocols: 3T-DTI, 3T-DKI, and 7T-DTI. The hypothesis was that 3T-DKI and 7T-DTI would show benefits compared with 3T-DTI, as these protocols are more resource-intensive in terms of either time (3T-DKI) or the use of a scarce but SNR-boosting resource.
(7T-DTI). We also hypothesized that a more extensive and, thus, computationally intensive processing pipeline that incorporates several state-of-the-art processing steps would be beneficial. To test these hypotheses, we performed a region-based effect-size analysis. Cohen’s $d$ was used as a measure of effect size, as it evaluates the difference in means between two populations normalized by their joint SD (58). The effect size analysis was applied to analyze the difference between patients with SLE and HC in 72 major WM fiber tracts obtained from TractSeg (59, 60).

2. MATERIALS AND METHODS

Figure 1 shows the workflow of this study. In this section, we describe each of those steps from data acquisition to effect size estimation.

2.1. Data Acquisition and Participants

Imaging was performed on two different systems (3T Siemens Skyra and 7T Philips Achieva) with three different protocols, referred to as 3T-DTI, 3T-DKI, and 7T-DTI. The image resolutions, $b$-value scheme, and repetition and echo times (TR and TE) were adjusted for each protocol and system independently, and are reported in Table 1. Generally, TE and TR were minimized. The resolution and other imaging parameters were adjusted to minimize artifacts such as signal bias due to the rectified noise floor (61). Specifically, as 3T-DKI includes the acquisition of high $b$-value data with lower SNR, the 3T-DKI was performed with a lower resolution than the 3T-DTI protocol in order to avoid noise-floor effects. A 7T-DKI protocol was not included due to limitations in total scan time per patient. The 7T-DTI protocol featured fewer encoding directions than the 3T-DTI protocol, also due to scan time prioritizations. The extent to which these protocol differences affected the results will be considered in the discussion.

In total, 106 female subjects were scanned. Out of these, 76 were SLE patients and 30 HC. The Regional Ethical Review Board in Lund, Sweden approved the studies on 3T and 7T (#2012/254, #2014/778, #2016/30, #2019/01953) and all participants gave written informed consent prior to the examinations. None of the controls had a history of neurologic, neurodegenerative, or psychiatric disorders. The 106 subjects were investigated at least one of the acquisition protocols (refer to Table 3). Out of the 30 HC, 13 were scanned with all the three different protocols, 20 with at least 3T-DTI, 20 with at least 3T-DKI, and 21 with at least 7T-DTI. Out of the 76 patients with SLE, 59 were scanned with all the three acquisition protocols, 63 with at least 3T-DTI, 56 with at least 3T-DKI, and 54 patients with at least 7T-DTI.

2.2. Processing Pipeline

To analyze the effect of the processing pipeline on the effect size in a group comparison, we built seven processing pipelines. These comprise some or all of three processing steps: denoising, correction for Gibbs-ringing artifacts, and correction of distortions due to head motion and eddy currents. For denoising, we used the method proposed by Veraart et al. (73), termed Marchenko-Pastur principal component analysis (MPPCA). This method is based on the idea of applying principal component analysis (PCA) within a local neighborhood of the voxel, in order to shrink the redundant components over which thermal noise is spread and instead reveal the signal-carrying principal components (74). In contrast to previous local PCA denoising approaches by Manjón et al. (75), MPPCA automatically estimates the number of eigenvalues associated with noise by using random matrix theory for noisy covariance matrices (76). For removal of Gibbs-ringing artifacts, which appear due to a k-space truncation (77), we used the method presented in Kellner et al. (78). In that study, Gibbs-ringing artifacts, most often appearing on sharp edges, are minimized by finding the optimal subvoxel-shift for pixels in the neighborhood of such sharp edges. Finally, dMRI data also suffer from subject motion as well as eddy current-induced artifacts due to the strong and rapidly switching diffusion encoding gradients. To correct for such distortions, we used eddy from FSL (79).

The seven pipelines that were examined in this study consisted of MPPCA for denoising, Gibbs for removal of Gibbs-ringing artifacts, Eddy for correction of motion and eddy current-induced distortions, MPPCA and Gibbs in combination, MPPCA and Eddy in combination, Gibbs and Eddy in combination and MPPCA and Gibbs and Eddy in combination (refer to phase 2 in Figure 1). For reference, we also investigated the effect of applying no processing at all (i.e., none). We also evaluated the impact of Gaussian smoothing, by smoothing the diffusion-weighted imaging data using kernels with SDs ranging from 0 to 1 in units of 0.1 (refer to phase 4 in Figure 1). The purpose of smoothing is to increase SNR prior to the estimation of diffusion scalar metrics (20). All three processing methods (MPPCA, Gibbs, and Eddy) were incorporated in recently published articles (20, 52) and smoothing has often been applied in studies involving DTI or DKI (20, 39, 68, 80–84).

2.3. Segmentation of WM Tracts

To obtain tract-specific parameter values, WM tract segmentation was performed using TractSeg (59, 60). This is a convolutional neural network-based segmentation approach that automatically segments 72 major WM tracts in the native space of the diffusion-weighted images. The algorithm was pretrained on reference segmentations of tracts for 105 subjects from the Human Connectome Project (85). The main benefit of TractSeg is that it is both fast and accurate (60). It achieves that by directly segmenting the tracts in the field of fiber orientation distribution function peaks without using tractography and image registration. A list of all 72 tracts can be found online (https://github.com/MIC-DKFZ/TractSeg).

2.4. Diffusion Parameter Estimation

We estimated the diffusion parameters using two approaches: DTI (86), which provided FA, MD, AD, and RD, and DKI (18), which provided MK, RK, and AK in addition to the parameters provided by DTI. For DTI, we used DTIFIT in FSL with weighted linear least squares. For DKI, we used the package dipy and its module DiffusionKurtosisModel, again with weighted linear least squares. Examples of parameter maps are shown in Figure 2.
The workflow. The diffusion MRI (dMRI) data were acquired using three different acquisition protocols: 3T-diffusion tensor imaging (DTI), 3T-diffusion kurtosis imaging (DKI), and 7T-DTI (left-most block). The three slices in the left block correspond to the same axial view of the dMRI volume with a $b$-value of 1,000 $s/mm^2$ in each of the three protocols. Subsequently, dMRI data were processed by seven different processing pipelines (a second block from the left). The processed data were then used by TractSeg to segment 72 major white matter (WM) tracts (top right-most block). The processed dMRI data were also fitted using either DTI or DKI (depending on the initial acquisition) to extract dMRI parameters. To assess the impact of smoothing prior to model fitting, this fitting step was performed with different degrees of smoothing, ranging from no to substantial smoothing (right-most block second from top). Finally, Cohen’s $d$ was computed from the average and SD of the parameters within the 72 WM tracts segmented earlier (bottom block).

| Image acquisition | #SLE, #HC | Mean age (std) of HC, SLE | Image resolution (isotropic, in $mm^3$) | $b$-values in $s/mm^2$ (# of directions) | TR/TE in ms/ms |
|--------------------|-----------|--------------------------|----------------------------------------|--------------------------------------|----------------|
| 3T-DTI             | 63, 20    | 37 (9), 36 (9)           | 2.0                                    | 0 (8), 1,000 (64)                     | 7,300/73       |
| 3T-DKI             | 56, 20    | 37 (9), 35 (9)           | 2.3                                    | 0 (3), 250 (6), 500 (6), 1,000 (20), 2750 (30) | 7,500/103      |
| 7T-DTI             | 54, 21    | 40 (10), 40 (9)          | 2.0                                    | 0 (3), 1,000 (30)                     | 8,816/62       |

Rows show the three protocols: # corresponds to the number of, SLE, patients with systemic lupus erythematosus; HC, human controls; TR, repetition time; TE, echo time. In total 106 subjects participated in this study: 76 patients with SLE and 30 HC. Out of the 106 subjects, 47 were scanned with all three image acquisition protocols (3T-DTI and 3T-DKI and 7T-DTI); 37 patients with SLE and 10 HC.
2.5. Effect Size Estimates

We evaluated for each tract and diffusion parameter the groupwise difference between SLE patients and HC by Cohen’s $d$ (58), which provides an effect size measure (87) defined as:

$$d = \frac{u_{HC} - u_{SLE}}{s}$$ (1)

with

$$s = \sqrt{\frac{(n_{HC} - 1)s_{HC}^2 + (n_{SLE} - 1)s_{SLE}^2}{n_{HC} + n_{SLE} - 2}}$$ (2)

where $u_{HC}$ and $u_{SLE}$ are the mean values of the parameters within the tract for the HC and SLE groups, respectively, and $s$ is the pooled SD of the two groups. Moreover, $n_{HC}$ and $n_{SLE}$ are the sizes and $s_{HC}^2$ and $s_{SLE}^2$ the variances of the HC and SLE groups, respectively. To make it easier to compare effect sizes across parameters, we changed the sign of the effect size estimate in MD, RD, and AD so that all parameters had positive effect size estimates. As Cohen’s $d$ is a standardized effect size measure, it can be more easily compared across studies and populations (88). Of note, for an unbalanced dataset like ours (more patients than controls), the pooled SD in the above Cohen’s $d$ formula mostly reflects the SD of the larger group (i.e., the patients), which tends to be more heterogeneous and, thus, have higher SD than the controls group. This will, in turn, provide lower effect sizes than was would have been observed in a balanced setting. However, this does not affect the comparison of effect sizes, which is the main topic of the study.

To assess statistical significance, we note that the 95% CI for Cohen’s $d$ in the absence of a true effect (a true $d$ of zero) for group sizes of 56 and 20, as in 3T-DKI spans the range where the magnitude of $d$ is below 0.54. Effect sizes larger than this can, thus, be considered significant on a 5% significance level (89). In the case of both 3T-DTI and 7T-DTI, the level of significance was set to 0.53, based on the corresponding sizes of the cohorts. Effect sizes were estimated in 72 tracts and, thus, multiple
comparison problem arises. We did not correct this, but note that a Bonferroni-like correction can be applied (90–92). Correcting for 72 independent tests corresponds to using the 99.9% CI (1 - 0.05/72), which in turn corresponds to a minimum threshold for a significant effect at \( d = 0.84 \). Another way to approach the multiple comparisons problem is to note that in the absence of a true effect in all 72 tracts, the probability to still identify more than seven tracts as significant on the uncorrected significance level (5%) is less than 5% (93). This means that observing more than seven tracts as significant on the uncorrected level indicates a true effect in at least some of those tracts.

3. RESULTS

The results that are primarily reported have been derived from comparisons using all subjects, but complementary analyses were also performed using only subjects with data from all three protocols. These latter results are reported in the Supplementary Material. For the effect size analysis using Cohen’s \( d \), we computed mean and SDs of each diffusion parameter (e.g., FA), that were weighted by each tract probability-mask. However, before the above multiplication, we first excluded the voxels that in the probability map had a probability value of less than 0.5.

3.1. WM Fiber-Tracts Segmentation

Figure 3 shows example segmentations of the cingulum and the fornix across different subjects. While the cingulum was segmented consistently, a large variation in segmentation performance is seen for the fornix. Figure 4 shows the segmentation performance for all tracts in terms of the coefficient of variation of the tract volume. The analysis was applied to HC only, as we expect the least variation in that cohort (Figure 4). The values were averaged for tracts found in both the left and right hemispheres, resulting in 41 rows. Three tracts exhibited excessively high variation in tract volume for data acquired with all three protocols (the fornix, the inferior cerebellar peduncle, and the superior cerebellar peduncle). For 7T-DTI, another two tracts showed high volume variation (the anterior commissure and the middle cerebellar peduncle). High variation in volume was defined as a coefficient of variation exceeding 0.25, which is considerably larger than expected from pure variation in anatomy (94). For reference, the coefficient of variation of the total brain volume, the total intracranial volume, the total WM volume, and the total gray matter volume extracted via MRI volumetry is approximately 0.07, 0.12, 0.08, and 0.07, respectively (95). The excessively large volume variation in the aforementioned tracts indicates that TractSeg struggled to reliably segment these across the cohort.

![Figure 3](image-url)  
**Figure 3** | Demonstration of how a tract's volume change across six exemplar human controls for each of the three examined acquisition protocols. We chose to present the variation in the cingulum (top 3 rows) and in the fornix (bottom three rows), as an example of a tract that does not and does, respectively, challenge TractSeg in segmenting it. Note that the fornix is a very small tract in contrast to the cingulum that has a more recognizable shape, facilitating in that way TractSeg on segmenting it and vice-versa for the fornix.
The fornix and the anterior commissures are two of the smallest tracts and are known to be challenging to segment with TractSeg (60), as these are both small structures. The cerebellar peduncles, although larger than the fornix and the anterior commissure, also presented a segmentation challenge possibly because the diffusion tensors change dramatically in the region where the peduncles cross (96). Due to the unreliable segmentation, these five tracts were excluded from further analysis (i.e., the fornix, commissure anterior, superior, inferior, and middle cerebellar peduncles). Note that high volume variation was associated mainly with 7T, as 7T-DTI gave a higher volume variation among the three protocols in most of the tracts regardless of the choice of the pipeline (results shown in Figure 4 were derived by pipeline VII, however, highly similar results were obtained with all seven pipelines). Some possible explanations for this are given later in the discussion.

### 3.2. Effects of Individual Pipelines

Figure 5 shows effect size estimates for different processing pipelines. Rows show results for different parameters (mean FA and mean MD), while columns show results from the three protocols (3T-DTI, 3T-DKI, and 7T-DTI). The figure shows that the choice of processing pipeline had a smaller effect than the choice of acquisition protocol. MPPCA on average (across all tracts and acquisition protocols) reduced the effect sizes (by 0.005, 0.018, 0.014, and 0.018 for FA, MD, AD, and RD, respectively). However, MPPCA had a positive impact on all diffusion kurtosis parameters, with the strongest effect in RK (+0.023). Gibbs-ringing removal had a mixed impact on the effect sizes for DTI parameters (-0.010, +0.003, +0.023, and -0.007 for FA, MD, AD, and RD, respectively), but a positive effect on all DKI parameters related to diffusivity (MD, AD, and RD), with the highest increase seen in AD (+0.022). Finally, eddy current and motion correction on average had a positive effect on all
FIGURE 5 | Impact of the processing pipeline. The impact of the processing pipelines (x-axis) is evaluated through effect size estimates (Cohen’s $d$ scores, y-axis).
Top row: effect size estimates in mean fractional anisotropy (FA) from 3T-DTI (top left), 3T-DKI (top center), and 7T-DTI (top right). Bottom row: effect size estimates with mean diffusivity (MD) from 3T-DTI (top left), 3T-DKI (top center), and 7T-DTI (top right). No smoothing was applied. Each dot represents one tract. An orange dotted line in each plot defines the threshold in effect size above which the result is considered significant.

TABLE 2 | The impact of each processing step (Marchenko-Pastur principal component analysis (MPPCA), Gibbs, and Eddy) on the effect size.

| Parameter | Protocol | MPPCA | Gibbs | Eddy |
|-----------|----------|-------|-------|------|
| FA        | 3T-DTI   | -0.005| -0.028| +0.069|
|           | 3T-DKI   | -0.003| -0.013| +0.059|
|           | 7T-DTI   | -0.009| -0.017| -0.039|
| MD        | 3T-DTI   | -0.026| +0.009| +0.068|
|           | 3T-DKI   | -0.011| +0.006| 0.000|
|           | 7T-DTI   | -0.029| -0.015| -0.013|
| AD        | 3T-DTI   | -0.019| +0.027| +0.029|
|           | 3T-DKI   | -0.011| +0.018| -0.021|
|           | 7T-DTI   | -0.020| +0.004| +0.022|
| RD        | 3T-DTI   | -0.026| -0.010| +0.081|
|           | 3T-DKI   | -0.010| -0.005| +0.021|
|           | 7T-DTI   | -0.024| -0.020| -0.026|
| MK        | 3T-DTK   | +0.020| -0.025| -0.022|
| AK        | 3T-DKI   | +0.005| -0.002| +0.016|
| RK        | 3T-DKI   | +0.010| -0.002| -0.009|
| Mean effect over all parameters | Over all protocols | -0.011 | -0.005 | +0.016 |

A positive value denotes an increase in effect size, whereas a negative value denotes a decrease. The rows correspond to the mean of each diffusion parameter.

DTI parameters and increased their effect sizes by 0.031, 0.0003, 0.006, and 0.010 for FA, MD, AD, and RD, respectively. Table 2 shows another overview of the influence of the processing steps on effect sizes, where the effect of each specific step can be judged by comparing pipelines with and without that step. Overall, the incorporation of eddy current and motion correction increased the effect size by 0.016 on average, while the incorporation of Gibbs-ringing removal and MPPCA reduced effect sizes by 0.005 and 0.011 on average, respectively.

Figure 6 depicts the change in effect sizes over the whole brain due to each method. Among all tracts, the strongest negative impact of MPPCA on the effect size was seen in the isthmus, the middle longitudinal fascicle, and the thalamo-parietal (top panel of brains in Figure 6). In these tracts, effect sizes decreased by more than 0.04 due to MPPCA. Moreover, the inclusion of MPPCA did not lead to a substantial increase in effect sizes in any of the tracts. The inclusion of Gibbs removal caused a slight increase in effect size in the corticospinal tracts (+0.018) and had a less detrimental effect on effect sizes in general compared with MPPCA. The inclusion of Eddy reduced effect sizes in the splenium and the left middle longitudinal fascicle by more than 0.03. However, it also increased effect sizes by more than 0.06 in the left striato-fronto-orbital, the left thalamo-poscentral, and the left striato-premotor (bottom panel of brains in Figure 6).
Overall, Eddy was the most impactful processing step. Therefore, we performed the subsequent analysis with a pipeline that included Eddy. Four pipelines met that criterion (pipelines III, V, VI, and VII). In the end, we chose pipeline VII, which combines all three tested processing steps, since that pipeline was the closest among the four to the most preferable processing scheme in the literature (20, 52).

3.3. Effect of Smoothing

Figure 7 shows the influence of smoothing on the effect size. Four main patterns were identified. The first and dominant pattern comprised a decline in effect size with greater smoothing. This pattern applied to the mean FA (Figure 7A), the SD of FA, the mean MD (Figure 7C), the SD of AD, and the mean MK (refer to Figure 7D) and RK. The second pattern was one where smoothing had little influence on the effect size and applied to the SD of MD, the mean and SD of RD, and the mean AK (refer to Figure 7E). The third pattern was one where the effect size increased slightly with greater smoothing up to a kernel with a SD of 0.4–0.5 followed by a decline for greater smoothing. This applied to the SD of the three DKI parameters (MK, AK, and RK, refer to Figure 7F). The fourth pattern comprised a slight increase with smoothing and was noted only for the case of mean AD (refer to Figure 7B). Note that the smoothing only affected parameter estimation but not the tract segmentation as smoothing was applied after TractSeg and before parameter estimation. This analysis suggests that smoothing was overall detrimental for the effect sizes for all three protocols and, therefore, subsequent analyses were performed without smoothing.

3.4. Effects of Magnetic Field Strength and Acquisition Protocol

Figure 8 shows effect sizes for data acquired with different diffusion protocols (DTI and DKI) and different magnetic field strengths (3T and 7T) using the mean (Figure 8A) and the SD (Figure 8B) of the parameters within the tracts.

Among all parameters from all protocols, the three that had the largest number of tracts with an effect size above 0.54 (significant before correction for multiple comparisons) were the mean and SD of FA and the mean RD, all derived by 3T-DTI. 3T-DTI yielded clearly higher effect sizes than 3T-DKI and 7T-DTI for the mean FA, with an average difference of 0.26 and 0.20, respectively. The FA was the parameter with the highest effect size.
FIGURE 7 | Influence of Gaussian smoothing. Different levels of Gaussian smoothing (sigma, x-axis) were evaluated by the effect size estimates (Cohen’s $d$, y-axis). (A–F) shows the effect on mean FA from 3T-DTI, mean axial diffusivity (AD) from 3T-DTI, mean MD from 7T-DTI, mean MK from 3T-DKI, mean AK from 3T-DKI, and SD of RK from 3T-DKI. The box plots show Cohen’s $d$ scores among tracts. Overall, smoothing had a detrimental influence on effect sizes. An orange dotted line in each plot defines the threshold in effect size above which the result is considered significant.

overall. Among the DKI parameters, RK showed on average the largest effect sizes (refer to far right plots of Figures 8A,B).

Furthermore, the within-tract SD of the FA and RK exhibited similar if not higher effect sizes than the mean (Figure 8B). In some tracts, the SD of FA was the only parameter that exhibited effect sizes above the significance threshold. The SD of FA, when derived by the 3T-DTI protocol, was the parameter for which the most tracts showed significant effects (31 out of 41 or 76% of all tracts), followed by the same parameter when derived by 7T-DTI and 3T-DKI (51 and 32% of all tracts, respectively; see Figure 8B). In 35 out of the 41 tracts, the effect size from 3T-DTI was higher than that derived by 3T-DKI. Effect sizes from 3T-DTI were also higher than those from 7T-DTI in 37 out of the 41 WM tracts. The fact that 3T-DTI yielded the highest number of tracts with a significant effect size can be appreciated also from Figures 9, 10, where brighter colors indicate higher effect sizes.

Finally, we also evaluated all results using only the subjects that were scanned with all three protocols (refer to Figures 11–15 in Appendix). The outcome of that evaluation was similar to the one presented in this section (3T-DTI yielded on average the highest effect sizes), although the margin between 7T-DTI and 3T-DTI was smaller.

4. DISCUSSION

Our analysis of WM pathology in SLE assessed by dMRI showed that the diffusion protocol had the strongest influence on effect sizes among the ones examined: diffusion protocol, magnetic field strength, and processing pipeline. The two DTI protocols (3T-DTI and 7T-DTI) yielded higher effect sizes than the 3T-DKI protocol for most parameters. In only three out of the eight parameter-wise comparisons, did DKI yield higher effect sizes: the mean AD and the SD of MD and RD (refer to Figure 8). Overall, FA was the parameter displaying the highest effect sizes in all the protocols. Using 3T-DTI to compute FA, 76% of all tracts showed significant effects. The corresponding number for 7T-DTI was 51% and 32% for 3T-DKI. The 3T-DKI protocol provided three unique parameters (MK, AK, and RK), but these generally showed low effect sizes. Only RK exhibited effect sizes as high as any of the DTI parameters. These results were consistent across the seven tested processing pipelines. Interestingly and opposite to our initial hypothesis, no substantial increase in sensitivity came with the use of more demanding acquisitions (DKI and 7T). Does this finding generalize across pathologies and across variations in the acquisition protocols? This will be discussed below.

Our observation that DKI is less sensitive than DTI to WM pathology in SLE is not in accordance with findings in other pathologies. For example, DKI parameters revealed differences to a broader extent than DTI parameters across age groups (30), between patients with Parkinson’s disease and HC (68), and between patients with MS and HC (31). DKI also demonstrated a sensitivity superior to that of DTI in Alzheimer’s disease (34) and temporal lobe epilepsy in children (62). On the other hand,
FIGURE 8 | Evaluation of the acquisition protocol on the effect size. Columns show different diffusion parameters obtained with different protocols 3T-DTI, 3T-DKI, or 7T-DTI, shown in cyan, gold, and green, respectively. The mean (A) and SD (B) of each diffusion parameter, derived for each protocol were examined. Pipeline VII was used in all cases. No smoothing was applied. The swarm and box plots show the distribution and quartiles (together with the median) of Cohen’s d values among tracts, respectively. Overall, 3T-DTI was the acquisition protocol that yielded the three most sensitive diffusion parameters [mean FA, SD of FA, and mean radial diffusivity (RD)], and exhibited effect sizes above 0.54 a threshold of significance before correcting for multiple comparisons) in the largest number of tracts. Moreover, the results between 3T-DTI and 7T-DTI were more similar than those between 3T-DTI and 3T-DKI, which implies that the choice of the diffusion protocol impacted the effect size analysis more than the choice of the magnetic field strength. Kurtosis parameters did not yield significant effect sizes, with the exception of the SD of RK, which had significant effects in a small number of tracts. An orange dotted line in each plot defines the threshold in effect size above which the result is considered significant.
FIGURE 9 | Effect size of diffusion tensor parameters (FA, MD, AD, and RD) over the whole brain, when derived by the different acquisition protocols (3T-DTI, 3T-DKI, and 7T-DTI). Effect sizes in the depicted brains range from 0.0 (deep purple areas) to 1.0 (yellow areas). Overall, FA and RD were the parameters with the highest effect sizes across all three acquisition protocols, but the FA derived by 3T-DTI was the diffusion indices with the highest number of areas in the brain exceeding the value 0.54 of effect size. In the case of FA, 3T-DTI (top panel, first row) yielded the highest number of tracts with significantly large effect size than 3T-DKI (top panel, second row) and 7T-DTI (top panel, third row) since the brains in the top panel appear brighter than the brains in the second and third panels.
MD from DTI showed a greater extent of differences across a WM skeleton than any DKI parameter in subjects with sport-related concussion (35). Overall, there seems to be a discrepancy in findings among different diseases (refer to Table 3). This discrepancy could be due to the specific effect of the pathology on the WM microstructure. For example, any tissue alteration that results in constant FA and MD but different MK would be detected by DKI but not DTI. This could for example happen due to axonal degeneration in regions with high orientation dispersion when the intra and extra axonal water have similar isotropic diffusivities. In such situations, FA would be low due to the orientation dispersion and MD would be insensitive to the axonal water fraction. However, MK would be sensitive as the diffusion heterogeneity would change. Although not considering this situation exactly, Szczepankiewicz et al. (98) considers similar scenarios. Finding a scenario in which FA and/or MD are sensitive while MK is not, as was the case in the present study, is more challenging and may point to another reason as an explanation of our findings.

A second reason for a discrepancy in findings with DTI and DKI could be specific differences in the acquisition protocol. In our case, the image resolution was lower in the DKI protocol than in the DTI protocol (2.3 mm × 2.3 mm × 2.3 mm vs. 2 mm × 2 mm × 2 mm). Voxel size is known to affect group level results for DTI parameters (99). An increase in voxel size results in a decrease in FA and an increase in MD, AD, and RD due to both an increase in partial volume effects and an elevated SNR (100). In this study, we found that synthetically reducing the resolution by smoothing generally reduced effect sizes. Whether DTI is more sensitive than DKI when acquired in the same resolution needs to be further investigated, however, this is not trivial as the high $b$-values employed in DKI attenuate the signal considerably. This often necessitates a reduction in the spatial resolution to avoid the noise floor (101). In turn, this may result in reduced effect sizes. However, as DKI demands a higher baseline SNR than DTI, this is an inherent limitation of DKI. High-performance gradient coils can partially alleviate this limitation by enabling shorter echo times and, thus, higher baseline SNR (102).

The field-strength analysis showed that 3T-DTI yielded smaller variation in tract volume and higher effect sizes than 7T-DTI. It could be argued that the latter was due to the shorter scan time of the 7T-DTI protocol, which featured fewer diffusion encoding directions than the 3T-DTI protocol (30 vs. 64, respectively). It, thus, had a slight disadvantage in terms of...
**TABLE 3** Overview of studies comparing 3T-DTI to 3T-DKI.

| References      | Pathology | DKI resolution in mm | Processing pipeline | Analysis method | Sensitivity metric                                                                 | Most sensitive       | Comment                                                                 |
|-----------------|-----------|----------------------|---------------------|-----------------|------------------------------------------------------------------------------------|----------------------|------------------------------------------------------------------------|
| Gao et al. (62) | Epilepsy  | 2.5 x 2.5 x 2.5      | Eddy                | TBSS            | PASW Statistics (patients vs controls)                                               | DKI                  | Small number of subjects                                               |
| Kamagata et al. (38) | PD        | 3.0 x 3.0 x 3.0      | -                   | TBSS            | Student’s t-test (PD vs. HC)                                                        | DKI                  | DKI useful for evaluating crossing fibers                               |
| Zhu et al. (32) | Schizophrenia | 2.0 x 2.0 x 2.0  | Eddy                | TBSS            | % of significantly different skeleton-voxels (schizophrenia vs. HC)                 | none                 | DKI complementary to DTI                                                |
| Coutu et al. (28) | aging    | 2.0 x 2.0 x 2.0      | Eddy                | TBSS            | correlation of diffusion parameters with aging                                     | none                 | DKI complementary to DTI                                                |
| Zhang et al. (63) | PD        | 1.9 x 1.9 x 3.0      | -                   | ROI-based       | student's t-test (PD vs. HC)                                                        | DKI                  | ROI's based on population-based template                                |
| Billiet et al. (64) | aging | 2.5 x 2.5 x 2.5      | Eddy                | ROI-based       | quadratic correlation coefficients of metrics with age                             | DTI                  | dtMRI acquired 6 months after injury                                     |
| Lancaster et al. (35) | mTBI     | 3.0 x 3.0 x 3.0      | Eddy                | TBSS            | significant different skeleton-voxel values (mTBI vs. HC)                          | DTI                  |                                                                           |
| Chen et al. (3-4) | AD        | 1.8 x 1.8 x 1.8      | Eddy                | ROI-based       | classification accuracy (AD, HC)                                                    | DKI                  | ROIs manually drawn                                                      |
| Grinberg et al. (30) | aging  | 1.9 x 1.9 x 1.9      | BckgNoise-Eddy      | TBSS            | Cohen’s d                                                                          | DKI                  | DTI, DKI varied depending on anatomy                                     |
| Chung et al. (65) | IWM       | 2.5 x 2.5 x 2.5      | MPPCA-Gibbs-Eddy-Outliers | TBSS           | voxel-wise correlation with LNS                                                     | DKI                  | MK and AWF the only sensitive parameters                                |
| Karlsen et al. (66) | mTBI     | 2.5 x 2.5 x 2.5      | Eddy                | TBSS            | Welch’s t-test (mTBI vs. HC)                                                        | none                 | Combined utility of DTI and DKI suggested                               |
| Tan et al. (67)  | Astrocytoma | 2.5 x 2.5 x 6.0  | Eddy                | ROI-based       | t-test (patients vs. HC, via SPSS)                                                  | DKI                  | Manual estimation of each parameter’s value                             |
| De Santis et al. (31) | MS       | 1.5 x 1.5 x 1.5      | Eddy                | TBSS            | ANOVA                                                                              | DKI                  | -                                                                      |
| Kamiya et al. (69) | PD        | 3.0 x 3.0 x 3.0      | MPPCA-Gibbs-Eddy-B1 | TBSS and ROI-based | significant skeleton-points, correlation with age                                   | DKI                  | Multidimensional diffusion encoding used                                |
| Yang et al. (69) | BD        | 2.0 x 2.0 x 2.0      | Eddy                | TBSS            | Independent-samples t-test (BD vs. HC)                                              | DKI                  | Higher fidelity in widespread regions in DKI than DTI                   |

TBSS, Tract-based Spatial Statistics; PASW, Predictive Analytics Software; PD, Parkinson’s Disease; HC, human controls; Eddy, eddy-current and motion correction; LNS, Letter Number Sequencing; AD, Alzheimer’s Disease; mTBI, mild Traumatic Brain Injury; BckgNoise, correct for background noise ([70]); IWM, Impaired Working Memory; MPPCA, Marchenko-Pastur Principal Component Analysis; Gibbs, Gibbs’ artefact removal; Outliers, a processing step to remove outliers in DTI or DKI fitting ([71]); AWF, Axonal Water Fraction; MS, Multiple Sclerosis; ANOVA, ANOVA statistical package; BD, Bipolar Disorder; B1, corrected for B1 inhomogeneity ([72]).
of protocol performance. An additional analysis of a version of the 3T-DTI data subsampled to have only 30 directions did not substantially degrade its performance, however (refer to Figure 17 in Appendix). Note that studies have investigated the effect the number of gradient directions has on the accuracy of direction-sensitive diffusion parameters such as FA, AD, and RD (99, 103, 104). These studies show that above a certain number of directions [approximately 25 (103)] the accuracy in those parameters does not seem to improve substantially from an increased number of encoding directions (99). We, thus, expect both protocols to be equally accurate in terms of parameter estimation. Another aspect that could have contributed to lower effect sizes at 7T-DTI compared to 3T-DTI is the shorter T2* relaxation times at 7T. This reduces the intensity of k-space lines far from the center and, thus, leads to some image blurring already in the image acquisition step, which reduces effect sizes.

Regarding the high values of coefficient variation in tract volume in 7T-DTI, one possible reason could be the higher B1 heterogeneity at higher fields (49). This effect causes low signal intensities especially inferior in the brain and the lateral sides of the insula. This might explain why we notice the biggest difference in variation in volume between 3T and 7T acquisitions in the inferior cerebellar peduncle, the inferior longitudinal fascicle, the uncinate fascicle, and the striato-fronto-orbital tracts (Figure 4). Factors other than B1 homogeneity could also be considered, such as the field-of-view or the number of encoding directions. For example, Güllmar et al. (105) reported that the size ratio of the structure to-be-segmented and the size of the input samples (field-of-view) might have an effect on the performance of TractSeg. An additional analysis on the variation in tract volume, comparing the aforementioned three versions of the 3T-DTI protocol to the 7T-DTI one, showed that despite resampling the 3T-DTI to having the equal number of directions and field-of-view, 7T-DTI still shows higher variation in volume than 3T-DTI in most of the tracts (Figure 16 in Appendix). Therefore, B1 inhomogeneity should be the primary cause of the increased volume variation.

A benefit of 7T MRI is that it has a higher baseline SNR than 3T MRI. Here, we might have undermined the sensitivity of our 7T-DTI protocol by acquiring images with the same resolution as in the 3T-DTI protocol (2.0 mm isotropic) rather than utilizing the higher baseline SNR for a higher resolution. Of note, De Santis et al. compared 3T-DTI and 7T-DTI at a higher image resolution than ours (1.5 mm isotropic) and found slightly higher effect sizes at 7T compared to 3T (31). One future direction could be the fusion of 3T and 7T, exploiting the perks of both worlds, with the high angular and spatial resolution, respectively (97). Apart from these image-protocol-related topics (refer to Table 4), there may also be microstructure-related differences between dMRI at 3T and 7T, as relaxation times may change by different amounts with a field strength in different compartments. Interestingly, the highest effect sizes in MD and AD were found with the 7T-DTI protocol. This might indicate that 3T and 7T are sensitive to different aspects of the pathophysiology in SLE. Overall, the lack of a clear advantage with using UHF dMRI in our study agrees with the main message of a recent review in which the author states that diffusion

---

| Table 4 | Overview of studies comparing 3T-DTI to 7T-DTI. |
|------------------|---------------------------------|
| **References**   | **Pathology** | **7T resolution in mm** | **Processing pipeline** | **Analysis method** | **Sensitivity metric** | **Most sensitive** | **Comment**  |
| Sotiropoulos et al. (97) | HCP | 1.05x1.05x1.05 | Eddy and correction for gradient non-linearities | Voxel-wise deconvolution to assess crossing fibers | Percentage of voxels with at least 2 crossing fibers | None | Suggestion for a fusion, exploiting higher angular contrast of 3T-DTI and higher spatial resolution of 7T-DTI. Of note: spatial resolution at 7T was higher than at 3T. 4T (1.05 mm isotropic) compared to 1.5T (1.25 mm isotropic). |
| De Santis et al. (31) | MS | 1.5x1.5x1.5 | Eddy and correction for gradient non-linearities | Tensor-based Spatial Statistics | ANOVA | ANOVA | None | 7T | 7T-DTI and 3T-DTI were acquired at 15mm isotropic resolution. |

HCP, Human Connectome Project; Eddy, eddy-current and motion correction; MS, Multiple Sclerosis; TBSS, Tract-based Spatial Statistics; ANOVA, ANOVA statistical package.
imaging at UHF, though still a worthwhile pursuit, has manifold associated challenges and converting the potential of higher field strengths into “better” diffusion imaging is by no means a straightforward task (106). More study is needed on a 7T-DTI protocol that leverages its benefits (higher SNR) and addresses its weaknesses [e.g., enhancement of B1+ homogeneity using parallel transmit (pTx) RF coils and RF pulse design approaches (107)].

In the analysis of processing pipelines, we noticed three consistent patterns of interest. The first was the minuscule differences across the data processing pipelines. Correcting for distortions from motion and eddy currents (Eddy) was the most beneficial, whereas gross smoothing reduced effect sizes by up to 20%. Interestingly, smoothing has been applied in many studies involving DTI or DKI (20, 39, 68, 80–84). In Maximov et al. (20), it was proposed that the pipeline most sensitive to pathology in terms of aging combines corrections for motion and eddy-current induced distortions, susceptibility deformations, denoising, bias field and Gibbs-ringing removal, together with field mapping and spatial smoothing. However, that study did not examine each step individually, but only cumulatively. In contrast to this, our study suggests that the choice of the processing pipeline does not play a crucial role, although smoothing should be avoided and motion and eddy current correction should be included. The latter is in line with our previous results, which showed that a motion and eddy current correction method capable of dealing with high b-value data reveals significant differences where a simpler and worse one did not (108). Note that this is the first time the study-level impact of individual state-of-the-art processing steps in dMRI has been analyzed.

We identified four primary strengths of the current study. The first strength is the relatively large sample sizes (31 HC and 77 patients with SLE). Second, all processing steps were evaluated independently. Third, the segmentation of the tracts was performed by an automatic pre-trained method, which allowed us to not only investigate tracts over the whole brain but also eliminate bias from subjective tracking. Fourth, the statistical analysis of the tracts took place in the native space of each subject instead of a template space. Previous studies have predominantly used the Tract-Based Spatial Statistics (TBSS) pipeline to perform voxel-wise statistical-between-group comparison of DTI/DKI metrics on MNI152 space (109). However, by deriving WM skeletons from segmentations computed by thresholding FA maps, the TBSS approach lacks the ability to distinguish certain adjacent WM tracts, such as the inferior longitudinal and inferior fronto-occipital fasciculi and, thus, has limited capacity for anatomical specificity (110). Moreover, TBSS requires an accurate non-linear coregistration of the FA maps onto the MNI152 standard space and is, therefore, prone to misregistration errors that bias the final outcome of the study (111).

The study also has few limitations, and here, we consider four of them. First, we only examined one disease paradigm (SLE). Even though patients with SLE manifest a variety of neuropsychiatric symptoms that resemble many other neurodegenerative diseases (112, 113), the question is which of our results generalize to other diseases. Our results comparing DTI and DKI do not seem to generalize across diseases (Table 3), but we do expect that other pathologies will also exhibit large variations in effect sizes due to protocol, and this may be important to consider in meta-analyses. Furthermore, we do expect that our results concerning different pipelines will generalize, as well as our observation of reduced effect sizes from smoothing. Second, the pipelines we considered included neither outlier detection (52) nor harmonization (114). Regarding the former, however, the manual inspection did not reveal any clear outliers in our data, while harmonization was not necessary for our study, since we ran separate analyses on the different protocols and investigated between-group effect sizes per protocol: Cohen’s $d$ is a metric of groupwise differences in the mean normalized by the joint SD. When computed per protocol, inter-protocol biases do not need to be considered. Third, the resolution of DKI was lower than that of DTI (2.3 mm isotropic vs. 2.0 mm isotropic). As mentioned above, the poorer resolution might explain the lower effect sizes with DKI, in particular, given that smoothing led to reduced effect sizes. Fourth, not all subjects were scanned with all three protocols, which might have induced some systematic sample-related differences in the results. However, the majority of subjects overlapped and the overall picture did not change when including only matching subjects (refer to Supplementary Material).

5. CONCLUSION

In conclusion, effect sizes for detecting WM changes in patients with SLE were higher for DTI than DKI and higher for 3T than 7T. However, our results suggest that adjustments could be made to improve the protocols. For example, the sensitivity of 7T-DTI could potentially be enhanced by leveraging the higher baseline SNR of 7T for higher image resolution. Similarly, high-performance gradient coils could be utilized to reduce echo times and, thereby support a higher image resolution in DKI. Among the processing choices, eddy current and motion correction increase effect sizes, while no clear benefits seemed from denoising (MPPCA) and Gibbs-ringing removal. Smoothing was clearly detrimental for the effect sizes. However, the choice of diffusion protocol had a much greater impact than the choice of processing strategy.

DATA AVAILABILITY STATEMENT

Data are available upon request to the authors. A formal data sharing agreement is required prior to sharing data. Raw imaging data cannot be shared due to legal concerns.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Regional Ethical Review Board in Lund, Sweden. The patients/participants provided their written informed consent to participate in this study.
AUTHOR CONTRIBUTIONS

EK contributed to the development of the pipelines, the conception and design of the analysis, the interpretation of the results, analysis of data and drafted the work. SW contributed to the design of the pipelines, the interpretation of data for the work and the revision of the work for important intellectual content. TR contributed to the acquisition of data for the work and the revision of the work for important intellectual content. MC supervised the design of the pipelines, the interpretation of data for the work and the revision of the work for important intellectual content. LK contributed to the revision of the work for important intellectual content. AW contributed to the development of the pipelines. TR contributed to the acquisition of data for the work and the revision of the work for important intellectual content. PS contributed to the acquisition of data for the work, the interpretation of data for the work and the revision of the work for important intellectual content. MN contributed to the conception and design of the analysis, the interpretation of data for the work and the revision of the work for important intellectual content. All authors contributed to the article and approved the submitted version.

FUNDING

The study was supported by Regional Research Funds (RegSkane-824651) (PS), SUS Foundation and Donations funds (PS), Alfred Østerlund Foundation (PS), Swedish Rheumatism Association R-568371 (PS), and King Gustaf V’s 80-year Foundation (FAI-2019-0559) (PS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2022.837385/full#supplementary-material

REFERENCES

1. Lebel C, Gee M, Camicioli R, Wieler M, Martin W, Beaulieu C. Diffusion tensor imaging of white matter tract evolution over the lifespan. Neuroimage. (2012) 60:340–52. doi: 10.1016/j.neuroimage.2011.11.094
2. Schneider DK, Galloway R, Bazarjan JJ, Diekfuus JA, Dudley J, Leach JL, et al. Diffusion tensor imaging in athletes sustaining repetitive head impacts: a systematic review of prospective studies. J Neurotrauma. (2019) 36:2831–49. doi: 10.1089/neu.2019.6398
3. Mahan MY, Rafter DJ, Truwit CL, Oswood M, Samadani U. Evaluation of diffusion measurements reveals radial diffusivity indicative of microstructural damage following acute, mild traumatic brain injury. Magn Reson Imaging. (2020) 77:137–47. doi: 10.1016/j.mri.2020.12.012
4. Kubicki M, McCarley R, Westin CF, Park HJ, Maier S, Kikinis R, et al. A review of diffusion tensor imaging studies in schizophrenia. J Psychiatr Res. (2007) 41:15–30. doi: 10.1016/j.jpsychires.2005.05.005
5. Cetin-Karayumak S, Di Biase MA, Chunga N, Reid B, Somes N, Lyall AE, et al. White matter abnormalities across the lifespan of schizophrenia: a harmonized multi-site diffusion MRI study. Mol Psychiatry. (2020) 25:3208–19. doi: 10.1038/s41388-019-0509-y
6. Schwartz ST, Abaei M, Gouta V, Morgan PS, Bajaj N, Auer DP. Diffusion tensor imaging of nigral degeneration in Parkinson’s disease: a region-of-interest and voxel-based study at 3 T and systematic review with meta-analysis. Neuroimage Clin. (2013) 3:481–8. doi: 10.1016/j.nicl.2013.10.006
7. Surova Y, Nilsson M, Lätt J, Lampinen B, Lindberg Ö, Hall S, et al. Disease-specific structural changes in thalamus and dentatorubrothalamic tract in progressive supranuclear palsy. Neuroradiology. (2015) 57:1079–91. doi: 10.1007/s00234-015-1563-z
8. Inglese M, Bester M. Diffusion imaging in multiple sclerosis research and clinical implications. NMR Biomed. (2010) 23:865–72. doi: 10.1002/nbm.1515
9. Hughes M, Sundgren P, Fan X, Foerster B, Nan B, Welsh R, et al. Diffusion tensor imaging in patients with acute onset of neuropsychiatric systemic lupus erythematosus: a prospective study of apparent diffusion coefficient, fractional anisotropy values, and eigenvalues in different regions of the brain. Acta radiol. (2007) 48:213–22. doi: 10.1080/02841850601105825
10. Zivadinov R, Shucard J, Hussein S, Durfee J, Cox J, Bergland N, et al. Multimodal imaging in systemic lupus erythematosus patients with diffuse neuropsychiatric involvement. Lupus. (2013) 22:675–83. doi: 10.1177/0961233313486193
11. Schmidt-Wilde T, Cagnoli P, Wang P, Schultz T, Lotz A, Mccune WJ, et al. Diminished white matter integrity in patients with systemic lupus erythematosus. Neuroimage Clin. (2014) 5:291–7. doi: 10.1016/j.nicl.2014.07.001
12. Ercan E, Ingo C, Tritanon O, Magro-Checa C, Smith A, Smith S, et al. A multimodal MRI approach to identify and characterize microstructural brain changes in neuropsychiatric systemic lupus erythematosus. Neuroimage Clin. (2015) 8:337–44. doi: 10.1016/j.nicl.2015.05.002
13. Ercan E, Magro-Checa C, Valabregue R, Branzoli F, Wood ET, Steup-Beekman GM, et al. Glial and axonal changes in systemic lupus erythematosus measured with diffusion of intracellular metabolites. Brain. (2016) 139:1447–57. doi: 10.1093/brain/aww631
14. Nystedt J, Nilsson M, Jönsen A, Nilsson P, Bengtsson A, Lilja Å, et al. Altered white matter microstructure in lupus patients: a diffusion tensor imaging study. Arthritis Res Ther. (2018) 20:21. doi: 10.1186/s13075-018-1516-0
15. Kozora E, Filley C, Erkan D, Ulug A, Vo A, Ramon G, et al. Longitudinal evaluation of diffusion tensor imaging and cognition in systemic lupus erythematosus. Lupus. (2018) 27:1810–8. doi: 10.1177/0961231817793215
16. Hui ES, Cheung MM, Qi L, Wu EX. Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis. Neuroimage. (2008) 42:122–34. doi: 10.1016/j.neuroimage.2008.04.237
17. Jensen JH, Helpen J. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. NMR Biomed. (2010) 23:696–710. doi: 10.1002/nbm.1518
18. Jensen JH, Helpen J, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med. (2005) 53:1432–40. doi: 10.1002/mrm.20508
19. Heidemann RM, Porter DA, Anwander A, Feiwetier T, Heberlein K, Knösche TR, et al. Diffusion imaging in humans at 7T using readout-segmented EPI and GRAPPA. Magn Reson Med. (2010) 64:9–14. doi: 10.1002/mrm.22480
20. 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) 40:4146–62. doi: 10.1002/hbm.24691
21. Liang MH, Corzillius M, Bae SC, Lew RA, Fortin PR, Gordon C, et al. The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. Arthritis Rheum. (1999) 42:599–608. doi: 10.1002/1529-0131(199904)42:4<599::AID-ANR2>3.0.CO;2-F
22. Zimny A, Szmyrka-Kaczmarek M, Szewczyk P, Bladowska J, Pokryszko-Dragan A, Grusza E, et al. In vivo evaluation of brain damage in the course of systemic lupus erythematosus using magnetic resonance spectroscopy, perfusion-weighted and diffusion-tensor imaging. Lupus. (2014) 23:10–19. doi: 10.1177/0961233313511556

Komaropoulos et al. Factors Affecting Sensitivity of dMRI April 2022 | Volume 13 | Article 837385
Khatami M, Schmidt-Wilcke T, Sundgren PC, Abbasloo A, Schöllkopf B, Schultz T. BundleMap: anatomically localized features from dMRI for detection of disease. In: *International Workshop on Machine Learning in Medical Imaging*. Cham: Springer (2019). p. 52–60.

Lu H, Jensen JH, Ramani A, Helpert JA. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR Biomed.* (2006) 19:236–47. doi: 10.1002/nbm.1020.

Westin CF, Knutsson H, Pasternak O, Szczepankiewicz F, Ozsaran E, van Westen D, et al. Q-space trajectory imaging for multidimensional diffusion MRI of the human brain. *Neuroimage*. (2016) 135:345–62. doi: 10.1016/j.neuroimage.2016.02.039.

Henriques RN, Jespersen SN, Shemesh N. Correlation tensor magnetic resonance imaging. *Neuroimage*. (2020) 211:116605. doi: 10.1016/j.neuroimage.2020.116605.

Veraart J, Poot DH, Van Hecke W, Blockx I, Van der Linden A, Verbey M, et al. More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging. *Magn Reson Med.* (2011) 65:138–45. doi: 10.1002/mrm.22603.

Coutu JP, Chen JY, Rosas HD, Salat DH. Non-Gaussian water diffusion in aging white matter. *Neurobiol Aging*. (2014) 35:1412–21. doi: 10.1016/j.neurobiolaging.2013.12.001.

Paydar A, Fieremans E, Nwankwo J, Lazar M, Sheth H, Adisetiyo V, et al. Diffusional kurtosis imaging of the developing brain. *Am J Neuroradiol*. (2014) 35:808–14. doi: 10.3174/ajnr.A3764.

Grimberg F, Maximov II, Farber E, Neuner I, Amort I, Thoennell H, et al. Diffusion kurtosis metrics as biomarkers of microstructural development: a comparative study of a group of children and a group of adults. *Neuroimage*. (2017) 144:12–22. doi: 10.1016/j.neuroimage.2016.08.033.

De Santis S, Bastiani M, Drobay A, Kolber P, Zipp F, Pracht E, et al. Characterizing microstructural tissue properties in multiple sclerosis with diffusion MRI at 7 T and 3 T: the impact of the experimental design. *Neuroscience*. (2019) 403:17–26. doi: 10.1016/j.neuroscience.2018.03.048.

Zhu J, Zhao C, Qin W, Wang D, Ma X, Zhou Y, et al. Performances of diffusion kurtosis imaging and diffusion tensor imaging in detecting white matter abnormality in schizophrenia. *Neuroimage Clin*. (2015) 7:170–8. doi: 10.1016/j.neuroimageclinb.2014.12.008.

Praet J, Manyakov NV, Mvecheni L, Mai Z, Terzopoulos V, de Backer S, et al. Diffusion kurtosis imaging allows the early detection and longitudinal follow-up of amyloid-β-induced pathology. *Alzheimers Res Therapy*. (2018) 10:1–16. doi: 10.1186/s13195-017-0329-8.

Choi S, Cunningham DT, Aguila F, Corrigan JD, Bogner J, Mysiw WJ, et al. DTI at 7 and 3 T: systematic comparison of SNR and its impact on quantitative metrics. *Magn Reson Imaging*. (2011) 29:739–751. doi: 10.1016/j.mri.2011.02.009.

Tannous J, Godlewski BR, Tirumalaraju V, Soares JC, Covey PJ, Selvaraj S. Stress, inflammation and hippocampal subfields in depression: A 7 Tesla MRI Study. *Transl Psychiatry*. (2020) 10:1–7. doi: 10.1038/s41398-020-0759-0.

Berron D, Vievge P, Hoekhappeler A, Pluta J, Ding SL, Maass A, et al. A protocol for manual segmentation of mediatal temporal lobe subregions in 7 Tesla MRI. *Neuroimage Clin*. (2017) 15:466–82. doi: 10.1016/j.nicl.2017.05.022.

Choi S, Cunningham DT, Aguila F, Corrigan JD, Bogner J, Mysiw WJ, et al. DTI at 7 and 3 T: systematic comparison of SNR and its impact on quantitative metrics. *Magn Reson Imaging*. (2011) 29:739–751. doi: 10.1016/j.mri.2011.02.009.

Berron D, Vievge P, Hoekhappeler A, Pluta J, Ding SL, Maass A, et al. A protocol for manual segmentation of mediatal temporal lobe subregions in 7 Tesla MRI. *Neuroimage Clin*. (2017) 15:466–82. doi: 10.1016/j.nicl.2017.05.022.

Cunningham D, Choi S, Corrigan J, Bogner J, Mysiw WJ, et al. DTI at 7 and 3 T: systematic comparison of SNR and its impact on quantitative metrics. *Magn Reson Imaging*. (2011) 29:739–751. doi: 10.1016/j.mri.2011.02.009.

Tannous J, Godlewski BR, Tirumalaraju V, Soares JC, Covey PJ, Selvaraj S. Stress, inflammation and hippocampal subfields in depression: A 7 Tesla MRI Study. *Transl Psychiatry*. (2020) 10:1–7. doi: 10.1038/s41398-020-0759-0.

Berron D, Vievge P, Hoekhappeler A, Pluta J, Ding SL, Maass A, et al. A protocol for manual segmentation of mediatal temporal lobe subregions in 7 Tesla MRI. *Neuroimage Clin*. (2017) 15:466–82. doi: 10.1016/j.nicl.2017.05.022.

Cunningham D, Choi S, Corrigan J, Bogner J, Mysiw WJ, et al. DTI at 7 and 3 T: systematic comparison of SNR and its impact on quantitative metrics. *Magn Reson Imaging*. (2011) 29:739–751. doi: 10.1016/j.mri.2011.02.009.

Berron D, Vievge P, Hoekhappeler A, Pluta J, Ding SL, Maass A, et al. A protocol for manual segmentation of mediatal temporal lobe subregions in 7 Tesla MRI. *Neuroimage Clin*. (2017) 15:466–82. doi: 10.1016/j.nicl.2017.05.022.

Cunningham D, Choi S, Corrigan J, Bogner J, Mysiw WJ, et al. DTI at 7 and 3 T: systematic comparison of SNR and its impact on quantitative metrics. *Magn Reson Imaging*. (2011) 29:739–751. doi: 10.1016/j.mri.2011.02.009.

Berron D, Vievge P, Hoekhappeler A, Pluta J, Ding SL, Maass A, et al. A protocol for manual segmentation of mediatal temporal lobe subregions in 7 Tesla MRI. *Neuroimage Clin*. (2017) 15:466–82. doi: 10.1016/j.nicl.2017.05.022.

Cunningham D, Choi S, Corrigan J, Bogner J, Mysiw WJ, et al. DTI at 7 and 3 T: systematic comparison of SNR and its impact on quantitative metrics. *Magn Reson Imaging*. (2011) 29:739–751. doi: 10.1016/j.mri.2011.02.009.
Collier Q, Veraart J, Jeurissen B, den Dekker AJ, Sijbers J. Iterative reweighted.

Kamiya K, Kamagata K, Ogaki K, Hatano T, Ogawa T, Takeshige-Amano H, Chung S, Fieremans E, Kucukboyaci NE, Wang X, Morton CJ, Novikov DS, et al. Working memory and brain tissue microstructure: white matter and postconcussional syndrome. *Neuroimage*. (2019) 97:568–81. doi: 10.1016/j.neuroimage.2019.04.038.

Billiet T, Vandenbulcke M, Mädler B, Peeters R, Dhollander T, Zhang H, Zhang G, Zhang Y, Zhang C, Wang Y, Ma G, Nie K, et al. Diffusion tractography of substantia nigra is a sensitive method for early diagnosis and disease evaluation in Parkinson’s disease. *Parkinsons Dis*. (2015) 2015:207624. doi: 10.1155/2015/207624.

Karlsen RH, Einarsen C, Moe HK, Håberg AK, Vik A, Skandsen T, et al. Diffusion kurtosis imaging of substantia nigra is a sensitive method for early diagnosis and disease evaluation in Parkinson’s disease. *Parkinsons Dis*. (2015) 2015:207624. doi: 10.1155/2015/207624.

Billiet T, Vandenbulcke M, Mädler B, Peeters R, Dhollander T, Zhang H, et al. Age-related microstructural differences quantified using myelin water imaging and advanced diffusion MRI. *Neurobiol Aging*. (2015) 36:2107–21. doi: 10.1016/j.neurobiolaging.2015.02.029.

Chung S, Fieremans E, Kucukboyaci NE, Wang X, Morton CJ, Novikov DS, et al. Working memory and brain tissue microstructure: white matter tract integrity based on multi-shell diffusion MRI. *Sci Rep*. (2018) 8:1–7. doi: 10.1038/s41598-018-21428-4.

Karlsen RH, Einarsen C, Moe HK, Häberg AK, Vik A, Skandsen T, et al. Diffusion kurtosis imaging in mild traumatic brain injury and postconcussional syndrome. *J Neurosci Res*. (2019) 97:568–81. doi: 10.1002/ijr.24383.

Tan Y, Zhang H, Wang X, Qin J, Wang L, Yang G, et al. Comparing the value of DKI and DTI in detecting isocitrate dehydrogenase genotype of astrocytomas. *Clin Radiol*. (2019) 74:314–20. doi: 10.1016/j.crad.2018.12.004.

Kamiya K, Kamagata K, Ogaki K, Hatano T, Ogawa T, Takeshige-Amano H, et al. Brain white-matter degeneration due to aging and Parkinson disease as revealed by double diffusion encoding. *Front Neurosci*. (2020) 14:1091. doi: 10.3389/fnins.2020.00581.

Yang J, Jiang X, Wei S, Deng X, Zhu Y, Chang M, et al. White matter tracts in Bipolar Disorder patients: a comparative study based on diffusion kurtosis and tensor imaging. *J Affect Disord*. (2021) 292:45–55. doi: 10.1016/j.jad.2021.05.030.

André ED, Grinberg F, Farerre E, Maximov II, Shah NJ, Meyer C, et al. Influence of noise correction on intra- and inter-subject variability of quantitative metrics in diffusion kurtosis imaging. *PLOS ONE*. (2014) 9:e94531. doi: 10.1371/journal.pone.0094531.

Collier Q, Veraart J, Jeurissen B, den Dekker AJ, Sijbers J. Iterative reweighted linear least squares for accurate, fast, and robust estimation of diffusion magnetic resonance parameters. *Magn Reson Med*. (2015) 73:2174–84. doi: 10.1002/mrm.25351.

Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: improved N3 bias correction. *Med Image Anal*. (2010) 14:293–307. doi: 10.1016/j.media.2010.02.006.

Veraart J, Novikov DS, Christiaens D, Ades-Aron B, Sijbers J, Fieremans E. Denoising of diffusion MRI using random matrix theory. *Neuroimage*. (2016) 142:394–406. doi: 10.1016/j.neuroimage.2016.08.016.

Hotelling H. Analysis of a complex of statistical variables into principal components. *J Educ Psychol*. (1933) 24:417. doi: 10.1037/h0071325.

Manjón JV, Coupé P, Concha L, Buades A, Collins DL, Robles M. Diffusion weighted image denoising using overcomplete local PCA. *PLOS ONE*. (2013) 8:e73021. doi: 10.1371/journal.pone.0073021.

Marcenko VA, Pastur LA. Distribution of eigenvalues for a random matrix. *Math USSR Sbornik*. (1967) 1:457. doi: 10.1070/SM1967v001n04ABEH001994.

Veraart J, Fieremans E, Jeleczu JO, Knoll F, Novikov DS. Gibbs ringing in diffusion MRI. *Magn Reson Med*. (2018) 79:630–1. doi: 10.1002/mrm.25866.

Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn Reson Med*. (2016) 76:1574–81. doi: 10.1002/mrm.26054.
healthy volunteers and in brain tumors. *Neuroimage.* (2015) 104:241–52. doi: 10.1016/j.neuroimage.2014.09.057

99. Barrio-Arranz G, de Luis-García R, Tristán-Vega A, Martín-Fernández M, Aja-Fernández S. Impact of MR acquisition parameters on DTI scalar indexes: a tractography based approach. *PLoS ONE.* (2015) 10:e0137905. doi: 10.1371/journal.pone.0137905

100. Papinutto ND, Maule E, Jovicich J. Reproducibility and biases in high field brain diffusion MRI: An evaluation of acquisition and analysis variables. *Magn Reson Imaging.* (2013) 31:827–39. doi: 10.1016/j.mri.2013.03.004

101. Vis G, Nilsson M, Westin CF, Szczepankiewicz F. Accuracy and precision in super-resolution MRI: Enabling spherical tensor diffusion encoding at ultra-high b-values and high resolution. *Neuroimage.* 245:118673. doi: 10.1011/2021.03.1743519

102. Setsompop K, Kimmilrong R, Eberlein E, Witzel T, Cohen-Adad J, McNab JA, et al. Pushing the limits of in vivo diffusion MRI for the human connectome project. *Neuroimage.* (2013) 80:220–33. doi: 10.1016/j.neuroimage.2013.05.078

103. Correia MM, Carpenter TA, Williams GB. Looking for the optimal DTI acquisition scheme given a maximum scan time: are more b-values a waste of time? *Magn Reson Imaging.* (2009) 27:163–75. doi: 10.1016/j.mri.2008.06.011

104. Giannelli M, Cosottini M, Michelassi MC, Lazzarotti G, Belmonte G, Bartolozzi C, et al. Dependence of brain DTI maps of fractional anisotropy and mean diffusivity on the number of diffusion weighting directions. *J Appl Clin Med Phys.* (2010) 11:176–90. doi: 10.1120/jacmp.v11i1.2927

105. Güllmar D, Jacobsen N, Deistung A, Timmann D, Ropele S, Reichenbach JR. Investigation of biases in convolutional neural networks for semantic segmentation using performance sensitivity analysis. *Zeitschrift für Medizinische Physik.* (2022) doi: 10.1016/j.zemedi.2021.11.004

106. Gallichan D. Diffusion MRI of the human brain at ultra-high field (UHF): a review. *Neuroimage.* (2018) 168:172–80. doi: 10.1016/j.neuroimage.2017.04.037

107. Waiczies S, Els A, Kuchling J, Bloch KM, Pankowska A, Waiczies H, et al. Magnetic resonance imaging of multiple sclerosis at 7.0 Tesla. *J Vis Exp.* (2021) 168:e62142. doi: 10.3791/62142

108. Nilsson M, Szczepankiewicz F, van Westen D, Hansson O. Extrapolation-based references improve motion and eddy-current correction of high B-value DWI data: application in Parkinson's disease dementia. *PLoS ONE.* (2015) 10:e0141825. doi: 10.1371/journal.pone.0141825

109. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage.* (2006) 31:1487–505. doi: 10.1016/j.neuroimage.2006.02.024

110. Zhang H, Awate SP, Das SR, Woo JH, Melhem ER, Gee JC, et al. A tract-specific framework for white matter morphometry combining macroscopic and microscopic tract features. *Med Image Anal.* (2010) 14:666–73. doi: 10.1016/j.media.2010.05.002

111. Schwarz CG, Reid RI, Gunter JL, Senjem ML, Przybelski SA, Zuck SM, et al. Improved DTI registration allows voxel-based analysis that outperforms tract-based spatial statistics. *Neuroimage.* (2014) 94:65–78. doi: 10.1016/j.neuroimage.2014.03.026

112. Bertsisas GK, Boupas DT. Pathogenesis, diagnosis and management of neuropsychiatric SLE manifestations. *Nat Rev Rheumatol.* (2010) 6:358. doi: 10.1038/nrrheum.2010.62

113. Ainiala H, Loukkola J, Peltola J, Korpela M, Hietaharju A. The prevalence of neuropsychiatric syndromes in systemic lupus erythematosus. *Neurology.* (2001) 57:496–500. doi: 10.1212/WNL.57.3.496

114. Fortin JP, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, et al. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage.* (2017) 161:149–70. doi: 10.1016/j.neuroimage.2017.08.047

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Kornaropoulos, Winzeck, Rumetshofer, Wikstrom, Knutsson, Correia, Sundgren and Nilsson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**