Published in final edited form as:

J Magn Reson Imaging. 2016 December ; 44(6): 1608–1618. doi:10.1002/jmri.25308.

Quantifying the Impact of Underlying Measurement Error on Cervical Spinal Cord Diffusion Tensor Imaging at 3T

Samantha By, M.S.^{1,2}, Alex K. Smith, M.S.^{1,2}, Lindsey M. Dethrage, M.S.^{2,3}, Bailey D. Lyttle, B.S.², Bennett A. Landman, Ph.D.^{1,2,3,4}, Jeff L. Creasy, M.D.³, Siddharama Pawate, M.D.⁵, and Seth A. Smith, Ph.D.^{1,2,3}

¹Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA

²Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, TN, USA

³Department of Radiology and Radiological Sciences, Vanderbilt University, Nashville, TN, USA

⁴Department of Electrical Engineering, Vanderbilt University, Nashville, TN, USA

⁵Department of Neurology, Vanderbilt University, Nashville, TN, USA

Abstract

Purpose—To empirically characterize and quantify the impact of gradient weighting schemes on the appearance and fidelity of diffusion tensor imaging of the human spinal cord in vivo in clinically relevant scan time equivalents (STE).

Materials & Methods—In five healthy controls at 3 tesla, we evaluated test-retest reproducibility and performed voxel-wise analysis of DTI-derived indices (fractional anisotropy (FA), mean (MD), axial (AD) and radial (RD) diffusivity) in the cervical spinal cord to assess spatial dependencies of measurement error and differences across three different sampling schemes (6, 15, and 32 directions) at STE of 4.5, 9 and 18 minutes. A subjective assessment was also performed.

Results—With six directions, column-specific errors are highest (effect size=2.9%, 4.4%, 7.2% for FA in dorsal column, lateral column, and gray matter) and different than the 15-direction scheme ($p<0.05$). STE sequences with 15 and 32 directions exhibited small differences in error ($p>0.05$). For FA and AD, measurement errors are prevalent in gray matter, while partial volume effects with CSF heavily influence RD. Measurement errors decreased with increasing scan time ($p<0.01$), albeit with diminishing returns at scan times longer than 9 minutes ($p<0.05$).

Conclusion—A 15-direction scheme of 9 minutes yields measurements of the cervical spinal cord with low error.

Keywords

spinal cord; magnetic resonance imaging; diffusion tensor imaging

INTRODUCTION

Diffusion tensor imaging (DTI) offers an opportunity to probe tissue microstructure through estimating three eigenvalues from the diffusion tensor, where λ_1 is the axial diffusivity (AD), and the mean of λ_2 and λ_3 is known as the radial diffusivity (RD). Other scalar quantities may be calculated from these eigenvalues such as the fractional anisotropy (FA), a measure of diffusion eccentricity, and the mean diffusivity (MD), the average diffusion in a voxel irrespective of direction (1,2). These quantitative measures have demonstrated promise in characterizing demyelination and axonal damage of white matter in the brain (3,4). There has been increased motivation to expand the application of DTI to study the spinal cord, which is a smaller, and arguably more challenging, tissue, and to evaluate its relationship to the biological and clinical manifestations of diseases. However, clinical implementation of human spinal cord DTI has been hindered by a lack of in vivo characterization of the impact of diffusion weighting choice and scan time.

The spinal cord is a small central nervous system structure that is somatotopically arranged, comprised of segmented tracts that directly communicate with the brain. The two major functions subtended by the spinal cord are motor and sensory function: sensory function is primarily conveyed from the dorsal columns, whereas motor pathways stem from the lateral and ventral columns (5). The integrity of the spinal cord microstructure is vital to neurological function, and damage by neurodegenerative diseases or trauma can have a devastating impact on day-to-day functions such as walking, bowel and bladder function, and sensation. Thus, there is a need for improved quantitative magnetic resonance imaging (MRI) measurements as they may reveal deeper insight on the subtle pathological changes within the human spinal cord that precede neurological dysfunction.

Specifically for DTI in the human spinal cord in vivo, a detailed set of benchmarks for DTI acquisitions could improve our understanding and interpretation of DTI-derived indices through characterizing the impact of (and interaction of anatomy with) noise, DTI-derived index variation (both spatially and temporally) and diffusion weighting scheme. In the human spinal cord, the feasibility of such studies has been hindered by four considerable challenges which must be overcome to generate high quality data: 1) the small size of the spinal cord (~1.5cm) and its substructures (< 5mm), 2) constant physiological motion, 3) field inhomogeneity near bone/tissue interfaces, and 4) cerebrospinal fluid pulsation (6). Simulation studies have identified that estimation of diffusion tensors is heavily influenced by the noise inherent in a DTI acquisition, and consequently, the signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) of the diffusion-weighted image can impact the accuracy and reproducibility of the derived indices (7–9). Similarly, in the brain, in vivo studies corroborate these findings (10–12). However, in the human spinal cord, the sources and distributions of noise vary compared to the brain, and are not well studied in vivo, even though, theoretically, a set of cylindrically oriented tensors may offer some improved insight. As with the brain, one has the choice of increasing the directional resolution of the gradient scheme or increasing the number of averages to increase the SNR for a given scan time (10); however, because of the spinal cord's unique anatomical and physiological differences due to motion, spatial field gradients, largely anisotropic tensors, and cylindrical geometry, selecting more gradients over longer scan time may not provide the expected

increase in fidelity of the derived tensors as can be predicted by studies in the brain, or through simulation.

Lee *et al.* studied DTI acquisition schemes for the cervical cord in the sagittal orientation (13) and focused on FA, however with improved coil and sequence design, higher resolution in the axial orientation can be achieved, which allows a study of both high and low anisotropy tissues (white and gray matter, respectively) (14). A second study suggests no benefit in acquiring more than 15 directions for the estimation of FA (15). With considerable attention being paid to the study of the directional diffusivities for their relationship to axonal and myelin integrity (3,14), a more comprehensive study of all DTI-derived indices, rather than just FA, is warranted. Moreover, for clinical studies when scan time is limited, we seek to quantify how choice of diffusion gradient scheme influence the accuracy and reproducibility of DTI-derived indices for both gray and white matter in the cervical spinal cord at 3T, a step that has not yet been investigated.

Therefore, the objective of this study is to characterize the source of measurement error in clinically feasible, well-balanced diffusion schemes for the cervical spinal cord.

MATERIALS AND METHODS

Data Acquisition

Five healthy volunteers participated in this study (3 male/2 female, ages: 25–36 years). Local institutional review board approval and written informed consent were obtained prior to imaging. Imaging was acquired using a 3.0T whole body MR scanner (Philips Achieva, Best, Netherlands). A quadrature body coil was used for excitation and a 16-channel SENSE neurovascular coil was used for reception. Three DTI datasets were acquired in each scan session, and each volunteer was rescanned within ten days following the initial scan, yielding a total of 30 DTI scans.

Each DTI sequence was acquired with a single-shot echo planar imaging (EPI) using a reduced field-of-view outer volume suppression technique (16) in the axial plane at an effective b -value of 750 s/mm². The protocol included a cardiac-gated spin-echo acquisition with the following relevant parameters: SENSE factor=1.5, flip angle=90°, TR=5 beats (~5000 ms), TE=50 ms, resolution=1×1mm² slice thickness=5mm, number of slices=14, coverage=C2–C5, FOV=64mm × 48mm, and diffusion gradient times of $\delta = 24.4$ ms, $\delta = 12.8$ ms. Three scan time equivalent scans were acquired in each session with a minimally weighted image ($b = 0$) acquired in each dynamic: 6 non-collinear diffusion-weighted directions acquired 9 times (18:15 min), 15 directions with 4 acquisitions (17:15 min), and 32 directions with 2 acquisitions (17:15 min). The direction schemes were chosen from the default orientation scheme from the scanner, such that directions were sampled uniformly around a sphere. Repeats of a diffusion-weighted scheme for each scan time equivalent acquisition were acquired using multiple dynamics to avoid scanner calibration, re-shimming, and power optimization changes. Consequently, each volume of the dynamic independently contributed to the tensor calculation, and no averaging of repeated volumes was applied prior. Schemes consisting of 6, 15, and 32 directions were chosen, as these are standard default values on most clinical scanners and are therefore most readily available for

clinical implementation. Additionally, a high-resolution ($0.6 \times 0.6 \times 5 \text{ mm}^3$) multi-slice, multi-echo gradient echo (mFFE) (17) anatomical image was acquired (TR/TE/ TE = 700/7.2/8.8 ms, $\alpha = 28^\circ$, number of slices=14, 6:10 min) over the same volume to improve co-registration and to serve as a template for segmentation. The total scan time for one session was 1:02.19, including the mFFE and other survey planning and calibration scans.

Data Processing

To provide a benchmark for comparison, all diffusion-weighted (DW) images within each session were combined to compute a so-called gold standard (10). Therefore, for each acquisition a total of 193 diffusion-weighted acquisitions entered into the tensor calculation. Within each session, the three acquired scan time equivalent scans (6, 15, and 32 gradient directions) were broken down into two additional comparisons. Each protocol included the same number of $b = 0$ images (Table 1: Overview of Diffusion Sequences). For each DTI set, each diffusion weighted acquisition was registered to the mFFE (Figure 1a) using the FLIRT package from FSL v5.0.2.1 (FMRIB, Oxford, UK) (18). The $b = 0$ was directly registered to the mFFE, whereas all diffusion-weighted images were registered to the prior diffusion-weighted volume, and this transformation was concatenated with the $b = 0$ and mFFE transformation. Registration was limited to translation (in x and y) and rotation (± 5 degrees about the z -axis) in plane. The rotational component of the correction from the registration procedure was applied to the DTI gradient table (19). As shown in Figure 1b, regions of interest (ROIs) of white matter (WM) and gray matter (GM) were automatically segmented from the co-registered mFFE image using a slice-based groupwise multi-atlas procedure designed specifically for the spinal cord (20). The segmented ROIs were then transferred to the registered/calculated diffusion index maps (FA, MD, AD, RD) for each scan time equivalent scan and gold standard (Figure 1c).

Data Analysis

Following registration, the diffusion tensor calculation was estimated using Camino using a nonlinear fit (21). Fractional anisotropy (FA), mean (MD), axial (AD) and radial diffusivity (RD) maps were calculated from the eigenvalues of the diffusion tensor. CNR of each map was calculated as:

$$CNR = \frac{\bar{x}_{GM} - \bar{x}_{WM}}{\sqrt{\sigma_{GM}^2 + \sigma_{WM}^2}} \quad (1)$$

where $\bar{x}_{ROI}$ is the mean value of the ROI and σ_{ROI} is the variance of the ROI. A subjective image assessment was performed to qualitatively evaluate the effects of gradient direction scheme on DTI-derived maps.

Test-Retest Reproducibility—Test-retest reproducibility of each index map from each diffusion-weighted direction scheme was assessed using Bland-Altman (22). With Bland-Altman, each subject from each time point contributes one data point in the analysis per index; the goal is to ascertain whether the 95% confidence interval for the mean difference overlaps zero (no difference at the $\alpha = 0.05$ level) and to provide an estimate of the variation

over time. This was performed for both segmented white and gray matter. The normalized Bland-Altman difference (D_{BA}), 95% confidence interval for the difference, and the limits of agreement were used as reliability metrics. D_{BA} was defined as:

$$D_{BA} = \frac{D_{12}}{M} * 100\% \quad (2)$$

where D_{12} is the mean difference of the two sessions and M is the mean DTI value. Secondly, a non-parametric Wilcoxon signed rank was performed on each mean DTI index in white matter and gray matter at the full scan time equivalent to test whether a difference between the two time points was observed at the $\alpha = 0.05$ level.

Voxel-Wise Analysis—To identify the spatial dependency of the underlying error in DTI measurements, a voxel-wise analysis was also performed. The calculated maps from each control from the first scan were registered to one another using the mFFE anatomical images, and each DTI-derived index map was averaged over all controls, resulting in a single FA, MD, AD, and RD map representing the mean over all five controls. In order to analyze the effect of diffusion scheme for each acquisition against the gold standard to characterize the accuracy of the DTI indices, each DTI index from each sampling scheme was subtracted from the gold standard map. Accuracy was quantified as the bias, or the difference from the gold standard; precision was calculated as the standard deviation of the difference between the chosen scheme's derived maps and the gold standard (10,11). Root mean square error (RMSE) from the gold standard was chosen as the metric for quantifying the overall error. Additional ROIs were manually drawn on the mean mFFE of all controls using MIPAV (Figure 2) to comprehensively assess the spatial dependencies of error in white matter and gray matter: dorsal column (DC), right lateral column (RLC), left lateral column (LLC) and gray matter (GM). The effect size (Ω) of these measurements is reported as the difference in mean RMSE values (23). To assess inter-rater variability, a second rater manually placed ROIs and Bland Altman was used to calculate whether the 95% confidence interval of the difference between the two raters was significantly different from 0. Additionally, the intraclass correlation coefficient (ICC) of rater 1 against rater 2 for each DTI-derived index was computed. The 95% confidence intervals (CI) for the ICC was calculated using a bootstrap procedure.

Cross-Sectional Analysis—The goal of this analysis was to determine whether there was a statistically significant difference in measurement error (i) between gradient weighting schemes (6, 15, and 32), (ii) at different scan time equivalents (4.5, 9, and 18 min) and (iii) between ROIs (DC, RLC, LLC and GM). First, a non-parametric Wilcoxon signed rank test was performed on the mean RMSE for FA, MD, AD, and RD in four different ROIs between the different gradient schemes. This analysis was repeated for each scan time equivalent breakdown (4.5, 9, 18 minutes) to observe the effect of scan time on choice of sampling scheme. Second, a non-parametric Wilcoxon signed rank test was performed on the mean RMSE for the same diffusion scheme at different scan time equivalents to determine whether increased scan time would mitigate measurement error. Third, a non-parametric Wilcoxon signed rank test was performed on the mean RMSE for the same scheme in different tracts.

For all comparisons, each subject contributed four RMSE values (across four slices) to the Wilcoxon analysis, and only the first visit (session) was studied. To reduce the possibility of spurious significances, the significance threshold was chosen to be $\alpha = 0.01$.

RESULTS

Gradient Direction Scheme: Effects on DTI-derived maps

Representative DTI maps of a healthy volunteer at the C3/C4 level are displayed in Figure 3, demonstrating the impact of the number of gradient directions at different scan time equivalents. The SNR of the spinal cord of each $b = 0$ image was approximately 13. Qualitatively, the gray/white matter contrast within the spinal cord for FA, AD, and RD can be observed in all schemes, but there is noticeably poorer separation between each tissue type in the 6-direction scheme. This is particularly pronounced on the AD and RD images, where the central gray matter butterfly pattern is difficult to appreciate, specifically in the ventral horns, and there is less contrast between cerebrospinal fluid (CSF) and the spinal cord. However, as expected, the MD is not directionally sensitive and has less dependency on the number of directions (24).

Figure 3 also lists the CNR between the gray and white matter of the spinal cord for each DTI metric. In general, the CNR decreases as scan time decreases with the exception of MD. At the full scan time equivalent, the CNR of the different schemes are comparable. While the contrast between gray and white matter generally increases with more gradient directions, multiple averaging produces lower noise, and therefore, the CNR of the 6-direction scheme for FA and RD is high. As the scan time decreases, the CNR for the 6-direction scheme monotonically decreases, for FA, AD, and RD, while the CNR for the 15-direction and 32-direction schemes remains relatively consistent. At a scan time above 4.5 minutes, however, the effect of a reduced scan time on CNR is minimal for all maps sampled with the 15-direction scheme. The 32-direction scheme is more affected as there is less averaging to account for the variability in image quality.

Reproducibility

The reproducibility of each gradient scheme across two time points, choosing the intermediate scan time of 9 minutes, using Bland Altman and Wilcoxon signed rank are summarized in Table 2. The normalized mean differences of the two visits were small (below 14%) for all indices, with the largest D_{BA} being produced from the 32-direction scheme. In addition, no significant differences between the two visits were detected with Wilcoxon signed rank for any sampling scheme. The 15-direction scheme resulted in the lowest D_{BA} for the most comparisons (FA: WM $D_{BA}=0.09\%$, GM $D_{BA}=2.43\%$; MD: WM $D_{BA}=2.28\%$, GM $D_{BA}=0.07\%$; RD: WM $D_{BA}=3.1\%$), indicating a smaller variability across different time points with the 15-direction scheme overall.

Spatial Dependency of Error

Figure 4 provides a group-wise analysis of the error using data from all five controls at C3/C4. The spatial dependency and artifact contribution of each gradient scheme in the spinal cord can be clearly highlighted in the bias maps (left columns) relative to the mean

gold standard. The mean map points out any residual misalignment between multiple subjects. In addition, the RMSE (top) and bias (bottom) plots (right columns) demonstrate the spatial dependency from four different ROIs, with the error bars representing inter-subject variability. In the top panel, the difference in FA (FA) from the gold standard is shown. It is clear that the difference map is most ordered in the 6-direction scheme, with the gray matter visually distinct from the other substructures. The 15-direction scheme and 32-direction scheme produce maps where the cord's general features cannot be easily identified until we reach a low SNR regime of 4.5 minutes. Additionally, it should be noted that there is an upward bias in the mean FA with 6-directions (11) as indicated by the uniformly positive FA values, but this bias diminishes and accuracy in FA estimation is improved with an increase in the number of gradient directions. These trends are also seen in the AD maps, but the effects are negligible for MD. The RD maps indicate that bias is attributed from partial volume effects at the boundaries of the cord and the CSF, rather than from the gray matter structure as seen in FA and AD.

Inter-rater difference RMSE was low: D_{BA} of 0.67% (FA), 3.08% (MD), 0.07% (AD), and 5.38% (RD). The ICC (95% CI ICC) between rater 1 and rater 2 was 0.94 (0.92, 0.95), 0.87 (0.84, 0.90), 0.95 (0.93, 0.96), and 0.80 (0.74, 0.84) for FA, MD, AD, and RD respectively. Notably, we find that increased scan time from 4.5 to 18 minutes leads to reduced RMSE (e.g., for 15 directions $p=0.0004$, 0.0004, 0.001, 0.0001 for DC, RLC, LLC, GM respectively) with moderate effect size for FA (e.g., for 15 directions, η^2 [effect size]=0.02, 0.03, 0.03, 0.05 or 3%, 4.4%, 4.4%, 9% for DC, RLC, LLC, GM). MD was generally stable at a scan time greater than 4.5 minutes ($\eta^2 < 0.05$, 0.03, 0.08 or 5%, 3%, 8% for 6, 15, 32 directions) while AD and RD mirrored the effects seen in FA. Bias contributed to approximately half of the total error FA and AD, approximately 1/4 of the error to RD, and nominally to MD (Figure 4).

For all indices, however, the reduction in RMSE from 9 minutes to 18 minutes is minimal for the 15-direction scheme (for FA, $|\eta^2| < 3.4\%$ and $\eta^2 = -3.7\%$; for MD, $|\eta^2| < 3\%$ and $\eta^2 = -2.4\%$; for AD, $|\eta^2| < 3.7\%$ and $\eta^2 = -3.1\%$; for RD, $|\eta^2| < 5.9\%$ and $\eta^2 = -3.6\%$ for WM and GM respectively), and remains lower than the inter-subject variability. Scan time equivalent (STE) sequences with 15 and 32 directions exhibited small differences in error (e.g., for 9 minutes, $\eta^2 < 1\%$ and $\eta^2 = 1.8\%$ for WM and GM at a $p > 0.39$ and $p = 0.82$ in FA with similar relative differences seen in AD, RD and MD). However, the 6-direction scheme was substantively different than 15 directions. For FA at 9 minutes, $\eta^2 < 4.4\%$ and $\eta^2 = 5.5\%$ for WM and GM ($p < 0.01$), and the difference in error was seen to increase with increased scan time; at 18 minutes, $\eta^2 < 4.4\%$ and $\eta^2 = 7.2\%$ WM and GM ($p < 0.01$).

For FA, the error between GM and DC was significantly different ($p=0.001$, 0.01, 0.02 and $\eta^2=0.02$, 0.03, 0.01 or 3.3%, 4.9%, 1.6% for 6, 15, and 32 directions). It should also be noted that at 9 minutes, the difference in RMSE between the RLC and LLC is lowest for all DTI metrics with the 15-directions (e.g. for FA, $\eta^2=0.08$, 0.007, 0.03 or 11.8%, 1.0% or 4.4% for 6, 15, and 32 directions); no difference is detected between the two columns (e.g. for FA, $p=0.94$, 0.37, 0.23 for 6, 15, and 32 directions).

DISCUSSION

In this work, we investigated the accuracy and precision of FA, MD, AD, and RD in scan time equivalent acquisitions for the cervical spinal cord. The goal was to provide an estimate of the benefits and consequences when designing pulse sequences to be deployed in the clinic for evaluation of various disease states, highlighting the magnitude of error for the most commonly used gradient schemes. From our findings, we observe that at a clinically relevant scan time of 9 minutes, a 15-direction gradient scheme produces the lowest error in diffusion measurements of all gradient schemes tested and provides DTI maps with high and reproducible contrast. Furthermore, we find that there is no benefit in using scan times over 9 minutes with the 15-direction scheme, as the reduction in RMSE from 9 to 18 minutes is lower than the inter-subject variability and any benefits in doubling scan time are negligible. Therefore, we conclude that a 15-direction gradient scheme of 9 minutes can accurately quantify tissue microstructure of the cervical spinal cord at 3T.

Recent studies of the spinal cord indicate the large disagreement in the applied number of directions (25–28). When establishing a DTI protocol for clinical implementation, it is critical to consider the number of gradient directions of the DTI scheme in terms of the overall goal of the study. Consistent with other studies, we observed that a scan time equivalent gradient scheme of 6-directions is significantly worse than a gradient scheme of 15- or 32-directions in estimation for FA (15), but we further reveal that these trends are also observed for AD and RD. Therefore, when referencing the literature and performing future studies, comparisons of protocols using different gradient schemes should be cautiously analyzed. Simulations of the brain have identified a scheme of 20 directions is optimal for robust anisotropy estimation, and at least 30 are required for estimation of mean diffusivity (29). Our choice to compare 6-, 15- and 32-direction schemes enabled us to test whether these findings could be easily translatable to the cervical spinal cord on most clinical MRI scanners. Furthermore, we investigated the error in gradient schemes at three different scan time equivalents. The 18-minute scan time equivalent demonstrates the most extreme case we are capable of performing, providing differences in gradient schemes at a high SNR regime while minimizing variability (10). While increased scan time led to reduced RMSE, the trends observed at the full scan time equivalent were mirrored at the other scan time equivalents, albeit with reduced effect. We use this finding to guide the user to implement the maximum available scan time equivalent when possible to minimize spurious conclusions.

In terms of contrast, a balance between acquisition time and gradient scheme was observed. With 32-directions, the CNR was lowest for FA and RD. This can be attributed to the fact that in the 32-direction gradient scheme there is an increased sampling along the longitudinal (and primary diffusion) axis, yielding increased diffusion and an attenuation of signal resulting in increased noise. The CNR for the 6-direction schemes should not be overanalyzed, as the known upward bias in FA in lower SNR regimes may incorrectly detect contrast (11). Furthermore, it was observed that multiple averages minimize variability caused by image corruption due to patient movement and inherent inconsistencies in hardware of the MR system. It should also be noted that choice of b -value will influence

CNR, and in this study, we chose a b -value of 750 s/mm^2 to provide high gray and white matter contrast while maintaining sufficient SNR, as suggested by Summers *et al.* (30).

Assessment of reproducibility of DTI measurements is vital when evaluating the clinical utility of DTI in the spinal cord. In order to perform longitudinal or multisite studies, the variation of the measurement error of different tracts must be known in order to determine whether observed differences are normal or abnormal. We found that the mean difference between two scans was small, but the 32-direction scheme produced the largest D_{BA} values due to the increased variability when no averaging was applied for this scheme at the 9-minute scan time equivalent. Cardiac triggering was applied in this study to reduce signal intensity variations (30), and may have strengthened the reproducibility of these measurements.

Given these combined results, we conclude that in clinical studies where the goal is a general characterization of quantitative indices in each spinal cord tissue type, a 15-direction scheme is sufficient for reliable tensor estimation and the time saved from sampling fewer gradient directions can be leveraged for greater SNR and sensitivity. In general, for clinically relevant applications (acquisition time $< 10 \text{ min}$), we recommend the use of a 15-direction gradient scheme with two averages (9 minutes) for accurate estimation of FA, AD, and RD. When investigating the intricacies of the spinal cord microstructure in greater detail, however, a higher sampling rate (32 directions over 15) may be beneficial to resolve boundaries of the spinal cord, as fewer directions may lead to greater variability in the principal eigenvector (PEV) orientation, which is an important consideration for tractography. This may be important for tissues that are damaged or when following the longitudinal evolution of diseases. For example, when it is desired to observe features such as crossing fibers, schemes with higher angular resolution such as HARDI and DBSI have been implemented, as an increase in the number of gradient directions is necessary to distinguish partial volume effects (31–33).

While others have reported optimization of DTI protocols in the spinal cord (13,15), this study provides two main distinctions. First, this study provides a comprehensive analysis using all DTI-derived indices in the spinal cord, where it is not transparent which direction scheme is optimal for minimizing variance in the derived indices while maximizing contrast. This type of experiment provides important findings, as acquisition time can be used more effectively to provide higher sensitivity, resolution, or SNR rather than to acquire unnecessary gradient directions (redundant data) for relatively oriented tensors such as in the spinal cord. Furthermore, the consideration of diffusivities other than FA is integral to comprehensively assessing pathological processes. Oh *et al.* (34) observed RD to be the most distinguishable index in their study of MS patients versus healthy controls, but current studies on optimization of DTI in the spinal cord have only considered the effect of acquisition parameters on FA. We demonstrated that accurate estimation of RD is directionally sensitive, and therefore, findings from this study can offer new opportunities for comprehensively studying the role of the spinal cord in diseased states. Second, this type of study not only provides an optimal protocol for spinal cord DTI, but additionally, it quantifies the impact of the sacrifices (i.e. lower number of gradient directions or fewer excitations) that must be made for clinical implementation of spinal cord DTI on DTI-

derived indices. Given the RMSE and bias of each gradient scheme, the results of this study can also be used to build the foundation to perform power tests to calculate sample sizes for future diffusion studies. In addition, this study implements more recent findings of DTI that have not been thoroughly investigated for protocol optimization such as imaging the spinal cord axially (14), the use of cardiac gating (30), reduced field-of-view imaging (16,35,36), and voxel-wise analysis of individual tracts (37). The voxel-wise analysis of multiple tracts highlights where the measurement is specifically failing, and the spatial dependencies of error. It should be noted that the ventral column could not be reliably determined due to partial volume effects.

This study had several limitations. First, we did not perform any fiber tractography. While an increase in angular resolution may be necessary for improved tractography of the cord, since the spinal cord primarily runs in the rostral-caudal direction, we have chosen to focus on the scalar indices under the assumption that tractography may be able to be performed with probabilistic methods, or even fewer directions than in the brain (38). This study could also be improved by using a larger sample size of healthy controls and including abnormal patients. Patients with neurological disorders may have movement impairments, which may additionally affect the spatial dependency of error in our DTI-derived measurements which we sought to minimize. However, it is noted that larger studies in patient populations would provide improved insight into the direct clinical impact that these optimization strategies may have. Technological advancements, such as simultaneous multislice (SMS) imaging, may help expedite diffusion scans and further minimize errors (39). Finally, it should be noted that the reduced field-of-view excitation scheme chosen for this work is currently not product on all clinical scanners, however it can easily be implemented using 2D excitation, outer volume suppression methods or saturation bands, which are currently available features on most clinical scanners without further software modification.

Looking forward, though these guidelines were optimized for the cervical spine, they can easily be adapted for the thoracic and lumbar spine or any anatomy where their fibers align along a single, largely coherent direction (peripheral nervous system, optic nerve). When considering smaller structures, the SNR is significantly reduced due to a need for higher resolution. In the condition where the SNR is not as high as it is in the cervical cord, the directional dependencies and need for more averages may play a role in choosing the best sequence to obtain quality indices. Moreover, it may be advantageous to develop an optimized gradient direction scheme, rather than using the basic schemes available on the scanner, for more precise and accurate tensor estimation (40).

In conclusion, the results of this study provide the underlying error and variation of different gradient schemes on estimation of FA, MD, AD, and RD. Taken together, the observed results demonstrate the efficiency of the 15-direction scheme in minimizing error when characterizing the spinal cord's tracts overall.

Acknowledgments

Grant Support: NIH/NINDS R21 NS087465-01, NIH/NEI R01 EY023240, DoD W81XWH-13-0073, NMSS RG-1501-02840

We would like to thank Ms. Kristen George-Durrett, Ms. Clair Kurtenbach, Ms. Leslie McIntosh, and Mr. Chris Thompson, who have provided invaluable assistance with scheduling and subject assistance.

References

1. Alexander AL, Lee JE, Lazar M, Field AS. Diffusion tensor imaging of the brain. *Neurotherapeutics*. 2007; 4:316–329. [PubMed: 17599699]
2. Basser PJ, Pierpaoli C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of magnetic resonance Series B*. 1996; 111:209–219. [PubMed: 8661285]
3. Song SK, Yoshino J, Le TQ, et al. Demyelination increases radial diffusivity in corpus callosum of mouse brain. *Neuro Image*. 2005; 26:132–140. [PubMed: 15862213]
4. Bozzali M, Falini A, Franceschi M, et al. White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. *Journal of neurology, neurosurgery, and psychiatry*. 2002; 72:742–746.
5. Kandel, ER.; Schwartz, JH.; Jessell, TM. *Principles of neural science*. McGraw-hill; New York: 2000.
6. Clark CA, Werring DJ. Diffusion tensor imaging in spinal cord: methods and applications - a review. *NMR Biomed*. 2002; 15:578–586. [PubMed: 12489104]
7. Anderson AW. Theoretical analysis of the effects of noise on diffusion tensor imaging. *Magn Reson Med*. 2001; 46:1174–1188. [PubMed: 11746585]
8. Skare S, Hedehus M, Moseley ME, Li TQ. Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. *J Magn Reson*. 2000; 147:340–352. [PubMed: 11097823]
9. Jones DK, Basser PJ. Squashing peanuts and smashing pumpkins": how noise distorts diffusion-weighted MR data. *Magn Reson Med*. 2004; 52:979–993. [PubMed: 15508154]
10. Landman BA, Farrell JAD, Jones CK, Smith SA, Prince JL, Mori S. Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1. 5T. *Neuro Image*. 2007; 36:1123–1138. [PubMed: 17532649]
11. Farrell JAD, Landman BA, Jones CK, et al. Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1. 5T. *J Magn Reson Im*. 2007; 26:756–767.
12. Ni H, Kavcic V, Zhu T, Ekholm S, Zhong J. Effects of number of diffusion gradient directions on derived diffusion tensor imaging indices in human brain. *AJNR*. 2006; 27:1776–1781. [PubMed: 16971635]
13. Lee JW, Kim JH, Kang HS, et al. Optimization of acquisition parameters of diffusion-tensor magnetic resonance imaging in the spinal cord. *Invest Radiol*. 2006; 41:553–559. [PubMed: 16772848]
14. Wheeler-Kingshott CAM, Hickman SJ, Parker GJM, et al. Investigating Cervical Spinal Cord Structure Using Axial Diffusion Tensor Imaging. *Neuro Image*. 2002; 16:93–102. [PubMed: 11969321]
15. Santarelli X, Garbin G, Ukmar M, Longo R. Dependence of the fractional anisotropy in cervical spine from the number of diffusion gradients, repeated acquisition and voxel size. *Magnetic resonance imaging*. 2010; 28:70–76. [PubMed: 19577395]
16. Wilm BJ, Svensson J, Henning A, Pruessmann KP, Boesiger P, Kollias SS. Reduced field-of-view MRI using outer volume suppression for spinal cord diffusion imaging. *Magn Reson Med*. 2007; 57:625–630. [PubMed: 17326167]
17. Held P, Dorenbeck U, Seitz J, Fründ R, Albrich H. MRI of the abnormal cervical spinal cord using 2D spoiled gradient echo multiecho sequence (MEDIC) with magnetization transfer saturation pulse. A T2* weighted feasibility study. *Journal of neuroradiology*. 2003; 30:83–90.
18. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. *Fsl*. Neuroimage. 2012; 62:782–790. [PubMed: 21979382]

19. Leemans A, Jones DK. The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn Reson Med*. 2009; 61:1336–1349. [PubMed: 19319973]
20. Asman AJ, Bryan FW, Smith SA, Reich DS, Landman BA. Groupwise multi-atlas segmentation of the spinal cord's internal structure. *Med Image Anal*. 2014; 18:460–471. [PubMed: 24556080]
21. Cook, PA.; Bai, Y.; Nedjati-Gilani, S., et al. ISMRM PottSMot. Camino: Open-Source Diffusion-MRI Reconstruction and Processing. Seattle, WA, USA: 2006. p. 2759
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1:307–310. [PubMed: 2868172]
23. Murphy RK, Sun P, Xu J, et al. Magnetic Resonance Imaging Biomarker of Axon Loss Reflects Cervical Spondylotic Myelopathy Severity. *Spine*. 2015
24. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. Diffusion tensor MR imaging of the human brain. *Radiology*. 1996; 201:637–648. [PubMed: 8939209]
25. Choe AS, Belegu V, Yoshida S, et al. Extensive neurological recovery from a complete spinal cord injury: a case report and hypothesis on the role of cortical plasticity. *Front Hum Neurosci*. 2013; 7:290–305. [PubMed: 23805087]
26. Koskinen E, Brander A, Hakulinen U, et al. Assessing the state of chronic spinal cord injury using diffusion tensor imaging. *Journal of neurotrauma*. 2013; 30:1587–1595. [PubMed: 23758292]
27. Kamble RB, Venkataramana NK, Naik AL, Rao SV. Diffusion tensor imaging in spinal cord injury. *The Indian Journal of Radiology & Imaging*. 2011; 21:221–224. [PubMed: 22013299]
28. Song T, Chen W-J, Yang B, et al. Diffusion tensor imaging in the cervical spinal cord. *European Spine Journal*. 2011; 20:422–428. [PubMed: 20938788]
29. Jones DK. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. *Magn Reson Med*. 2004; 51:807–815. [PubMed: 15065255]
30. Summers P, Staempfli P, Jaermann T, Kwiecinski S, Kollias S. A preliminary study of the effects of trigger timing on diffusion tensor imaging of the human spinal cord. *AJNR American journal of neuroradiology*. 2006; 27:1952–1961. [PubMed: 17032874]
31. Cohen-Adad J, Benali H, Hoge RD, Rossignol S. In vivo DTI of the healthy and injured cat spinal cord at high spatial and angular resolution. *Neuroimage*. 2008; 40:685–697. [PubMed: 18201909]
32. Wang X, Cusick MF, Wang Y, et al. Diffusion basis spectrum imaging detects and distinguishes coexisting subclinical inflammation, demyelination and axonal injury in experimental autoimmune encephalomyelitis mice. *NMR Biomed*. 2014; 27:843–852. [PubMed: 24816651]
33. Cohen-Adad J, El Mendili MM, Lehericy S, et al. Demyelination and degeneration in the injured human spinal cord detected with diffusion and magnetization transfer MRI. *Neuro Image*. 2011; 55:1024–1033. [PubMed: 21232610]
34. Oh J, Zackowski K, Chen M, et al. Multiparametric MRI correlates of sensorimotor function in the spinal cord in multiple sclerosis. *Mult Scler J*. 2012; 19:427–435.
35. Dowell NG, Jenkins TM, Ciccarelli O, Miller DH, Wheeler-Kingshott CAM. Contiguous-slice zonally oblique multislice (CO-ZOOM) diffusion tensor imaging: Examples of in vivo spinal cord and optic nerve applications. *J Magn Reson Im*. 2009; 29:454–460.
36. Saritas EU, Cunningham CH, Lee JH, Han ET, Nishimura DG. DWI of the spinal cord with reduced FOV single-shot EPI. *Magn Reson Med*. 2008; 60:468–473. [PubMed: 18666126]
37. Toosy AT, Kou N, Altmann D, Wheeler-Kingshott CA, Thompson AJ, Ciccarelli O. Voxel-based cervical spinal cord mapping of diffusion abnormalities in MS-related myelitis. *Neurology*. 2014; 83:1321–1325. [PubMed: 25186861]
38. Lebel C, Benner T, Beaulieu C. Six is enough? Comparison of diffusion parameters measured using six or more diffusion-encoding gradient directions with deterministic tractography. *Magn Reson Med*. 2012; 68:474–483. [PubMed: 22162075]
39. Setsompop K, Cohen-Adad J, Gagoski BA, et al. Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *Neuro Image*. 2012; 63:569–580. [PubMed: 22732564]
40. Jones DK, Horsfield MA, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn Reson Med*. 1999; 42:515–525. [PubMed: 10467296]

Figure 1. Flowchart describing image processing scheme

A) Each scan session acquired an anatomical image, which was used for registration and segmentation of white and gray matter. A 2D affine registration was performed from DTI to mFFE space. These segmented ROIs (B) were then transferred to the registered DTI set (C).

Figure 2. Additional ROIs overlaid on mFFE for error map processing

Yellow: gray matter, light blue: right lateral column, green: left lateral column, dark blue: dorsal column

Figure 3. DTI maps of a representative healthy volunteer acquired from different gradient schemes at different scan time equivalents

Contrast-to-noise ratio (CNR) of each DTI map is listed below each map. At the full scan time equivalent, the CNR of the different schemes are comparable. However, as scan time decreases, in general, the CNR drops for the 6- and 32-direction scheme, while the 15-direction scheme with 2 averages remains relatively stable.

Figure 4. Error bias maps, calculated as the difference from the gold standard, using all control data for gradient schemes of 6, 15 and 32 directions at 4.5, 9, and 18 minutes

Increased scan time leads to reduced RMSE. MD was relatively stable, while the effect size for FA, AD and RD was moderate. Scan time equivalent sequences of 15 and 32 exhibited small differences in error, but 6-directions was significantly different from 15-directions.

Table 1

Overview of Diffusion Sequences

	Directions	Number of Repetitions	Added b0s	Scan Time
Acquired	6	9		18:15
	15	4		17:15
	32	2		17:15
Full STE (~18 min)	6	9	0	
	15	4	5	
	32	2	7	
Processing Breakdown				
Half STE (~9 min)	6	5	0	
	15	2	3	
	32	1	4	
Shortest STE (~4.5 min)	6	3	0	
	15	1	3	
	32	-	-	

Table 2

Mean ($\pm$ standard deviation) of DTL-derived parameters over all participants for both scans at the 9-minute breakdown for white matter (WM) and gray matter (GM)

Results from Bland Altman are also listed, including the mean difference ($\pm$ standard deviation), 95% confidence interval for the difference, the Bland Altman limit of agreement statistic, the normalized difference, and the Wilcoxon signed rank (WSR) p -value between the two scans.

EA	Scan 1		Scan2		Bland Altman			WSR	
	<i>mean$\pm$std</i>	<i>mean$\pm$std</i>	<i>mean$\pm$std</i>	<i>mean$\pm$std</i>	<i>difference (D)</i>	<i>95% CI</i>	<i>LOA</i>	<i>D_{BA}(%)</i>	<i>p</i>
WM									
WM 6	0.73 $\pm$ 0.05	0.74 $\pm$ 0.07	0.74 $\pm$ 0.07	0.74 $\pm$ 0.07	-0.008	[-0.04, 0.02]	[-0.15, 0.13]	1.09	0.65
WM 15	0.68 $\pm$ 0.05	0.68 $\pm$ 0.07	0.68 $\pm$ 0.07	0.68 $\pm$ 0.07	-0.001	[-0.03, 0.03]	[-0.12, 0.12]	0.09	0.91
WM 32	0.69 $\pm$ 0.04	0.66 $\pm$ 0.10	0.66 $\pm$ 0.10	0.66 $\pm$ 0.10	0.036	[0.0, 0.07]	[-0.12, 0.19]	5.4	0.05
GM 6	0.61 $\pm$ 0.07	0.62 $\pm$ 0.06	0.62 $\pm$ 0.06	0.62 $\pm$ 0.06	-0.017	[-0.04, 0.01]	[-0.12, 0.09]	2.84	0.2
GM 15	0.55 $\pm$ 0.08	0.54 $\pm$ 0.08	0.54 $\pm$ 0.08	0.54 $\pm$ 0.08	0.013	[-0.01, 0.04]	[-0.10, 0.12]	2.43	0.41
GM 32	0.57 $\pm$ 0.09	0.56 $\pm$ 0.10	0.56 $\pm$ 0.10	0.56 $\pm$ 0.10	0.014	[-0.01, 0.04]	[-0.10, 0.13]	2.53	0.26
MD ($\mu\text{m}^2/\text{ms}$)									
WM 6	1.03 $\pm$ 0.12	1.00 $\pm$ 0.08	1.00 $\pm$ 0.08	1.00 $\pm$ 0.08	0.032	[-0.02, 0.08]	[-0.18, 0.24]	3.15	0.22
WM 15	1.00 $\pm$ 0.07	0.97 $\pm$ 0.06	0.97 $\pm$ 0.06	0.97 $\pm$ 0.06	0.022	[-0.01, 0.05]	[-0.11, 0.15]	2.28	0.28
WM 32	0.96 $\pm$ 0.06	1.02 $\pm$ 0.13	1.02 $\pm$ 0.13	1.02 $\pm$ 0.13	-0.053	[-0.11, 0.01]	[-0.31, 0.20]	5.3	0.11
GM 6	0.99 $\pm$ 0.12	0.99 $\pm$ 0.09	0.99 $\pm$ 0.09	0.99 $\pm$ 0.09	0.004	[-0.03, 0.04]	[-0.16, 0.17]	0.42	0.91
GM 15	0.94 $\pm$ 0.06	0.94 $\pm$ 0.07	0.94 $\pm$ 0.07	0.94 $\pm$ 0.07	0.001	[-0.04, 0.04]	[-0.16, 0.16]	0.07	0.94
GM 32	0.92 $\pm$ 0.07	0.94 $\pm$ 0.11	0.94 $\pm$ 0.11	0.94 $\pm$ 0.11	-0.022	[-0.08, 0.03]	[-0.26, 0.22]	2.34	0.55
AD ($\mu\text{m}^2/\text{ms}$)									
WM 6	2.11 $\pm$ 0.27	2.06 $\pm$ 0.17	2.06 $\pm$ 0.17	2.06 $\pm$ 0.17	0.052	[-0.02, 0.13]	[-0.27, 0.37]	2.48	0.39
WM 15	1.92 $\pm$ 0.13	1.89 $\pm$ 0.13	1.89 $\pm$ 0.13	1.89 $\pm$ 0.13	0.035	[-0.02, 0.09]	[-0.19, 0.26]	1.82	0.23
WM 32	1.89 $\pm$ 0.12	1.91 $\pm$ 0.10	1.91 $\pm$ 0.10	1.91 $\pm$ 0.10	-0.014	[-0.07, 0.04]	[-0.26, 0.23]	0.75	0.65
GM 6	1.79 $\pm$ 0.23	1.79 $\pm$ 0.20	1.79 $\pm$ 0.20	1.79 $\pm$ 0.20	-0.009	[-0.07, 0.05]	[-0.27, 0.25]	0.48	0.55
GM 15	1.62 $\pm$ 0.13	1.59 $\pm$ 0.12	1.59 $\pm$ 0.12	1.59 $\pm$ 0.12	0.029	[-0.05, 0.11]	[-0.30, 0.35]	1.83	0.53
GM 32	1.60 $\pm$ 0.19	1.59 $\pm$ 0.13	1.59 $\pm$ 0.13	1.59 $\pm$ 0.13	0.013	[-0.09, 0.11]	[-0.43, 0.45]	0.8	0.88
RD ($\mu\text{m}^2/\text{ms}$)									
WM 6	0.49 $\pm$ 0.08	0.47 $\pm$ 0.10	0.47 $\pm$ 0.10	0.47 $\pm$ 0.10	0.022	[-0.03, 0.08]	[-0.21, 0.26]	4.63	0.46

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

FA	Scan 1		Scan2		Bland Altman			WSR	
	mean±std		mean±std		difference (D)	95% CI	LOA	D _{BA} (%)	p
WM 15	0.53±0.08		0.52±0.10		0.016	[-0.03, 0.06]	[-0.17, 0.20]	3.1	0.41
WM 32	0.50±0.07		0.57±0.18		-0.07	[-0.14, 0.00]	[-0.38, 0.24]	13.36	0.09
GM 6	0.59±0.11		0.58±0.08		0.01	[-0.03, 0.05]	[-0.15, 0.18]	1.78	0.5
GM 15	0.61±0.10		0.62±0.10		-0.016	[-0.05, 0.2]	[-0.16, 0.12]	2.54	0.37
GM 32	0.57±0.08		0.61±0.15		-0.039	[-0.09, 0.01]	[-0.24, 0.17]	6.57	0.16