AJNR

This information is current as of August 8, 2025.

Test-Retest and Interreader Reproducibility of Semiautomated Atlas-Based Analysis of Diffusion Tensor Imaging Data in Acute Cervical Spine Trauma in Adult Patients

D.J. Peterson, A.M. Rutman, D.S. Hippe, J.G. Jarvik, F.H. Chokshi, M.R. Reyes, C.H. Bombardier and M. Mossa-Basha

Test-Retest and Interreader Reproducibility of Semiautomated Atlas-Based Analysis of Diffusion Tensor Imaging Data in Acute Cervical Spine Trauma in Adult Patients

D.J. Peterson, A.M. Rutman, D.S. Hippe, J.G. Jarvik, F.H. Chokshi, M.R. Reyes, C.H. Bombardier, and M. Mossa-Basha

ABSTRACT

BACKGROUND AND PURPOSE: DTI is a tool for microstructural spinal cord injury evaluation. This study evaluated the reproducibility of a semiautomated segmentation algorithm of spinal cord DTI.

MATERIALS AND METHODS: Forty-two consecutive patients undergoing acute trauma cervical spine MR imaging underwent 2 axial DTI scans in addition to their clinical scan. The datasets were put through a semiautomated probabilistic segmentation algorithm that selected white matter, gray matter, and 24 individual white matter tracts. Regional and white matter tract volume, fractional anisotropy, and mean diffusivity values were calculated. Two readers performed the nonautomated steps to evaluate interreader reproducibility. The coefficient of variation and intraclass correlation coefficient were used to assess test-retest and interreader reproducibility.

RESULTS: Of 42 patients, 30 had useable data. Test-retest reproducibility of fractional anisotropy was high for white matter as a whole (coefficient of variation, 3.8%; intraclass correlation coefficient, 0.93). Test-retest coefficient-of-variation ranged from 8.0%–18.2% and intraclass correlation coefficients from 0.47–0.80 across individual white matter tracts. Mean diffusivity metrics also had high test-retest reproducibility (white matter: coefficient-of-variation, 5.6%; intraclass correlation coefficient, 0.86) with coefficients of variation from 11.6%–18.3% and intraclass correlation coefficients from 0.57–0.74 across individual tracts, with better agreement for larger tracts. The coefficients of variation of fractional anisotropy and mean diffusivity both had significant negative relationships with white matter volume (26%–27% decrease for each doubling of white matter volume, $P < .01$).

CONCLUSIONS: DTI spinal cord segmentation is reproducible in the setting of acute spine trauma, specifically for larger white matter tracts and total white or gray matter.

ABBREVIATIONS: CV = coefficient of variation; ICC = intraclass correlation coefficient; FA = fractional anisotropy; MD = mean diffusivity

D TI is a technique that provides microstructural evaluation not afforded by conventional MR imaging techniques.[1] In various disease states, DTI has been extensively investigated in brain applications and can detect abnormalities in otherwise normal-appearing brain regions[2,3] and is able to predict outcomes.[4] Early DTI use shows promise in detecting spinal cord abnormalities associated with spinal cord injury,[5,6] demyelinating diseases,[7]

spondylotic myelopathy,[8] HIV myelopathy,[9] and various inflammatory and vascular myelopathies.[10] In acute spinal cord trauma, DTI has shown value in assessing microstructural injury, differentiating between hemorrhagic and nonhemorrhagic contusions, and strong correlation with clinical injury scores.[5]

Similar to brain DTI, tract-based white matter analysis of the spinal cord may offer additional insight into white matter characteristics in both healthy and diseased states.[11-13] Current methods of evaluating spine DTI data, however, are either purely qualitative assessments or labor-intensive hand-drawn ROIs that may be prone to reader-related variability/imprecision and poor reproducibility. In brain DTI, completely automated methods are available to reliably parcel the brain,[14] with application to clinical care.[15] Recently, a set of tools has been released as part of the

Indicates open access to non-subscribers at www.ajnr.org

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photos.

http://dx.doi.org/10.3174/ajnr.A5334

"Spinal Cord Toolbox" that can allow for spinal cord registration, segmentation, and parcellation.[16] The Spinal Cord Toolbox (https://www.nitrc.org/projects/sct/) has been applied to flaccid myelitis on T2-weighted imaging,[17] functional imaging of the spine,[18] and T2*, DTI, and inhomogeneous magnetization transfer sequences in healthy patients at a range of ages.[19] To date, evaluation of the reproducibility of spinal cord segmentation and analysis algorithms such as the Spinal Cord Toolbox when using DTI sequences has been lacking. In addition, the reproducibility of DTI in the setting of acute spinal cord trauma has yet to be evaluated. Determination of these characteristics is particularly important in the setting of trauma evaluation, where the presence of factors such as pain or cognitive dysfunction from associated injuries, medication effect, susceptibility artifact from metallic fusion hardware, or the presence of external lines may impact image acquisition and interpretation.

In this study, we evaluated the test-retest reproducibility of a semiautomated atlas-based technique for extracting tract-specific and level-specific diffusion metrics in patients with acute cervical spine trauma. Furthermore, we also assessed the influence of reader-induced variability on the parcellation process.

MATERIALS AND METHODS

Subjects

After institutional review board approval (Harborview Medical Center), 42 consecutive patients presenting with acute cervical spinal trauma were prospectively recruited through an institutional review board–approved waiver of consent and scanned by using an imaging protocol that included 2 separate axial DTI acquisitions. Inclusion criteria were: 1) clinical concern for cervical spinal cord injury, 2) undergoing MR imaging of the cervical spine within 72 hours of initial injury, and 3) adult patient ≥18 years of age. The exclusion criteria were: 1) spine surgery or hardware for treatment of spinal injury within the scanning field, 2) pregnancy, and 3) known spinal cord disease or previous injury that would affect DTI metrics.

MR Imaging Acquisition

MR imaging scanning was performed on a 3T Trio scanner (Siemens, Erlangen, Germany). The scanning protocol included 2D sagittal T1 FLAIR, T2, STIR, axial T2, and 2-axial DTI acquisitions. For the current study, axial DTI and sagittal STIR sequences were used in processing and analysis, and thus, the parameters are listed. Axial DTI sequences are single-shot echo-planar acquisitions with reduced field of view in the anteroposterior dimension and 10 directions of diffusion, which were acquired during the same scan session with the following parameters: TR, 2600 ms; TE, 90 ms; 0.85×0.85 mm in-plane resolution; 200 mm × 100 mm field of view; section thickness, 5 mm; 0 intersection gap; 6 averages; bandwidth, 1766 Hz/pixel; and generalized autocalibrating partially parallel acquisition, 2. For each acquisition, images were acquired with spherically distributed b-vectors at a b-value of 750 seconds/mm^2, along with 6 interspersed minimally weighted B0 volumes. A total of 22–28 sections were acquired for the DTI scan with 11–14 cm of coverage in the foot-to-head direction depending on needed coverage, with a scan time of 3 minutes 30 seconds to 4 minutes 28 seconds. The advanced shim mode and dynamic field correction options were activated to reduce B0 and eddy current distortions, respectively. Axial DTI ex-

tended from the foramen magnum to the C7–T1 vertebral body level in the craniocaudal direction. The sagittal STIR sequence had the following parameters: TR, 3700 ms; TE, 47 ms; field of view, 220 × 220 mm; TI, 230 ms; section thickness, 3 mm; in-plane resolution, 1 × 0.7 mm; parallel imaging acceleration, 2; 2 averages; and bandwidth, 252 Hz/pixel. The axial DTI scans were performed at the beginning and end of the MR imaging scan. Before performance of the final DTI scan, the scanner table was removed from the MR imaging scanner, with removal of the detachable coils, and the patient was repositioned in craniocaudal and right-to-left directions. The coils were subsequently placed again and the table reintroduced into the scanner. Localizers were repeated and the DTI scan field of view was repositioned for the second scan. Patients were not removed from the scanner table or room because of concern for patient safety relating to injuries that would limit patient mobility and function and could potentially result in additional patient discomfort and/or injury. In this setting, we felt that patient repositioning and relocalization would be sufficient for reproducibility assessment.
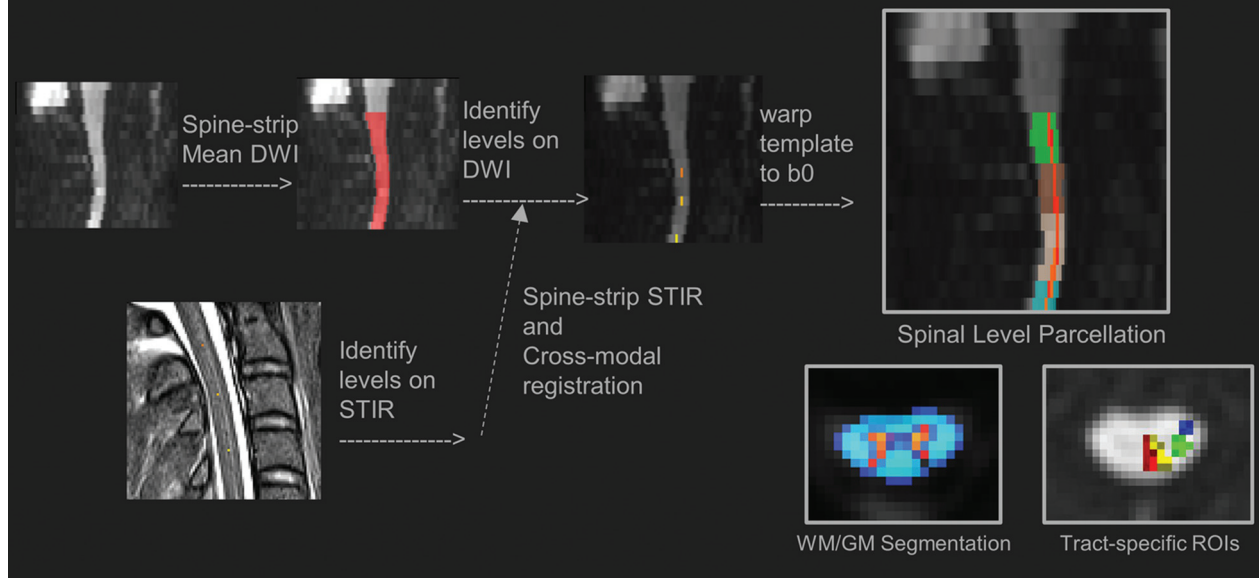
Analysis Pipeline

An image analysis pipeline was constructed, which takes an axial DTI dataset and a sagittal STIR image as input and gives vertebral level-specific DTI metrics in total gray matter and white matter, regional white matter (dorsal, lateral, and ventral), and within 30 labeled white matter tracts (see On-line Tables 1–3 for individual white matter tracts analyzed). The core of the parcellation procedure is a coordinate transformation that maps the subject-space spinal image onto a labeled template. This transformation is the concatenation of a section-wise spine-straightening transformation, then an affine transformation, and then a final nonlinear warp. This warp is a symmetric normalization transformation[20] as implemented in Advanced Normalization Tools (http://stnava.github.io/ANTs/)[21] by using mutual information as the cost function. The template used is the MNI-poly-AMU template with labeled probabilistic ROIs. This template was created by labeling, co-registering and averaging high-quality T2 images from 16 healthy patients. A complete description of the template is available in Fonov et al[22] and in the Spinal Cord Toolbox on-line documentation (https://sourceforge.net/p/spinalcordtoolbox/wiki/MNI-Poly-AMU/).

The full parcellation pipeline includes the following steps (summarized in Fig 1), with command-line utilities in single quotes: 1) Diffusion datasets were corrected for motion and eddy current–induced distortions by using 'sct_dmri_moco'. 2) The tensor was calculated by using FSL's 'dtifit' (http://fsl.fmrib.ox.ac.uk/fsl/fsl-4.1.9/fdt/fdt_dtifit.html) by using weighted least squares fitting. 3) A spinal cord–stripping routine ('sct_propseg') was then run separately on both the STIR image and the mean DWI. This routine models the spinal cord surface as a tubular mesh, which is deformed until it matches the edges of the spinal cord. 4) At this point, manual intervention is required to identify the vertebral levels, which is done on the STIR image by placing single-point 3D ROIs, performed by a reader blinded to clinical information and other imaging findings. 5) After step 3 is finished, the vertebral levels are mapped onto the diffusion images by crossmodal registration of the STIR to the mean DWI by using

FIG 1. Atlas-based parcellation pipeline of DTI of the cervical spinal cord.

'sct_register_multimodal'. After determining the STIR-to-mean DWI transformation, the same transformation was applied to the 3 manually defined point ROIs, yielding mean DWI with labeled vertebral levels. 6) The MNI-poly-AMU spinal cord atlas[22] was then registered to the diffusion dataset with 'sct_warp_template'. This utility finds the subject-to-template composite transformation described above, then applies the reverse transformation to the template, thereby mapping the atlas ROIs into individual subject space. Registration and white matter–gray matter segmentation at the level of spinal cord injury is shown in On-line Fig 1. 7) Spine and tract-level fractional anisotropy (FA) and mean diffusivity (MD) values were obtained with 'sct_extract_metrics' by using the default maximum-likelihood method, which is a Bayesian parameter estimation method that has been shown to yield more accurate values than a simple weighted average.[23] Further documentation on these utilities and a description of their default parameters is available on the Spinal Cord Toolbox Web site (https://sourceforge.net/p/spinalcordtoolbox/wiki/tools/). The processing pipeline was coded by using makefiles and run through 'make,' following the approach described by Askren et al.[24] The pipeline was run on a Debian 7 workstation, with FSL version 2.0.9 (http://www.fmrib.ox.ac.uk/fsl), and Spinal Cord Toolbox version 2.2.

To evaluate the influence of manual intervention, a second reader, a neuroradiologist with 6 years' experience interpreting spine MR imaging, independently repeated step 4 (the only step requiring manual intervention) while also blinded to clinical information and other imaging findings.

In 8 cases, the fully automated spinal cord–stripping routine failed to produce a satisfactory mask because of poor initial estimates of the center of the spinal cord. In these cases, 3 single-point ROIs were placed within the spinal cord to seed the propagating segmentation.

Statistical Analysis
Pairs of measurements from each reader or each acquisition were compared to assess test-retest and interreader reproducibility of region volume, FA, and MD. Reproducibility was assessed visually by using Bland-Altman plots. Test-retest and interreader reproducibility were also summarized by using the within-paired standard deviation, the coefficient of variation (CV) (calculated as [100% × the within-paired standard deviation]/mean) and the intraclass correlation coefficient (ICC). Qualitative interpretation of ICC values are as follows: 0–.20 = poor; 0.21–0.40 = fair; 0.41–0.60 = moderate; 0.61–0.80 = good; and 0.81–1 = excellent.[8] For test-retest reproducibility, these metrics were calculated separately for each reader, and then the metrics from the readers were averaged. Similarly, for interreader reproducibility, these metrics were calculated separately for each acquisition, and then the metrics from the acquisitions were averaged. Standard errors of these metrics were calculated by using the nonparametric bootstrap, where patients were resampled to account for dependence between vertebral levels of the same patient.[25] These standard errors were multiplied by 2 to represent the approximate 95% CI and presented as metric (2× standard error) to show the precision of the metric estimates. A generalized estimating equations log-linear model was used to estimate the linear trend between CV and white matter volume. All statistical calculations were conducted with the statistical computing language R (version 3.1.1; http://www.r-project.org/). Throughout, 2-sided t tests were used, with statistical significance defined as $P < .05$.

RESULTS
Patient Data
Of the 42 patients recruited, 12 had insufficient image quality because of motion ($n = 9$), susceptibility from metal artifact ($n = 2$),
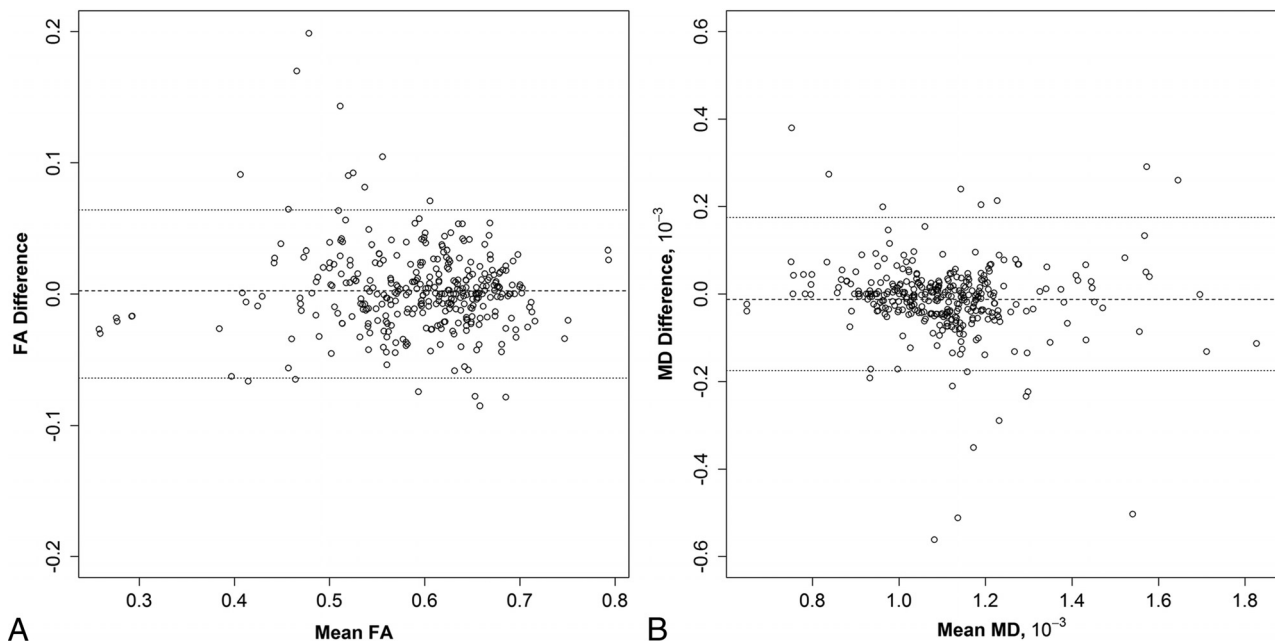
FIG 2. *A,* Bland-Altman plots of scan and re-test measurements of FA. The *dashed line* indicates the mean difference between measurements, and the *dotted lines* indicate the limits of agreement. Each reader's measurements are shown separately. *B,* Bland-Altman plots of scan and re-test measurements of MD. The *dashed line* indicates the mean difference between measurements, and the *dotted lines* indicate the limits of agreement. Each reader's measurements are shown separately.

or coverage ($n = 1$) and were excluded. Among the 30 remaining patients, 9 (30%) were woman, with ages ranging from 18–91 years (median, 41 years). At the time of imaging, 6 patients had an acute spinal cord contusion, 1 had moderate degenerative stenosis, and the remaining patients had no evidence of spinal cord injury or appreciable abnormality on conventional MR imaging or diffusion trace maps. Both readers could extract DTI metrics from 174 cervical levels (mean, 5.8 ± 1.1 per subject) from both scans.

These 174 levels were segmented into up to 30 white matter tracts. However, the left and right lateral reticulospinal, medial reticulospinal, and medial longitudinal fasciculus tracts could not be segmented on >95% of levels because of small size relative to the imaging resolution and were excluded, leaving 24 of 30 tracts available for analysis (On-line Fig 2). Among the 24 remaining tracts, the following tracts could not be segmented on some levels (reported by number not segmentable on the left/right sides out of 174 total levels): fasciculus cuneatus (1/0), fasciculus gracilis (1/0), rubrospinal (3/1), spino-olivary (0/1), tectospinal (33/35), lateral vestibulospinal (64/69), ventrolateral reticulospinal (67/72), ventral corticospinal (74/73), and ventral reticulospinal (85/88) tracts. After excluding these tract levels, of the original 4176 potential tract segments (174 × 24), there were 3509 total white matter tract segments per level available for analysis.

Test-Retest Reproducibility

White matter volume and the volume of individual white matter tracts had test-retest CVs of 7.7% and 13.2%, respectively (On-line Table 1). The volume of the ventral reticulospinal tract had the highest CV of 36.9%, but this was also the smallest tract assessed (mean volume, 5 ± 3 mm³).

The test-retest reproducibility of FA metrics was high for the white matter as a whole (CV, 3.8%; ICC, 0.93) and, to a lesser
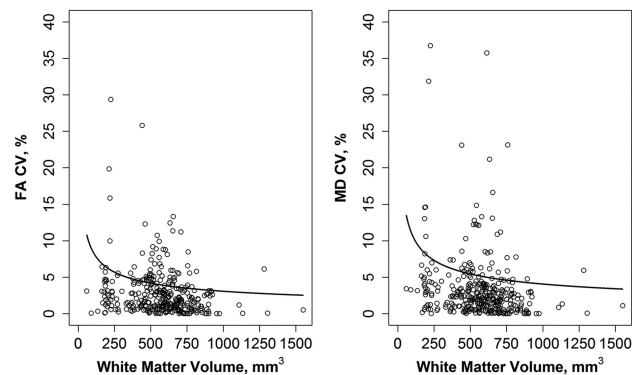
FIG 3. CV of FA (*left panel*) and MD (*right panel*), calculated per test-retest pair, versus the white matter volume. The generalized estimating equations–based estimate of mean CV is shown by the *bold curve.* CV increased with decreasing white matter tract volume for both FA ($P = .008$) and MD ($P < .001$).

extent, among all individual white matter tracts as a group (CV, 10.8%; ICC, 0.81) (On-line Table 2, Fig 2A). Across the individual white matter tracts, the test-retest CV ranged from 8.0% (fasciculus cuneatus) to 18.2% (ventral reticulospinal tract), and the test-retest ICC ranged from 0.47 (ventral reticulospinal tract) to 0.80 (lateral corticospinal tract). As noted, the ventral reticulospinal tract was the smallest tract on average, and in general, reproducibility improved with increasing white matter volume of the analyzed tract (Fig 3). The test-retest CV of FA decreased by 26.6% (95% CI, 22.6%–30.3%; $P = .008$) for each doubling of white matter volume.

Similarly, the test-retest reproducibility of MD metrics was high for the white matter as a whole (CV, 5.6%; ICC, 0.86) and for the individual white matter tracts (CV, 14.5%; ICC, 0.75) (On-line Table 3, Fig 2B), though the CV estimates ($P < .001$ and $P = .035$) and ICC estimates ($P < .001$ and $P < .001$) of MD over the

white matter and individual tracts were statistically significantly lower than those for FA. For each tract, the test-retest CV ranged from 11.6% (fasciculus cuneatus) to 18.3% (ventral reticulospinal tract), and the test-retest ICC ranged from 0.57 (ventral reticulospinal tract) to 0.74 (spino-olivary tract). As with FA, the MD of the ventral reticulospinal tract had the lowest reproducibility among the tracts assessed, likely because of its small size. Similar to FA, the test-retest CV of MD decreased by 25.5% (95% CI, 23.9%–27.1%; $P <$ .001) for each doubling of white matter volume (Fig 3).

Interreader Reproducibility
Interreader reproducibility of FA metrics was generally numerically higher than the corresponding test-retest reproducibility estimates (On-line Table 2), with the interreader CV <10% and ICC >0.80 for all individual tracts except the ventral reticulospinal tract (CV, 13.0%; ICC, 0.73) and ventral corticospinal tract (CV, 11.2%; ICC, 0.76). Interreader reproducibility of MD metrics was also numerically higher than the corresponding test-retest reproducibility estimates (On-line Table 3), though the interreader reproducibility of MD of the whole white matter and individual tracts as a group were statistically significantly lower than those for FA ($P <$ .01 for all comparisons). The interreader CV and ICC of MD were <15% and >0.70, respectively, for all individual white matter tracts except for the rubrospinal tract (CV, 16.9%; ICC, 0.65).

DISCUSSION
We report the first study to evaluate the test-retest and interreader reproducibility of semiautomated atlas-based segmentation of DTI of the cervical spinal cord in patients with acute trauma. Atlas-based parcellation of spinal cord DTI data shows good to excellent test-retest reproducibility for volume of gray matter, white matter (total), and individual tracts. Test-retest reproducibility for FA and MD for individual white matter tracts ranged from moderate to good, whereas the test-retest reproducibility was good to excellent for total white matter, gray matter, and white matter regions (ventral, lateral, and dorsal white matter stations). Cervical spinal cord tract-specific diffusion metrics are especially reproducible within the larger, major white matter tracts, with a lower degree of reproducibility in smaller white matter tracts. This is likely a product of in-plane image resolution, and with higher resolution acquisitions, DTI metrics of smaller white matter tracts of interest would likely be more reproducible and analyzable. Estimates of test-retest variability can be used for sample size planning in future longitudinal studies that use spinal cord DTI to measure outcomes. The manual step of identifying the vertebral levels and repeated cord segmentation introduces only limited variability in the extracted metrics of moderately sized and larger parcels, as shown by overall good to excellent interreader agreement.

A few prior studies have investigated the reproducibility of DTI of the spinal cord mainly by using manual ROI evaluation. In an analysis of 40 healthy control patients, Brander et al[26] used whole cord and right, left, and posterior manual ROIs as well as tractography-based analysis for quantitative DTI metric assessments. There was excellent and good intrareader and good interreader agreement for whole cord FA and ADC values, respectively, when using ICC. There was excellent intrareader

agreement for tractography-based analysis for all metrics. In 10 pediatric patients with chronic spinal cord injury, Mulcahey et al[6] assessed the scan-rescan reproducibility of DTI metrics in pediatric patients with chronic spinal cord injury by using whole cord manual ROIs drawn at each level of the cervical spinal cord. The ICC ranged from 0.50–0.89 for FA depending on the cervical spinal cord level, MD ranged from 0.80–0.95, axial diffusivity ranged from 0.82–0.94, and radial diffusivity ranged from 0.82–0.94. Smith et al[27] evaluated the scan-rescan and interreader reproducibility of DTI in 9 health volunteers, with manual ROIs placed over the right and left side of the spinal cord and over the dorsal columns. There was no significant difference between readers or between scans for each ROI placement. The normalized Bland-Altman difference for interreader assessment was 1.89%–2.06% and for test-retest evaluation was 2.38–4.54%. These studies showed interreader, intrareader, and scan-rescan reproducibility of spinal cord DTI by using manual ROI assessment in healthy volunteers and patients with chronic spinal cord injury, similar to our results that show reproducibility of DTI metrics by using semiautomated segmentation in the setting of acute trauma.

Our study is the first to evaluate spine DTI reproducibility in patients with acute cervical spine trauma in a clinical setting. This study indicates that the use of spine DTI in clinical patients with acute cervical spine trauma is feasible and reproducible. Although only 30 of 42 DTI cases had 2 sets of useable DTI data, 4 of the 9 cases discarded because of motion artifact degrading 1 of the DTI acquisitions had a useable first DTI scan. Thus, in clinical use for spinal cord assessment, 34 of 42 cases (80%) would have been adequate. The spine DTI reproducibility studies cited above relied on manual ROIs that are more labor intensive and cumbersome and do not provide tract-specific information compared with atlas-based segmentation evaluation algorithms. Frequently, the ROIs used were whole cord, which provides limited data on disease impact on white matter specifically, especially considering the variations in DTI metrics between white and gray matter as well as the potential microstructural differences in disease influence between these tissues. Cheran et al[5] previously established the value of DTI of the spine in acute trauma by using whole cord ROIs (gray and white matter included), with correlation of DTI values to clinical scores. The current study uses a segmentation algorithm that provides tract-specific metrics that can confer increased specificity with respect to clinical injury scores compared with whole-cord ROIs.

Establishing the reproducibility of DTI of the spine in a clinical environment has the potential to further its use in disease states in which it has shown promise. DTI can detect spinal cord abnormalities in multiple sclerosis,[7] neuromyelitis optica,[10] HIV myelopathy,[9] and spondylotic myelopathy[8] when conventional MR imaging appears normal and could potentially better guide treatment. In addition, DTI provides quantitative microstructural data that may show improved association with clinical presentation compared with conventional imaging, and its inclusion may be able to better predict outcomes.[28]

This investigation has some limitations. First, 12 of 42 patients had to be excluded for unusable image sets. Imaging patients with acute spine trauma is challenging because motion artifact can be significant, and complete compliance cannot be guaranteed. Ex-

tended scan sessions can introduce significant discomfort, and repeating motion-corrupted scans was not feasible. Furthermore, having 1 corrupted run (out of 2) was enough to disqualify a patient from this analysis. A second limitation is that patients were not removed from the table or room between DTI scans. To attenuate this limitation, patients had the coils removed and were repositioned on the table and relocalized, with subsequent repositioning of the axial DTI field of view. Considering significant patient injuries and limited patient mobility and function, we felt this was an adequate step for reproducibility while keeping subject safety in mind. The DTI sequences were not cardiac-gated for control of spinal cord motion because of the time cost and lack of clinical feasibility in this population; cardiac gating could double scan time in some cases in a population that often requires immediate medical care and is motion-prone.

CONCLUSIONS

This work is an initial step toward using automated parcellation of spinal cord DTI in acute traumatic cervical spinal cord injury. The established test-retest and interreader reproducibility of these measures may inform the development of future studies focused on DTI as an imaging biomarker in diagnostic and therapeutic interventions in this patient population.

REFERENCES

1. Basser PJ, Pierpaoli C. **Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI.** *J Magn Reson B* 1996;111:209–19 CrossRef Medline
2. Brandstack N, Kurki T, Tenovuo O. **Quantitative diffusion-tensor tractography of long association tracts in patients with traumatic brain injury without associated findings at routine MR imaging.** *Radiology* 2013;267:231–39 CrossRef Medline
3. Gratsias G, Kapsalaki E, Kogia S, et al. **A quantitative evaluation of damage in normal appearing white matter in patients with multiple sclerosis using diffusion tensor MR imaging at 3 T.** *Acta Neurol Belg* 2015;115:111–16 CrossRef Medline
4. Yuh EL, Cooper SR, Mukherjee P, et al. **Diffusion tensor imaging for outcome prediction in mild traumatic brain injury: a TRACK-TBI study.** *J Neurotrauma* 2014;31:1457–77 CrossRef Medline
5. Cheran S, Shanmuganathan K, Zhuo J, et al. **Correlation of MR diffusion tensor imaging parameters with ASIA motor scores in hemorrhagic and nonhemorrhagic acute spinal cord injury.** *J Neurotrauma* 2011;28:1881–92 CrossRef Medline
6. Mulcahey MJ, Samdani A, Gaughan J, et al. **Diffusion tensor imaging in pediatric spinal cord injury: preliminary examination of reliability and clinical correlation.** *Spine* 2012;37:E797– 803 CrossRef Medline
7. Naismith RT, Xu J, Klawiter EC, et al. **Spinal cord tract diffusion tensor imaging reveals disability substrate in demyelinating disease.** *Neurology* 2013;80:2201– 09 CrossRef Medline
8. Curvo-Semedo L, Lambregts DM, Maas M, et al. **Rectal cancer: assessment of complete response to preoperative combined radiation therapy with chemotherapy– conventional MR volumetry versus diffusion-weighted MR imaging.** *Radiology* 2011;260: 734– 43 CrossRef Medline
9. Mueller-Mang C, Law M, Mang T, et al. **Diffusion tensor MR imaging (DTI) metrics in the cervical spinal cord in asymptomatic HIV-positive patients.** *Neuroradiology* 2011;53:585–92 CrossRef Medline
10. Rivero RL, Oliveira EM, Bichuetti DB, et al. **Diffusion tensor imaging of the cervical spinal cord of patients with neuromyelitis optica.** *Magn Reson Imaging* 2014;32:457– 63 CrossRef Medline
11. O'Donnell LJ, Westin CF, Golby AJ. **Tract-based morphometry for white matter group analysis.** *NeuroImage* 2009;45:832– 44 CrossRef Medline
12. Reich DS, Smith SA, Jones CK, et al. **Quantitative characterization of the corticospinal tract at 3T.** *AJNR Am J Neuroradiol* 2006;27: 2168–78 Medline
13. Yushkevich PA, Zhang H, Simon TJ, et al. **Structure-specific statistical mapping of white matter tracts.** *Neuroimage* 2008;41:448– 61 CrossRef Medline
14. Fischl B. **FreeSurfer.** *Neuroimage* 2012;62:774– 81 CrossRef Medline
15. Brewer JB. **Fully-automated volumetric MRI with normative ranges: translation to clinical practice.** *Behav Neurol* 2009;21:21–28 CrossRef Medline
16. De Leener B, Lévy S, Dupont SM, et al. **SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data.** *Neuroimage* 2017;145:24– 43 CrossRef Medline
17. McCoy DB, Talbott JF, Wilson M, et al. **MRI atlas-based measurement of spinal cord injury predicts outcome in acute flaccid myelitis.** *AJNR Am J Neuroradiol* 2017;38:410–17 CrossRef Medline
18. Eippert F, Kong Y, Winkler AM, et al. **Investigating resting-state functional connectivity in the cervical spinal cord at 3T.** *Neuroimage* 2017;147:589– 601 CrossRef Medline
19. Taso M, Girard OM, Duhamel G, et al. **Tract-specific and age-related variations of the spinal cord microstructure: a multi-parametric MRI study using diffusion tensor imaging (DTI) and inhomogeneous magnetization transfer (ihMT).** *NMR Biomed* 2016;29:817–32 CrossRef Medline
20. Tustison NJ, Avants BB. **Explicit B-spline regularization in diffeomorphic image registration.** *Front Neuroinform* 2013;7:39 CrossRef Medline
21. Avants BB, Tustison NJ, Stauffer M, et al. **The Insight ToolKit image registration framework.** *Front Neuroinform* 2014;8:44 CrossRef Medline
22. Fonov VS, Le Troter A, Taso M, et al. **Framework for integrated MRI average of the spinal cord white and gray matter: the MNI-Poly-AMU template.** *Neuroimage* 2014;102 Pt 2:817–27 CrossRef Medline
23. Lévy S, Benhamou M, Naaman C, et al. **White matter atlas of the human spinal cord with estimation of partial volume effect.** *Neuroimage* 2015;119:262–71 CrossRef Medline
24. Askren MK, McAllister-Day TK, Koh N, et al. **Using make for reproducible and parallel neuroimaging workflow and quality-assurance.** *Front Neuroinform* 2016;10:2 CrossRef Medline
25. Davison AC, Hinkley DV. *Bootstrap Methods and Their Application.* Cambridge: Cambridge University Press; 1997
26. Brander A, Koskinen E, Luoto TM, et al. **Diffusion tensor imaging of the cervical spinal cord in healthy adult population: normative values and measurement reproducibility at 3T MRI.** *Acta Radiol* 2014; 55:478– 85 CrossRef Medline
27. Smith SA, Jones CK, Gifford A, et al. **Reproducibility of tract-specific magnetization transfer and diffusion tensor imaging in the cervical spinal cord at 3 Tesla.** *NMR Biomed* 2010;23:207–17 CrossRef Medline
28. Agosta F, Absinta M, Sormani MP, et al. **In vivo assessment of cervical cord damage in MS patients: a longitudinal diffusion tensor MRI study.** *Brain* 2007;130:2211–19 CrossRef Medline