

UCSF

UC San Francisco Previously Published Works

Title

Toward Precision and Reproducibility of Diffusion Tensor Imaging: A Multicenter Diffusion Phantom and Traveling Volunteer Study

Permalink

<https://escholarship.org/uc/item/1zb9q35b>

Journal

American Journal of Neuroradiology, 38(3)

ISSN

0195-6108

Authors

Palacios, EM

Martin, AJ

Boss, MA

et al.

Publication Date






2017-03-01

DOI

10.3174/ajnr.a5025

Peer reviewed

# Toward Precision and Reproducibility of Diffusion Tensor Imaging: A Multicenter Diffusion Phantom and Traveling Volunteer Study

 E.M. Palacios,  A.J. Martin,  M.A. Boss,  F. Ezekiel,  Y.S. Chang,  E.L. Yuh,  M.J. Vassar,  D.M. Schnyer,  C.L. MacDonald,  K.L. Crawford,  A. Irimia,  A.W. Toga,  P. Mukherjee, and the TRACK-TBI Investigators



## ABSTRACT

**BACKGROUND AND PURPOSE:** Precision medicine is an approach to disease diagnosis, treatment, and prevention that relies on quantitative biomarkers that minimize the variability of individual patient measurements. The aim of this study was to assess the intersite variability after harmonization of a high-angular-resolution 3T diffusion tensor imaging protocol across 13 scanners at the 11 academic medical centers participating in the Transforming Research and Clinical Knowledge in Traumatic Brain Injury multisite study.

**MATERIALS AND METHODS:** Diffusion MR imaging was acquired from a novel isotropic diffusion phantom developed at the National Institute of Standards and Technology and from the brain of a traveling volunteer on thirteen 3T MR imaging scanners representing 3 major vendors (GE Healthcare, Philips Healthcare, and Siemens). Means of the DTI parameters and their coefficients of variation across scanners were calculated for each DTI metric and white matter tract.

**RESULTS:** For the National Institute of Standards and Technology diffusion phantom, the coefficients of variation of the apparent diffusion coefficient across the 13 scanners was <3.8% for a range of diffusivities from 0.4 to  $1.1 \times 10^{-6}$  mm<sup>2</sup>/s. For the volunteer, the coefficients of variations across scanners of the 4 primary DTI metrics, each averaged over the entire white matter skeleton, were all <5%. In individual white matter tracts, large central pathways showed good reproducibility with the coefficients of variation consistently below 5%. However, smaller tracts showed more variability, with the coefficients of variation of some DTI metrics reaching 10%.

**CONCLUSIONS:** The results suggest the feasibility of standardizing DTI across 3T scanners from different MR imaging vendors in a large-scale neuroimaging research study.

**ABBREVIATIONS:** AD = axial diffusivity (units of  $10^{-3}$  mm<sup>2</sup>/s); CoV = coefficient of variation; FA = fractional anisotropy; MD = mean diffusivity (units of  $10^{-3}$  mm<sup>2</sup>/s); NIST = National Institute of Standards and Technology; PVP = polyvinylpyrrolidone; RD = radial diffusivity (units of  $10^{-3}$  mm<sup>2</sup>/s); TBI = traumatic brain injury; TRACK-TBI = Transforming Research and Clinical Knowledge in Traumatic Brain Injury

The new paradigm of precision medicine relies on quantitative biomarkers validated for their accuracy and efficacy. They must also be standardized so that precise and reproducible mea-

surements can be made despite differences in instrumentation and procedures.<sup>1</sup> Medical imaging can provide many clinically useful biomarkers; however, the precision of its metrics must first be established in large-scale multisite studies that encompass the range of scanning hardware and software used to acquire the data.

DTI studies of patients with traumatic brain injury (TBI) show alterations of white matter microstructure in many tracts, across the range of severe, moderate, and mild TBI.<sup>2-5</sup> However, TBI is a highly heterogeneous pathology in cause, severity, and clinical course. New standardized clinical and imaging biomarkers are needed to enhance outcome prediction and for triage to the most appropriate therapeutic interventions. Transforming Research and Clinical Knowledge in TBI (TRACK-TBI), a National Insti-

Received May 1, 2016; accepted after revision October 10.


From the Departments of Radiology and Biomedical Imaging (E.M.P., A.J.M., F.E., Y.S.C., E.L.Y., P.M.), Neurological Surgery and Brain and Spinal Injury Center (M.J.V.), and Bioengineering and Therapeutic Sciences (P.M.), University of California, San Francisco, San Francisco, California; National Institute of Standards and Technology (M.A.B.), Boulder, Colorado; Brain and Spinal Cord Injury Center (E.L.Y., M.J.V., P.M.), San Francisco General Hospital and Trauma Center, San Francisco, California; Department of Psychology (D.M.S.), University of Texas, Austin, Texas; Department of Neurological Surgery (C.L.M.), University of Washington, Seattle, Washington; and Mark and Mary Stevens Neuroimaging and Informatics Institute (K.L.C., A.I., A.W.T.), University of Southern California, Los Angeles, California.

This research was funded by National Institutes of Health grant U01 NS086090.

Grant funding support for investigators: A.J.M., NIH R01EB012031, NIH U01NS086090, NIH R21CA186140, NIH R01CA194533, Voyager Therapeutics, Quest Diagnostics Incorporated; E.L.Y., NIH U01 NS086090, DoD W81XWH-14-2-0176, D.M.S., DoD-Navy W911QY-14-C-0070, NIH U01NS086090, NIH R01AG046460; C.L.M., NIH R01HD083126, DoD PT108802-SCI04833-SI, NIH U01NS086090, DoD W81XWH-14-2-0176; A.I., NIH R44 NS 081792, NIH U01NS086090; A.W.T., NIH R01MH094343, NIH P41EB015922, NIH U01NS086090; and P.M., NIH R24MH106096, NIH R01HD072074, NIH U01 NS086090, DoD W81XWH-14-2-0176, DoD W81XWH-13-1-0494, GE 114-2014-GES-0001.

The contribution of the NIST is not subject to copyright in the United States.

Please address correspondence to Pratik Mukherjee, MD, PhD, Department of Radiology and Biomedical Imaging, University of California, San Francisco, China Basin, 185 Berry St, San Francisco, CA 94107; e-mail: pratik.mukherjee@ucsf.edu

 Indicates open access to non-subscribers at [www.ajnr.org](http://www.ajnr.org)

<http://dx.doi.org/10.3174/ajnr.A5025>

tutes of Health–funded study, began in October 2013 with the goal of enrolling 3000 patients with TBI at 11 enrollment sites across the United States. The objective was to create a large, high-quality data base that integrates clinical, imaging, proteomic, genomic, and outcome measures to establish more precise methods for TBI diagnosis and prognosis.<sup>6</sup> The first critical element for a multicenter imaging study is to minimize the intersite variability. Biomarkers derived from DTI metrics must be optimally robust with respect to differences in the acquisition hardware and software across medical centers.<sup>7</sup> Intersite differences can originate from variations in scanner manufacturer; hardware characteristics such as field strength, gradient strength, and speed; and the type of radiofrequency coil, software version, site-specific quality control procedures, and adherence to the research protocol.<sup>8</sup>

Diffusion phantom and human volunteer studies are needed to assess the reproducibility and variability of the imaging data. Previous studies have shown DTI variability on test-retest studies,<sup>9–20</sup> but only 3 previous articles<sup>7,21,22</sup> have studied systematic DTI variability due to intrinsic differences between different MR imaging systems in an isotropic phantom and in a human volunteer, either within a single site or across sites. However, these prior investigations were limited to relatively few sites in the case of multicenter studies, to 1.5T scanners or a combination of 1.5T and 3T scanners, to a DTI protocol with low angular resolution, and, in 1 case, to nonharmonized DTI by using different protocols at different sites.

In this study, we assessed the intersite precision of DTI metrics from a harmonized high-angular-resolution DTI protocol across thirteen 3T scanners at the 11 sites enrolling patients for the TRACK-TBI study. High-angular-resolution diffusion imaging, which requires the acquisition of  $\geq 30$  diffusion gradient directions per scan, has been proposed to improve the accuracy and reliability of DTI metrics compared with standard DTI acquisitions with fewer directions.<sup>23,24</sup> In addition to the traveling human volunteer, data were acquired from a novel isotropic diffusion phantom developed at the National Institute of Standards and Technology (NIST), which allows precise assessment of the range of diffusivities that might be encountered in healthy and pathologic brain tissue.

## MATERIALS AND METHODS

All study procedures were approved by the institutional review boards of all 11 enrollment sites of the TRACK-TBI multicenter study.

The initial diffusion phantom and human brain data were acquired from all 13 scanners at the 11 sites within a 4-month time period at the beginning of the study. The acquisition, processing, and analysis of the isotropic diffusion phantom data were performed by a PhD MR imaging physicist (A.J.M.), who is a Professor of Radiology with 25 years of experience in industry and academia. The traveling volunteer brain imaging data were processed and analyzed by a PhD scientist (E.M.P.) with 8 years of experience in diffusion tensor image processing and analysis in TBI. All of the imaging analysis was overseen by a board-certified neuro-radiologist (P.M.) with 20 years of experience with DTI research

and who is the director of the imaging core and a Principal Investigator of the multicenter study.

### **NIST Isotropic Diffusion Phantom**

Diffusion MR imaging was acquired from a prototype isotropic diffusion phantom in a 3D-printed shell developed at the NIST and from the brain of a traveling healthy volunteer (male, 49 years of age) on thirteen 3T MR imaging scanners at 11 different sites by using 8- or 12-channel head radiofrequency coils. Isotropic phantoms are used to calibrate diffusivities, so a full DTI acquisition or identical parameters with the traveling volunteer sequences is not a requirement. The main objective of the isotropic phantoms is to verify the precision of the principal axis diffusivities because all DTI metrics are computed from these diffusivities.

The diffusion phantom was scanned at  $b=0$ , 500, and 900  $\text{s}/\text{mm}^2$  with 1.1-mm in-plane resolution and 5-mm sections, by using 3 orthogonal directions and 4 signal averages. An 11-section protocol was used and the image stack was prescribed such that the sections were centered on the phantom and intersected vials within the phantom cross-sectionally.

### **Traveling Volunteer**

The traveling volunteer was scanned once at each center during the same session as the NIST phantom. All volunteer scans were acquired within a 4-month interval. Additionally, from one of the sites (site 3), 2 DTI scans were acquired on the same day from the human traveling volunteer, 30 months after the first DTI acquisition from this volunteer. The characteristics of the scanners for each site and parameters of acquisition from the traveling volunteer in each site are reported in Table 1. The DTI protocol is based on the “Enhanced DTI” standard established by the Alzheimer Disease Neuroimaging Initiative 2 study for GE Healthcare scanners ([http://adni.loni.usc.edu/wpcontent/uploads/2010/05/ADNI2\\_GE\\_22\\_E\\_DTI.pdf](http://adni.loni.usc.edu/wpcontent/uploads/2010/05/ADNI2_GE_22_E_DTI.pdf)), except that 64 diffusion directions were used at most sites instead of 60 directions to accommodate the Siemens in-product DTI sequence in specifying the number of directions. In brief, the protocol consists of a multislection spin-echo echo-planar sequence with 2.7-mm isotropic spatial resolution and a  $b$ -value of 1300  $\text{s}/\text{mm}^2$ , with 8 additional brain volumes acquired at  $b=0$   $\text{s}/\text{mm}^2$ .

All the images were inspected, and none had any major artifacts. One of the advantages of using high-angular-resolution diffusion imaging is that a corrupted image in any single direction has a minimal effect on the fitting or metrics. Thus, images from high-angular-resolution diffusion imaging produce more accurate and reliable results than lower angular resolution acquisitions.

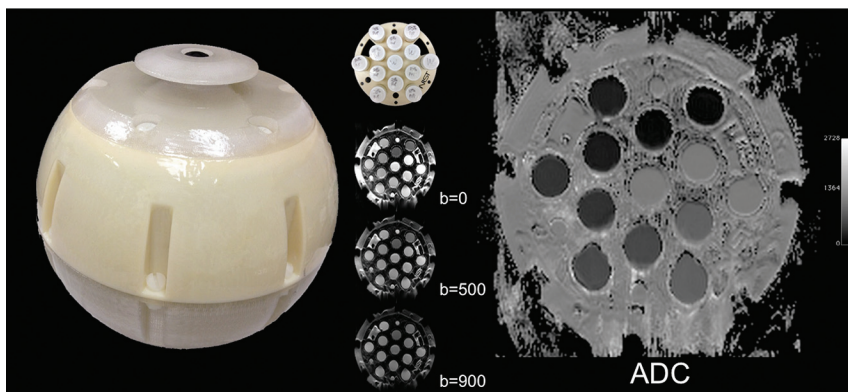
### **Preparation and Processing of the NIST Isotropic Diffusion Phantom**

We established an array of thirteen 20-mL vials containing variable amounts of the polymer polyvinylpyrrolidone (PVP) (Fig 1).<sup>25</sup> PVP containing vials, with mass fractions of 0%, 10%, 20%, 30%, 40%, and 50% were arranged in a plane, with 0% PVP (deionized water) at the center and increasing concentrations of PVP in 2 rings around the central vials. Vials were contained in a spheric water-tight vessel. The phantom was transported between

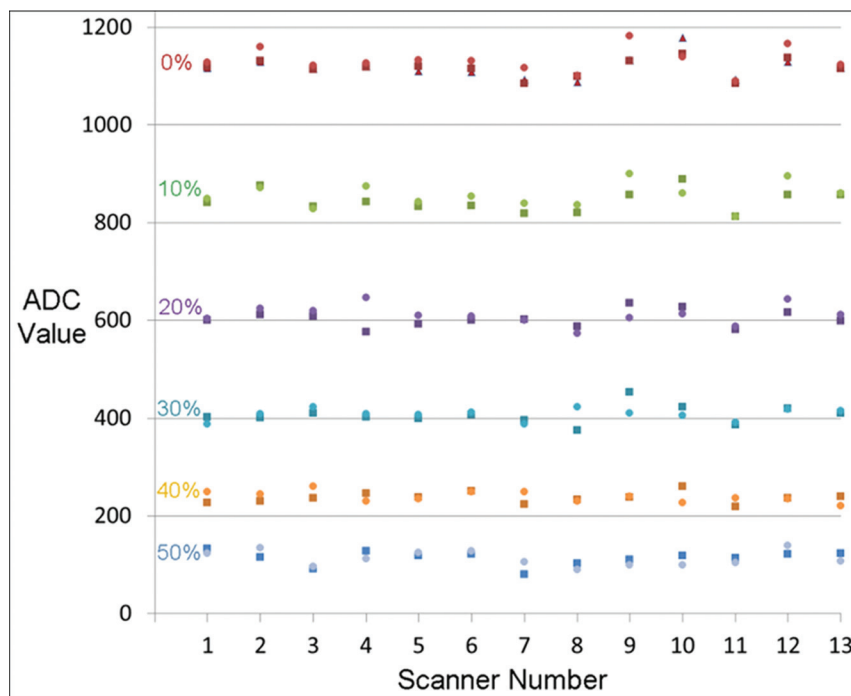
**Table 1: DTI acquisition protocol for the 11 academic medical centers<sup>a</sup>**

Site	Scanner	Vendor	Model and Field (Tesla)	Software Version	Head Coil (Channels)	Voxel Size	Flip Angle	TR (ms)	TE (ms)	Directions	Sections	Matrix	b-Value	B0
1	Trio	Siemens	Trio 3T	syngo MR B17	12	2.7	90°	9000	94	64	64	128 × 84	1300	8
2a	Skyra	Siemens	Skyra 3T	syngo MR D13	12	2.7	90°	9000	92	64	59	128 × 84	1300	8
2b	Trio	Siemens	Trio 3T	syngo MR B17	12	2.7	90°	9000	94	64	59	128 × 84	1300	8
3	MR750	GE	MR750 3T	DV24.0_RO1_1344	8	2.7	90°	9050	81	60	59	128 × 128	1300	8
4	Signa	GE	Signa 3T	HD16.0_V02_1131a	8	2.7	90°	14,000	83	60	54	128 × 128	1000	8
5	Trio	Siemens	Trio 3T	syngo MB B17	12	2.7	90°	9000	94	64	59	128 × 84	1300	8
6	Trio	Siemens	Trio 3T	syngo MR B17	12	2.7	90°	9000	94	64	59	128 × 84	1300	8
7a	MR750	GE	MR750 3T	DV24.0_RO1_1344	8	2.7	90°	9000	81	60	59	128 × 128	1300	8
7b	Trio	Siemens	Trio 3T	syngo MR B17	12	2.7	90°	9000	94	64	59	128 × 84	1300	8
8	Skyra	Siemens	Skyra 3T	syngo MR D13	20	2.7	90°	9000	92	64	59	128 × 84	1300	8
9	Signa	GE	Signa 3T	HD16.0_V02_1131a	8	2.7	90°	14,000	88	60	54	128 × 128	1300	8
10	Achieva	Phillips	Achieva 3T	U3.2.3, U3.2.3.1	8	2.7	90°	9000	71	64	59	128 × 125	1300	8
11	Ingenia	Phillips	Ingenia 3T	U5.1.2, U5.1.2.0	15	2.7	90°	9000	94	64	59	128 × 126	1300	8

<sup>a</sup> Sites 2 and 7 participated in the study with 2 different scanners each.



**FIG 1.** Prototype isotropic diffusion phantom in a 3D-printed shell from the National Institute of Standards and Technology.



**FIG 2.** ADC measurements obtained across all 13 scanners in TRACK-TBI. The measured ADC in units of  $10^{-6} \text{ mm}^2/\text{s}$  for vials containing 0%–50% PVP are plotted for each TRACK-TBI MR imaging system. The NIST phantom allows 3 measures of 0% PVP and 2 measures of each of the 10%–50% PVP concentrations. Data are shown for the central vial (triangle), inner vials (squares), and outer vials (circles) of the phantom.

sites in a dry state in a foam enclosure. On arrival at each site, the phantom was initially packed with ice and placed in a refrigerator. In parallel, a slurry of ice and water was prepared and allowed to equilibrate for 1–2 hours. Ice water was then added to the phantom along with as much additional ice as could be accommodated; the phantom was returned to the refrigerator and left overnight (8–10 hours). In the morning, additional ice was packed into the phantom and left in an insulated container until imaged (within 2 hours).

### DTI Preprocessing and Analysis for the Traveling Volunteer

DTI preprocessing and analysis were performed by using tools from the Oxford Centre for Functional MR Imaging of the Brain (FMRIB) Software Library (FSL (<http://www.fmrib.ox.ac.uk/fsl>)).

First, images were corrected for eddy current distortions and motion by using an average of the 8  $b=0 \text{ s/mm}^2$  volumes as a reference. The registered images were skull-stripped by using the FSL Brain Extraction Tool (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET>). All the resulting brain masks were visually inspected. Fractional anisotropy (FA) maps were calculated by using the FMRIB Diffusion Toolbox (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT>).<sup>26</sup> After calculation of the FA map, a voxelwise statistical analysis of the FA data was performed by using Tract-Based Spatial Statistics (TBSS; <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS>).<sup>27,28</sup> Within the TBSS pipeline, FA data were aligned into the common FMRIB58 FA template, which is in Montreal Neurological Institute 152 at-

las standard space, by using the nonlinear registration algorithm FMRIB Nonlinear Registration Tool (FNIRT; <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT>).<sup>29</sup>

Next, a mean FA image was created from the images for all the subjects' serial scans at different sites in this common space and thinned to generate a mean FA white matter skeleton that represented the center of all tracts common to the entire group of scans. The FA white matter skeleton was thresholded to  $FA > 0.2$  to exclude gray matter and voxels containing partial volume effects with gray matter. This process ensured that each dataset skeleton was in the group space while also representing the center of the subject's unique white matter bundles. The aligned FA volume was then projected onto the skeleton by filling the skeleton with FA values from the nearest relevant tract center. This was achieved for each skeleton voxel by searching perpendicular to the local skeleton structure for the maximum value in the FA image of the subject. Mean FA, mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD) values were obtained from the FA skeleton map of each scan.

Specific fasciculi were studied by using masks obtained from the FSL Jülich histologic atlas (<http://neuro.debian.net/pkgs/fsl-juelich->

histological-atlas.html) mapped on the standard Montreal Neurological Institute space and resampled to 1-mm resolution. Binary mask images from the fasciculi of interest were used to mask the individual FA (skeletonized) maps previously registered to the Montreal Neurological Institute standard space by using the nonlinear tools in the TBSS procedure. We included 15 large main fasciculi. Mean FA, MD, AD, and RD values were obtained from each subject's white matter skeleton and each of the skeletonized ROIs.

### Statistical Analysis

**NIST Isotropic Diffusion Phantom.** ADC images were reconstructed, and ROI analysis was performed. ADC values were quantified in the center section, which corresponded to the central plane of the phantom. A circular ROI measuring 1.2–1.3 cm<sup>2</sup> was placed in the center of each of the 13 vials, and the mean ADC value was determined. A mean ADC and its coefficient of variation (CoV), which is obtained from calculating the SD divided by the average, were established for each vial across all sites.

**Traveling Volunteer.** The CoV was calculated for each DTI metric in the whole white matter skeleton and for each fasciculus to summarize the amount of variation across scanners at the 11 sites. Furthermore, we also calculated a normalized CoV for each of the preselected ROIs, in which each DTI metric was divided by its corresponding mean value per site, as has been performed in prior DTI studies of TBI.<sup>30,31</sup>

## RESULTS

### NIST Isotropic Diffusion Phantom

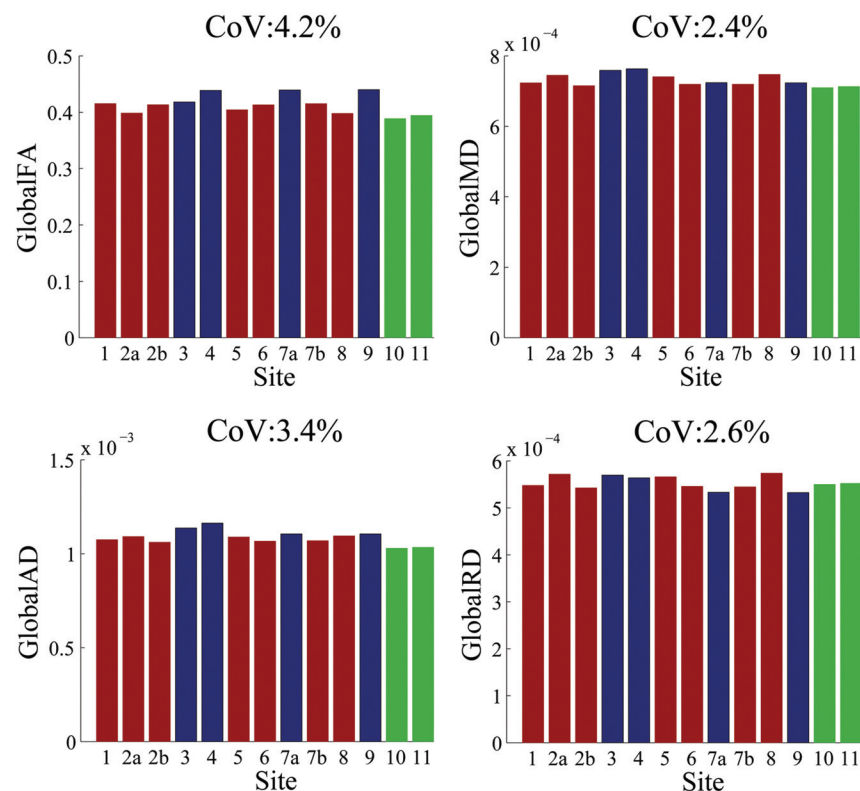
There were 3 ADC measures for the three 0% PVP vials and 2 measures each for the paired 10%, 20%, 30%, 40%, and 50% PVP vials. The differences between the inner and outer rings of vials were not statistically significant ( $P > .05$ ). Figure 2 shows the consistency of ADC measurements within and across systems. Table 2 provides the mean, SD, and CoV for each concentration of PVP. Data are shown for the full set of TRACK-TBI systems and for a subset of systems that shared the same manufacturer and model (Siemens Trio). The latter is the most common scanner type in TRACK-TBI study, with no more than 2 of any other system. The CoV is  $<3.8\%$  for concentrations of PVP of  $<30\%$  w/w. The CoV is slightly decreased if a single scanner type is considered. At 40% and 50% w/w PVP, the diffusivity decreased and the resulting CoV became larger.

### Traveling Volunteer

Whole-brain DTI measurements across sites are shown in Fig 3. The CoV for global white matter skeleton DTI parameters within and across scanners was

**Table 2: Mean ADC, SD, and CoV for each concentration of PVP**

Vial	All MR Imaging Systems (n = 13)			Trio Only (n = 5)		
	Mean ADC	SD	CoV	Mean ADC	SD	CoV
0% PVP	1123	23	2.1%	1105	15	1.4%
10% PVP	850	24	2.8%	829	13	1.6%
20% PVP	607	19	3.1%	596	14	2.3%
30% PVP	408	16	3.8%	399	16	4.0%
40% PVP	238	11	4.6%	237	13	5.5%
50% PVP	113	15	13.0%	105	16	15.2%



**FIG 3.** Global non-normalized DTI parameters and CoV across sites. Red indicates Siemens; Blue, GE Healthcare; Green, Phillips Healthcare. By order of sites, see Table 1.



**Table 3: Mean DTI parameters across all 13 scanners and their CoV**

	FA	CoV% <sup>a</sup>	CoV% <sup>b</sup>	MD	CoV% <sup>a</sup>	CoV% <sup>b</sup>	AD	CoV% <sup>a</sup>	CoV% <sup>b</sup>	RD	CoV% <sup>a</sup>	CoV% <sup>b</sup>
Global	0.41	4.16	NA	0.73	2.44	NA	1.08	3.41	NA	0.55	2.57	NA
SLFl	0.44	2.64	3.93	0.70	4.14	2.34	1.06	4.04	1.51	0.52	4.66	3.75
SLFr	0.43	3.38	2.44	0.71	2.56	0.99	1.07	3.38	1.07	0.53	2.60	1.68
SFOl	0.46	3.60	4.47	0.65	4.68	3.30	1.00	3.94	2.56	0.47	5.86	4.80
SFOr	0.45	4.72	4.22	0.67	4.94	3.62	1.03	4.82	2.55	0.50	5.90	5.19
CSTl	0.56	5.18	2.36	0.69	6.95	6.10	1.16	7.10	6.20	0.45	8.67	7.04
CSTr	0.58	5.39	3.22	0.67	6.61	5.71	1.14	6.61	5.37	0.43	8.97	7.46
CINGl	0.54	4.35	3.47	0.72	3.57	1.95	1.21	3.28	1.93	0.48	5.68	3.85
CINGr	0.50	4.42	2.69	0.74	3.77	1.66	1.20	4.43	1.79	0.51	4.39	2.65
GenuCC	0.67	5.29	3.52	0.76	4.04	4.45	1.47	3.10	1.99	0.41	10.11	9.82
BodyCC	0.66	2.60	3.32	0.82	3.05	1.97	1.56	2.66	1.97	0.44	5.41	4.18
SpleniumCC	0.74	2.56	2.20	0.70	3.11	2.26	1.44	2.76	1.68	0.32	7.30	5.81
ACRl	0.45	5.10	2.84	0.73	2.88	2.17	1.13	3.40	1.63	0.53	4.00	3.15
ACRr	0.44	5.22	3.15	0.72	4.09	3.05	1.13	4.75	2.39	0.52	4.56	3.93
PCRl	0.46	2.18	3.94	0.74	3.64	1.89	1.15	3.38	1.45	0.54	4.22	3.17
PCRr	0.46	2.86	2.30	0.74	2.82	1.64	1.13	2.83	1.99	0.54	3.49	1.79
ALICl	0.55	4.84	2.49	0.70	3.75	2.42	1.19	4.91	2.10	0.45	5.05	3.88
ALICr	0.54	6.22	3.10	0.71	3.63	2.35	1.20	4.85	2.22	0.46	5.55	4.09
PLICl	0.66	2.99	4.05	0.65	6.23	4.47	1.24	6.16	3.36	0.36	4.38	3.42
PLICr	0.64	3.78	3.54	0.67	4.38	2.78	1.25	4.44	1.84	0.38	6.95	5.67
PTRl	0.55	2.13	3.26	0.80	4.40	2.65	1.36	4.78	2.16	0.52	4.37	3.38
PTRr	0.52	3.08	1.96	0.78	3.42	1.31	1.29	3.96	1.01	0.52	3.76	1.97
ECl	0.38	7.31	4.06	0.76	2.77	1.89	1.10	3.67	1.22	0.60	3.77	2.78
ECr	0.37	7.75	4.12	0.78	2.73	1.39	1.11	4.20	1.55	0.62	3.20	1.77
UNCl	0.42	7.66	6.05	0.77	5.66	4.90	1.14	6.00	4.27	0.58	7.03	6.55
UNCr	0.43	5.57	2.94	0.75	5.97	5.00	1.14	6.60	4.84	0.56	6.29	5.43

**Note:**—r indicates right; l, left; SLF, superior longitudinal fasciculus; SFO, superior fronto-occipital fasciculus; CST, corticospinal tract; CC, corpus callosum; CING, cingulum corticospinal tract; ACR, anterior corona radiata; PCR, posterior corona radiata; PLIC, posterior limb of internal capsule; ALIC, anterior limb of internal capsule; PTR, posterior thalamic radiation; EC, external capsule; UNC, uncinate fasciculus; NA, not applicable.

<sup>a</sup> CoV across sites.

<sup>b</sup> CoV across sites normalized by the mean FA.

<5%, showing consistency comparable with the NIST diffusion phantom results. Table 3 and Fig 4 report the mean of the DTI parameters of each of the 14 white matter tracts (left and right sides) and the CoV extracted from the DTI parameters and their normalized values. For FA, MD, AD, and RD values, the CoV was  $\leq 5\%$  for most tracts both within and across scanners (Tables 3 and 4). The CoV of normalized DTI values was less than that of non-normalized DTI values.

## DISCUSSION

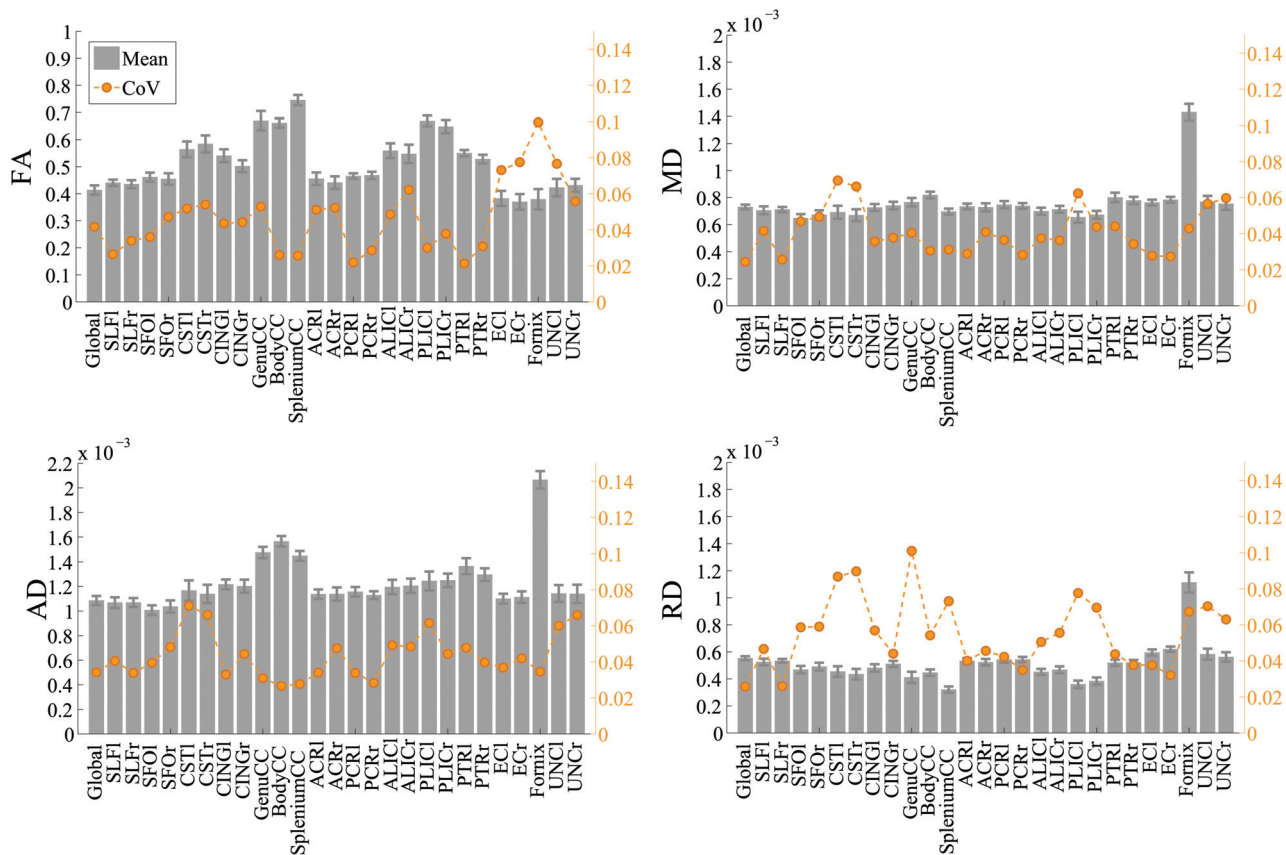
In this study of intersite variability of DTI by using both an isotropic phantom and a human volunteer, the 3T high-angular-resolution diffusion imaging protocol yielded CoVs below 5% for FA, MD, AD, and RD, averaged over the white matter skeleton of the whole brain. This finding demonstrates acceptably low intersite variation despite the 3 different scanner vendors and 6 different scanner models across the 11 centers. When examining the DTI parameters for specific tracts, we found better measurement precision in large central tracts than in small peripheral white matter, with reduced variability (lower CoV) when normalizing the values by the global DTI parameters per site. The low intersite variability of diffusion metrics was confirmed by ADC results across all 13 scanners from the NIST diffusion phantom over a range of values characteristic of the healthy human brain, as well as high and low values caused by pathology such as TBI.

The NIST has developed a series of MR imaging isotropic diffusion phantoms to enable stable measurements to mimic human tissue responses to MR imaging in a predictable and repeatable way for use in calibrating MR imaging scanners. The NIST isotro-

pic diffusion phantom in the present study contained variable amounts of the PVP, which, at different concentrations, has shown properties similar to those of healthy brain tissue.<sup>25</sup> The different concentrations in each of the vials corresponded with the expected degrees of diffusivity and showed minimal variation between scanners. The relatively poor CoV seen at very low diffusivities likely comes from the selected range of b-values, which were limited to values of 0, 500, and 900 s/mm<sup>2</sup>. Incorporation of a larger b-value is likely to substantially improve the variance in high-PVP-concentration vials, and we have recently adapted our protocol to include a b=2000 s/mm<sup>2</sup> value as well.

The results of our study, with CoVs ranging from 2% to 5%, generally concur with those of prior investigations of intersite DTI measurements in healthy subjects.<sup>9,10,18-20</sup> In a very recent study, Kamagata et al<sup>17</sup> scanned 7 volunteers at 2 sites, both with 3T Achieva scanners (Philips Healthcare). With several analytic methods, including voxelwise comparison with TBSS, atlas-based analysis, and tract-specific analysis, they reported a CoV intersite variability of <4% in the case of tract-specific analysis and a CoV of <6% by using the atlas-based analysis. Even if the parameters of acquisition, scanners, and approaches of all these studies vary, the CoV seems to be a reliable and stable measure for intersite variation. However, all these studies compared groups of different healthy subjects, adding confounding factors due to intersubject biologic variability. In our study, this large source of variation was eliminated by using the same subject across all 13 scanners.

MR imaging phantoms enable more precise measurement of diffusivity values against known standards than is possible by us-



**FIG 4.** Graphic representation of CoVs across all 13 scanners. The coefficient of variation (percentage) is in orange. Mean non-normalized DTI parameters are in gray. r indicates right; l, left; SLF, superior longitudinal fasciculus; SFO, superior fronto-occipital fasciculus; CST, corticospinal tract; CC, corpus callosum; CING, cingulum corticospinal tract; ACR, anterior corona radiata; ALIC, anterior limb of internal capsule; PCR, posterior corona radiata; PLIC, posterior limb of internal capsule; PTR, posterior thalamic radiation; EC, external capsule; UNC, uncinata fasciculus.

ing a living human brain. To our knowledge, only 3 multisite studies have previously assessed the intersite reliability of DTI measures in both a traveling volunteer and an isotropic diffusion phantom. In the most recent study, Grech-Sollars et al<sup>22</sup> assessed the reproducibility of DTI in an ice water phantom and in 9 healthy volunteers across 8 scanners (1.5T Avanto; 1.5T Symphony; Siemens; 3T Achieva) at 5 sites, all by using different DTI acquisition parameters. The phantom used in that study consisted of 5 tubes filled with distilled water and 1 tube with sucrose. Despite the considerable variability across scanners and DTI protocols, the authors reported robustness across 1.5T and 3T scanners with CoVs computed from the non-normalized DTI values ranging from 1% to 7.4%.

Zhu et al<sup>21</sup> evaluated the intra- and intersite differences in DTI measurements in 3 centers, all with 1.5T GE Healthcare Signa scanners, in a travelling human volunteer and a phantom fabricated with cells in a cylindrical polycarbonate container filled with 3 chemicals (cyclic alkanes). Their normalized bias scores results suggested that the intersite variation, though relatively small among scanners of the same vendor, significantly affects DTI measurement accuracy and precision. In Walker et al,<sup>8</sup> the investigators proposed a 2-step analysis for assessing phantom data in multisite DTI studies. These results were part of a pediatric brain development study at 5 sites with 1.5T scanners and the American College of Radiology MR imaging-accreditation phantom, con-

sisting of 10 mmol of nickel chloride and sodium chloride. The results suggested that initial outlier identification is important for accurate assessment of inter- and intrasite variability and for identification of problems with data acquisition.

Our study offers several advances over these previous multisite DTI standardization studies. The novel NIST diffusion phantom has a much larger and more granular range of diffusivities, from 0.1 to 1.1 mm<sup>2</sup>/s with 6 distinct ADC values, compared with phantoms used in prior reports. This feature allows more precise calibration of the DTI values expected in healthy and pathologic brain tissue. We restricted the study to 3T scanners to avoid systematic differences due to field strength and because the much higher signal-to-noise ratio compared with 1.5T scanners enables measurement of DTI metrics such as FA with much less uncertainty.<sup>18</sup> Our use of a harmonized high-angular-resolution diffusion imaging protocol with 60+ diffusion directions also allows more precise quantitation of DTI metrics.<sup>23,24</sup> A prior multicenter DTI standardization study of 2 traveling volunteers in five 3T scanners (3 Trio scanners and 2 Signa scanners) also used a high-angular-resolution diffusion imaging protocol with 33 diffusion directions and found good concordance of FA and MD values across sites.<sup>19</sup> Consistent with our results, these authors concluded that the comparability of DTI measures between different magnets supports the feasibility of multisite clinical trials by using DTI as an outcome measure. However, they reported

**Table 4: Intrascanner DTI measures for site 3 and their CoV**

	FA	CoV% <sup>a</sup>	CoV% <sup>b</sup>	MD	CoV% <sup>a</sup>	CoV% <sup>b</sup>	AD	CoV% <sup>a</sup>	CoV% <sup>b</sup>	RD	CoV% <sup>a</sup>	CoV% <sup>b</sup>
Global	0.41	0.73	NA	0.76	0.16	NA	1.13	0.19	NA	0.57	0.42	NA
SLFl	0.43	0.52	0.53	0.74	1.30	1.17	1.12	1.12	1.26	0.55	1.50	1.10
SLFr	0.43	3.08	3.79	0.73	1.12	1.19	1.10	0.27	0.16	0.55	2.04	2.38
SFOl	0.44	4.03	3.33	0.70	4.31	4.16	1.08	2.75	2.82	0.51	6.08	5.65
SFOr	0.45	1.94	1.43	0.73	0.19	0.13	1.13	0.85	0.66	0.53	1.28	0.94
CSTl	0.59	3.81	3.19	0.68	1.16	1.03	1.20	1.46	1.44	0.42	4.50	4.08
CSTr	0.61	3.62	2.94	0.65	2.18	2.02	1.16	0.57	0.59	0.40	4.65	4.23
CINGl	0.52	1.75	1.23	0.77	1.34	1.27	1.27	1.34	1.15	0.52	2.13	1.92
CINGr	0.48	2.65	2.23	0.78	1.40	1.26	1.24	1.38	1.37	0.54	2.49	2.16
GenuCC	0.67	0.94	0.21	0.80	0.60	0.45	1.55	0.40	0.26	0.42	1.80	1.39
BodyCC	0.54	0.46	0.29	0.87	1.79	1.64	1.62	1.58	1.69	0.49	2.18	1.76
SpleniumCC	0.75	0.97	0.27	0.70	1.86	1.72	1.47	1.07	1.20	0.31	3.73	3.31
ACRl	0.46	0.97	0.80	0.77	0.05	0.11	1.20	0.42	0.26	0.55	0.43	0.40
ACRr	0.46	0.57	0.97	0.77	0.27	0.21	1.21	0.06	0.16	0.55	0.50	0.57
PCRl	0.47	0.22	0.55	0.75	0.72	0.79	1.17	0.83	0.97	0.54	0.63	0.62
PCRr	0.46	1.02	0.64	0.74	1.13	1.11	1.14	1.34	1.15	0.54	1.29	1.22
ALICl	0.56	3.00	2.44	0.75	1.81	1.71	1.29	1.51	1.32	0.48	3.60	3.31
ALICr	0.55	2.87	2.28	0.77	0.92	1.07	1.33	2.73	2.66	0.50	1.49	1.09
PLICl	0.66	1.29	0.64	0.72	2.94	2.82	1.36	2.20	2.36	0.40	4.22	3.82
PLICr	0.66	0.82	0.98	0.71	0.70	0.57	1.35	0.18	0.35	0.40	1.74	1.56
PTRl	0.55	1.36	0.74	0.83	1.71	1.68	1.42	1.01	1.20	0.53	2.67	2.41
PTRr	0.53	0.67	1.13	0.80	1.46	1.31	1.33	1.71	1.80	0.53	1.25	0.84
ECl	0.41	3.16	2.45	0.79	1.12	0.96	1.17	1.14	0.99	0.60	2.23	1.81
ECr	0.40	1.70	1.17	0.81	0.48	0.34	1.18	0.37	0.18	0.63	1.05	0.69
Fornix	0.39	5.87	5.17	0.13	1.13	1.00	2.01	1.05	0.89	0.10	3.19	2.78
UNCl	0.44	6.63	5.96	0.84	6.53	6.38	1.29	4.17	4.22	0.62	9.17	8.75
UNCr	0.44	4.17	3.53	0.84	3.47	3.36	1.28	2.00	2.15	0.61	5.13	4.76

**Note:**—r indicates right; l, left; SLF, superior longitudinal fasciculus; SFO, superior fronto-occipital fasciculus; CST, corticospinal tract; CC, corpus callosum; CING, cingulum corticospinal tract; ACR, anterior corona radiata; PCR, posterior corona radiata; PLIC, posterior limb of internal capsule; ALIC, anterior limb of internal capsule; PTR, posterior thalamic radiation; EC, external capsule; UNC, uncinate fasciculus; NA, not applicable.

<sup>a</sup> CoV across sites.

<sup>b</sup> CoV across sites normalized by the mean FA.

greater intersite variability in DTI metrics than we found, with an average white matter CoV of 6.8% for FA and 4.1% for MD.

Despite having many more scanners (13 versus 5) and many more scanner models (6 versus 2) than Fox et al,<sup>19</sup> we found a global white matter CoV of 4.2% for FA and 2.4% for MD. This finding is likely due to the hand-drawn ROI measurements of Fox et al, which introduce intrarater and interrater variability, compared with the fully automated DTI measurements of the present study using TBSS, which has the added advantage of evaluating white matter throughout the entire brain instead of only a few selected regions. In addition, when estimating the FA CoV of a specific tract (ie, the anterior corona radiata) from a prior DTI study of mild TBI,<sup>5</sup> we found a 5% CoV for the anterior corona radiata, known to be commonly injured after mild TBI,<sup>2</sup> across the pooled group of patients with complicated mild TBIs and controls, with a statistically significant mean difference of 3% between patient and control groups at  $P < .05$ . Thus, the interscanner variation of FA observed in this tract in the current study across 13 different 3T scanners of various types is similar to the intersubject variability of FA on a single scanner.<sup>5</sup>

A limitation of our study is that only a single volunteer could be assessed at all sites, given the travel expenses involved in flying multiple volunteers to be scanned at 11 different medical centers located throughout the United States. Also, intrasite scan-rescan reliability was obtained at only 1 site, though with the same human subject scanned across all of the sites. Our results indicate, as expected, that the variability of DTI metrics is less within the same scanner (Table 4) than across different scanners (Table 3). There

were small variations in the DTI protocol performed at each site due to hardware or software idiosyncrasies of each vendor platform; however, low intersite variability was achieved despite these minor protocol differences. This study was designed to measure the magnitude of variability, but the many possible sources of variation due to scanner hardware and software factors, as well as dozens of pulse sequence parameters, are beyond the scope of this study. This subject remains an open area in the DTI literature, which requires future investigation. The application of newly developed methods to retrospectively compensate for intersite differences in diffusion MR imaging acquisition by using correction factors may further reduce the variability of DTI metrics in multisite studies and improve reproducibility,<sup>20</sup> including for longitudinal data.<sup>32</sup> This application is particularly important for the investigation of smaller peripheral white matter tracts in which intersite variation is higher than in large central white matter regions.

## CONCLUSIONS

By using a novel isotropic diffusion phantom and a traveling volunteer, we have demonstrated the feasibility of standardizing DTI on 3T scanners from all 3 major MR imaging vendors across the 11 patient enrollment sites of the TRACK-TBI study. Our findings support the viability and reproducibility of large-scale multisite DTI projects to validate diffusion-based biomarkers for TBI and many other neurologic and psychiatric disorders in the coming era of precision medicine.



## ACKNOWLEDGMENTS

We gratefully acknowledge the MR imaging technologists and study coordinators from all sites of the TRACK-TBI study. We also acknowledge the pioneering efforts of the Quantitative Imaging Biomarkers Alliance of the Radiological Society of North America, especially their contributions to previous designs of an isotropic diffusion phantom that led to the development of the NIST diffusion phantom.

Data used in preparation of this article were obtained from the Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) project. As such, there are investigators within TRACK-TBI that contributed to the design and implementation and/or provided data but did not participate in analysis or writing of this report. A complete listing of TRACK-TBI investigators can be found at: <http://tracktbi.ucsf.edu/organization-and-leadership>.

Disclosures: Alastair Martin—RELATED: Grant: National Institutes of Health, Comments: NIH U01NS086090 (Principal Investigator, Dr. Geoffrey T. Manley) "Transforming Research and Clinical Knowledge in Traumatic Brain Injury"; Support for Travel to Meetings for the Study or Other Purposes: National Institutes of Health, Comments: NIH U01NS086090 (Principal Investigator, Dr. Geoffrey T. Manley) "Transforming Research and Clinical Knowledge in Traumatic Brain Injury."\* Michael A. Boss—UNRELATED: Patents (Planned, Pending, or Issued): National Institute of Standards and Technology, Comments: The phantom shell design was submitted for a patent, "MRI Phantom, Method for Making Same and Acquiring an MRI," application No. 20150323639, The author has not received any payment related to this patent application.\* Frank Ezekiel—RELATED: Grant: National Institutes of Health.\* David M. Schnyer—RELATED: Grant: National Institute of Neurological Disorders and Stroke\*; UNRELATED: Grants/Grants Pending: National Institute of Mental Health.\* Christine L. MacDonald—RELATED: Grant: National Institutes of Health, Comments: 1U01NS086090-01 "Transforming Research and Clinical Knowledge in Traumatic Brain Injury";\* UNRELATED: Consultancy: Sinapis Pharma, Comments: I received payment for consultancy to Sinapis Pharma unrelated to this article; Employment: University of Washington, Comments: I am employed at the University of Washington; Grants/Grants Pending: R01HD083126, Comments: I received an ROI grant from National Institutes of Health/National Institute of Neurological Disorders and Stroke for research unrelated to this article; Payment for Lectures including Service on Speakers Bureaus: University of Tennessee Health Sciences, Neurology Grand Grounds University of Hawaii, Pathology Chronic Traumatic Encephalopathy (CTE) Conference, Comments: I was an invited speaker to university and symposium lectures; I received payment for invited speaking engagements during the prior 36-month period unrelated to this article. Karen L. Crawford—RELATED: Grant: National Institutes of Health\*; UNRELATED: Employment: University of Southern California; Grants/Grants Pending: National Institutes of Health, Alzheimer's Association, Michael J. Fox Foundation.\* Arthur W. Toga—RELATED: Grant: National Institutes of Health\*; UNRELATED: Employment: University of Southern California; Grants/Grants Pending: National Institutes of Health, Alzheimer's Association, Michael J. Fox Foundation.\* Pratik Mukherjee—UNRELATED: Board Membership: GE-National Football League Medical Advisory Board, Comments: No payment has been received by me or my institution in the prior 36 months for my service on this board; Payment for Lectures including Service on Speakers Bureaus: GE Healthcare, Comments: Webinar. \*Money paid to the institution.

## REFERENCES

- Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372:793–95 CrossRef Medline
- Niogi SN, Mukherjee P, Ghajar J, et al. Extent of microstructural white matter injury in postconcussive syndrome correlates with impaired cognitive reaction time: a 3T diffusion tensor imaging study of mild traumatic brain injury. *AJNR Am J Neuroradiol* 2008;29:967–73 CrossRef Medline
- Newcombe V, Chatfield D, Outtrim J, et al. Mapping traumatic axonal injury using diffusion tensor imaging: correlations with functional outcome. *PLoS One* 2011;6:e19214 CrossRef Medline
- Palacios EM, Sala-Llonch R, Junque C, et al. White matter integrity related to functional working memory networks in traumatic brain injury. *Neurology* 2012;78:852–60 CrossRef Medline
- Yuh EL, Cooper SR, Mukherjee P, et al; TRACK-TBI INVESTIGATORS. Diffusion tensor imaging for outcome prediction in mild traumatic brain injury: a TRACK-TBI study. *J Neurotrauma* 2014;31:1457–77 CrossRef Medline
- Yue JK, Vassar MJ, Lingsma HF, et al; TRACK-TBI Investigators. Transforming research and clinical knowledge in traumatic brain injury pilot: multicenter implementation of the common data elements for traumatic brain injury. *J Neurotrauma* 2013;30:1831–44 CrossRef Medline
- Hsu W, Markey MK, Wang MD. Biomedical imaging informatics in the era of precision medicine: progress, challenges, and opportunities. *J Am Med Inform Assoc* 2013;20:1010–13 CrossRef Medline
- Walker L, Curry M, Nayak A, et al; Brain Development Cooperative Group. A framework for the analysis of phantom data in multicenter diffusion tensor imaging studies. *Hum Brain Mapp* 2013;34:2439–54 CrossRef Medline
- Pagani E, Hirsch JG, Pouwels PJ, et al. Intercenter differences in diffusion tensor MRI acquisition. *J Magn Reson Imaging* 2010;31:1458–68 CrossRef Medline
- Vollmar C, O'Muircheartaigh J, Barker GJ, et al. Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0T scanners. *Neuroimage* 2010;51:1384–94 CrossRef Medline
- Jansen JF, Kooi ME, Kessels AG, et al. Reproducibility of quantitative cerebral T2 relaxometry, diffusion tensor imaging, and 1H magnetic resonance spectroscopy at 3.0 Tesla. *Invest Radiol* 2007;42:327–37 CrossRef Medline
- Bisdas S, Bohning DE, Besenski N, et al. Reproducibility, interrater agreement, and age-related changes of fractional anisotropy measures at 3T in healthy subjects: effect of the applied b-value. *AJNR Am J Neuroradiol* 2008;29:1128–33 CrossRef Medline
- Takao H, Hayashi N, Kabasawa H, et al. Effect of scanner in longitudinal diffusion tensor imaging studies. *Hum Brain Mapp* 2012;33:466–77 CrossRef Medline
- Magnotta VA, Matsui JT, Liu D, et al. Multicenter reliability of diffusion tensor imaging. *Brain Connect* 2012;2:345–55 CrossRef Medline
- Huang L, Wang X, Baliki MN, et al. Reproducibility of structural, resting-state BOLD and DTI data between identical scanners. *PLoS One* 2012;7:e47684 CrossRef Medline
- Veenith TV, Carter E, Grossac J, et al. Inter subject variability and reproducibility of diffusion tensor imaging within and between different imaging sessions. *PLoS One* 2013;8:e65941 CrossRef Medline
- Kamagata K, Shimoji K, Hori M, et al. Intersite reliability of diffusion tensor imaging on two 3T scanners. *Magn Reson Med Sci* 2015;14:227–33 CrossRef Medline
- Polders DL, Leemans A, Hendrikse J, et al. Signal to noise ratio and uncertainty in diffusion tensor imaging at 1.5, 3.0, and 7.0 Tesla. *J Magn Reson Imaging* 2011;33:1456–63 CrossRef Medline
- Fox RJ, Sakaie K, Lee JC, et al. A validation study of multicenter diffusion tensor imaging: reliability of fractional anisotropy and diffusivity values. *AJNR Am J Neuroradiol* 2012;33:695–700 CrossRef Medline
- Mirzaalian H, de Pierrefeu A, Savadjiev P, et al. Harmonizing diffusion MRI data across multiple sites and scanners. *Med Image Comput Assist Interv* 2015;9349:12–19 CrossRef Medline
- Zhu T, Hu R, Qiu X, et al. Quantification of accuracy and precision of multi-center DTI measurements: a diffusion phantom and human brain study. *Neuroimage* 2011;56:1398–411 CrossRef Medline
- Grech-Sollars M, Hales PW, Miyazaki K, et al. Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *NMR Biomed* 2015;28:468–85 CrossRef Medline
- Mukherjee P, Berman JI, Chung SW, et al. Diffusion tensor MR imaging and fiber tractography: theoretic underpinnings. *AJNR Am J Neuroradiol* 2008;29:632–41 CrossRef Medline
- Mukherjee P, Chung SW, Berman JI, et al. Diffusion tensor MR imaging and fiber tractography: technical considerations. *AJNR Am J Neuroradiol* 2008;29:843–52 CrossRef Medline

25. Pierpaoli C, Sarlls J, Nevo U, et al. **Polyvinylpyrrolidone (PVP) water solutions as isotropic phantoms for diffusion MRI studies.** In: *Proceedings of the Annual Meeting and Exhibition of the International Society for Magnetic Resonance in Medicine*, Honolulu, Hawaii. April 18–24, 2009:1414
26. Behrens TE, Johansen-Berg H, Woolrich MW, et al. **Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging.** *Nat Neurosci* 2003;6:750–57 CrossRef Medline
27. Smith SM, Jenkinson M, Johansen-Berg H, et al. **Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data.** *Neuroimage* 2006;31:1487–505 CrossRef Medline
28. Smith SM, Jenkinson M, Woolrich MW, et al. **Advances in functional and structural MR image analysis and implementation as FSL.** *Neuroimage* 2004;23(suppl 1):S208–19 CrossRef Medline
29. Andersson JL, Jenkinson M, Smith S. **Non-linear registration, aka spatial normalisation.** <http://www.fmrib.ox.ac.uk/analysis/techrep/>. FMRIB technical report TR07JA2. Published June 28, 2007
30. Galanaud D, Perlberg V, Gupta R, et al; Neuro Imaging for Coma Emergence and Recovery Consortium. **Assessment of white matter injury and outcome in severe brain trauma: a prospective multicenter cohort.** *Anesthesiology* 2012;117:1300–10 CrossRef Medline
31. Dinkel J, Drier A, Khalilzadeh O, et al; for NICER (Neuro Imaging for Coma Emergence and Recovery) Consortium. **Long-term white matter changes after severe traumatic brain injury: a 5-year prospective cohort.** *AJNR Am J Neuroradiol* 2014;35:23–29 CrossRef Medline
32. Venkatraman VK, Gonzalez CE, Landman B, et al. **Region of interest correction factors improve reliability of diffusion imaging measures within and across scanners and field strengths.** *Neuroimage* 2015;119:406–16 CrossRef Medline