

# Title

Data Descriptor

## In vivo human whole-brain Connectom diffusion MRI dataset at 760 $\mu\text{m}$ isotropic resolution

### Authors

Fuyixue Wang<sup>1,2\*</sup>, Zijong Dong<sup>1,3</sup>, Qiyuan Tian<sup>1</sup>, Congyu Liao<sup>1</sup>, Qiuyun Fan<sup>1</sup>, W. Scott Hoge<sup>4</sup>, Boris Keil<sup>5</sup>, Jonathan R. Polimeni<sup>1,2</sup>, Lawrence L. Wald<sup>1,2</sup>, Susie Y. Huang<sup>1,2</sup> & Kawin Setsompop<sup>1,2</sup>

### Affiliations

1. Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States
2. Harvard-MIT Health Sciences and Technology, MIT, Cambridge, MA, United States
3. Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, United States
4. Department of Radiology, Brigham and Women's Hospital, Boston, MA, United States
5. Department of Life Science Engineering, Institute of Medical Physics and Radiation Protection, Giessen, Germany

\*Corresponding author(s): Fuyixue Wang (fuyixue@mit.edu)

### Abstract

We present a whole-brain in vivo diffusion MRI (dMRI) dataset acquired at 760  $\mu\text{m}$  isotropic resolution and sampled at 1260 q-space points across 9 two-hour sessions on a single healthy subject. The creation of this benchmark dataset is possible through the synergistic use of advanced acquisition hardware and software including the high-gradient-strength Connectom scanner, a custom-built 64-channel phased-array coil, a personalized motion-robust head stabilizer, a recently developed SNR-efficient dMRI acquisition method, and parallel imaging reconstruction with advanced ghost reduction algorithm. With its unprecedented resolution, SNR and image quality, we envision that this dataset will have a broad range of investigational, educational, and clinical applications that will advance the understanding of human brain structures and connectivity. This comprehensive dataset can also be used as a test bed for new modeling, sub-sampling strategies, denoising and processing algorithms, potentially providing a common testing platform for further development of in vivo high resolution dMRI techniques. Whole brain anatomical T<sub>1</sub>-weighted and T<sub>2</sub>-weighted images at submillimeter scale along with field maps are also made available.

### Background & Summary

Diffusion magnetic resonance imaging (dMRI)<sup>1</sup> has proven to be an invaluable tool in the study of brain structural connectivity and complex tissue microarchitecture. It has also offered the ability to investigate a variety of neurological, neurodegenerative and psychiatric disorders<sup>2-10</sup>. In recent years, a number of dMRI data repositories<sup>11-18</sup> have been published to enhance the understanding of human brain, including large-scale population study projects launched across the world, such as the Human Connectome Project (HCP)<sup>19-22</sup> and the UK Biobank project<sup>23</sup>. However, most of these studies acquired in vivo dMRI datasets at a relatively low spatial resolution of 1.25 - 3 mm, due to technical difficulties in achieving higher spatial resolution that include: i) the inherent low SNR of

dMRI acquired with smaller voxels; ii) the distortion/blurring artifacts from longer EPI readouts; and iii) the inevitable subject motion during long scan time required for obtaining multiple diffusion directions. Submillimeter isotropic resolution dMRI datasets have therefore been mostly confined to ex-vivo studies<sup>24-26</sup> where there are no scan time and motion constraints. While ex vivo dMRI plays an important role in studying brain structures with its higher achievable resolution, the alteration of tissue properties due to fixation may fundamentally change the dMRI measurement and affect its ability to reflect accurate values in the living human brain. In vivo dMRI serves as a non-invasive probe of tissue structure in different brain pathologies, and can also be combined with functional and behavioral data for further investigation of structure-function relationships. The availability of submillimeter resolution in vivo dMRI datasets would enable more detailed examination of the brain's fine-scale structures, such as probing short-association fibers at the gray-white matter boundary<sup>27</sup> or studying the cytoarchitecture of thin cortical layers (< 1-4 mm in thickness)<sup>26,28</sup>. Submillimeter resolution in vivo dMRI would also facilitate the development and optimization of new sampling, modeling and processing methods specifically targeted at recovering information about tissue microstructure from such high resolution in vivo dMRI datasets.

With the advent of advanced hardware<sup>29-31</sup>, acquisition and reconstruction strategies<sup>32-37</sup>, submillimeter isotropic resolution dMRI is becoming increasingly feasible<sup>38-42</sup>, presenting new opportunities for studying brain structure at the mesoscopic scale in vivo. In this work, we aim to create the first publicly available in vivo whole brain dMRI reference dataset acquired at submillimeter isotropic resolution of 760  $\mu\text{m}$ , i.e. at 7.7 times smaller voxel size than conventional dMRI acquisition at  $\sim 1.5$  mm resolution. To enable high-SNR at this extreme resolution, state-of-the-art hardware and acquisition techniques were used along with a carefully designed diffusion protocol targeting at a single-subject scan across 9 two-hour sessions. The data were acquired on the MGH-USC 3T Connectom scanner equipped with high-strength gradient ( $G_{\text{max}} = 300$  mT/m)<sup>29,30,43</sup> and a custom-built 64-channel phased-array coil<sup>31</sup>, using the recently developed SNR-efficient simultaneous multi-slab imaging technique termed gSlider-SMS<sup>40,41</sup>. To mitigate potential image artifacts and loss in resolution, a custom-made form-based headcase (<https://caseforge.co>) that precisely fit the shape of subject's head and the coil was also used to ensure minimal motion contamination, consistency of subject positioning and same field shim settings across sessions. Advanced parallel imaging reconstruction with ghost-reduction algorithm (Dual-Polarity GRAPPA<sup>44,45</sup>) and reversed phase-encoding acquisition were also performed to mitigate ghosting artifacts and image distortion. We acquired a total of 1260 q-space points including 420 directions at  $b=1000$  s/mm<sup>2</sup> and 840 directions at  $b=2500$  s/mm<sup>2</sup> across the 9 sessions to ensure high angular resolution. An optimized preprocessing pipeline was employed afterwards to correct for intensity drifts, distortions and motion while preserving as much high-resolution information as possible.

We envision that this reference dMRI dataset will have a broad range of investigational, educational, and clinical applications that will advance the understanding of human brain structure. In particular, the data should be useful in the exploration of mesoscale structures revealed through submillimeter in vivo dMRI, and in the evaluation of structural brain connectivity that should improve with reduced intravoxel fiber complexity using smaller voxels and with higher intravoxel fiber crossing resolution using high angular resolution<sup>46</sup>. The data may also serve as a test bed for new modeling, data sub-sampling, denoising and processing algorithms, potentially providing a common testing platform for further technical development of in vivo high resolution dMRI.

## Methods

### Participant

The participant was a young healthy Caucasian man (born in 1989). He was fully instructed about the scans and gave written informed consent for participation in the study. The experiments were performed with approval from the institutional review board of Partners Healthcare.

## Data acquisition

All data were acquired in 9 two-hour scan sessions on the MGH-USC 3T Connectom scanner equipped with a maximum gradient strength of 300 mT/m and a maximum slew rate of 200 T/m/s<sup>29,30,43</sup>, using a custom-built 64-channel phased-array coil<sup>31</sup>. A custom-designed personalized motion-robust stabilizer (<https://caseforge.co>), that was precisely fit to the shape of the subject's head and the inside of the coil, was used to ensure the same subject positioning across sessions and minimal motion contamination within sessions (Fig. 1a). Field shim settings were kept identical across sessions to minimize differences in image distortion, with the goal of minimizing reliance on image post-processing that could introduce blurring.

To obtain high spatial resolution dMRI data, the gSlider-SMS<sup>40</sup> sequence (Fig. 1a) was used in our study. gSlider-SMS is a recently developed self-navigated simultaneous multi-slab acquisition that combines blipped-CAPI<sup>47</sup> controlled aliasing with a slab-selective radio frequency encoding, which has been demonstrated to provide high SNR-efficiency for high resolution dMRI. The gSlider-SMS protocol to acquire these data was constructed to achieve an isotropic submillimeter resolution with whole brain coverage, high SNR, minimal reconstruction errors, reduced distortion and blurring as well as tolerable scan time per volume. Motivated by this goal, the decision of the final imaging parameters was made by evaluation of data quality of pilot scans and conclusions from prior studies. The 0.76-mm isotropic resolution was chosen as it provides a good trade-off in term of resolution and scan time to achieve reasonable SNR. A gSlider encoding factor of 5 and a multi-band acceleration factor of 2 were used to achieve high SNR efficiency by acquiring 10 imaging slices simultaneously per EPI-readout across two thin RF-encoded slabs. In-plane acceleration of 3 was selected to mitigate EPI distortion and blurring by reducing the effective echo spacing from 1.02 ms to 0.34 ms, while ensuring good reconstruction performance with low noise amplification when the reconstruction was performed with Dual polarity GRAPPA to also mitigate ghosting artifacts. The detailed imaging parameters are listed in Figure 2a.

To achieve good SNR and high angular resolution, a total of 2808 volumes of dMRI data were acquired across the 9 sessions, consisting of 144  $b = 0$  s/mm<sup>2</sup> images, 420  $b = 1000$  s/mm<sup>2</sup> diffusion-weighted images (DWIs) and 840  $b = 2500$  s/mm<sup>2</sup> DWIs, along with their paired reversed phase-encoding volumes. The two shells at  $b = 1000$  s/mm<sup>2</sup> and  $2500$  s/mm<sup>2</sup> are shown in Figure 2b, where the 420 directions at  $b = 1000$  s/mm<sup>2</sup> represented a subset of the 840 directions at  $b = 2500$  s/mm<sup>2</sup>. The diffusion vectors were designed using the algorithm described in Emmanuel Caruyer *et al.*<sup>48</sup> (publicly available tool at <http://www.emmanuelcaruyer.com/q-space-sampling.php>), which not only provides uniform angular coverage for each shell and for overall multishell sampling, but is also able to provide uniform incremental angular distribution. The TEs of  $b = 1000$  s/mm<sup>2</sup> and  $b = 2500$  s/mm<sup>2</sup> were matched. In each session, 312 dMRI volumes were acquired, including 280 DWIs (140 unique diffusion directions with reversed phase-encoding acquisition) and 32  $b = 0$  s/mm<sup>2</sup> images (interleaved with every 10 diffusion images). These data were further divided and acquired in four acquisition blocks, so the calibration data could be updated every ~25 mins and the paired reversed phase-encoding volumes could be interleaved.

In each session, a low-resolution field map and a structural scan ( $T_1$  or  $T_2$ -weighted image) were also acquired as time permitted. The low-resolution field maps were acquired using a two-TE gradient echo field mapping sequence: TR = 351 ms; TEs = 4.14, 6.6 ms; FOV = 220 × 220 mm<sup>2</sup>; voxel size = 2 × 2 mm<sup>2</sup>; number of slices = 38; slice thickness = 2.5 mm; slice gap = 1.3 mm; flip angle = 45°; bandwidth = 500 Hz/Pixel; total acquisition time = 79 s. The  $T_1$ -weighted images were

acquired with the Multi-echo Magnetization-Prepared Rapid Acquisition Gradient Echo (MEMPRAGE)<sup>49</sup> sequence at both 1-mm isotropic and 0.7-mm isotropic resolutions. The 1-mm MEMPRAGE protocol was acquired using the following parameters: TR = 2530 ms; TI = 1100 ms; number of echoes = 4; TEs = 1.61, 3.47, 5.34, 7.19 ms; flip angle = 7°; FOV = 256 × 256 × 208 mm<sup>3</sup>; bandwidth = 650 Hz/Pixel; GRAPPA factor along phase encoding (PE) direction = 2; total acquisition time = 6 minutes. The 0.7-mm MEMPRAGE protocol: TR = 2530 ms; TI = 1110 ms; number of echoes = 4; TE = 1.29, 3.08, 4.83, 6.58 ms; flip angle = 7°; FOV = 256 × 256 × 146 mm<sup>3</sup>; bandwidth = 760 Hz/Pixel; GRAPPA factor along PE = 2; total acquisition time = 8.38 minutes. T<sub>2</sub>-weighted data were acquired with a T<sub>2</sub>-SPACE<sup>50</sup> sequence at 0.7-mm isotropic resolution with the following parameters: TR/TE = 3200/563 ms; FOV = 224 × 224 × 180 mm<sup>3</sup>; bandwidth = 745 Hz/Pixel; GRAPPA factor along PE = 2; total acquisition time = 8.4 minutes. All T<sub>1</sub>- and T<sub>2</sub>-weighted data were acquired with image readout in the head-foot (HF) direction, phase-encoding (PE) in the anterior-posterior (AP) direction, and partition in the left-right (LR) direction. One 1-mm isotropic resolution MEMPRAGE, three 0.7-mm isotropic resolution MEMPRAGE and three 0.7-mm isotropic resolution T<sub>2</sub>-SPACE images were acquired in session 3-9, respectively.

### Data reconstruction

The image reconstruction for dMRI was performed on raw k-space data in MATLAB (Mathworks, Natick, MA, USA) (Fig. 1c). First, dual-polarity-GRAPPA<sup>44,45</sup> (DPG) was used for parallel imaging reconstruction of both in-plane and multi-band accelerations. DPG has been validated to reduce Nyquist ghost artifacts of EPI effectively by correcting for higher-order phase errors<sup>44,45</sup>. Subsequently, POCS<sup>51</sup> was used for partial Fourier reconstruction to recover high spatial information from partially acquired k-space data. After DPG and POCS, diffusion background phase-corruptions were removed using the real-valued diffusion algorithm to provide real-value data and avoid magnitude bias in subsequent post-processing steps<sup>52</sup>. Finally, gSlider's slab-to-slice reconstruction<sup>40,41</sup> was performed to reconstruct thick RF-encoded slabs (3.8 mm) into thin slices (0.76 mm), where T<sub>1</sub> recovery was incorporated into the reconstruction model to mitigate slab-boundary artifacts<sup>53</sup>.

The image reconstruction of the T<sub>1</sub>- and T<sub>2</sub>-weighted images and field maps was performed by the vendor's (Siemens Healthcare, Erlangen, Germany) online reconstruction (including vendor's implementation of parallel imaging reconstruction, fast Fourier transformation, coil combination, root mean square combination of different echo images and calculation of field maps) and was finally stored in Digital Imaging and Communications in Medicine (DICOM) format.

### Data preprocessing

Preprocessing corrections across the reconstructed image series were performed to correct for signal drift, susceptibility, eddy currents, and gradient nonlinearity induced distortions, and subject motion. An optimized pipeline was used where the warp fields for susceptibility, eddy-current distortion, gradient nonlinearity and motion were estimated, combined, and applied to each image volume in a one-step resampling process to ensure minimal smoothing effects due to interpolation<sup>54,55</sup>. The FMRIB Software Library (FSL)<sup>56-58</sup> was used to perform these estimations and corrections. The preprocessing pipeline is shown in Fig. 1d and described in detail below:

i) First, intensity normalization and signal drift<sup>59,60</sup> correction were performed. The non-DW images ( $b = 0 \text{ s/mm}^2$ ) were used to normalize the signal intensity of the data acquired within and across the different acquisition blocks (~25 minutes each). The signal drift within each 25-minute acquisition block was corrected using scaling factors obtained from fitting a second-order polynomial on the mean signals of the non-DW images as a function of the scanned non-DW and DW images. Specifically, there were 8 non-DW images interleaved throughout each 25-minute acquisition block

(one non-DW volume every 10 DW volumes, 3.2 minutes). The mean signals of the non-DW volumes were used to fit polynomial coefficients using MATLAB's *polyfit* function, which were then fed into the *polyval* function to obtain scaling factors for DW volumes in between those non-DW volumes. The signal intensity differences across acquisition blocks were normalized to the mean signal intensity of the first session.

ii) Susceptibility-induced and eddy current-induced distortion estimations and corrections were performed using topup<sup>56,61</sup> and the eddy-current correction toolbox<sup>62-64</sup> in the FSL<sup>56-58</sup>. The warp fields for each volume were obtained and stored for later use in the combined one-step resampling correction. Note that topup and eddy-current correction were performed for each individual data subset (2 acquisition blocks with paired reversed phase encoding block), such that a field map could be obtained for each data subset to account for field changes due to field drift along time or subject position changes. In addition to updating the field map, the interaction between susceptibility-induced field changes and subject motion within each data block was also considered by using a dynamic correction algorithm<sup>64</sup> as implemented in the FSL eddy toolbox<sup>62-64</sup>.

iii) The warp field for gradient nonlinearity correction<sup>54,65</sup> was calculated for the diffusion data. Here, the gradient nonlinearity correction was performed after topup and eddy correction<sup>54</sup>. This is because of the consideration that the voxel displacements due to susceptibility and eddy current distortions are much larger than those due to within-block rigid motion. Therefore, to obtain better estimation of these larger voxel displacements, the original phase-encoding direction was maintained prior to the estimation and correction for susceptibility and eddy current-induced distortions<sup>54</sup>. The effects of performing gradient-nonlinearity at a later step on the accuracy of within-block rigid motion estimation were considered to be small, because we assume that the confined motion within each acquisition block would only cause small differences in image deformation caused by gradient non-linearity.

iv) The next step estimated larger motion parameters across different data subsets such that the images from different acquisition blocks and sessions could be registered and aligned. This was performed using FLIRT<sup>66-68</sup> in FSL<sup>56-58</sup> with 6 degrees of freedom (DOF) rigid motion.

v) Finally, the warp fields for susceptibility, eddy-current distortion, gradient nonlinearity and motion from steps ii-iv were combined and applied to each image volume in a one-step resampling process to ensure minimal smoothing effects due to interpolation<sup>54,55</sup>. The Jacobian matrix computed from the combined warp field was used to correct for the signal intensity affected by the transformation.

All 18 data subsets from 9 sessions and their corresponding b-value and b-vector files were concatenated together and stored in compressed Neuroimaging Informatics Technology Initiative (NIfTI) format for downloading (Fig. 1d, data output B) and further analysis. Freesurfer<sup>69</sup> functions were used to load and save NIfTI data.

To provide flexibility for the end users to use their own preprocessing pipeline of choice, minimally preprocessed data are also made available for download in compressed NIfTI format (Fig. 1c, data output A). With the minimally preprocessed data, only the gradient nonlinearity correction has been performed without any other preprocessing corrections. The rationale in performing this correction is that the gradient nonlinearity coefficients needed in this correction are considered proprietary information by Siemens, and are not publicly available.

T<sub>1</sub> and T<sub>2</sub>-weighted images are available for download in compressed NIfTI format (Fig. 1c, data output A). After online reconstruction, the DICOM data were converted to NIfTI format using dcm2nii<sup>70</sup> tool of MRICron (<http://www.mricron.com>), and MRI gradient nonlinearity correction was performed. Subsequently, face and ear regions were masked to de-identify the high-resolution structural images using the defacing tool provided by FSL<sup>56-58</sup>. The averaged images of the three



0.7-mm  $T_1$  and  $T_2$ -weighted images are also provided after registration, averaging and de-identification. The low-resolution field maps are provided after gradient nonlinearity correction.

## Data Records

All data records listed in this section are publicly available in dryad<sup>71,72</sup>.

### dMRI data

- i) The preprocessed dMRI data are saved under 'preprocessedDWIData.zip'<sup>71</sup>. All the images in 9 sessions are saved in a single NIfTI file. The corresponding b-value and b-vector files are also provided.
- ii) The minimally preprocessed dMRI data are saved under 'reconDWIData.zip'<sup>72</sup>. 18 individual data subsets are saved in the compressed NIfTI format (paired phase encoding direction concatenated together), encoded by labels identifying scan session and data subset number. The corresponding b-value and b-vector files are also provided in the folder.

### Structural data

- i)  $T_1$ -weighted images are saved under 'T1.zip'<sup>71</sup> in compressed NIfTI format, containing one 1-mm MEMPRAGE image (echo combined), three 0.7-mm MEMPRAGE images (echo combined) and their average image.
- ii)  $T_2$ -weighted images are saved under 'T2.zip'<sup>71</sup> in compressed NIfTI format, containing three 0.7-mm T2-SPACE images and their average image.
- iii) Field maps for all 9 sessions are saved under 'FieldMap.zip'<sup>71</sup> in compressed NIfTI format, labeled by numbers identifying scan session.

Supplementary to the above data, a readme text file<sup>71</sup> is included in the repository with more information of the datasets.

## Technical Validation

### Assessment of DW image quality and head motion

The single and mean DWIs at  $b = 1000 \text{ s/mm}^2$  and  $b = 2500 \text{ s/mm}^2$  in three orthogonal views are shown in Figure 3a, which highlight the high image quality of this isotropic resolution dataset in all three spatial orientations with minimal reconstruction and ghosting artifacts. The high isotropic spatial resolution of the data reveals detailed structures of the brain (e.g., thalamus, globus pallidus, cerebellum) and sharp delineation of the gray-white matter boundaries (apparent in mean DWIs).

Subject motion between the 2808 dMRI volumes across the 9 two-hour sessions was evaluated in Figure 3b. The motion parameter estimates from different sessions acquired on different dates were concatenated together in a chronological order and separated by dashed lines. Within each two-hour session, the translations along the LR and the HF directions as well as the rotations around the AP (roll) and the HF axes (yaw) were all confined within a  $\pm 1 \text{ mm/degree}$  range. The rotations around the LR axis was slight larger but was still within a  $\pm 2\text{-degree}$  range. The estimates of translational motion along the AP direction appeared to be the highest and contained fluctuations and large jumps between the 25-minute acquisition blocks with opposing phase-encoding directions. This likely resulted from the inclusion of some eddy-current induced translation along the phase-encoding direction, which is known to be difficult to differentiate from actual subject motion since they affect the data in the same way. However, whether this translation is included in the eddy

current-induced estimate or the subject movement-induced estimate will not affect the image motion correction. Across different scan sessions, the abrupt changes in motion parameters represented differences in head position placements across sessions. With the use of the personalized head motion stabilizer, the position changes across different sessions were minimized to within  $\pm 2$  mm/degrees.

### **Assessment of diffusion modelling outcomes of the multi-session dataset**

The quality of the multi-session dataset was also assessed by evaluating the diffusion modelling outcomes. Specifically, the diffusion tensor (DT) model and a multi-shell extension of the ball & stick model<sup>73-75</sup> implemented in the bedpostx tool in FSL were used. The necessity and benefits of the multi-session data acquisition were illustrated by comparing the above diffusion outcomes obtained using data from 1, 3 and 9 sessions (total dMRI data acquisition time of 1.6, 4.8 and 14.5 hours). As shown in Figure 4, increasing the amount of data significantly increased the SNR and quality of the fractional anisotropy (FA) maps. The FA maps obtained using data from all 9 sessions revealed exquisite detail within fine-scale structures, including the radiating fiber bundles within the corpus callosum, highlighted as red stripes on the FA map (Fig. 4, zoomed-in area, arrow-A); the  $\sim 1$ -voxel-width line of reduced-FA separating the corticospinal tract from an adjacent tract (Fig. 4, zoomed-in area, arrow-B), highlighting the low level of blurring and high spatial accuracy of this data acquired across 9 two-hour sessions; and the gray matter bridges that span the internal capsule, giving rise to the characteristic stripes seen in the striatum (Fig. 4, zoomed-in area, arrow-C). In addition, the increased SNR and the denser q-space sampling from the multi-session data dramatically increased the ability to resolve three-way crossing fibers (see Fig. 4 for increased volume fraction of the third fiber shown in red-yellow on top left) and revealed complex intermingling fiber bundles within the centrum semiovale (see Fig. 4 fiber orientation maps modelling 3 fiber compartments per voxel), where the callosal fibers, the corticospinal tract and the superior longitudinal fasciculus intersect.

### **Assessment of spatial resolution**

A comparison of diffusion results at different spatial resolutions was performed to validate the benefit of using high spatial resolution acquisition. The dMRI data at 0.76 mm isotropic resolution was compared with downsampled data at 1-mm, 1.25-mm and 1.5-mm isotropic resolution (using sinc interpolation on dMRI data, i.e. cropping in the kspace of the dMRI data). Figure 5 shows the colored-FA maps of four zoomed-in areas (yellow squares) at those spatial resolutions. Improved delineation of fiber bundles and more detailed structures can be captured under higher spatial resolution. Figure 6 shows the comparison of DTI tensors at 1.5 mm and 0.76 mm isotropic resolution, which consistently highlights the improved capability of the submillimeter dMRI data to visualize subcortical white matter as it turns into the cortex. A clear dark band with reduced FA at the gray-white junction can be seen at 0.76-mm resolution (red arrows, top and middle rows). In addition, 0.76-mm data detected more sharp-turning fibers such as those connecting cortical regions between adjacent gyri (red arrows, bottom row).

### **Code Availability**

The MATLAB code for the preprocessing steps described above are included in the released dataset. The code make use of the functions in the FSL<sup>56-58</sup> toolbox and Freesurfer<sup>69</sup> that are freely available for download at <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslInstallation> and <https://surfer.nmr.mgh.harvard.edu/fswiki/DownloadAndInstall>.

## Acknowledgements

This work was supported by the NIH (R01-EB020613, R01-EB019437, R01-MH116173, U01-EB025162, and U01-EB026996, U01-MH093765, P41-EB015896, P41-EB030006, K23-NS096056) and NIH shared instrumentation grants (S10-RR023401, S10-RR023043, and S10-RR019307). We thank Dr. Cornelius Eichner for sharing his python-based code for optimized preprocessing of dMRI data with one-step resampling.

## Author contributions

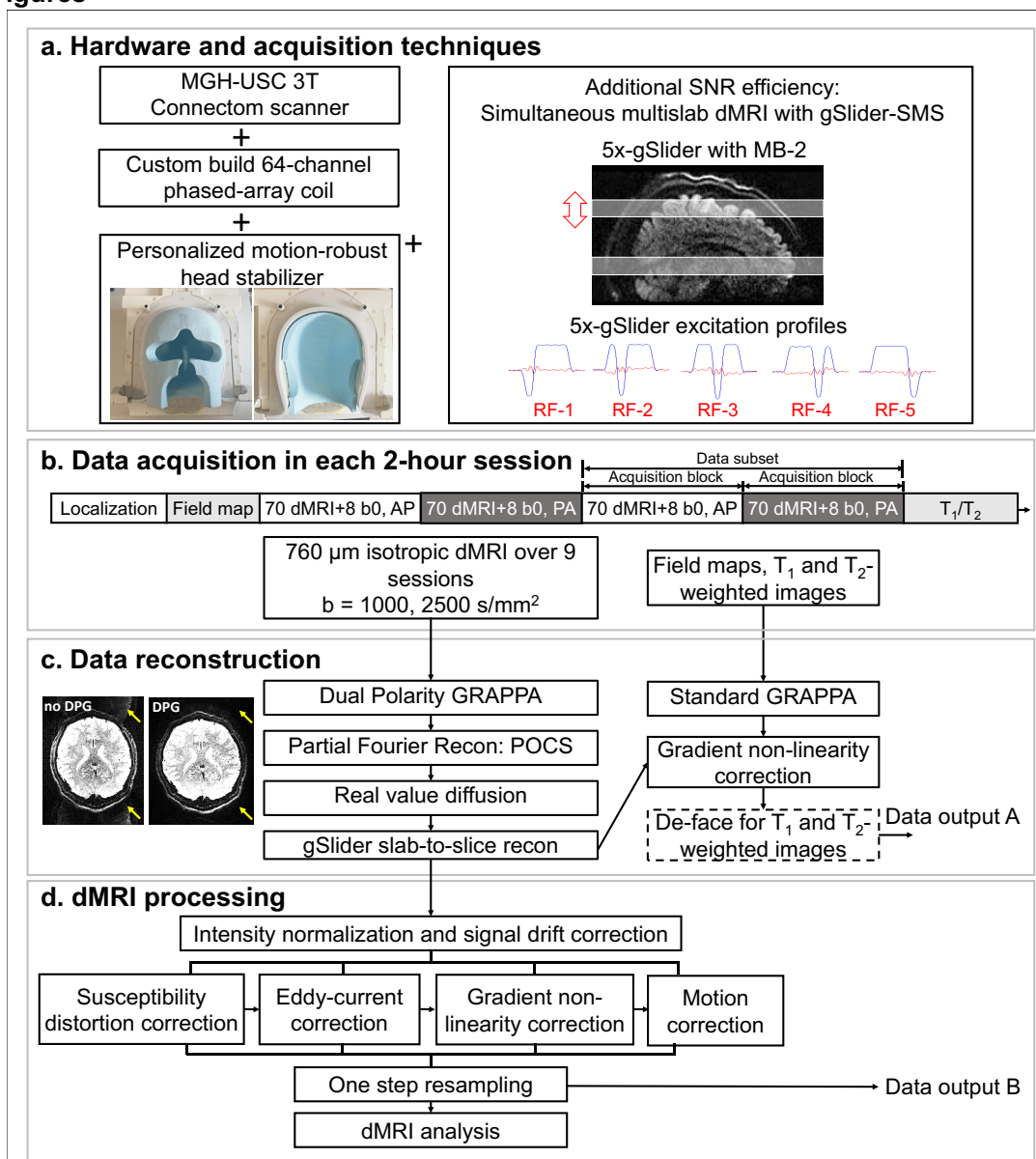
F.W. conceptualized and conducted the study, acquired and processed the data, and wrote the manuscript. Z.D. helped with the data acquisition and the data processing, and contributed to the manuscript. Q.T., C.L., and Q.F. helped with the data processing. W.S.H. helped with the DPG reconstruction. B.K. provided the 64-channel phased-array coil. J.R.P., L.L.W., and S.Y.H. provided the conceptual discussion and contributed to the manuscript. K.S. conceptualized and supervised the study, and contributed to the manuscript.

## Additional information

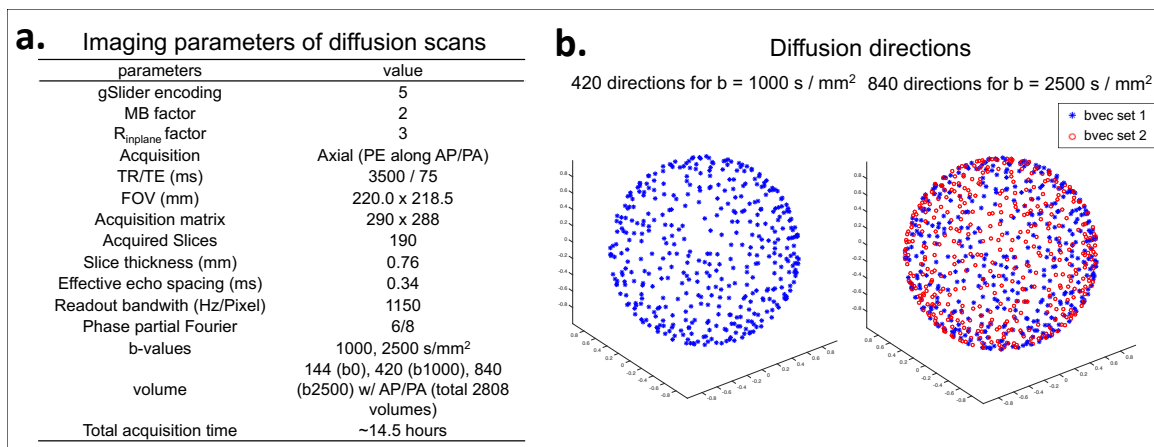
**Competing financial interests:** the authors declare no competing financial interest.



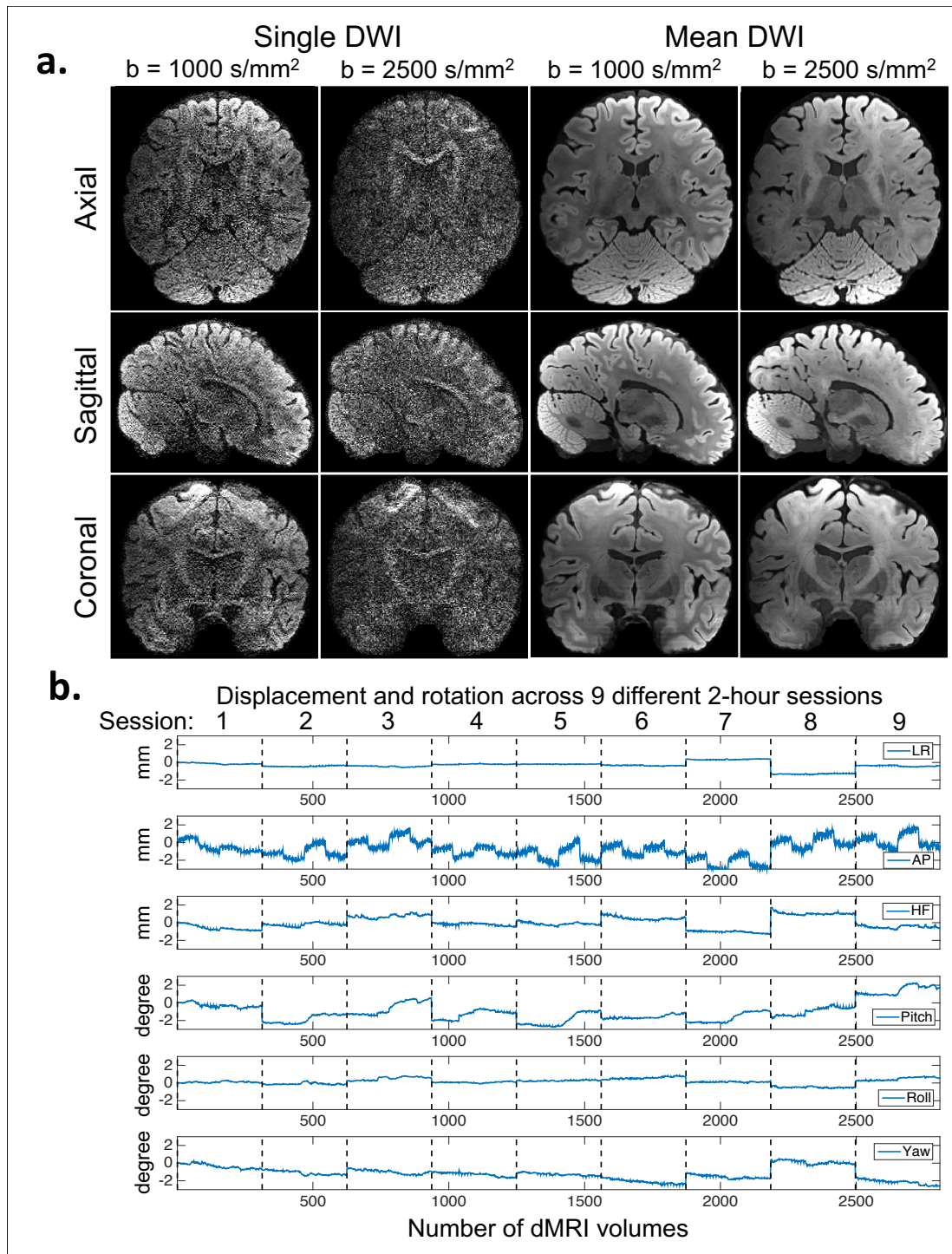
## Figures



**Figure 1. Data acquisition, reconstruction and preprocessing workflow.** (a) Hardware and acquisition techniques used to acquire the data. The Connectom scanner, the custom-built 64-channel phased-array coil and the gSlider-SMS acquisition technique provide high SNR, while the personalized motion-robust stabilizer ensures consistent positioning across multiple scan sessions and minimal motion contamination within each session. (b) Data acquisition pipeline in each 2-hour session. The dMRI acquisition was divided into 4 acquisition blocks, each of which comprised 70 diffusion volumes and 8 interleaved  $b = 0 \text{ s/mm}^2$  volumes. Two acquisition blocks with paired reversed phase encoding composed a data subset with a unique set of diffusion directions. 6 data subsets were acquired at  $b = 1000 \text{ s/mm}^2$  in the first 3 sessions, and 12 data subsets at  $b = 2500 \text{ s/mm}^2$  in the remaining 6 sessions. (c) Data reconstruction workflow. dMRI data were reconstructed using Dual Polarity GRAPPA, partial Fourier reconstruction, real value diffusion and gSlider slab-to-slice reconstruction, while field maps and structural images were reconstructed using standard GRAPPA reconstruction implemented on the scanner. Minimally preprocessed data were then provided after the gradient non-linearity correction (Data output A). (d) dMRI preprocessing with one-step resampling. The preprocessed data were obtained after intensity normalization to correct for signal drift, along with susceptibility-induced and eddy current-induced distortion correction, gradient non-linearity correction and motion correction, all of which were performed together using a one step resampling process to minimize interpolation-induced image blurring (Data output B).

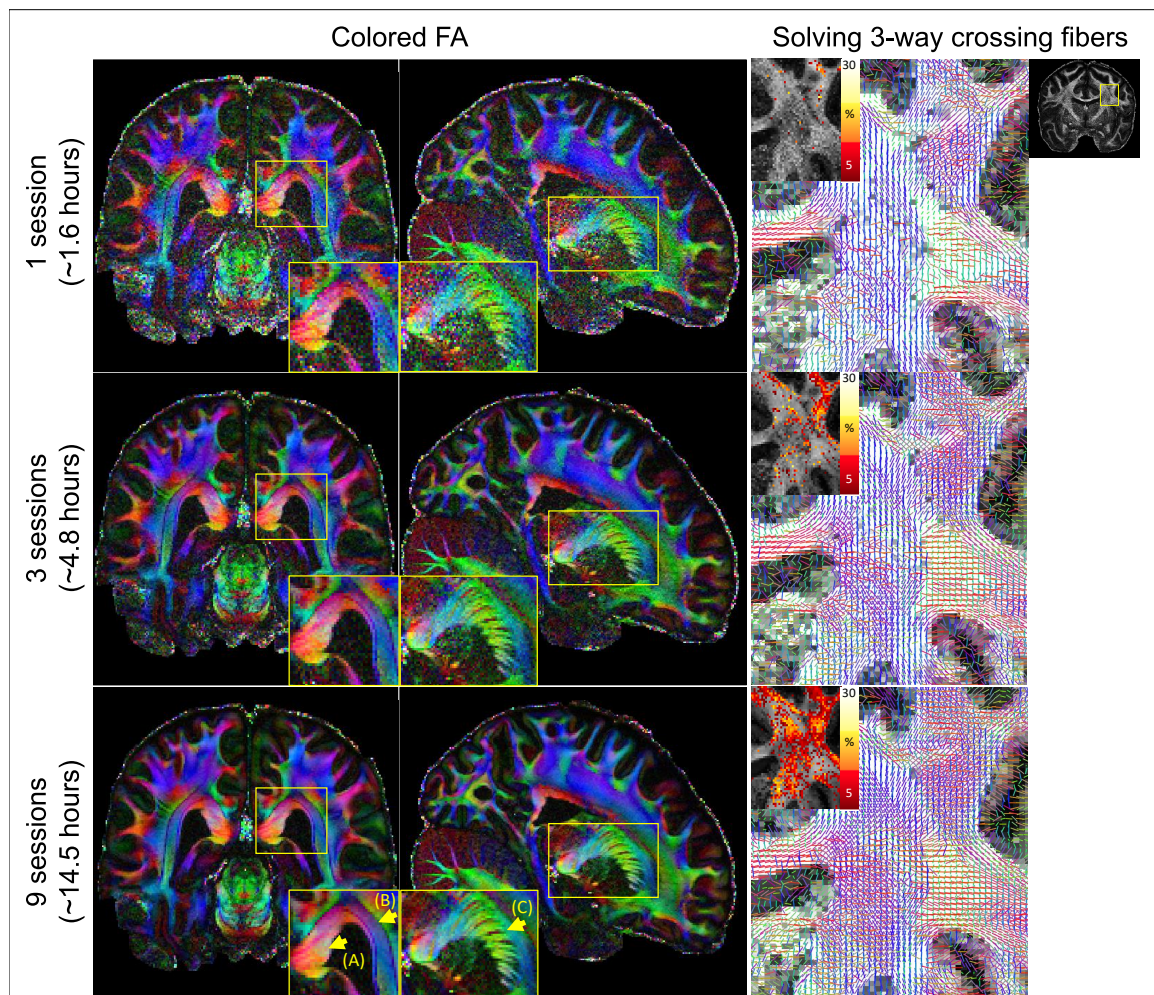


**Figure 2. Imaging parameters for the dMRI dataset.** (a) Detailed imaging parameters for gSlider-SMS acquisition protocol. (b) Diffusion directions of the datasets. The 420 directions at b = 1000 s/mm<sup>2</sup> is a subset of the 840 directions for b = 2500 s/mm<sup>2</sup> (shown in blue).

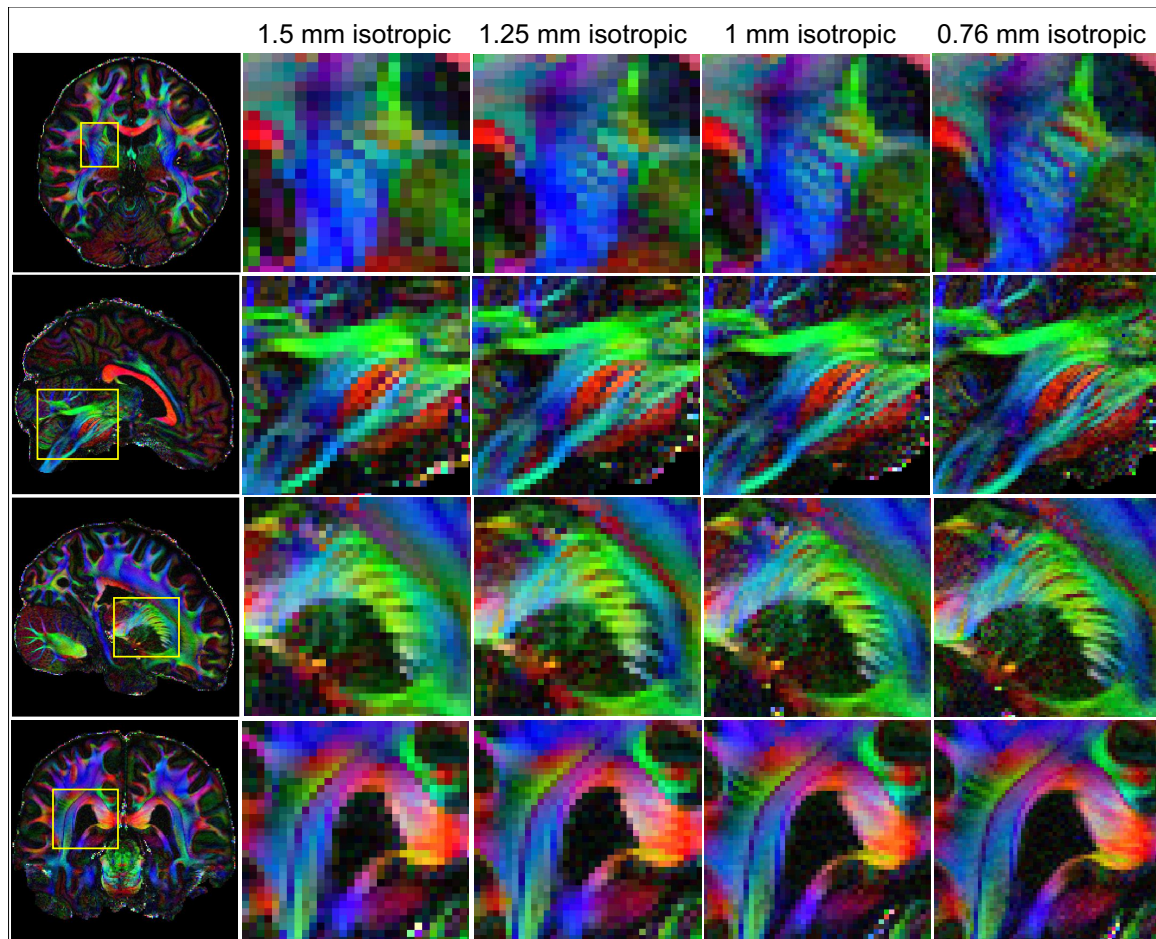


**Figure 3. DW images and subject movement measurements.** (a) Single DWI and mean DWI images at different b-values shown in three orthogonal views. (b) subject movement measurements of all 2808 dMRI volumes acquired in 9 two-hour sessions. From top to bottom are translations along left-right (LR), anterior-posterior (AP), head-foot (HF) and rotations around LR (Pitch), AP (Roll), HF (Yaw) axis. The motion parameters were estimated relative to the first  $b = 0 \text{ s/mm}^2$  volume in the first session. dMRI volumes acquired in different sessions on different dates are separated by black dashed lines. As shown, both within-session movements and across-session head position changes were kept small with the use of the personalized motion stabilizer.



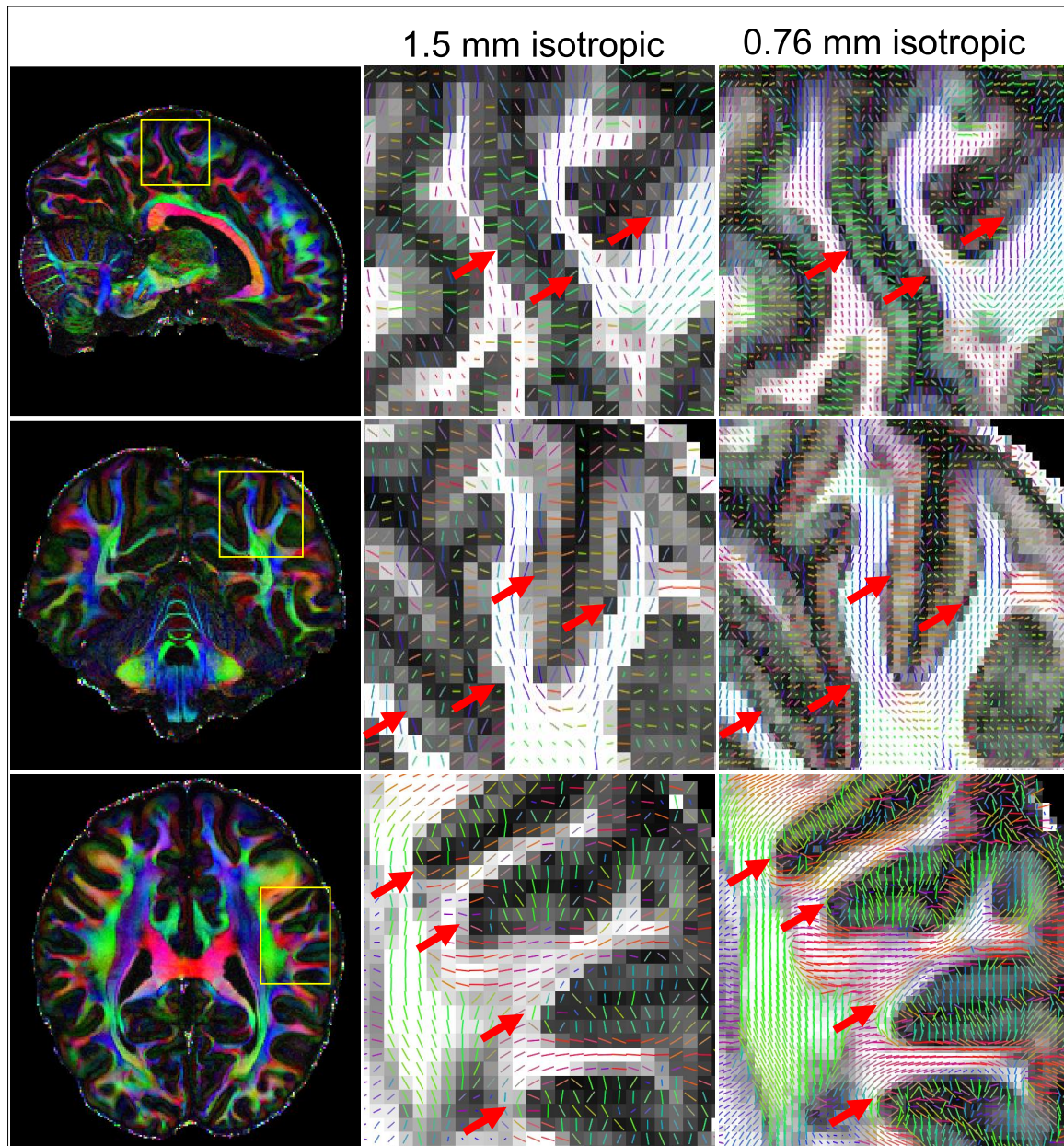


**Figure 4. Comparison of FA maps and crossing fiber analysis results using data from 1, 3 and 9 acquisition sessions.** On the left, colored FA maps in three orthogonal views are shown for data acquired in 1 session (top), 3 sessions (middle) and 9 sessions (bottom). The 9-session data provides high SNR FA and reveals exquisite fine-scale structures as shown in the zoomed-in areas (yellow boxes), including (A) the radiating fiber bundles within the corpus callosum; (B) the 1-voxel-width line of reduced-FA separating the corticospinal tract from an adjacent tract, highlighting the low level of blurring and high spatial accuracy of this data; and (C) the gray matter bridges that span the internal capsule. On the right, zoomed-in fiber orientations (three fiber compartments per voxel) within centrum semiovale (yellow box) overlaid on gray-scale maps of the total anisotropic volume fraction are shown for those data as well. In addition, the volume fraction of the third fiber are shown in red-yellow on the top left of each fiber orientation map to visualize the sensitivity of solving 3-way crossing fibers. Both fiber orientation and third-fiber volume fraction are shown when the respective volume fraction is higher than 5%. The denser q-space sampling and higher SNR of the 9-session data increases its ability to solve three-way crossing fibers and to reveal complex intermingling fiber bundles within the centrum semiovale.



**Figure 5. Comparison of dMRI results at different spatial resolutions.** (a) Colored FA maps and their zoomed-ins at different brain regions (yellow boxes) and at different isotropic resolutions.





**Figure 6. Comparison of DTI principal directions at 1.5-mm and 0.76-mm resolution.** Zoomed-in FA and DTI principal vector maps are shown in the area highlighted by yellow boxes. High resolution dMRI data shows improved ability to visualize subcortical white matters as it turns into the cortex. A clear dark-band with reduced FA at the grey-white junction can be seen (red arrows, top and middle rows) at 0.76 mm resolution, and more sharpening-turning fibers such as those connecting cortical regions between adjacent gyri (red arrows, bottom row) can be detected.



## References

- 1 Le Bihan, D. & Breton, E. J. C. r. d. I. A. d. s. S., Mécanique, Physique, Chimie, Sciences de l'univers, Sciences de la Terre. Imagerie de diffusion in vivo par résonance magnétique nucléaire. **301**, 1109-1112 (1985).
- 2 Moseley, M. *et al.* Diffusion-weighted MR imaging of acute stroke: correlation with T2-weighted and magnetic susceptibility-enhanced MR imaging in cats. *American Journal of Neuroradiology* **11**, 423-429 (1990).
- 3 Warach, S., Chien, D., Li, W., Ronthal, M. & Edelman, R. J. N. Fast magnetic resonance diffusion-weighted imaging of acute human stroke. **42**, 1717-1717 (1992).
- 4 Mangeat, G. *et al.* Changes in structural network are associated with cortical demyelination in early multiple sclerosis. **39**, 2133-2146 (2018).
- 5 Malayeri, A. A. *et al.* Principles and applications of diffusion-weighted imaging in cancer detection, staging, and treatment follow-up. **31**, 1773-1791 (2011).
- 6 Jones, D. K. *et al.* Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia. **27**, 230-238 (2006).
- 7 Sussmann, J. E. *et al.* White matter abnormalities in bipolar disorder and schizophrenia detected using diffusion tensor magnetic resonance imaging. **11**, 11-18 (2009).
- 8 Pasternak, O., Kelly, S., Sydnor, V. J. & Shenton, M. E. J. N. Advances in microstructural diffusion neuroimaging for psychiatric disorders. **182**, 259-282 (2018).
- 9 Goveas, J. *et al.* Diffusion-MRI in neurodegenerative disorders. **33**, 853-876 (2015).
- 10 Bozzali, M. *et al.* White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. **72**, 742-746 (2002).
- 11 Eickhoff, S., Nichols, T. E., Van Horn, J. D. & Turner, J. A. Sharing the wealth: neuroimaging data repositories. *Neuroimage* **124**, 1065 (2016).
- 12 Walker, L. *et al.* The diffusion tensor imaging (DTI) component of the NIH MRI study of normal brain development (PedsDTI). *Neuroimage* **124**, 1125-1130 (2016).
- 13 Hodge, M. R. *et al.* ConnectomeDB—sharing human brain connectivity data. *Neuroimage* **124**, 1102-1107 (2016).
- 14 Hsu, Y. C., Lo, Y. C., Chen, Y. J., Wedeen, V. J. & Isaac Tseng, W. Y. NTU-DSI-122: A diffusion spectrum imaging template with high anatomical matching to the ICBM-152 space. *Human brain mapping* **36**, 3528-3541 (2015).
- 15 Varentsova, A., Zhang, S. & Arfanakis, K. Development of a high angular resolution diffusion imaging human brain template. *NeuroImage* **91**, 177-186 (2014).
- 16 Mori, S., Wakana, S., Van Zijl, P. C. & Nagae-Poetscher, L. *MRI atlas of human white matter*. (Elsevier, 2005).
- 17 Poldrack, R. A. *et al.* Long-term neural and physiological phenotyping of a single human. *Nature communications* **6**, 1-15 (2015).
- 18 Froeling, M., Tax, C. M., Vos, S. B., Luijten, P. R. & Leemans, A. "MASSIVE" brain dataset: Multiple acquisitions for standardization of structural imaging validation and evaluation. *Magnetic resonance in medicine* **77**, 1797-1809 (2017).
- 19 Uğurbil, K. *et al.* Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage* **80**, 80-104 (2013).
- 20 Setsompop, K. *et al.* Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *NeuroImage* **80**, 220-233, doi:10.1016/j.neuroimage.2013.05.078 (2013).
- 21 McNab, J. A. *et al.* The Human Connectome Project and beyond: initial applications of 300 mT/m gradients. *NeuroImage* **80**, 234-245, doi:10.1016/j.neuroimage.2013.05.074 (2013).
- 22 Van Essen, D. C. *et al.* The WU-Minn human connectome project: an overview. *Neuroimage* **80**, 62-79 (2013).
- 23 Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. **19**, 1523-1536 (2016).
- 24 Miller, K. L. *et al.* Diffusion imaging of whole, post-mortem human brains on a clinical MRI scanner. **57**, 167-181 (2011).
- 25 Dyrby, T. B. *et al.* An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. **32**, 544-563 (2011).
- 26 Leuze, C. W. *et al.* Layer-specific intracortical connectivity revealed with diffusion MRI. **24**, 328-339 (2014).
- 27 Fan, Q. *et al.* High b-value and high Resolution Integrated Diffusion (HIBRID) imaging. *NeuroImage* **150**, 162-176 (2017).
- 28 McNab, J. A. *et al.* Surface based analysis of diffusion orientation for identifying architectonic domains in the in vivo human cortex. **69**, 87-100 (2013).
- 29 McNab, J. A. *et al.* The Human Connectome Project and beyond: initial applications of 300mT/m gradients. *Neuroimage* **80**, 234-245 (2013).

- 30 Setsompop, K. *et al.* Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *Neuroimage* **80**, 220-233 (2013).
- 31 Keil, B. *et al.* A 64-channel 3T array coil for accelerated brain MRI. *Magnetic resonance in medicine* **70**, 248-258 (2013).
- 32 Chen, N.-k., Guidon, A., Chang, H.-C. & Song, A. W. A robust multi-shot scan strategy for high-resolution diffusion weighted MRI enabled by multiplexed sensitivity-encoding (MUSE). *Neuroimage* **72**, 41-47 (2013).
- 33 Wu, W. *et al.* High-resolution diffusion MRI at 7T using a three-dimensional multi-slab acquisition. *NeuroImage* **143**, 1-14 (2016).
- 34 Dong, Z. *et al.* Interleaved EPI diffusion imaging using SPIR i T-based reconstruction with virtual coil compression. *Magnetic resonance in medicine* **79**, 1525-1531 (2018).
- 35 Dong, Z. *et al.* Tilted-CAIPI for highly accelerated distortion-free EPI with point spread function (PSF) encoding. *Magnetic resonance in medicine* **81**, 377-392 (2019).
- 36 Frost, R., Jezzard, P., Porter, D. A., Tijssen, R. & Miller, K. in *Proceedings of the 21st Annual Meeting of ISMRM, Salt Lake City, USA, (Abstract 3176)*.
- 37 Engström, M. & Skare, S. Diffusion-weighted 3D multislab echo planar imaging for high signal-to-noise ratio efficiency and isotropic image resolution. *Magnetic resonance in medicine* **70**, 1507-1514 (2013).
- 38 Song, A. W., Chang, H.-C., Petty, C., Guidon, A. & Chen, N.-K. Improved delineation of short cortical association fibers and gray/white matter boundary using whole-brain three-dimensional diffusion tensor imaging at submillimeter spatial resolution. *Brain connectivity* **4**, 636-640 (2014).
- 39 Chang, H.-C. *et al.* Human brain diffusion tensor imaging at submillimeter isotropic resolution on a 3Tesla clinical MRI scanner. *NeuroImage* **118**, 667-675 (2015).
- 40 Setsompop, K. *et al.* High-resolution in vivo diffusion imaging of the human brain with generalized slice dithered enhanced resolution: Simultaneous multislice (gSlider-SMS). *Magnetic Resonance in Medicine* (2017).
- 41 Wang, F. *et al.* Motion-robust sub-millimeter isotropic diffusion imaging through motion corrected generalized slice dithered enhanced resolution (MC-gSlider) acquisition. *Magnetic resonance in medicine* **80**, 1891-1906 (2018).
- 42 Halder, J. P., Liu, Y., Liao, C., Fan, Q. & Setsompop, K. Fast submillimeter diffusion MRI using gSlider-SMS and SNR-enhancing joint reconstruction. *Magnetic Resonance in Medicine* **84**, 762-776 (2020).
- 43 Fan, Q. *et al.* MGH-USC Human Connectome Project datasets with ultra-high b-value diffusion MRI. *Neuroimage* **124**, 1108-1114 (2016).
- 44 Hoge, W. S. & Polimeni, J. R. Dual-polarity GRAPPA for simultaneous reconstruction and ghost correction of echo planar imaging data. *Magnetic resonance in medicine* **76**, 32-44 (2016).
- 45 Hoge, W. S., Setsompop, K. & Polimeni, J. R. Dual-polarity slice-GRAPPA for concurrent ghost correction and slice separation in simultaneous multi-slice EPI. *Magnetic resonance in medicine* **80**, 1364-1375 (2018).
- 46 Calabrese, E. *et al.* Investigating the tradeoffs between spatial resolution and diffusion sampling for brain mapping with diffusion tractography: Time well spent? **35**, 5667-5685 (2014).
- 47 Setsompop, K. *et al.* Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magnetic Resonance in Medicine* **67**, 1210-1224 (2012).
- 48 Caruyer, E., Lenglet, C., Sapiro, G. & Deriche, R. Design of multishell sampling schemes with uniform coverage in diffusion MRI. *Magnetic resonance in medicine* **69**, 1534-1540 (2013).
- 49 van der Kouwe, A. J., Benner, T., Salat, D. H. & Fischl, B. Brain morphometry with multiecho MPRAGE. *Neuroimage* **40**, 559-569 (2008).
- 50 Lichy, M. P. *et al.* Magnetic resonance imaging of the body trunk using a single-slab, 3-dimensional, T2-weighted turbo-spin-echo sequence with high sampling efficiency (SPACE) for high spatial resolution imaging: initial clinical experiences. *Investigative radiology* **40**, 754-760 (2005).
- 51 Haacke, E., Lindskog, E. & Lin, W. A fast, iterative, partial-Fourier technique capable of local phase recovery. *Journal of Magnetic Resonance (1969)* **92**, 126-145 (1991).
- 52 Eichner, C. *et al.* Real diffusion-weighted MRI enabling true signal averaging and increased diffusion contrast. *NeuroImage* **122**, 373-384 (2015).
- 53 Liao, C. *et al.* High-fidelity, high-isotropic-resolution diffusion imaging through gSlider acquisition with and T1 corrections and integrated  $\Delta B_0/R_x$  shim array. *Magnetic resonance in medicine* **83**, 56-67 (2019).
- 54 Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80**, 105-124 (2013).
- 55 Eichner, C. *et al.* in *Proceedings of the 27th Annual Meeting of ISMRM, Montreal, Canada*. 1047.
- 56 Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23**, S208-S219 (2004).
- 57 Woolrich, M. W. *et al.* Bayesian analysis of neuroimaging data in FSL. *Neuroimage* **45**, S173-S186 (2009).

- 58 Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. Fsl. *Neuroimage* **62**, 782-790 (2012).
- 59 Vos, S. B. *et al.* The importance of correcting for signal drift in diffusion MRI. *Magnetic resonance in medicine* **77**, 285-299 (2017).
- 60 Jeurissen, B., Leemans, A., Tournier, J. D., Jones, D. K. & Sijbers, J. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human brain mapping* **34**, 2747-2766 (2013).
- 61 Andersson, J. L., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* **20**, 870-888 (2003).
- 62 Andersson, J. L. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* **125**, 1063-1078 (2016).
- 63 Andersson, J. L., Graham, M. S., Zsoldos, E. & Sotiropoulos, S. N. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage* **141**, 556-572 (2016).
- 64 Andersson, J. L., Graham, M. S., Drobniak, I., Zhang, H. & Campbell, J. Susceptibility-induced distortion that varies due to motion: correction in diffusion MR without acquiring additional data. *Neuroimage* **171**, 277-295 (2018).
- 65 Jovicich, J. *et al.* Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage* **30**, 436-443 (2006).
- 66 Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825-841 (2002).
- 67 Jenkinson, M. & Smith, S. A global optimisation method for robust affine registration of brain images. *Medical image analysis* **5**, 143-156 (2001).
- 68 Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* **48**, 63-72 (2009).
- 69 Fischl, B. FreeSurfer. *Neuroimage* **62**, 774-781 (2012).
- 70 Li, X., Morgan, P. S., Ashburner, J., Smith, J. & Rorden, C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of neuroscience methods* **264**, 47-56 (2016).
- 71 Wang, F. *et al.* Data from: In vivo human whole-brain Connectom diffusion MRI dataset at 760  $\mu$ m isotropic resolution (PART I). *Dryad* <https://doi.org/10.5061/dryad.nzs7h44q2> (2020).
- 72 Wang, F. *et al.* Data from: In vivo human whole-brain Connectom diffusion MRI dataset at 760  $\mu$ m isotropic resolution (PART II). *Dryad* <https://doi.org/10.5061/dryad.rjdfn2z8g> (2020).
- 73 Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34**, 144-155 (2007).
- 74 Jbabdi, S., Sotiropoulos, S. N., Savio, A. M., Graña, M. & Behrens, T. E. Model-based analysis of multishell diffusion MR data for tractography: how to get over fitting problems. *Magnetic resonance in medicine* **68**, 1846-1855 (2012).
- 75 Hernández, M. *et al.* Accelerating fibre orientation estimation from diffusion weighted magnetic resonance imaging using GPUs. *PloS one* **8**, e61892 (2013).